MOLECULAR MAPPING AND CLONING OF POLYEMBRYONY (*OsPE*) INSERTIONAL MUTANT IN BASMATI RICE

A THESIS

Submitted in partial fulfilment of the requirements for the award of the degree of

DOCTOR OF PHILOSOPHY

in

BIOTECHNOLOGY

By

ANJU BHALLA

DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE-247 667 (INDIA)

APRIL, 2008

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **MOLECULAR MAPPING AND CLONING OF POLYEMBRYONY (*OsPE*) INSERTIONAL MUTANT IN BASMATI RICE** in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy and submitted in the Department of Biotechnology of the Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during the period from April 2003 to April 2008 under the supervision of Dr. H. S. Dhaliwal and Dr. G. S. Randhawa, Professors, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee, India.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institute.

(ANJU BHALLA)

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

(G. S. Randhawa)
Supervisor

(H. S. Dhaliwal)
Supervisor

Date: 5/4/2008

The Ph.D. Viva-Voce Examination of **Ms. Anju Bhalla**, Research Scholar has been held on .....

Signature of Supervisors

Signature of External Examiner

CONTENTS

Abstract

In plant functional genomics, mutants play an important role in studying gene functions. Rice (*Oryza sativa* L.) is recognized as an important model plant among cereals because of its small genome size and the development of efficient transformation system. T-DNA insertional mutagenesis is one of the most important approaches of finding and cloning new genes. It is believed that the T-DNA insertions in the genome are random, and that the insertions can be stably transmitted.

Analysis of the flanking sequences of three T-DNA/*Ds* insertional mutants of *Oryza sativa*, cultivar Basmati 370: B-4-1 (Polyembryony; *OsPE*), B-2-2 and B-8-7 has been done. Sequence analysis showed that the T-DNA insertion was in the promoter region in *OsPE*, in the coding sequence in B-2-2 and in the junk region of B-8-7. PCR amplification with 950bp *hpt* gene amplification (used as selectable marker during transformation) confirmed T-DNA insertions in all the three mutants and Southern blot analysis showed that the mutant phenotype in polyembryony is due to single copy T-DNA insertion. The present investigation aims at mapping and cloning the candidate gene for multiple embryo formation in one of the independent T-DNA insertional mutants.

OsPE mutant represents the first reported example of high frequency and heritable polyembryony in rice caused by insertional mutagenesis. Polyembryony has been described previously in several crops and cereal plants, but the underlying genetic causes of twinning in such cases have been complex and difficult to resolve. This homozygous and fertile polyembryonic mutant showed resistance to hygromycin

and the twin/triplet seedlings at a frequency of 15-20%. Multiple embryos with independent root and shoot axis were attached to single scutellum.

A mapping population was developed by crossing *OsPE* with a non-Basmati fine rice cultivar PR106 and a large F_2 was obtained. Morphological data for plant height, number of tillers per plant and *hpt* gene amplification was collected for F_2 population and for *hpt* gene (χ^2= 1.156 at 1 *df* and p< 0.05). The F_2 phenotypic data for polyembryony unlike *hpt* data did not segregate in the expected 3:1 ratio. Out of 175 F_2 plants only 14 showed polyembryony (twin/triplet) phenotype. All the 14 plants which showed polyembryonic phenotype gave PCR amplification for *hpt* gene as well. This may be due to incomplete penetrance and variable expressivity in the inheritance of polyembryonic phenotype in F_2 population. Penetrance here refers to the presence of polyembryony while expressivity as single, twin, triplet and quadruplet embryos. The *OsPE* mutant was highly fertile with somatic chromosomal number 2n= 24 and normal 12 IIs at meiotic metaphase I.

A set of 98 markers showing polymorphism between parental lines, Basmati 370 and PR106 were used in the *OsPE*/PR106 mapping population for identifying markers associated with the *OsPE* gene through Bulk Segregant Analysis. RM14645 (5.79cM) and RM14667 (2.17cM) markers showed linkage with *OsPE* gene on chromosome 3. The mapping of *OsPE* on rice chromosome 3 using BSA has been further confirmed with other reverse genetic approaches including genome walking and TAIL-PCR involving amplification of T-DNA flanking region using T-DNA and adaptor based primers.

To carry out genome walking and TAIL-PCR, nested primers were designed for T-DNA right and left borders. Using T-DNA right border nested primers along with adaptor primer (Genome Walking) and arbitrary degenerate primer (TAIL-PCR),

PCR amplification was obtained. The PCR product was purified and sequenced. The location of *OsPE* gene on chromosome was determined using *japonica* rice cultivar Nipponbare sequence. Sequencing with T-RB2/AP2 (Genome Walking) and T-RB3 (TAIL-PCR) showed significant alignment with the *Oryza sativa* cv. Nipponbare with only single hit at chromosome 3. The T-DNA insertion is present in the promoter region of the candidate gene Os03g0241300. The length of the candidate gene is 2.8Kb with 2 UTRs (1473bp and 193bp), 2 Exons (154bp and 935bp), and one Intron (81bp). To confirm the insertion on chromosome 3, primers were designed for rice genome flanking the T-DNA borders. PCR was done in combination with the T-DNA (T-RB) based primers and genome-specific primers. Amplification of expected size was obtained in polyembryonic mutant but not in Basmati 370. PCR with genome-specific primers gave amplification in Basmati 370 and not in *OsPE* because of the insertion of approximately 10Kb fragment of the T-DNA cassette. All the 14 polyembryonic *hpt* positive F_3 progenies also showed amplification with the T-DNA specific (T-RB1) and genome specific (PE-RB1) primers.

The protein searched corresponding to the *OsPE* candidate gene sequence has been reported as a hypothetical protein (Os03g0241300) in *Oryza sativa*. *OsPE* gene lacked functional homologs in other species. Although using KEGG search around 401 sequences orthologous to *OsPE* were found through out the living kingdom but majority had low identity value (> 40%). It most closely matched with the *Vitis* and *Arabidopsis* hypothetical and unknown proteins, respectively. No *OsPE* paralog was found in rice. No conserved domains were found in the protein coded by *OsPE*. Availability of homologs genes of known function, specific tissue expression patterns and conserved domains would have supported assigning function to this novel gene, *OsPE*.

RT-PCR using reverse transcribed mRNA and primer pair designed on intron exon junction, the expression of *OsPE* gene was found in Basmati 370 shoots and not in roots. The expression profile of the candidate gene suggested by EST counts showed maximum ESTs in seed (11/32357) followed by flower (23/136502) and panicle (23/132789). The polyembryony is associated with seed where maximum ESTs were found for the *OsPE* gene.

Full length *OsPE* gene was cloned in Basmati 370 on the basis of the sequences obtained by genome walking and TAIL-PCR. The nucleotide sequence of the candidate gene *OsPE* in *japonica* rice was used to design overlapping PCR primers and the *OsPE* gene was amplified in Basmati 370 genomic DNA as the template. The sequence obtained using corresponding forward and reverse primers were aligned to determine the sequence errors and full length *OsPE* gene in Basmati 370 was obtained. BLAST search was done for homologous sequence in Nipponbare *japonica* rice. Ten SNPs (Single Nucleotide Polymorphism) and one gap were found in the *OsPE* candidate gene sequence as compared to Nipponbare sequence thus suggesting a total of 1% dissimilarity between the two sequences. The predicted protein was searched and variation was found in only two amino acids. With the minor differences in coding region the *OsPE* gene is expected to translate a similar protein in Basmati 370.

The *OsPE* gene BLAST alignment against rice FST database (Rice Functional Genomics Dataabase RiceGE) showed two hits. The *OsPE* gene has insertion in the promoter region whereas FST searched insertions are found to be in the coding region. No phenotypic description or associated annotations related to these insertions were reported.

Identification for flanking sequences for two more independent T-DNA insertional mutants, B-2-2 and B-8-7 was done using TAIL-PCR. The insertion in chromosome 11 of B-2-2 is present in the gene that codes for Pectin methylesterase (PME).

The combined use of Southern blot, BSA, genome walking, TAIL-PCR, RT-PCR techniques and bioinformatics led to the identification of a candidate gene controlling the multiple embryo formation in Basmati rice. A detailed molecular characterization and mapping of the *OsPE* insertional mutant is available through this study. It can thus be speculated that *OsPE* gene controls the number of embryos or embryological divisions as the phenotype of the *OsPE* insertional mutant shows polyembryony with no sterility. The validation of the *OsPE* and other candidate genes can be done using two element *Ac/Ds* transposon system and RNAi approach.

Acknowledgements

I believe God has directly helped me in this endeavour of mine i.e. in the form of people who have seen me through all the while. The acknowledgement of their contributions has been indelibly imprinted on my mind.

The first name among them that strikes me is the scientist who put the meaning of 'super' into the word supervisor, **Prof H. S. Dhaliwal**, for his endless enthusiasm and support over the last few years. His total commitment, excellent guidance, constructive criticism, judicious examination and a fatherly trust in me leave me short of words to articulate my gratitude for him. You have been a great mentor sir!

I am also very grateful to my supervisor **Prof. G. S. Randhawa** who has helped me sharpen and strengthen my research skills and develop my thinking as a confident and sincere researcher and individual.

I convey my thanks and regards to Prof. B. M. J. Pereira Head, Department of Biotechnology, Indian Institute of Technology Roorkee. I highly appreciate the cooperation and affection extended to me by all the faculty and staff of Department of Biotechnology, Indian Institute of Technology Roorkee.

My special thanks to Dr. Kuldeep Singh, Molecular Geneticist, Punjab Agricultural University, Ludhiana for providing space and freedom to work in his laboratory.

I would like to thank Dr. Kianian, Department of Plant Sciences, North Dakota State University, USA for all those right words of advice. I extend my gratitude to all my friends and coworkers at the Department of Plant Sciences, North Dakota State

University, Dr. Rishi, Dr. Farhad and Dr. Khwaja (Postdoc Fellows), Allen and Justin (Technicians), including all present and former WGE group members.

The financial assistance by CSIR (India) and NSF (USA) is highly acknowledged.

A special thanks to the mutant plants I worked with and sacrificed.

Meanwhile, I would not like to forget, my fellow scholars in the department for their wonderful helping attitude, and for providing an affable environment at the workplace. I would particularly like to remember my labmates, Osman, Deepak, Mankesh, Javed, Naincy, Nidhi, Neelam, Vijay and Manish on this occasion. The four of us (Rice group) were HSD's first Ph.D. students from IIT Roorkee, and the uphill struggle of building a research lab was made easier because it was done with good friends. We made lab an enjoyable place each and every day. I will cherish all moments spent with those people. I am going to miss them all, they were great friends.

As an important phase in my life comes to an end, I will miss my coffee buddies Seema and Anindita, and the good times we shared together.

Special thanks to Dr. Anvita, Dr. Sumati, Anjana, Anna (Vishwa) and Divya who have been incomparable companions.

To all my relatives at Roorkee, especially Mukut *Chacha* for enabling me to laugh aloud and have a sense of humour about life. For always being at five minutes distance (either from the Department or hostel) and during that time they always kept their home open to me. You have been a great teacher and friend to me. To Jeet, Channi and Pammi *Chacha* for their help through all my highs and lows during my stay at Roorkee. I need their blessings as I go into the real world.

Thanks to Amit and Kheema, for believing and giving me wings to fly. Also my hearty feelings towards cute Amulya for loving me so much.

To my parents, thank you for pushing me to excel in everything that I do. They have been so loving and supportive in countless ways. The sacrifice they have made from the very first day of my journey of life. I do not want to sum up their love and my feelings for them in words and it is to them that I dedicate this work.

Dated: **(Anju Bhalla)**

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS USED

T-DNA	Transfer Deoxyribonucleic Acid
OsPE	*Oryza sativa* Polyembryony (mutant)
RB	T-DNA Right Border
LB	T-DNA Left Border
hpt	Hygromycin Phosphotransferase (gene)
T-RB	T-DNA Right Border
T-LB	T-DNA Left Border
AP	Adaptor Primer
AD	Arbitrary Degenerate Primer
SPT	Streptomycin Phosphotransferase
NPT	Neomycin Phosphotransferase
F_1	First Filial Generation
F_2	Second Filial Generation
F_3	Third Filial Generation
PCR	Polymerase Chain Reaction
TAIL-PCR	Thermal Asymmetric Interlaced Polymerase Chain Reaction
iPCR	Inverse Polymerase Chain Reaction
RT-PCR	Reverse Transcriptase –Polymerase Chain Reaction
Ds	Dissociation Element
Tos17	Transposon *Oryza sativa*
RAP-DB	The Rice Annotation Project Database
KEGG	Kyoto Encyclopedia of Genes and Genomics
IRFGC	International Rice Functional Genomics Consortium
IRGSP	International Rice Genome Sequencing Project, 2005
SSR	Simple Sequence Repeats
BSA	Bulk Segregant Analysis
FST	Flanking Sequence Tags
EST	Expressed Sequence Tags

EHA101	*Agrobacterium tumefaciens* strain
$Hm^R Ds$	Plasmid
PAU	Punjab Agricultural University, Ludhiana
NCBI	National Center for Biotechnology Information
BLAST	Basic Local Alignment and Search Tool
EDTA	Ethylene Diamine Tetraacetic Acid
α-P^{32}dCTP	Radiolabelled ^{P32} Cytosine Triphosphate
TBE	Tris Borate EDTA
TAE	Tris Acetate EDTA
dNTPs	Nucleotide Triphosphates
TE	Tris EDTA
DEPC	Diethyl Pyrocarbonate
SSC	Saline Sodium Citrate
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
KCl	Potassium Chloride
$MgCl_2$	Magnesium Chloride
CTAB	Cetyl Trimethyl Ammonium Bromide
SDS	Sodium dodecyl sulfate
PAGE	Poly-acrylamide Gel Electrophoresis
EtBr	Ethidium Bromide
mRNA	Messenger Ribonucleic acid
cDNA	Complimentary DNA
DNase	Deoxyribonuclease
RNase	Ribonuclease
PMC	Pollen Mother Cell
ppm	Parts Per Million
T_m	Melting Temperature
bp	Base Pair
cM	Centimorgan
%	Percentage

χ^2	Chi Square
df	Degree of Freedom
µg	Microgram
µm	Micrometer
mm	millimetre
Fig.	Figure
mg	Milligram
ng	Nanogram
ml	Millilitre
mM	millimole
°C	Degree Centigrade
U	Units
v/v	Volume/Volume
w/v	Weight/Volume
sec	Second
h	Hour
min	Minute
g	Gram
l	Litre
M	Molarity
µ	micron
α	Alpha
β	Beta

CHAPTER 1

INTRODUCTION

1. Introduction

Rice is life of over two billion children, women and men on our planet. Being the most important cereal crop it attracted major attention in functional genomics. It has become a model for monocot plants because of its relatively small genome size among cereals (389Mb); ease of transformation; availability of detailed physical and dense molecular marker maps, existence of similarities in gene sequence and gene function among all cereals and grasses as well as its economic importance. Any understanding of rice genome is directly applicable to other cereals. With 370Mb (95%) of complete sequence available, assigning a function to unknown or predicted genes becomes the major task of functional genomics (Delseny *et al.*, 2001; Jeon *et al.*, 2000; An *et al.*, 2003). Knockout or gain of function by a gene by insertional mutagenesis using T-DNA or transposon with "single copy insertion" is of great value to address its function through the mutant phenotype (Jeong *et al.*, 2006; Itoh *et al.*, 2007).

The relationship of the genetic mechanisms regulating the development and function of polyembryony in plants is largely unknown. It may be due to unique set of genes and mechanisms that regulate polyembryony. Although polyembryony is wide spread but, it is infrequent in higher plants. It represents a fascinating and unusual example of embryogenesis gone awry. Polyembryony has been described in several crop and cereal plants, but the underlying genetic causes of twining in such cases have been difficult to resolve. The recovery of polyembryony in insertional mutant may prove to be a useful tool for understanding and developing apomictic seed production in rice hybrids.

In addition it will provide a powerful strategy for understanding the genetic mechanisms underlying plant embryo development.

Apomictically produced plants are genetically identical to the female parent. Therefore, if apomixis can be introduced in crop plants it would provide an inexpensive way to perpetuate a given genotype and maintenance of heterosis through successive seed generations. Apomixis would also simplify commercial hybrid seed production because isolation would not be necessary to produce F_1 or maintain and multiply parental lines. Apomixis was once regarded as an evolutionary dead end because the meiotic segregation and recombination on the female side is eliminated. However, now it is being regarded as a potential powerful genetic tool for use in crop improvement. Inspired by the potential benefits of apomictic reproduction, plant breeders are making serious attempts to introduce this trait into crop plants (Khush *et al.*, 1994).

Due to enormous potential economic and social benefits of apomixis for plant breeding and seed production various attempts were undertaken to transfer genes conferring apomixis from natural apomicts into their sexual relatives (Naumova, 1997). However, apomictic maize lines have not yet been obtained (Savidan, 2001). Some of the most extensive attempts to introgress apomixis into *Pennisetum glaucum* have been carried out with the tertiary gene pool relatives of pearl millet (Dujardin and Hanna, 1989).

A rice polyembryonic seedling mutant (Ximei *et al.*, 2006) has been isolated in the M_2 generation of IR36-4-4X rice following 25 KeV-N ion beam treatment. Another example of polyembryony in rice has been reported in APIV by Lu and Liu (2002) in which genetic difference for the structure of embryo sac was found in a single panicle.

Fertilization and development of the poly egg apparatus resulted in polyembryony in APIV. No genetic mechanism for multiple embryo formation has been described. Apomixis, asexual reproduction of seeds can revolutionize agriculture if its genetic basis can be elucidated. However, the genetic control of natural apomixis has remained obscure until quite recently, owing to all the complications of Mendelian genetics, such as epistatic gene interactions, components that are expressed sporophytically and gametophytically, expression modifiers, polyploidy, aneuploidy, segregation distortion, suppressed recombination, etc. that seem to have accumulated during the evolution of apomixis (Ozias-Akins and van Dijk, 2007).

Insertional mutagenesis has provided a more rapid and direct way to clone a gene. As the sequence of the T-DNA inserted element is known, the gene into which it is inserted can be recovered by means of genome walking, TAIL-PCR, iPCR approaches. Thus, both forward and reverse genetics approaches can be used to identify gene function in mutants. Chengkun *et al.* (2007) reported an efficient procedure to construct an indexed and region specific insertional mutant library of rice. The procedure makes use of efficient long PCR based high throughput indexing coupled with a random but anchored population of *Ds* transposants. Latest gene constructs by Upadhyaya *et al.* (2006) are the most suited for producing single copy insertions. Use of such T-DNA/*Ds*/*Tos*17 lines will allow individual groups to concentrate on particular chromosomes or regions within the chromosomes and perform localized insertional mutagenesis.

The genome sequencing of rice has facilitated large scale analysis of gene functions. Mutants for functional genomics studies have been produced by several

groups. To provide easy access to such resources, The Rice Annotation Project Database (RAP-DB) has integrated the mutant information created by ten independent groups compiled under SALK Database. All the flanking sequences that were generated by *Tos*17, T-DNA and *Ds* insertions were compared with the rice genome sequence and the positions of genes disrupted by different methods were simultaneously displayed in the RAP-DB update 2008. In order to facilitate functional analysis of rice genes around 141,254 FSTs are available with T-DNA; 46,083 from T-DNA with enhancers (Nobuta *et al.*, 2007; Hsing *et al.*, 2007), 13,309 with *Ac/Ds,Spm/dSpm* (Kumar *et al.*, 2005) and 17,937 FSTs with *Tos17* (Miyao *et al.*, 2003, 2007).

There is an urgent need for collaboration in building tools and resources, especially for assembling a set of lines with mutations in all the predicted 40,000 genes together with integrated databases containing all the relevant information about each gene. Towards this goal, an International Rice Functional Genomics Consortium (IRFGC) has been formed to provide a much needed platform for information and resource sharing.

Finally, although rice is an important food crop in its own right, it is also a model for other cereals. Discoveries in rice can be applied to other cereals such as maize, wheat and barley. With sequencing project nearly completed in maize and taken up in wheat, the functional genomics findings in rice will assist gene discovery and breeding in other cereals using the power of comparative genomics (Upadhyaya, 2007).

In this study polyembryonic mutant (*OsPE)* represents the first reported example of high frequency and heritable polyembryony in rice caused by insertional mutagenesis. Keeping the above information in mind, the present investigation is aimed for

characterization of T-DNA insertional mutants in Basmati 370 (Dhaliwal *et al.*, 2001) and molecular mapping and cloning of *OsPE* mutant with the following objectives:

1. Phenotypic screening of insertional mutants.
2. To confirm insertion in selected mutants.
3. To analyse the copy number.
4. To map *OsPE* locus using BSA approach.
5. To identify and characterize T-DNA flanking gene in Basmati 370.
6. Homolog search using FSTs and other rice databases.
7. Expression of *OsPE* gene by RT-PCR.
8. To clone *OsPE* gene in Basmati 370.

CHAPTER 2

LITERATURE REVIEW

2. Review of Literature

Insertional mutagenesis is one of the most useful methods of reverse genetics for analyzing gene function. When foreign DNA is inserted into a gene, it not only creates a mutation but also tags the affected gene, facilitating its isolation and characterization (Azpirez-Leehan and Feldmann, 1997). Hybrid breeding for exploiting heterosis is expected to overcome the existing yield barriers in cereals including rice (Khush, 1994). The recovery of a multiple seedlings insertional mutant due to polyembryony may prove to be a useful tool for understanding and developing apomictic seed production in rice hybrids.

The relevant literature on the present study has been reviewed under the following heads.

1. Rice as staple food and a model crop
2. Basmati rice
3. Polyembryony and its types
4. Polyembryonic mutants
5. Apomixis
6. Mapping strategy
7. Insertional mutagenesis in Rice
8. *Agrobacterium*- mediated gene transformation
9. Screening of T-DNA insertional mutants
10. Pectin methylesterase

2.1 Rice as staple food and a model crop

Rice remains the most important staple food on the planet since it feeds roughly half the population on a daily basis. Approximately 750 million of the world's poorest people depend on it for survival. According to FAO, the global rice requirement in 2025 will be of the order of 800 million tonnes. At the moment, the production is less than 600 million tonnes. It is predicted that world will need an additional 50 million tonnes of rice annually from 2006-2015 (about 9% of the current production) to meet the expected requirement. Rice occupies a pivotal role in indian agriculture. It is the staple food for more than 70 percent indians, and a source of livelihood for 120-150 million rural households. In India rice is cultivated round the year in one or other part of the country, in diverse ecological situations over 44 million hectare with a production of around 91 million tonnes in 2006-2007 (Rice Today, IRRI 2006), representing the largest area and the second highest production in the world. At the current rate of population growth of 1.5 percent, rice requirement by 2020 would be around 110 million tonnes. Therefore, a three pronged strategy will have to be promoted for ensuring the security of adequate rice availability; enhancement of productivity, profitability and sustainability of rice yield. The use of new technologies including biotechnology and genetic engineering becomes important in this context.

Rice is a model plant among cereal crops because of its small genome size, efficient plant transformation, construction of physical maps, large-scale analysis of expressed sequence tags (ESTs), and extensive synteny shared with other cereals, availability of large number of FSTs and economic importance. Completion of the genome sequence of rice has provided new reverse genetics tools for assigning function

to genome sequences (Kumar and Hirochika, 2001; Feng *et al.*, 2002; Sasaki, 2003; Yu *et al.*, 2005).

2.2 Basmati Rice

Basmati, the unique aromatic quality rice is a nature's gift to indian sub-continent. Basmati rice is known to have originated in undivided India and Pakistan Punjab. Epicureans acclaimed its delightful fragrance, taste and texture which makes it the best among the aromatic rice of the world. Typically the delicately curved, long grained, highly aromatic rice which elongate and cook soft and fluffy were the ones which were categorized as Basmati and enjoy privileged treatment both in domestic and international markets commanding three times higher price. In the export markets still the traditional tall Basmati varieties, Taroari Basmati followed by Basmati 370 and Type3 (Dehraduni) have maintained supremacy over other varieties due to their exclusive quality features.

Basmati cultivation is confined to northwest indian states- Punjab, Haryana, Uttarakhand and western Uttar Pradesh and to a limited area in Hiamchal Pradesh, Jammu & Kashmir, Delhi and Rajasthan. Among the four Basmati producing states about 35% of total Basmati area is in Haryana, 17% in Uttarakhand, 3-4% and 6-8% of total rice area in Uttar Pradesh and Punjab, respectively (Rani, 2005).

Basmati 370 released in 1933 by the Rice Experimental Station Kala Shah Kaku (now in Pakistan) is the oldest documented Basmati rice variety. Basmati 370 is photoperiod sensitive and flowers in the first week of October and matures in the first week of November. Basmati 370 possesses long slender clear translucent grains with strong pleasant aroma. Cooked rice is almost of double the length of uncooked raw

milled rice, non sticky, moist and tender to eat with excellent palatability. Average yield of Basmati 370 is about 12 qt of paddy per acre.

The most important thrust areas in Basmati rice are the multiple crosses involving traditional Basmati types to develop better Basmati genotypes, development of hybrids in Basmati, marker assisted selection to combine Basmati grain with *sd1* dwarfing gene, photoperiod insensitivity, bacterial blight resistance and desirable amylose content (Singh *et al.*, 2001; Sanchez *et al.*, 2000; Jena *et al.*, 2005; Sharma *et al.*, 2005(a); Singh *et al.*, 2007; Cheema *et al.*, 2007; Khush and Singh, 2007).

2.3 Polyembryony and its types

Polyembryony has been defined as the occurrence of more than one embryo in a seed. However the additional embryos do not always mature. Therefore, if mature seeds are taken into account the percentage of polyembryony in a species would be far less than its actual frequency. Therefore, polyembryony includes all the actual occurrence of two or more proembryos or embryos in a developing ovule. Except for a few taxa (*Citrus*, *Mangifera*), polyembryony occurs only as an abnormal feature. About 244 species from 140 genera belonging to 59 angiosperm families are reported to exhibit polyembryony. Polyembryony in angiosperms may arise by:

2.3.1 Cleavage of proembryo

Cleavage and proliferation of zygote or its derivatives leading to the establishment of separate embryonal primordial is widespread among gymnosperms. In angiosperms this feature is less frequent. Among angiosperms cleavage polyembryony is

quite common in orchids. In *Eulophia epidendraea*, Swamy (1943) recorded three different modes of supernumerary embryo formation:

1. The zygote divides irregularly to form a mass of cells of which those lying toward the chalazal end grow simultaneously and give rise to many embryos.
2. The proembryo forms small buds or outgrowths which may themselves function as embryos.
3. The filamentous embryo becomes branched, and each branch forms an embryo.

While cleavage polyembryony arising during seed development is known in many orchids, the formation of plural embryos during seed germination is known only in *Vanda*. (Rao, 1965). Suspensor polyembryony is a common feature in the genus *Exocarpus*. As many as six embryos may develop simultaneously in an ovule by the proliferation of the suspensor cells. Eventually, however, only one of them takes the lead and reaches maturity. Rarely, two germinable embryos may be formed in a seed.

2.3.2 Formation of embryos by cells of the embryo sac other than the egg

In this category the most common source of additional embryo is the synergid. Depending on whether it arises from fertilized synergid or unfertilized synergid, the embryo may be diploid or haploid. In *Aristolochia bracteata*, *Poa alpine* and *Sagittaria graminea* besides the egg and the polars, one or both the synergids may get fertilized. This can be brought about either by the entry of more than one pollen tube into the embryo sac or by the presence of additional sperms in the same pollen tube. In this situation the zygotic as well as the synergid embryos are diploid. Embryos

arising from unfertilized synergids are known in *Argemone mexicana* and *Phaseolus vulgaris*. Formation of embryos from antipodals is rather rare. It has been observed in *Paspalum scrobiculatum, Ulmas americana* and *U. glabra.* The antipodal cells may divide a few times to form proembryo- like structures. However, they fail to grow into an adult embryo and there is no report of antipodals forming germinable embryos. Most of the reports concerning endosperm cells forming embryos are doubtful. However, *Bracheria setigera*, an apomictic species, is the only example where endosperm has been reported to produce triploid embryos and seedlings (Muniyamma, 1978).

2.3.3 Development of more than one embryo sac within the same ovule

Multiple embryo sacs in an ovule may arise from

1. Derivatives of the same megaspore mother cell,
2. Derivatives of two or more megaspore mother cells, or
3. Nucellar cells.

Formation of twin embryo sacs within an ovule is known in *Casuarina equisetifolia*, *Citrus* and *Poa pratensis*. In *Pennisetum ciliare* 22 percent seeds contain twin embryos. The normal embryo sac develops only up to the 4- nucleate stage, and the multiple embryos are formed by aposporous embryo sacs. The members of the family Loranthaceae lack a conventional ovule. Numerous embryo sacs develop concurrently in the same ovary and their tips, carrying the egg apparatus; grow up to various heights in the style. After fertilization, the embryos grow downward and enter the composite endosperm in the ovarian cavity. All

embryos but one collapse during seed development, resulting in monoembryonate seeds. Occasionally however, two or more embryos may mature. In *Scurrula pulverulenta* about 2 percent seeds are polyembryonate (Bhojwani, 1968). In such seeds there are either two fully developed green embryos, or there are more then two embryos, of which only one is well developed, and the other under-developed and non- green. The only exception in this family is *Struthanthus vulgaris* which has only one embryo sac and thus does not exhibit polyembryony.

2.3.4 Activation of some sporophytic cells of the ovule

The embryos arising from the maternal sporophytic tissues (outside the embryo sac) are called adventive embryos. The only maternal tissues which are known to form adventive embryos are the nucellus and the integuments. Besides the more popular examples of *Citrus* and *Mangifera*, nucellar polyembryony occurs in *Opuntia dillenii* and *Trillium undulatum*. Some species of *Citrus* are monoembryonate (*C. grandis, C. limon*) while others are polyembryonate (*C. microcarpa, C. reticulate*). Seeds with as many as 40 embryos have been reported in *C. unshiu*. In polyembryonate species the adventive embryos arise by the proliferation of the nucellar cells. With rare exceptions (*Trillium undulatum*) nucellar embryos arise from the micropylar half of the nucellus. In *Mangifera* the nucellar cells destined to form adventive embryos can be distinguished from other cells of the nucellus by their dense cytoplasm and starchy contents. The adventive embryos do not show synchronous development. A single seed may show embryos at various stages of development.

2.4 Polyembryonic mutants

2.4.1 Rice

A polyembryonic seedling mutant (Ximei *et al.*, 2006) has been isolated in the M_2 generation of the IR36-4-4X rice following 25 KeV-N ion beam treatment. The major agronomic properties and the polyembryonic seedling frequency of the mutant line (named IR 36-Shuang) were investigated from 2003 to 2005. The plant height, the panicle length and 1000 grain weight of IR36- Shuang were lower than those of its control by 35.41%, 5.06% and 15.72%, respectively. Moreover the seed setting percentage decreased by 12.39% compared with that in normal IR36-4X plants. The average percent of the polyembryonic seedlings in the IR36-Shuang line was significantly higher than that in the control, IR36-X with an average of 13.24 %. The process of embryogenesis of IR36-Shuang was generally identical with that of the normal rice. However, a certain rate of additional embryo-like structure could be found in the embryo sacs of IR36-Shuang. The additional embryo in IR36-Shuang might arise from the double set of embryo sacs in a single ovary, antipodal cells or endosperm cells. Evidence from the embryological studies confirms that the additional embryos were derived from cells that may not involve the events of fertilization that defines sexual embryo development. Recent study has indicated that apomixis is genetically controlled and it is likely that many of the genes controlling sexual development are also responsible for the induction of apomixes (Spillane, 2001).

A spontaneous mutant, polyembryonic rice APIV, was found in which there were more than one egg cells in the embryo sac, in some florets (Mu *et al.*, 1998). The

number of egg cells were different among different embryo sacs, in some florets. In this poly egg apparatus embryo sac, three eggs were in majority. Besides the poly egg apparatus embryo sacs, a new variant embryo sac referred to as 'double set of embryo sac', was also found in APIV. The egg cells were far from micropylar end in the 'double set of embryo sac'. According to genetic analysis, gametophytic genotypes, rather than sporophytic ones probably controlled the production of polyeggs. Genetic polymorphism for the structure of embryo sac was found in a single panicle of APIV. Fertilization and development of the poly egg apparatus resulted in polyembryony in APIV.

2.4.2 *Arabidopsis*

Embryo-defective mutant of *Arabidopsis* exhibits frequent polyembryony following embryogenic transformation of the suspensor (Daniel and David, 1994). Supernumerary embryos of *twin* mutant arise through a novel mechanism, transformation of cells within the suspensor, a differentiated structure established early in embryogenesis. Transformed suspensor cells appear to duplicate the patterns of cell division and developmental pathways characteristic of zygotic embryogenesis. The polyembryonic mutant embryos exhibit a number of developmental defects, including irregular patterns of cell division and abnormal morphology.

The TWIN locus therefore, appears to be required for normal development of the embryo proper as well as suppression of embryogenic potential in the suspensor. The development of viable embryos in *twin* demonstrates that cells of the *Arabidopsis*

suspensor can successfully establish embryonic polarity and complete the full spectrum of developmental programmes normally restricted to the embryo proper.

The *twin* phenotype indicates that disruption of a single genetic locus can result in the conversion of a single terminally differentiated cell type to an embryonic state.

