MOLECULAR DOCUMENTATION OF *NJAVARA* TYPES OF RICE (*Oryza sativa* L.) FOR CULTIVAR IDENTIFICATION

By

SHAREESH. N

THESIS

Submitted in partial fulfilment of the
requirement for the degree of

Master of Science in Agriculture

Faculty of Agriculture
Kerala Agricultural University

Department of Plant Breeding and Genetics

COLLEGE OF HORTICULTURE
VELLANIKKARA, THRISSUR - 680 656
KERALA, INDIA

2007

DECLARATION

I, **Shareesh N. (2005-11-116)** hereby declare that this thesis entitled **'Molecular documentation of *Njavara* types of rice (*Oryza sativa* L.) for cultivar identification'** is a bonafide record of research work done by me during the course of research and this thesis has not previously formed the basis for the award to me of any degree, diploma, associateship, fellowship or other similar title of any other University or Society.

Vellanikkara
Date: 6-11-07

SHAREESH.N

CERTIFICATE

Certified that this thesis, entitled **'Molecular documentation of *Njavara* types of rice (*Oryza sativa* L.) for cultivar identification'** is a bonafide record of research work done independently by **Mr. Shareesh N.** under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to him.

Dr.C.R.Elsy
Chairman, Advisory Committee
Associate Professor
Department of Plant Breeding and Genetics
College of Horticulture
Kerala Agricultural University
Thrissur, Kerala

Vellanikkara
Date: 6-11-07

CERTIFICATE

We, the undersigned members of the Advisory Committee of **Mr. Shareesh, N** a candidate for the degree of **Master of Science in Agriculture** with major in Plant Breeding and Genetics, agree that this thesis entitled **"Molecular documentation of *Njavara* types of rice (*Oryza sativa* L.) for cultivar identification"** may be submitted by **Mr. Shareesh, N** in partial fulfilment of the requirement for the degree.

Dr. C. R. Elsy
Chairperson, Advisory Committee
Associate Professor
Department of Plant Breeding and Genetics
College of Horticulture
KAU. P.O. Thrissur

Dr. V.V.Radhakrishnan
(Member)
Professor & Head
Dept. of Plant Breeding and Genetics
College of Horticulture

Dr. Dijee Bastian 2.11.07
(Member)
Assistant Professor (SS)
Dept. of Plant Breeding and Genetics
College of Horticulture

Dr. A. Augustin
(Member)
Professor
Centre for Plant Biotechnology and
Molecular Biology
College of Horticulture

EXTERNAL EXAMINER
Dr. K. Thiyagarajan

Acknowledgement

I would like to express my deep and sincere gratitude to my chairperson, ***Dr. C. R. Elsy****, Associate Professor, Department of Plant Breeding and Genetics, College of Horticulture, Kerala Agricultural University. Her wide knowledge and her logical way of thinking have been of great value for me. Her understanding, encouraging and personal guidance have provided a good basis for the present thesis.*

I am deeply grateful to my committee member, ***Dr.A.Augustin****, Associate Professor , Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Kerala Agricultural University for his detailed and constructive comments, and for his important support throughout this work.*

I wish to express my warm and sincere thanks to ***Dr.V.V.Radhakrishnan*** *Professor and Head , Department of Plant Breeding and Genetics, College of Horticulture, Kerala Agricultural University who gave me important guidance during my first steps into research. His ideas and concepts have had a remarkable influence on my entire research.*

I am deeply indebted to my committee member, ***Dr.Dijee Bastain,*** *Assistant Professor(SS) for her helpful suggestions and encouragement during the time of research and thesis writing.*

With deep reverence, I express my sincere thanks to ***Dr.D. Girija*** *,Associate Professor, CPBMB, for her treasured technical guidance, sustained interest, kind concern and ever-willing help rendered for the successful completion of the research work. She has been a support during each step of the way and my unfading gratitude and obligation remain with her forever.*

Heartfelt thanks are due to ***Dr. P. A. Nazeem****, Professor and Head,* ***Dr. Keshavachandran****, Professor , Centre for Plant Biotechnology and Molecular Biology for their precious suggestions and timely help rendered during my difficult moments.*

My warm thanks to ***Dr.K.T.Presannakumari,*** *Professor ,* ***Dr. R.Sujatha,*** *Assistant Professor (SS),* ***Dr. Jjjy Joseph,*** *Assistant Professor,* ***Dr. Sreenivasan,*** *Associate Professor and* ***Dr. Nandini K,*** *Professor, Department of Plant Breeding and Genetics for the support and guidance during the study.*

A special word of thanks is due to ***Dr.P.K.Rajeevan,*** *Professor and Associate Dean ,College of Horticulture for his splendid support and encouragement throughout the postgraduate programme.*

Thanks are due to ***Dr. Jose Mathew*** *,Professor and Head, Cashew Research Station, Madakathara for his advices and support .*

I warmly thank ***Dr. Sanal kumar.P*** *for his valuable advice and friendly help. His extensive discussions around my work and interesting explorations have been very helpful for this study.*

My sincere thanks are due to all other departmental referees for their detailed review, constructive criticism and excellent advice during the preparation of this thesis.

I owe my loving thanks to my class mates ***Liffey Zacharia Antony , Abdulla Fayas , Vinesh kumar , Kishore, Praveen, Ratheesh*** *and my beloved friends* ***Nibin, Varun , Ramya Haridas*** *for their loving support and help during the course work and thesis submission.*

My friends from the Centre for Plant Biotechnology and Molecular Biology had supported me in my research work. I want to thank them for all their help, support, interest and valuable hints. Especially I am obliged to ***Jithesh, Aneesh, Atheesh, Arun, Vipin, Firoz, Beena, Sheeba, Praveena, Rekha , Sruthi and Jose.***

I also want to thank ***Mr. Sreekumar,*** *Farm Officer, CPBMB for all his assistance during the photo session for thesis work.*

I wish to extend my warmest thanks to all those who have helped me with my work in the Department of Plant Breeding and Genetics especially my seniors ***Namboodiri Raji, Gayathri.G, Vidhu Fransis, Sani George, Marimuthu*** *and my junior* ***Divya Balakrishnan.***

A special word of thanks is due to ***Sreejith.P.V, Johnkutty George,*** *and* ***Jaisal Babu, Sajnanad*** *for their splendid support and encouragement during the course of study and thesis work.*

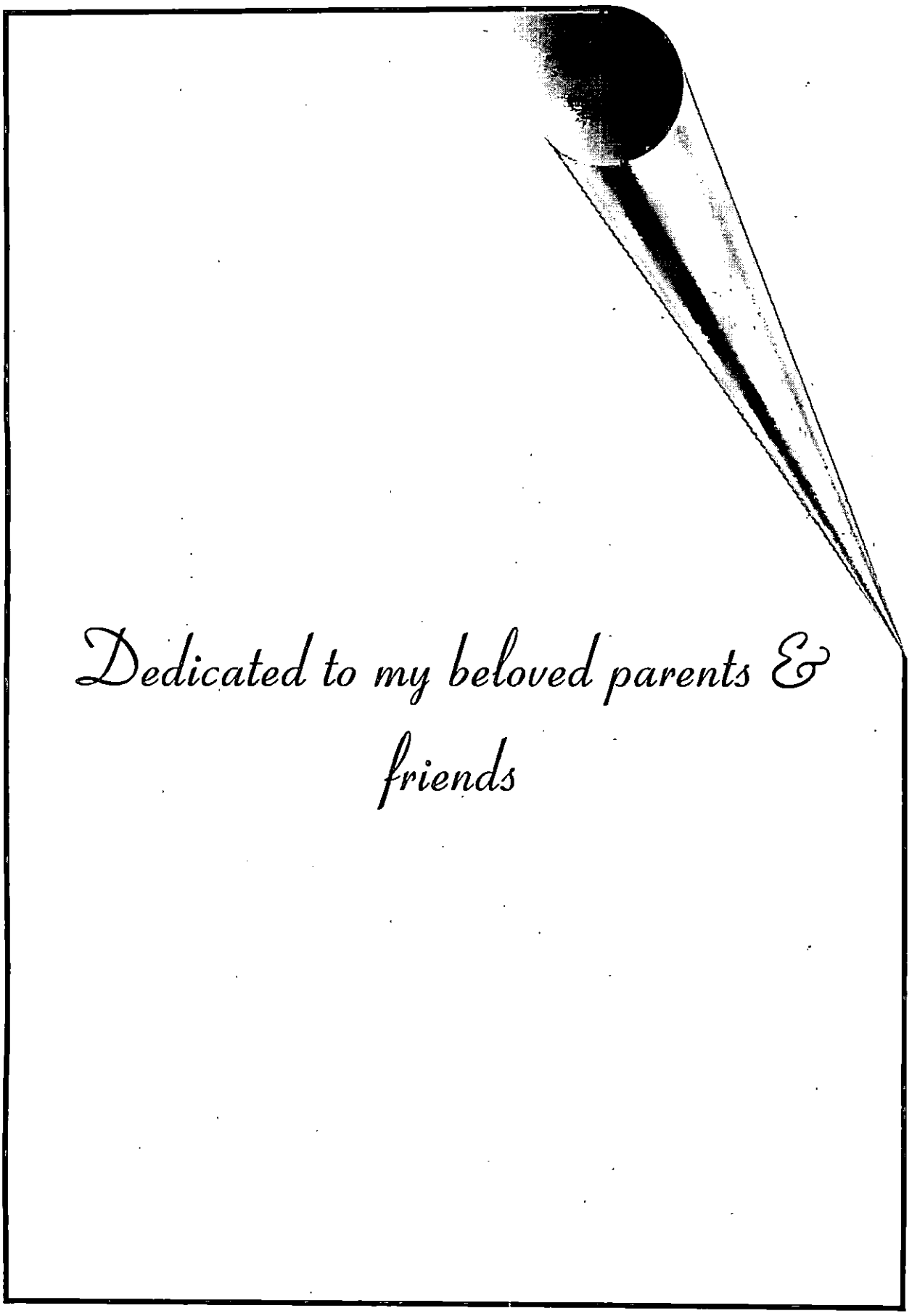
I am forever beholden to my beloved ***Achan, Amma, Ashi, Dhina and Dhannu*** *for their moral support, constant prayers, personal sacrifices and unceasing encouragement during the research work.*

Thanks are due to ***Mr. Santhosh,*** *Students Computer Centre, College of Horticulture for prompt help in the preparation of the manuscript.*

The financial support of Kerala Agricultural University through fellowships is gratefully acknowledged.

Above all I bow my head before GOD ALMIGHTY for blessing me with health, strength and confidence to get through all tedious circumstances and for showering flowers of success throughout my academic career.

Shareesh N.



Dedicated to my beloved parents & friends

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

Introduction

1. Introduction

Rice (*Oryza sativa* L.) is an ancient crop domesticated more than 8000 years ago, evolved along with man and nature adapting to diverse ecologies. It is also the principle food crop, grown across 18 different countries feeding more than half of the world's population (Virmani, 1999). In Kerala, rice is grown in an area of 2.8 lakh hectares with an annual production of 5.7 lakh tonnes during 2003-2004 (Government of Kerala, 2004).This nutritious cereal crop provides 20 per cent of the calories and 15 per cent of protein to world's population. Rice is one of three major dietary staple food in the world and has a highly syntenic genomic and gene structure with respect to the other two major food crops viz., maize and wheat.

India is known for its rich medicinal and aromatic rice types like Basmati and other fine grain aromatic types grown in north-west regions of the country. Kerala is regarded as the granary of aromatic rices like *Gandhasala* and *Jeerakasala* of Wyanad district which comes under the high ranges at an elevation of about 750m above MSL (Kumary *et.al.*, 2003).

Kerala is the cradle of Ayurveda and ayurvedic treatments. Ayurvedic treatise (*Indian Materia Medica*) speaks of cultivars like '*Njavara*' and '*Gathuhan*' for its rich medicinal properties. In Kerala, important medicinal rice cultivars grown are *"Njavara"*, *"Chennellu"* and *"Rakthasali"* and among these *Njavara* cultivated and conserved by farmers of Kerala is unique. *Njavara* cultivar maturing in sixty days, has medicinal property in redressing 'thridosha' , root cause of all body ailments (Balachandran ,2007) . Boluses of cooked *Njavara* rice are used for '*Navarakizhi*', used in the Ayurvedic treatment of various skeletal and muscular diseases, paralysis, seiatica etc. *'Njavarakizhi'* and '*Njavaratheppu'* are two major treatments in Ayurveda in conditions of arthritis, paralysis, neurological complaints, degeneration of muscles, tuberculosis etc. It is also used in religious ceremonies and certain preparations of *Njavara* are made as

offerings in temples. Ancient Ayurvedic texts distinguish yellow glumed and black glumed cultivars of *Njavara*. Black glumed *Njavara* has been used in Ayurveda treatment from the age of Charaka during BC.600. In recent days, the awareness to use *Njavara* rice, in *'NjavaraKanji'* (a special herbal preparation with rice) has shown a significant increase.

As the area under traditional cultivars is reducing in fast pace due to the spread of high yielding and improved varieties, there is every chance that this unique rice cultivar may become extinct in the near future. Although some studies were conducted on this cultivar, little or no efforts were taken for the molecular documentation of this unique cultivar. Molecular tools like Random Amplified Polymorphic DNA (RAPD) analysis can be employed to document the variability in crop plants. The RAPD technique has several advantages such as sampling of a relatively unbiased portion of the genome, ease of use, low cost, and use of a small amount of plant material (Fritsh and Rieseberg, 1996). A detailed characterisation of this cultivar is also very essential to study the level of diversity existing within the cultivar and to establish an index of genetic similarities among different ecotypes. Molecular characterisation of the cultivar will help in the protection of the intellectual property rights of farming communities over this unique endemic cultivar. Moreover, the collection and conservation of the cultivar needs immediate attention so as to bring it under the preview of the 'Biological Diversity Act, 2002' and to protect the 'farmers' rights' under 'Plant Variety Protection and Farmers' Rights Act, 2001'.

Hence the present study was taken up with the following objectives:

i) To develop molecular markers for *Njavar*a genotypes.
ii) Molecular cloning and characterisation of unique markers in *Njavara.*

Review of Literature

2. REVIEW OF LITERATURE

Njavara (*Oryza sativa* L.) is exceptionally an unique medicinal rice cultivar indigenous to Kerala. Medicinal value and extremely short growth duration make this cultivar "unique". *Njavara* is known as '*Shashtikam*', '*Shashtikasali*' or '*Snighdathandulam*' in Ayurveda. '*Shashtikam*' denotes sixty days and this peculiar rice cultivar has a life cycle of about sixty days. Buchnan (1807) reported that *Njavara* had been under cultivation in Kerala in the early 19th century. The ancient Indian documents state that this cultivar had been under cultivation in Kerala for about 2500 years since the time of Sushruthacharya , the great Indian pioneer in medicine and surgery (Kumary, 2004). *Njavara* ecotypes, carefully conserved by farmers over generations are valuable sources of genes in breeding programmes.

2.1 Njavara –the medicinal rice cultivar

Kerala is the cradle of Ayurvedic and traditional system of medicines and *Njavara* is widely used in the traditionally much acclaimed herbal therapeutic systems since the time of Sushruthacharya (2500 B.C.) he highlighted *Njavara* grain as a curative base for circulatory, respiratory as well as the digestive ailments. Traditional Ayurvedic system of medicine utilizes this cultivar for specific treatments including *'Panchakarma'* from time immemorial.

Two types of *Njavara* have been mentioned in the ancient Ayurvedic text '*Ashtatanga Hridayam*' (Vagbhata, 500 B.C.) viz., the white glumed and black glumed, both of which are used in Ayurvedic system of medicine. *'Karkidakakanji'* or '*Marunnukanji*' is a rejuvenating health drink consumed in Kerala in the month of '*Karkidakam*' of S.W. Monsoon season. It's a gruel prepared from rice grains with several other herbal ingredients and few spices.

Manakkodan (1949) reported that *Njavara* can be used for acclaimed therapy of joint pain, rheumatism and hemiplegia.

'*Njavarakizhi*' is a system in "*Panchakarma Chikilza*" of Ayurveda ('*Keraleeya chikilsavidhi*'), were a herbal preparation made out of *Njavara* grains is used to massage the body, after thorough application of oil all over the body. This procedure offers better physical consistency, better nerve strength and complexion improvement. '*Njavarakizhi*' is also used as a treatment for neurological disorders, rheumatism, arthritis and emaciation of limb (Dash, 1992). This treatment is restorative and nourishing.

Studies by Menon and Potty (1996) revealed that this is a unique cereal having high content of free amino acids. It is assumed that sulphur containing Methionine (a free amino acid) may be the constituent imparting medicinal property to this cultivar.

'*Shastika thilakam*' a preparation from the bran of *Njavara* is used to give glowing colour and brightness to the skin (Jayakumar, 2003).

Moos (2004) cited that *Njavara* rice with other herbals can be given as a special and safe food to patients undergoing various diseases. Porridge of *Njavara* grains in milk is traditionally given as a special food for invalids.

Siddiq (2004) reported that there is no rice cultivar, either singly or in groups which has been extensively used in any system of medicine as *Njavara* in the Ayurveda system of Kerala.

New born babies are fed with a dish called 'angri' made of *Njavara* flour and dried powder of banana (MSSRF, 2005).

According to ancient Ayurvedic records, indigenous preparation using *Njavara* along with *'Kurunthotty' (Sida rhombifolia var.retusa)* rejuvenates the muscles and nerves. Local healers claim that *Njavara* could cure hemorrhoids and decoctions made out of the root of *Njavara* could solve urinary problems of children. *Njavara payasam,* a sweet dish made out of jaggery and *Njavara* increases mother's milk. *Njavara* rice increases semen and fertility in male and is recommended for childless couples. *Njavara* flakes powdered with roots of *Aswagandha* (*Withania somnifera*) and sugar increase vigour, body weight and act as an aphrodisiac. *Njavara* is recommended as a safe food to diabetic patients (Hali, 2006).

Balachandran (2007) reported that *Njavara* maturing in sixty days had medicinal property in redressing 'thridosha' the root cause of all body ailments. This also enriches the body elements to exclude toxic metabolites, to strengthen, regenerate and energize body and regulate blood pressure.

2.2 Other medicinal rices in Kerala

Kerala has an immense wealth of medicinal rice cultivars and almost half a dozen medicinal rice varieties still exist in Kerala including *Chennellu, Kunjinellu, Erumakkari, Karuthachembavu* and *Kavunginpoothala* (Kumary, 2004). *Chennellu* and *Kunjinellu* are cultivars indigenous to North Kerala and are used locally in treatment of diarrhoea and vomiting and also given to patients recovering from jaundice. *Erumakkari* and *Karuthachembavu* are the traditional rice cultivars indigenous to South Kerala. The gruel made by the poached grains of *Karuthachambavu* is used to treat nausea, vomiting and stomach pains. *Erumakkari* is used for treatment of cough. A wild species of rice named *Annoori* is used by Kani tribes for treatment of small pox (Sujith kumar, 1999). *Kavunginpoothala* popular in Palakkad district is given to diabetic patients to reduce discomfort (Kumary, 2007).

2.3 Conservation and documentation

Njavara is oftenly regarded as "magical rice" with very special health and therapeutical qualities. Many workers in rice sector consider it as a precious gift of God to our country. Since Ayurveda, is becoming more and more popular within and outside the country, there is more demand for *Njavara.*

Arunachalam (2000) suggested the participatory conservation to rescue and revitalize biodiversity in local communities especially in *Njavara.* Participatory conservation is essentially an approach bridging farmer's knowledge and formal (scientific) theory of conservation, sustainable use and benefit sharing.

Prabhakaran (2004) reported that the *Njavara* cultivars are facing extinction. Farmers are facing a great difficulty in getting pure *Njavara* seeds.

Agrobiodiversity Centre of M.S. Swaminathan Research Foundation at Kalpetta in Waynad district in Kerala had done yeomen service in the revival of *Njavara* cultivation in Kerala and had succeeded in gaining global attention towards this unique rice cultivar (Hali, 2006).

Many agricultural products including *Njavara* needs proper documentation according to Pordie (2006). Research will be needed for the validation of claims and for ensuring that the claims printed on marketing labels are rooted on scientific data.

Geographical indication could be an interesting tool to protect unique agricultural goods produced locally. Palakkad *Njavara* Rice Farmers Society had attempted to protect community rights on *Njavara* cultivar by registering it as a Geographical Indication (GOI, 2007).

2.4 Morphological characterization

Different morphotypes of *Njavara* such as *Njavara* golden (awnless and awned) and *Njavara* black (awnless and awned) are conserved in different localities in Kerala.

Elsy *et al.* (1992) studied the morpho-physiological characters of black glumed *Njavara* and revealed that it exhibited extra short growth duration of 69 days with a good source-sink relationship. It also exhibited a high photosynthetic ability.

Menon and Potty (1996) evaluated the qualitative variations existing between the black and golden yellow glumed *Njavara* types and reported that the black glumed type, though less productive, had better quality compared to yellow glumed type.

Reddy (2000) evaluated the general morphology of different *Njavara* ecotypes and found distinct variability in qualitative and quantitative traits among themselves and with the traditional variety, PTB-10.

Studies conducted by Sreejayan *et al.* (2003) suggested that *Njavara* was genetically diverse and was a composite of distinct morphotypes.

Twelve ecotypes of *Njavara* including yellow type and black type were reported by Kumary (2004).

High variability was reported by Kumar (2005) in *Njavara* genotypes for lemma and palea colour viz., brown furrow on straw background, gold furrow on straw background and black furrow/patches on straw background.

Based on the morphological features Thomas (2006) grouped *Njavara* as long yellow, short yellow, intermediate yellow and short black. He also stated that *Njavara* represented an ancient gene pool still remaining unadulterated.

2.5 Molecular characterization

Molecular markers are heritable differences in nucleotide sequences of DNA at corresponding position on homologous chromosomes of two different individuals, which follow a simple Mendelian pattern of inheritance. These differences are detected by employing techniques like RAPD (Random Amplified Polymorphic DNA), AFLP (Amplified Fragment Length Polymorphism), RFLP (Restriction Fragment Length Polymorphism), SNP (Single Nucleotide Polymorphism), STS (Sequenced Tagged Sites) etc.

2.5.1 RAPD markers

RAPD analysis, a PCR-based molecular marker technique, was developed independently by Welsh and McClelland (1990) and Williams *et al.* (1990). RAPD polymorphisms are detected by using oligonucleotides of various sequences as primers in a PCR reaction.

Williams *et al.* (1990) reported that RAPD could be used to construct genetic maps in a variety of species. Traits of interest to plant breeders, including disease resistance, had been successfully mapped in crop plants using RAPD markers (Rafalski *et al.*, 1991 and Hadrys *et al.*, 1992).

Waugh and Powell (1992) explained the basis and applications of RAPD markers, highlighting differences with RFLP markers. RFLPs can detect only allelic variants whereas, RAPDs can detect more loci, will give markers over the whole genome, are easier to use, require only 10-50 ng of crude DNA.

The RAPD technique had several advantages over isozyme and DNA markers, such as speed, low cost and the use of small amount of plant material (Huff *et al.*, 1993). In recent years, RAPD analysis had become a popular method for estimating genetic diversity and relatedness in plant populations, cultivars and germplasm accessions.

RAPDs were generally found to be more cost and time efficient for studies involving small sample sizes (Ragot and Hoisington, 1993). Newburg and Ford (1993) identified that the technique of RAPD for the analysis of variation at the DNA level offered advantages over other methods in speed, technical simplicity and the frequency of polymorphisms.

Yu *et al.* (1994) suggested that RAPD markers could be used to detect genetic variation among upland and lowland rice cultivars. Forty two random primers were used for the study. Polymorphism were obtained and dendrogram depicted the genetic distance of rice cultivars.

Sixty random primers were used to amplify DNAs of cultivars viz., Nipponbare (*japonica*) and Kasalath (*indica*) of F_2 population generated by a single cross between these parents. 102 RAPD markers mapped on all 12 chromosomes of rice were obtained. The loci of the RAPD markers were determined into the framework of RFLP linkage map (Monna, 1994).

Genetic relationships between 37 geographically diverse rice varieties were analysed using arbitrary oligonucleotide primers in the RAPD method. Sufficient polymorphism was detected to allow identification of individual varieties. Banding patterns confirmed the close relativeness of commercial Australian and USA lines (Ko *et al.*, 1994).