Another *twn2* mutant of *Arabidopsis* exhibited an early defect in embryogenesis in which the apical growth of the zygote was arrested after one or two divisions (Zhang and Somerville, 1997). The basal cell that normally gave rise to the suspensor proliferated abnormally and gave rise to multiple embryos. The *twn2-1* mutation, which is the only known allele, was caused by a T-DNA insertion in the 5' untranslated region of a putative valyl-tRNA synthetase gene, *val*RS.

The insertion causes reduced transcription of the *val*RS gene reproductive tissues and developing seeds. Expression analysis suggested that enhancer elements in the first two introns interact with the border of the T-DNA to cause the altered pattern of expression of the *val*RS gene in the *twn2* mutant. The polyembryonic phenotype resulted from arrested development of the apical cells, followed by abnormal proliferation of the basal cells to produce secondary embryos.

2.5 Apomixis

Apomixis or asexual reproduction through seeds refers to a method of reproduction in which the embryo develops from an unfertilized egg and results in a progeny genetically identical to the mother plant. Apomixis may be obligate or

facultative, depending on whether seed production is entirely asexual or in part sexual. (Richards, 1986; Bashaw and Hanna, 1990).

2.5.1 Inheritance of apomixis in Gramineae

2.5.1.1 *Brachiaria*

Both sexual diploid and apomictic tetraploid cytotypes exist in *Brachiaria brizantha*. Rare sexual diploid accessions of *B. brizantha* were used for chromosome doubling by *in vitro* techniques for genetic studies with tetraploid apomicts (Pinheiro *et al.*, 2000). None of the induced tetraploids showed any evidence of aposporous embryo sac formation (Araujo *et al.*, 2005). Molecular mapping of apomixis has not been as extensive in *Brachiaria* as in some other grass species. The approach used included crosses between induced 4X sexual *B.ruziziensis* and apomictic 4X *B. brizantha* (Pessino *et al.*, 2001). Using probes mapped as RFLPs on maize and rice plus anonymous AFLP markers, apomixis was found to be linked with 2 of 46 AFLP markers plus 4 maize, 2 rice and 1 oat marker (Pessino *et al.*, 1998). Maize markers showed synteny with maize chromosome 5 and rice chromosome 2 (www.gramene.org). The recent focus then shifted from mapping of the trait to analysis of gene expression. From 39 primer combinations that amplified 2500 fragments, 31 were indicative of differential expression and two were confirmed to be specific to the apomictic plants in amplified cDNA. Neither of these two fragments however showed segregation in the mapping population and thus linkage with mode of reproduction could not be determined. (Leblanc *et al.*, 1995; Rodrigues *et al.*, 2003)

2.5.1.2 *Panicum*

In *Panicum* mapping approach was used (Ebina *et al.,* 2005). Apomixis co-segregated with 9 AFLP markers at 1.4cM but when same cluster of markers applied to *Pennisetum*, *Cenchrus* and *Paspalum*, it suggested that gene(s) for apomixis reside in chromosomal region where recombination is low. Gene expression in apomictic *Panicum* ovules has also been studied (Chen *et al.,* 2005; Chen *et al.,* 1999).

2.5.1.3 *Paspalum*

In *Paspalum* tetraploid cytotypes are apomictic whereas diploid cytotypes are sexual. A segregation ratio of 1 apomictic to 2.8 sexuals observed in *P. notatum* (Martinez *et al.,* 2001) does not clearly fit any model of inheritance where apomixis is dominant. Furthermore, many markers in this group (7AFLPs, 2RAPDs, and 2RFLPs out of 19) did not recombine with the trait, indicating a suppression of recombination in the region of the genome that controls apospory (Martinez *et al.,* 2003; Stein *et al.,* 2004). The suppression of recombination may be due to a chromosomal rearrangement.

Molecular genetic analysis in *Paspalum simplex* indicated that the apomixis controlling locus (ACL) was syntenic with a portion of rice chromosome 12 (Pupilli *et al.,* 1997). The degree of synteny was similar in apomictic *P. malacophyllum* but different in *P. notatum*, where the distal markers were replaced by markers rice chromosome 2 (Pupilli *et al.,* 2004), suggesting that translocation has occurred with in the species. By extrapolating from rice data, the size of the ALC in *P. notatum* was estimated to be 8-10cM in rice but in all three species of *Paspalum* recombination was suppressed.

Extensive AFLP screening has been conducted in *P. simplex*. FISH and BAC clones containing the hemizygous marker resulted in the distal region of chromosome 12. Based on gene sequence discovered however, numerous rearrangements due to transposable element insertions or deletions led to the conclusion that the apomixis-associated genes probably lacked coding capacity.

2.5.1.4 *Pennisetum*

The genus *Pennisetum* contains the sexual diploid forage and grain crop, pearl millet (*P. glaucum*), the sexual tetraploid biomass crop (*P. purpureum*) and a large number of apomictic tertiary gene pool species, some of which are grown as ornamental crops. One apomictic relative *P. ciliare* is also known as *Cenchrus ciliaris*. Some of the most extensive attempts to introgress apomixis into a crop species have been carried out with tertiary gene pool relatives of pearl millet. The crossing data led to the conclusion that apomixis is dominant, inherited and heterozygous (Dujardin and Hanna, 1989). To map the trait of apomixis, RFLPs or SCARs (Sequence characterized Amplified Regions) developed from sequenced RAPD markers were identified in apomictic parent (Ozias-Akins *et al.*, 1998; Ozias-Akins *et al.*, 1993). Mapping with these targeted markers showed that none of them recombined with apomixis out of a population of almost 400 individuals. The cytogenetic map generated with the *Pennisetum* BACs showed that markers were distributed throughout the distal region of the short arm of the ASGR (Apospory Specific Genome Region). Carrier chromosome appears to contain an insertion or duplication that results in a measurable difference in size when compared with its homologous or homeologous partners (Akiyama *et al.*, 2004; Akiyama *et al.*, 2006).

The search for genes differentially expressed between ovaries from sexual and apomictic plants of *C. ciliaris* has yielded two genes (*Pca21*) and (*Pca24*). Transcripts from both could be detected by northern blot analysis of ovary tissue much more abundantly in apomicts then sexuals (*Pca21*) or exclusively in apomicts (*Pca24*). *In-situ* hybridization also demonstrated expression of both genes in embryo sacs. Preliminary linkage analysis indicated that neither gene was located in the ASGR.

2.5.1.5 *Poa*

The apomixis associated trait of parthenogenesis rather than unreduced embryo sac formation, has most frequently been mapped in *Poa*. Parthenogenetic and non parthenogenetic bulks allowed the identification of AFLP and RAPD markers that appeared to be linked to the trait (Barcaccia *et al.*, 2007). Parthenogenesis is only a part of the apomixis process and may be linked or unlinked with reduced embryo sac development. A complex genetic model for mode of reproduction was proposed and tested in *Poa* (Matzk *et al.*, 2004) using the Flow cytometric and seed screen (FCSS) (Matzk *et al.*, 2000). Four gene models that included apospory initiator (*Ait*) gene and apospory preventer (*Apr*) gene as well as parthenogenesis initiator (*Pit*) and parthenogenesis preventer (*Ppv*) genes. It was concluded that incomplete penetrance of *Ait* was responsible for the erroneous classification. A fifth gene that affected megaspore development (*Mdv*) was postulated to account for significant female sterility in such plants. Albertini *et al.* (2001) concluded that apospory and parthenogenesis were linked. Further gene expression pattern was examined in detail in sexual, apomictic and recombinant genotypes (Albertini *et al.*, 2005). None

displayed expression confined to either the apomictic or sexual genotype rather than expression patterns were differentially modulated in each mode of reproduction.

2.5.1.6 *Tripsacum*

Tripsacum dactyloides is diplosporous and form eight nucleate embryo sac from MMCs that either do not enter meiosis or exit meiosis through first division restitution. The heterochronic exit from meiosis is proposed to be a key feature of diplospory that distinguishes it from apospory, where a change in non-sporogenic nucellar cell fate occurs (Gramanelli *et al.,* 2003). Inheritance of apomeiosis in *Tripsacum* is consistent with the segregation of a single gene. Gramanelli *et al.* (1998) showed that syntenic chromosomal regions of diploid sexual *Tripsacum* and maize displayed much more recombination among molecular markers than the same region in tetraploid apomictic *Tripsacum*. Six of the 25 RFLP probes distributed across the genome identified markers significantly linked with fertility on 6 linkage groups. One of these linkage groups was *Tripsacum* (*Tr16*) which is syntenic with maize chromosome 6.

2.5.2 Genetic control of Apomixis

The independent control of apomeiosis and parthenogenesis is confirmed in several species *e.g.* in *Allium*, *Poa*, *Erigreron*, *Taraxacum*, *Hieracium*, *Hypericum* and *Potentilla*. Whether apomixis is indeed under genetic control has been questioned by Carman's duplicate-gene asynchrony hypothesis (Carman 1997), which implies epigenetic rather than genetic control and requires a genome duplication event through either polyploidization or paleopolyploidization. The relationship between

gametophytic apomixis and polyploidy is striking. Heritable epigenetic changes are based on covalent modifications of nucleotides such as methylation and chromatid remodeling are known to play a role in autonomous endosperm development (Lohe and Chaudhary 2002).

Apomixis is a developmentally complex trait as reflected in its genetics. Several genetic complications seem to be accumulated in explaining inheritance of apomixis. Epistatic gene interactions, components that are expressed sporophytically and gametophytically, expression modifiers, polyploidy, segregation distortion and suppressed recombination. Despite additional complications such as low pollen fertility, interspecific crosses, and genome instability that occur in some species, a large number of informative genetic studies have been conducted. In general apomixis genes are usually dominant (Ozias-Akins *et al.*, 2007).

The use of genetic markers and new phenotyping methods have been instrumental in unravelling the genetics of apomixis. Genes that influence the penetrance of apomixis have not yet been genetically characterized. Comparative sequencing of apomixis-specific chromosomal region and their sexual homologs may help guide apomixis gene identification. Once apomixis controlling sequences have been identified, the task of introducing them into sexual genotypes while retaining function will be a technological challenge for the future.

2.6 Molecular Mapping

The construction of saturated linkage maps in rice have made it possible to map and tag genes of economic importance with molecular markers (Mohan *et al,*

1997; Kumar, 1999). Most of the initial gene annotations have been done using map based cloning. However not many reports are there for T-DNA insertion chromosomal location using BSA.

2.6.1 Molecular markers

The DNA sequence that frequently varies in length between individuals of a species, which can be seen after cleavage by a restriction enzyme or *in vitro* DNA amplification by polymerase chain reaction. Molecular genetic markers exhibit simple Mendelian inheritance making it convenient to employ them in mapping genome and tagging genes of importance.

Molecular markers are scored as the presence or absence of a DNA fragment. It represents only one allele at a given locus and therefore cannot distinguish the heterozygotes from homozygotes, whereas co-dominant markers are scored by detecting the alternate fragment of the marker allele making it possible to distinguish heterozygotes from homozygotes. Molecular genetic markers have several advantages as compared to the biochemical and morphological markers such as high abundance, high degree of polymorphism, co-dominant nature in most of the cases, no pleiotropic effect, and independent of the environment and phenology of the crop and thus are useful for precision plant breeding (Melchinger, 1990). The analysis of DNA sequence variation is of major importance in genetic studies and molecular markers are a useful tool for assaying genetic variation (Balyan *et al.,* 2005; Lakshmi *et al.,* 2002). A variety of molecular genetic markers, including restriction fragment length polymorphism (RFLPs); random amplification of polymorphic DNAs (RAPDs)

(Singh *et al.*, 2006 a); amplified fragment length polymorphisms (AFLPs) (Prashanth *et al.*, 2002) and microsatellites or simple sequence repeats (SSRs) have been developed in different crop plants (Philips and Vasil, 2001; Varshney *et al.*, 2004; Sharma *et al.*, 2005(b); Hossain *et al.*, 2004; Kalavacharla *et al.*, 2006; Kumar *et al.*, 2007). Among these, SSR markers are more useful for a variety of applications in plant breeding.

2.6.2 Microsatellites or SSRs and molecular map of rice

SSRs are usually 1-6 bp long and embedded in DNA with unique sequences. Therefore, SSRs can be amplified *in vitro* using the polymerase chain reaction with primers developed according to the flanking region of the locus. The resulting PCR product possesses electrophoretic mobility that differs according to the number of repeated DNA units in the SSR allele(s) present. SSR markers are useful for a variety of application in plant breeding and genetics because of their reproducibility, multi allelic nature, co-dominant inheritance, relative abundance and dense genome coverage. Microsatellite markers are one of the most widely used DNA markers for mapping and tagging of genes controlling important agronomic traits (Röder *et al.*, 1998). SSR markers have been extensively used for developing comprehensive linkage map in rice. Molecular maps are important to know the location of genes corresponding to phenotypical traits and their usefulness (Singh *et al.*, 2006 (b); Zhu *et al.*, 2005). The first molecular genetic map of rice was developed at Cornell University by McCouch *et al.* (1988) and an RFLP map by Saito *et al.* (1991). In rice around 2240 microsatellite marker (SSR) loci were developed and experimentally

validated by McCouch *et al.* (2002), in addition to the previously mapped 500 SSR markers (Panaud *et al.*, 1995; Cho *et al.*, 2000; Temnykh *et al.*, 2000, 2001 and Chen *et al.*, 2002, 2007). Recent advances in the construction of saturated linkage maps in rice have made it possible to map and tag genes of economic importance with molecular markers.

2.6.3 Bulk Segregant Analysis (BSA)

Michelmore *et al.* (1991) described BSA for identifying markers linked to the target gene or genomic region. BSA is rapid procedure to identify markers in specific regions of the genome. The method involved comparing two pooled DNA samples of the individuals from a segregating population originating from a single cross. Within the bulk the individuals are identical for a trait. Two such bulks contrasting for a particular trait are analysed to identify markers that distinguish them. Markers that are polymorphic between the pools will be genetically linked to the loci and used for generating detailed genetic maps for several species. It has simply been done by analyzing the segregation of randomly selected molecular markers in single segregating population.

During the last decade different technologies such as Restriction Fragment Length Polymorphism (RFLP), Random Amplified Polymorphic DNA (RAPD), Sequence Tagged Sites (STS), Amplified Fragment Length Polymorphism (AFLP), Simple Sequence Repeats (SSR) or Microsatellites have been developed and applied to various crop species. In the last few years newer approaches such as Single Nucleotide Polymorphism (SNP), Expressed Sequence Tags (ESTs) and Flanking

Sequence Tags (FSTs) have been widely used for various applications. Extensive efforts so far have revealed the function of only a handful of rice genes and most of these have been identified through laborious map based cloning.

2.7 Insertional mutagenesis in rice

Rice lines are being generated using several experimental approaches to randomly tag genes by DNA insertional elements. Such a mutant population with gene knockout and over expressions will prove to be highly useful in gene annotation based on phenotypes. DNA elements that can insert randomly within chromosome to disrupt gene function include the T-DNA of *Agrobacterium tumefaciens*, transposons *Ds*, *dSpm* and the *Tos17* retrotransposon (Nobuta *et al.*, 2007; Jeong *et al.*, 2006; An *et al.*, 2005; Sallaud *et al.*, 2004; Hsing *et al.*, 2007). So far 172,500 FSTs have been generated (Jung *et al.*, 2008). Using this database of FSTs 27,551 (48%) of 57,142 rice loci (given by IRGSP and TIGR) contain insertions in the genic region or 5' untranslated regions of rice genome.

Another type of insertional mutagenesis lead to gain of function. Such 'activated tagged' lines carry strong enhancers near one end. These tags can activate genes within a few kilobases in chromosomal DNA. These tags can activate genes that are located up to 10Kb away from the enhancer sequence and are independent of the direction of transcription, resulting in increased expression of the targeted gene.

As biological data of rice continue to increase, the Rice Annotation Project Database (RAP-DB) has supplied the most up-to-date information (Jung *et al.*, 2008). IRGSP has released 77,763 flanking sequence tags generated by 10 independent

functional genomics groups (NIAS; *Tos17*), (CGT; *Ds*), (GSND; *Ds*), (NUS; *Ds*), (UCD; *Ds*), (CSIRO; T-DNA), (Genoplante; T-DNA), (POSTECH; T-DNA), (RMD; T-DNA), (TRIM; T-DNA), (Goff, 1999; Hiei *et al.,* 1994; Shimamoto and Kyozuka, 2002; Kellogg, 2001; Londo *et al.,* 2006; Khush, 1997; Ouyang *et al.,* 2007; Sasaki *et al.,* 2002; Yuan *et al.,* 2005; Itoh *et al.,* 2007; Ohyanagi *et al.,* 2006; Zhang *et al.,* 2006; IRGSP, 2005).

These combined total of the five mutant populations listed in table 2.1 are approximately 5,00,000 lines, with one to ten copies of mutations per line. If the insertions were evenly distributed across the genome, this FST collection would be predicted to the target insertion in 99% of the rice loci.

Table 2.1 Summary of rice insertional mutant resources

	Mutagen				
	T-DNA	*Ac/Ds, Spm/dSpm*	T-DNA with enhancer	FOX system	*Tos17*
Method	*Agrobacterium*	*Agrobacterium*, crossing or selfing	*Agrobacterium*	*Agrobacterium*	Tissue culture
Copies per line	1-2	1-7	1-2	1	5-10
Distribution	Genic regions preferred	Genic regions highly preferred	Genic regions preferred	Not analysed	Genic regions preferred
FSTs	141,254	13,309	46,083	8,225	17,937
Target site of insertion	Genic	Genic	Genic or intergenic	Full-length cDNA	Genic except intron

Source: Jung *et al.,* 2008

To provide easy access to the resource generated by above 10 groups it has been integrated under RAP-DB. All the flanking sequences that were tagged by

Tos17, T-DNA and *Ds* were compared with the rice genome so that the positions of genes disrupted by different methods were simultaneously displayed in the RAP-DB. These flanking sequences have been linked to the web pages of the mutant providers (RAP-DB updates 2008).

2.8 *Agrobacterium* mediated gene transformation

The natural mechanism of *Agrobacterium*-mediated gene transfer has been widely utilized to genetically modify monocots and dicots (Hansen andWright, 1999; Fukuoka *et al.*, 2000). *Agrobacterium tumefaciens* is a soil-borne pathogen that has the capacity to infect plants and transfer a segment of DNA (T-DNA) from its tumor inducing (Ti) plasmid to the plant cell. The T-DNA then integrates into the plant nuclear genome where the genes on the T-DNA are expressed. *A. tumefaciens* has a wide host range although some economically important crops remain recalcitrant (DeCleene and DeLey, 1976; Komari and Kubo, 1999). Advances in transformation techniques have resulted in high transformation frequencies in maize and rice ranging from 5% to 30% (Hiei *et al.*, 1994; Ishida *et al.*, 1996; Zhao *et al.*, 2001; Frame *et al.*, 2002).

The mechanism of DNA transfer is principally, the bacterium transferring a piece of its plasmid DNA into the plant cells infected, where it integrates into the nuclear genome and expresses its own genes, whose products disrupt the hormonal balance within the plant cells and induce their proliferation to form tumors. In addition, it also produces enzymes to synthesize opines, which the bacteria can use for their own nutrition. The T-DNA is located on a large plasmid called Ti (tumor-inducing) plasmid,

which also contains other functional parts for virulence (*vir*), conjugation (*con*) and the origin of its own replication (*ori*). In the natural infection by wild type bacteria, the T-DNA and the *vir* genes are essential for inducing plant tumors. The *vir* region is about 30 kb and encodes at least 10 operons (*virA-virJ*) whose products are vital to T-DNA processing and transfer. Any gene located in the T-DNA region in principle can be transferred, but they themselves are dispensable for this process. Only the 25bp direct repeats at the right and the left borders are necessary of which 14 base pairs are completely conserved and cluster as two separate groups. While both ends of the T-DNA seem to be involved in integration, the right one appears more important with less extent of deletion (McClean, 1998).

2.8.1 *Agrobacterium*-mediated gene transformation in rice

Initially, it was believed that only dicots, gymnosperms and a few monocot species could be transformed by *Agrobacterium* but later achievements totally changed this view by showing that many "recalcitrant" species not included in its natural host-range such as monocots and fungi can now be transformed (Davey *et al.*, 1989). Efficient transformation of rice mediated by *Agrobacterium* was developed by Hiei *et al.* (1994) and Rashid *et al.* (1996). The use of both a super-binary vector and a transformation stimulator, such as acetosyringone during co-cultivation has led to the improvements in transformation efficiency in rice. Hiei *et al.* (1994) substituted the wound response by adding acetosyringone to the co-cultivation medium and found that the temperature (22-28°C) during co cultivation was critical for rice transformation. Since the first production of transgenic rice plants from both *japonica* (Toriyama *et al.*, 1988; Zhang and Wu,

1988; Zhang *et al.*,1997) and *indica* rice (Peng *et al.*, 1992) several reports are being made for successful plant regeneration with direct and indirect DNA delivery techniques in rice.

Wide differences exist between the tissue culturability of *indica* and *japonica* rice, the former being less responsive than the later. Thus, it is important to establish parameters for the transformation of popular *indica* rice varieties. Morphologically normal, fertile transgenic plants were obtained by co-culturing embryogenic calli of the Bangladeshi *indica* rice cultivars BR26 and Binni with *Agrobacterium tumefaciens* strain LBA4404 carrying the super binary vector pTOK233 (Al-Forkan, 2004). Mohanty *et al.* 1999 reported *Agrobacterium*-mediated gene transformation of Pusa Basmati and Rashid *et al.* 1996 reported 22% transformation efficiency of Basamti 370 which was as high as reported in *japonica* rice and dicots by using *Agrobacterium tumefaciens* strain EHA101 (pIG1121Hm).

2.9 Screening of T-DNA insertion mutants

With the completion of rice genome sequencing, attention is turning increasingly to the identification of genes and determining their functions. Insertional mutagenesis with transposons and *Agrobacterium* T-DNA has become the most widely used approach in the reverse genetics. Several recent reports (Jeong *et al.*, 2006; Upadhyaya *et al.*, 2006) describe the collection of T-DNA and transposon insertional lines in rice. The application of insertional mutagenesis is thus an attractive approach for functional genomics that reduces the steps needed to establish a link to a given gene and its function (Spradling *et al.*, 1995).

Insertional mutagenesis in rice has become particularly attractive due to the development of effective methods for amplification of sequences adjoining the insertion. In this way, the knock out genes can be identified in reference to the published gene sequences. There are several strategies for recovering the unknown sequences flaking insertions of known sequences such as T-DNA and transposons (Hui *et al.*, 1998). Among them Thermal Asymmetric Interlaced PCR (TAIL- PCR) (Liu *et al.*, 1995; Liu and Whittier, 1995), Inverse Polymerase Chain Reaction (iPCR), (Ochman *et al.*, 1989; Gasch *et al.*, 1992); Plasmid rescue (Behringer and Medford, 1992; Feldmann, 1991) and Genome Walking (Siebert *et al.*, 1995). TAIL-PCR and Genome Walking have been used in the present investigation and hence reviewed below.

2.9.1 Genome Walking

Identification of unknown genomic regions can be achieved through the screening of the genomic libraries using identified DNA probes (Rishi *et al.*, 2004). The accuracy depends on how much DNA sequence is available as probes for screening. Using the genomic DNA of interest pools of restriction digested, adaptor ligated genomic DNA fragments are generated which are referred to as Genome Walker libraries. After the libraries have been constructed, it consists of two PCR amplifications per library. The first or primary PCR uses the outer adaptor primer (AP1) and an outer gene specific primer (GSP1). The primary PCR mixture is then diluted and used as a template for a secondary or "Nested" PCR with nested adaptor primer (AP2) and nested gene specific primer (GSP2). This generally produces a single major PCR product. Each of the DNA

fragment which begins in a known sequence at the end of GSP2 and extended to the unknown adjacent genomic DNA can then be cloned and further analyzed.

2.9.2. TAIL-PCR

This method was developed by Liu *et al.*, (1995) to isolate genomic sequences flanking T-DNA insertions from transgenic lines of *Arabidopsis thaliana*. TAIL-PCR utilizes a set of nested sequence specific primers together with a shorter arbitrary primer of lower melting temperature. Annealing steps can be carried out at different temperatures so that both primers or alternatively, only the specific primer will function well. Interspersing high and reduced stringency cycles enables target sequences to be amplified preferentially over non- specific products (Liu *et al.*, 1995). Each step of the programme contains two basic cycles, one is amplification generated by the specific primer at a high Tm, while the other is generated through the non-specific primers at a low Tm. TAIL-PCR was initially used to isolate flanking sequences of T-DNA in transgenes and insert ends of YAC clones (Tsugeki *et al.*, 1996; Liu and Whittier, 1995). Subsequently, it was adopted to clone promoter and genomic sequences of genes (Wang *et al.*, 2004; Nakazaki *et al.*, 1998; Terauchi and Kahl, 2000). Because of the complexity of some plant genomes, this method did not always work well due to low amplification efficiency (Michiels *et al.*, 2003). Some improvements were made, most of which focused on the use of specific primers, including longer primers with higher specificity and higher annealing temperature (Sessions *et al.*, 2002; Terauchi and Kahl, 2000).

All the rice flanking sequence tags (FSTs) are publicly available at the Rice Functional Genomic Express Database (RiceGE) developed by Salk institute. Knockout

lines in genes of interest can be identified by carrying out a simple blast search using TIGR or NCBI locus name.

2.10 Pectin methylesterase

Pectin methylesterase EC:3.1.1.11 (PME) catalyses the de-esterification of pectin into pectate and methanol. In plants, pectin methylesterase plays an important role in cell wall metabolism during fruit ripening, organ abscission, senescence, pollen tube growth, seed germination and hypocotyl elongation. PMEs are also produced by phytopathogenic micro-organisms during plant infection and by symbiotic microorganisms during their interaction with plants. (Micheli, 2001). The first crystal structure of a plant pectin methylesterase was reported by Johansson *et al.,* (2002) in carrot. The gene encoding pectin methylesterase has also been cloned from *Erwinia chrysanthemi* B374. Expression of PME in *Escherichia coli* allowed the enzyme to be characterized. The structural gene was sequenced and consists of a 1098bp open reading frame encoding a polypeptide of 39,318 Daltons, which includes an amino-terminal signal peptide.

CHAPTER 3

MATERIALS AND METHODS

3. Materials and Methods

3.1 Materials

The biological and chemical materials obtained, generated and used in the present work are mentioned in the following sections.

3.1.1 Plant materials

Three independent insertional mutants viz; B-4-1, B-2-2 and B-8-7 in the present study were generated in superfine quality *indica* rice cultivar Basmati 370 by transformation using T-DNA/*Ds* through *Agrobacterium tumefaciens* strain EHA101 containing Hm^R*Ds* cloned in the plasmid. These T-DNA/*Ds* insertional Basmati 370 transgenic lines were obtained from PAU Ludhiana (Dhaliwal *et al.*, 2001). The details of Hm^R*Ds* construct used for transformation of Basmati 370 are given in (Fig. 3.1).

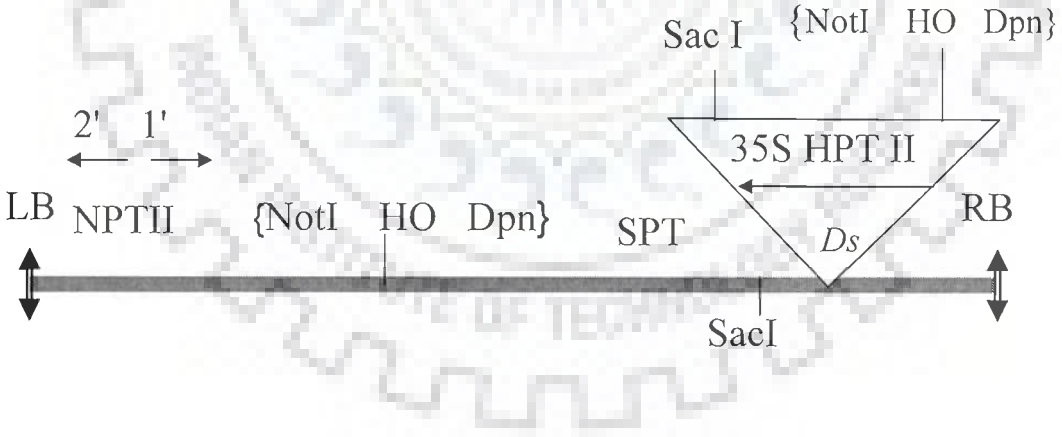

Fig. 3.1 Schematic representation of the Hm^RDs construct that was introduced into used for *Agrobacterium* mediated transformation of Basmati 370. LB: Left border, RB: Right border, SPT: streptomycin phosphotransferase , NPT: Neomycin phosphotransferase, HPT: Hygromycin phosphotransferase. *Not*I, *HO* nuclease and *Dpn/Cla*I methylase : Rare cutter sites within the T-DNA and within the *Ds* element. *Sac*I is the unique restriction site in the construct.

Hm^R*Ds* was made by replacing the internal Hind III (position 1783)- XhoI (Position 3557) fragment of *Ac* with a 1.6Kb 35S, hygromycin phosphotransferase (HPT), OcS 3' fusion. The 'simple *Ds*' and 'Hm^R*Ds*' were cloned into the binary vector pCLD III (Dean *et al.,* 1992) which carries the left border (LB) and right border (RB) of the *Agrobacterium* T-DNA, NPT II (neomycin phosphotransferase), selectable marker and a 35S-SPT fusion. The Hm^R*Ds* also carried three rare cutter sites, NotI, HO, nuclease and Dpn/ClaI methylase (Weil and McClelland, 1989) within the T-DNA and within the *Ds* element.

3.1.2 Generation of mapping population

Crosses were made between *OsPE* mutant with distantly related non-Basamti rice cultivar PR106. The F_1 plants were selfed and the 200 F_2 plants were forwarded to F_4 generation through single seed descent method. These plants were grown under containment of greenhouse at IIT Roorkee with recommended package of practices for rice cultivation. The seedlings of mutants, F_1 and subsequent generations were screened for hygromycin resistance during germination in petriplates in water containing 80 mg/L hygromycin.

3.1.3 Molecular Biology Kits

QIAquick gel extraction kit	Qiagen, Valencia, CA, USA
TRIzol plus RNA Purification kit	Invitrogen-life technologies, CA
RNeasy MinElute cleanup kit	Qiagen, Valencia, CA, USA
Superscript™ II RT kit	Invitrogen- life technologies, CA
Sequencing Quick Start Kit	Beckman Coulter, Fullerton, CA

3.1.4 Enzymes

Restriction endonucleases	New England Biolabs, Ipswich, MA
Taq DNA polymerase	Bangalore Genei, Bangalore
Reverse transcript, Superscript II	Invitrogen-life technologies, CA
DNAse I	Promega Madison, WI, USA
T4 DNA ligase	Bangalore Genei, Bangalore

3.1.5 Antibiotics

Hygromycin	Sigma-Aldrich, St Louis, MO, USA

3.1.6 Molecular weight Markers

1Kb DNA and 100bp ladder	Bangalore Genei, New England Biolabs

3.1.7 Membrane

Hybond-N	GE Biosciences, Piscataway, NJ, USA

3.1.8 Radioisotopes

α-P^{32}dCTP was handled in accordance to the North Dakota Radiological Health Rules, Section 33-10-04.1-14, "Waste Disposal" North Dakota State University, ND, USA.

3.1.9 Buffers and solutions

10X TBE buffer	1L 108g Tris 55g Boric acid 40 ml of 0.5M EDTA (pH 8.0)
TE buffer	10 mM Tris-HCl (pH 8.0) 1 mM EDTA (pH8.0)
20X SSC (pH 7.0)	1L 175.3 g NaCl 88.2 g Saline Sodium citrate pH was adjusted to 7.0 and autoclaved
Denaturation buffer (For Southern)	1.5M NaCl

	0.5N NaOH These were made from sterile 5M NaCl and 10 N NaOH stocks respectively
Neutralization buffer (For Southern)	1.5M NaCl 1.0M Tris (pH 7.4) These were made from sterile 5M NaCl and 10 N NaOH stocks respectively
Pre-hybridization buffer	6X SSC 5X Denhardt's reagent
Plant genomic DNA extraction buffer	2% Cetyl Trimethyl Ammonium Bromide (CTAB) 50mM EDTA (pH 8.0) 100mM Tris-HCl (pH 8.0) 4M NaCl Autoclaved and 1% β-mercaptoethanol was added before use
Phenol: Chloroform	Equal quantities of Tris saturated phenol (pH 7.6) and chloroform were mixed and stored in amber coloured bottles.
3M Sodium acetate(pH 5.2)	40.8g of sodium acetate .$3H_2O$ was initially added to 80ml water and the volume adjusted to 1L and sterilized by autoclaving.
5X RNA gel loading buffer	0.2 M MOPS (pH 7.0) 36% Formaldehyde 75% Deionized Farmamide
10X DNA loading dye	0.4% Bromophenol blue 0.4% Xylene cyanol FF 50% Glycerol in Distilled water

3.2. Methods

3.2.1 Recording of phenotypic data

Five plants from each row of parents and 200 F_2 plants were scored for following traits as per (Xiao *et al.*, 1998).

Tiller number: As the number of tillers bearing grain filled panicles of a plant.

Plant height: Average height (cm) from the ground to the tip of the panicle excluding the awns.

Days to heading: Number of days from date of sowing in the field until 10% of the panicles had emerged.

In addition, data for the frequency of polyembryony was also taken for parents, F_2 and subsequent generations.

3.2.2 DNA isolation

Leaf tissues for DNA isolation were collected from 5-6 week old plants from field. DNA was isolated by CTAB method as described by (Murray and Thompson 1980). Approximately 5g of leaf tissues were frozen in liquid nitrogen, ground with a mortar and pestle, and transferred to 30ml polypropylene tubes. 1% of β-Mercaptoethenol was added to the extraction buffer. Extraction buffer (10-15ml) preheated to 60°C was added to frozen tissues, mixed to form slurry and incubated at 60°C for 45min. Equal volume of 25:24:1 solution of Phenol:chloroform:isoamyl alcohol was added, mixed slowly and centrifuged at 8000X g for 10min. The upper phase was removed, and the DNA was precipitated with equal volume of isopropanol. The DNA was re-precipitated with 2 volumes of cold 95% (v/v) ethanol, dried, washed with 70% ethanol, and dissolved in TE buffer [10mM Tris-HCl (pH 8.0) and 1mM EDTA (pH 8.0)]. This DNA solution was given an RNase treatment (with

DNase free RNase A) at 37°C for 1hour and extracted with 24:1 Chloroform:Isoamylalcohol. The DNA was reprecipitated and quantified on a 0.9% agarose gel in 1X TAE [40mM Tris-acetate, 1mM EDTA (pH 8.0)] buffer.

3.2.3 *hpt* amplification

The *hpt* fragment was amplified with 20ng of DNA, 0.2mM of each of the dNTPs, 10pmole of forward and reverse primer (5′ GTC TGT CGA GAA GTT TCT GAT CG 3′ and 5′ GCG AGT ACT TCT ACA CAG CCA T 3′), 0.6 Units of Taq DNA polymerase and 1X PCR buffer (50mM KCl, 10mM Tris-HCl, 1.5mM $MgCl_2$) in a total reaction volume of 20µl. The PCR amplification was performed in PTC-200 thermalcycler (M J Research, USA), initial denaturation step of 94°C for 5min followed by 25 cycles of denaturation at 94°C for 30sec, annealing at 58°C for 20sec and extension at 72°C for 1min with final extension of 72°C for 5min.

3.2.4 Southern Blotting and hybridization

1.5µg of genomic DNA (mutant and wild type) was digested with the 10U of four restriction enzymes (*HindIII, DraI, EcoRV* and *EcoRI*) separately at 37°C overnight. The resultant fragments were separated by electrophoresis through 0.9% agarose gel, using the 1X TAE buffer at 20V for 24 hours until the bromophenol blue dye had migrated to the appropriate position. The DNA was electrophoresed at low voltage for longer period for better separation of fragments. The DNA in the gel was stained with ethidium bromide, rinsed in distilled water and photographed. The gel was destained in distilled water and DNA was purinated in 0.25N HCl for ten minutes and then neutralized with 0.4N NaOH. The DNA from the gel was transferred to a Biodyne/Hybond N^+ membrane (Amersham Inc.) by the disposable gel transfer

system using 0.4N NaOH solvent. Three sheets of Whatman filter papers (3MM) were placed in such a way that the edges were dipped into the tray containing 0.4M NaOH for capillary transfer. Gel was placed on filterpaper followed by Biodyne/Hybond N^+ membrane and three sheets of Whatman filter papers presoaked in 0.4N NaOH and paper towels (15cm thick) overnight. This arrangement allowed the DNA to transfer to the membrane. The *hpt* (950bp) probe was labelled from specific PCR product amplified by using *hpt* specific primers (5′ GTC TGT CGA GAA GTT TCT GAT CG 3′ and 5′ GCG AGT ACT TCT ACA CAG CCA T 3′). The PCR product was eluted form gel and purified using the QIAGEN QIAquick PCR purification kit (USA). PCR product (25ng) was made up to 8µl with sterilized double distilled water for 10min to denature. 2µl of hexanucleated buffer; 3µl of dNTP (AGT); 2µl of Klenow fragment and 5µl of radiolabelled α-P^{32}dCTP were added to the PCR product and incubated for 1 hour and purified through spin columns containing Sephadex G50. The blot was prehybridized with prehybridization buffer at 65°C for 2 hours. 30µl of buffer was added to the radiolabelled probe and denatured it by adding equal volume of 0.4N NaOH . Denatured probe was added to the blot and incubated for 20 hours at 65°C. Hybridization membrane was washed at 65°C for 30min each in 2X SSC and 1X SSC followed by 1 hour in 0.5X SSC. All washing solutions contained 0.1% (w/v) SDS. The hybridized blot was exposed to X-Ray film for 7 days which revealed the specific hybridization signals.

3.2.5 Bulk Segregation Analysis

3.2.5.1 Detection of polymorphism by SSRs

Two hundred and nine rice microsatellites SSRs (simple sequence repeats) primer pairs were selected representing all the 12 chromosomes of rice covering both

arms for initial parental polymorphism between Basmati 370 and PR106. Information regarding chromosomal location, SSR sequence of primers as given by Temnykh *et al.*, (2000 and 2001) and Cho *et al.*, (2000). PCR was carried out in 20µl reaction containing 50ng of DNA, 0.1mM of each of the dNTPs, 0.25µM of forward and reverse primers, 0.6U of Taq DNA polymerase and 1X PCR buffer (50mM KCl, 10mM Tris-HCl, 1.5mM $MgCl_2$) and the condition using 94°C for 5min, followed by 35 cycles of denaturation at 94°C for 30sec, annealing at 55°C for 30sec and extension at 72°C for 2min with final extension at 72°C for 7min.

3.2.5.2 Resolution of amplified SSRs product

After the completion of the reaction 2µl of 10X loading dye (0.4% w/v bromophenol blue, 0.4% w/v Xylene Cyanol FF, 50% Glycerol) was added. The PCR product was loaded on 2.5% LMP high resolution superfine agarose (Promega) prepared with 0.5X TBE buffer. Gels were stained with 1µg/ml ethidium bromide and then visualized under UV light and photographed using UVP Gel documentation system (Model GDS 7600) with GRAB-IT software programme (Annotating Grabber 32-Bits).

3.2.5.3 Identification of linked microsatellite marker and linkage analysis

Two bulks were made from the selected rows of segregating F_3 mapping population of *OsPE* and PR106. Bulks were prepared from 10 homozygous *hpt* +ve polyembryonic lines (bulk 1) and *hpt* -ve single embryo seedlings (Bulk 2). An equal concentration (5 µg) of DNA from these F_3 progenies were pooled. These two bulks and parents were used to identify putative markers linked to polyembryonic mutant.

The microsatellite markers distinguishing parents and bulks were used to analyse individual plants constituting each bulk and further tested on the F_3 progenies.

3.2.6 Genome Walking

3.2.6.1 Restriction digestion

Highly purified and intact genomic DNA of *OsPE* was used for restriction digestion. Five different blunt end restriction enzymes were used to generate genome walking libraries. These enzymes were selected on the basis of restriction map of

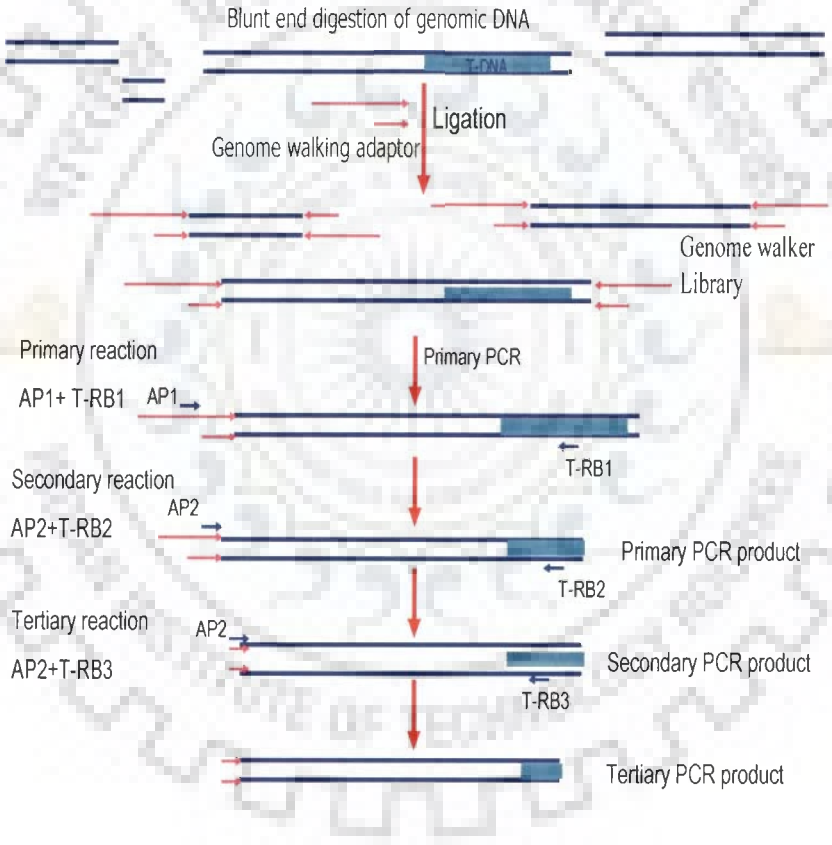

Fig. 3.2 Schematic diagram explaining Genome Walking technique

random rice BAC clones. For complete digestion, 500ng of genomic DNA was restricted with 30U of *EcoRV, DraI, RsaI, HpaI* and *HaeIII* separately at 37°C overnight. The digested DNA was ethanol precipitated after Phenol: Chloroform (1:1) extraction. The DNA pellet was re-suspended in water and 500ng of completely digested DNA was used to ligate with the "Genome Walker Adaptor".

3.2.6.2 Primers

Following are the T-DNA specific, genome specific and adaptor specific primers used for genome walking experiment. The specificity of each primer was confirmed by comparing its sequence with all the predicted rice coding sequences using the BLAST tool at NCBI to ensure that each pair targets a unique site within the set of predicted rice coding sequences.

Gene- specific primers

T-DNA left border nested primers

T-LB 1: (5′ -TGGGTATCTGGGAATGGCGAAATA-3′)

T-LB 2: (5′ -CAAGGCATCGATCGTGAAGTTT-3′)

T-LB 3: (5′ -AATGTAGACACGTCGAAATAAAGA-3′)

T-LB 4: (5′ -CATGTAGATTTCCCGGACATGAAGCCATTT-3′)

T-DNA right border nested primers

T-RB 1: (5′ -GGGGCATCGCACCGGTGAGTAAT-3′)

T-RB 2: (5′ -AGCGAATTTGGCCTGTAGACCTCA-3′)

T-RB 3: (5′ -TATTCGGGCCTAACTTTTGGTGTG-3′)

T-RB 4: (5′ -GGCCTGTAGACCTCAATTGCGAGCTTTCTA-3′)

3.2.6.3 Genome walker adaptor

Genome walker adaptor was constructed by mixing 48bp sequence and 3′ modified 8bp sequence followed by heating at 95°C for 10 minutes and then kept at room temperature for one hour.

GenomeWalker Adaptor:

5'–GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGT–3'
3'–**H2N**-CCCGACCA-PO4–5′

Adaptor-specific primers

AP1: (5′ -GTAATACGACTCACTATAGGGC-3′)

AP2: (5′ -ACTATAGGGCACGCGTGGT-3′

Genome-specific nested primers

PE-RB1: (5′ -GGCATAGAATTTGCAGGTCATGAAG-3′)

PE-RB2: (5′ -TGTAAATCCGATGGTTCAACAAACC-3′)

PE-LB1: (5′ -CTGTTTACAATTCGTGCTGCATTTG-3′)

PE-LB2: (5′ -AAGTTTCCGCTCGGTATACAACTCC-3′)

3.2.6.4 Ligation

Completely digested 500ng of DNA was used to ligate with the "Genome Walker Adaptor". The ligated products were made to final volume of 80µl, out of which 1µl containing approximately 6.25ng of template DNA was used for the first Genome Walking PCR.

3.2.6.5 Genome Walking PCR Conditions

The PCR reactions were carried out in a total volume of 20µl containing the following to a final concentration: 1X Taq DNA polymerase buffer, 1.5mM $MgCl_2$, 0.2mM of each (dATP, dTTP, dCTP, dGTP), Taq polymerase and 0.2µM of each primer. T-DNA specific primer (T-RB1: 5′GGGCATCGCACCGGTGAGTAAT-3′) and Adaptor specific primer (AP1: 5′ -GTAATACGACTCACTATAGGGC-3′) were used for the primary PCR. The PCR product was separated on agarose gel. The primary PCR product was diluted fifty fold and 1µl of diluted PCR product was used as template for the secondary PCR.

The secondary PCR reaction was carried in a total volume of 50µl containing the following to a final concentration; 1X Taq DNA Polymerase buffer, 1.5mM $MgCl_2$, 0.2µM of each primer (adaptor-specific and T-DNA specific), 1µl diluted first PCR reaction as template and Taq DNA polymerase. The T-DNA specific primer (T-RB2: 5′-AGCGAATTTGGCCTGTAGACCTCA-3′) and an adaptor-specific primer (AP2: 5′-ACTATAGGGCACGCGTGGT-3′) were used. The amplified product from the second PCR reaction was resolved using 1% (w/v) Agarose gel and 1X TBE buffer.

Tertiary PCR was performed using T-DNA specific primer (T-RB3: 5′ -TATTCGGGCCTAACTTTTGGTGTG -3′) and an adaptor-specific primer (AP2: 5′ -ACTATAGGGCACGCGTGGT-3′) and template of fifty fold diluted secondary PCR product with similar set of PCR conditions using Applied Biosciences 2720 Thermal Cycler, USA. The PCR conditions used are given in Table 3.1.

Table3.1 PCR Cycles used for Genome Walking

Reaction	Cycle number	Thermal cycling conditions
Primary	1	94°C (3min)
	10	94°C (30sec), 72°C - 67°C (45sec) (-0.5°C / per cycle), 72°C (2.5min)
	30	94°C (30sec), 67°C (45sec), 72°C (2.5min)
	1	72°C (7min)
	1	4°C (Forever)
Secondary	1	94°C (3min)
	10	94°C (30sec), 72°C - 67°C (45sec) (-0.5°C / per cycle), 72°C (2.5min)
	30	94°C (30sec), 67°C (45sec), 72°C (2.5min)
	1	72°C
	1	4°C (Forever)
Tertiary	1	94°C (3min)
	35	94°C (30sec), 72°C - 67°C (45sec) (-0.5°C / per cycle), 72°C (2.5min)
	1	72°C
	1	4°C (Forever)

3.2.6.6 PCR conditions used with genome specific and T-DNA specific primers

The insertion was confirmed using T-DNA specific and genome specific primers using Basmati 370 and mutant DNA template. For this three T-DNA specific primers (T-RB1, T-RB2 and T-RB3) and two nested genome specific primers (PE-RB1 and PE-RB2) were used. Conditions used for 20µl reaction were 94°C for 3 min; [Touchdown PCR: 94°C, 30sec; 70°C to 65°C (-0.5°C /cycle), 45sec; 72°C, 150sec] for 10 cycles; [94°C for 30sec; 65°C for 45sec; 72°C for 150sec] for 30 cycles and an additional cycle of 72°C for 7minutes using Applied Biosciences 2720 Thermal Cycler, USA.

3.2.7 TAIL PCR

3.2.7.1 Primers

The sequence of T-DNA specific primers used for TAIL-PCR are as follows:

T-RB1: 5'-GGGGCATCGCACCGGTGAGTAAT-3' (T_m=63.7°C)

T-RB2: 5'-AGCGAATTTGGCCTGTAGACCTCA-3' (T_m=60.4°C)

T-RB3: 5'-TATTCGGGCCTAACTTTTGGTGTG-3' (T_m=57.4°C)

AD1: 5'-NTCGASTWTSGWGTT-3' (64-fold degeneracy, average T_m=45.3°C)

Where, N= A/C/G/T, S= C/G, and W= A/T

3.2.7.2 TAIL-PCR procedure

20µL of primary TAIL-PCR reaction contained 1X PCR buffer (10mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$), 200µM each of dNTPs, about 200ng of genomic DNA, 0.8 unit of Taq polymerase, 0.15µM T-RB1 and AD primer (1.5µM for AD1), 1µL of the primary PCR product was diluted with 49µL of distilled water and 2µL of diluted DNA was used as template for secondary reaction. 20µL of secondary reaction

contained 1X PCR buffer, 0.6 unit of Taq polymerase, 200μM each of dNTPs, 0.2μM T-RB2 and the same AD primer used in the primary reaction. 100μL of tertiary PCR mixture contained 1X PCR buffer, 0.6 unit of Taq polymerase, 200μM each of dNTPs, 0.2μM T-RB3 and 1.5 μM AD primer used in the primary reaction. 1μL of the secondary PCR product was diluted with 9μL distilled water and 2μL of the diluted DNA was used as template DNA in tertiary PCR reaction. Primary, secondary and tertiary reactions were executed as summarized in Table 3.2 in Biorad thermal cycler.

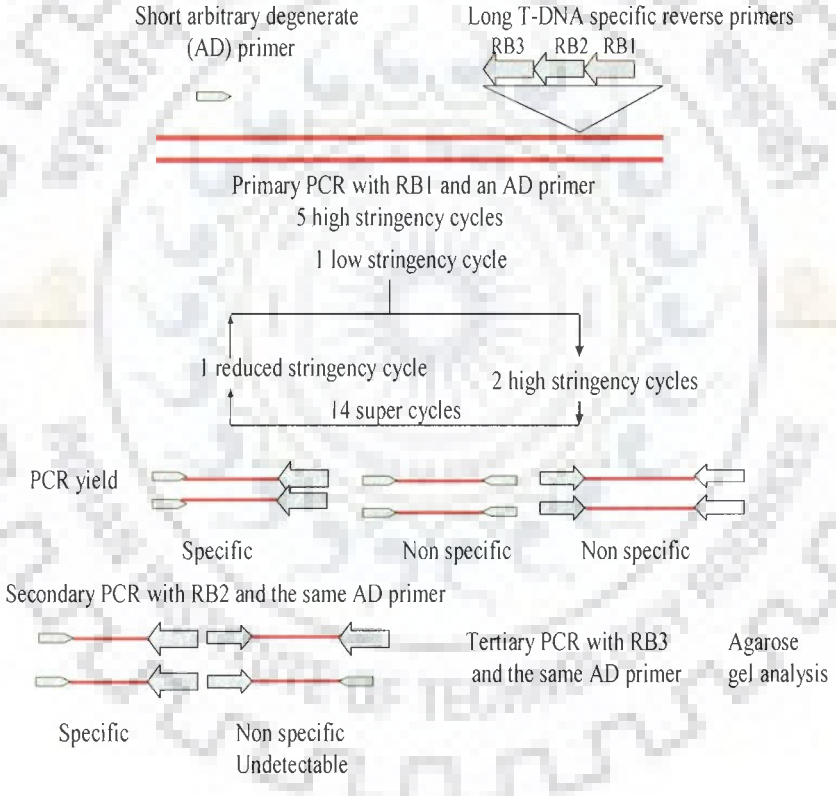

Fig.3.3 Schematic diagram explaining TAIL-PCR

Table3.2 PCR Cycles used for TAIL-PCR

Reaction	Cycles	Temperature(°C)	Time
Primary reaction	1	93 95	1 min. 1 min.
	5	94 62 72	1 min. 1 min. 2.30 min.
	1	94 25 72	1 min. 3 min. 2.30 min.
	15	94 68 72 94 68 72 94 44 72	30 sec. 1 min. 2.30 min. 30 sec. 1 min. 2.30 min. 30 sec. 1 min. 2.30 min.
	1	72	5 min.
Secondary reaction	1	94	1 min.
	15	95 63 72 94 63 72 94 44 72	10 sec. 1 min. 2 min. 10 sec. 1 min. 2 min. 10 sec. 1 min. 2 min.
	1	72	5 min.
Tertiary reaction	1	95	3 min.
	30	94 44 72	20 sec. 1 min. 2 min.
	1	72	10 min.

3.2.8 Gel elution and purification of PCR product

The DNA fragment was excised from the agarose gel with a clean and sharp scalpel. Three volumes of Buffer QG (provide in the QIAquick gel extraction kit) was added to 1 volume of the gel. Incubated for 10min at 50°C. After the gel slice had dissolved completely, 1 gel volume of isopropanol was added. QIAquick spin column was placed in the 2ml collection tube and centrifuged for 1min. 0.5ml BufferQG was added and centrifuged for 1min followed by the addition of 0.75ml Buffer PE (provided in the kit) and centrifuged for 1min with and additional run of 1min. Placed the column in the clean 1.5ml microcentrifuge tube. Elution of DNA was done by adding 20µl water to the centre of the QIAquick membrane and centrifuged the column for 1 minute.

3.2.9 Direct Sequencing of PCR product

The sequencing reactions were carried in a total volume of 20µl using ECQ Dye Terminator Cycle Sequencing with Quick Start Kit (Beckman Coulter) containing the following to a final concentration; 1.8Kb gel eluted fragment of second Genome Walking PCR (65ng-130ng); sequencing primer T-RB2/AP2 (5µM) and DTCS Quick Start master mix (8µl). The thermal cycling program was; 96°C for 30sec, 50°C for 20sec, 60°C for 4min for 30 cycles. 5µl of stop solution/ Glycogen mixture (2µl of 3M Sodium acetate (pH 5.2), 2µl of 100mM Na_2EDTA (pH 8.0) and 1µl of 20mg/ml of glycogen) was added to the sequencing PCR product. For ethanol precipitation 30µl of cold 95% v/v ethanol was added followed by centrifugation at 14,000rpm at 4°C for 15 minutes. Rinsed with 70% (v/v) ethanol and vacuum dried; re-suspended the sample in 40µl of the sample loading solution (provided in the kit). The re-suspended samples were transferred to the appropriate wells of the CEQ

sample plate overlaid with a drop of light mineral oil. The sample plate was loaded into the Beckman Coulter sequencing machine.

3.2.10 RT-PCR

3.2.10.1 Total RNA isolation

Roots and shoots of the polyembryonic mutant were freeze dried after 6 and 12 days of seed germination. The frozen specimen was removed from the freezer and was placed into TRIZOL reagent in 1.5ml eppendorf and homogenized using a tissue homogenizer and incubated for 5min at room temperature. Added 0.1ml chloroform and was shaken vigorously for 15-30sec by hand and incubated at room temperature for 2-3min. Centrifuged at 12000X g for 15min at 4°C, removed the upper colourless layer into a fresh tube. Added 1/10 volume of sodium acetate and two volumes of ethanol followed by incubation for 30min at 20°C. Centrifuged at 12000X g for 15min at 4°C, pellet was washed with 70% ethanol and resuspended in 50µl RNase free water. RNA was quantified by using Nanodrop method.

3.2.10.2 DNase treatment and cleaning of RNA

DNA was removed by DNase treatment. 10µl reaction consisted of 8µl RNA in DEPC treated water, 1U/µg RQ1 RNase free DNase and 1µl 10X reaction buffer was incubated at 37°C for 30min. Resulted RNA was Phenol:chloroform extracted and precipitated with sodium acetate and ethanol. RNA was pelleted and washed with 70% ethanol and resuspended in DEPC treated water.

RNA was further cleaned up using RNeasy MinElute cleanup kit. 350µl buffer RLT was added to 100µl of RNA and mixed thoroughly, followed by addition of 250µl of 100% ethanol and mixed thoroughly. Added 700µl of sample to an RNeasy

MinElute spin column in a 2ml collection tube and centrifuged for 15sec at 8000X g. The flow through was discarded and 500µl of 80% ethanol was added to the spin column and centrifuged for 2min at 8000X g. The flow through was discarded cap of the spin column was opened and centrifuged at high speed for 5min and discarded the flow through and collection tube. Transfered the spin column in new 1.5ml collection tube, pipetted out 14µl of RNase free water directly on to the centre of the silica gel membrane, centrifuged to the maximum speed to elute RNA and quantified by using Nanodrop method (NanoDrop Tech).

3.2.10.3 Resolution of total RNA

The electrophoretic unit was treated with 1N NaOH overnight. All the solutions were prepared in DEPC treated water. 2µl of total RNA mixed with 2µl of STS loading dye was loaded on the gel and run at 50V. Gel was visualized and photographed.

3.2.10.4 cDNA synthesis

The mRNA was reverse transcribed using the superscript™ II RT kit (Invitrogen). 20µl reaction was set using 1µg of total RNA; 500µg/ml oligo (dT) $_{12-18}$ and 1µl dNTP mix (10mM each) was heated to 65°C for 5min and quick chilled on ice. Followed by brief centrifugation, added 4µl of 5X first strand buffer; 2µl of 0.1M DDT; 40U/µl RNase OUT. Mixed the contents of the tube thoroughly and incubated at 42°C for 2 minutes. Added 1µl (200 U) of Super Script II RT and mixed by pipetting gently up and down and incubated at 42°C for 90min. Inactivated the reaction by heating at 70°C for 15min. 2µl of this cDNA was used for the RT-PCR.

The sequences of Forward and Reverse primers used for amplification of Actin I gene were [Actin I Fwd: ATCCTTGTATGCTAGCGGTCGA; Actin I Rev: ATCCAACCGGAGGATAGCATG]. The primer pair was designed using PRIMER3 on intron exon junctions.

3.2.10.5 RT-PCR conditions

PCR reaction was carried out in 20µl volume containing 20ng of DNA; 0.2mM of each of the dNTPs; 10pmole of forward and reverse primers (RTP-F: 5′ CAAGCTAAGAATGGTAACATCC3′; RTP-R: 5′ GCTGGGGTTGGCATACCTTG 3′); 0.8U of Taq DNA polymerase and 1X PCR buffer (50mM KCl, 10mM Tris-HCl, 1.5mM $MgCl_2$). The PCR amplification was performed in PTC-200 Thermal cycler (M J Research, USA) with the thermal profile; initial denaturation step of 94°C for 5min, followed by 35 cycles of denaturation at 94°C for 20sec, annealing at 62°C for 20sec and extension at 72°C for 30sec with final extension of 72°C for 5min.

3.2.11 Meiotic analysis

Meiosis was studied after fixing the panicles in Carnoy's fixative (ethanol: chloroform: glacial acetic acid; 6:3:1) for 24 hours followed by 70% ethanol. Subsequently Pollen Mother Cells (PMCs) squash preparations were stained with 2% acetocarmine (Herr, 1971).

3.2.12 Bioinformatics tools

The links for various bioinformatics tools used in the present study are:

Homology search (NCBI BLAST)
http://www.ncbi.nlm.nih.gov/blast

Genomic region of rice (Mapview, Entrez and Gramene)
http://www.ncbi.nlm.nih.gov/mapview/maps.
http://www.ncbi.nlm.nih.gov/sites/entrez
http://www.gramene.org

Primer designing (Primer3)
http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi

Multi-alignment of protein sequence (ClustalW)
http://www.ebi.ac.uk/clustalw/

EST profile data for specific gene (UniGene),
http://www.ncbi.nlm.nih.gov/UniGene/

Homolog search (dpget-bin)
http://www.genome.jp/dbget-bin

Prediction of introns, exons and amino acid sequence (GENSCAN),
http://genes.mit.edu/GENSCAN

Rice Functional Genomic Express Database
http://signal.salk.edu/cgi-bin/RiceGE

CHAPTER 4

RESULTS

4. Results

Polyembryony has been described previously in several crops and cereal plants, but the underlying genetic causes of twinning in such cases have been complex and difficult to resolve. *OsPE* mutant therefore, represents the first reported example of high frequency and heritable polyembryony in rice caused by insertional mutagenesis. The combined use of Southern blot, BSA, Genome Walking, TAIL-PCR, RT-PCR techniques and bioinformatics led to the identification of a candidate gene controlling multiple embryo formation in Basmati 370. The results obtained with respect to aforesaid objectives are presented as under:

4.1 *OsPE* mutant

4.1.1 Polyembryony frequency

The frequency of Polyembryonic seeds including twins, triplet and rare quadruplets varied from 9.8 to 21.8% (Table 4.1; Fig. 4.1 and Fig. 4.2). Over 1000 mutant seeds from subsequent generations yielded similar results, confirming that variable penetrance and expressivity were the general characteristics of the *OsPE*. The polyembryonic trait of *OsPE* was heritable. Dissection of embryos from seeds with multiple shoots and roots during germination clearly indicated that the twin/ triplet embryogenic axis with independent root and shoot were developed from a single scutellum (Fig. 4.3). Each plantlet in the polyembryonic seed had its own root and shoot axis. All the plantlets did not emerge simultaneously. There was slight delay in the emergence of the additional roots and shoots. The additional plantlets when separated two weeks after germination gave rise to independent rice plants (Fig 4.4).

Table 4.1. Percentage of polyembryony among different seeds and plants of *OsPE* mutant.

Source of seed	Total no. Seeds sown	No. of seeds germinated	No. of seedlings with Single embryo	No. of seedlings with twin embryo	No. of seedlings with triplet embryos	Polyembryony%
Basmati 370 control	50	50	-	-	-	0
OsPE (Single) 1	800	741	658	122	20	19.7
OsPE (Single) 2	700	680	572	96	12	15.9
OsPE (Twin) 1	80	71	56	11	4	21.0
OsPE (Twin) 2	43	40	36	6	-	15.0
OsPE (Twin) 3	70	63	55	8	-	12.7
OsPE (Twin) 4	53	51	46	5	-	9.8
OsPE X PR106 F_3	138 lines	138 lines	61 lines	12 lines	2 lines	10.1

Fig. 4.1 Seeds showing twin/triplet plantlets in *OsPE* mutant

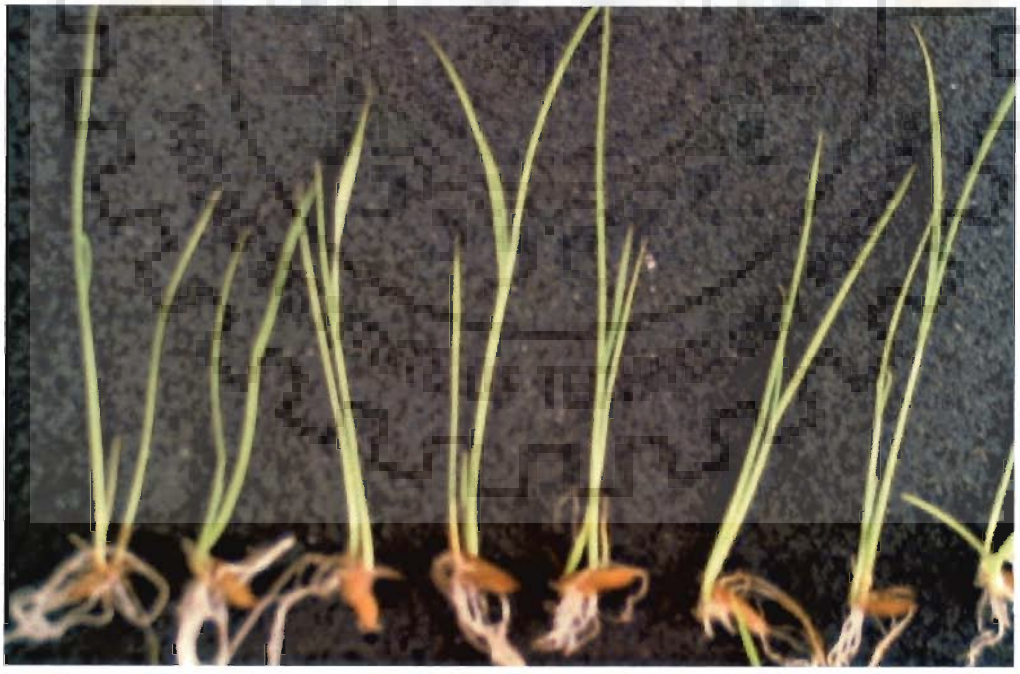

Fig. 4.2 Seeds showing twin/triplet seedlings of *OsPE* mutant

Fig. 4.3 Multiple embryos of *OsPE* mutant with independent shoot and root axis attached to a single scutellum. (Green arrows: Shoot; Yellow arrows: Roots)

Fig. 4.4 Plants separated from different twin/ triplet polyembryonic mutant *OsPE*

4.1.2 Association of *hpt* gene and hygromycin resistance with *OsPE*

The hygromycin resistance of the *OsPE* mutant was tested by growing the seeds in the petriplate containing different doses of hygromycin. At 80 ppm *OsPE* mutant germinated whereas Basmati 370 seeds which were used as control, showed no germination at all under similar conditions (Fig 4.5). Further confirmation of *hpt* gene was done by PCR amplification of the 950bp fragment of *hpt* gene using *hpt* F: (5′ GTC TGT CGA GAA GTT TCT GAT CG 3′) and *hpt* R: (5′ GCG AGT ACT TCT ACA CAG CCA T 3′) primers (Fig. 4.6). The expression of the polyembryony among *OsPE* plants varied considerably from single embryo to triplet. Out of 800 *OsPE* seeds grown using 80 ppm hygromycin solution, 122 were twin and 20 triplet and the rest were single. A rare occurrence of quadruplet embryo was also recorded.

Polyembryonic seedlings were also obtained in population generated from a cross *OsPE* X PR106. Extensive study on germination of seeds of F_3 and F_4 progenies indicated that twins/triplet and single plants consistently gave rise to comparable numbers of polyembryonic seeds following self-pollination. The *hpt* amplification was also done from F_2 plants generated from a cross *OsPE* X PR106.