Chen (1995) carried out RAPD analysis for dihaploid (DH) plants which were cultured *in vitro* and then naturally doubled. Rice molecular linkage maps were constructed using 52 RAPD markers.

Polymorphism was highest between members of the *indica* and *japonica* subspecies. Within the *japonica* subspecies, RAPDs had been shown to be polymorphic between the tropical and temperate types (Mackill, 1995).

Virk *et al.* (1996) identified associations between various quantitative traits and RAPD molecular markers using multiple regression analysis on diverse Asian rice (*Oryza sativa* L.) germplasm. This had allowed the prediction, for other samples of germplasm, of performance for traits such as culm length and number, days to flowering, grain width, and panicle and leaf length using only RAPD marker data. Such predictive capability was possible because of the availability of extensive diversity held in gene banks, and could be used in the future to facilitate the exploitation of that biodiversity.

RAPD analysis was done by Ishii *et al.* (1996) to study the phylogenetic relationship in A-genome species of rice. Twenty nine accessions were selected and their total DNA was used as template for PCR with 14 decamor primers. Average number of total amplified bands for each accession using fourteen primers were noted and dendrogram showing genetic relationship was constructed.

Four rice varieties viz. IR 8, Iratom 24, Iratom 38, Nizersail and Binasail were assessed for polymorphism in their DNA using a PCR-based method. Cellular DNA was extracted by a modified version of the CTAB (hexadecyltrimethyl ammonium bromide) method and PCR was carried out using random primers. One (OPT_1) of the seven primers used, generated RAPD markers which were polymorphic at one or two loci between genotypes and parent variety.

The amplification products were up to 1.3 kb and the average number of products per primer was 4.6 (Hakim, 1996).

Martin *et al.* (1997) carried out RAPD analysis using 93 *Oryza* accessions (*O. meridionalis*, *O. glumaepatula*, *O. nivara* and *O. rufipogon*) in the Genetic Resource Centre at IRRI. The RAPD data proposed that the relationship between accessions designated as *O. nivara* and *O. rufipogon* is complex, although it appeared that some mis-identification had also occurred for these two taxa. The results indicated that RAPD technology could be used as a fast and accurate method to assist in the validation of the identification of wild *Oryza* species.

From the survey of 389 arbitrary primers done by Subudhi *et al.* (1997) in bulked segregant analysis, four RAPD markers were identified in which three viz., OPF_{182600}, OPB_{19750} and $OPAA_{7550}$, were linked to *tms*3 (t) in repulsion phase and one, $OPAC_{3640}$ was linked to *tms*3(t) in coupling phase. The *tms*3 (t) gene was flanked by OPF_{182600} and $OPAC_{3640}$ on one side and by $OPAA_{7550}$ and OPB_{19750} on the other side. All four markers were low-copy sequences and two of them (OPF_{182600} and $OPAC_{3640}$) detected polymorphism when the markers were used to probe the genomic blots.

RAPD analysis was carried out by Parsons *et al.* (1997) to assess the genetic variation between samples of *Oryza sativa* from 19 localities in Bangladesh and Bhutan .It was revealed that 94 reproducible marker bands were amplified with a set of 14 arbitrary sequences of which 47 (50 per cent) were polymorphic.

RAPD analysis with nine 10-mer random primers were done to analyse seven male sterile rice lines, 12 maintainer lines and 12 restorer lines. This study generated 118 fragments of which 60 detected polymorphisms (Zhuang *et al.*, 1998).

Genetic variation within and between eight natural populations of *Oryza rufipogon* from China and Brazil was investigated at DNA level by analysis of RAPD fragments, with 20 selected primers. In this study 95 discernible DNA fragments were produced and 78 (21%) were polymorphic, which indicated high level of genetic variation in natural population. Studies also revealed that Chinese population showed greater polymorphism than Brazil population (Song *et al.*, 1999).

Genetic relationship between seven *japonica,* two *indica* and one tropical *japonica* rice varieties were analysed using 10 arbitrary primers, of which 77.4% were polymorphic. The banding patterns confirmed that all the seven *japonica* types were closely related, with similarity indices of 50-85%. Two *indica* varieties were grouped separately (Nadarajan *et al.*, 1999).

Rana *et al.* (1999) assessed the diversity of Indian rice genotypes through RAPD analysis. Twelve rice genotypes including three Basmati *indica,* eight non-Basmati *indica,* and one *japonica* were taken for the study. Thirteen primers generated a total of 146 DNA fragments, of which 91 were polymorphic. The size of fragments ranged from 0.24 to 3.7 kb. The average polymorphic value between Basmati and non-Basmati genotypes was 0.232, whereas among Basmati genotypes was 0.144 and among non-Basmati was 0.212. The polymorphic value between Azucena (*japonica*) and *indica* genotypes was 0.254.

Genetic variability in Indian scented rice germplasm was studied by Raghunathachari *et al.* (2000). Ten arbitrary oligonucleotide primers, applied to 18 accessions, produced a total of 144 bands of which 95.1 per cent were polymorphic. The size range of amplified DNAs was between 0.5 kbp and 4 kbp. This study revealed that scented rice varieties under cultivation with similar norms are genetically quite different.

Genetic relationship between twenty-three red rice accessions and a rice cultivar, Bengal was analysed by Estominos *et al.* (2000) through DNA fingerprinting. Polymerase Chain Reaction (PCR) with decamer primers generated random amplified DNA fragments showing polymorphisms among the fragments. Preliminary cluster analysis revealed that the 24 accessions could be separated into five groups with Bengal as the only rice cultivar in Group 1.

Qian *et al.* (2001) employed RAPD profiling for studying genetic variation within and between five populations of *Oryza granulata* from two regions of China. 199 reproducible bands with 30.65 per cent polymorphism were obtained using 20 RAPD primers. RAPD analysis revealed a low level of genetic diversity in wild populations of *O. granulata.* The RAPD markers revealed 73.85 per cent of the total genetic diversity residing between two regions, whereas only 19.45 per cent and 6.70 per cent were present between populations within regions and within a population respectively.

Identification and classification of 48 aromatic rice genotypes were done by Choudhury *et al.* (2001) using 58 random decamer primers. Most of the primers (96.5 per cent) detected polymorphism among the genotypes. Out of 465 amplified bands, 314 were polymorphic. Cluster analysis based on Jaccard's similarity coefficient using UPGMA grouped all traditional tall, photosensitive, low yielding, long grained 'Basmati' aromatic lines together. Different clusters were seen for short grained aromatic cultivars with high level of average similarity among themselves.

Total DNA from two WA type CMS lines viz. Zhenshan 97A, Longtepu A and their maintainers Zhenshan 97 B and Longtepu B was extracted by CTAB method. One hundred primers were used for RAPD screening to distinguish CMS lines (A) and maintainer (B) plants at seedling stage (Masahiko, 2002). Results showed that under conditions of 37°C annealing temperature and 1.5 mM $MgCl_2$ conc., in Zhenshan 97A and Longtepu A there was a 1600 bp

fragment in product amplified by primer OPA12, which was absent in Zhenshan 97B and Longtepu B. DA type CMS line Xieqingzao A also produced this band and was absent in Xieqingzao B. The results indicated that 1600 bp derived from CMS mitochondrial DNA could be used to distinguish A and B plants at seedling stage.

Sandhu *et al.* (2002) reported the use of RAPD markers for the analysis of herbicide - resistant Brasilian rice lines produced via mutagenesis. Rice line, IAC 103 after mutagenesis was checked for herbicide resistance. Eighty operon technology primers were tested and 10 were selected for detailed study. Resistant and susceptible lines produced variation in banding patterns.

Studies were carried out by Mane *et al.* (2003) to validate molecular markers known to be associated with maximum root length. Seventy nine diverse lines and six checks were used for the analysis. Variation was observed for several root morphological traits viz., maximum root length, total root number, root volume etc. RAPD and microsatellite primers, which have been established as markers for maximum root length, were used for screening rice genotypes. Single-marker analysis showed significant association between the markers and maximum root length.

Ravi *et al.* (2003) reported the use of RAPD and SSR markers in rice genetic diversity analysis of 40 cultivated varieties and five wild relatives. The accessions were evaluated for polymorphism after amplification with 36 decamer primers and 38 SSR primer pairs. Out of 499 RAPD markers obtained, 90 percentage showed polymorphism. The average Polymorphism Information Content (PIC) value was 0.578 and it ranged from a low of zero to a high of 0.890. The two marker systems contrasted most notably in pair-by-pair comparisons of relationships.

Cheng *et al.* (2004) reported the genetic diversity among 14 populations (bulk samples) and within two populations (individual samples) of wild species of rice, *Oryza granulata* using RAPD and ISSR amplification markers. The percentage of polymorphic bands was 59 per cent for RAPD and 64 per cent for ISSR, among the 14 populations.

RAPD analysis for the assessment of intra-collection diversity was carried out by Thomas *et al.* (2004). Genomic DNA isolated from five individuals each from 28 collections were subjected to analysis using a selected set of nine primers. In the 15 collections, the individuals analysed yielded monotypic profiles with respect to the primers examined, whereas more than one electrophoretically distinct genotype (electromorph) were discernible in the remaining 13 collections. Genetic relationships between the electromorphs were examined along with six other rice varieties using 1106 AFLP markers, generated by 12 primer combinations. The results of cluster analysis revealed 80 per cent similarity.

Saker *et al.* (2005) reported that genetic variability and relationships among seven Egyptian rice genotypes namely Giza 178, Giza 177, Giza 175, Giza 171, Giza 172, Sakha 102 and Sakha 101 were established by eight RAPD primers, six SSR and eight AFLP primer combinations. The level of polymorphism revealed by RAPD, SSR and AFLP were 72.2, 90 and 67.9% respectively. The highest genetic relationship as revealed by combined RAPD, SSR and AFLP was detected between Giza 175 and Giza 177 (83.4%), while the lowest was found between Giza 178 and Sakha 101 (61.5%). Giza 171 and Giza 172 were clustered together and genotypes Giza 175 and Giza 177 in the same cluster after cluster analysis.

Seven *Njavara* (N_1-N_7) genotypes along with two check varieties viz., PTB-10 and Karavella were subjected to RAPD analysis by Kumar (2005). Thirty eight decamer primers under OPE and OPP series were screened for amplification, out of which 11 primers that gave good amplification were

selected. Amplification with primers OPE 6, OPP 6 and OPP 11 exhibited unique bands for *Njavara* genotypes. The dendrogram data revealed that the *Njavara* genotypes were grouped in one cluster whereas check variety in another.

Genetic diversity of 38 aromatic rice lines and 2 non-aromatic lines was assessed using RAPD technique (Ghose *et al.*, 2005). Five RAPD primers, amplified 44 DNA bands ranging from 500 bp to 3.5 kb in size. Of the 44 DNA bands, 41 were polymorphic. The analysis indicated considerable amount of genetic diversity within the genotypes.

Brandolin *et al.* (2006) carried out the chemometrical characterization of four Italian rice varieties based on genetic and chemical analyses. It was found that genomic data gave the best discrimination based on varieties, indicating that RAPD assays could be useful in discriminating among closely related species.

Wu *et al.* (2006) extracted genomic DNA from 29 samples of rice seed, including jasmine rice 'KDML 105', 'KDML 105' derived varieties, non aromatic Thailand rice, and *japonica* rice and carried out RAPD analysis. The degree of band sharing revealed that RD15, CNTLR85033 and CNT 87040 were found to be closest to 'KDML 105'. Four RAPD fragments clearly distinguished jasmine rice from others.

To accumulate a genetic understanding to establish a methodology for non-farm conservation, RAPD markers were used by Fukuoka (2006). Eighteen RAPD markers detected 38 genotypes among 320 aromatic rice samples grown at 23 sites of farmers' fields.

RAPD assay and related other fingerprinting techniques have been employed to detect the genotoxin - induced DNA damage and mutations (Liu *et al.*, 2007). 12 RAPD primers of 50-70% GC content were found to produce unique polymorphic band patterns and a total of 180 bands of 179-3056 bp were

obtained for control root tips of rice seedlings. Results produced by these primers indicated that changes in RAPD profiles of root tips after Cadmium treatment including modifications in band intensity and gain or loss of bands in comparison with control.

2.5.2 Other molecular markers

Other than RAPD markers, a number of DNA based marker techniques have been developed since the early '80's. Such common markers are RFLPs, AFLPs and SSRs, also known as micro satellites. Single Nucleotide Polymorphisms (SNPs) are the newest type (Lalitha, 2005).

Wang and Tanksley (1989) analyzed 70 varieties of rice, representing the breadth of the species *Oryza sativa*, by using 10 rice RFLP markers. Polymorphism was detected for all probes, and 58 of 70 varieties tested could be uniquely distinguished from one another by combining all probe-enzyme combinations. Within-population variation, usually in the form of homozygous variant alleles, was found for 26% of the rice varieties. Based on genetic distance calculations, the ratio of the genetic variation within rice varieties was estimated to be around 12 to one. An RFLP based dendrogram was constructed depicting genetic distances among these rice varieties.

Microsatellites (Akkaya *et al.*, 1992) and AFLP (Zabeau and Vos, 1993) markers have been proposed for gene mapping in species where RFLPs are limited. Studies in rice indicated that micro satellite markers are highly polymorphic (Wu and Tanksley, 1993; Zhaw and Kochert, 1993; Yang *et al.*, 1994). RFLP maps developed at Cornell (Causse *et al.*, 1994) and in Japan (Kurata *et al.*, 1994) continued to form the basis of gene mapping research in rice.

Transfer of useful genes from wild relatives of crop plants has relied upon successful conventional crossing or the availability of the cloned gene.

Co-bombardment of rice callus with total genomic DNA from wild rice (*Zizania palustris*) and a plasmid containing a gene confirming hygromycin resistance allowed recovery under selection of transgenic plants with grain characteristics from wild rice. AFLP analysis suggested that a significant amount of DNA from *Zizania* was introduced by this procedure. One plant had 16 of a possible 122 *Zizania* specific AFLP markers detected with the primers used (Abedinia *et al.*, 2000).

The genetic diversity of 26 accessions of weedy rice and six Uruguayan cultivars were analysed using AFLP (Faderici, 2001). Abundant polymorphisms were found among samples tested. Using different methods of analysis three groups of samples were revealed. A relationship was found between three groups and morphological traits. The results suggested that weedy rice adopted either to the natural environment or to cultivation. The AFLP technique is very effective for assessing genetic diversity within weedy rice and will be very useful for finger printing of local cultivars of rice.

Ishii *et al.* (2001) studied the nuclear and chloroplast microsatellite variation in A-genome species of rice. Total DNA from 29 cultivars and 30 accessions of wild A-genome species was used as a template for PCR to detect 24 nuclear and 10 chloroplast micro satellite loci. Microsatellites amplified clearly in all 59 accessions, with an average of 18.4 alleles per locus. The Polymorphism Information Content (PIC) value ranged from 0.85 to 0.94 with an average of 0.89. The magnitude of diversity was much greater for nuclear microsatellites than for chloroplast microsatellite. For both nuclear and chloroplast microsatellite, *O. sativa* showed the highest similarity values to *O. rufipogon*, and *O. glaberima* was most similar to *O. barthii*. These data supported previous evidence that these cultures originated from the corresponding wild ancestral species.

Thirty-eight rice cultivars of particular interest to U.S. breeding programmes and two wild species accessions (*O. rufipogon* and *O. nivara*) were

evaluated by means of 111 microsatellite markers distributed over the whole rice genome. A total of 753 alleles were detected, and the number of alleles per marker ranged from 1 to 17, with an average of 6.8. A positive correlation was found between the number of alleles per locus and the maximum number of repeats within a microsatellite marker (Junjian *et al.*, 2002). Compared to *indica* cultivars, the *japonica* group showed significantly higher genetic diversity on chromosomes 6 and 7, and considerably lower diversity on chromosome 2. All rice cultivars and lines could be uniquely distinguished, and the results of groups corresponded exactly to the *indica* and *japonica* subspecies, with *japonica* divided into temperate and tropical types.

A total of 2,417 new di-, tri- and tetra-nucleotide SSR markers, representing 2,243 unique loci had been developed and experimently validated for rice. The majority (92%) of primer pairs were developed in regions flanking perfect repeats greater than 24 bp in length (McCough *et al.*, 2002).

Zeng *et al.* (2004) analyzed the genetic diversity among rice genotypes with different adaptations to saline soils using SSR markers. Thirty-three rice genotypes were considered for the study. Plants were treated with sodium chloride and calcium chloride of 6:1 molar ratio (6.5 ds/m). A great amount of genetic diversity was identified using 25 micro satellite markers among the rice genotypes.

DNA polymorphism between two major *japonica* rice cultivars, Nipponbare and Koshihikari was identified by AFLP. Eighty four polymorphic AFLP markers were obtained by analysis with 360 combinations of primer pairs. Nucleotide sequences of 73 markers, 29 from Nipponbare and 44 from Koshihikari, were determined, and 46 AFLP markers could be assigned to rice chromosomes based on sequence homology to the rice genome sequence. Out of 46 primer pairs, 44 amplified single DNA fragments, six of which showed different sizes between Nipponbare and Koshihikari, yielding codominant SCAR

markers. DNA fragments amplified by 13 primer pairs showed polymorphism by CAPs, and polymorphism of those amplified by 13 other primer pairs were detected by PCR-RF-SSCP (PRS). The procedure of conversion of AFLP markers to the sequence specific markers used in this study enabled efficient sequence specific marker production for closely related cultivars (Shirasawa *et al.*, 2005).

Adreit (2006) developed nine new microsatellite markers for rice blast population studies. 18 markers were used in multiplex PCR to characterize six populations from different geographical origins. The average number of alleles per locus across populations ranged from 1.2 to 7 and the total number of alleles detected from 2 to 19. Based on this large range of polymorphism, this set of markers was expected to be useful for different kind of population studies at different geographical scales.

Sequence Tagged Microsatellite (STMs) markers were employed on 48 elite Indian rice genotypes comprising eight indigenous aromatic land races and 16 improved aromatic varieties of Basmati types and 24 non-aromatic rice varieties to undertake genetic relationship studies (Randhawa *et al.*, 2006). A total of 203 alleles were detected at the 63 STMS loci, 196 (96.65%) of which were polymorphic. The number of alleles per locus ranged from 1 to 11, with average of 3.11 polymorphic alleles per locus. The polymorphic information content values ranged from 0 to 0.989 with an average of 0.806. The indigenous aromatic land races of Basmati types showed low level of polymorphism compared with improved Basmati and non aromatic genotypes. Out of 128 alleles in the Basmati land races, 90 (70. 31%) were polymorphic, whereas in the improved Basmati, 154 out of 168 (91.67%) were polymorphic and in non aromatic genotypes 159 out of 177 (89.83%) alleles showed polymorphism. The STMs markers RM 166_{200}, RM 153_{200} and RM 110_{182} distinguished indigenous aromatic land races of Basmati types from improved Basmati.

Genetic relationship between *Njavara* genotypes were examined by Thomas *et al.* (2006) following unweighted pair group method with arithmetic mean (UPGMA) cluster analysis of 664 AFLP fragments generated by 11 primer combinations. High congruence was observed between AFLP and morphology-based grouping. Uniformity and stability of AFLP based classification was confirmed by tracking the inheritance of 260 AFLP fragments generated by five primer combinations across two generations in three offspring each. *Njavara* genotypes formed a distinct group when the genetic relationship between *Njavara* and 81 other traditional rice varieties cultivated in Kerala was assessed using six microsatellite loci and 14 morphological traits. The analysis revealed that six genetically distinct genotypes existed in *Njavara* types.

Giarrow *et al.* (2007) assessed the genetic diversity in Argentine rice cultivars with SSR markers. Sixty-nine accessions were surveyed with 26 SSR markers revealing the genomic relationship among cultivars. A total of 219 polymorphic bands were detected. Cluster analysis based on pair wise comparisons of cultivar genetic similarities resolved the *O. sativa* accessions into two major *O. sativa* groups, *indica* and *japonica*, and the *japonica* group into the subgroups, tropical and temperate. These clusters agreed with the pedigree information available on the accessions and almost all Argentina-released cultivars grouped within the *japonica* cluster.

2.5.3 Molecular cloning and sequencing

The first amino acid sequence of a mature 16-KD BBI (Bowman Birk Inhibitor) protein from rice was reported 19 years ago for the rice bran trypsin inhibitor (RBTI) by Tashiro *et al.* (1987).

Mundy and Chua (1988) characterized an ABA-responsive rice gene, *rab-21* (now called *rab-16*), that is induced by ABA and osmotic stress in vegetative tissues. Yamaguchi *et al.* (1989) cloned and sequenced four members

of the rice *rab* (*rab*-16A-D) gene family. Each gene contained a small intron. They all were transcriptionally active and encoded proteins of 15500-16800 Dd.

Reece *et al.* (1990) cloned and sequenced four genomic actin genes, *Act1*, *Act2*, *Act3* and *Act7* in rice. A sequence alignment between a 1.5 kb *Act1* cDNA clone (McElroy *et al.*, 1990) and the four genomic clones were made. The analysis identified three introns located in the same position in all four rice actin coding regions, and the presence of a 5'-noncoding exon, separated by an intron, from the first translated exon of the *Act1*. This was one of the few reported cases of a plant gene containing such a 5'-noncoding exon. Analysis of other rice cDNA clones showed that the rice actin gene family was composed of at least eight unique members. DNA sequence comparison of the coding regions showed that the rice actin genes were highly diverged from each other.

De Pater *et al.* (1990) characterized a rice gene that was highly expressed only in shoots of rice. By differential screening of a cDNA library of two week old rice seedlings, a shoot specific gene, *COS5*, was identified. The mRNA corresponding to this gene displayed an expression pattern similar to that of *rbcS* genes. The mRNA (800 bases) was light inducible and encoded by a single gene. The genomic clone (GOS5) was also isolated which contained two introns.

Kim *et al.* (1999) carried out the cloning and sequencing of molecular marker associated with cold sensitivity at seedling stage in rice. A molecular linkage map consisting of 85 RAPDs and 54 RFLP markers was constructed by using the F_2 population which covered the rice genome at intervals of 10.2 cM on average. In the QTL analysis, one QTL for cold sensitivity was detected on chromosome 5. The DNA marker associated with the QTL was confirmed to closely linked with the cold sensitivity of rice from the RAPD analysis for the cold tolerance of 21 rice cultivars. The marker was cloned from RAPD fragments amplified from genomic DNA of Dular, an *indica* variety. The complete

nucleotide sequence revealed that the putative open reading frame was 511 base pairs and encoded 169 amino acids. Alignment of amino acid and nucleotide sequence between the DNA markers and Database of Gene Bank revealed 79% at nucleotide level with the rice cDNA (C26347) and 60% at amino acid level with hypothetical protein of Maize. This result provided the basis for marker-aided selection for the improvement of cold tolerance in rice.

Sucrose-6^F-phosphate phosphohydrolase enzyme was purified to homogeneity from rice leaves and partially sequenced. The rice leaf enzyme was a dimmer with a native molecular mass of 100kDa and a subunit molecular mass of 50 kDa. Three peptides from cleavage of the purified rice with endoproteinase Lys-c showed similarity to the deduced amino acid sequences of three predicted open reading frames in the *Arabidopsis thaliana* genome and one in the genome of the Cyanobactor *Synechocystis* sp. Pec6803, as well as cDNA clones from Arabidopsis, maize and other species in the Gene Bank database of expressed sequence tags. The putative maize spp cDNA clone contained an ORF encoding a 420 amino acid polypeptide. A PSI-BLAST search of the Gene Bank database indicated that maize spp. is a member of the haloacid dehalogenase hydrolase / phosphatase super family (John *et al.*, 2000).

Two Bacterial Artificial Chromosome (BAC) libraries had been constructed from the rice cultivar Nipponbare. Over 1000 BAC end sequences had been generated from these libraries and at a current total of 28 Mbp, representing 6.5% of the total rice genome sequence (Presting *et al.*, 2000).

Qu *et al.* (2003) identified seven BBI (Bowman-Birk Inhibitor) genes from *japonica* rice (*Oryza sativa* subsp. *japonica* var. Teqing). All the genes identified were found in a single cluster on the southern end of the long arm of rice chromosome one. Four of the seven BBI genes had repetitive cysteine-rich domains, whereas one had a truncated domain with only one reactive site. Different members of the rice BBI gene family displayed different expression

patterns during rice seed germination, and wounding induced the expression of rice BBI transcripts. It was found that the mRNA of rice BBI genes was present in abundant amounts in scutellar epithelium and aleurone layer cells. Over expression of *RBBI 2-3* in transgenic rice plants resulted in resistance to the fungal pathogen *Pyricularia oryzae*, indicating that proteinase inhibitors conferred resistance against the fungal pathogen.

A cDNA encoding enzyme of PPP, *6*-phosphogluconate dehydrogenase (6PGDH), was isolated from rice by Huang *et al. (2003)* and designated as *Os6PGDH*. The *Os6PGDH* encoding protein is a cytosolic isoenzyme according to the absence of plastid transit peptide at the N-terminus. The full-length cDNA of 1751 bp encoded 480 amino acids and its putative protein sequence was 94%, 84% and 83% identical to maize, spinach and alfalfa 6PGDHs respectively. Comparison of the cloned mRNA sequence with that of the genomic sequence from the Rice Genome Project showed a simple genomic organization devoid of introns in the translated region of the gene. It was suggested that 6PGDH in plant may play an important role in cell division and salt response.

The DNA sequence of 106 BAC/PAC clones in the minimum tilling path (MTP) of the long arm of rice chromosome 11, between map positions 57.3 and 116.2 cM, had been assembled to phase 2 or PLN level. This region had been sequenced to 10x redundancy by the Indian Initiative for Rice Genome Sequencing (IIRGS). The region, excluding overlaps, had been predicted to contain 2,932 genes using different software (Singh *et al.*, 2004). A gene by gene BLAST N search of the NCBI wheat EST database of over 420000 cDNA sequences revealed that 1,143 of the predicted rice genes (38.9%) had significant homology to wheat ESTS.

Wang *et al.* (2004) reported the cloning and sequencing of 5187-770 to develop a dominant SCAR (sequence characterized amplified region). An *indica* TGMS rice mutant, OA15-1 was crossed with a fertile *indica* line "GuiSi-8" to

map the gene responsible for the TGMS. An RAPD marker, 5187-770 linked to the TGMS gene at a distance of 1.3 cM in coupling phase was identified and cloned. Homology search against rice genome DNA sequence database indicated that 5187-770 was located on the short arm of chromosome 3 and close to centromere as a single cope sequence. This SCAR marker could be used in marker assisted transfer of this gene to different genetic document.

The cDNA of 2 starch-branching enzymes (SBE) genes, *Sbe*1 and *Sbe* 2, encoding SBE I and SBE III were cloned using RT-PCR from a template cDNA library derived from the total mRNAs of *japonica* rice cv. Wuyunjing 7 immature seeds. The size of the cloned *sbe*l and *sbe*3 cDNAs, including the entire coding sequence, were 2490 and 2481 bp, respectively. There were only 4 base-pair differences observed in the *Sbe*1 cDNA and the reported *sbe*1 (Gene Bank Accesson Number D11082), resulting in changes of 2 deduced amino acids (Yang *et al.*, 2004).

Xiang *et al.* (2005) developed a pair of PCR primers based on the conserved sequences of cab gene family in various rice cultivars. A fragment of the cab gene (appr. 350 bp) was amplified from the first strand of rice cDNA and PCR product as probe to screen the cDNA library of rice. When the largest inserted cDNA was sequenced, a full length cDNA of cab was obtained. The *cab-n8* gene was 1128 bp long, contained an open reading frame encoding 244 amino acids and one stop codon from the 55^{th} to the 789^{th} position. The DNA and amino acid sequences of *cab-n8* and *cab-27* were highly homologous (97%).

Liu *et al.* (2005) studied the genetic and physical mapping of *Pi36(t)*, a novel rice blast resistance gene located on rice chromosome 8. Blast resistance in *indica* cultivar (cv.) QbI was inherited as a single dominant gene in two F_2 populations, F_2-1 and F_2-2, derived from crosses between the donor cultivar and two susceptible *japonica* cvs.viz., Aichi Asahi and Lijiomgreintuanheigu (LTH). To rapidly determine the chromosomal location of the resistance (R) gene

detected in QbI, RAPD analysis was performed in the F_2-1 population using bulked-segregant analysis combination with recesisive-class analysis. One of the three linked markers identified, BA 1126550 was cloned and sequenced. The R gene locus was roughly mapped on rice chromosome 8 by comparison of the BA 1126550 sequence with rice sequences in the database. This novel R gene is tentatively designated as *Pi 36(t)*.

Molecular characterization and differential expression of cytokinin-response type-A response regulators in rice were carried out by Jain *et al.* (2006). By analysis of the whole genome sequence of rice, ten genes encoding type-A response regulators based on their high sequence identity within the receives domain were identified. The exon-intron organization, intron-phasing as well as chromosome location of all the RT-PCR amplified rice response regulator (OSRR) genes had been analysed.

Olsen *et al.* (2006) reported that an intron 1 splice donor site mutation of the *Waxy* gene was responsible for the absence of amylose in glutinous rice varieties. *Waxy* DNA sequence analyses indicated that the splice donor mutation is prevalent in temperate *japonica* rice varieties, but rare or absent in tropical *japonica*, *indica*, *aus*, and aromatic varieties. Sequence analysis across a 500-kb genomic region centered on *Waxy* revealed patterns consistent with a selective sweep in the temperate *japonica* associated with the mutation. The size of the selective sweep (>250 kb) indicated very strong selection in this region, with an inferred selection coefficient that was higher than similar estimates from maize domestication genes or wild species. These findings demonstrated that selection pressures associated with crop domestication regimes could exceed by one to two orders of magnitude those observed for genes under even strong selection in natural systems.

Mahmood *et al.* (2007) reported the cloning and sequence analysis of germin - like protein gene 2 promoter from *Oryza sativa* L. ssp. *indica*. During the

study, about 1107 bp 5' region of *Os RGLP2* gene was amplified, cloned and sequenced. The sequence analysis by BLAST showed that this promoter sequence have five common regions (CRI-CRS) of different sizes, which were repeated at 3-6 other locations in 30 kb region in which this gene driven by its promoter was located. Analysis of these common regions located on *OsRGLP2* indicated presence of many elements including those for light responsiveness, dehydration, stresses etc. Analysis of the 30 kb germin/GLP clustered region by GENSCAN detected each gene to have a putative 40 bp promoter containing TATA box and Dof factor which turned out to be a part of CR2.

2.5.4 Bowman Birk Inhibitor Proteins in crops

The Bowman-Birk Inhibitor (BBI) family is a typical canonical serine proteinase inhibitor. They are found in the seeds of leguminous (dicots) and graminaceous (monocot) plants (Birk, 1985).The family is named after the workers who first isolated (Bowman, 1946) and characterized (Birk, 1963) a member of this family from soybean. Soybean BBI is perhaps the most studied member of this family and often referred to as 'classical BBI'.

Soybean diet rich in protease inhibitors (e.g. BBI) lowered breast tumor incidence in irradiated rats (Troll *et al.*, 1980). BBI proteins are important since they are highly stable to both cooking temperature and digestion (Yavelow *et.al.*, 1985).

The suppression of malignant transformation by BBI and other anticarcinogenic protease inhibitors was studied by St.Clair *et al.* (1991) and suggested that it could be due to the effects of BBI on the expression on c-*myc* and other genes or oncogenes involved in the initiation, promotion or progression of the malignant phenotype.

Kennedy (1993) reported that Bowman-Birk inhibitor (BBI), a soybean derived protease inhibitor suppressed tumorigenesis in several experimental

carcinogenesis model systems. He also demonstrated that BBI can maintain its cancer chemopreventive potency even when administered as a dietary supplement.

Bowman-Birk trypsin inhibitor consisting of 125 amino acid residues was studied by Song *et al.* (1999) in barley seeds (BBBI). Qu *et al.* (2003) had identified seven Bowman Birk trypsin inhibitor genes in *japonica* cultivar of rice through molecular cloning and functional analysis. Feng *et al.* (2005) reported the presence of Bowman-Birk protease Inhibitors in sunflower (SFTI-1), with 14 amino acid residues.

Materials and Methods

3. MATERIALS AND METHODS

The molecular studies on *Njavara* genotypes were carried out in the Department of Plant Breeding and Genetics and Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara during 2005-2007.

3.1 MATERIALS

Seven *Njavara* accessions viz., N1, N2, N3, N4, N5, N6 and N7 along with PTB-10 and Karavella as check varieties formed the material for this study. The details of these genotypes are given in the Table 1

Table 1. Details of the genotypes used in the study

Sl. No.	Genotypes	Particulars
1. *Njavara* types	N1	NIC R18383
2 "	N2	NIC18430
3 "	N3	IC203771
4 "	N4	IC203767
5 "	N5	Kottakal Yellow type
6 "	N6	Aluva type
7 "	N7	Thrissur type
8. Check variety	PTB-10	From RARS, Pattambi
9 "	Karavella	From RARS, Pattambi

3.2 METHODS

3.2.1 GRAIN COLOUR CHARACTERIZATION

The *Njavara* types were characterized for grain colour according to Standard Evaluation System (SES) for rice (IRRI ,1995) for the following parameters

3.2.1.1 *Lemma and palea colour**: At maturity stage, when the terminal spikelets were ripened, the colour of lemma and palea was classified into 12 classes as

Code	Colour description
0	Straw
1	Gold and / or gold furrows on straw
2	Brown spots on straw
3	Brown furrows on straw
4	Brown (tawny)
5	Reddish to light purple
6	Purple spots on straw
7	Purple furrows on straw
8	Purple
9	Black
10	White
11	Black furrows/patches on straw background*

* Lemma and palea colour described in SES for Rice by IRRI (1995) was not accounting all the colour/shades exhibited by *Njavara* genotypes. Hence a new class (11) was added as black furrows/patches on straw background.

3.2.1.2 *Seed coat (bran) colour*: At maturity stage, rice (dehulled grains) was classified according to the standard system as

Code	Colour
1	White
2	Light brown
3	Speckled brown
4	Brown
5	Red
6	Variable purple
7	Purple

3.2.2 MOLECULAR CHARACTERIZATION

RAPD analysis was adopted for molecular characterization.

3.2.2.1 Isolation of genomic DNA

Modified Dellaporta method (Kumar, 2005) was used to isolate genomic DNA .Tender leaves were taken from the selected plants using sterile blades. The leaf samples were collected on ice and then wiped with cotton soaked in 70 percent alcohol and immediately used for extraction.

Reagents used

For the isolation of genomic DNA from rice, the following stock solutions were used.

Extraction buffer

a. 100 mM Tris (pH 8.0)
b. 50 mM EDTA (pH 8.0)
c. 500mM NaCl
d. 10 mM β-Mercaptoethanol (added immediately before use)

20% SDS

5 M Potassium acetate (pH 5.5)

TE buffer

a. 10 mM Tris pH 8.0
b. 1 mM EDTA pH 8.0

Isopropanol

Ethanol (70% and 100%)

3.2.2.1.1 Procedure

Seven *Njavara* accessions were sown and seedlings were raised in separate plastic trays. Tender leaf sample (one gram) was collected on ice, cut into

pieces with a sterile blade and transferred to a mortar and ground into fine powder using liquid nitrogen. Powder was transferred to a 50ml centrifuge tube kept on ice containing 7ml extraction buffer and 20µl β-Mercaptoethanol.500µl of 20 per cent SDS was added to the tube, mixed well and incubated at 65^{0}c for 20minutes. To the above suspension, 2.5ml of 5M potassium acetate was added, shaken vigorously and left on ice for 20 minutes with periodic shaking. The mixture was then centrifuged at 12,000 rpm at 4^{0}c for 10min and supernatant was filtered using sterile muslin cloth into a clean 50ml centrifuge tube. The filtrate was mixed with 5ml ice cold isopropanol and was incubated at -20°c for 30 min. After the incubation the mixture was centrifuged at 12,000 rpm for 10 min. at 4^{0}c to pellet the DNA.The supernatant was discarded and the DNA in the pellet form was washed with 70 per cent 200µl alcohol by centrifuging at 5000 rpm for 5 min. and airdried. The pellet was then resuspended in 300µl TE buffer and transferred to a eppendorff tube.

3.2.2.2 Agarose gel electrophoresis of DNA samples

Materials and equipments used

The following were the materials and equipments used

Agarose

TAE buffer 50X

a.	Tris base	- 242 g
b.	Glacial acetic acid	- 57.1 ml
c.	EDTA (0.5 M)	- 100 ml (pH 8.0)

Made up with distilled water to 1 litre.

Gel loading dye 2X (100 ml)

a.	Glycerol	- 40 ml
b.	4X TAE buffer	- 50 ml
c.	Bromophenol	- 0.5 %

Ethidium bromide solution (0.1 %)

Electrophoresis unit, power supply unit, *casting tray and comb*

3.2.2.2.1 Procedure

Gel buffer (TAE 1X) was taken in a conical flask (100 ml for large gel and 30 ml for small). Agarose (1 % for DNA and 1.2 % for RAPD samples) was weighed, added to the flask, stirred and boiled with frequent stirring till the agarose dissolved completely. Ethidium bromide was added into the flask and it was allowed to cool to 65°C. The open end of the gel casting tray was sealed with cello tape and placed on a horizontal surface and the comb was placed properly on the tray. The dissolved agarose was poured gently into the tray. The gel was allowed to solidify for 30 minutes and then the comb was removed carefully. The gel was then placed in the electrophoresis unit (Genei) with the well side directed towards cathode. 1X TAE buffer was added to cover the gel with a few mm of buffer. 10 µl of DNA sample (15 µl in case of RAPD products) was pipetted out onto a parafilm and mixed well with 4 µl of loading dye. The samples were then loaded carefully into the well by using micropipette. Standard DNA molecular weight markers were also added in one well. The cathode and anode of the electrophoresis unit were then connected to the power supply and the gel was run at constant voltage (60 mA). The power supply was turned off when the loading dye moved to required distance (1 ½ to 2 hours).

3.2.2.3 Gel documentation

The gel was taken from electrophoresis unit and viewed under UV light in a UV transilluminator. The ethidium bromide stain intercalates between the nitrogen bases of DNA and fluoresces in orange colour under UV light. The image of the gel was monitored and stored in a gel documentation system (Alpha Imager-2000, Alpha Infotech, USA).

3.2.2.4 Quantification of DNA

The quantity of isolated DNA was evaluated using a spectrophotometer

(Spectronic R Genesys 5). 1 µl of the sample were diluted 2.5 times using sterile water and the optical density was determined by reading the absorbance at two specific wave length viz., 260 nm and 280 nm. The 260/280 ratio was then calculated to check the purity. Pure DNA gives the ratio 1.8. The quantity of DNA in the pure sample was calculated as per the equation,

The OD value at 260 nm = 1 is equivalent to 50 x 2.5 µg of DNA per ml of the sample under study. Therefore, OD at 260 x 50 x 2.5 gives the quantity of DNA in µg per ml.

3.2.2.5 RAPD analysis

RAPD is a technique in which a single short oligonucleotide primer, which binds to many different loci, is used to amplify random sequences from a template DNA. The number of amplified products in RAPD depends on the length of primer and the size of the target genome, and is based on the probability that a given DNA sequence (complementary to that of the primer) will occur in the genome on opposite strands of the DNA, in opposite orientation within a distance readily amplifiable by PCR. The variation in RAPD profile in the form of presence or absence of band results from variation in primer binding sites. The products can be easily separated by standard electrophoresis techniques and visualized by ultraviolet illumination of ethidium bromide stained gels. PCR amplification process involves repeated thermal cycles.

The procedure reported by Raghunathachari *et al.*, 2000 and slightly modified (Kumar, 2005) was used for amplification of rice DNA.

The thermal cycles included,

Step no.	Temperature (°C)	Duration (min)	Steps involved	No. of cycles
1	94	3	Initial denaturation	1
2	92	1	Denaturation	40
3	37	1	Annealing	
4	72	2	Extension	
5	72	5	Final extension	1

The reaction mixture (20µl) consisted the following:

1. 10X assay buffer - 2.0 µl
2. dNTPs Mix - 1.0 µl
3. Mgcl2 - 1.0 µl
4. Taq DNA polymerase - 0.3 µl (1 unit)
5. Primer - 1.5 µl (5-10 p moles)
6. Template DNA - 2.0 µl
7. Sterile Milli-Q water - 12.2µl

A master mix without the template DNA was prepared using the reaction mixture for the required number of reactions. From this master mix, 18 µl was pipetted into each PCR tube and 2 µl of the template DNA was added. The PCR tubes were loaded in the Thermal Cycler (PTC 200, MJ Research, USA) and the programme was run. The programme was completed in 3 hours 45 minutes. The amplified products may be electrophoresed on 1.2 per cent agarose gel. The gel was viewed under UV light and documented.

3.2.2.5.1 Screening of random primers for RAPD

A total of 25 decamer primers (Table 2) under Operon series were screened for amplification of genomic DNA extracted from *Njavara* samples, using the Thermal Cycler mentioned under RAPD. These included 22 primers under OPA, OPN and OPF series and three primers viz., OPP 6, OPP 11, OPE 6 already reported to be efficient in amplification of genomic DNA from *Njavara* (Kumar, 2005). From these, 10 primers that gave good amplification (3-6 bands) were selected and utilized for further characterization of nine genotypes. The total number of bands along with the number of polymorphic bands obtained in all nine genotypes with each of the ten primers tried were recorded.

3.2.2.5.2 Data Analysis

The pattern of DNA amplification for the 10 primers was scored as 1 or 0 by the presence or absence of bands respectively and the data was fed to the NTSYS PC 2.0 software package.

The DNA fingerprint data were used to construct dendrogram by employing Unweighted Pair Group Method of Arithmetic Averages (UPGMA) using NYSTS pc version 2.01 programme (Rohlf, 1998) using SAHN coefficient.

Table 2. List of primers used for screening

Sl No.	Primer code	Primer sequence
1	OPA-1	CAGGCCCTTC
2	OPA-2	TGCCGAGCTG
3	OPA-3	AGTCAGCCAC
4	OPA-4	AATCGGGCTG
5	OPA-5	AGGGGTCTTG
6	OPA-6	GGTCCCTGAC
7	OPA-7	GAAACGGGTG
8	OPA-8	GTGACGTAGG
9	OPA-9	GGGTAACGCC
10	OPA-10	GTGATCGCAG
11	OPN-1	CTCACGTTGG
12	OPN-4	GACCGACCCA
13	OPN-5	ACTGAACGCC
14	OPN-6	GAGACGCAC
15	OPN-9	TGCCGGCTTG
16	OPN-12	CACAGACACC
17	OPN-14	TCGTGCGGGT
18	OPN-16	AAGCGACCTG
19	OPN-18	GGTGAGGTCA
20	OPN-19	GTCCGTACTG
21	OPF-6	GGGAATTCGG
22	OPF-20	GGTCTAGAGG
23	OPP-6	GTGGGCTGAC
24	OPP-11	AACGCGTCGG
25	OPE-6	AAGACCCCTC

3.2.2.6 DNA Gel Extraction

Gel extraction method (spin protocol by Axygen Biosciences) was used to purify specific bands when there were more than one band in the agarose gel. DNA gel extraction employed optimized reagents in combination with a convenient spin column to purify DNA fragments from either TAE or TBE agarose gels . Each column bound up to 8µg of DNA. DNA fragments in a size range of 70bp upto 10Kb could be efficiently recovered.

Materials used

The following were the materials used for DNA gel extraction

Axyprep columns

Microfuge tube (2ml)

Microfuge tube (1.5 ml)

Buffer DE-A : Gel solubilization buffer stored at room temperature

Buffer DE-B : Binding buffer stored at room temperature

BufferW1 : Wash buffer stored at room temperature

Buffer W2 : Desalting buffer (100 per cent ethanol added before use)

(Concentrate)

Eluent : 2.5mM Tris HCl (pH8.5).Stored at room temperature.

Isopropanol

Heated water bath

3.2.2.6.1 Procedure

a) Agarose gel slice containing the DNA fragment of interest was excised with a clean, sharp scalpel under ultraviolet illumination. The excised gel slice was briefly placed on tissue paper to remove the residual buffer. Gel slice was transferred to a preweighed 1.5 ml microfuge tube. The weight of the gel was regarded as equivalent to the volume (e.g. 100mg of gel was equivalent to a 100µl volume)

b) Added a 3 times sample volume of Buffer DE-A.

c) Resuspended the gel in Buffer DE-A by vortexing and heated at 75^0c until the gel was completely dissolved (typically, 6-8 minutes). Heated at 40^0c when low-melt agarose gel was used. Intermittent vortexing (every 2-3 minutes) accelerated gel solubilization. Gel was completely dissolved for complete DNA fragment recovery.

d) Added 0.5 Buffer DE- A volume of Buffer DE-B and mixed. If the DNA fragment was less than 400 bp, supplemented further with a 1 × sample volume of isopropanol.

e) Placed an AxyPrep column into a 2 ml microfuge tube. The solubilized agarose was transferred from step 4 into the column and centrifuged at 12, 000 g for 1 minute.

f) Discarded the filtrate from the 2 ml microfuge tube. Returned the AxyPrep column to the 2 ml microfuge tube and added 500µl of Buffer W1. Centrifuged at 12,000 g for 30 sec.

g) Discarded the filtrate from the 2 ml microfuge tube. Returned the AxyPrep column to the 2 ml microfuge tube and added 700µl of Buffer W2. Centrifuged at 12,000 g for 30 sec.

h) Discarded the filtrate from 2 ml microfuge tube. The AxyPrep column was placed back into the 2 ml microfuge tube and centrifuged at 12, 000 g for 1 minute.

i) Transferred the AxyPrep column to a clean 1.5 ml microfuge tube (provided). To elute the DNA, 25-30 µl of Eluent was added to the centre of the membrane and allowed to stand for 1 minute at room temperature and then centrifuged at 12, 000 g for 1 minute.

3.2.2.7 Cloning of the eluted DNA

Molecular cloning refers to the procedure of isolating a defined sequence and obtaining multiple copies of it *in vivo*.Cloning is frequently employed to amplify DNA fragments containing genes,but it can be used to amplify any DNA

sequence such as promoters, non- coding sequences, chemically synthesized oligonucleotides and randomly fragmented DNA.

StrataClone PCR Cloning protocol was used for cloning the eluted DNA. The Competent cells were provided in a convenient single-tube transformation format.

Media and reagents used

LB Agar (per Liter)

a) 10 g of NaCl
a) 10 g of tryptone
c) 5 g of yeast extract
d) 20 g of agar

Added deionized water to a final volume of 1 litre. Adjusted the pH to 7.0 with 5 N NaOH .Then autoclaved and poured into petri dishes (~25 ml/100-mm plate).

LB–Ampicillin Agar (per liter)

1 liter of LB agar , autoclaved and cooled to 55°C was added to 10 ml of 10-mg/ml filter-sterilized ampicillin.Poured into petri dishes(~25 ml/100-mm plate)

2% X-Gal (per 10 ml)

0.2 g of 5-bromo-4-chloro-3-inodlyl-β-Dgalactopyranoside(X-Gal) and 10 ml of dimethylformamide (DMF) .Stored at –20°C then spreaded 40 µl per LB-agar plate.

SOB Broth (per liter)

20.0 g of tryptone
5.0 g of yeast extract
0.5 g of NaCl
Added deionized H_2O to a final volume of 1 litre then autoclaved .Then added the following filter-sterilized supplements prior to use:

10.0 ml of 1 M $MgCl_2$
10.0 ml of 1 M $MgSO_4$

SOC Broth (per 100 ml)
This medium was prepared immediately before use.
2 ml of filter-sterilized 20% (w/v) glucose or 1 ml of filter-sterilized 2 M glucose
SOB medium (autoclaved) to a final volume of 100 ml

3.2.2.7.1 Ligating the Insert

A ligation procedure was employed whereby the amplified fragment was inserted into a vector. The method is descibed below .

a) Prepared the ligation reaction mixture by combining (in order) the following components:

i) 3 µl StrataClone™ Cloning Buffer
ii) 2 µl of PCR product (5–50 ng, typically a 1:10 dilution of a robust PCR reaction) or 2 µl of StrataClone™ Control Insert
iii) 1 µl StrataClone™ Vector Mix

b) Mixed gently by repeated pipetting, and then incubated the ligation reaction at room temperature for 5 minutes. When the incubation was completed, the reaction was placed on ice.