4.1.3 Inheritance of *OsPE* gene

A population was developed by crossing *OsPE* with a non Basmati fine rice cultivar and a large F_2 was obtained. Morphological data for plant height and number of tillers was collected for 202 F_2 plants. The F_2 population with 143 tall: 59 dwarf plants segregated in 3tall: 1dwarf ratio (χ^2 =1.90 at 1 *df* and $p<0.05$) (Fig. 4.7). Normal distribution was obtained for tiller number (Fig. 4.8). The segregation pattern of F_2 plants

was also analysed on the basis of hygromycin resistance and amplification of *hpt* gene sequence (Table 4.2). A representative gel picture showing amplification profile of *hpt* positive and *hpt* negative individuals in F_2 plants is presented in (Fig 4.9).

The goodness of fit of the observed number of *hpt* +ve and *hpt* –ve plants were tested against the expected segregation ratio for a single gene using χ^2. Out of 138 F_2 plants, 96 were *hpt* positive while 42 were *hpt* negative showing a good fit to the expected 3:1 ratio (χ^2= 1.156 at 1 *df* and p< 0.05) . The F_2 phenotypic data for polyembryony unlike *hpt* data was not segregating in the expected 3:1 ratio. Out of 175 F_2 plants only 14 showed polyembryony (twin/triplet phenotype) showing a significant deviation for 3:1 ratio (χ^2= 419.22 at 1 *df* and p< 0.05). All the fourteen plants which showed polyembryonic phenotype were giving PCR amplification for *hpt* gene as well.

4.2. Meiotic analysis

The *OsPE* mutant was highly fertile with typical Basmati type seeds. The behaviour of pollen mother cells (PMCs) was studied. The meiotic analysis for PMC of mutants indicated diploid (2n=24) chromosomal number. Meiotic cell divisions were normal with 12 IIs at metaphase I and tetrads (Fig. 4.10).

4.3 Southern blot and hybridization

The genomic DNA was digested with four restriction enzymes (*Hindlll, DraI, EcoRV* and *EcoRI*), separated by agarose gel electrophoresis, blotted on Nylon membrane and hybridized with 950bp *hpt* probe. Only one band in each lane of DNA restricted with *Hindlll, DraI, EcoRV* and *EcoRI* was found indicating only single insertion of the cassette (Fig. 4.11). In *OsPE* mutant, only one copy of the Hm^R*Ds* insert was found.

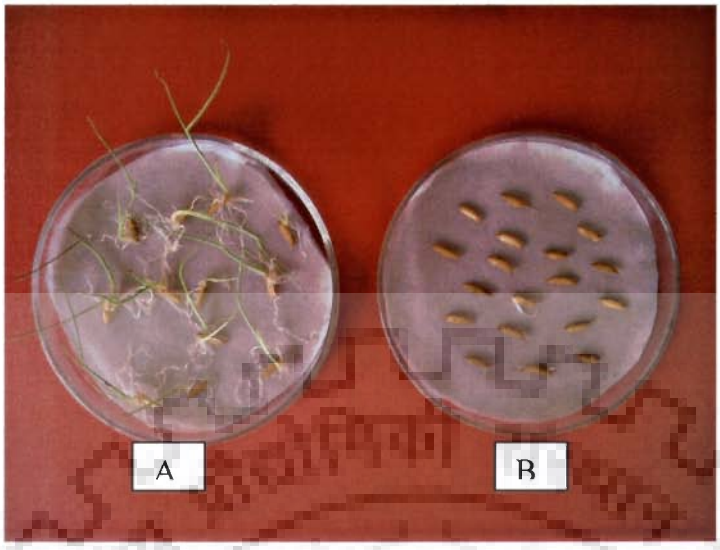

Fig. 4.5 Germination of *OsPE* mutant and Basmati 370 on 80 ppm hygromycin solution A: Polyembryonic mutant B:Basmati 370

Fig. 4.6 PCR amplification of 950 bp *hpt* fragment in insertional mutants Insertional Mutants (Lanes 1,2,4); *OsPE* (Lane 3); Basmati 370 (Lane 5); Lane M: Molecular weight marker 1 Kb

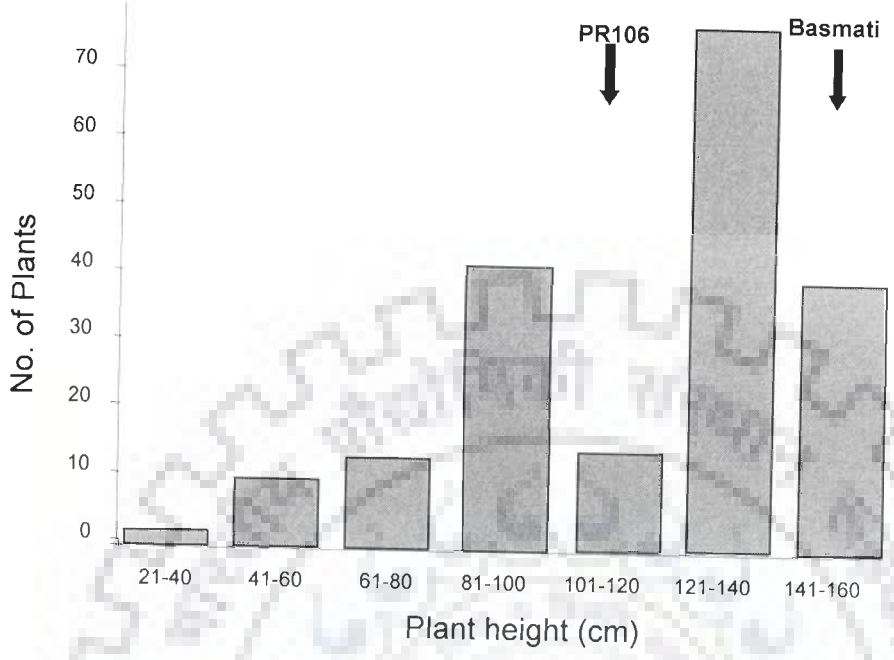

Fig. 4.7 Frequency distribution of *OsPE*/PR106 F_2 plants for plant height

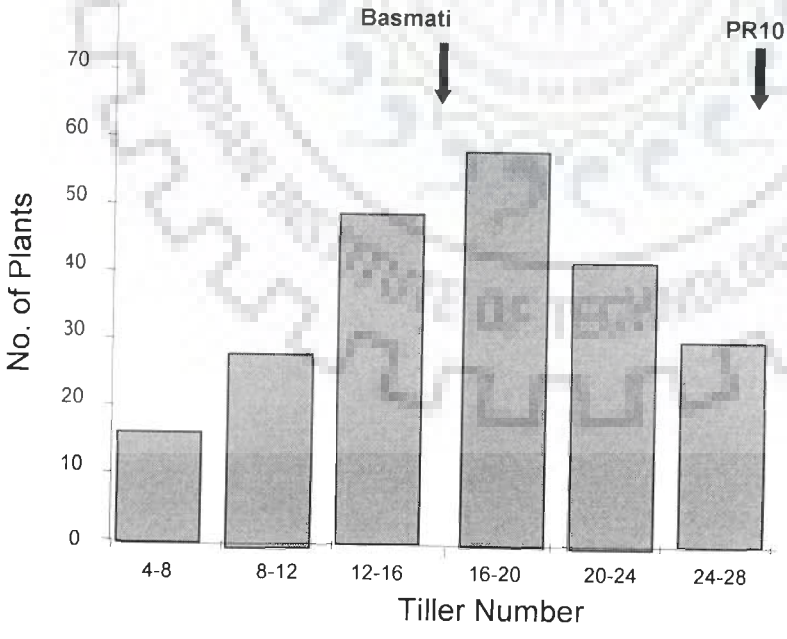

Fig. 4.8 Normal distribution of *OsPE*/PR106 F_2 plants for tiller number

Table 4.2 Segregation data of F_2 plants/F_3 progenies of *OsPE*/PR106 for seed germination and *hpt* amplification

Cross	F_2 (*OsPE* / PR106)									
	Germination of 138 F_3 progenies on 80 ppm hygromycin solution					*hpt* gene amplification (138 F_2 plants)				
OsPE / PR106	No. of F_3 lines germinated	No. of F_3 lines grown	Expected ratio	χ^2 value		No. of F2 *hpt* +ve plants	No. of F2 *hpt* -ve plants	Expected ratio	χ^2 value	
				Calculated	Table (p>0.05)				Calculated	Table (p>0.05)
	96	138	3:1	1.15	3.841	96	42	3:1	1.156	3.841

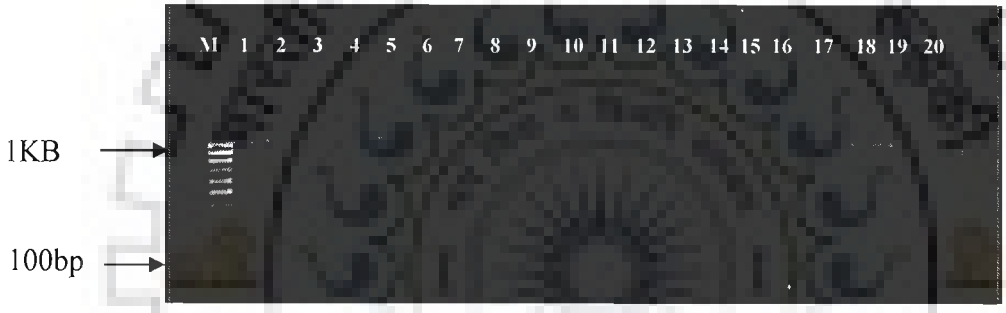

Fig. 4.9 PCR amplification of 950 bp *hpt* fragment in *OsPE*/PR106 F_2 population Lane 1: *OsPE* mutant ; Lane 2: Basmati370; Lanes: 3-20F_2 Plants; Lane M: Molecular weight marker 100bp

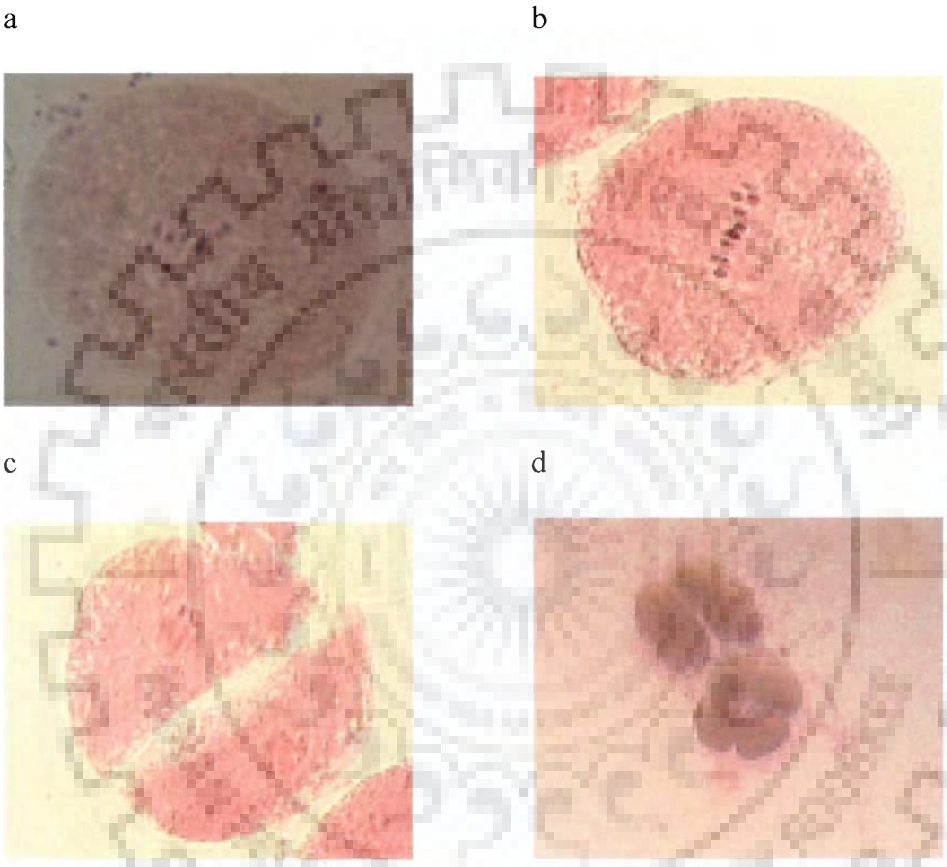

Fig. 4.10 PMC Meiosis in *OsPE*, a: Diakinesis 12IIs ; b: Metaphase-I, 12IIs; c: Telophase-I; d: Tetrad stage

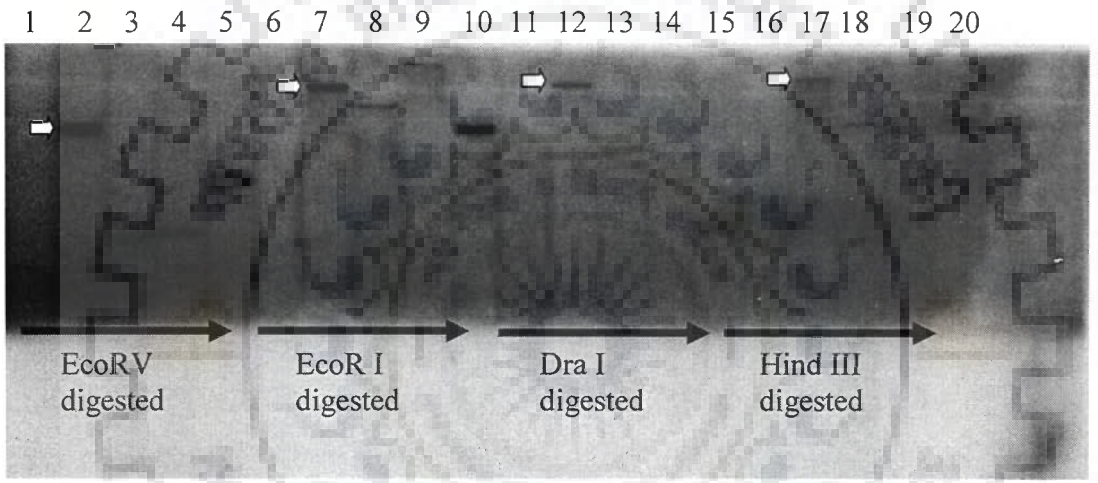

Fig. 4.11 Autoradiograph of Southern hybridization of Basmati and various insertional mutants restricted with four restriction enzymes and probed with 950bp *hpt* product
Lanes 1, 6,11,16: Basmati 370; Lanes 2,7, 12, 17: *OsPE* (arrows)

4.4 Bulk Segregant Analysis (BSA)

4.4.1 Parental polymorphism survey

Parental polymorphism survey was conducted for identifying the SSR markers polymorphic between the Basmati 370 and PR106. 209 SSR primer pairs uniformly distributed across 12 rice chromosomes were selected based on the rice map of Temnykh *et al.* (2000) and IRGSP 2005. These markers were selected at an average interval of 10cM covering the entire length of 1500cM of rice genome. Of the 209 primer pairs tested, 98 (46.8%) exhibited polymorphism (Fig. 4.12). The remaining primer pairs amplified more than one product or were monomorphic.

4.4.2 Chromosomal location and mapping of *OsPE* locus

Mapping of the *OsPE* mutant was carried out following BSA approach. A set of 98 markers showing polymorphism between the parental lines Basmati 370 and PR106 were used for identifying markers associated with the *OsPE* locus through BSA. Ten *hpt* +ve and equal number of *hpt* –ve F_3 progenies from the cross *OsPE* / PR106 were used for constituting the +ve and –ve bulks. The *hpt* +ve lines were also showing polyembryonic phenotype.

The RM14667 and RM14645 markers showed clear polymorphism between bulks. The gel picture showing the results of BSA using RM14667 and RM 14645 markers is given in (Fig 4.13). Both these markers are located on chromosome 3. These two putative markers showing polymorphism in bulks were further used on debulks, the individual plants constituting the bulks. Out of ten plants in +ve debulk for RM14667 only one recombinant (Lane 7) was found and for RM14645 one heterozygous (Lane 6)

and one recombinant (Lane 9) were found. For the 12 progenies in the –ve debulk amplified with RM14667 and RM14645, no recombinant was found. The amplification obtained clearly indicated close association of these markers with *OsPE* locus (Fig 4.14 and 4.15). These two putative markers were then used for genotyping on F_2 population of 138 plants. The linkage analysis showed 3 and 8 recombinants for RM14667 and RM14645, respectively (Table 4.3). The co-segregation analysis of individual markers and *OsPE* genotype of 138 F_2 plants showed that the marker loci of RM14645 and RM14667 were linked to *OsPE* at 5.79cM and 2.17cM, respectively (Fig 4.16). The distance between SSR markers and T-DNA was calculated based on the recombination frequency. The data obtained using the SSR markers clearly indicated that the *OsPE* mutant is located on chromosome 3.

In the present study BSA approach has been further confirmed with other reverse genetic approaches such as genome walking and TAIL-PCR for amplification of T-DNA flanking region, sequencing and BLAST search to reach the exact location of insertion on chromosome 3.

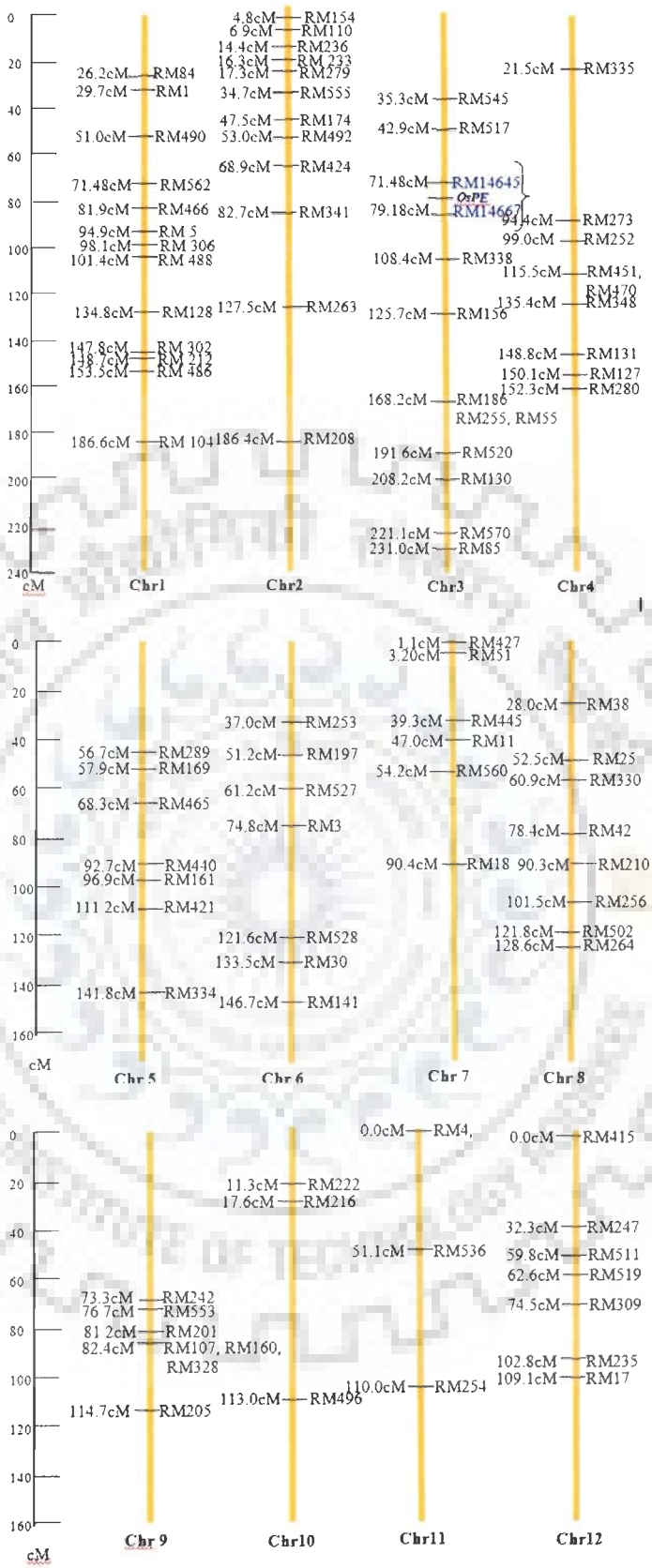

Fig. 4.12 Map position of SSR markers on 12 rice chromosomes showing polymorphism between Basmati 370 and PR106 and RM14645 and RM14667 on chromosome 3 linked to *OsPE* mutant (In blue colour)

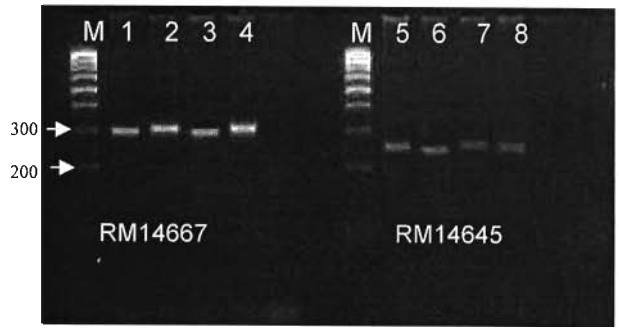

Fig. 4.13 PCR amplification with RM14667 and RM14645 markers on bulks; Lanes 1 and 6: PR106, Lanes 2 and 5: Basmati 370, Lanes 3 and 8: -ve bulk, Lane 4 and 7: +ve bulk, M: Molecular weight marker

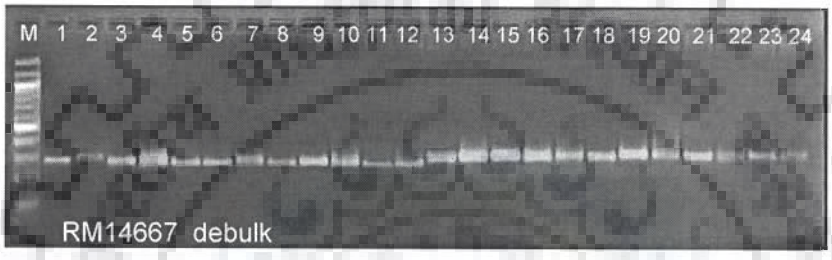

Fig. 4.14 PCR amplification of individual plants constituting the bulk with RM14667; Lane 1: PR106, Lane 2: Basmati 370, Lanes 3-12: -ve debulked lines, Lanes 14-24: +ve debulked lines, M: Molecular weight marker

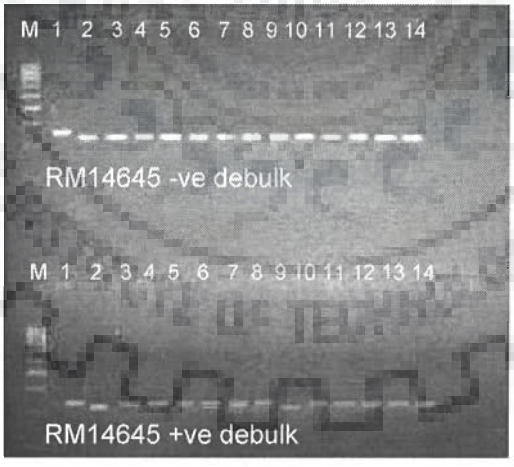

Fig. 4.15 PCR amplification of individual plants constituting the bulk with RM14645.
Upper row, Lane 1: Basmati 370, Lane 2: PR106, Lanes 3-14: -ve debulked lines. Lower row, Lane 1: Basmati 370, Lane 2: PR106, Lanes 3-14 : +ve debulked lined. M: Molecular weight marker

Table 4.3 Recombination frequency in 138 F_2 plants with the two linked markers

Marker	Basmati 370 allele but *hpt* -ve	Homozygous for PR106 allele but *hpt* +ve	Total number of recombinants	Recombination frequency (%) in cM
RM14645	2	6	8	5.79
RM14667	1	2	3	2.17

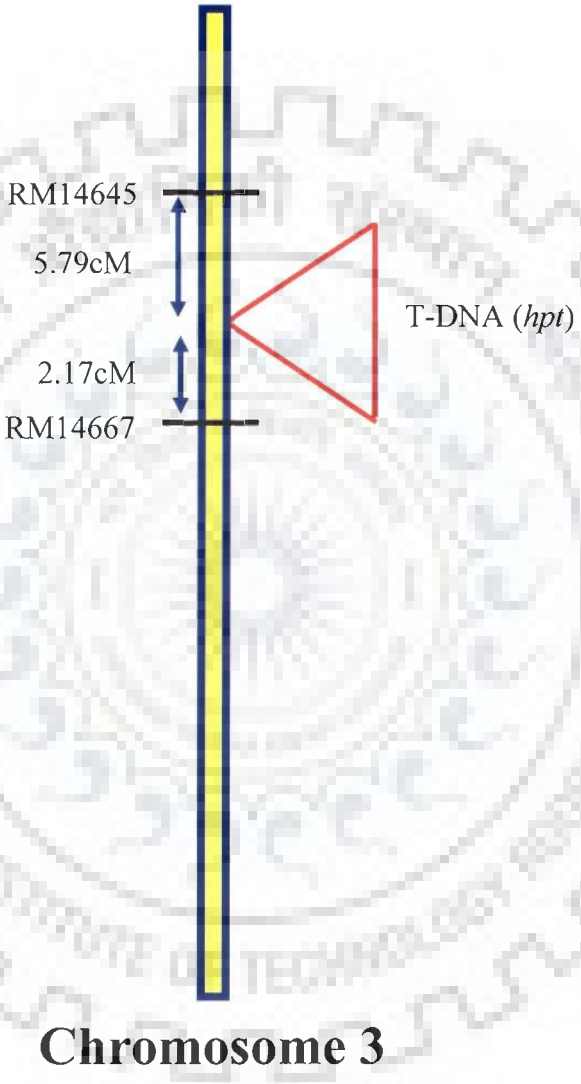

Fig. 4.16 Linkage map of chromosome 3 showing the genetic distance of linked markers RM14645 and RM14667 in relation to T-DNA insertional site of *OsPE* mutant

4.5 Genome walking

Three T-DNA specific nested primers (T-RB1, T-RB2 and T-RB3) were designed and used for genome walking for identifying sequence flanking T-DNA insertion (Fig 4.17). The Primer3 programme was used to design these primers. Five Genome walking libraries using *EcoRV*, *DraI*, *RsaIII*, *HaeIII* and *HpaI* were developed. The digested DNA was purified and ligated to the adaptors. The libraries were then used to perform primary, secondary and tertiary PCR. The primary PCR gel showed a smear (Fig 4.18 lane 1). The secondary PCR using the T-RB2/AP2 primers and diluted primary PCR product as template, amplified the desired product of 1.9Kb. Tertiary nested PCR resulted in a single band of 1.8Kb (Fig 4.18 lane 2). Out of five genome walking libraries, *DraI* and *RsaIII* resulted in amplification of 200bp and 300bp fragments but no amplification was observed in the *HaeIII* and *HpaI* libraries. PCR amplified product of 1.9Kb was observed in *EcoRV* library.

```
Right Border Sequence
GGATCCCTGAAAGCGACGTTGGATGTTAACATCTACAAATTGCCTTTTCT

TATCGACCATGTACGTAAGCGCTTACGTTTTTGGTGGACCCTTGAGGAAA

CTGGTAGCTGTTGTGGGCCTGTGGTCTCAAGATGGATCATTAATTTCCAC
                      >>RB1
CTTCACCTACGATGGGGGGCATCGCACCGGTGAGTAATATTGTACGGCTA
           >>RB2
AGAGCGAATTTGGCCTGTAGACCTCAATTGCGAGCTTTCTAATTTCAAAC
          >>RB3
TATTCGGGCCTAACTTTTGGTGTGATGATGCTGACTGGCAGGATATATAC

CGTTGTAATTGCCTAACTTTTGGTGTGATGATGCTGACTG
```

Fig. 4.17 T-DNA right border sequences with three T-RB nested primers in colour

4.5.1. Insertion at Chromosome 3

Assigning location and function to the obtained sequence is an essential step. The PCR- amplified products were resolved with low melting Agarose 1% (w/v) and 1.9Kb DNA was eluted and purified using the QIAGEN QIAquick PCR purification kit (USA). The gel eluted PCR products were sequenced using the T-DNA specific primer [T-RB 2 (5′ AGCGAATTTGGCCTGTAGACCTCA-3′)] and adaptor-specific primer [AP2 (5′ - ACTATAGGGCACGCGTGGT-3′)]. While isolating the genomic sequence for *OsPE* only one amplicon was obtained with one set of primers. The sequencing with T-RB2 and AP2 showed the 556 and 559 nucleotide sequences respectively (Fig. 4.19 and Fig. 4.21). The chromosomal location of the *OsPE* gene was determined using sequence of *japonica* rice cultivar Nipponbare. The similarity searches of the T-DNA flanking sequences (T-RB2 and AP2) using BLAST (Altschul, 1990) were done (www.ncbi.nlm-.nih.gov/BLAST) (Fig. 4.20 and Fig. 4.22). BLAST results of AP2 showed 87% identity with *Oryza sativa* cv. Nipponbare chromosome 3 whereas T-RB2 showed highly significant alignment of 94% with rice cv. Nipponbare chromosome 3. Both showed significant alignment with the *Oryza sativa* cv. Nipponbare with only single hit at chromosome 3. Sequence obtained with T-RB2 corresponds to the gene present in the contig NC_008396.1.