3.2.2.7.2 Transforming the competent cells

Transformation is the genetic alteration of a cell resulting from the uptake and expression of foreign genetic material (DNA). The protocol is mentioned below.

a) Thawed one tube of StrataClone Solo Pack competent cells on ice for each ligation reaction.
b) Added 1 µl of the cloning reaction mixture to the tube of thawed competent cells and mixed gently.
c) Incubated the transformation mixture on ice for 20 minutes. During the incubation period, SOC medium was pre-warmed to 42°C.
d) Heat-shock were given to transformation mixture at 42°C for 45 seconds.
e) Then incubated the transformation mixture on ice for 2 minutes.
g) Added 250 µl of pre-warmed SOC medium to the transformation reaction mixture and allowed the competent cells to recover for 1 hour at 37°C with agitation.
h) During the outgrowth period, LB–ampicillin plates were prepared for blue white color screening by spreading 40 µl of 2% X-gal on each plate.
i) Plated 5 µl and 100 µl of the transformation mixture on the LB– ampicillin–X-gal plates and incubated the plates overnight at 37°C.

3.2.2.7.3 Analyzing the transformants

White or light blue colonies were picked up for plasmid DNA analysis. Colonies harboring plasmids containing typical PCR product inserts would be white.

3.2.2.7.4 Confirming the presence of insert DNA

The number of colonies obtained and the cloning efficiency depend upon the size, amount, sequence, and purity of the PCR product used for ligation. For typical PCR products, the standard protocol produces hundreds of colonies for analysis. The presence of insert DNA was confirmed through RAPD analysis (with OPE 6 primer) using isolated recombinant plasmid DNA as given below

3.2.2.7.5 Isolation of recombinant plasmid

For the isolation of recombinant plasmid alkali lysis method was used.

Materials used

Solution I- Resuspension buffer

i) 50 mM glucose

ii) 25 mM Tris - Cl (pH 8.0)

iii) 10 mM EDTA (pH 8.0)

Solution I was prepared in batches of approximately 100 ml, autoclaved for 15 minutes at 10 lb/sq. in on liquid cycle and stored at 4°C.

Solution II- Lysis buffer

i) 0.2N NaOH

ii) 1 % SDS

Solution III- Neutralization buffer

i) 5 M Potassium acetate (60 ml)

ii) Glacial acetic acid (11.5 ml)

iii) Water (28.1ml)

3.2.2.7.5.1 Harvesting and lysis of bacteria

i) Harvesting

a) A single bacterial colony was transferrd into 3 ml of LB medium containing the appropriate antibiotic in a loosely capped 15 ml tube and incubated the culture overnight at 37°C with vigorous shaking.

b) Poured 1.5 ml of the culture into a microfuge tube and centrifuged at 12,000 G for 30 seconds at 4°C in a microfuge. Stored the remainder of the culture at 4°C.

c) Removed the medium by aspiration, leaving the bacterial pellet as dry as possible.

ii) Lysis by alkali

a) Resuspended the bacterial pellet (obtained in step 3) in 100µl of ice -cold solution I by vigorous vortexing.

b) Added 200µl of freshly prepared solution II. The tube was closed tightly, and the contents were mixed by inverting the tube rapidly five times. Made sure that the entire surface of the tube came in contact with solution II. Stored the tube on ice

c) Added 150µl of ice cold solution III. Closed the tube and vortexed it gently in an inverted position for 10 seconds to disperse solution III through the viscous bacterial lysate. Stored the tube on ice for 3-5 minutes.

d) Centrifuged at 12000 rpm for 5 minutes at 4°C in an microfuge and transferred the supernatant to a fresh tube.

e) Precipitated the double stranded DNA with 2 volumes of ethanol at room temperature mixed by vortexing, and allowed the mixture to stand for 2 minutes at room temperature.

f) Centrifuged at 2000G for 5 minutes at 4°C in a microfuge.

g) Removed the supernatant by gentle aspiration. Allowed the tube to stand in an inverted position on a paper towel to allow all of the fluid to drain away.

h) Rinsed the pellet of double stranded DNA with 1 ml of 70% ethanol at 4°C. Removed the supernatant as described in step 8 and allowed the pellet of nucleic acid to dry in the air for 10 minutes.

i) Redissolved the nucleic acids in 50µl of TE containing DNAase free pancreatic RNAase. Vortexed briefly. Stored the DNA at -20°C.

The cloned fragment were confirmed through RAPD analysis with OPE 6 primer. The molecular weight of cloned RAPD product was compared with original amplicon.

3.2.3 Sequence analysis of the clone

3.3.3.1 Sequencing

Sequencing is the determination of the primary structure (or primary sequence) of an unbranched biopolymer. Sequencing results in a symbolic linear depiction known as a sequence which succinctly summarizes much of the atomic-level structure of the sequenced molecule.DNA sequencing is the process of determining the nucleotide order of a given DNA fragment.

The insert were sequenced at DNA sequencing services of Genei, Bangalore using M13 Forward universal primers.

3.2.3.2 Sequence analysis using bioinformatics tools

Different bioinformatics tools were used to analyse the query sequence.

3.2.3.2.1 BLAST (Basic Local Alignment Search Tool)

BLAST is defined as a sequence algorithm optimized for speed and used to search sequence databases for optimum local alignments to a query.The commonly used BLAST programme for nucleotide analysis is BLASTn

(Nucleotide- Nucleotide Blast). It compares gene sequence to nucleotide sequence database and calculates the statistical significance of matches. It also infers functional and evolutionary relationship between sequences. The matching hits are selected based on their e value. E value is the number of hits that are expected to occur by chance. Low e value score is a good hit to a model whereas a higher score is a worse hit.

Protocol

a) Accessed the NCBI homepage from the website (www.ncbi.nlm.nih.gov)
b) Selected the BLASTn tool from it.
c) The sequence of the selected gene was copied and pasted
d) The other parameters were set for default values/settings
e) BLAST search was run
f) The format page was displayed which had the details of the conserved domains
g) The FORMAT button was selected
h) The results were displayed which contained the description of the query sequence and as a list of related hits based on e value. The hit with the lowest e value were selected.

3.2.3.2.2 ORF finder

To search for Open Reading Frames (ORF) in the DNA sequence ORF finder tool in National Centre for Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov) was used. ORF is a graphical analysis tool which finds all open reading frames of a selectable minimum size in a user's sequence or in a sequence already in the database. This tool identifies all open reading frames using the standard or alternative genetic codes. The deduced amino acid sequence could be saved in various formats and searched against the sequence database using the WWW. BLAST server.

Protocol

a) Accessed the NCBI homepage from the website (www.ncbi.nlm.nih.gov).

b) Selected the ORF finder from it.

c) Submitted the sequence in the FASTA format.

d) Clicked on 'ORF find'.

e) A graphical output indicating the location of each ORF was found.

3.2.3.2.3 GENSCAN

GENSCAN is a computer programme for gene identification. It is designed to predict complete gene structures. GENSCAN can identify introns, exons, promoter sites and polyA sites.

Protocol

a) Accessed the genscan web server at MIT from the site http://genes.mit.edu / GENSCAN.html.

b) Pasted the query nucleotide sequence in the search box.

c) Clicked "run genscan" button.

d) The genscan results for the query nucleotide sequence was obtained. The output contained the translated protein sequence and exon positions.

e) Saved the results.

3.2.3.2.4 Vector screening

VecScreen is a system for quickly identifying segments of a nucleic acid sequence in query sequence that may be of vector origin.

Protocol

a) Accessed the NCBI homepage from the website (www.ncbi.nlm.nih.gov).

b) Selected the VecScreen from it.

c) Submitted the sequence in the FASTA format.

d) Clicked on 'Run VecScreen'.

e) Distribution of Vector Matches on the query sequence were displayed.

3.2.3.2.5 Aligning cloned vector sequence with query sequence

3.2.3.2.5.1 CLUSTALW

CLUSTALW is a general purpose multiple sequence alignment program for DNA or proteins. It produces biologically meaningful multiple sequence alignments of divergent sequences. It calculates the best match for the selected sequences, and lines them up so that the identities, similarities and differences could be seen.

Protocol

a) Accessed the EMBL-EBI homepage from the website (www.ebi.ac.uk).
b) Selected the CLUSTALW from it.
c) Submitted the sequences in the FASTA format.
d) Clicked on 'Run'.
e) Sequence alignment were displayed and saved.

3.2.3.2.5.2 BLAST 2 SEQUENCES

This tool produces the alignment of two given sequences using BLAST engine for local alignment.

Protocol

a) Accessed the NCBI homepage from the website (www.ncbi.nlm.nih.gov).
b) Selected the BLAST 2 SEQUENCES from it.
c) Submitted the sequences in the FASTA format.
d) Clicked on 'Align'.
e) The aligned sequences were displayed.
f) Saved the results.

Results

4. RESULTS

The results of the experiments on 'Molecular documentation of *Njavara* types of rice for cultivar identification' are described in this chapter.

4.1 GRAIN COLOUR CHARACTERIZATION

Grain colour characterization of seven *Njavara* genotypes along with two check varieties were recorded and are presented Table 3.

Table.3. Variability of lemma, palea and seed coat colour in *Njavara* types

Sl. No	Genotypes used for the study	Lemma and palea colour	Seed coat colour (Bran)
1.	N1	Black including black patches on straw and black furrows on straw background	Red
2.	N2	Black including black patches on straw and black furrows on straw background	Light brown
3.	N3	Gold furrows on straw	Red
4.	N4	Brown furrows on straw	Brown
5.	N5	Gold furrows on straw	Red
6.	N6	Black (light shade)	Brown
7.	N7	Brown furrows on straw	Red
8.	PTB-10	Brown furrows on straw	Red
9.	Karavella	Brown furrows on straw	Red

High variability was noticed among *Njavara* genotypes with respect to lemma and palea colour.Two major classes based on glume colour variability could be distinguished as yellow (straw colour) and black *Njavara* types. Detailed characterization revealed that yellow glume colour could be further grouped into gold furrows on straw (N3 and N5) and brown furrows on straw (N4 and N7). Three genotypes viz., N1, N2 and N6 exhibited a lemma and palea colour dominated by black. N1 and N2 exhibited variations in black colour for lemma and palea as pure black, black furrows/ patches on straw background whereas N6 exhibited light shade of black.

Seed coat colour (bran) also varied in these genotypes. The genotypes N1, N3, N5 and N7 exhibited red seed coat colour. Light brown seed coat colour was noticed in N2. The genotype N4 and N6 showed brown seed coat colour. The seed coat colour for PTB-10 and Karavella was red.

The morphotypes in *Njavara* are presented in Plates 1-2 along with check varieties (Plate 3) used for the study. The results of this study confirmed the existence of greater variability in *Njavara* at different parts of the state.

4.2 MOLECULAR CHARACTERIZATION

4.2.1 RAPD Marker Analysis

The results of the experiments conducted for the extraction of genomic DNA of seven *Njavara* genotypes and two check varieties and the evaluation using RAPD assay are presented below.

Plate 1. Njavara- Black types

Plate 2. Njavara- Yellow types

Plate 3. Check varieties used in the study

4.2.1.1 Genomic DNA isolation

The genomic DNA was isolated from tender leaf tissues of seven *Njavara* genotypes and two check varieties using modified Dellaporta method. The DNA extracted when tested on 0.7 per cent agarose gel electrophoresis revealed the quality and yield of DNA.

The quantity and quality of isolated DNA are presented in the Table 4. The ratio of absorbance at 260/280 was between 1.84 and 1.96. The quantity of the DNA extracted varied between 925.1 to 1594.0 ng/µl of leaf tissue. High recovery was obtained for N7 (1594 ng/µl of DNA extract) whereas the quality was good for the genotypes N2, N4, N6 and Karavella. The electrophoretic profile showed good clear bands. The gel picture showing genomic DNA is presented in the Plate 4.

Table 4. Quantity and quality of genomic DNA isolated from rice genotypes

Sl. No.	Genotypes	260/280 ratio	Quantity (ng/µl)	Quality
1.	N1	1.91	925.1	Average
2.	N2	1.85	985.7	Good
3.	N3	1.95	1248.5	Average
4.	N4	1.86	1164.6	Good
5.	N5	1.92	1447.9	Average
6.	N6	1.84	1268.4	Good
7.	N7	1.96	1594.0	Average
8.	PTB-10	1.91	1000.7	Average
9.	Karavella	1.86	949.6	Good

4.2.1.2 Screening of random primers

DNA sample from a representative *Njavara* genotype were amplified with twenty two random decamer primers {10 from Operon Primer A (OPA) series, 10 from Operon Primer N (OPN) series and two from Operon Primer F (OPF) series} for screening the primers. The results of screening are given in Tables 5 and 6.

4.2.1.2.1 OPA series

The results of screening of 10 primers of OPA series are presented in Table 5 and Plate 5. The number of bands ranged between zero and eight. Out of the ten primers screened, five viz., OPA 1, 4, 6, 7 and 9 gave good amplification and the number of amplification products were also more for these primers with OPA 1 and 4 producing eight bands whereas OPA 6 producing seven bands. Five bands were produced by OPA 7 and OPA 9. OPA3 gave an amplification of average quality and others (OPA 2, 5, 8, 10) gave poor quality or no amplification at all. Repeated tests gave similar results. OPA 1, 4, 6, 7 and 9 were selected for further studies based on distinct banding pattern with good quality amplification, reproducibility and stability of expression.

4.2.1.2.2 OPN and OPF series

Results of the screening with 10 primers in OPN and OPF series are given in Table 6. Number of bands produced by the primers varied from zero to six (Plate 6). Three primers viz., OPN 4, 16 and 18 gave good amplification with five, six and seven bands respectively. OPN 6 and 12 gave amplification of average quality. Seven primers viz., OPN 1, 5, 9 and 14 did not give any amplification. Similarly primers under OPF series (OPF 6 and 20) gave no amplification .Subsequent trials gave same results. OPN 6 and 18 were selected for further study due to better stability, distinct banding pattern with good quality amplification and reproducibility.

OPA 1, 4, 6, 7, 9, OPN 6, OPN 18 found to be better performing in amplification of *Njavara* DNA and OPP 6, OPP 11, OPE 6 earlier reported to be good in amplification were selected for further studies. Details of these primers are given in Table 7

Plate 4. Isolated Genomic DNA

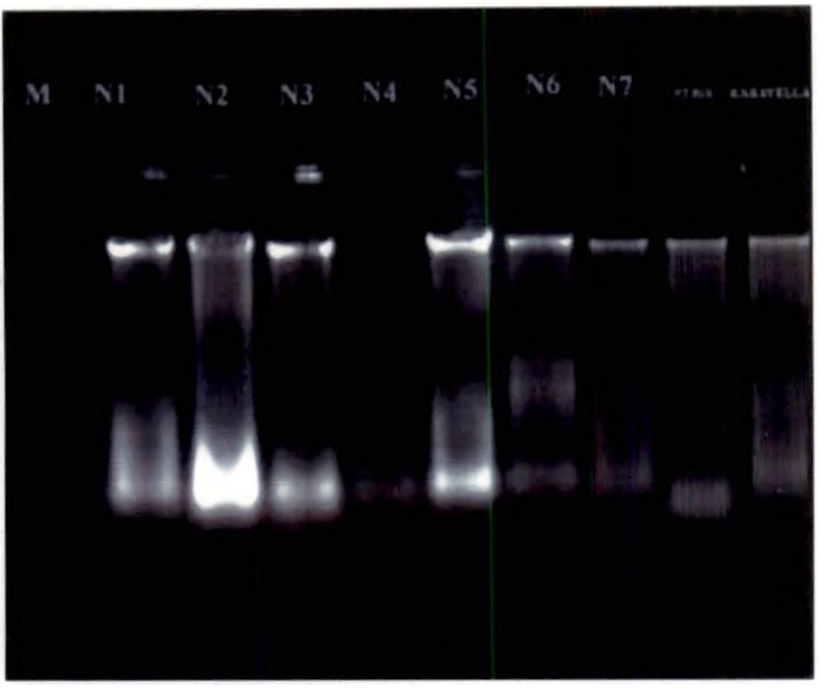


Plate 5. Screening with OPA primers

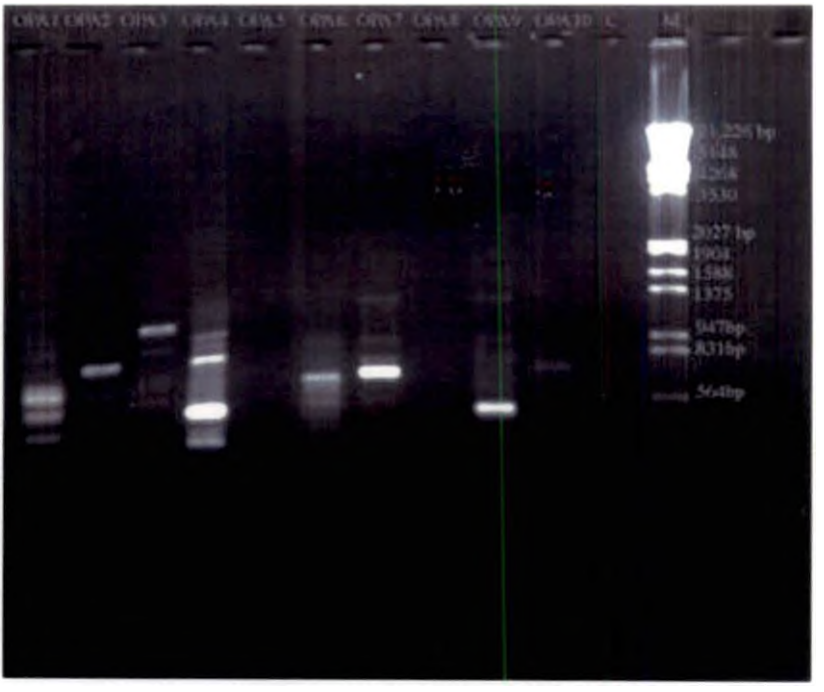


Plate 6. Screening of rice (N7) genomic DNA with OPN & OPF primers

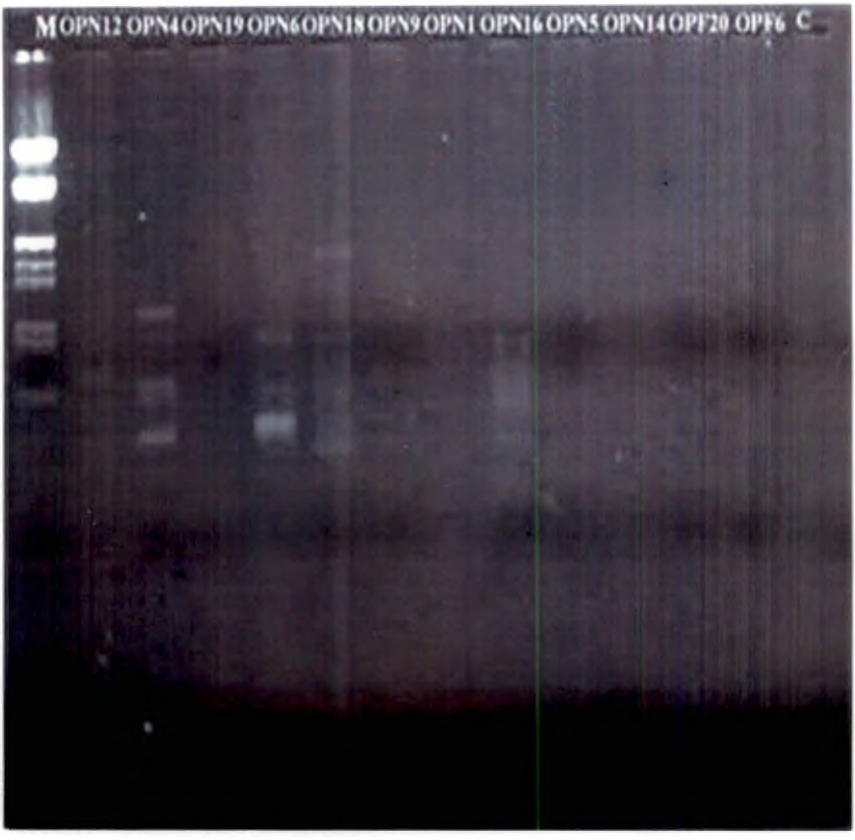


Table 5. Amplification pattern of rice (N7) genomic DNA with different decamer primers under OPA series

Sl. No.	Primer code	No. of bands	Quality of amplification
1	OPA 1	8	Good
2	OPA 2	2	Poor
3	OPA 3	4	Good
4	OPA 4	8	Good
5	OPA 5	0	Poor
6	OPA 6	7	Good
7	OPA 7	5	Good
8	OPA 8	0	Poor
9	OPA 9	5	Good
10	OPA 10	1	Poor

Table 6. Amplification pattern of rice (N7) genomic DNA with different decamer primers under OPN & OPF series

Sl. No.	Primer code	No. of bands	Quality of amplification
1.	OPN 1	0	Poor
2.	OPN 4	5	Good
3.	OPN 5	0	Poor
4.	OPN 6	4	Good
5.	OPN 9	0	Poor
6.	OPN 12	4	Good
7.	OPN 14	0	Poor
8.	OPN 16	7	Good
9.	OPN 18	6	Good
10.	OPN 19	0	Poor
11.	OPF 6	0	Poor
12.	OPF 20	0	Poor

Table 7. List of selected decamer primers used for RAPD analysis

Sl. No.	Primer code
1	OPA 1
2	OPA 4
3	OPA 6
4	OPA 7
5	OPA 9
6	OPN 6
7	OPN 18
8	OPP 6*
9	OPP 11*
10	OPE 6*

* Primers already reported to be good in amplifying *Njavara* genotypes (Kumar 2005)

4.2.1.3 Screening of *Njavara* genotypes with selected primers

The genomic DNA isolated from seven *Njavara* genotypes and two check varieties were subjected to RAPD analysis using 10 selected primers in order to assess the intra genotype heterogenity. The 10 random primers used for amplification gave 364 scorable amplified products of which 352 (96.7 per cent) were polymorphic and 12 bands were monomorphic (3.2 per cent).

Amplification with primer OPA1 resulted in polymorphic amplification in all seven *Njavara* genotypes and two check varieties .Molecular weight of bands ranged between 0.58 kb and 3.53 kb (Table 8, Plate 7). Amplification pattern of N2 and N5 gave maximum of seventeen bands. N4 and N7 gave minimum number of amplified products.

The primer OPA 4 gave bands with molecular weight ranging between 0.53 kb and 0.96 kb (Table 9, Plate 8). The genotypes N3, N4, N5, N6, N7, PTB-10 and Karavella gave uniform electrophoretic pattern. N1and N2 produced minimum number of amplified products.

Primer OPA 6 gave a RAPD profile without any monomorphic band (Plate 9). Amplification with this primer resulted in polymorphic amplification in all seven *Njavara* genotypes and two check varieties (Table 10). The bands ranged between 2.74 kb and 5.19 kb. The genotypes N6 and Karavella gave minimum number of amplified products.

All the seven *Njavara* genotypes and two check varieties gave one monomorphic band for primer OPA7 (Plate 10, Table 11). Molecular weight of the monomorphic band for primer was found to be 0.43 kb.