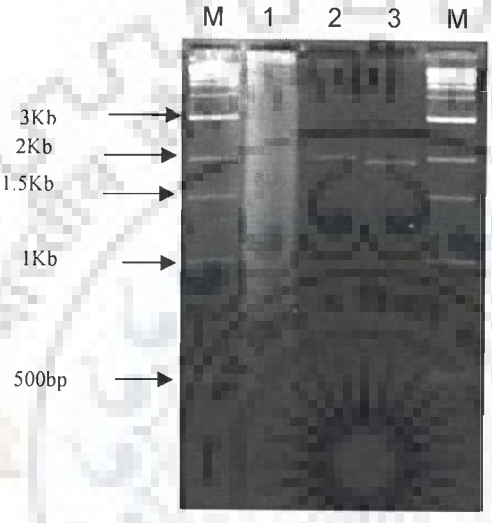

Fig. 4.18 PCR amplification of *EcoR*V digested library for identifying T-DNA flanking region in *OsPE* using T-RB and AP primers. Lane1: Primary PCR (T-RB1/AP1); Lane 2: Secondary PCR (T-RB2/AP2); Lane 3: Tertiary PCR (T-RB3/AP2); Lane M: Molecular weight marker 1kb

>T-RB2:*OsPE*

AGAGGCTTATCCTGGTCATGGCAAGCATAGGCGGGCCTAACTTTTGGTGTGATCGAT
GCTCATGCCGGGCTCATGCCGGGCTGGCACGGGAGGAGGCCAGGAGGGGGAGGAGC
GGCGTCCCACGTGCCGCCACGCGATTGCCGATGCTTCTAGGTCGGTTCCCCGCCGCG
CGGGCCCCACCAACCAGCGAGAGAAAACACGTTCCGCGCGCTACGCGCTCACTCAA
CCCTCCCCCTCTCCCTCGCTTACGGGGTGGGTCCCACACCACGGGCCAATCAGGGCC
CCGCCGCCTTCACCACCCTCGCCTGCGTTGGCCGACCAAAGCCACCCCGCCCTTTCCT
TTACGCCCACTGACGTGTGGGCCCCGAATCTCCCCCGGGTCCCCGTTGTTTTTTTTTG
TTGATTACCTGGTTCTCCGAAGCCCATTCAACCCCTGGGGTTAAAAATGGGGAAGGG
AAGGGCCCCCCCCCAGTTTTGGGTTCACCCCCCGCCCNANTTNTTGTGCTTCCCTTCCT
CTCCGGCTAATAAGTATTGGCCTCACCCNCCCCCCCTCAAAA

Fig. 4.19 T-DNA flanking sequence using T-RB2 (T-DNA based primer) for *OsPE* mutant

```
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 3
Score =  481 bits (260),  Expect = 3e-134
Identities = 300/317 (94%), Gaps = 12/317 (3%)
Strand=Plus/Plus

Query  68       GGCTCATGCCGGGCTGGCACGGGAGGAGGCCAGGAGGGGGAGGAGCGGCGTCCCACGTGC  127
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7447382  GGCTCATGCCGGGCTGGCACGGGAGGAGGCCAGGAGGGGGAGGAGCGGCGTCCCACGTGC  7447441

Query  128      CGCCACGCGATTGCCGATGCTTCTAGGTCGGTTCCCCGCCGCGCGGGCCCCACCAACCAG  187
                 |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  7447442  CGCCACGCGATTGCCGATGCTTCTAGGTCGGTTCCC-GCCGCGCGGGCCCCACCAACCAG  7447500

Query  188      CGAGAGAAAACACGTTCCGCGCGCTACGCGCTCACTCAACCCTCCCCCTCTCCCTCGCTT  247
                 ||||||||||||||||| || ||||||||||||||||||||| ||||||||||||||||||
Sbjct  7447501  CGAGAGAAAACACGTTTCGTGCGCTACGCGCTCACTCAACC-TCCCCCTCTCCCTCGCTT  7447559

Query  248      ACGGGGTGGGTCCCACACCACGGGCCAATCAGGGCCCCGCCGCCTTCACCACCCTCGCCT  307
                 ||||| |||||||||||||||||||||||||||||||| ||||| |||||||||| ||| |
Sbjct  7447560  ACGGG-TGGGTCCCACACCACGGGCCAATCAGGGCCC-GCCGC-TTCACCACCCCCGC-T  7447615

Query  308      GCGTTGGCCGACCAAAGCCACCCCGCCCTTTCCTTTACGCCCACTGACGTGTGGGCCCCG  367
                 |||||| || |||||||||||||| ||| ||||||||||||||||||||||||||||| |
Sbjct  7447616  GCGTTG-CCTACCAAAGCCACCC-GCC-TTTCCTTTACGCCCACTGACGTGTGGGCCC-G  7447671

Query  368      AATCTCCCCCGGGTCCC  384
                 ||||| ||| || ||||
Sbjct  7447672  AATCTTCCC-GG-TCCC  7447686
```

Fig. 4.20 BLAST result of T-DNA flanking sequence using T-RB2 (Genome Walking PCR) showing 94% identity with rice cultivar Nipponbare at Chromosome 3

>**AP2:*OsPE***

GTTTTTCCCCTGTGTTAGCACCATTGTGGAAGAGCTTGGCCTGTGATCTCTTTGTCCTC
TTTCTTCTTTTCGGCTTTTTCCCTTTCTTTCTGGAGCTTCAGCCACCAGAGCAATCGTC
AGTTCTGATTTAGAACGGACGACGGGAATGCAATGCAGAATTTTCTTTGAAATATAT
AGTAATGATGAAGTACCAATTTACACAAGCACACAGCAGGAAACAAATGTTGTGCT
ATCAGTTACAGAAGATAACACAATTAGGGCACGGACACTTTCACATAGAATAATTTG
GATAATCCTTAATATATTTGTAAACATGGGTTAGGTCACTGGAACTGGTAAACCCCC
TTACAATTCATAACCAAGCCACCTCTTGGGGCAACAGAGAATAATCTTTTTTTTCTTG
CAGTTAGTTTTCCNAACTAAAAACAAGGATAAGTTCAGTTAACTATTTTCTATTATTG
GTGAATTGTCCAAAAATTTTTTACCCACCAACATTGTACATTCTCCACAAATTAAATG
CCAANTATTGCCTTTCNCNAAAATACCTTACAAACCGGTTCATGTTTAAC

Fig. 4.21 T-DNA flanking sequence using AP2 for *OsPE* mutant

```
 Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 3
 Score =  547 bits (296),  Expect = 3e-154
 Identities = 447/510 (87%), Gaps = 51/510 (10%)
 Strand=Plus/Minus

Query  19       CACCATTGTGGAAGAGCTTGGCCTGTGAtct-ctttgtccnctttcnctttcggctt  77
                 ||||| ||||||||||||||  ||||| | || |||| | ||||||||||||||| || ||
Sbjct  7449191  CACCA-TGTGGAAGAGCTT-TCCTGTCA-CTCCTTT-T-CTCTTTCTTCTTTTCAGCCTT  7449137

Query  78       ttncnttttttttcGAGCTTCAGCCACCAGAGCAATCGTCAGTTCTGATTTAGAACGGA  137
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449136  TTCCCTTTCTTTCTGGAGCTTCAGCCACCAGAGCAATCGTCAGTTCTGATTTAGAACGGA  7449077

Query  138      CGACGGGAATGCAATGCAGAATTTTCTTTGAAATATATAGTAATGATGAAGTACCAATTT  197
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449076  CGACGGGAATGCAATGCAGAATTTTCTTTGAAATATATAGTAATGATGAAGTACCAATTT  7449017

Query  198      ACACAAGCACACAGCAGGAAACAAATGTTGTGCTATCAGTTACAGAAGATAACACAATTA  257
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449016  ACACAAGCACACAGCAGGAAACAAATGTTGTGCTATCAGTTACAGAAGATAACACAATTA  7448957

Query  258      GGGCACGGACACTTTCACATAGAATAATTTGGATAATCCTTAATATATTTGTAAACATGG  317
                 |||||||||||||| |||||||||||||| |||||||||||||||||| | |||||
Sbjct  7448956  GGGCACGGACACTT-CACATAGAATAATTTG-ATAATCCTTAATATATTTGTTA-CATGG  7448900

Query  318      GTTAGGTCACTGGAACTGGTAAACCCCCTTACAATTCATAACCAAGCCACCTCTTGGGGC  377
                  |||| |||| | ||||| ||| |||  ||||||| |||||| |||| || ||||||  |
Sbjct  7448899  -TTAG-TCACAG-AACTG-TAA-CCC--TTACAAT-CATAAC-AAGC-AC-TCTTGG--C  7448853

Query  378      AACAGAGAATAATCtttttttCTTGCAGTTAGTTTTCCNAACTAAAAACAAGGATAAGT  437
                 |||||||||||| ||      || ||||| ||||||| |||||| |||||||| ||
Sbjct  7448852  AACAGAGAATAA-CTAAAC---CT-GCAGT-AGTTTTCG-AACTAAA--CAAGGATA-GT  7448803

Query  438      TCAGTTAACTATTTTCTATTATTGGTGAATTGTCCAAAAATTTTTTACCCACCAACATTG  497
                  |||||| |||||| |||| || |||||||  | | |||| ||||   ||| || |||||  |
Sbjct  7448802  -CAGTTA-CTATTT-CTAT-AT-GGTGAAT-G-C-AAAA-TTTT--ACC-AC-AACAT-G  7448757

Query  498      TACATTCTCCACAAATTAAATGCCAANTAT  527
                 ||||| ||| |||||| || ||| || |||
Sbjct  7448756  TACAT-CTC-ACAAAT-AA-TGC-AAGTAT  7448732
```

Fig. 4.22 BLAST result of T-DNA flanking sequence using AP2 (Genome Walking PCR) showing 87% identity with rice cultivar Nipponbare at Chromosome 3

4.6 Thermal Asymmetric Interlaced PCR (TAIL-PCR)

Another technique to identify and isolate the flanking region to a known insertion in the genome is TAIL-PCR. Three nested T-DNA primers (T-RB1, T-RB2 and T-RB3) and an arbitrary degenerate (AD) primer was designed to perform consecutive TAIL-PCR reactions in *OsPE* mutant. These nested primers together with AD primer having a lower Tm (melting temperature) was used for relative amplification efficiencies of specific and non- specific products.

The primary reaction was carried out using T-RB1 and AD1 primers. The primary PCR gel showed multiple bands due to amplification of non specific products, as AD1 primer annealed throughout the genome due to low stringency cycle. The secondary TAIL-PCR was performed using T-RB2/AD1 primers and 50 times diluted primary PCR product as template. By resolving secondary PCR product on 1.2% agarose gel showed amplification of single fragment of 600bp. The specific product was preferentially amplified over non specific ones because of two high stringency PCR cycles interspersed by one low stringency cycle. The specific amplification was also obtained in tertiary TAIL-PCR using T-RB3/ AD1 primers and ten times diluted secondary PCR product as template (Fig. 4.23). Amplified product of tertiary TAIL-PCR was run on 1.2% agarose gel, eluted and purified. About 8ng of purified PCR product of tertiary TAIL-PCR was used for sequencing reaction using T-RB3 primer. The sequence result showed 511nucleotide sequence (Fig. 4.24). When similarity search was done using NCBI BLAST it hit only chromosome 3 at 7447382 position with 92% identity with *Oryza sativa* cv. Nipponbare (Fig. 4.25). Thus, TAIL-PCR also validated the T-DNA insertion on chromosome 3 as shown by Genome Walking experiment.

The gene starts at the position 7447811 and stops at 7450646 (Fig 4.26). Although the insertion is in the promoter region at 777382 which is 425bp upstream the gene Os03g0241300. *OsPE* candidate gene was mapped to chromosome 3 with 2 UTRs (1473bp and 193bp), 2 Exons (154bp and 935bp), and 1 Intron (81bp) (Fig 4.27, Fig. 4.28 and Table 4.4). The length of *OsPE* candidate gene was 2.8Kb which encodes for a hypothetical protein of 362 amino acids.

4.7. Confirmation of T-DNA insertion on chromosome 3

To confirm the insertion on chromosome 3, primers were designed for rice genome flanking the T-DNA borders (Genome-specific primers, PE-RB1, PE-RB2/PE-LB1, and PE-LB2) (Fig.4.29). PCR was done in combination with the T-DNA (T-RB) based primers and genome-specific primers (Table 4.5) and amplification of expected size was obtained in *OsPE* mutant, but not in Basmati 370 (Fig 4.30a). PCR with genome-specific primers gave amplification in Basmati 370 and not in polyembryonic mutant because of the insertion of approximately 10Kb fragment of the cassette which cannot be amplified by the Taq polymerase used for the this PCR (Fig 4.30b). All the 14 polyembryonic *hpt* positive F_3 lines also showed amplification with the T-DNA specific (T-RB1) and genome specific (PE-RB1) primers (Fig 4.31). Later T-DNA specific and genome specific primers were applied on the F_2 population. Out of 138 F_2 individuals 49 showed amplification.No results were obtained using T-LB primers. This may be because of the loss of T-DNA left border during transformation.

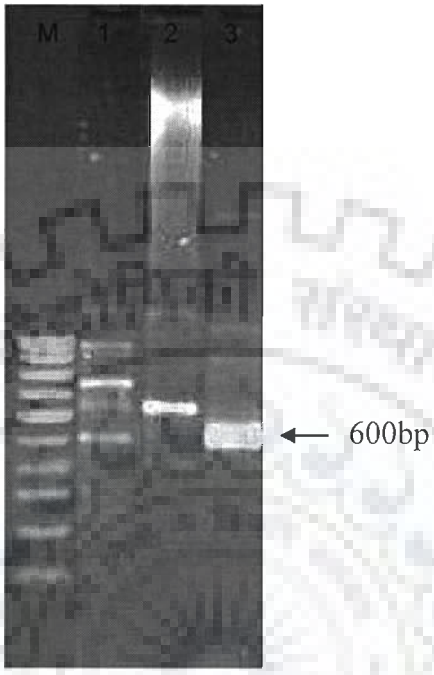

Fig. 4.23 PCR amplification of Primary, Secondary and Tertiary TAIL-PCR for identifying T-DNA flanking region in *OsPE*. Lane1: Primary PCR (T-RB1/AD1); Lane 2: Secondary PCR (T-RB2/AD1); Lane 3: Tertiary PCR (T-RB3/AD1); Lane M: Molecular Weight Marker 1Kb

>TRB3:*OsPE*

TGGCAAGCATAGGCGGGCCTAACTTTTGGTGTGATCGATGCTCATGCCGGGCTCATG
CCGGGCTGGCCGGGAGGAGGCCAGAGGGGGAGGAGCGGCGTCCCACGTGCCGCCAC
GCGATTGCCGATGCTTCTAGGTCGTTCCCCGCCGCGCGGGCCCCACCAACCAGCGAG
AGAAACACGTTCCGCGCGCTACGCGTCACTCAACCCTCCCCCTCTCCCTCGCTTACGG
GTGGGTCCCACACCACGGGCCAATCAGGGCCCCGCCGCCTTCACCACCCTCGCCTGC
GTTGGCCGACCAAAGACCCCGCCCTTTCCTTTACGCCCACTGACGTGTGGGCCCGGG
AATCTCCCCCGGGTCCCCGTTGTTTTTTTTTGTTGATTACCTGGTTCTCCGAACTGCCC
AATTCAACCCCTGGGGTTAAAAATGGGGAAGGGAAGGGCCCCCCCCAGTTTTGGGTT
CACCCCCCGCCCNANTTNTTGTGCCCTTCCTCTCCGGCTAATAAGTATTGGCCTCACC
CNCCCCCCCTCAAAA

Fig. 4.24 T-DNA flanking sequence using T-RB3 (Tertiary TAIL- PCR product) for *OsPE* mutant

```
 Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 3
 Score =  440 bits (238),  Expect = 4e-122
 Identities = 293/316 (92%), Gaps = 17/316 (5%)
 Strand=Plus/Plus

Query  50       GGCTCATGCCGGGCTGGC-CGGGAGGAGGCCA-GAGGGGGAGGAGCGGCGTCCCACGTGC  107
                ||||||||||||||||| |||||||||||| |||||||||||||||||||||||||||
Sbjct  7447382  GGCTCATGCCGGGCTGGCACGGGAGGAGGCCAGGAGGGGGAGGAGCGGCGTCCCACGTGC  7447441

Query  108      CGCCACGCGATTGCCGATGCTTCTAGGTCG-TTCCCCGCCGCGCGGGCCCCACCAACCAG  166
                |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||||
Sbjct  7447442  CGCCACGCGATTGCCGATGCTTCTAGGTCGGTTCCC-GCCGCGCGGGCCCCACCAACCAG  7447500

Query  167      CGAGAGAAA-CACGTTCCGCGCGCTACGCG-TCACTCAACCCTCCCCCTCTCCCTCGCTT  224
                ||||||||| |||||| || |||||||||| |||||||||| ||||||||||||||||||
Sbjct  7447501  CGAGAGAAAACACGTTTCGTGCGCTACGCGCTCACTCAACC-TCCCCCTCTCCCTCGCTT  7447559

Query  225      ACGGGTGGGTCCCACACCACGGGCCAATCAGGGCCCCGCCGCCTTCACCACCCTCGCCTG  284
                |||||||||||||||||||||||||||||||||||||| ||||| |||||||||| ||| ||
Sbjct  7447560  ACGGGTGGGTCCCACACCACGGGCCAATCAGGGCCC-GCCGC-TTCACCACCCCCGC-TG  7447616

Query  285      CGTTGGCCGACCAAAGAC-CCCGCCCTTTCCTTTACGCCCACTGACGTGTGGGCCCGGGA  343
                ||||| || ||||||| | |||||| |||||||||||||||||||||||||||||||  |
Sbjct  7447617  CGTTG-CCTACCAAAGCCACCCGCC-TTTCCTTTACGCCCACTGACGTGTGGGCCCG--A  7447672

Query  344      ATCTCCCCCGGGTCCC  359
                |||| ||| || ||||
Sbjct  7447673  ATCTTCCC-GG-TCCC  7447686
```

Fig. 4.25 BLAST result of T-DNA flanking sequence using T-RB3 (TAIL-PCR) showing 92% identity with rice cultivar Nipponbare at Chromosome 3

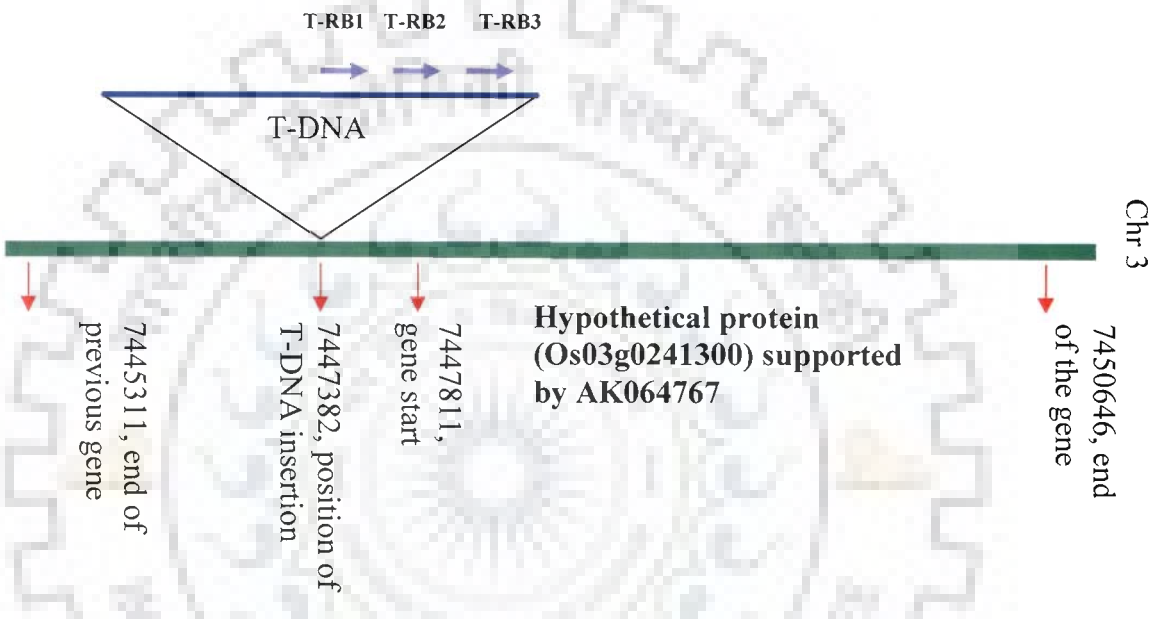

Fig. 4.26 Diagrammatic representation of T-DNA insertion on chromosome 3 in Polyembryony mutant of Basmati 370

>ref|NC_008396.1 Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 3

Fig. 4.27 Sequence of candidate gene *OsPE* including the promoter region in *Oryza sativa* cv. Nipponbare genomic DNA Chromosome 3

(Red: end of previous gene; green: Promoter of candidate gene); Blue: UTR region; Orange Underlined : Exon; black: Intron

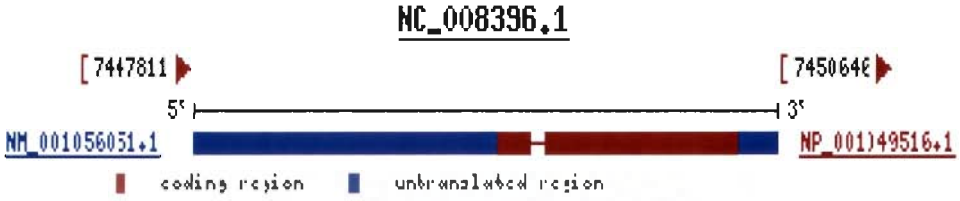

Fig. 4.28 Diagrammatic representation of the candidate gene *OsPE* showing UTRs, coding and non coding regions

Table 4.4 Total size (bp) of UTRs, Exons and Intron present in the candidate gene *OsPE*

7447811-7450646	NM_001056051.1	Os03g0241300	Size
7447811-7449283	Exon	UTR	1473bp
7449284-7449437	Exon	Exon	154bp
7449438-7449518	Intron		81bp
7449519-7450453	Exon	Exon	935bp
7450454-7450646	Exon	UTR	193bp

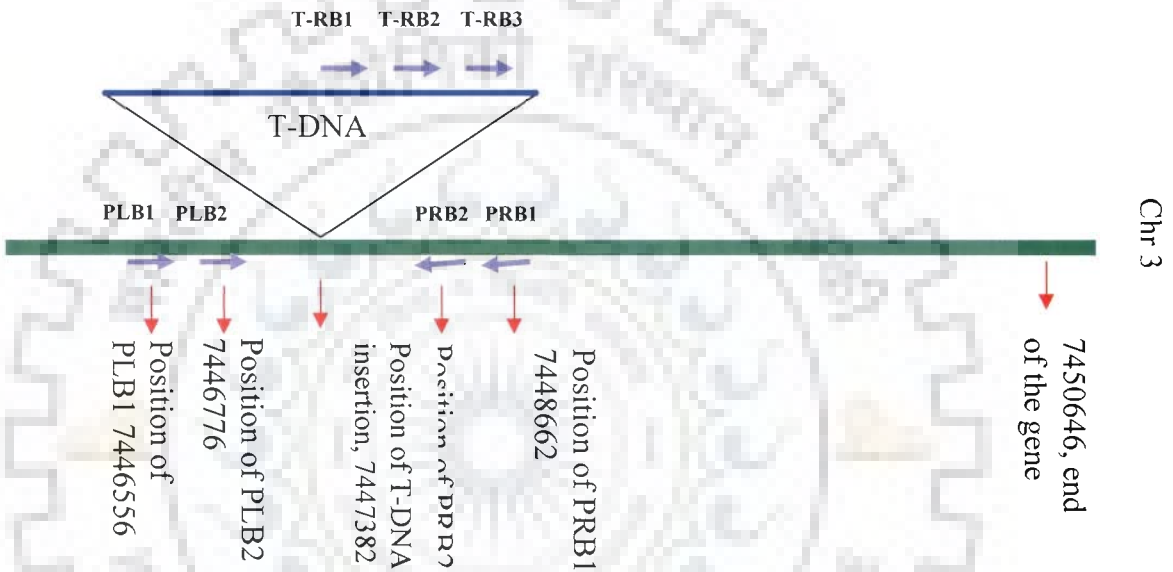

Fig. 4.29 Diagrammatic representation of designing rice genome specific primers and expected product size (Table4.5) in Polyembryony mutant of Basmati 370

Table 4.5 Primer combination of rice genome specific and T-DNA specific PCR and expected product size in Basmati 370 and *OsPE*

Primer combination	Expected size Basmati370 (bp)	Expected size in Polyembryony mutant(bp)
P-RB1 + T–RB1	-	1420
P-RB1 + T–RB2	-	1380
P-RB1 + T–RB3	-	1340
P-RB1 + T–RB4	-	1370
P-RB2 + T–RB1	-	893
P- RB2 + T–RB2	-	853
P-RB2 + T–RB3	-	810
P-RB2 + T–RB4	-	843
P-RB1 + P-LB1	2106	>10kb
P-RB1 + P-LB2	1886	>10kb
P-RB2 + P-LB1	1579	>10kb
P-RB2 + P-LB2	1359	>10kb

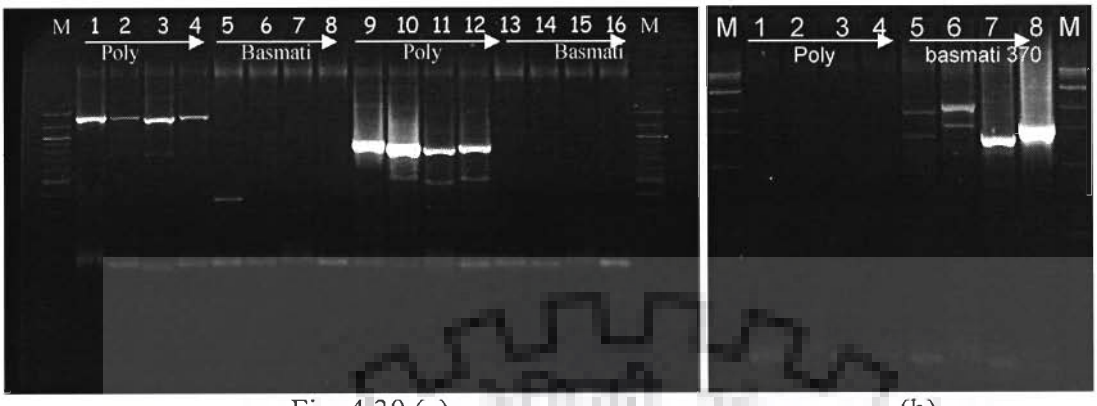

Fig. 4.30 (a) (b)

Fig. 4.30 (a) PCR amplification of *OsPE* and Basmati 370 genomic DNA using T-DNA specific and genome specific primers. Only mutant lines showed amplification. Lanes 1-4and 9-12: P-RB2/T-RB1,2,3,4; Lanes 5-8 and 13-16: P-RB1/T-RB1,2,3,4; M: Molecular weight marker 100bp.

(b) PCR amplification of *OsPE* and Basmati 370 DNA using genome specific primers (T-RB/LB). Only Basmati 370 showed amplification. Lanes 1,5: P-RB1/P-LB2; Lanes 2,6: P-RB1/P-LB1; Lanes 3,7: P-RB2/P-LB2; Lanes 4,8: P-RB2/P-LB1; Lane M: Molecular weight marker 1 Kb

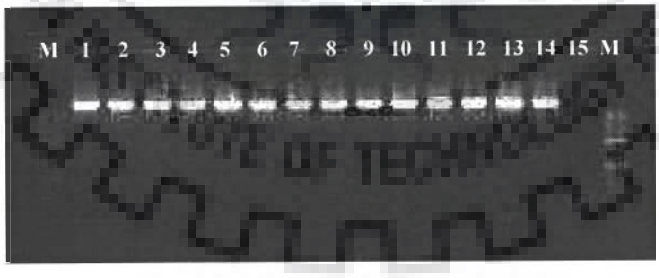

Fig. 4.31 PCR amplification of T-RB1/P-RB1 in polyembryonic lines from F_2 Population. Lanes 1-14: T-RB1/P-RB1 +ve; Lane M: Molecular weight marker 100bp

The T-DNA insert is found to be present in the promoter region of the candidate gene Os03g0241300. Corresponding mRNA and protein sequences were searched and the corresponding sequences are given in Fig 4.29 and 4.30, respectively. The function of Os03g0241300 (protein) has been reported as a hypothetical protein in *Oryza sativa*. At nucleotide level *OsPE* shared identity with *Arabidopsis, Zea mays* and *Vitis vinifera*. When the protein sequence was submitted to protein BLAST, the significant sequences that were found in the database matched mainly with *Oryza sativa* and little with *Arbidopsis, Vitis vinifera* and *Zea mays* (Fig. 4.34). The sequence showed some significant alignment with *Zea mays* as a stress inducible protein; hypothetical protein in *Vitis vinifera* and unknown protein in *Arabidopsis thaliana*. No conserved domains have been found in this protein. No *OsPE* paralog was found in rice. *OsPE* gene with evidence of transcription lacked functional homologs in other species. Although using KEGG search around 401 orthologs sequence were found throughout the living kingdom but majority had low identity value (> 40%) and most closely matches were with *Vitis* and *Arabidopsis* with hypothetical and unknown proteins respectively.

4.8 Cloning of *OsPE* gene of Basmati 370

On the basis of the sequences obtained by genome walking and TAIL-PCR BLAST search was done for the homologous sequence in Nipponbare *japonica* rice. The nucleotide sequence of the candidate gene *OsPE* in *japonica* rice was used to design seven pairs of overlapping PCR primers and the *OsPE* gene was amplified with Basmati 370 genomic DNA as template. These seven amplified products of *OsPE* gene in Basmati 370 was purified, gel eluted and sequenced using both forward and reverse primers. The sequences obtained using various forward and reverse primer pairs were aligned and the

sequence of full length *OsPE* gene in Basmati 370 was obtained as shown in Fig. 4.35. The sequence thus obtained was used for BLAST against *japonica* rice (Fig. 4.36). Ten SNPs and one gap were found in the *OsPE* candidate gene sequence as compared to Nipponbare sequence and thus leading to a total of 1% dissimilarity between the two sequences. The protein sequence was predicted from nucleotide sequence using GENSCAN (http://genes.mit.edu/GENSCAN). The variation was found only in two amino acids.

4.9 RT- PCR (Reverse Transcriptase PCR)

To examine the expression of *OsPE* gene in different plant organs, reverse transcription PCR analysis was performed with total RNA isolated from roots and shoots (Fig 4.37). The transcript of *OsPE* gene was detected in shoot.

To analyze the expression of *OsPE* gene, a relative expression of a reference gene (as control) was required. The reference gene model used was actin I (Os03g50890). Before applying those primers on RNA, it was applied on the genomic DNA to check any amplification. This is to avoid any non-specific amplification of the total RNA if contaminated with genomic DNA. No amplification was obtained with genomic DNA of the mutant (Fig 4.38). With RT-PCR the expression was found in Basmati 370 shoots (Fig 4.39). It can thus be speculated that *OsPE* gene is expressing in the shoot.

Large number of Expressed Sequence Tags (ESTs) were available in rice. Total EST database was categorized into eight pools corresponding to the different parts of the plant where expression was detected. Expression profile of the candidate gene suggested by analysis of EST counts showed maximum ESTs in the seed followed by flower and panicle. The Gene EST/ Total EST in pools are given in Table 4.6.

>gi|115451830|ref|NM_001056051.1| Oryza sativa (japonica cultivar-group) Os03g0241300 (Os03g0241300) mRNA, complete cds

GAGGCTCCCTTTTTTCTCTTCTTCCCCTCCCGCAACCCTAATTTTTTTTCCCCATCCGCGAAACCCTAGC
CGCCACGCGAAACCAAATCCCGCCGCGCGGGATCCTTTTCCGCCGGATTCCACCCGCGAATCGGGGTTCC
CCTTACGATTCGCGGGCGGATTAGCGCGAGGCGCGCCTCCCCCTACCTCTGTGTGATCCGGGGGTGAGGT
TAGGCCGGACGCCGGGGCATCAGCCATGTCGAGGTGCTTCCCCTACCCGCCGCCGGGGTACGTGCGAAAC
CCAGTGGTGGCCGTGGCCGCGGCCGAAGCGCAGGCGACCACTAAGGTTTGTTGAACCATCGGATTTACAC
ACGCACGTGCCGGATCATTTGCTCTTGCCTGTTGGTTTTGATCGGATCTGTTGGTTGTGCGTGTGTGATT
TGGGGATCGCACGTGCGGGGAAGCTAACCTTTGCATGGATAACTTGAGATTTGTGAGGCCGCGCTTCGAC
CAGATCGGTCGCCAATCTTTTAGTGGCTGACCGTGGAAAGAGGATATTACTGACCTTCGGTTTGCTAATT
TTGGTTGTGCCGTTGAATCTGAAATAACCAGAATAGTCATGGGGAAAAAAGTCTGATCTGGAAGGTTCGA
ATTACATTTCTATATATTGTTGTGCTCCCAGACGATGGTTGCAAGAAATCACTCATGCTGGATAAAATTG
TGGATGTAAGAGTCTGCAGTCGTTAAAATCTGGAAACAGCACATTTTGCCGTAGTAAATTTGAATCCATG
TTGCTGTCTCGTTATTGGTGTGTTACGAGTAACCTGTGTGTTGTTATCTCCGCTTGGACTAGATTCCAAG
TAATCCAGTGCCTTCATGACCTGCAAATTCTATGCCTATGAAGTAACATGAACAGTTTGTATGTATGTAT
TCTGTTGATGCATACTTGCATTATTTGTGAGATGTACATGTTGTGGTAAAATTTTGCATTCACCATATAG
AAATAGTAACTGACTATCCTTGTTTAGTTCGAAAACTACTGCAGGTTTAGTTATTCTCTGTTGCCAAGAG
TGCTTGTTATGATTGTAAGGGTTACAGTTCTGTGACTAACCATGTAACAAATATATTAAGGATTATCAAA
TTATTCTATGTGAAGTGTCCGTGCCCTAATTGTGTTATCTTCTGTAACTGATAGCACAACATTTGTTTCC
TGCTGTGTGCTTGTGTAAATTGGTACTTCATCATTACTATATATTTCAAAGAAAATTCTGCATTGCATTC
CCGTCGTCCGTTCTAAATCAGAACTGACGATTGCTCTGGTGGCTGAAGCTCCAGAAAGAAAGGGAAAAGG
CTGAAAAGAAGAAAGAGAAAAGGAGTGACAGGAAAGCTCTTCCACATGGTGAGATATCCAAGCATTCAAA
GCGAACCCACCACAAGAAGAGAAAACATGAAGACATCAATAATGCTGATCAGAAGTCCCGGAAGGTTTCC
TCCATGGAACCTGGTGAGCAATTGGAGAAGAGTGGACTCTCAGAAGAGCATGGAGCTCCTTGCTTTACTC
AGACAGAGCATGGCTCTCCAGAGAGTTCACAGGACAGCAGCAAGAGAAGAAAGGTTGTGTTACCCAGTCC
TAGCCAAGCTAAGAATGGTAACATCCTTCGAATAAAGATAAGAAGAGATCAAGATTCTTCAGCTTCCCTT
TCGGAGAAATCTAATGTTGTACAAACACCAGTTCATCAAATGGGATCAGTTTCATCTCTGCCAAGTAAGA
AAAACTCAATGCAACCACACAACACCGAAATGATGGTGAGAACAGCATCAACCCAGCAGCAAAGCATCAA
AGGTGATTTTCAAGCAGTACCGAAACAAGGTATGCCAACCCCAGCAAAAGTCATGCCAAGAGTCGATGTT
CCTCCATCTATGAGGGCATCAAAGGAAAGGATTGGCCTTCGTCCTGCAGAGATGTTGGCCAATGTTGGTC
CTTCACCCTCCAAGGCAAAACAGATTGTCAATCCTGCAGCTGCTAAGGTTACACAAAGAGTTGATCCTCC
ACCTGCCAAGGCATCTCAGAGAATTGATCCTCTGTTGCCATCCAAGGTTCATATAGATGCTACTCGATCT
TTTACGAAGGTCTCCCAGACAGAGATCAAGCCGGAAGTACAGCCCCCAATTCTGAAGGTGCCTGTGGCTA
TGCCTACCATCAATCGTCAGCAGATTGACACCTCGCAGCCCAAAGAAGAGCCTTGCTCCTCTGGCAGGAA
TGCTGAAGCTGCTTCAGTATCAGTAGAGAAGCAGTCCAAGTCAGATCGCAAAAAGAGCCGCAAGGCTGAG
AAGAAAGAGAAGAAGTTCAAAGATTTATTTGTTACCTGGGATCCTCCGTCTATGGAAATGGATGATATGG
ATCTCGGGGACCAGGATTGGCTGCTTGATAGTACGAGGAAACCTGATGCTGGCATTGGCAACTGCAGAGA
AATTGTTGATCCACTTACTTCTCAATCAGCAGAGCAGTTCTCATTGCAGCCTAGGGCGATTCATTTACCA
GACCTTCATGTCTATCAGTTGCCATATGTGGTTCCATTCTAGGTTTGTGTAGTGAGATGGAGTAGGTGAG
AAGTAGAGAGATGTTGGGAGAGAGCTGTGTGGGTCTGGGAGATTATGGTTCCCTGCCACAGTTTTCCAGC
TTTGTTCCAGAGCGTTTCTTGTTTCAGGTTGCTGAGCCTGTCAAACTTCATGTAATGTTGTATTTTTGCT
TCGTTTATTGCAGCAATTGAACTTC

Fig. 4.32 Complete cDNA sequence corresponding to *OsPE* (candidate gene) in *Oryza sativa* cv.Nipponbare genomic DNA chromosome 3