Plate 7. RAPD profile for Njavara genotypes with primer OPA 1

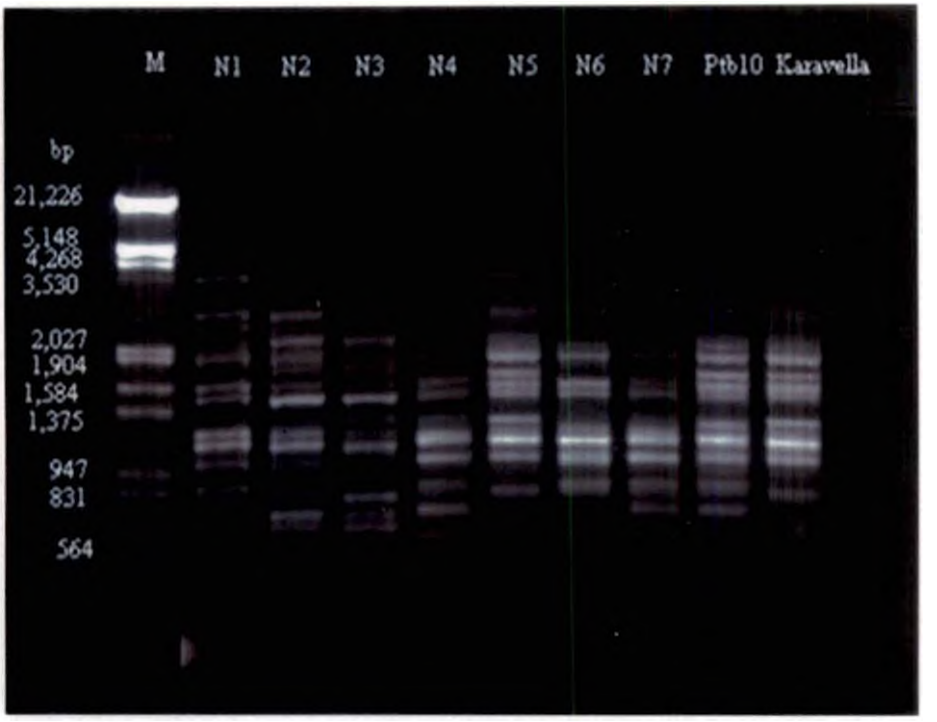


Plate 8. RAPD profile for Njavara genotypes with primer OPA 4

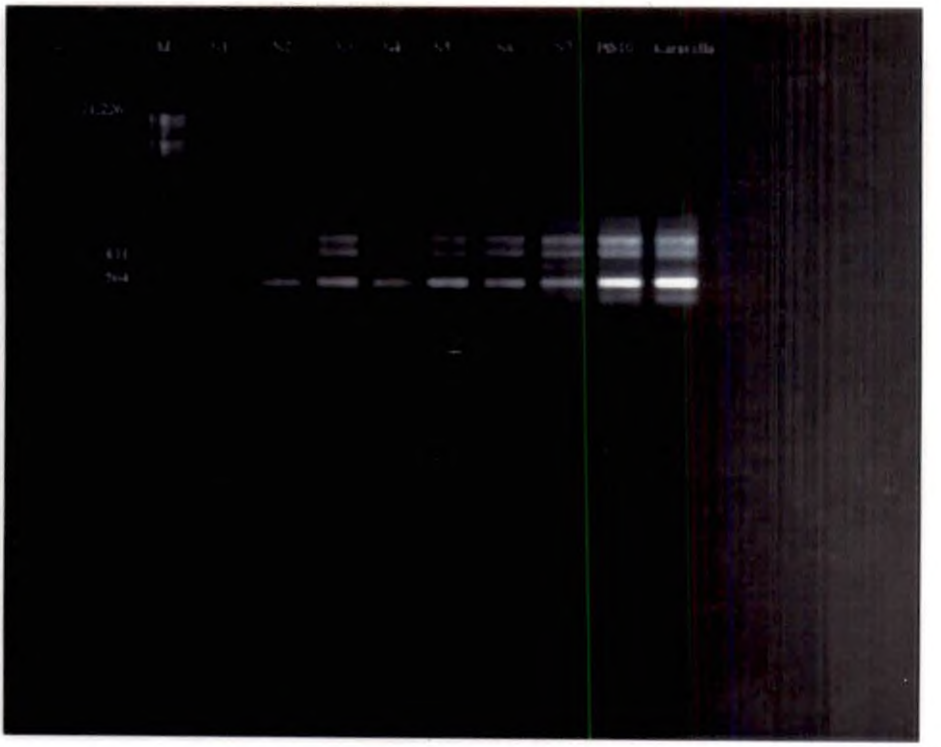


Plate 9. RAPD profile for Njavara genotypes with primer OPA 6

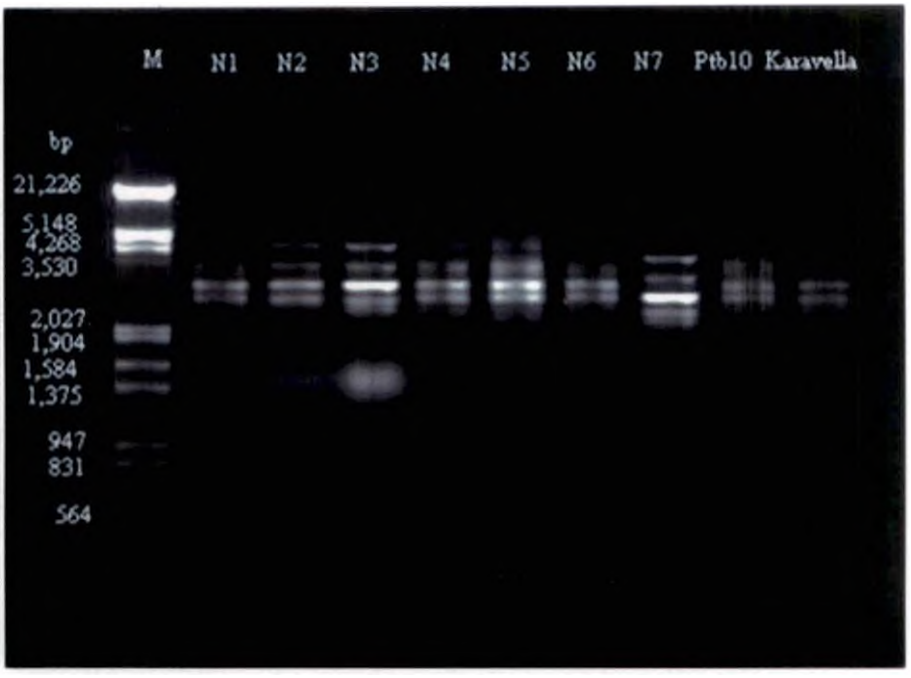


Plate 10. RAPD profile for Njavara genotypes with primer OPA 7

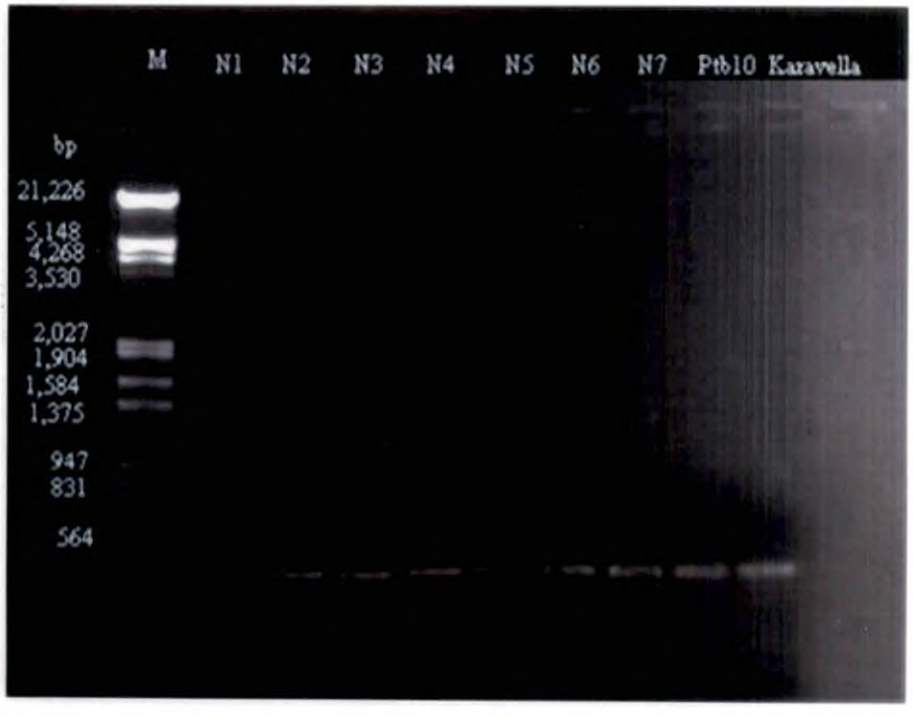


Table 8. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPA 1

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	10	4	6
2.	N2	17	4	13
3.	N3	11	4	7
4.	N4	10	4	6
5.	N5	17	4	13
6.	N6	12	4	8
7.	N7	9	4	5
8.	PTB-10	12	4	8
9.	Karavella	12	4	8
Total		110	36	74

Table 9. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPA 4

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	2	1	1
2.	N2	2	1	1
3.	N3	4	1	3
4.	N4	3	1	2
5.	N5	5	1	4
6.	N6	5	1	4
7.	N7	5	1	4
8.	PTB-10	4	1	3
9.	Karavella	4	1	3
Total		34	9	25

Table 10 RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPA 6

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	6	2	4
2.	N2	6	2	4
3.	N3	5	2	3
4.	N4	5	2	3
5.	N5	6	2	4
6.	N6	4	2	2
7.	N7	5	2	3
8.	PTB-10	4	2	2
9.	Karavella	3	2	1
Total		44	18	26

Table 11. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPA 7

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands
1.	N1	1	1
2.	N2	1	1
3.	N3	1	1
4.	N4	1	1
5.	N5	1	1
6.	N6	1	1
7.	N7	1	1
8.	PTB-10	1	1
9.	Karavella	1	1
Total		9	9

Amplification with primer OPA 9 produced one monomorphic band (Table 12) in all the seven *Njavara* genotypes and two check varieties. The molecular weight of bands ranged between 0.297kb and 1.9 kb (Plate 11). N1 gave maximum number of ten bands. The genotypes N4, N5 and N7 gave uniform electrophoretic pattern whereas N3 gave minimum number of amplified products. Polymorphic bands were obtained for N1, N2, N4, N5, N6 and N7.

All the seven *Njavara* genotypes and two check varieties gave one monomorphic band for primer OPN 6 (Plate 12). The molecular weight of bands ranged between 0.53 kb and 1.59 kb (Table 13). Amplification pattern of N7 gave maximum of six bands. PTB-10 exhibited a unique medium molecular weight band (1.59 kb).

Primer OPN 18 gave an RAPD profile with one monomorphic band (1.9kb).The genotypes N2, N3 and N4 showed uniform electrophoretic banding pattern. The molecular weight of bands ranged between 0.44 kb and 1.9 kb (Table 14). Amplification pattern of N6 gave maximum of six bands. PTB-10 and Karavella exhibited unique bands of molecular weight of 0.81kb and 1.2 kb (Plate 13) resp.

Amplification with primer OPP 6 produced banding pattern without any monomorphic bands (Plate 14). The molecular weight of bands ranged between 0.715 kb and 1.5 kb (Table 15). The genotypes N2 and N3 showed uniform electrophoretic banding pattern, with ten bands.

Electromorphic banding pattern for primer OPP 11 gave high amplification for all genotypes except N3 (Plate 15). Two monomorphic bands (0.94 kb and 2.1 kb) were shown with seven *Njavara* genotypes and two check varieties. N5 exhibited maximum number of thirteen bands (Table 16). N6 and N7 exhibited uniform banding pattern.

Amplification with primer OPE 6 resulted in unique amplification bands of molecular weights of 1.375 kb, 1.29 kb and 0.44 kb in all seven *Njavara* genotypes. These bands were absent in check varieties selected (Plate 16). The molecular weight of bands ranged between 0.44 kb and 1.375 kb (Table 17).

When the amplification pattern was compared, primer OPA 1and OPA11 gave maximum number of amplified products viz., 17 and 13 for the genotypes under study. OPN 6 exhibited one unique band for PTB-10.The number of amplicons shared by *Njavara* types were three with OPE 6 primer(Table 18).

Plate 11. RAPD profile for Njavara genotypes with primer OPA 9

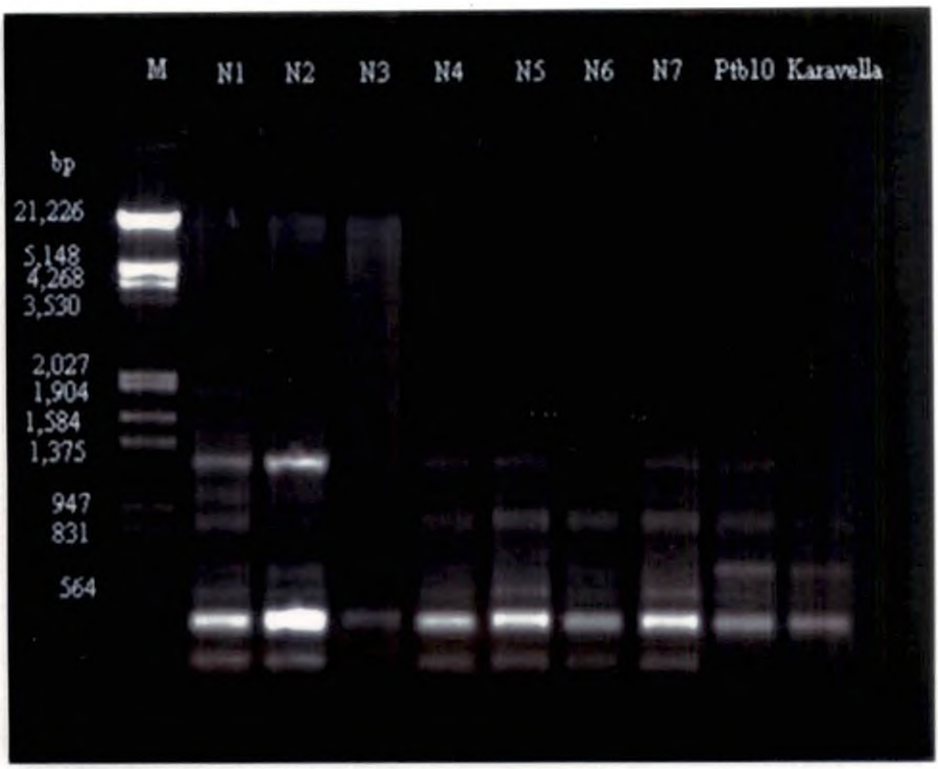


Plate 12. RAPD profile for Njavara genotypes with primer OPN 6

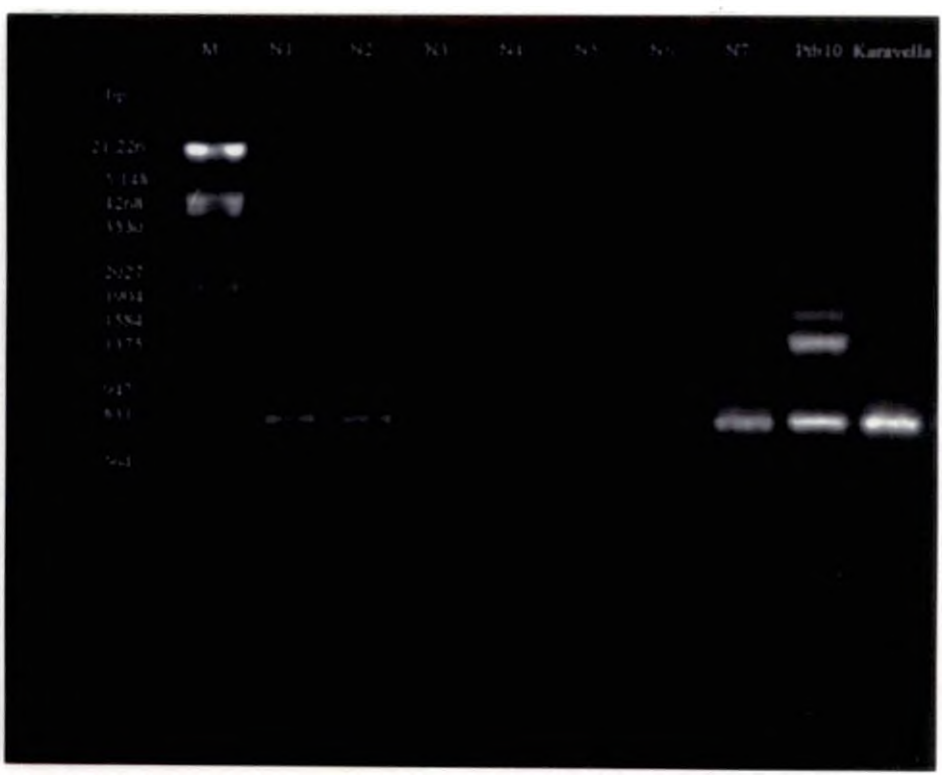


Plate 13. RAPD profile for Njavara genotypes with primer OPN 18

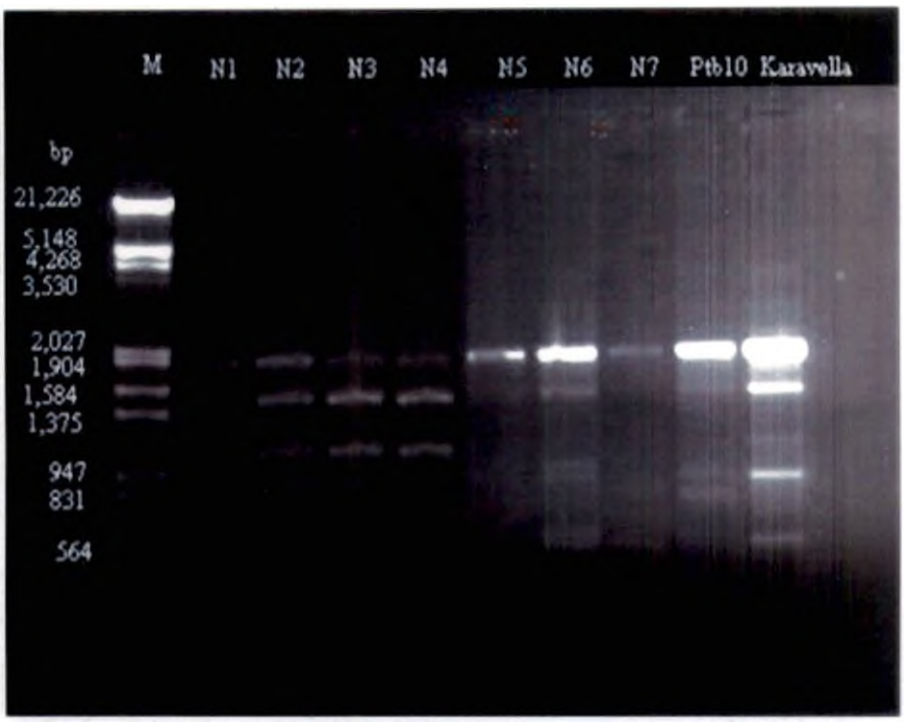


Plate 14. RAPD profile for Njavara genotypes with primer OPP 6

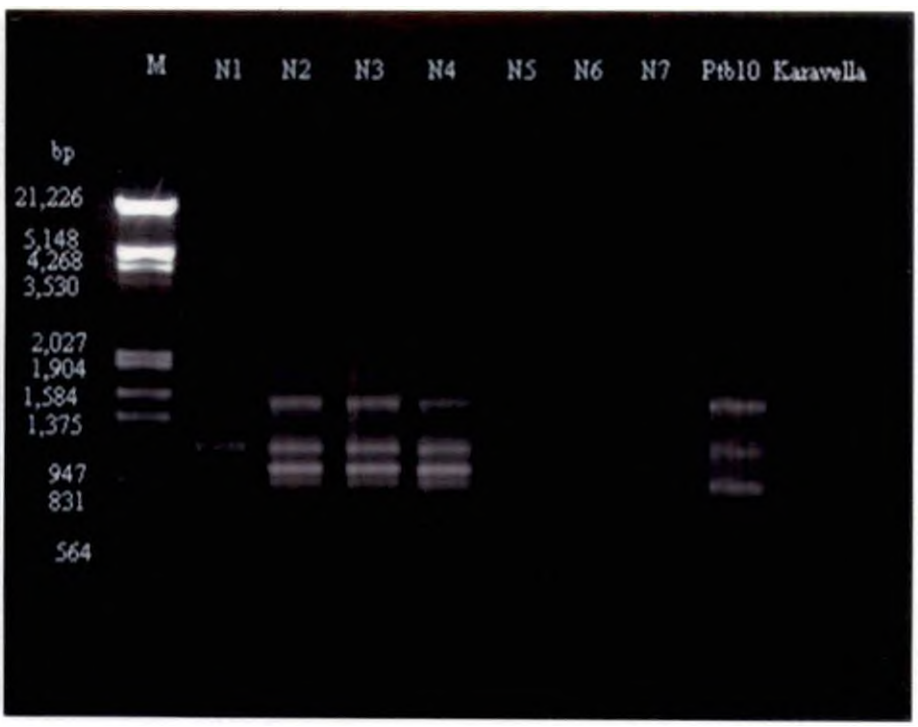


Plate 15. RAPD profile for Njavara genotypes with primer OPP 11

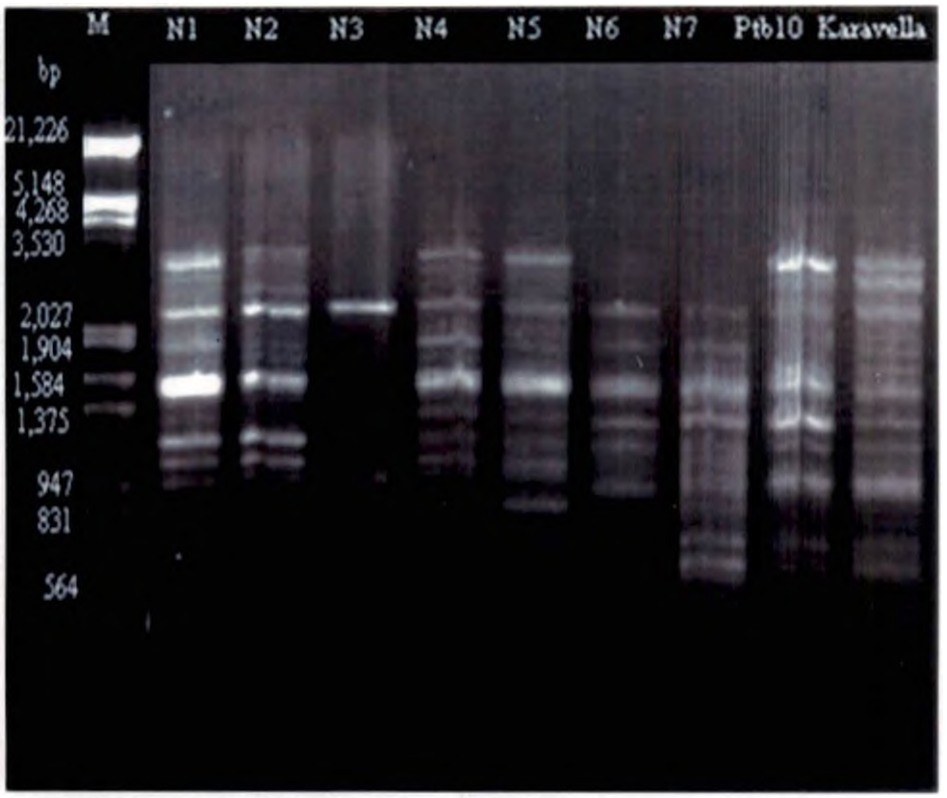


Plate 16. RAPD profile for Njavara genotypes with primer OPE 6

Table 12. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPA 9

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	3	2	1
2.	N2	3	2	1
3.	N3	4	2	2
4.	N4	4	2	2
5.	N5	3	2	1
6.	N6	3	2	1
7.	N7	3	2	1
8.	PTB-10	7	2	5
9.	Karavella	4	2	2
Total		34	18	16

Table 13. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPN 6

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	1	1	0
2.	N2	1	1	0
3.	N3	1	1	0
4.	N4	1	1	0
5.	N5	1	1	0
6.	N6	1	1	0
7.	N7	4	1	3
8.	PTB-10	3	1	2
9.	Karavella	2	1	1
Total		15	9	6

Table 14. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPN 18

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	1	1	0
2.	N2	4	1	3
3.	N3	3	1	2
4.	N4	3	1	2
5.	N5	1	1	0
6.	N6	4	1	3
7.	N7	1	1	0
8.	PTB-10	4	1	3
9.	Karavella	5	1	4
Total		26	9	17

Table 15. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPP 6

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	5	3	2
2.	N2	9	3	6
3.	N3	9	3	6
4.	N4	5	3	2
5.	N5	4	3	1
6.	N6	4	3	1
7.	N7	4	3	1
8.	PTB-10	4	3	1
9.	Karavella	3	3	0
Total		47	27	20

Table 16. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPP 11

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	10	2	8
2.	N2	12	2	10
3.	N3	3	2	1
4.	N4	13	2	11
5.	N5	13	2	11
6.	N6	10	2	8
7.	N7	10	2	8
8.	PTB-10	13	2	11
9.	Karavella	13	2	11
Total		32	18	14

Table 17. RAPD profile for seven *Njavara* genotypes and two check varieties using primer OPE 6

Sl. No.	Genotypes	Maximum no. of bands per sample	No. of monomorphic bands	No. of polymorphic bands
1.	N1	4	0	4
2.	N2	3	0	3
3.	N3	3	0	3
4.	N4	3	0	3
5.	N5	3	0	3
6.	N6	4	0	4
7.	N7	3	0	3
8.	PTB-10	4	0	4
9.	Karavella	1	0	1
Total		28	10	18

4.2.1.4 Molecular genetic relationship

Dendrogram was constructed for *Njavara* genotypes from pooled RAPD data using UPGMA clustering. At similarity coefficient 0.49 the dendrogram got divided into two clusters, one large cluster with all seven *Njavara* genotypes and other small cluster with check varieties, PTB-10 and Karavella (Fig. 1). The *Njavara* genotypes were subdivided into two clusters at 0.81 similarity coefficients with N4, N5, N6 and N7 in one cluster and N1, N2 and N3 in another distinct cluster. N2 and N3 clustered at 1.0 similarity coefficient along with N1 at similarity coefficient of 0.88. N4 formed different cluster from N5, N6 and N7 at similarity coefficient of 0.88. N5 and N6 came under one cluster at a similarity coefficient of 0.90. At similarity coefficient of 0.86, separate cluster was obtained for N7

4.2.2 CLONING OF THE UNIQUE BAND

The N5 genotype was subjected to RAPD analysis with primer OPE 6 which gave unique bands for *Njavara* and then gel electrophoresis was carried out to elute the unique DNA band for further cloning and sequencing.

4.2.2.1 CLONING

The amplified unique band from N5 was eluted from the gel using DNA Gel Extraction Spin Protocol and the presence of DNA was checked through gel electrophoresis, which confirmed the presence of DNA (Plate 17). The eluted product was discrete and repeatable. This product was further used for cloning and sequencing.

4.2.2.1.1 Transformation of DNA

The selected amplicons from N5 was inserted in StrataClone PCR Cloning Vector, pSC-A and the ligated product was transferred to the competent cells. Both blue and white colonies were obtained when the transformed cells were cultured in ampicillin media overlaid with X gal (5-bromo 4-chloro 3-indolyl β-D galactosidase) and IPTG (isopropyl thiogalactosidase). Presence of white colonies indicated a successful transformation. Plasmid DNA isolated from white colonies confirmed the transformation (Plate 18).

Table 18. Amplification pattern by selected primers

Primer	Maximum no. of amplicons per sample	No. of amplicons shared by *Njavara* types alone	No. of amplicons shared by check types and all *Njavara* types	No. of amplicons shared by check types alone
OPA 1	17	0	4	0
OPA 4	5	0	1	0
OPA 6	6	0	2	0
OPA 7	1	0	1	0
OPA 9	7	0	2	0
OPN 6	4	0	1	1
OPN 18	5	0	1	0
OPP 6	9	0	3	0
OPP 11	13	0	2	0
OPE 6	4	3	0	0

Fig.1.Dendrogram of *Njavara* genotypes from pooled RAPD data using UPGMA clustering

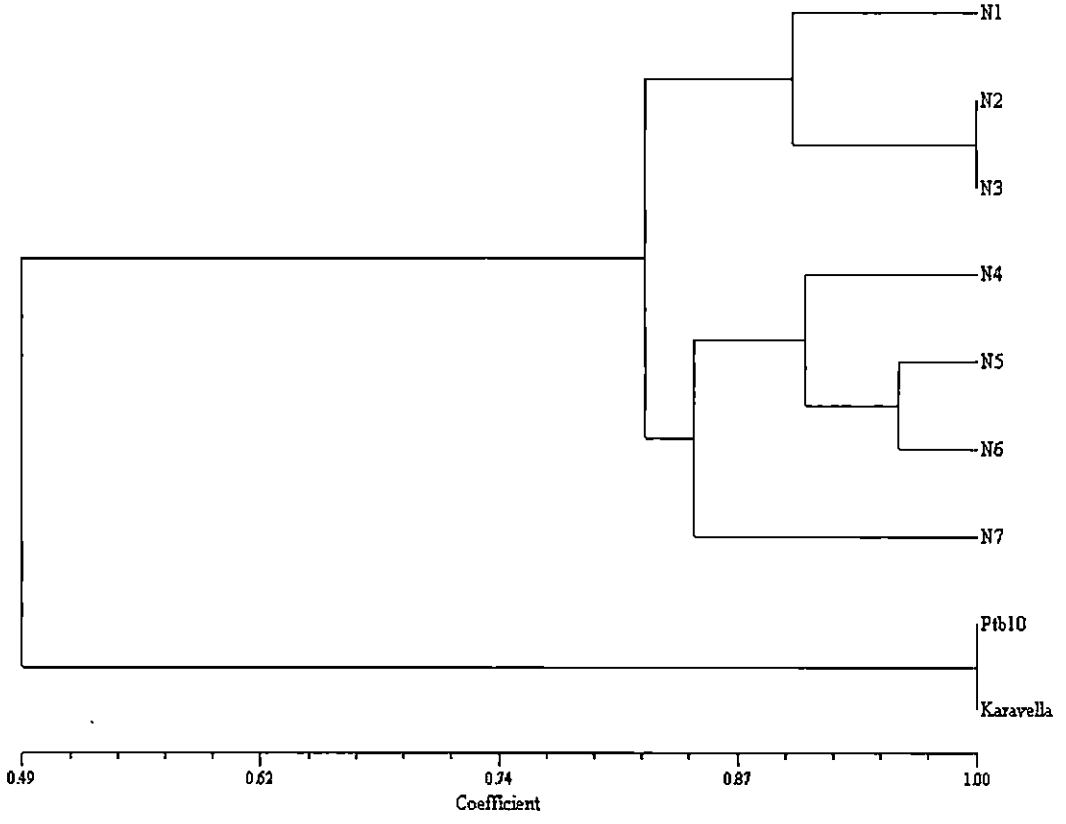


Plate 17. Amplicon after the elution

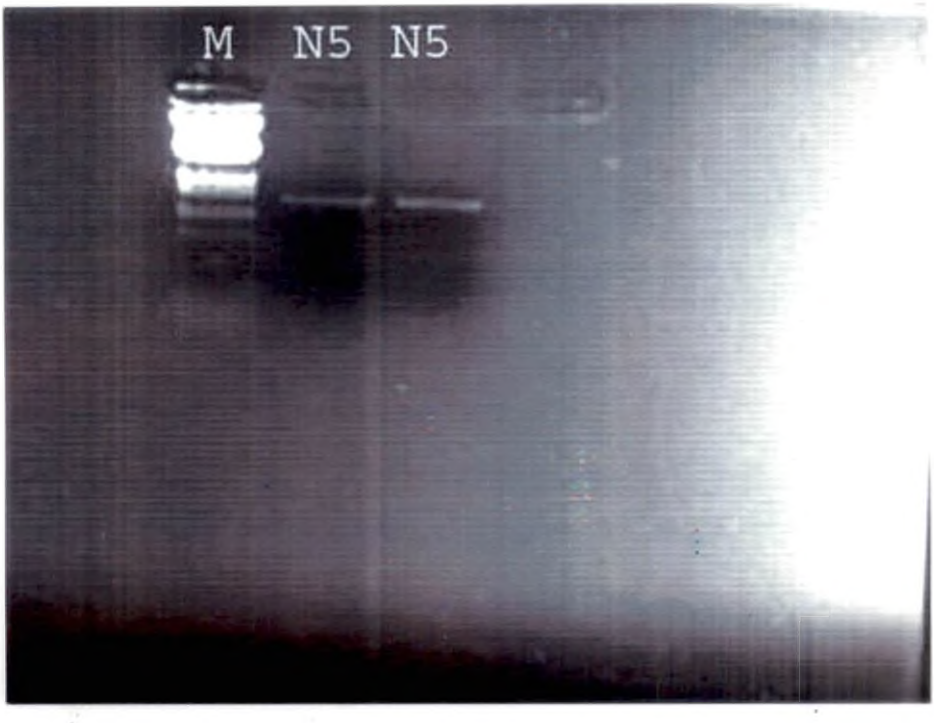


Plate 18. Plasmid DNA isolated from the transformed colonies

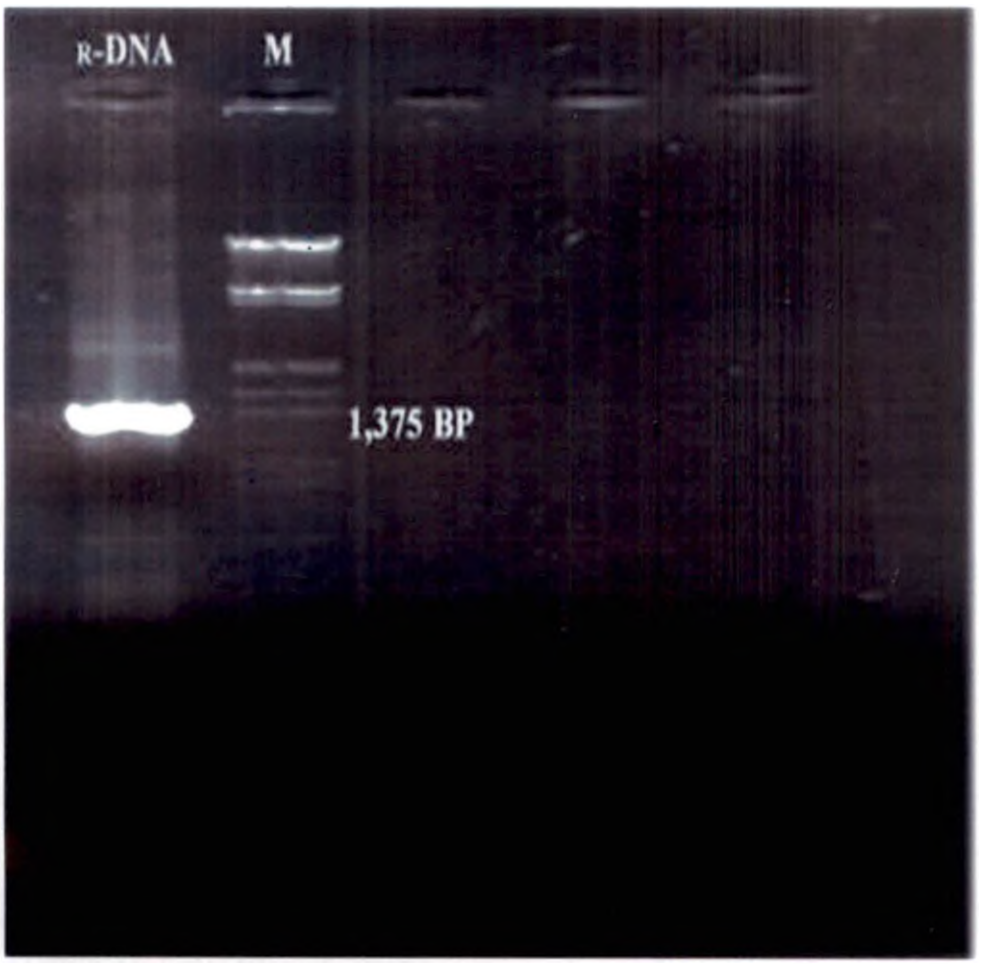
4.2.2.1.2 Screening for recombinant plasmid

The blue colonies and white colonies picked up from the transformation plate were multiplied separately in the LB broth and plasmids isolated were electrophoresed. The plasmid isolated from white colonies had a higher molecular weight when compared to the plasmid isolated from blue colonies. This confirmed the presence of insert in the plasmid.

4.2.2.1.3 Detection of the insert by RAPD amplification

Plasmid DNA isolated from colonies were subjected to RAPD analysis using primer OPE 6 for confirmation of recombination Amplifications were observed in the plasmids isolated from the white colony (recombinant), after RAPD reaction with OPE 6. Amplification was not detected in the plasmids isolated from blue colonies . The RAPD amplification confirmed presence of insert in the recombinant plasmid. There was a single band of required size (1.37 kb) confirming the presence of insert in the plasmid.The results of confirmation of recombinanat plasmid are presented in the Plate 19.

Plate 19. RAPD amplification of recombinant plasmid



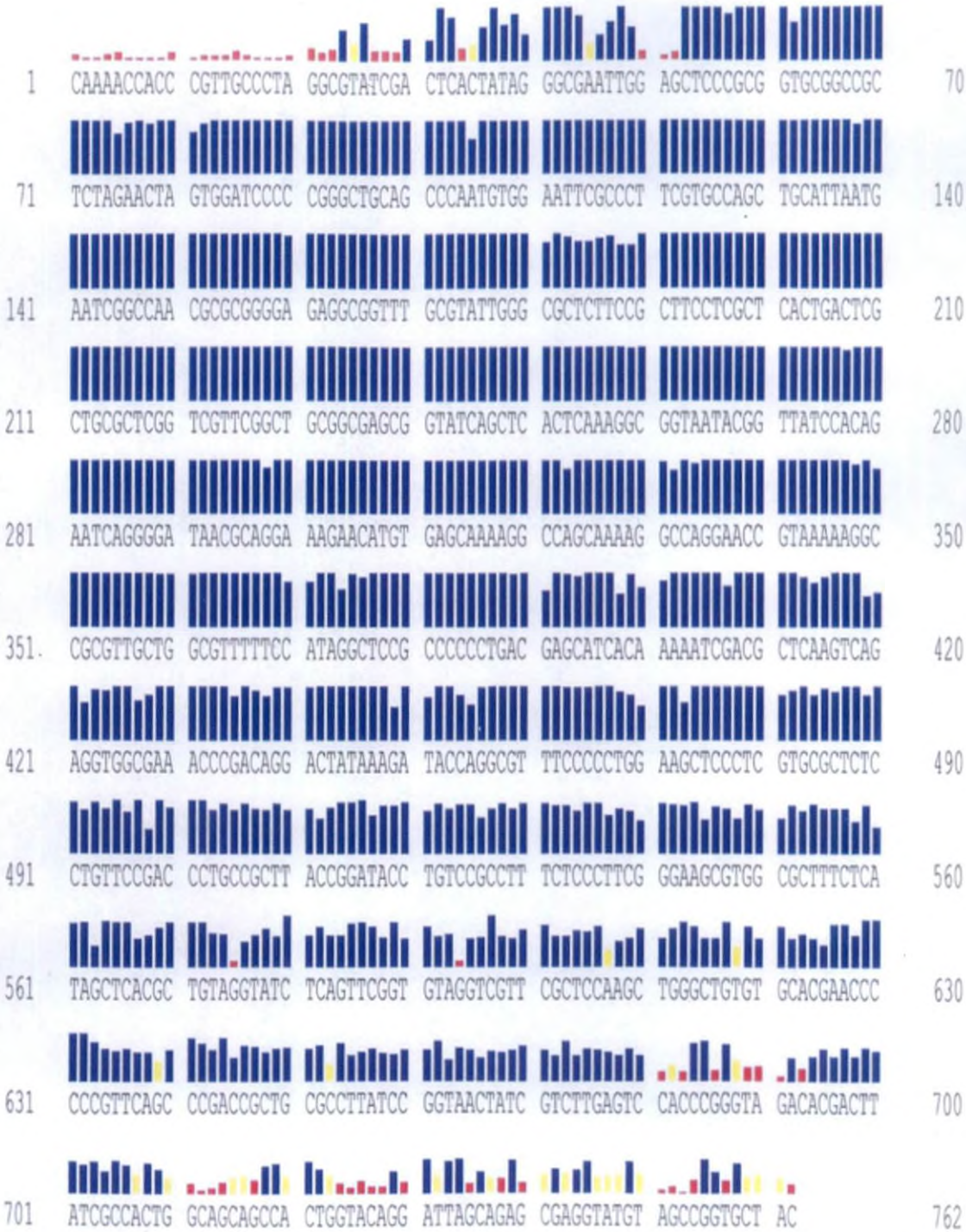
4.2.2.2 Sequence analysis of the clone

4.2.2.2.1 Sequencing

The insert were sequenced at DNA sequencing services, Genei, Bangalore using M13 FP universal primer. The results of the sequence are presented in the graphical form in Fig.2. The sequence was found to be 762 bp long as given below.

>
CAAAACCACCCGTTGCCCTAGGCGTATCGACTCACTATAGGGCGAATT
GGAGCTCCCGCGGTGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCT
GCAGCCCAATGTGGAATTCGCCCTTCGTGCCAGCTGCATTAATGAATC
GGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCG
GTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG
GATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCA
GGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC
CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC
CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG
CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAG
GTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC
GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGT
CTTGAGTCCACCCGGGTAGACACGACTTATCGCCACTGGCAGCAGCCA
CTGGTACAGGATTAGCAGAGCGAGGTATGTAGCCGGTGCTAC //

Fig . 2. Graphical representation of inserted sequence



4.2.2.2.2 Sequence analysis using bioinformatics tools

In the present study, homology of nucleotide sequence of cloned product from *Njavara* with other reported sequences was done.

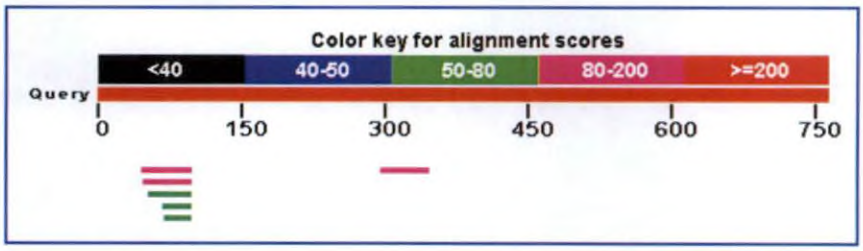
4.2.2.2.2.1 BLAST N

When the nucleotide sequence was subjected to BLAST N homology search (Fig.3), it revealed the maximum identity with genes from *Oryza sativa* (*japonica* cultivar-group) mitochondrial gene , *O.. sativa* mRNA for chilling-inducible protein, *O. sativa* rbbi2-5 gene for putative Bowman Birk tryspin inhibitor, *O. sativa* rbbi2-3 gene for putative Bowman Birk trypsin inhibitor, *O. sativa* rbbi2-4 gene for putative Bowman Birk trypsin inhibitor and *O. sativa* (*japonica* cultivar-group) mRNA for chilling tolerance related protein . Maximum identity of 100 percent was shown to genes from *O. sativa* (*japonica* cultivar-group) mitochondrial gene, *O. sativa* rbbi2-4 gene for putative Bowman Birk trypsin inhibitor and *O.sativa* (*japonica* cultivar-group) mRNA for chilling tolerance related protein .

4.2.2.2.2.2 ORF finder

The sequences were also examined for Open Reading Frames (Plate 20 & 21).The sequence was found to have 5 ORF's. The longest ORF had 180 bases encoding 59 amino acids (Plate 20) in the predicted coding region. Among the amino acids, serine was occurring more frequently than other amino acids. ORF in +3 reading frame was found to be of a residue length of 102 bases, encoding 33 amino acids (Plate 21).

Fig. 3. Results of homology search in cloned fragment obtained through BLAST N



Accession	Description	Max score	Total score	Query coverage	E value	Max identity
D13107.1	*Oryza sativa* (*japonica* cultivar-group) mitochondrial gene for tRNA-Asn, complete sequence, clone:pMTVSCT15	97.1	97.1	6%	4e-18	100%
Z54153.1	*O.sativa* mRNA for chilling-inducible protein	84.2	84.2	6%	3e-14	94%
AJ277472.1	*Oryza sativa* rbbi2-5 gene for putative Bowman Birk tryspin inhibitor	80.5	80.5	6%	4e-13	94%
AJ277470.1	*Oryza sativa* rbbi2-3 gene for putative Bowman Birk trypsin inhibitor	78.7	78.7	5%	2e-12	97%
AJ277471.1	*Oryza sativa* rbbi2-4 gene for putative Bowman Bird trypsin inhibitor	56.5	56.5	3%	7e-06	100%
D10859.1	*Oryza sativa* (*japonica* cultivar-group) mRNA for chilling tolerance related protein, complete cds, clone:pBC121	54.7	54.7	3%	3e-05	100%

Plate 20. Longest Open reading frame in the sequence

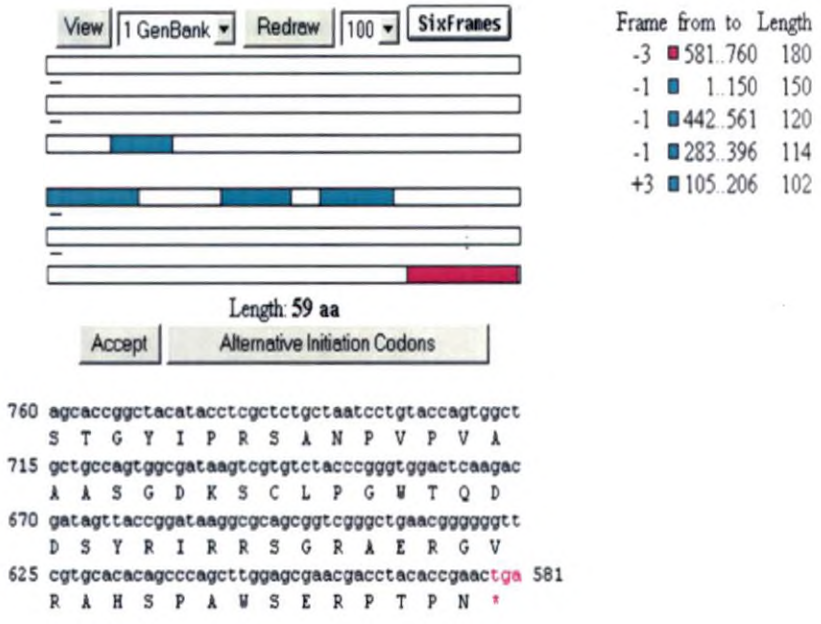
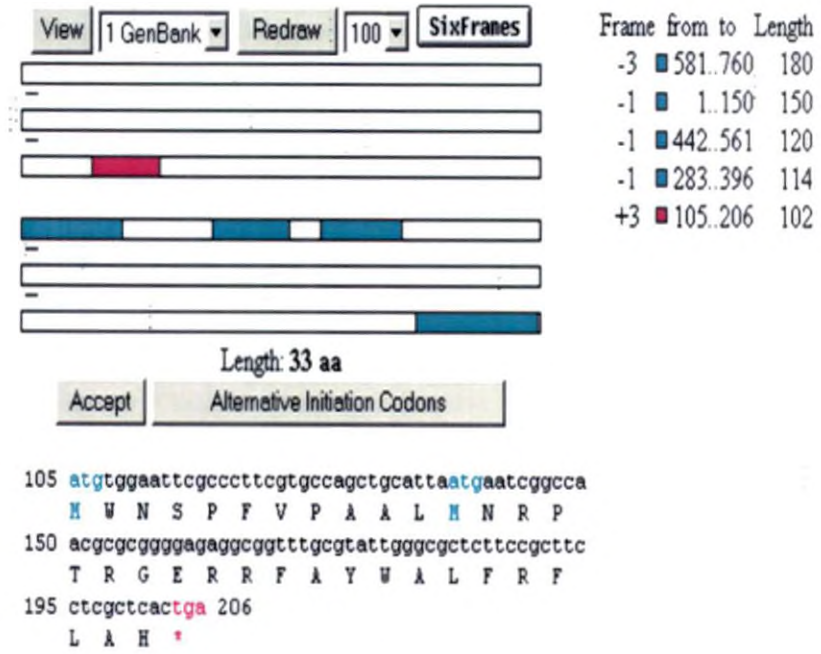


Plate 21. Open reading frame in the + 3 reading frame of the sequence



4.2.2.2.2.3 GENSCAN

Gene prediction analysis were done with the help of GENSCAN tool. It showed the presence of two internal exons from the clone with a length of 91 & 129 residues (Fig.4). Of this the first exon was ranging from 76-166 bases and second from 595-623.These are only fragments of exons predicted from these sequence and will not represent the entire coding sequence of full length gene. The internal exons are shown in Fig.5