>gi|115451831|ref|NP_001049516.1| Os03g0241300 [Oryza sativa (japonica cultivar-group)]
MEPGEQLEKSGLSEEHGAPCFTQTEHGSPESSQDSSKRRKVVLPSPSQAKNGNILRIKIRRDQDSSASLS
EKSNVVQTPVHQMGSVSSLPSKKNSMQPHNTEMMVRTASTQQQSIKGDFQAVPKQGMPTPAKVMPRVDVP
PSMRASKERIGLRPAEMLANVGPSPSKAKQIVNPAAAKVTQRVDPPPAKASQRIDPLLPSKVHIDATRSF
TKVSQTEIKPEVQPPILKVPVAMPTINRQQIDTSQPKEEPCSSGRNAEAASVSVEKQSKSDRKKSRKAEK
KEKKFKDLFVTWDPPSMEMDDMDLGDQDWLLDSTRKPDAGIGNCREIVDPLTSQSAEQFSLQPRAIHLPD
LHVYQLPYVVPF

Fig. 4.33 Protein sequence corresponding to *OsPE* (candidate gene) in *Oryza sativa* cv.Nipponbare genomic DNA chromosome 3

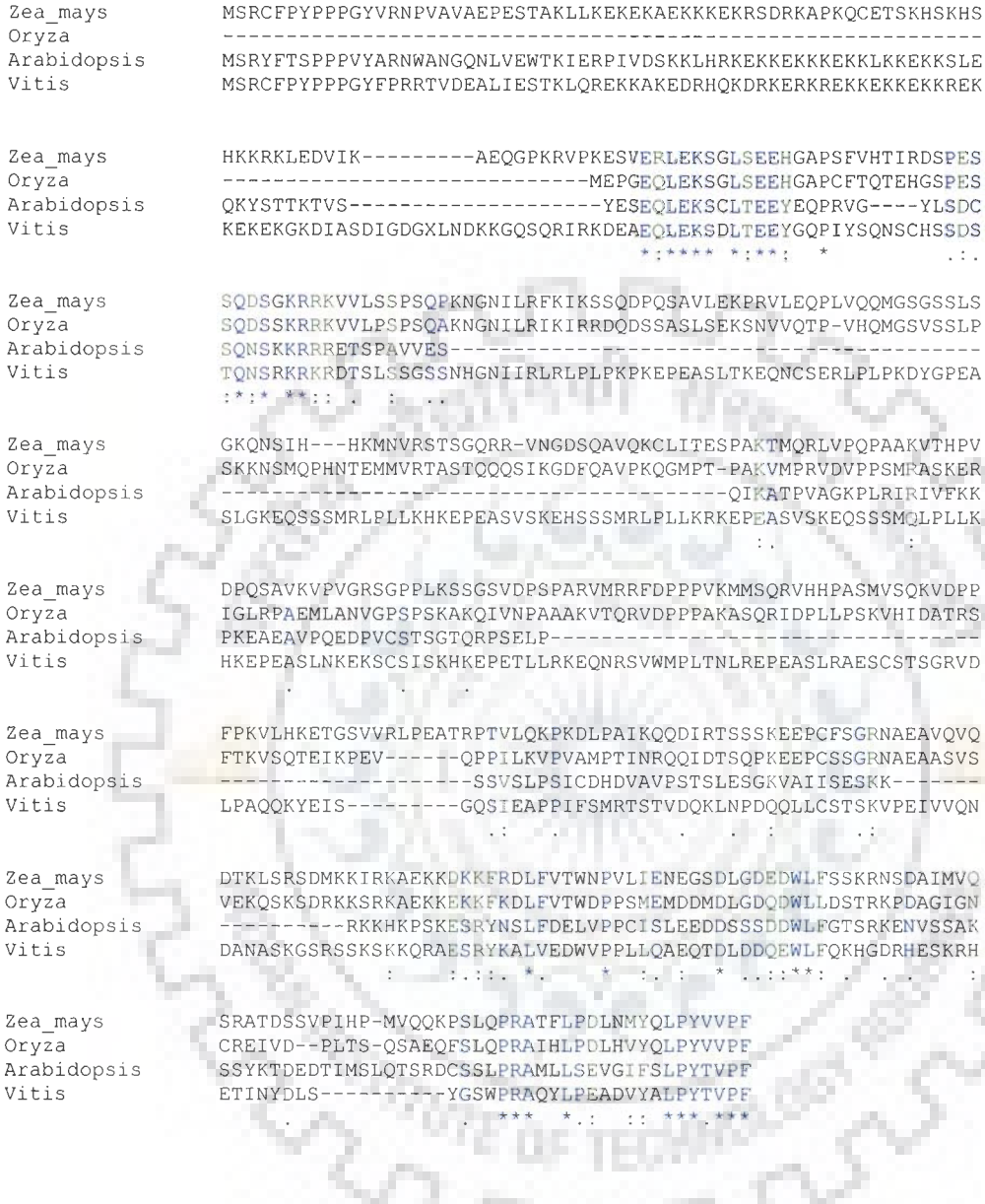

Fig. 4.34 Protein BLAST of *Oryza sativa* (Os03g0241300) with *Arabidopsis, Zea mays* and *Vitis vinifera*

```
Consensus key
* - single, fully conserved residue
: - conservation of strong groups
. - conservation of weak groups
  - no consensus
CLUSTAL W (1.81) multiple sequence alignment
```

Fig. 4.35 *OsPE* candidate gene sequence in Basmati 370 showing SNPs (in colour) with the corresponding Nipponbare sequence

Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 3
 Score = 5155 bits (2791), Expect = 0.0
 Identities = 2822/2836 (99%), Gaps = 5/2836 (0%)
 Strand=Plus/Plus

```
Query  1        AGGCTCCCTTTTTTCTCTTCTTCCCCTCCCGCAACCCTAAtttttttCCCCATCCGCGA  60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7447812  AGGCTCCCTTTTTTCTCTTCTTCCCCTCCCGCAACCCTAATTTTTTTTCCCCATCCGCGA  7447871

Query  61       AACCCTAGCCGCCACGCGAAACCAAATCCCGCCGCGCGGGATCCTTTTCCGCCGGATTCC  120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7447872  AACCCTAGCCGCCACGCGAAACCAAATCCCGCCGCGCGGGATCCTTTTCCGCCGGATTCC  7447931

Query  121      ACCCGCGAATCGGGGTTCCCCTTACGATTCGCGGGCGGATTAGCGCGAGGCGCGCCTCCC  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7447932  ACCCGCGAATCGGGGTTCCCCTTACGATTCGCGGGCGGATTAGCGCGAGGCGCGCCTCCC  7447991

Query  181      CCTACCTCTGTGTGATCCGGGGGTGAGGTTAGGCCGGACGCCGGGGCATCAGCCATGTCG  240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7447992  CCTACCTCTGTGTGATCCGGGGGTGAGGTTAGGCCGGACGCCGGGGCATCAGCCATGTCG  7448051

Query  241      AGGTGCTTCCCCTACCCGCCGCCGGGGTACGTGCGAAACCCAGTGGTGGCCGTGGCCGCG  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448052  AGGTGCTTCCCCTACCCGCCGCCGGGGTACGTGCGAAACCCAGTGGTGGCCGTGGCCGCG  7448111

Query  301      GCCGAAGCGCAGGCGACCACTAAGGTTTGTTGAACCATCGGATTTACACACGCACGTGCC  360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448112  GCCGAAGCGCAGGCGACCACTAAGGTTTGTTGAACCATCGGATTTACACACGCACGTGCC  7448171

Query  361      GGATCATTTGCTCTTGCCTGTTGGTTTTGATCGGATCTGTTGGTTGTGCGTGTGTGATTT  420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448172  GGATCATTTGCTCTTGCCTGTTGGTTTTGATCGGATCTGTTGGTTGTGCGTGTGTGATTT  7448231

Query  421      GGGGATCGCACGTGCGGGGGAAGCTAACCTTTGCATGGATAACTTGAGATTTGTGAGGCCG  480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448232  GGGGATCGCACGTGCGGGGGAAGCTAACCTTTGCATGGATAACTTGAGATTTGTGAGGCCG  7448291

Query  481      CGCTTCGACCAGATCGGTCGCCAATCTTTTAGTGGCTGACCGTGGAAAGAGGATATTACT  540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448292  CGCTTCGACCAGATCGGTCGCCAATCTTTTAGTGGCTGACCGTGGAAAGAGGATATTACT  7448351

Query  541      GACCttcggtttgctaattttggttgtgccgttgaatctgaaataaccagaaTAGTCATG  600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448352  GACCTTCGGTTTGCTAATTTTGGTTGTGCCGTTGAATCTGAAATAACCAGAATAGTCATG  7448411

Query  601      GGGAAAAAAGTCTGATCTGGAAGGTTCGAATAACATTTCTATATATTGTTGTGCTCCCAG  660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448412  GGGAAAAAAGTCTGATCTGGAAGGTTCGAATTACATTTCTATATATTGTTGTGCTCCCAG  7448471

Query  661      ACGATGGTTGCAAGAAATCACTCATGCTGGATAAAATTGTGGATGTAAGAGTCTGCAGTC  720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448472  ACGATGGTTGCAAGAAATCACTCATGCTGGATAAAATTGTGGATGTAAGAGTCTGCAGTC  7448531

Query  721      GTTAAAATCTGGAAACAGCACATTTTGCCGTAGTAAATTTGAATCCATGTTGGTGTCTCG  780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448532  GTTAAAATCTGGAAACAGCACATTTTGCCGTAGTAAATTTGAATCCATGTTGCTGTCTCG  7448591

Query  781      TTATTGGTGTGTTACGAGTAACCTGTGTGTTGTTATCTCCGCTTGGACTAGATTCCAAGT  840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448592  TTATTGGTGTGTTACGAGTAACCTGTGTGTTGTTATCTCCGCTTGGACTAGATTCCAAGT  7448651

Query  841      AATCCAGTGCCTTCATGACCTGCAAATTCTATGCCTATGAAGTAACATGAACAGTTTGTA  900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448652  AATCCAGTGCCTTCATGACCTGCAAATTCTATGCCTATGAAGTAACATGAACAGTTTGTA  7448711

Query  901      TGTATGTATTGTGTTGATGCATACTTGCATTATTTGTGAGATGTACATGTTGTGGTAAAA  960
                |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448712  TGTATGTATTCTGTTGATGCATACTTGCATTATTTGTGAGATGTACATGTTGTGGTAAAA  7448771
```

```
Query  961      TTTTGCATTCACCATATAGAAATAGTAACTGACTATCCTTGTTTAGTTCGAAAACTACTG  1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448772  TTTTGCATTCACCATATAGAAATAGTAACTGACTATCCTTGTTTAGTTCGAAAACTACTG  7448831

Query  1021     CAGGTTTAGTTATTCTCTGTTGCCAAGAGTGCTTGTTATGATTGTAAGGGTTACAGTTCT  1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448832  CAGGTTTAGTTATTCTCTGTTGCCAAGAGTGCTTGTTATGATTGTAAGGGTTACAGTTCT  7448891

Query  1081     GTGACTAACCATGTAACAAATATATTAAGGATTATCAAATTATTCTATGTGAAGTGTCCG  1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7448892  GTGACTAACCATGTAACAAATATATTAAGGATTATCAAATTATTCTATGTGAAGTGTCCG  7448951

Query  1141     TGCCCTAATTGTGTTATCTTCTGTAACTGATAGCCCAACATTTGTTTCCTGCTGTGTGCT  1200
                |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  7448952  TGCCCTAATTGTGTTATCTTCTGTAACTGATAGCACAACATTTGTTTCCTGCTGTGTGCT  7449011

Query  1201     TGTGTAAATTGGTACTTCATCATTACTATATATTTCAAAGAAAATTCTGCATTGCATTCC  1259
                |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct  7449012  TGTGTAAATTGGTACTTCATCATTACTATATATTTCAAAGAAAATTCTGCATTGCATTCC  7449071

Query  1260     CGTCGTCCGTTCTAAATCAGAACTGACGATTGCTCTGGTGGCTGAAGCTCCagaaagaaa  1319
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449072  CGTCGTCCGTTCTAAATCAGAACTGACGATTGCTCTGGTGGCTGAAGCTCCAGAAAGAAA  7449131

Query  1320     agg-aaaggctgaaaagaagaaagagaaaaggagtgacaggaaagCTCTTCCACATGGTG  1378
                ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449132  GGGAAAAGGCTGAAAAGAAGAAAGAGAAAAGGAGTGACAGGAAAGCTCTTCCACATGGTG  7449191

Query  1379     AGATATCCAAGCATTCAAAGCGAACCCACCACAAGAAGAGAAAACATGAAGACATCAATA  1438
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449192  AGATATCCAAGCATTCAAAGCGAACCCACCACAAGAAGAGAAAACATGAAGACATCAATA  7449251

Query  1439     ATGCTGATCAGAAGTCCCGGAAGGTTTCCTCGATGGAACCTAGTGAGCAATTGGAGAAGA  1498
                |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Sbjct  7449252  ATGCTGATCAGAAGTCCCGGAAGGTTTCCTCCATGGAACCTGGTGAGCAATTGGAGAAGA  7449311

Query  1499     GTGGACTCTCAGAAGAGCATGGAGCTCCTTGCTTTACTCAGACAGAGCATGGCTCTCCAG  1558
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449312  GTGGACTCTCAGAAGAGCATGGAGCTCCTTGCTTTACTCAGACAGAGCATGGCTCTCCAG  7449371

Query  1559     AGAGTTCACAGGACAGCAGCAAGAGAAGAAAGGTTGTGTTACCCAGTCCTAGCCAAGCTA  1618
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449372  AGAGTTCACAGGACAGCAGCAAGAGAAGAAAGGTTGTGTTACCCAGTCCTAGCCAAGCTA  7449431

Query  1619     AGAATGGTGAGGCCCTTTCTTGCATTTGTCTTCTTTTAGCTGGTGATGTTGTATTGGTTT  1678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct  7449432  AGAATGGTGAGGCCCTTTCTTGCATTTGTCTTCTTTTAGCTGGTGATGTTGAATTGGTTT  7449491

Query  1679     GACTTATCCTGAATTATCATGTTGCAGGTAACATCCTTCGAATAAAGATAAGAAGAGATC  1738
                |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  7449492  GACTTATCCTGAATTATCATCTTGCAGGTAACATCCTTCGAATAAAGATAAGAAGAGATC  7449551

Query  1739     AAGATTCTTCAGCTTCCCTTTCGGAGAAATCTAATGTTGTACAAACACCAGTTCATCAAA  1798
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449552  AAGATTCTTCAGCTTCCCTTTCGGAGAAATCTAATGTTGTACAAACACCAGTTCATCAAA  7449611

Query  1799     TGGGATCAGTTTCATCTCTGCCAAGTAAGAAAAACTCAATGCAACCACACAACACCGAAA  1858
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449612  TGGGATCAGTTTCATCTCTGCCAAGTAAGAAAAACTCAATGCAACCACACAACACCGAAA  7449671

Query  1859     TGATGGTGAGAACAGCATCAACCCAGCAGCAAAGCATCAAAGGTGATTTTCAAGCAGTAC  1918
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449672  TGATGGTGAGAACAGCATCAACCCAGCAGCAAAGCATCAAAGGTGATTTTCAAGCAGTAC  7449731

Query  1919     CGAAACAAGGTATGCCAACCCCAGCAAAAGTCATGCCAAGAGTCGATGTTCCTCCATCTA  1978
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  7449732  CGAAACAAGGTATGCCAACCCCAGCAAAAGTCATGCCAAGAGTCGATGTTCCTCCATCTA  7449791
```

```
Query  1979     TGAGGGCATCAAAGGAAAGGATTGGCCTTCGTCCTGCAGAGATGTTGGCCAATGTTGGTC  2038
Sbjct  7449792  TGAGGGCATCAAAGGAAAGGATTGGCCTTCGTCCTGCAGAGATGTTGGCCAATGTTGGTC  7449851

Query  2039     CTTCACCCTCCAAGGCAAAACAGATTGTCAATCCTGCAGCTGCTAAGGTTACACAAAGAG  2098
Sbjct  7449852  CTTCACCCTCCAAGGCAAAACAGATTGTCAATCCTGCAGCTGCTAAGGTTACACAAAGAG  7449911

Query  2099     TTGATCCTCCACCTGCCAAGGCATCTCAGAGAATTGATCCTCTGTTGCCATCCAAGGTTC  2158
Sbjct  7449912  TTGATCCTCCACCTGCCAAGGCATCTCAGAGAATTGATCCTCTGTTGCCATCCAAGGTTC  7449971

Query  2159     ATATAGATGCTACTCGATCTTTTACGAAGGTCTCCCAGACAGAGATCAAGCCGGAAGTAC  2218
Sbjct  7449972  ATATAGATGCTACTCGATCTTTTACGAAGGTCTCCCAGACAGAGATCAAGCCGGAAGTAC  7450031

Query  2219     AGCCCCCAATTCTGAAGGTGCCTGTGGCTATGCCTACCATCAATCGTCAGCAGATTGACA  2278
Sbjct  7450032  AGCCCCCAATTCTGAAGGTGCCTGTGGCTATGCCTACCATCAATCGTCAGCAGATTGACA  7450091

Query  2279     CCTCGCAGCCCAAAGAAGAGCCTTGCTCCTCTGGCAGGAATGCTGAAGCTGCTTCAGTAT  2338
Sbjct  7450092  CCTCGCAGCCCAAAGAAGAGCCTTGCTCCTCTGGCAGGAATGCTGAAGCTGCTTCAGTAT  7450151

Query  2339     CAGTAGAGAAGCAGTCCAAGTCAGATCGCAAAAAAGAGCCGCAAGGCTGAGAAGAAAGAGA  2398
Sbjct  7450152  CAGTAGAGAAGCAGTCCAAGTCAGATCGCAAAAAGAGCCGCAAGGCTGAGAAGAAAGAGA  7450211

Query  2399     AGAAGTTCAAAGATTTATTTGTTACCTGGGATCCTCCGTCTATGGAAATGGATGATATGG  2458
Sbjct  7450212  AGAAGTTCAAAGATTTATTTGTTACCTGGGATCCTCCGTCTATGGAAATGGATGATATGG  7450271

Query  2459     ATCTCGGGGACCAGGATTGGCTGCTTGATAGTACGAGGAAACCTGATGCTGGCATTGGCA  2518
Sbjct  7450272  ATCTCGGGGACCAGGATTGGCTGCTTGATAGTACGAGGAAACCTGATGCTGGCATTGGCA  7450331

Query  2519     ACTGCAGAGAAATTGTTGATCCACTTACTTCTCAATCAGCAGAGCAGTTCTCATTGCAGC  2578
Sbjct  7450332  ACTGCAGAGAAATTGTTGATCCACTTACTTCTCAATCAGCAGAGCAGTTCTCATTGCAGC  7450391

Query  2579     CTAGGGCGATTCATTTACCAGACCTTCATGTCTATCAGTTGCCATATGTGGTTCCATTCT  2638
Sbjct  7450392  CTAGGGCGATTCATTTACCAGACCTTCATGTCTATCAGTTGCCATATGTGGTTCCATTCT  7450451

Query  2639     AGGTTTGTGTAGTGAGATGGAGTAGGTGAGAAGTAGAGAGATGTTGGGAGAGAGCTGTGT  2698
Sbjct  7450452  AGGTTTGTGTAGTGAGATGGAGTAGGTGAGAAGTAGAGAGATGTTGGGAGAGAGCTGTGT  7450511

Query  2699     GGGTCTGGGAGATTATGGTTCACTGCCACAGTTTTCCAGCATTGTTCCAGAGCGTTTCTT  2758
Sbjct  7450512  GGGTCTGGGAGATTATGGTTCCCTGCCACAGTTTTCCAGCTTTGTTCCAGAGCGTTTCTT  7450571

Query  2759     GTTTCAGGTTGCTGAGCCTGTCAAACTTCATGTAATGTTGTATTTAAGCTTCGTTTATTG  2816
Sbjct  7450572  GTTTCAGGTTGCTGAGCCTGTCAAACTTCATGTAATGTTGTATTTTTGCTTCGTTTATTG  7450631

Query  2817     CAAGCAATTGAACTTC  2832
Sbjct  7450632  CA-GCAATTGAACTTC  7450646
```

Fig. 4.36 BLAST result of sequence for *OsPE* candidate gene in Basmati 370 and Nipponbare

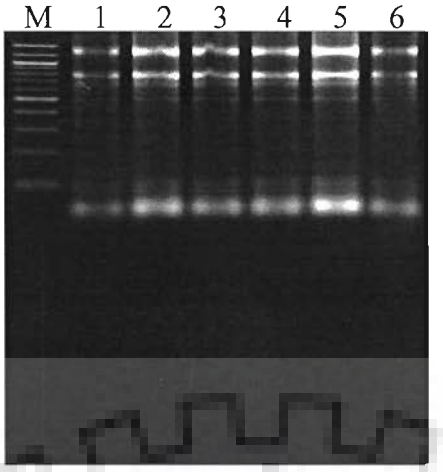

Fig. 4.37 Total RNA Lane 1: Basmati 370 Shoots; Lane 2: Basmati roots; Lane 3: Other insertional mutant, roots; Lane 4: Other insertional mutant, shoots; Lane 5: *OsPE* Roots; Lane 6: *OsPE* Shoots; Lane M: molecular weight marker 1Kb

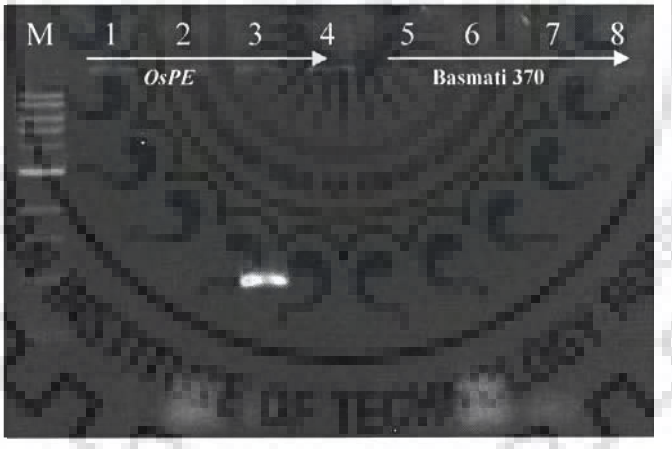

Fig. 4.38 RT-PCR primer check on genomic DNA. Lane 1: *OsPE* with RT-Specific primer; Lane 2: *OsPE*, actinI primer; Lane 3: *OsPE*, *hpt* primer; Lane 4: -ve control; Lane 5: Basmati 370 with RT- Specific primer; Lane 6: Basmati 370, actinI primer; Lane 7: Basmati 370, *hpt* primer; Lane 8: -ve control; Lane M: Molecular weight marker 100bp

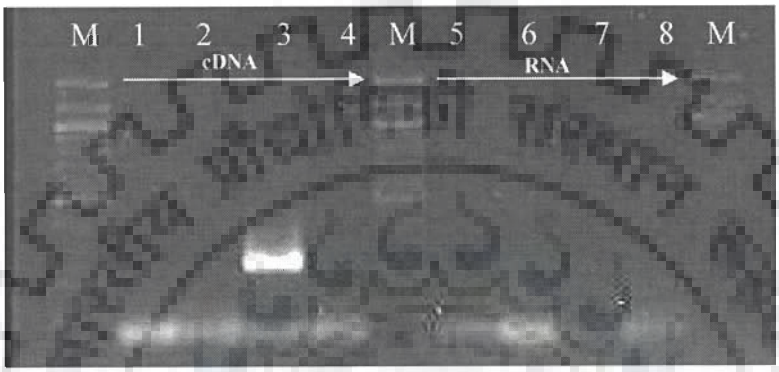

Fig. 4.39 RT-PCR expression in Basmati 370 shoots
Lane 1: *OsPE*(cDNA) with RT-Specific primer; Lane 2: *OsPE*(cDNA), actinI primer; Lane 3: Basmati (cDNA) with RT-Specific primer; Lane 4: Basmati 370(cDNA) with actinI primer. Lanes 5-8 same set of primers repeated with total RNA as template. Lane M: Molecular weight marker 100bp

Table 4.6 Rice ESTs corresponding to candidate gene *OsPE*

Pool name	Transcripts per million	Gene EST/Total EST in pool
Callus	42	7/164802
Flower	168	23/136502
Leaf	69	12/171750
Panicle	173	23/132789
Root	87	6/68202
Seed	339	11/32357
Stem	102	13/126881
Vegetative meristem	0	0/4594

4.10 Blast of *OsPE* sequence with Rice FSTs

Insertional mutant databases containing FSTs for rice have been compiled through the efforts of 11 institutes in 7 countries. So far 1,72,500 FSTs have been generated via insertional mutagenesis in 27,551 (48%) of 57,142 rice loci (given by the International Rice Genome Sequencing Project and the current TIGR rice genome pseudomolecules release) (Jung *et al.*, 2008). All the rice flanking Sequence Tags (FSTs) are publically available at Rice Functional Genomics Express Database (Rice GE) developed by the SALK institute as well as at NCBI. The BLAST alignment of the sequence of *OsPE* gene against FST database showed two hits. The *OsPE* gene has insertion in the promoter region whereas FST searched insertions were found to be in the coding region (Fig. 4.40). No phenotypic description or associated annotations related to these insertions was reported. It can provide a link between the candidate genes for polyembryony in *OsPE* mutant by comparing the phenotype in other rice cultivars. This represents a straight forward approach to assign function to rice gene sequences. In addition no hits were obtained from *Tos*17, UCD, POSTECH and *Oryza* Tag Lines whereas RMD, TRIM and SHIP databases showed around 21 hits in the nearby region but not within the *OsPE* gene.

4.11 TAIL-PCR for other insertional mutants

In addition to polyembryonic mutants, identification of flanking sequences for two more independent T-DNA insertional mutants (B-2-2, B-8-7) was done using TAIL-PCR. TAIL-PCR was performed using T-DNA specific primers (T-RB1, T-RB2, T-RB3) and arbitrary degenerate primer (AD1) (Fig. 4.41). The PCR amplified products were resolved with low melting 1% agarose (w/v), eluted and purified. The gel eluted PCR

products were sequenced using T-RB2 primers for both the insertional mutants. Sequences obtained for B-8-7 and B-2-2 are given in Fig. 4.42 and Fig. 4.43.

The insertion in B-8-7 mutant was found to be present in the junk region at position 27661397 to 276616602 in chromosome 3 when searched for identity using BLAST against *Oryza sativa* cv. Nipponbare (Fig. 4.44).

The sequencing result of B-2-2 with T-RB2 showed 902 nucleotide long sequence. The chromosomal location for B-2-2 insertion was determined using NCBI, BLAST search with available data for *japonica* rice cultivar Nipponbare. The BLAST result showed 87% identity with chromosome 11 of rice genome (Fig. 4.45). The sequence corresponds to the contig region NC_008404.1. The insertion is present in the coding region of gene NC_008494.1. at position 26784525 to 26783660. The gene codes for protein Pectin methylesterase (PME). The length of the gene was 2,185bp with five exons and five introns. The length of mRNA was 1749bp which codes for 423 amino acids. The gene Pectin methylestrase contains putative pectinestrase domain. EST counts retrieved from the database suggests that the identified knockout gene B-2-2 expresses only in panicles and flowers.

Genomic fragments encoding PME and its isoforms have been isolated from several higher plants such as *Brassica, Petunia, Arabidopsis* and *Oryza sativa.* Around 143 protein matching counts were present in *Oryza sativa.* Several Pectin methylestrase isoforms with different pH, isoelectric points and molecular mass are present in the same plant. Not much variation has been observed in the mutant with respect to wild type except 10-15% smaller grains.

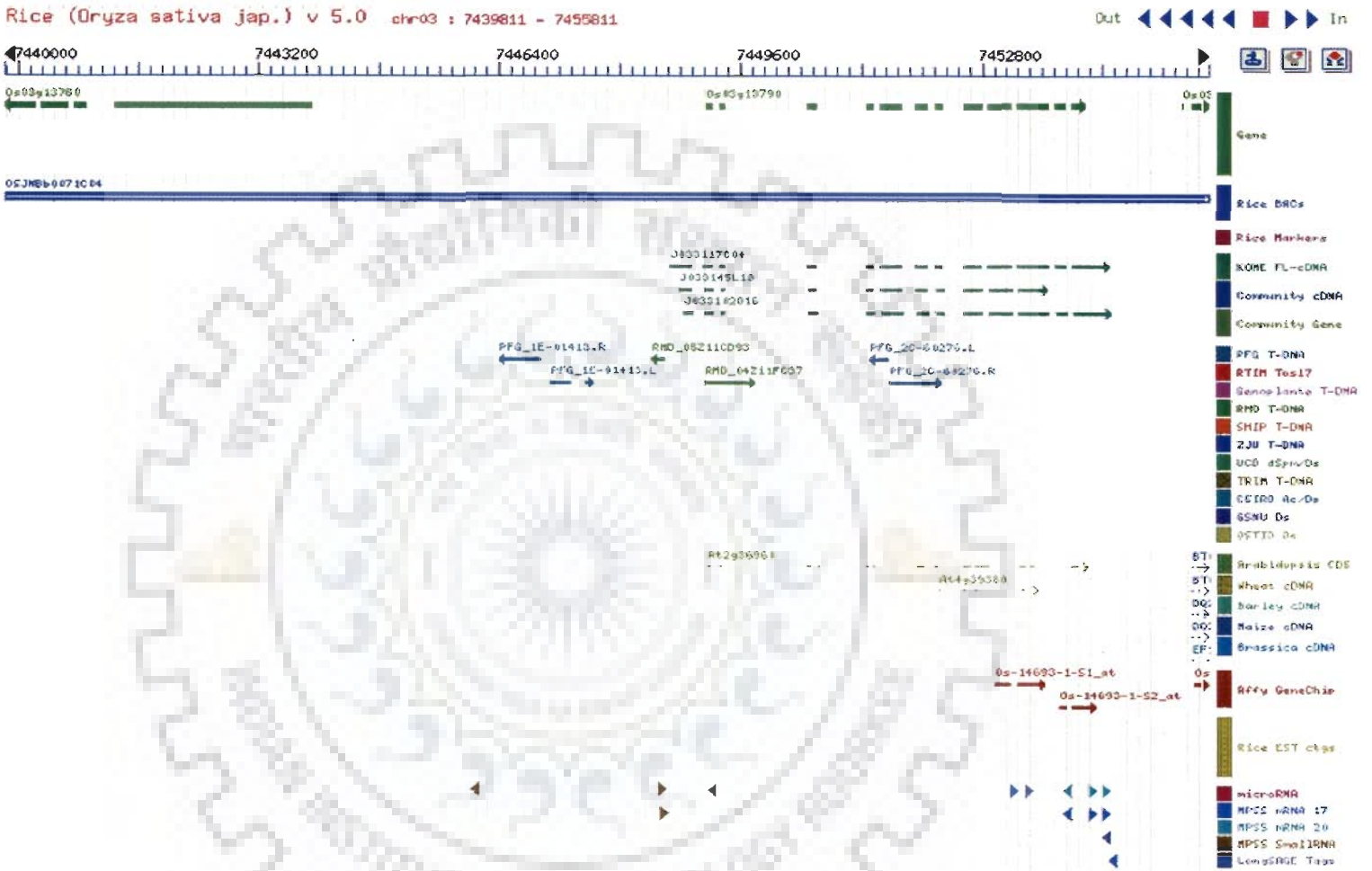

Fig. 4.40 FSTs in rice chromosome 3 in the *OsPE* region (7439811-7455811)

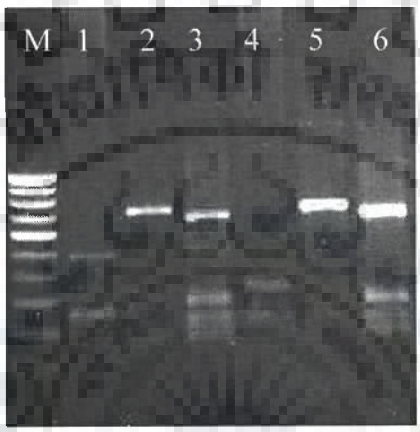

Fig. 4.41 PCR amplification of Primary, Secondary and Tertiary TAIL-PCR for identifying T-DNA flanking region in B-2-2 and B-8-7
B-2-2 Lane1: Primary PCR (T-RB1/AD1); Lane 2: Secondary PCR (T-RB2/AD1); Lane 3: Tertiary PCR (T-RB3/AD1); B-8-7 Lane 4: Primary PCR (T-RB1/AD1); Lane 5: Secondary PCR (T-RB2/AD1); Lane 6: Tertiary PCR (T-RB3/AD1)
Lane M: Molecular Weight Marker 1Kb

> B-8-7_RB3-1
GGGATCCAGGCGAAAGACTGTTCTTCTATCACGAGACGAGTACTCCGCGCTGTGAGTAGGTTGATGGGACG
TAGGATCAGAAAACGCGTTCTTGCCTGTCCTGGTTTCACTAACCAAATCGAGTATGAAAGAACTTAAACAC
AGGCACAATTTGATCGTATAATCTCTCAAAATAATTTGGTGGACAAAATTTCTAGACGACCAAACATCGTG
TCGGCAACAGGGTGTTCATGTCACTCGACAACCGACAAAGCTAGCCCCGAAAAACCCCATCTTTGGCCAGC
ATAACCCTTGGAAAGAAGGCGGAATAAAACCGAAACTTCCGGTAACTATACTAAAGGCTTAAACGCGGTAT
TAAACTCACCCGTATGTGAACCACGATCAGGGGGTCCATATACATAACGGACTACCCCAGCACATGTAACC
GTGTGTAAGACCAATTTGTAGACACCGTGCAATATGGAGCTTTTGCCCAAAACACCAAGTGATCATCAGGT
GCCGTCCCAAGTGCAGCGAGATGAGCTTCAGCGCTCATACTCATTCACTGTTTAAGTTGAATACTAACACA
AAAGCGGAAACGCTACATGTTTAAACAACAATGTCTATCCTCTCTAATCTCCAAATCTTCTATTCACCAGC
ACTAATCCCAGGTGGATTACGCCACCTAATTTGCAAACAAAAGACCATAGTCCACTCCCCAGAGAAAGACA
TCAACGCGATACAGAGTCGAACGCCTCGGCTAAATAGCGGAGATTTCCCATCCTGGGTGCAAATCGTTAAA
CCAAAAATCTGTTCTTATCAAACTTCTACAGGGTACTCTAACAAGAAAGGAGAATCCGCTGCATGTGCTTT
TACCTCTCATCCACCACAACACCAACAGTGGCCGTGACATATAAAACGATCCGAGGGGCCGCG

Fig. 4.42 T-DNA flanking sequence using T-RB2 (T-DNA based primer in TAIL-PCR) of B-8-7 mutant

> B-2-2_RB3-1
AAATCGCTCTACTACTTTTGTTTTTACATTGATTAATTAATTAACACTCTCTCTTTTTCATATTATGAATC
GCTTTGATTTCTTTTGCATTTTTTTCAAGTTTAACCAAATTTTATTGGAAAACATATCAACACAATTTTTT
AAGAAAAGAACAAACATATTATAAAAATATATTCAACTTTACATTTAGTTAAACTATTTTTTAAAAATGCT
GCTAAATTTTTTGAAAAAAAATTGATCAAATATAAATAAGTTTGACAAAAAAAAAATCAAAACAAAGTTACA
TTTTATATGAAATGGAAGGGGTAATCAATTAATTATGCTAGTACAAATTAATATATTCTAATTGTTGGTGT
AACATCAAATAGCGGGACATCAACGTCATCTATCTCTGTTTCTACGCAGAGCAACTCAACCGCGGAAGCCC
CGACGGCGAGTAGTACCTCAACTGCATCGGCTAATAAAAAGGAAAGCGATTCAACTGCAGAAGCGCCGACG
GCCACAAGAACCTCACCTGCGTCGGCGACACCGAGTAATCCACTGCGGCAGCATAGGCGACTTATTACTCG
AGTTCAGACTCACGGTCAAAAGAGAAGTGTCAATACAAGTCCAGCTGCAACAACGCGCCCGACGCTGATGA
ATCCGTTTCATTGGGGGCACACAAGTTCTGTTTTCGAAAGCACTGTTAATCGTTCATTTTAGCGGGTAAAT
GTGGAGCGGAATAAATTGAAATCGGTGCTTACGTAGATGAGTGCCTCGCTGTAAAGCTGATTTTTACTTAC
CAATTTCGTTCGGCTTTTGATGCGGCACCGACCACTCAAAAACTATGAGTTAGTTCTAACATGGTTTCCTT
GATGACAACGGACTTTTGAAGACTAGTGCTACATCGTTATGGGATAAATTACTTCCAGAGAAAGATAACGA
TCACAAATTTGGTTTCAAACGGGAATAACACTTGTGAAACAGGGTCTACCTCTGGTACAGTGCATAAATAT
GTGGTACAAAGGTATGTTGCTTTTCTAACAGGAGTTGTTCTCCTGAGATGAGGACCACATACACCTCGGAT
TCAATGCAAAAACAGGCCCCGAAAAGAAACGCGTTTTTCATTATGCATTACACGAAGCGTGAATAAGAATT
GTCTATTGGAAGAGCCCACGGATTATTGGACACATCCAGATGCGGCTGACGTCAGCGTAGAATGAA

Fig. 4.43 T-DNA flanking sequence using T-RB2 (T-DNA based primer in TAIL-PCR) of B-2-2 mutant

```
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 3
Identities = 189/209 (90%), Gaps = 8/209 (3%)
 Strand=Plus/Plus

Query  41        TACTCC-GCGCTG-TGAGTAGGTTGATG--GGACGTAGGATCAGAAAAC-GCGTTCTTGC  95
                 ||||||  ||| | |||| | || ||||  |||||| ||||| |||||| | | |  |||
Sbjct  27661397  TACTCCAACGC-GTTGAGAAAGTGGATGTTGGACGTCGGATC-GAAAACGGAGAT-GTGC  27661453

Query  96        CTGTCCTGGTTTCACTAACCAAATCGAGTATGAAAGAACTTAAACACAGGCACAATTTGA  155
                  ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27661454  ATGTGCTGGTTTCACTAACCAAATCGAGTATGAAAGAACTTAAACACAGGCACAATTTGA  27661513

Query  156       TCGTATAATCTCTCAAAATAATTTGGTGGACAAAATTTCTAGACGACCAAACATCGTGTC  215
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27661514  TCGTATAATCTCTCAAAATAATTTGGTGGACAAAATTTCTAGACGACCAAACATCGTGTC  27661573

Query  216       GGCAACAGGGTGTTCATGTCACTCGACAA  244
                 |||||||||||||||| ||| |||||||
Sbjct  27661574  GGCAACAGGGTGTTCATTTCAGTCGACAA  27661602
```

Fig. 4.44 BLAST result of T-DNA flanking sequence of B-8-7 mutant using T-RB2 (TAIL-PCR) showing 90% identity with rice cultivar Nipponbare at Chromosome 3

```
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 11
 Score =  974 bits (527),  Expect = 0.0 Identities = 788/902 (87%), Gaps = 65/902 (7%)
 Query  10        TACTACTTTTGTTTTTACATTGATTAATTAATTAACactctctctttttcatattatgaa  69
                  ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  26784525  TACTACTTTTGTTTTTACATCGATTAATTAATTAACACTCTCTCTTTTTCATATTATGAA  26784466

Query  70        tcgctttgatttctttttgcatttttttcaagtttaaccaaaattttattggaaaacatatc  129
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26784465  TCGCTTTGATTTCTTTTGCATTTTTTTCAAGTTTAACCAAATTTTATTGGAAAACATATC  26784406

Query  130       aacacaatttttttaagaaaaagaacaaacatattataaaaatatattcaactttacattta  189
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26784405  AACACAATTTTTTAAGAAAAGAACAAACATATTATAAAAATATATTCAACTTTACATTTA  26784346

Query  190       gttaaactatttttttaaaaatgctgctaaatttttttgaaaaaaaaattgatcaaatataaa  249
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26784345  GTTAAACTATTTTTTAAAAATGCTGCTAAATTTTTTGAAAAAAAATTGATCAAATATAAA  26784286

Query  250       taagtttgacaaaaaaaaatcaaaacaaagttacatttatatgaaatggaaggggtaAT  309
                  |||||||||| |||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  26784285  TAAGTTTGAC-AAAAAAAATCAAAACAAAGTTACATTTTATATGAAATGGAAGGTGTAAT  26784227

Query  310       CAATTAATTATGCTAGTACAAATTAATATATTCTAATTGTTGGTGTAACATCAAATAGCG  369
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26784226  CAATTAATTATGCTAGTACAAATTAATATATTCTAATTGTTGGTGTAACATCAAATAGCG  26784167

Query  370       GGACATCAACGTCATCTATCTCTGTTTCTACG-CAGAGCAACTCAACCGCGGAAGCCCCG  428
                  ||||||||||||||||| ||| | ||| ||| | ||||||||||||||||||||||| |||
Sbjct  26784166  GGACATCAACGTCATCTAGCTCAGCTTCAACGTC-GAGCAACTCAACCGCGGAAGCACCG  26784108

Query  429       ACGGCGAGTAGTACCTCAACTGCATCGGCTAATAAAAAGGAAAGCGATTCAACTGCAGAA  488
                  |||||||||||||| ||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  26784107  ACGGCGAGTAGTAACTCAACTGCATCGGCTAATAAAAAGGAAAGCAATTCAACTGCAGAA  26784048

Query  489       GCGCCGACGGCCACAAGA-ACCTCACCTGCGTCGGCGAC-ACCGAGTAAT-CCACTGCGG  545
                  ||||||||||| || ||  || ||| ||||||| ||||| || ||||||| ||||||||
Sbjct  26784047  GCGCCGACGGCGACTAGCTAC-TCAACTGCGTCAGCGACGAC-GAGTAATTCCACTGCGT  26783990

Query  546       CAGC-ATAGGCGACTTATTACTCGAGTTCAGACTC-AC-GGTCAAAAGAG-AAGT-GTCA  600
                  |||| | |||||||| ||||||||||||||| | || || || |||||| | |||| || |
Sbjct  26783989  CAGCGA-AGGCGACTAATTACTCGAGTTCA-AGTCCACCGG-CAAAAG-GCAAGTAGTTA  26783934

Query  601       ATACAAGTCCAGCTGCAACAACG-CGCCCGACGCTGATGAATC-CGTTTCATTGGGGGCA  658
                  ||||||| |||| || || ||| |||| ||||| |||||||| |||| |||||||| ||
Sbjct  26783933  ATACAAGTGCAGCGGCGACGACGGCGCC-GACGC-GATGAATCGCGTT-CATTGGGGCCA  26783877

Query  659       CAC-AAGTTCTGTTTTCGAAAGCACTGTT-AATCGTTCATT-TTAGCGGGTAAATGTGGA  715
                  || || |||||||| ||| || | ||| ||| |||||| || | |  | ||| |||
Sbjct  26783876  -ACCAA-TTCTGTTTTGGAA-GCCC-GTTTAATTGTTCATTCTTTGTG--TCAATTTGG-  26783824

Query  716       GC-GGAATAAATTGAAATCGGT-G-CTTACGTAGATGAGTGC-C-T-CGCTGTAA-AGCT  768
                  || ||| ||| | || |||  | || || |||||  | | | | || ||||| || |
Sbjct  26783823  GCCGGATTAA-TC-AA-TCGGCCGGCT-ACTTAGATGTTT-CACGTACG-TGTAACAGTT  26783770

Query  769       G-ATTTTTACTTACCAATTTCGTTCGG--CTTTTGATGC-G-GCACCGACCA-CT-CAAA  821
                  | ||||| ||| ||| ||||||| ||  ||||| || | | | ||| | | | || ||||
Sbjct  26783769  GTATTTT--CTTCCCA-TTTCGTTTGGGTCTTTT-AT-CAGTGTAC-G-C-ATCTTCAAA  26783718

Query  822       AACTATGAGTTAGTTCTA-ACATGGTTTCCTTGATGACAACGGACTTTTGAAGACTAGTG  880
                  | |||| ||| | || || ||||| |||||||  ||| || | ||| |||||||| |||
Sbjct  26783717  A-CTAT-AGTGAATTTTACACATGTTTTCCTTT-TGAAAATG-ACTCTTGAAGACTTGTG  26783662

Query  881       CT  882
                  ||
Sbjct  26783661  CT  26783660
```

Fig. 4.45 BLAST result of T-DNA flanking sequence of B-2-2 mutant using T-RB2(TAIL-PCR) showing 87% identity with rice cultivar Nipponbare at Chromosome 11

We searched the database using BLAST tool to identify any mutant (in other species) corresponding to *OsPE*. Homologies with genes of known functions, specific tissue expression patterns, and conserved motifs would have supported assigning function to *OsPE*, but in the absence of conserved domains in the protein coded by *OsPE* and phylogenetic relationship within *Oryza sativa* indicate this to be a novel gene. Southern hybridization indicated the presence of single insertion. BSA was done for linkage analysis of *OsPE* locus. Genome walking and TAIL-PCR confirmed the position of insertion. RT-PCR confirmed the gene expression in the shoot region. A detailed molecular characterization and mapping of the *OsPE* insertional mutant is available through this study and RNAi and complementation analysis will greatly help in validation of the assigned function to this novel gene.

CHAPTER 5

DISCUSSION

5. Discussion

Polyembryony (*OsPE*) mutant in the present study was generated from superfine quality *indica* rice cultivar Basmati 370, through transformation with *Agrobacterium tumefaciens* containing Hm^R*Ds* construct. The independent insertional mutant showed resistance to hygromycin (80 ppm) and the twin/triplet seedlings occurred at the frequency of 15-20%. Polyembryonic seeds were also obtained in some F_2 plants generated from a cross *OsPE*/PR106. The segregation pattern of the F_2 plants was also analyzed on the basis of hygromycin resistance and PCR amplification of *hpt* gene fragment. The goodness of fit to the expected 3:1 ratio was observed among *hpt* +ve and *hpt* –ve plants suggesting single insertion. Southern hybridization in *OsPE* using *hpt* probe also indicated single T-DNA insertion. But the F_3 seed data for polyembryony unlike the *hpt* data did not segregate in the expected 3:1 ratio suggesting variable penetrance and expressivity for the polyembryonic insertional mutant.

The penetrance is a qualitative and expressivity is a quantitative extent of expression of a gene. Penetrance relates to whether a phenotype is expressed for a particular genotype whereas expressivity relates to the degree to which a phenotype is expressed after penetrance. Expressivity is thus dependent on penetrance (Kral, 1999). In the homozygous *OsPE* insertional mutant, confirmed on hygromycin resistance and *hpt* amplification, there was a maximum penetrance of 21% while the expressivity in these seeds with polyembryony ranged from twin, triplet to a rare quadruplet embryos. In 138 F_3 progenies, 96 gave *hpt* amplification whereas only 14 lines among them showed polyembryony. The *OsPE* gene in F_2 plants therefore had a penetrance of 14.5%. There can be many factors (genetic/environmental) that affected the expression of

polyembryony in *OsPE* and thus the deviation from the ratio expected by Mendelian inheritance. Recently incomplete penetrance was also reported in *Arabidopsis* mutant for *DYAD/SWITCH1* (*SWI1*) in which diploid plants produced triploid progeny (Maruthachalam *et al.*, 2008).

The origin of multiple embryos in *OsPE* whether zygotic or apomictic is not clear. Multiple embryos with independent root and shoot axis were clearly attached to a single scutellum suggesting that they might have originated from a single zygotic embryo. The *OsPE* was successfully crossed with PR106 and the F_2 population segregated for *hpt*, polyembryony, plant height and other morphological traits differentiating the parents suggests that the *OsPE* could be at the most a facultative apomict. The twin/triplet plants from the polyembryonic seeds within the fourteen F_3 progenies were morphologically similar further suggests that the multiple embryos might have arisen through proliferation from a single zygotic embryo. In seeds with polyembryony, embryos differed in size and early seedling vigour but later on grew to fertile plants with similar morphology.

The meiotic analysis of the pollen mother cells of *OsPE* indicated diploid (2n= 24) chromosomal number and normal meiotic cell divisions with 12 IIs at metaphase I and tetrads. In case of sporophytic apomicts (diplosporous and aposporous) embryo sac, meiotic tetrad is generally not observed. They are mosty polyploid, yet several members of the same or closely related species are commonly diploids. Apomixis has the potential to revolutionize agriculture if its genetic basis is known. However, the genetic control of natural apomixis has remained obscure until quite recently, owing to all the complications of Mendelian genetics, such as epistatic gene interactions, compoments that are expressed sporophytically and gametophytically, expression modifiers,

polyploidy, aneuploidy, segregation distortion, suppressed recombination etc, that seem to have accumulated during the evolution of apomixis (Ozias-Akins and van Dijk, 2007). Attempts to introgress apomictic trait from the wild relatives into maize and pearl millet (Savidan, 2000, 2001) have not been successful.

BSA using the rice SSR markers in *OsPE*/PR106 mapping population clearly indicated that the *OsPE* gene is located on chromosome 3 and showed linkage with two flanking markers RM14667 (2.17cM) and RM14645 (5.79cM). The SSR markers are co-dominant, multi- allelic and available at a high density in the rice genome (one SSR per 157 Kb; McCouch *et al.,* 2002). These markers can be used for identifying the target region associated with the trait of interest. BSA approach has been used for molecular tagging genes for several characters in rice namely blast resistance (Liu *et al.,* 2005; Sharma *et al.,* 2005; Chen *et al.,* 2007), bacterial blight resistance (Rao 2003; Yang *et al;* 2003; Qi *et al.,* 2006) and for gall midge resistance (Biradar *et al.,* 2004). In the present study the mapping of *OsPE* on rice chromosome 3 using BSA has been further confirmed with other reverse genetic approaches including genome walking and TAIL-PCR involving amplification of T-DNA flanking region using T-DNA and adaptor based primers.

To carry out genome walking and TAIL-PCR, nested primers were designed for T-DNA right and left borders. No results were obtained using T-LB primers. This may be because of the modification of T-DNA left border. According to McClean (1998) the only structural requirement for T-DNA transfer is the TR direct repeat border. The T-DNA transfer is in the right to left direction. The genes required for oncogenic functions are linked with right border. Thus, when right border is used for T-DNA transfer, the

genes required for the oncogenic functions are transferred, but when the left border is used in lieu of the right border, transfer did not include these genes and transformation did not occur. Many a times left border was found to have deletions during transformation. In *Aschyta rabiei*, the T-DNA left border region contained deletion of upto 72bp upon integration (White and Chen, 2006).

Using T-DNA right border nested primers along with adaptor primer (Genome Walking) and arbitrary degenerate primer (TAIL-PCR), PCR amplification was obtained. The PCR product was purified and sequenced. The location of *OsPE* gene on chromosome was determined using *japonica* rice cultivar Nipponbare sequence. Sequencing with T-RB2/AP2 (Genome Walking) and T-RB3 (TAIL-PCR) showed significant alignment with the *Oryza sativa* cv. Nipponbare with only single hit at chromosome 3. The T-DNA insertion is present in the promoter region of the candidate gene Os03g0241300. The length of the candidate gene is 2.8Kb with 2 UTRs (1473bp and 193bp), 2 Exons (154bp and 935bp), and one Intron (81bp).

To confirm the insertion on chromosome 3, primers were designed for rice genome flanking the T-DNA borders. PCR was done in combination with the T-DNA (T-RB) based primers and genome-specific primers. Amplification of expected size was obtained in polyembryonic mutant but not in Basmati 370 confirming the insertion at the specific region of chromosome 3. PCR with genome-specific primers gave amplification in Basmati 370 and not in *OsPE* because of the insertion of approximately 10Kb fragment of the T-DNA cassette which cannot be amplified by the Taq polymerase used for the this PCR.

The protein searched corresponding to the *OsPE* candidate gene sequence has been reported as a hypothetical protein (Os03g0241300) in *Oryza sativa*. *OsPE* gene lacked functional homologs in other species. Although using KEGG search around 401 sequences orthologous to *OsPE* were found through out the living kingdom but majority had low identity value (> 40%). It most closely matched with the *Vitis* and *Arabidopsis* hypothetical and unknown proteins, respectively. No *OsPE* paralog was found in rice. No conserved domains were found in the protein coded by *OsPE*. Availability of homologs genes of known function, specific tissue expression patterns and conserved domains would have confirmed assigning function to this novel gene, *OsPE*.

Using the Promoter2 (Knudsen, 1999) and the Dragon Promoter Finder (DPF) (Bajic *et al.*, 2002) no TATA Box was found in the promoter region of *OsPE* where insertion was present. It was found that the gene starts at 425bp downstream the promoter region where transcription start site (TSS) was present. Many highly expressed genes contain a strong TATA Box in their core promoter. However, in some large groups of genes, like the housekeeping and photosynthesis genes, the TATA Box is often absent, and the corresponding promoters are referred to as TATA-less promoters (Shahmuradov, 2005). In these promoters, the exact position of the transcription start point may be controlled by the nucleotide sequence of the transcription initiation region (Inr) or the recently found downstream promoter element (DPE), which is typically observed 30bp downstream of the TSS (Smale, 1997; Burke and Kadonga, 1997).

RT-PCR using reverse transcribed mRNA and primers pair designed on intron/ exon junction, the expression of *OsPE* gene was found in Basmati 370 shoots and not in roots. The expression profile of candidate gene suggested by EST counts showed

maximum ESTs in seed (11/32357) followed by flower (23/136502) and panicle (23/132789). The polyembryony is associated with seed where maximum ESTs were found for the *OsPE* gene.

Full length *OsPE* gene was cloned in Basmati 370 on the basis of the sequences obtained by genome walking and TAIL-PCR. The nucleotide sequence of the candidate gene *OsPE* in *japonica* rice was used to design overlapping PCR primers and the *OsPE* gene was amplified in Basmati 370 genomic DNA as the template. The sequence obtained using corresponding forward and reverse primers were aligned to determine the sequence errors and full length *OsPE* gene in Basmati 370 was obtained. BLAST search was done for homologous sequence in Nipponbare *japonica* rice. Ten SNPs (Single Nucleotide Polymorphism) and one gap were found in the *OsPE* candidate gene sequence as compared to Nipponbare sequence thus suggesting a total of 1% dissimilarity between the two sequences. The predicted protein was searched and variation was found in only two amino acids. The reported SNP varies form 0.53% to 0.78% between *japonica* and *indica* rice subspecies (IRGSP 2005). With the minor differences in coding region the *OsPE* gene is expected to translate a similar protein in Basmati 370.

The *OsPE* gene BLAST alignment against rice FST database (Rice Functional Genomics Dataabase RiceGE) showed two hits. The *OsPE* gene has insertion in the promoter region whereas FST searched insertions were found to be in the coding region. No phenotypic description or associated annotations related to these insertions have been reported.

In addition to polyembryonic mutants, identification for flanking sequences for two more independent T-DNA insertional mutants (B-2-2, B-8-7) was done using TAIL-

PCR. The insertion in B-8-7 mutant has been found to be present in the junk region where as in B-2-2 the insertion is present in the coding region of chromosome 11. The insertion in chromosome 11 of B-2-2 is present in the gene that codes for Pectinmethylestrase (PME). The genomic regions encoding PME and its isoforms have been isolated from several higher plants such as *Brassica, Petunia, Arabidopsis* and *Oryza sativa*. All the three insertional mutants studied in the present investigation were found to have T-DNA insertions in genes as reported earlier in *Agrobacterium* mediated transformation (Nobuta *et al.*, 2007; Hsing *et al.*, 2007 and Jeon *et al.*, 2000).

The present knowledge and material can be used for generating additional insertions in the same allele or nearby sites using *Ac/Ds* system. The T-DNA cassette in the insertional mutants has Hm^{R}*Ds*. If rice with *Ac* element is crossed with these insertional mutants, *Ds* can transpose to nearby sites to generate additional insertional mutants for further insight in the candidate genes. This strategy has been used successfully to establish large collections of insertional mutants in *Arabidopsis* (Alonso and Ecker, 2006; Wing, *et al.*, 2005; Lipshutz, *et al.*, 1999) and rice (Kolesnik *et al.*, 2004).

The validation of the candidate gene can also be done using RNAi technology. RNA silencing has emerged as a major functional genomic tool to establish correlation between gene and its function. RNA silencing has been successfully used in plants to suppress specific gene functions (Singh *et al.*, 2006; Anand *et al.*, 2003; Miki *et al.*, 2005). Most of the silencing experiments have been carried out for the model plant *Arabidopsis*. Thus, the *OsPE* and other candidate genes can be validated using RNAi approach.

REFERENCES

References

Akiyama, Y., Conner, J.A., Goel, S., Morishige, D.T., Mullet, J.E. 2004. High-resolutionphysical mapping in *Pennisetum squamulatum* reveals extensive chromosomal heteromorphism of the genomic region associated with apomixis. *Plant Physiol.* **134:** 1733–1741.

Akiyama, Y., Goel, S., Chen, Z., Hanna, W., Ozias-Akins, P. 2006. *Pennisetum squamulatum*: Is the predominant cytotype hexaploid or octaploid? *J. Hered.* **97:** 521–524.

Albertini, E., Marconi, G., Reale, L., Barcaccia, G., Porceddu, A. 2005. SERK and APOSTART. Candidate genes for apomixis in *Poa pratensis*. *Plant Physiol.* **138:** 2185–2199.

Albertini, E., Porceddu, A., Ferranti, F., Reale, L., Barcaccia, G. 2001. Apospory and apomictic mode of reproduction in maize-*Tripsacum* hybrids using maize RFLP markers. *Theor. Appl. Genet.* **90:**1198–1203.

Al-Forkan, M., Power, B., Anthony, P., Lowe, K.C., Davey, M.R. 2004. *Agrobacterium*-mediated transformation of Bangladeshi Indica rices. *Cell. Mol. Biol. Lett.* **9:** 287-300.

Alonso, J.M., and Ecker, J.R. 2006. Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nature Rev. Genet.* **7:** 524–536.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. 1990 Basic local alignment search tool. *J Mol Biol.* **215:** 403–410.

An, G., Lee, S., Kim, S.H., Kim, S.R. 2005. Molecular genetics using T-DNA in rice. *Plant Cell Physiol.* **46:**14–22.

An, S., Park, S., Jeong, D.H., Lee, D.Y, Kang, H.G., Yu, J.H., Hur, J., Kim, S.R., Kim, Y.H . 2003. Generation and analysis of end sequence database for T-DNA tagging lines in rice. *Plant Physiol.* **133**: 2040–2047.

Anand, A., Trick, H.N., Gill, B.S., Muthukrishnan, S. 2003. Stable transgene expression and random gene silencing in wheat. *Plant Biotechnol J.* **1:** 241-251.

Anonymous 1996. Polyembryony in plants: a weapon in the war over offspring numbers. *Tree.* vol 11. no1.

Anonymous. The Rice Annotation Project Database (RAP-DB): 2008 update *Nucleic Acids Research* .2007. 1–6. Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License.

Araujo, A., Nobrega, J., Pozzobon, M., Carneiro, V. 2005. Evidence of sexuality in induced tetraploids of *Brachiaria brizantha* (Poaceae). *Euphytica.* **144:** 39–50.

Azpiroz-Leehan, R., Feldmann, K.A. 1997. T-DNA insertion mutagenesis in *Arabidopsis:* going back and forth. *Trends Genet.* **13:** 152-156.

Bajic, V.B., Seah, S.H., Chong, A.G., Koh, J.L., Brusic, Y. 2002. Dragon promoter finder: recognition of vertebrate RNA polymerase II promoters. *Bioinformatics* **18**: 198-199.

Balyan, H.S., Gupta, P.K., Rustgi, S., Bandopadhyay, R., Goyal, A., Singh, A., Kumar, A., Kumar, N., Sharma S. 2005. Development and use of SSRs of bread wheat for genetic and physical mapping and transferability to the species of *Triticum-Aegilops* complex. *Czech J Genet Pl Breed* **3:** 141-144.

Barcaccia, G., Baumlein, H., Sharbel, T. 2007. Apomixis in St. John's wort (*Hypericum perforatum* L.): an overview and glimpse towards the future. *Plant Journal.***47:** 259–80.

Bashaw, E.C., and Hanna, W.W. 1990. Reproductive versatility in the grasses. In G. P. Chapman (ed.) Apomictic reproduction. Cambridge U. K. Cambridge University Press. Pp. 100-130.

Behringer, F.J., and Medford, J.I., 1992. A plasmid rescue technique for the recovery of plant DNA disrupted by T-DNA insertion. *Plant Mol. Biol. Report.* **10:** 190-198.

Bhojwani, S.S. 1968. Morphogenetic studies on cultured mature endosperm and embryo of some angiosperms. Ph.D. thesis, University of Delhi, Delhi, 85pp.

Biradar, S.K., Sunderaram, R.M., Thirumurugan, T., Bentur, J.S., Amudhan, S., Shenoy, V.V., Mishra, B., Bennett, J., Sarma, N.P. 2004. Identification of flanking SSR markers for major rice gall midge resistance gene *GmI* and their validation. *Theor and Appl Genet.* **109:** 1468-1473.

Burke, T.W. and Kandonaga, J.T. 1997. The downstream promoter element, DPE, is conserved from *Drosphila* to humans and recognized by TAFII60 of *Drosophila. Genes Dev* **11:** 3020-3031.

Carman, J.G. 1997. Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biol. J. Linn. Soc.* **61:** 51–94.

Cheema, K.K., Bains, N.S., Mangat, G.S., Das, A., Vikal, Y., Brar, D.S., Khush, G.S., Singh, K. 2007. Development of high yielding IR64 × *Oryza rufipogon* (Griff.) introgression lines and identification *Euphytica*: status Online.

Chen, L., Guan, L., Sio, M., Hoffmann, F., Adachi, T. 2005. Developmental expression of ASG-1 during gametogenesis in apomictic guinea grass (*Panicum maximum*). *J. Plant Phy.* **162:** 1141–48.

Chen, L., Miyazaki, C., Kojima, A., Saito, A., Adachi, T. 1999. Isolation and characterization of a gene expressed during early embryo sac development in apomictic Guinea grass (*Panicum maximum*). *J. Plant Physiol.* **154:** 55–62.

Chen, Q.H., Wang, Y.C., Li, A.N., Zhang, Z.N., Zheng, X.B. 2007. Molecular mapping of two cultivar-specific avirulence genes in the rice blast fungus *Magnoporthe grisea. Mol Genet and Genomics.* **227**: 137-148.

Chen, M., Presting, G., Barbazuk, W.B., Goicoechea, J.L., Blackmon, B., Fang, G., Kim, H., Frisch, D., Yu, Y., Sun, S., Higingbottom, S., Phimphilai, J., Phimphilai, D., Thurmond, S., Gaudette, B., Li, P., Liu, J., Hatfield, J., Main, D., Farrar, K., Henderson, C., Barnett, L., Costa, R., Williams, B., Walser, S., Atkins, M., Hall, C., Budiman, M.A., Tomkins, J.P., Luo, M., Bancroft, I., Salse, J., Regad, F., Mohapatra, T., Singh, N.K., Tyagi, A.K., Soderlund, C., Dean, R.A., Wing, R.A. 2002. An integrated physical and genetic map of the rice genome. *Plant Cell* **14:** 537–545.

Chengkun, H., Moul, D., Zhihong, L., Faping, D., Fengling, L., Ray, W. 2007. An efficient method for producing an indexed, insertional-mutant library in rice Genomics. *Elsevier* **89:** 532–540.

Cho, Y.G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., Park, W.D., Ayres, N., Cartinhour, S., McCouch, S. 2000. Diversity of microsatellite derived from genomic libraries and Genebank sequences in rice (*Oryza sativa* L.). *Theor and Appl Genet.* **100**: 713-722.

Daniel, M.V. and David, W.M, 1994. Embryogenic transformation of the suspensor in twin, a polyembryonic mutant of *Arabidopsis. Developmental biology.* **165**: 566-578.

Davey, M.R., Rech, E.L., Mulligan, B.J. 1989 Direct DNA transfer to plant cells. *Plant Mol. Biol.* **13:** 273-285.

Dean, C., Sjodin, C., Page, T., Jones, J., Lister, C. 1992. Behaviour of the maize transposable element *Ac* in *Arabidopsis thaliana. Plant J.* **2:** 69–82.

DeCleene, M. and De Ley, J. 1976. The host range of crown gall. *Bot Rev* **42:** 389–466.

Delseny, M., Salses, J., Cooke, R., Sallaud, C., Regad, F., Lagoda, P., Guiderdoni, E., Ventelon, M., Brugidou, C. and Ghesquiere, A. 2001 Rice genomics: present and future. *Plant Physiol. Biochem.* **39:** 324–334.

Dhaliwal, H.S., Das, A., Singh, A., Gupta, V.K. 2001. Isolation of insertional mutants in i*ndica* rice using *Ds* transposable element of maize. *Rice Genetics News Letter*. **18**: 98-99.

Dujardin, M. and Hanna, W.W. 1989. Crossability of pearl millet with wild *Pennisetum* species. *Crop Sci*. **29:** 77-80.

Ebina, M., Nakagawa, H., Yamamoto, T., Araya, H., Tsuruta, S. 2005. Co-segregation of AFLP and RAPD markers to apospory in Guineagrass (*Panicum maximum* Jacq.). *Grassland Sci.* **51:** 71–78.

Feldmann, K.A. 1991 T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *Plant J.* **1:** 71–82.

Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y.L.Y., Zhu, J., Liu, Y., Hu, X., Jia, P.Z.Y., Zhao, Q., Ying, K., Yu, S., Tang, Y., Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L.S., Fan, D., Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J.W.M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R., Yin, H., Cai, Z., Ren, S., Lv, G.G.W., Zhu, G., Tu, Y., Jia, J., Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu., Q.Z.X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S., Ni, L., Zhu, F., Chen, W., Lan, L.L.Y., Cheng, Z., Gu, M., Jiang, J., Li, J., Hong, G., Xue, Y., Han, B. 2002. Sequence and analysis of rice chromosome 4. *Nature* **420:** 316–320

Frame, B., Shou, H., Chikwamba, R., Zhang, Z., Xiang, C., Fonger, T., Pegg, S.E., Li, B., Nettleton, D., Pei, D., Wang, K. 2002. *Agrobacterium*-mediated transformation of maize embryos using a standard binary vector system. *Plant Physiol* **129:** 13–22.

Fukuoka, H., Ogawa, T., Mitsuhara, I., Iwai T., Isuzigawa, K., Nishizawa, Y., Gotoh, Y., Nishizawa, Y., Tagiri A., Ugaki, M. 2000. *Agrobacterium*-mediated transformation of monocot and dicot plants using the NCR promoter derived from soybean chlorotic mottle virus. *Plant Cell Rep* **19:** 815–820.

Gasch, A., Aoyama, T., Chua, N.H. 1992. Gene isolation with the polymerase chain reaction. In methods in *Arabidopsis* research (Konez, C., Chua, N.H., and Schell, J., eds) London: word scientific pp. 342- 356.