4.2.2.2.2.4 VecScreen

Vector screening using VecScreen software of NCBI revealed a strong match of (28-100 bp and 121 -762 bp) and suspect origin of (1-27 bp and 101-120 bp) with cloning vectors like pBR322 , pGEM-9Zf(-),pMC1neo, E.coli lactose operon with lacI, lacZ, lacY and lacA genes, pBluescript II KS(+) vector DNA, phagemid excised from lambda ZAPII etc (Fig. 6).The alignment is given below

> gnl|uv|J01749.1:1-4361-49 Cloning vector pBR322
Length=4410

Score = 1111 bits (554), Expect = 0.0
Identities = 581/586 (99%), Gaps = 2/586 (0%)
Strand=Plus/Plus

Query 179
GGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCG
GCTGCGGCGAG 238
Sbjct 2347
GGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCG
GCTGCGGCGAG 2406

Query 239
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAG
GGGATAACGCAG 298
Sbjct 2407
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAG
GGGATAACGCAG 2466

Fig.4. GENSCAN output

Predicted genes/exons:

Gn.Ex	Type	S	.Begin	...End	.Len	Fr	Ph	I/Ac	Do/T	CodRg	P....	Tscr..
1.01	Intr	+	76	166	91	1	1	15	2	103	0.300	0.56
1.02	Intr	+	595	723	129	0	0	13	52	146	0.461	10.57

Fig. 5. Graphical output of exons predicted from inserted sequence through GENSCAN

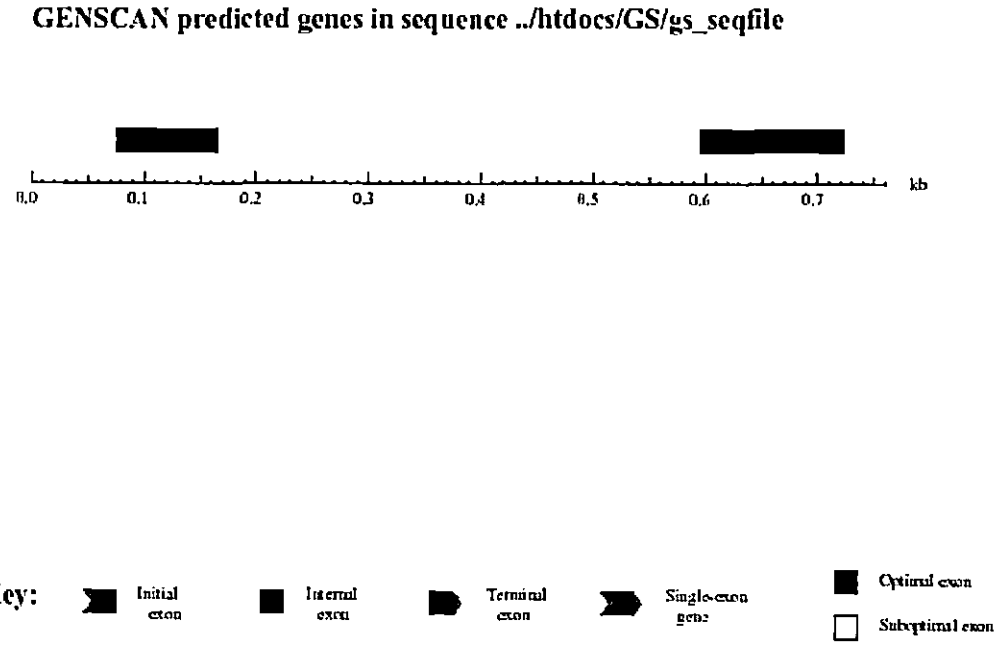
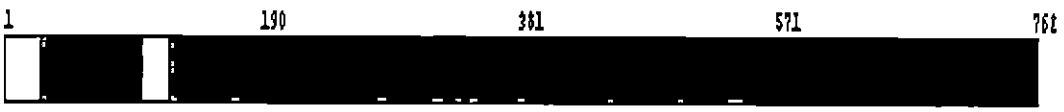


Fig. 6. Distribution of Vector Matches on the Query Sequence using VecScreen programme

Query=
Length=762

Distribution of Vector Matches on the Query Sequence



Match to Vector: Strong Moderate Weak

Segment of suspect origin:

Segments matching vector:
Strong match: 28-100, 121-762
Suspect origin: 1-27, 101-120

> gnl|uv|X65312.2:273-365 Cloning vector pGEM-9Zf(-)
Length=93

Score = 186 bits (93), Expect = 9e-45
Identities = 93/93 (100%), Gaps = 0/93 (0%)
Strand=Plus/Plus

Query 135
TTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCG
CTCTTCCGCTTC 194
Sbjct 1
TTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCG
CTCTTCCGCTTC 60

Query 195 CTCGCTCACTGACTCGCTGCGCTCGGTCGTTCG 227
Sbjct 61 CTCGCTCACTGACTCGCTGCGCTCGGTCGTTCG 93

> gnl|uv|U43611.1:1744-1837 Cloning vector pMC1neo
Length=94

Score = 176 bits (88), Expect = 1e-41
Identities = 94/95 (98%), Gaps = 1/95 (1%)
Strand=Plus/Plus

Query 132
GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGG
GCGCTCTTCCGC 191
Sbjct 1
GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATT-
GGCGCTCTTCCGC 59

Query 192 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTC 226
Sbjct 60 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTC 94

> gnl|uv|J01636.1:1-7477 E.coli lactose operon with lacI, lacZ, lacY and lacA genes
Length=7477

Score = 126 bits (63), Expect = 1e-26
Identities = 63/63 (100%), Gaps = 0/63 (0%)
Strand=Plus/Minus

Query 121
TCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGG
TTTGCGTATTGGG 180
Sbjct 1131
TCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGG
TTTGCGTATTGGG 1072

Query 181 CGC 183
Sbjct 1071 CGC 1069

> gnl|uv|X52327.1:570-844 pBluescript II KS(+) vector DNA, phagemid excised from lambda ZAPII
Length=275

Score = 116 bits (58), Expect = 1e-23
Identities = 73/76 (96%), Gaps = 3/76 (3%)
Strand=Plus/Plus

Query 28 CGACTCACTATAGGGCGAATTGGAGCTC--CCGCGGT-
GCGGCCGCTCTAGAACTAGTGG 84
Sbjct 62
CGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTC
TAGAACTAGTGG 121

Query 85 ATCCCCCGGGCTGCAG 100
Sbjct 122 ATCCCCCGGGCTGCAG 137

4.2.2.2.2.5 CLUSTALW

When the sequence were aligned with vector sequence used for cloning (pSC A vector) using CLUSTALW programme, it revealed poor homology between the sequences, as given below :

Alignment

* matching nucleotides

```
gene        CAAAACCACCCGTTGCCCTAGCCGT---ATCGACTCACTATAGGGCGAATTGGAGCTCC- 56
psca        -ATGACCATG-ATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGG 58
             *  ****    ** * * * ***    **  ** * *  **  * ***    ****

gene        CGCGGTGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGCCCAATGTGGAATTCG 115
psca        TACCGGGCCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCCACTGTGGAATTCG 118
              * * **  ** *** **  * *  * ***     *  **    *** ***********

gene        CCCTTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT 176
psca        CCCTTAAGGGCGAATTCCACATTGGGCTG--CAGCCCG-GGGGATCCACTAGTTCTAGAG 175
            *****   * *   * *   * **    *  ** * ** *****    *   **    *

gene        TGGGCGCTCT--------TCCGCTTCCTCGCTCACTGACTCGC--TGCGCTCGGTCGTTC 226
psca        CGGCCGCACCGCGGGAGCTCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTG 235
             ** *** *         ***  ***  *    * *** ***   * *** ** **  *
```

```
gene    GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA-CAGAATCA 285
psca    GCCGTCGTTTTACAACGTCG--TGACT--GGGAAAACCCTGGCGTTACCCAACTTAATCG 291
        * *  **     *    **   * ***    *       *   **** *** *  ****

gene    GGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAA 345
psca    CCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCG--AT 349
          * * **                          *** **  * *  * ****  *

gene    AAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA--CAAAA 403
psca    TAAATTTTGGTCATGAGATTATCAAAAAGGATCT-TCACCTAGATCCTTTTAAATTAAAA 408
         *      * *  **  ** *   * *** **   * **  **       * *   ****

gene    ATCGACGCTCAAGTCAG----AGGTG-GCGAAACCCGACAGGACTATAAAGATACCAGGC 458
psca    ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAA-T 467
        **  *   * ** ***     * **       *    **  *     * ** *****

gene    GTTTCCCCCTGGAAGCTCCCTCGT--GCGCTCT-CCTGTTCCGACCCTGC-CGCTTACCG 514
psca    GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT 527
        * **   *   ** ** **    * *** ***  ** ** **  * * *    ** **

gene    GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC--ACGCTG 572
psca    GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG 587
        **  **    **  *   *  ** ** * * * **   *  *   * ** *    ****

gene    TAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG-TGCACGAACCCC 631
psca    CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAG 647
         *   **  *           * **    ******   *   * *   *  ** * **

gene    CCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACT--ATCGTCTTGAGTCCACCCGGGT 689
psca    CCGGAAGGGCCGAGCGCAG-AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATT 706
        ***    * **** *** *   *  *** * ****  ***  * *   **** *   *

gene    AGACACGACTTATC-GCCACTGGCAGCAGC-CACTGGTACAGGATTAGCAGAGCGAGGTA 747
psca    AATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGT---- 762
        *       *      ** * *  ** **  * *  **  *   ** **  * **

gene    TGTAGCCGGTGCTAC---------------------------------------------- 762
psca    TGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAG 822
        *** ***  ******

gene    ------------------------------------------------------------
psca    CTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGT 882

gene    ------------------------------------------------------------
psca    TAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCAT 942
```

```
psca  GGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT 1002

gene  ------------------------------------------------------------
psca  GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTC 1062

gene  ------------------------------------------------------------
psca  TTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCAT 1122

gene  ------------------------------------------------------------
psca  CATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAG 1182

gene  ------------------------------------------------------------
psca  TTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGT 1242

gene  ------------------------------------------------------------
psca  TTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACG 1302

gene  ------------------------------------------------------------
psca  GAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTA 1362

gene  ------------------------------------------------------------
psca  TTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCC 1422

gene  ------------------------------------------------------------
psca  GCGCACATTTCCCCGAAAAGTGCCACCTTAATCGCCCTTCCCAACAGTTGCGCAGCCTGA 1482

gene  ------------------------------------------------------------
psca  ATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGC 1542

gene  ------------------------------------------------------------
psca  GCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTT 1602

gene  ------------------------------------------------------------
psca  CCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAG 1662

gene  ------------------------------------------------------------
psca  GGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTT 1722
```

```
gene  ----------------------------------------------------------
psca  CACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGT 1782

gene  ----------------------------------------------------------
psca  TCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATT 1842

gene  ----------------------------------------------------------
psca  CTTTTGATTTACAGTTAATTAAAGGGAACAAAAGCTGGCATGTACCGTTCGTATAGCATA 1902

gene  ----------------------------------------------------------
psca  CATTATACGAACGGTACGCTCCAATTCGCCCTTTAATTAACTGTTCCAACTTTCACCATA 1962

gene  ----------------------------------------------------------
psca  ATGAAATAAGATCACTACCGGGCGTATTTTTTGAGTTGTCGAGATTTTCAGGAGCTAAGG 2022

gene  ----------------------------------------------------------
psca  AAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGAGTACTGCGATGAGTGGCAGG 2082

gene  ----------------------------------------------------------
psca  GCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGTTGCTACGCCT 2142

gene  ----------------------------------------------------------
psca  GAATAAGTGATAATAAGCGGATGAATGGCAGAAATTCGAAAGCAAATTCGACCCGGTCGT 2202

gene  ----------------------------------------------------------
psca  CGGTTCAGGGCAGGGTCGTTAAATAGCCGCTTATGTCTATTGCTGGTTTACCGGTTTATT 2262

gene  ----------------------------------------------------------
psca  GACTACCGGAAGCAGTGTGACCGTGTGCTTCTCAAATGCCTGAGGCCAGTTTGCTCAGGC 2322

gene  ----------------------------------------------------------
psca  TCTCCCCGTGGAGGTAATAATTGACGATATGATCCTTTTTTTTCTGATCAAAAAGGATCTA 2382

gene  ----------------------------------------------------------
psca  GGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCA 2442

gene  ----------------------------------------------------------
psca  CTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCG 2502
```

```
gene  ------------------------------------------------------------
psca  CGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGA 2562

gene  ------------------------------------------------------------
psca  TCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAA 2622

gene  ------------------------------------------------------------
psca  TACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCC 2682

gene  ------------------------------------------------------------
psca  TACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTG 2742

gene  ------------------------------------------------------------
psca  TCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAAC 2802

gene  ------------------------------------------------------------
psca  GGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT 2862

gene  ------------------------------------------------------------
psca  ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCC 2922

gene  ------------------------------------------------------------
psca  GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTG 2982

gene  ------------------------------------------------------------
psca  GTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATG 3042

gene  ------------------------------------------------------------
psca  CTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCT 3102

gene  ------------------------------------------------------------
psca  GGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGA 3162

gene  ------------------------------------------------------------
psca  TAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCG 3222

gene  ------------------------------------------------------------
psca  CAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGC 3282

gene  ------------------------------------------------------------
psca  GCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAG 3342

gene  ------------------------------------------------------------
psca  TGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTT 3402

gene  ------------------------------------------------------------
psca  TATGCTCCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAA 3462

gene  -----
psca  CAGCT 3467
```

4.2.2.2.2.6 BLAST 2 SEQUENCES

BLAST 2 SEQUENCES resulted in local alignment between the vector sequence and query sequence, as given below

Sequence 1: Query sequence
Length = 762 (1 - 762)

Sequence 2: Vector sequence
Length = 3467 (1 - 3467)

```
 Score = 1190 bits (619),  Expect = 0.0
 Identities = 639/644 (99%), Gaps = 2/644 (0%)
 Strand=Plus/Minus

Query  121   TCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3315  TCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG  3256

Query  181   CGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3255  CGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCG  3196

Query  241   GTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGA  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  3195  GTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGA  3136

Query  301   AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTG  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3135  AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTG  3076

Query  361   GCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3075  GCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG  3016

Query  421   AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTC  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3015  AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTC  2956

Query  481   GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2955  GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG  2896

Query  541   GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2895  GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT  2836

Query  601   CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2835  CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC  2776

Query  661   GGTAACTATCGTCTTGAGTCC-ACCCGGGTAGACACGACTTATCGCCACTGGCAGCAGCC  719
              |||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||
Sbjct  2775  GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC  2716

Query  720   ACTGGT-AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC  762
              |||||| |||||||||||||||||||||||||||||||||||||
Sbjct  2715  ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC  2672

 Score =  181 bits (94),  Expect = 4e-42
 Identities = 94/94 (100%), Gaps = 0/94 (0%)
 Strand=Plus/Minus

Query  28   CGACTCACTATAGGGCGAATTGGAGCTCCCGCGGTGCGGCCGCTCTAGAACTAGTGGATC  87
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  217  CGACTCACTATAGGGCGAATTGGAGCTCCCGCGGTGCGGCCGCTCTAGAACTAGTGGATC  158

Query  88   CCCCGGGCTGCAGGCCAATGTGGAATTCGCCCTT  121
             ||||||||||||||||||||||||||||||||
Sbjct  157  CCCCGGGCTGCAGGCCAATGTGGAATTCGCCCTT  124
```

Discussion

5. DISCUSSION

Njavara is a medicinal rice cultivar traditionally used in Ayurveda, the most ancient health care system in India. The medicinal uses of *Njavara* in Ayurveda are well documented in the oldest Ayurvedic texts like "*Ashtanga Hridaya*" and "*Sushrutha Samhita*". Preliminary studies conducted revealed that *Njavara* is genetically diverse with distinct morphotypes, thus pointing to the need for molecular evaluation.

Recently, a range of DNA based markers have been employed for the study of plant diversity and each method has its own benefits and constraints. RAPD is well suited for genetic analysis because of its simplicity, speed, low cost, requirement of smaller quantities of DNA and relative abundance of the marker in the genome. Eventhough studies had been conducted for evaluating morphological and nutritional characterization of *Njavara* germplasm, genetic variation studies with molecular markers have been a subject of lesser consideration. Hence this study was aimed at developing a molecular tool to characterize this cultivar and to reveal its uniqueness through sequence analysis.

5.1 GRAIN COLOUR CHARACTERIZATION

The variability in selected cultivars was examined for salient grain colour characters like lemma and palea colour and seed coat colour .

The results showed variability among *Njavara* genotypes with respect to lemma and palea colour. Two major classes could be distinguished as yellow (straw colour) and black *Njavara* types. Detailed characterization revealed that yellow glume colour could be further grouped into gold furrows on straw (N3 & N5) and brown furrows on straw (N4 & N7). Three genotypes viz., N1, N2 and N6 exhibited a lemma and palea colour dominated by black. N1 and N2 exhibited variations in black colour for lemma and palea as pure black, black furrows/ black

patches on straw background whereas N6 exhibited light shade of black.Sreejayan *et al.* (2003) characterized *Njavara* types and observed black and gold and/or gold furrows on straw background colour for lemma and palea. In the present study none of the *Njavara* types exhibited uniform dark black lemma and palea colour. This showed significant chances of natural crossing between black type with yellow type resulting in varying morphotypes.

Seed coat colour (bran) also varied in collected *Njavara* types. Red, light brown and brown seed coat colours were observed among the genotypes. Most of the genotypes exhibited red seed coat colour. Kumar (2005) reported two types of seed coat in *Njavara* viz., red and brown.

The variation in grain colour indicated the need for molecular characterization of these morphotypes and their conservation for future breeding programmes.

5.2 RAPD ANALYSIS

RAPD technique is important due to its simplicity, efficiency, relative ease to perform and non-requirement of sequence information (Karp *et al.*, 1997).The identification of conserved rice germplasm at the species level can be effective using RAPD (Martin *et al.*, 1997; Williams *et al.*, 1990) markers. In the present study quantity of the DNA extracted varied between 925.1 to 1594.0 ng/µl of leaf tissue. DNA extracted where of good quality and quantity. Modified Dellaporta method was found to be efficient in extraction of DNA from rice.

5.2.1 Screening of random primers

The decamer primers supplied from Operon Technologies Inc. USA were used for the study. A total of 22 decamer primers, ten from OPA series, ten from OPN and two from OPF series were screened initially. Ten primers were selected

based on the number of bands, quality of amplification and stability of expression, for further anlalysis. The amplification with these primers proved the usefulness of Operon primers in rice DNA studies..Many workers have reported the use of Operon primers in rice (Erikson *et al.*, 1995, Hakim, 1996, Subudhi *et al.*, 1997, Rana *et al.*, 1999, Santhy *et al.*, 2000, Raghunathachari *et al.*, 2000, Lee *et al.*, 2001, Sandhu *et al.*, 2002, Jena *et al.*, 2003, Ravi *et al.*, 2003 and Wu *et al.*, 2004).

Among the OPA series, five primers gave good amplification in the present study (Plate 5). Maximum number of bands were eight (for OPA 1 and 4) followed by seven bands for OPA 6 and five bands for OPA 7 and 9.Based on the amplification pattern OPA 1 was rated as the best primer for DNA studies in *Njavara and* this was further confirmed in screening studies with more *Njavara* genotypes.

Primers from OPF series were poor in performance. Among the OPN series, two primers viz., OPN 6 and 18 gave good amplification with four and six bands. This indicated the suitability of OPN 18 for DNA studies in *Njavara.* In the present study seven primers along with three other primers viz., OPP 6, OPP 11, OPE 6 were selected for further screening based on superiority of primer in preliminary studies. Sandhu *et al.* (2002) carried out RAPD analysis by screening 80 primers on 21 rice lines and among these, 13 were selected as suitable on the basis of good DNA amplification. Nadarajan *et al.* (1999) screened ten rice varieties with twenty primers (OPA 1 - 20) and selected ten operon primers based on good amplification for further RAPD analysis.

5.2.2 Screening of *Njavara* genotypes with selected primers

The 10 selected primers were used for screening seven *Njavara* genotypes in order to assess the intra genotype heterogeneity. Of the total 364 scorable amplified products, 352 (96.7 per cent) were polymorphic and 12 bands were monomorphic (3.2 per cent).The molecular weight of amplified products ranged from 0.297 kb to 5.19 kb.

Ranghunathachari *et al.* (2000) employed RAPD profiling for the analysis of genetic variability in 18 accession of Indian scented rice. Polymerase Chain Reaction with ten arbitrary oligonucleotide primers produced a total of 144 different markers bands of which 95.1 per cent were polymorphic. The size range of amplified DNA were between 0.5 kb and 4 kb.

Amplification with primer OPA 1 indicated its possible use in DNA amplification of rice. Molecular weight of bands produced by this primer ranged between 0.58 kb and 3.53 kb (Plate 7). Amplification pattern of N2 and N5 gave maximum of seventeen bands (Table 8). N4 and N7 gave minimum number of amplified products. Eventhough OPA 1 resulted in maximum amplification pattern compared to other primers used in the study, it failed to reveal any unique band for *Njavara* types, or for check varieties.

The primer OPA 4 gave bands with moleular weights ranging between 0.53 kb and 0.96 kb (Plate 8). The genotypes N3, N4, N5, N6, N7, PTB-10 and Karavella gave uniform electrophoretic pattern whereas N1and N2 produced minimum number of amplified products, indicating the inefficiency of this primer as a molecular marker for identification of *Njavara* genotypes (Table 9).

Primer OPA 6 gave a RAPD profile without any monomorphic band (Plate 9). Amplification with this primer resulted in polymorphic amplification in all seven *Njavara* genotypes and two check varieties (Table 10). This primer, as in

the case of OPA 1 failed to give any unique banding pattern for *Njavara* genotypes.

All the seven *Njavara* genotypes and two check varieties gave one monomorphic band for primer OPA7 (Plate10, Table 11) revealing its inefficiency as molecular marker for *Njavara* identification.

Amplification with primer OPA 9 was comparable with that of OPA 1. OPA 9 produced one monomorphic band in all the seven *Njavara* genotypes and two check varieties. The molecular weight of bands ranged between 0.297 kb and 1.9 kb (Plate 11) .N1 gave ten bands. The genotypes N4, N5 and N7 gave uniform electrophoretic pattern whereas N3 gave minimum number of amplified products. Polymorphic bands were obtained for N1, N2, N4, N5, N6 and N7 genotypes with a molecular weight of 0.297 kb (Table 12).

All the seven *Njavara* genotypes and two check varieties gave one monomorphic band for primer OPN 6 (Plate 12). The molecular weight of bands ranged between 0.53 kb and 1.59 kb. Amplification pattern of N7 gave maximum of six bands (Table 13). PTB-10 exhibited a unique medium molecular weight band (1.59 kb) with OPN 6 indicating utilizing this primer as a marker to characterize PTB-10, the cosmopolitan traditional variety from Kerala which had been used in the development of many HYVs of rice released in the state.

Primer OPN 18 gave a RAPD profile with one monomorphic band (1.9 kb). The genotypes N2, N3 and N4 showed uniform electrophoretic banding pattern. The molecular weight of bands ranged between 0.44 kb and 1.9 kb. Amplification pattern of N6 gave maximum of six bands. PTB-10 and Karavella exhibited unique bands of molecular weights of 0.81kb and 1.2 kb respectively (Plate 13, Table 14).

Amplification with primer OPP 6 produced banding pattern without any monomorphic bands (Plate 14, Table 15). The molecular weight of bands ranged between 0.715 kb and 1.5 kb. The genotypes N2 and N3 showed uniform electrophoretic banding pattern .Eventhough amplification studies by Kumar (2005) with primer OPP6 produced unique bands for *Njavara* genotypes,in the present study this primer failed to produce any unique band for *Njavara.*

Electromorphic banding pattern for primer OPP 11 gave high amplification for all genotypes except N3 (Plate 15). Two monomorphic bands (0.94 kb and 2.1 kb) were shown with seven *Njavara* genotypes and two check varieties. N5 exhibited maximum number of thirteen bands (Table 16). N6 and N7 exhibited uniform banding pattern. Kumar (2005) reported that *Njavara* genotypes produced unique bands with OPP 11 primer. This indicated the usefulness of this primer in the genomic studies of rice .Its use in molecular characterization of *Njavara* need further studies.