Goff, S.A. 1999. Rice as a model for cereal genomics. *Curr. Opin. Plant Biol.* **2:** 86–89.

Grimanelli, D., Garcia, M., Kaszas, E., Perotti, E., Leblanc, O. 2003. Heterochronic expression of sexual reproductive programs during apomictic development in *Tripsacum*. *Genetics* **165:** 1521–31.

Grimanelli, D., Leblanc, O., Espinosa, E., Perotti, E., Gonzales, de., Leon, D., Savidan, Y. 1998.Mapping diplosporous apomixis in tetraploid *Tripsacum*: one gene or several genes? *Heredity* **80:** 33–39.

Hansen, G. and Wright, M.S. 1999. Recent advances in the transformation of plants. *Trends Plant Sci.* **4:** 226–231.

Herr, Jr J.M. 1971. A new clearing squash technique for the study of ovule development in angiosperms. *American J of Bot.* **58:** 785-790.

Hiei, Y., Ohta, S., Komari, T., Kumashiro, T. 1994. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J* **6:** 271–282.

Hossain, K.G., Kalavacharla, V., Lazo, G.R., Hegstad, J., Wentz, M.J., Simons, K., Gehlhar, S., Rust, J.L., Syamala, R.R., Obeori, K., Bhamidimarri, S., Karunadharma, P., Chao, S., Anderson, O.D., Qi, L.L., Echalier, B., Gill, B.S., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Miftahudin., Ross, K., Gustafson, J.P., Sidhu, D., Dilbirligi, M., Gill, K.S., Peng, J.H., Lapitan, N.L.V., Greene, R.A., Bermudez-Kandianis, C.E., Sorrells, M.E., Feril, O., Pathan, M.S., Nguyen, H.T., Gonzalez-Hernandez, J.L., Wennerlind, E.J., Anderson, J.A., Fenton, D., Close, T.J., McGuire, P.E., Qualset, C.O., Kianian, S.F. 2004. A chromosome bin map of 2148 EST loci of wheat homeologous group 7. *Genetics* **168:** 687-699.

Hsing, Y.I., Chern, C.G., Fan, M.J., Lu, P.C., Chen, K.T., Lo, S.F., Sun, P.K., Ho, S.L., Lee, K.W. 2007. A rice gene activation/ knockout mutant resource for high throughput functional genomics. *Plant Mol. Biol* **63:** 351–364.

Hui, Ek-W., Wang, P.C., Lo, S.J. 1998. Strategies for cloning unknown cellular flanking DNA sequences from foreign integrants. *Cell mol. Life Sci.* **54:** 1403-1411.

IRGSP. The map-based sequence of the rice genome. 2005. *Nature* **436:** 793–800

Ishida, Y., Saito, H., Ohta, S., Hiei, Y., Komari, T., Kumashiro, T. 1996. High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. *Nature Biotechnol* **14:** 745–750.

Itoh, T., Tanaka, T., Barrero, R.A., Yamasaki, C., Fujii,Y., Hilton, P.B., Antonio, B.A., Aono, B., Apweiler, A., Bruskiewich, R., Bureau, T., Burr, F., Oliveira, A.C., Fuks, G., Habara, T., Haberer, G., Han, B., Harada, E., Hiraki, A.T., Hirochika, H., Hoen, D., Hokari, H., Hosokawa, S., Hsing, Y., Ikawa, H., Ikeo, K., Imanishi, T., Jaiswal, P., Kanno, M., Kawahara, Y., Kawamura, T., Kawashima, H., Khurana, J.P., Kikuchi, S., Komatsu, S., Koyanagi, K.O., Kubooka, H., Lieberherr, D., Lin, Y.C., Lonsdale, D., Matsumoto, T., Matsuya, A., McCombie, W.R., Messing, J., Miyao, A., Mulder, N., Nagamura, Y., Nam, J., Namiki, N., Numa, H., Nurimoto, S., Donovan, C., Ohyanagi, H., Okido, T., Ota, S., Osato, S., Palmer, L.E., Quetier, F., Raghuvanshi,

S., Saichi, N., Sakai, H., Sakai, Y., Sakata, H., Sakurai, T., Sato, F., Sato, Y., Schoof, H., Seki, M., Shibata, M., Shimizu, Y., Shinozaki, K., Shinso, Y., Singh, N.K., White, B.S., Takeda, J., Tanino, M., Tatusova, T., Thongjuea, S., Todokoro, F., Tsugane, F., Tyagi, A.K., Vanavichit, A., Wang, A., Wing, R.A., Yamaguchi, K., Yamamoto, M., Yamamoto, N., Yu, Y., Zhang, H., Zhao, Q., Higo, Burr, B., Gojobori, T., Sasaki, T. 2007. Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*. *Genome Research*. **17:** 175–183.

Jena, K., Jeung, J., Lee, J., Brar, D. 2005. High resolution mapping of a new planthopper (BPH) resistance gene, Bph18(t), and marker- assisted selection for BPH resistance in rice (*Oryza sativa* L.). *Theor Appl Genet*. **110:** 1-10.

Jeon, J.S., Lee, S., Jung, K.H., Jun, S.H., Jeong, D.H., Lee, J., Kim, C., Jang, S., Yang, K .2000. T-DNA insertional mutagenesis for functional genomics in rice. *Plant J* **22**: 561–570.

Jeong, D.H., An, S., Park, S., Kang, H.G., Park, G.G., Kim, S.R., Sim, J., Kim, Y.O., Kim, M.K. 2006. Generation of a flanking sequence-tag database for activation-tagging lines in *japonica* rice. *Plant J* **45:** 123–132.

Johansson, K., El-Ahmad, M., Friemann, R., Jornvall, H., Markovic, O., Eklund, H. 2002. Crystal structure of plant pectin methylesterase. *FEBS letters* **514:** 243-249.

Jung, K.H., An, G., Ronald, P.C. 2008. Towards a better bowl of rice:assigning function to tens of thousands of rice genes. *Nat Rev Genet.* **9:** 91-101.

Kalavacharla, V., Hossain, K., Gu, Y., Riera-Lizarazu, O., Vales, M.I., Bhamidimarri, S., Gonzalez-Hernandez, J.L., Maan, S.S., Kianian, S.F. 2006. High-Resolution Radiation Hybrid Map of Wheat Chromosome 1D. *Genetics*. **173:** 1089-1099.

Kellogg, E.A. 2001. Evolutionary history of the grasses. *Plant Physiol.* **125:** 1198–1205.

Khush, G.S. 1994. Apomixis: Exploiting hybrid vigor in rice. eds pp. 39. (Los Banos, Philippines: IRRI).

Khush, G.S. 1997. Origin, dispersal, cultivation and variation of rice. *Plant Mol. Biol.* **35:** 25–34.

Khush, G.S., Brar, D.S., Bennett, J., Virmani, S.S. 1994. Apomixis for rice improvement. In: G.S. Khush (Ed.), Apomixis: Exploiting Hybrid Vigor in Rice. pp. 1–17. International Rice Research Institute, Los Banos, Philippines.

Khush, G.S. and Singh, K. 2007. Development of high yielding IR64 × *Oryza rufipogon* (Griff.) introgression lines and identification *Euphytica*: status Online.

Knudsen S. 1999. Promoter 2.0: for the recognition of PolII promoter sequences. *Bioinformatics* **15:** 356-361.

Kolesnik, T., Szeverenyi, I., Bachmann, D., Kumar, C.S., Jiang, S., Ramamoorty, S., Cai, M., Ma, Z.G., Sundaresan, V., Ramachandran, S. 2004. Establishing an efficient *Ac/Ds* tagging system in rice: large scale analysis of *Ds* flanking sequences. *Plant J.* **37:** 301-314.

Komari, T., and Kubo, T. 1999. Methods of genetic transformation: *Agrobacterium tumefaciens.* (In. I.K. Vasil, Molecular improvement of cereal crops. eds). Pp.49.

Kral, L. 1999. Penetrance and Expressivity. Australian Shepherd Health and Genetics Inforamtion Resource and Health Registry. Online.

Kumar, C.S., Wing, R.A., Sundaresan, V. 2005. Efficient insertional mutagenesis in rice using the maize *En/Spm* elements. *Plant J.* **44:** 879–892.

Kumar and Hirochika. 2001. Applications of retrotransposons as genetic tools in plant biology. *Trends Plant Sci* **6:** 127-134.

Kumar, L.S. 1999. DNA markers in plant improvement: An overview. *Biotechnol adv.* **17:** 143-182.

Kumar, N., Kulwal, P.L., Balyan, H.S., Gupta, P.K. 2007. QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. *Mol Breed.* **19:** 163-177.

Lakshmi, M., Parani, M., Rajalakshmi, S., Parida, A. 2002. Analysis of species relationship among seven nutritive millets using molecular markers. *J. Plant Biochem and Biotechnol.* **11:** 85-91.

Leblanc, O., Grimanelli, D., Gonzalez-de-Leon, D., Savidan, Y. 1995. Detection of the parthenogenesis may be uncoupled in *Poa pratensis*: a cytological investigation. *Sex. Plant Reprod.* **14:** 213–17.

Lipshutz, R.J., Fodor, S.P., Gingeras, T.R., Lockhart, D.J. 1999. High density synthetic oligonucleotide arrays. *Nature Genet.* **21:** 20–24.

Liu, X.Q., Wang, L., Chen, S., Lin, F., Pan, Q.H. 2005. Genetic and physical mapping of *Pi36(t)*, a novel rice blast resistance gene located on chromosome 8. *Mol Genet and Genomics.* **274 (4):** 394-401.

Liu, Y.-G., Mitsukawa, N., Oosumi, T., Whittier, R.F. 1995. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8:** 457–463.

Liu, Y.G., Whitter, R.F. 1995 Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* **25:** 674–681.

Lohe, A.R. and Chaudhury, A. 2002. Genetic and epigenetic processes in seed development. *Curr. Op. Plant Biol.* **5:** 19–25.

Londo, J.P., Chiang, Y.C., Hung., K.H., Chiang, T.Y., Schaal, B.A. 2006. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa. Proc. Natl Acad. Sci. USA* **103:** 9578–9583.

Lu, Y.G. and Liu, X.D. 2002. A mutant for variable egg number in rice *Rice Genetics Newsletter* pp. 15.

Martinez, E., Hopp, H., Stein, J., Ortiz, J., Quarin, C. 2003. Genetic characterization of apospory in tetraploid *Paspalum notatum* based on the identification of linked molecular markers. *Mol. Breeding* **12:** 319–27.

Martinez, E.J., Urbani, M.H., Quarin, C.L., Ortiz, J.P.A. 2001. Inheritance of apospory in Bahiagrass, *Paspalum notatum. Hereditas* **135:** 19–25.

Maruthachalam, R., Mohan, P.A., Marimuthu., Siddiqi, I. 2008. Gamete formation without meiosis in *Arabidopsi. Nature* **451:** 1121-1124.

Matzk, F., Meister, A., Schubert, I. 2000. An efficient screen for reproductive pathways using mature seeds of monocots and dicots. *Plant J.* **21:** 97–108.

Matzk, F., Prodanovic, S., Baumlein, H., Schubert, I. 2004. The inheritance of apomixis in *Poa pratensis* confirms a five locus model with differences in gene expressivity and penetrance. *Plant Cell* **17:** 13–24.

McClean, P. 1998. Analyzing plant gene expression with transgenic plants. Online course topics NDSU, USA.

McCouch, S.R., Teytelman, L., Xu, Y.B., Lobos, K.B., Clare, K., Walton, M., Fu, B.Y., Maghirang, R., Li, Z.K., Xing, Y.Z., Zhang, Q.F., Kono, I., Yano, M., Fjellstrom, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D., Stein, L. 2002. Development and mapping of 2,240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res* **9:** 199–207.

McCouch, S.R., Kochert, G., Yu, Z.H., Khush, G.S., Coffman, W.R., Tanksley, S.D. 1988. Molecular mapping of rice chromosomes. *Theo. Appl. Genet.* **76:** 815-829.

Melchinger, A.E. 1990. Use of the molecular markers in breeding for ologogenic disease resistance. *Plant Breeding* **104:** 1-19.

Micheli, F. 2001. Protein methylesterases: cell wall enzymes with important roles in plant physiology. *Trends in Plant Sci.* **6:** 414-419.

Michelmore, R.W., Paran, I., Kesseli, R.V., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci.* USA. **88:** 9828-9832.

Michiels, A., Tucker, M., Ende, W.V.D., Laere, A.V. 2003. Chromosomal walking of flanking regions from short known sequences in GC-rich plant genomic DNA. *Plant Mol Bio Rep* **21:** 295-302.

Miki, D., Itoh, R., Shimamoto, K. 2005. RNA Silencing of Single and Multiple Members in a Gene Family of Rice. *Plant Physiology* .**138:** 1903-1913.

Miyao, A., Iwasaki, Y., Kitano, H., Itoh, J.I., Maekawa, J., Murata, K., Yatou, O., Nagato, Y., Hirochika, H. 2007. A large-scale collection of phenotypic data describing an insertional mutant population to facilitate functional analysis of rice genes. *Plant Mol Biol* **63:** 625–635.

Miyao, A., Tanaka, K., Murata, K., Sawaki, H., Takeda, S., Abe, K., Shinozuka, Y., Onosato, K., Hirochika, H. 2003. Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. *Plant Cell* **15:** 1771–1780.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T.G., Yaun, M., Bhatia, C.R., Sasaki, T. Genome mapping, molecular markers and marker assisted selectin in crop plants. 1997. *Molecular breeding* **3:** 87-103.

Mohanty, A., Sharma, N.P., Tyagi, A.K. 1999. *Agrobacterium*-mediated high frequency transformation of an elite *indica* rice variety Pusa Basmati and transmission of the transgene to R2 progeny. *Plant Sci.* **147:** 125-135.

Mu, X.J., Shi, G.C., Zhu, Z.Q., Cai, X. 1998. Polyembryony in Rice Ap111 (Shuang 13). *Apomixis News Letter.* pp. 10.

Murray, M.G. and Thompson, W.F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic acid Res.* **8:** 4321-4325.

Muniyamma, M. 1978. Variations in microsporogenesis and the development of embryo sacs in *Echinochioa stagnina* (Gramineae), *Bot. Gaz.* **139:** 87-92.

Nakazaki, T. and Ikehashi, H. 1998. Genomic sequence and polymorphisms of a rice chitinase gene, Cht4. *Breeding Sci* **48:** 371-376.

Naumava, T.N. 1997. Apomixis in tropical fodder crops: Cytological and functional aspects. *Euphytica.* **96:** 93-99.

Nobuta, N., Venu, R.C., Lu, C., Bélo, A., Vemaraju K. 2007. An expression atlas of rice mRNAs and small RNAs. *Nature Biotechnol.* **25:** 473–477.

Ochman, H., Medhora, M., Garza, D., Hartl, D.L. 1989. Inverse polymerase chain reaction. In PCR: Applications and Protocols, M. Innis, D. Gelfand, J. Sninsky, and T. Whiíe, eds (New York: Academic Press), pp. 105-111.

Ohyanagi, H., Tanaka, T., Sakai, H., Shigemoto, Y., Yamaguchi, K., Habara, T., Fujii, Y., Antonio, B.A., Nagamura, Y., Imanishi, T., Ikeo, K., Itoh, T., Gojobori, T., Sasaki, T. 2006. The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucleic Acids Res.* **34:** D741–D744.

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thibaud-Nissen, F., Malek, R.L., Lee, Y., Zheng, L., Orvis, J., Haas, B., Wortman, W., Buell, C.R. 2007. The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res.* **35:** D883–887.

Ozias-Akins, P., Lubbers, E.L., Hanna, W.W., McNay, J.W. 1993. Transmission of the apomictic mode of reproduction in *Pennisetum*: co-inheritance of the trait and molecular markers. *Theor. Appl. Genet.* **85:** 632–38.

Ozias-Akins, P., Roche, D., Hanna, W.W. 1998. Tight clustering and hemizygosity of apomixis-linked molecular markers in *Pennisetum squamulatum* implies genetic control of apospory by a divergent locus which may have no allelic form in sexual genotypes. *Proc. Natl. Acad. Sci. USA* **95:** 5127–32.

Ozias-Akins, P. and van Dijk, P. 2007. Mendelian Genetics of Apomixis in Plants. *Annu. Rev. Genet.* **41**:509–37.

Panaud, O., Chen, X., McCouch, S.R. 1995. Frequency of microsatellite sequences in rice (*Oryza sativa* L.). *Genome* **38:** 1170–1176.

Peng, J., Kononowicz, H., Hodges, T.K. 1992. Transgenic *indica* rice plant. *Theor. Appl. Genet.* **83:** 855-863.

Pessino, S.C., Evans, C., Ortiz, J.P.A., Armstead, I., do Valle, C.B., Hayward, M.D. 1998. A genetic map of the apospory-region in *Brachiaria* hybrids: identification of two markers closely associated with the trait. *Hereditas* **128:** 153–58.

Pessino, S.C., Espinoza, F., Martinez, E.J., Ortiz, J.P.A., Valle, E.M., Quarin, C.L. 2001. Isolation of cDNA clones differentially expressed in flowers of apomictic and sexual *Paspalum notatum*. *Hereditas* **134:** 35–42.

Philips, R.L. and Vasil, I.K. 2001. DNA based markers in plants (2nd ed). Kluwer Academic Publishers pp 497.

Pinheiro, A.A., Pozzobon, M.T., do Valle, C.B., Penteado, M.I.O., Carneiro, V.T.C. 2000. Duplication of the chromosome number of diploid *Brachiaria brizantha* plants using colchicine. *Plant Cell Rep.* **19:** 274–78.

Prashanth, S.R., Parani, M., Mohanty, B., Parida, A. 2002. Genetic diversity in some cultivated varieties and landraces of *Oryza sativa* L from India as revealed by AFLP markers. *Genome* **45:** 451-459.

Pupilli, F., Martinez, E.J., Busti, A., Calderini, O., Quarin, C.L., Arcioni, S. 2004. Comparative mapping reveals partial conservation of synteny at the apomixis locus in *Paspalum* spp. *Mol. Genet. Genom.* **270:** 539–48.

Pupilli, F., Caceres, M.E., Quarin, C.L., Arcioni, S.1997. Segregation analysis of RFLP markers reveals a tetrasomic inheritance in apomictic *Paspalum simplex*. *Genome* **40:** 822–28.

Qi, H., Dongbo, L., Yongsheng, Z., Mugpu, T., Duanpin, Z., Xinghua, L. 2006. Fine mapping of *Xa2*, a bacterial blight resistance gene in rice. *Mol Breeding*. **17 (1):** 1-6.

Rani, S. 2005. Present status of Basmati research in India- Issues and future strategies. In the proceedings of the National symposium on Basmati Rice Research, SVBP Univ. of Agri and Tech. Modipuram (U.P.) 6-7 September 2005.

Rao, K.K. 2003. Molecular tagging of a new bacterial blight resistance gene in rice using RAPD and SSR markers. IRRN **28:** 16-17.

Rao, P.S. 1965. *In vitro* induction of embryonal proliferation in *Santalum album* L. *Phytomorphology*. **15:** 175-181.

Rashid, H., Yokoi, S., Toriyama, K., Hinata, K. 1996. Transgenic plant production mediated by *Agrobacterium* in *Indica* rice. *Plant Cell Rep.* **15:** 727-730.

Richards, A.J. 1986. Plant breeding systems. Unwin Hyman, London. 529 pp.

Rice today, International Rice Research Institute. Oct-Dec 2006, **5:** No.4.

Rishi, A.S., Nelson, N.D., Goyal, A. 2004. genome walking of large fragments: an improved method. *J of Biotechno*. **111:** 9-15.

Röder, M.S., Plaschke, J., Konig, S., Borner, A., Sorells, M.E., Tanksley, S.D., Ganal, M.W. 1998. Abundance, variability and chromosomal location of microsatellites in wheat. *Mol and Gen Genet*. **246:** 327-333.

Rodrigues, J.C.M., Cabral, G.B., Dusi, D.M.A., de Mello, L.V., Rigden, D.J., Carneiro, V.T.C. 2003. Identification of differentially expressed cDNA sequences in ovaries of sexual and apomictic plants of *Brachiaria brizantha*. *Plant Mol. Biol.* **53:** 745–57.

Sachs, E.S., Benedict, J.H., Stelly, D.M., Taylor, J.F., Altman, D.W., Berberich, S.A., Davis, S.K. 1998. Expression and segregation of genes encoding *Cry IA* insecticidal proteins in cotton. *Crop Sci.* **38:** 1-11.

Saito, A., Shimosaka, E., Hayano, Y., Saito, K., Matuura, S., Yano. 1991. Molecular mapping on chromosom 3 of rice *Jap J of Breed.* **41:**158-159.

Sallaud, C., Gay, C., Larmande, P., Bes, M., Piffanelli, P., Piegu, B., Droc, G., Regad, F., Bourgeois, E . 2004. High-throughput T-DNA insertion mutagenesis in rice: a first step towards *in silico* reverse genetics. *Plant J* **39:** 450–464.

Sanchez, A.C., Brar, D.S., Huang, N., Li, Z., Khush, G.S. 2000. Sequence tagged site marker-assisted selection for three bacterial blight resistance genes in rice. *Crop Sci.* **40:** 792–797.

Sasaki, T. 2003. Rice genome analysis: Understanding the genetic secrets of the rice plant. *Breeding Science* **53:** 281–289.

Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y., Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Ito, S., Ito, T., Ito, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, JH., Kim, HI., Eun, MY., Yano, M., Jiang, J., Gojobori, T. 2002. The genome sequence and structure of rice chromosome 1. *Nature* **420:** 312–316.

Savidan Y. 2000. Apomixis: genetics and breeding. *Plant Breed. Rev.* **18:**13–85.

Savidan, Y. 2001. Transfer of apomixis through wide crosses. In The Flowering of Apomixis:From Mechanisms to Genetic Engineering, ed. Y Savidan, J Carman, T Dresselhaus, pp. 153–67. Mexico, D.F.: CIMMYT, IRD, Eur. Comm. DG VI (FAIR).

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, J., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, J., Goff,

F.A. 2002. A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* **14:** 2985–2994.

Shahmuradov, I., Gammerman, A., Hancock, J.M., Bramley, P.M., Solovyer, V.V. 2005. PlantProm: a database of plant promoter sequences. *Nucleic Acids Res* **31:** 114-117.

Sharma, T.R., Shanker, P., Singh, B.K., Jana, T.K., Madhav, M.S., Gaikwad, K., Singh, N.K., Plaha, P., Rathour, R. 2005. Molecular mapping of rice blast resistance gene *Pik h* in the rice variety Tetep. *J Plant Biochem. Biotech* **14:**127-133 (a).

Sharma, T.R., Madhav, M.S., Singh, B.K., Shanker, P., Jana, T.K., Dalal, V., Pandit, A., Singh, A., Gaikwad, K., Upreti, H.C., Singh, N.K .2005. High-resolution mapping, cloning and molecular characterization of the gene of rice, which confers resistance to rice blast. *Mol Genet. Genomics*. **274:** 569-578 (b).

Shimamoto, K. and Kyozuka, J. 2002. Rice as a model for comparative genomics of plants. *Annu. Rev. Plant Biol.* **53:** 399–419

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., Lukyanov, S.A. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* **23:** 1087–1088.

Singh, D., Sirohi, A., Sharma, V., Kumar, S., Gaur, V., Chaudhary, A., Chaudhary, R.N .2006. RAPD based identification of farmers' collection of Kalanamak and traditional varieties of basmati rice. *Indian J. Crop Sci.*. **1:** 102-105 (a).

Singh, N.K., Vivek, D., Batra, K., Singh, B.K., Chitra, G., Singh, A., Ghazi, I.A., Yadav, M., Pandit, A., Dixit, R., Singh, P.K., Singh, H., Koundal, K.R., Gaikwad, K., Mohapatra, T., Sharma, T.R. 2007. Single-copy genes define a conserved order between rice and wheat for understanding differences caused by duplication, deletion, and transposition of genes. *Funct Integr Genomics* **7:** 17–35.

Singh, P.K., Gonzalez-Hernandez, J.L., Mergoum, M., Ali, S., Adhikari, T.B., Kianian, S.F., Elias, E.M., Hughes, G.R. 2006. Identification and molecular Mapping of a gene in tetraploid wheat conferring resistance to *Pyrenophora tritici-repentis* race 3. *Phytopathology* **96:** 885-889 (b).

Singh, S., Sidhu, J.S., Huang, N., Vikal, Y., Li, Z., Brar, D.S., Dhaliwal, H.S., Khush, G.S. 2001. Pyramiding three bacterial blight resistance genes (Xa5, Xa13 and Xa21) using marker-assisted selection into *indica* rice cultivar PR106. *Theor. Appl. Genet.* **102:** 1011–1015.

Small, S.T. 1997. Transcription initiation from TATA-less promoters within eukaryotes protein-coding genes. *Biochim.Biophys.Acta* **1351:** 73-88.

Spillane, C., Stimer, A., Grossniklaus, U. 2001. Apomixis in agriculture: the quest for clonal seeds. *Sex. Plant Reprod.* **14:**179–87.

Spradling, A.C., Stern, D.M., Kiss, I., Roote, J., Laverty, T., Rubin, G.M. 1995. Gene disruptions using P transposable elements: An integral component of the *Drosophila* genome project. *Proc. Natl. Acad. Sci USA* **92:** 10824-10830.

Stein, J., Quarin, C.L., Martinez, E.J., Pessino, S.C., Ortiz, J.P.A. 2004. Tetraploid races of *Paspalum notatum* show polysomic inheritance and preferential chromosome pairing around the apospory-controlling locus. *Theor. Appl. Genet.* **109:** 186–91.

Swamy, B.G.L., 1943. Gametogenesis and embryogeny in *Eulophia epidendra*. Fisher. *Proc. Indian. Acad. Sci.* **9:** 59-63.

Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., McCouch, S.R. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency length, variation transposon associations and genetic marker potential. *Genome Res* **11:** 1441–1452.

Temnykh, S., Park, W.D., Ayres, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y.G., Ishii, T., McCouch, S.R. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theor Appl Genet* **100:** 697–712.

Terauchi, R. and Kahl, G. 2000. Rapid isolation of promoter sequences by TAIL-PCR: the 5′ flanking regions of *Pal* and *Pgi* genes from yams (*Dioscorea*). *Mol Gen Genet* **263:** 554-560.

Toriyama, K., Arimoto, Y., Uchimiya, H., Hinata, K. 1988. Transgenic rice plants after direct gene transfer into protoplasts. *BioTech.* **6:** 1072-1074.

Tsugeki, R., Kochieva, E.Z., Fedoroff, N.V. 1996. A transposon insertion in the *Arabidopsis SSR16* gene causes an embryodefective lethal mutation. *Plant J.* **10:** 479–489.

Upadhyaya, N.M. 2007 eds. Rice functional Genomics: Challenges, Progress and Prospects. G. Springer- Verlag Berlin Heidelberg.

Upadhyaya, N.M., Zhu, Q.H., Dennis, L. 2006. A Gene Machine for Functional Genomics of Rice. RIRDC Publication No W05/194 RIRDC Project No CSP-8A.

Varshney, R.K., Prasad, M., Graner, A. 2004. Molecular marker maps of barley: a resource for intra and interspecific genomics. In: Molecular marker systems in plant breeding and crop improvement (eds.) Lorz, H and Wenzel, G. Springer- Verlag Berlin Heidelberg. pp 229-239

Wang, S., Sim, T.B., Kim, Y.S., Chang, Y.T. 2004. Tools for target identification and validation. *Curr. Opin. Chem. Biol.* **8:** 371–377.

Weil, M.D., and McClelland, M., 1989. Enzymatic cleavage of a bacterial genome at a 10-base pair recognition site. *Proc Natl. Acad Sci. USA* **1:** 51-55.

White D. and Chen W. 2006. Genetic transformation of *Ascochyta rabiei* using *Agrobacterium*-mediated transformation. *Current Genet.* **49:**272-280.

Wing, R.A., Luo, M., Kim, H., Yu, Y., Kudrna, D., Goicoechea, J., Wang, W., Nelson, W., Rao, K., Brar, D., Mackill, D., Han, B., Soderlund, C., Stein, L., SanMiguel, P., Jackson, S. 2005. The *Oryza* map alignment project: the golden path to unlocking the genetic potential of wild rice species. *Plant Mol. Biol.* **59:** 53–62.

Ximei, D., Qunce, I., Guoping, L., Xiuming, H., Guangyong,Q., Zengliang, Y. 2006. Study of Genetics and Embryology of Polyembryonic Mutant of Autotetraploid Rice Induced by N+ Beam Implantation. *Plasma Sci and Tech.* **8:** 6-10.

Xiao, J., Li, J., Grandillo, S., Ahu, S.N., Yuan, L., Tanksley, S.D., McCough, S.R. 1998. Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon. Genetics* **150:** 899-909.

Yang, Z., Sun, X., Wang, S., Zhang, Q. 2003. Genetic and physical mapping of a new gene for bacterial blight resistance in rice. *Theor and Appl genet.* **106:** 1467-1472.

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, Y., Tong, X., Cong, L., Geng, L., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, J., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, X., Xu, X., Zhai, X., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, Z., Tao, L., Ye, J., Tan, T., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, W Zhu, L., Yuan, L., Yang H. 2005. The genomes of *Oryza sativa*: a history of duplications. *PLoS Biol.* **3:** e38.

Yuan, Q., Ouyang, S., Wang, A., Zhu, W., Maiti, R., Lin, H., Hamilton, J., Haas, B., Sultana, R., Cheung, F., Wortman, J., Buell, C.R. 2005. The Institute for Genomic Research *Osa1* rice genome annotation database. *Plant Physiol.* **138:** 18–26.

Zhang, J., Li, C., Wu, C., Xiong, L., Chen, G., Zhang, Q., Wang, S . 2006. RMD: a rice mutant database for functional analysis of the rice genome. *Nucleic Acids Res* **34:** D745–D748.

Zhang, J.Z., and Somerville, C.R. 1997. Suspensor derived polyembryony caused by altered expression of valyl tRNA synthetase in the *twn2* mutant of *Arabidopsis*. *Proc. Nat. Acad. Sci. USA*. **94:** 7349-7355.

Zhang, J., Xu, R., Blackley, D., Elliott, M.C., Chen, D.F. 1997. *Agrobacterium*-mediated transformation of elite *Indica* and *Japonica* rice cultivars. *Mol. Biotechnol.* **8:** 223-231.

Zhang, W. and Wu, R. 1988. Efficient regeneration of transgenic plants from rice protoplasts and correctly regulated expression of the foreign gene in the plants. *Theor. Appl. Genet.* **76:** 835-840.

Zhao, Y., Christensen, S.K., Fankhauser, C., Cashman, J.R., Cohen, J.D., Weigel, D. Chory, J. 2001. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science*. **291:** 306–309.

Zhu, L.C., Smith, C.M., Fritz, A.K., Boyko, E., Voothuluru, P., Gill, B.S. 2005. Inheritance and molecular mapping of new greenbug resistance genes in wheat germplasms derived from *Aegilops tauschii*. *Theor Appl Genet*. **111:** 831-837.

LIST OF PUBLICATIONS

1. M Kumar, D Rajpurohit, P O Basha, **A Bhalla**, G S Randhawa and H S Dhaliwal. (2008) Genetic control of after-ripening seed dormancy in Basmati rice. (communicated)

2. **A Bhalla**, P O Basha, Mankesh Kumar, Kuldeep Singh, Deepak Rajpurohit G.S. Randhawa and H S Dhaliwal (2008) Molecular mapping of polyembryonic, Oligoculm and Dwarf insertional mutants in Basmati 370. (Ready for communication)

3. Characterization of *Ds* insertional Polyembryonic, Oligo culm, and Dwarf mutants in Basmati 370. **Anju Bhalla**, P. Osman Basha, Mankesh Kumar, Deepak Rajpurohit, G.S.Randhawa and H.S.Dhaliwal.
   Accepted for presentation at National biotechnology Conference-2006, Current Trends and Future Perspectives, 2-3 September 2006. Department of Biotechnology, IIT Roorkee and Indian Federation of Biotechnologists (IFB-India).

4. Characterization of *Ds* insertional Polyembryonic (*OsPE*) mutant of Basmati 370. **Anju Bhalla**, G.S. Randhawa and H.S. Dhaliwal.
   Accepted for presentation at National symposium on Basmati Rice Research: Current Trends and Future Prospects, 6-7 September, 2005. Department of Genetics and Plant Breeding SVBP university of Agriculture and Technology, Merrut.

5. Pyramiding of genes for dwarfing and bacterial blight resistance in Basmati type 3. Deepak Rajpurohit , P. Osman Basha, Mankesh Kumar, **Anju Bhalla** and H.S. Dhaliwal.
   Accepted for presentation at National symposium on Basmati Rice Research: Current Trends and Future Prospects, 6-7 September, 2005. Department of Genetics and Plant Breeding SVBP university of Agriculture and Technology, Merrut.

6. Genetic and biochemical studies on stress tolerance in *Rhizobium leguminosarum*. Ihsan Arfan Ali, **Anju Bhalla** and G.S. Randhawa
   Accepted for presentation at 3rd International Nitrogen Conference held at Nanjing, China during Oct 12 to Oct 16, 2004

7. Randhawa G S, G Shubha, Singh N K, Kumar A & **Bhalla A**. (2003) Recent advances in rhizobium-legume symbiosis, *Indian J Expt Biol*, **41**:1184.