Amplification with primer OPE 6 resulted in three unique bands in all seven *Njavara* genotypes (Plate 16, Table 17). These bands were absent in both the check varieties indicating its usefulness as a marker to identify *Njavara* genotypes. This is in confirmation to the reports of Kumar (2005) who reported the presence of unique bands for *Njavara* genotypes using the primer OPE 6. Of these unique bands two were of medium molecular weight (1.375 kb and 1.29 kb) .The, third band was of low molecular weight (0.44 kb). The unique high intensity amplified product of 1.375 kb was used for further cloning and sequencing studies.

When the amplification pattern were compared, primer OPA 1and OPA11 gave maximum number of amplified products viz., 17 and 13 for the genotypes under study. Hence can be utilized for amplification studies in rice.

OPN 6 exhibited one unique band for PTB-10, which can be utilized for identification of this variety.

From the above discussion it could be inferred that OPE 6 produced distinct bands for *Njavara* genotypes .It is also worthy to note that this unique band was present in all *Njavara* genotypes under study,eventhough lemma and palea colours of these types were different and was absent in the two check varieties studied. Hence this can be used as a molecular marker to identify *Njavara* genotypes from other rice types.

5.2.3 Molecular genetic relationship

The dendrogram constructed with RAPD markers revealed that *Njavara* genotypes were distinct from other check varieties (Fig 1).PTB-10 and Karavella entered into a single cluster.

The distinct clustering of *Njavara* genotypes along with unique banding pattern revealed that *Njavara* might have some peculiar characters which could be its medicinal property. The dendrogram drawn by Kumar (2005) also revealed the distinctness of *Njavara* from check varieties.

The dendrogram drawn by Raghunathachari *et al.* (2000) revealed the relative genetic similarities among the scented rice accessions as ranging from 25 to 77.5 per cent. Wu *et al.* (2004) carried out cluster analysis (UPGMA) for 14 samples of rice to determine genetic diversity among and within populations of species. The dendrogram indicated that the genetic similarity between populations ranged from 0.63 to 0.92.

5.3 Cloning and sequencing of the unique fragment

Molecular cloning refers to the procedure of isolating a defined DNA sequence and obtaining multiple copies of it *in vivo*. Cloning is frequently employed to amplify DNA fragments containing genes, but it can also be used to amplify any DNA sequence such as promoters, non-coding sequences, chemically synthesised oligonucleotides and randomly fragmented DNA.

5.3.1 Transformation of eluted DNA

Transformation is the genetic alteration of a cell resulting from the uptake and expression of foreign genetic material (DNA).

When Strataclone competent cells were transformed with ligated PCR product both blue and white colonies were observed. Transformed *E. coli* cells with non recombinant plasmid developed into blue colonies in the presence of chromogenic substrate X-gal (Ullman *et al.*, 1967, Hortwitz *et al.*, 1964).

Presence of white colonies in the transformed plate indicated successful recombination.

5.3.2 Screening for recombinant plasmids

The plasmid isolated from white colonies had a higher molecular weight when compared to the plasmid isolated from blue colonies. This confirmed the presence of insert in the plasmid.

5.3.3 Confirmation of recombination

Polymerase chain reaction carried out with RAPD primer yielded amplicon of desired size viz., 1.37 kb approximately (Plate 19). No amplicon was observed in negative control. This confirmed presence of insert in the vector.

5.4 Sequence analysis

5.4.1 Sequencing

The cloned fragments when sequenced by automated sequencing with M13 Forward primer provided the information about the region starting from forward primer. The sequence was found to be 762 bp long.

Sequence analysis encompasses the use of various bioinformatic methods to determine the biological function and/or structure of genes and the proteins they code.

Homology search through BLAST programme (http://www.ncbi.nlm.nih.gor/blastl) is a heuristic method to find the highest scoring locally optimal alignments between a query sequence and a database sequence (Attschul *et al.*, 1997). This can determine the sequence homology to predict the identity and function of query sequence. In the present investigation homology search showed high homology of the cloned DNA segment to *O. sativa* (*japonica* cultivar-group) mitochondrial gene for tRNA-Asn, complete sequence, clone: pMTVSCT15, *O.sativa* mRNA for chilling-inducible protein , *O. sativa rbbi2-5* gene for putative Bowman Birk tryspin inhibitor, *O. sativa rbbi2-3* gene for putative Bowman Birk trypsin inhibitor, *O. sativa rbbi2-4* gene for putative Bowmam Birk trypsin inhibitor, *O. sativa* (*japonica* cultivar) mRNA for chilling tolerance related protein, complete cds, clone:pBC121 (Fig. 3). Of the three hits that gave

maximum identity of 100 percent, one was *rbbi2-4* gene for putative Bowman Birk trypsin inhibitor. The segment also showed 94 percent identity with *O. sativa rbbi2-5* gene for putative Bowman Birk tryspin inhibitor and *O.sativa* mRNA for chilling-inducible protein . Bowman-Birk Trypsin Inhibitor protein, is also known to possess anti-inflammatory and anti-allergic properties in animals and was reported to be capable of imparting resistance to fungal pathogens and pests in crops .

Proteinase inhibitors are potential tools of crop improvement targeting plant protection and human nutrition. The Bowman Birk Inhibitor (BBI) family is a typical canonical serine proteinase inhibitor. They are found in the seeds of leguminous (dicots) and graminaceous (monocot) plants (Birk, 1985.). BBI proteins are important since they are highly stable to both cooking temperature and digestion (Yavelow *et.al.*, 1985). Qu *et al.* (2003) had identified seven Bowman Birk trypsin inhibitor genes in *japonica* cultivar of rice through molecular cloning and functional analysis.

Anticarcinogenic nature of Bowman Birk Inhibitor protein had been reported by many workers. Soybean diet rich in protease inhibitors (e.g. BBI) lowered the breast tumour incidence in irradiated rats (Troll *et al.*, 1980). Kennedy (1993) reported that Bowman Birk inhibitor suppressed tumorigenesis in several experimental carcinogenesis model systems and BBI could maintain its cancer chemopreventive property even when administered as a dietary supplement. Hence Sawey (2001) opined that BBI will be eminently suitable as a cancer chemopreventive agent for use in humans.The homology of cloned DNA fragments of *Njavara* (N5) with BBI genes (*rbbi 2-3, rbbi2-4* and *rbbi 2-5*) in BLASTN analysis is a preliminary support to the medicinal property (anticarcinogenic property) of this unique cultivar of Kerala and requires indepth studies for confirmation. Earlier studies about BBI had revealed that they are highly stable to cooking temperature.In Ayurveda system, for *'Panchakarma'* treatments, *Njavara* grains are boiled with herbal preparations and then used for

massaging to induce muscle and nerve rejuvenation.The presence of sequences related to the thermostable BBI protein justifies the above mentioned method of application of *Njavara* grains in '*Panchakarma*' treatment.

Open reading frame or ORF is a portion of an organism's genome which contains a sequence of bases that could potentially encode a protein. In a gene, ORFs are located between the start-code sequence (initiation codon) and the stop-code sequence (termination codon). It starts with an initiation codon and ends with a termination codon. The existence of an ORF, especially a long one, is usually a good indication of the presence of a gene in the surrounding sequence. Once the open reading frame is known, the DNA sequence can be translated into its corresponding amino acid sequence. Short ORFs can also occur by chance outside of genes. The present clone has five ORF's (Plate 20-21). The longest open reading frame had 180 bases encoding 59 amino acids (Plate 20) in the predicted coding region. Among the amino acids,serine was occurring more frequently than other amino acids. ORF in +3 reading frame was found to be of a residue length of 102 bases, encoding 33 amino acids (Plate 21).

The regions that actually encode the gene product, which can be much smaller than the introns, are known as exons. One single gene can lead to the synthesis of multiple proteins through the different arrangements of exons produced by alternative splicings.The exons present in the cloned sequence were analysed by 'Genscan' tool. It determined two internal exons from the clone with a length of 91 and 129 residues (Fig. 4). These are only fragments of exons predicted from these sequence and will not represent the entire coding sequence of full length gene. Present study revealed presence of ORF with occurrence of two internal exons and absence of introns. Boulter (1993) had reported that gene size and coding regions of proteinase inhibitors are small, devoid of introns. Indepth studies about gene fragments encoding active proteins, ORFs and exons are required to prove the usefulness of identified gene sequence in plant protection and health care systems.

However sequence analysis of the data with VecScreen of NCBI revealed that major portion of the sequence showed homology to sequences from Vectors like pBR322 , pGEM-9Zf(-), pMC1neo, E.coli lactose operon with lacI, lacZ, lacY and lacA genes, pBluescript II KS(+) vector DNA, phagemid excised from lambda ZAPII etc.There were also two regions of 27 kb and 20 kb indicating suspect origin.As the vector used in the study was pSCA , the presence of vector sequence from the above mentioned vectors are to be further studied and analysed.

When the alignment of vector pSCA sequence was compared with query sequence through CLUSTALW programme it revealed poor homology. But the vector sequence showed homology to query sequence when it was aligned with BLAST 2 SEQUENCES. Hence further studies in this regard is very essential to confirm the usefulness of the results.

5.4.2 Molecular marker for *Njavara* types

DNA fragment (1.37kb) eluted from amplified DNA of N5 *(Njavara)* genotype was used in the present study for gene cloning and sequencing. Screening of all other *Njavara* genotypes used in the study with OPE 6 through RAPD analysis had also revealed the presence of this unique DNA fragment. It is also worthy to note that this band was absent in the two traditional rice varieties viz., PTB-10 and Karavella used in the study. Kumar (2005) also had reported the presence of unique bands in *Njavara* when the primer OPE 6 was used for RAPD analysis.This pointed out to the possibility of using this DNA band generated in RAPD analysis as a molecular tool for the identification of this unique rice cultivar. Such molecular characterization also have significance in protecting the 'Intellectual Property Rights' of the farming community over this medicinal cultivar of Kerala especially in the era of globalization and 'biopiracy'.

Summary

6. SUMMARY

An investigation on 'Molecular documentation of *Njavara* type of rice for cultivar identification' was undertaken in the Department of Plant Breeding and Genetics and Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara from October 2005 to August 2007. Seven accessions of *Njavara* along with two check varieties (PTB-10 and Karavella) were included in the study. The salient findings are summarized below.

1. *Njavara* genotypes exhibited high variability with respect to lemma and palea colour. Two major classes could be distinguished as yellow (straw colour) and black *Njavara* types. Detailed characterization revealed that yellow type could be further grouped into gold furrow on straw (N3 & N5) and brown furrow on straw (N4 & N7). Three genotypes viz., N1, N2 and N6, exhibited a lemma and palea colour dominated by black. N1 and N2 exhibited variations in black colour for lemma and palea as pure black, black furrows/ black patches on straw background whereas N6 exhibited light shade of black.

2. Seed coat colour (bran) also varied in collected *Njavara* types. Red, light brown and brown seed coat colours were observed among the genotypes.

3. The method suggested by Dellaporta *et al.* (1983) with slight modifications was found effective in isolating good quality genomic DNA from *Njavara*. Good amplifications were observed for OPA 1, OPA 4, OPA 6, OPA7, OPA 9, OPN 6, OPN 18, OPP 6, OPP 11, and OPE 6. OPA 1, OPP 11 were found to be promising in the amplification of rice genomic DNA with maximum amplification.

4. The dendrogram with RAPD markers showed distinct clusters for *Njavara*. The *Njavara* genotypes were grouped in one cluster and check varieties in another.

5. The eluted product (fragment from N5) was effectively ligated in Strataclone vector and carried out transformation.

6. Sequencing of the clone with M 13 primer gave the sequence data of a gene segment of size 762 bp.

7. The homology search for clones showed that it has maximum identity with sequences from *Oryza sativa* (*japonica* cultivar-group) mitochondrial gene for tRNA-Asn, complete sequence, clone:pMTVSCT15, *O. sativa* mRNA for chilling-inducible protein, *O.sativa rbbi2-5* gene for putative Bowman Birk tryspin inhibitor, *O.sativa rbbi2-3* gene for putative Bowman Birk trypsin inhibitor, *O.sativa rbbi2-4* gene for putative Bowman Birk trypsin inhibitor and *O. sativa* (*japonica* cultivar-group) mRNA for chilling tolerance related protein, complete cds, clone:pBC121 (Fig. 3). Maximum identity of 100 percent was shown to genes from *O. sativa* (*japonica* cultivar-group) mitochondrial gene, *O. sativa* rbbi2-4 gene for putative Bowman Birk trypsin inhibitor and *O.sativa* (*japonica* cultivar-group) mRNA for chilling tolerance related protein.

8. The homology of cloned DNA fragments of *Njavara* (N5) with BBI genes (*rbbi 2-3*, *rbbi 2-4* and *rbbi 2-5*) is a preliminary indication of the medicinal property (anticarcinogenic) of this unique medicinal cultivar of Kerala. The present investigation indicated the presence of segments of BBI gene (not the full length gene) in *Njavara*. Further studies about full length gene and BBI protein activity are required to confirm the uniqueness of this medicinal rice cultivar.

9. The present investigation revealed homology of sequenced segment with mRNA for chilling tolerance related gene in *japonica* rice

10. The present clone has five ORF's (Plate 20-21). The longest open reading frame had 180 bases encoding 59 amino acids (Plate 20) in the predicted coding region. Among the amino acids, serine was occurring more frequently than other

amino acids. ORF in +3 reading frame was found to be of a residue length of 102 bases, encoding 33 amino acids.

11. GENSCAN tool determined two internal exons from the clone with a length of 91 and 129 residues (Fig. 4). These are only fragments of exons predicted from these sequence and will not represent the entire coding sequence of full length gene.

12. Sequence analysis of the data with VecScreen of NCBI revealed that major portion of the sequence showed homology to sequences from Vectors like pBR322 , pGEM-9Zf(-), pMC1neo, E.coli lactose operon, pBluescript II KS(+) vector DNA, phagemid excised from lambda ZAPII etc.There were also two regions of 27 kb and 20 kb indicating suspect origin. As the vector used in the study was pSCA and as the sequence analysis revealed homology with other vectors that were not used for the study, indepth studies are required in this area.

13. Alignment of sequences through CLUSTAL W programme revealed poor homology within query sequence and vector sequence used for cloning. Homology was observed between the sequences when BLAST 2 sequences programme was used. These results are to be confirmed through further studies.

14. Screening of *Njavara* genotypes with OPE 6 primer had revealed the presence of unique DNA fragment,which was absent in the two traditional rice varieties viz., PTB-10 and Karavella used in the study .This pointed out to the possibility of using this DNA band generated in RAPD analysis as a molecular tool for the identification of this unique rice cultivar. This is a confirmation result with regard to OPE 6 primer.

Suggested future line of work

1. Collection and conservation of all morphotypes of *Njavara* .
2. In depth studies for molecular characterization and sequencing.
3. Studies to prove medicinal property and other unique characters of *Njavara.*
4. Protection of this cultivar as a farmers variety under PPV & FR Act.

References

REFERENCES

Abedinia, M., Henry, R.J. and Blakeney, A.B. 2000. Accessing genes in the tertiary gene pool of rice by direct introduction of total DNA from *Zizania palustris* (wild rice). *Plant Mol. Biol. Rep.* 18:133-138

Adreit. 2006. Microsatellite markers for population studies of the rice blast fungus, Magnaporthe grisea. *Mol. Ecology Notes* (online Early Articles)

Akkaya, M.S., Bhagwat, A.A. and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean .*Genet.* 132:1131-1139

Altschul, S.F.,Thomas,L.,Madden, AA.,Schaffer, Zhang,Z., Miller,W.and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programes.*Nucleic Acids Res.*25:3389-3402

Arunachalam, A. 2000. Participatory conservation: a means of encouraging community biodiversity. *Plant Genet. Newsletter.*122 :1-6

Balachandran, P.V.2007. Rice scenario of Kerala and the future strategies. In: Muthunayagam,A.E.(ed.), *Proceeding of the nineteenth Kerala Science Congress* ; 29-31 January,2007;Kannur.Kerala state council for science, technology, and Environment, Government of Kerala,pp22-32.

Birk, Y.1985. *Int. J. Pept. Protein Res.*25:113-131

Birk,Y.,Gertler,A.,Khalef,S. 1963.*Biochem.J.*87:281-284.

Bostin, B., White, R.L., Bkolnick, M. and Davis, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American of Human Genet.* 32:314-331

Boulter, D. (1993). Insect pest control by copying nature using genetically engineered crops. *Biochem.* 34:1453-1466.

Brandolin. 2006. Chanometrical characterization of focus Italian Rice varieties based on genetic and chemical analysis. *J. Agric. Food Chem.* 54(26):9985-9991

Buchanan, F. 1807. *A journey from Madras through the countries of Mysore, Canara and Malabar*. Reprinted in 1988 by Asian Educational Sciences, New Delhi, India, 424p.

Changhu Wang, Peng Zhang, Zhenrong, M.A., Mingyong Zhana, Guchou Sun. and Dinghow Ling. 2004. Development of a genetic markers linked to a new therma sensitive male sterile gene in rice (*Oryza sativa* L.). *Euphytica*. 140(3) :217-222

Chen, P.W., Chow, S.H. and Chen, L.J. 1997. Nucleotide sequence of a cDNA (Acc No. U76001) encoding rice Bowman-Birk Proteinase inhibitor (PGR 97-015). *Plant Physiol*. 113: 663

Cheng, J. W., Zai, Q. C., Xing, Q. H., Shou, H. Y., Kai, M. C. and Chong, R. S. 2004. Genetic diversity among and within populations of *Oryza granulate* from Yunnan of China revealed by RAPD and ISSR markers: implications for conservation of the endangered species. *Pl. Sci*. 167 (1): 35-42

Chen-Hong and Zhu-Littuing. 1995. Construction of rice RAPD molecular linkage map. *Acta-Botanica-Sinica* . 37(9):677-684

Choudhury, P. R., Kohli, S., Srinivasan, K., Mohapatra, T. and Sharma, R. P. 2001 Identification and classification of aromatic rices based on DNA fingerprinting. *Euphytica*.118 (3): 243-251

Cousse, N.A., T.M., Fullon, Y.G. Cho, S.N. Ahn, J. Chunnongse, K.S. Wu, J.H. Xiao, Z.H., Yu, P.C. Ronald, S.E. 1994. Saturated Molec7ular map of the rice genome based on inter specific back cross population. *Genet.* 138: 1251-1274

Dash, 1992.Pancha karma therapy of Ayurvedic-massage therapy in Ayurveda. Concept publishing company, New Delhi.

De Pater, S., L. A. M. Hensgens and R. A. Schilperoort, 1990. Structure and expression of a light- inducible shoot-specific rice gene. *Plant Mol. Biol.* 15: 399-406.

Dellaporta, S. L., Wood, J. and Hicks, J. B. 1983. A plant DNA minipreparation: Version II. *Plant Mol. Biol. Reptr*. 1 (1): 19-21.

Deshpande, V.G. and Ranjekas, P.K., 1980. Repetitive DNA in three Graminae species with low DNA content. *Physiol Chem.* 361: 1223-1233

Elsy, C.R., Rosamma, C.A. and Potty, N.N. 1992. *Njavara* - a rice variety with special characters. *Oryza* 29:55-56

Erikson, J., Rogers, W. J. and Perera, A. L. T. 1995. Genetic screening of rice varieties using RAPD markers for indication of salt sensitivity/resistance. *Sri Lankan J. agric. Sci.* 32 (3): 16-29

Estorninos, L. E., Jr., Burgos, N. R., Gealy, D. R., Stewart, J. M. and Talbert, R. E. 2000. Genetic characterization of red rice populations using molecular markers. *Research Series - Arkansas Agric. Exp. Station.* 476: 76-84

Federici. 2001. Analysis of Uruguayan weedy rice genetic diversity using AFLP molecular markers. *J. of Biotech.*. 4 (3) :27-36

Feng,R.Q.,Sowen,Z.W.,Cheng,W.C.2005.Structural features and molecular evolution of Bowman-Birk protease inhibitors and their potential application.Acta Biochimica et Biophysica Sinica.37(5):283-292.

Fritsch, P. and Rieseberg, L.H.1996. The use of random amplified polymorphic DNA(RAPD) in conservation genetics. In: Molecular Genetic Approaches in Conservation. Smith, T.B. and Wayne, R.K. (eds).Oxford University press, London, pp 54-73.

Fukuoka 2006. Genetic organization of aromatic rice as revealed by RAPD markers: A case study in conserving crop genetic resources on farm *Euphytica.* 149(1-2):61-71

Ghose, T.K., Narottam Dey and Sudarsana Biswas. 2005. RAPD based genetic diversity analysis of aromatic rice (*Oryza sativa* L.). *Plant Cell Biotech. and Mol. Biol* .6(4): 133-142

Giarrow L.E., Marassi, M.A. and Salerno, G.L. 2007. Assessment of the genetic diversity in Argentine rice cultivars with SSR markers. *Crop Sci.* 47: 853-858

Government of India. 2007. GI. – application number 17. *Geographical indications J.* 17:32-42.

Government of Kerala. 2004. *Economic review 2004.* State Planning Board, Thiruvananthapuram, Kerala, India, 179 p.

Hakim, L. 1996. Detection of DNA sequence polymorphism between mutant and mother variety of rice by randomly amplified polymorphic DNA (RAPD). *J. Nucl. Agric. Biol.* 25 (2): 59-62

Hali, R. 2006. Kerala calling. *Njavara*: Kerala's own wonder Rice.2(27):33

Horwitz, A.J. and Chua, J., Curby, R.J., Tanson, A.J. and Fisher, B.E. 1964. Substrates for cytochemical demonstration of enzyme activity. *J.Med. Chem.*7:574-776

Huang, J., Zhang, H., Wang,J. and Yang, J. 2003. Molecular cloning and characterization of rice 6-phosphogluconate dehydrogenase gene that is up-regulated by salt stress. *Molecular Biol Rreports.30(4):* 223-227

Huff, D.R., Peakall, R., Smouse, P.T. 1993. RAPD variation within and among natural populations of out crossing buffalo grass (Buchloe daityloides). *Theor. Appl. Genet.*, 86:927-934

IRRI. 1995. *Standard Evaluation System for Rice.* International Rice Research Institute, Los Banos, Philippines, 56p.

Ishii, T., Nakano, T., Maeda, H. and Kamijima, O. 1996. Phylogenetic relationships in A-genome species of rice as revealed by RAPD analysis. *Genes and Genetic Systems.* 71 (4): 195-201

Ishii, T., Xu, Y. and Mc Couch, S.R. 2001. Nuclear and Chloroplast - microsatellite variation in A- genome species of rice. *Genome* .44 (4): 658-666

Jain, M.,Tyagi,A.K.,Jitendra,P.2006.Molecular characterization and differential expression of cytokinin-response type-A response regulators in rice. *Plant Physiol.*6:1

Jayakumar,V.2003.Economics of commercial production and utilization of medicinal rice (Oryza sativa L.) Njavara. Msc thesis. Kerala Agriculture University, Thrissur, p61.

Jena, K. K., Pasalu, I. C., Rao, Y. K., Varalaxmi, Y., Krishnaiah, K., Khush, G. S. and Kochert, G. 2003. Molecular tagging of a gene for resistance to brown plant hopper in rice (*Oryza sativa* L.). *Euphytica.* 129 (1): 81-88

John, E. Lunn, Anthony R. Ashtam, Marshall D. Hetch and Hans W. Heldt. 2000. Purification, molecular cloning and sequence analysis of sucrose - 6^{th} phosphate phosphohydrolase from plants. *Proc. Natl. Acad. Sci.* 97(23):12914-12919

Junjian, N., Peles, M., Colonit, G.I., David, J. and Mac Kil, 2002. Evaluation of genetic diversity in rice subspecies using Microsatellite markers. *Crop Sci. Soc. of America.* 42: 601-607

Karp A, Kresovich S, Bhat K V, Ayad W G and Hodgkin T 1997 Molecular tools in plant genetic resources conservation: a guide to the technologies; in *IPGRI Technical Bulletin* No. 2. International Plant Genetic Resources Institute, Rome, Italy

Kennedy, A.R. 1993. Cancer prevention by protease inhibitors .*Prev.Med.* 22:796-811

Kennedy, A.R.1993.Anticarcinogenic activity of protease inhibitors. Troll W. Kennedy, A.R. eds. Protease Inhibitors as cancer chemopreventive agents. Plenum Publishing .New York,65-91pp.

Kim, K., Jac Keun Sohn, Yong Shem Kevon, Mooyoung Evn and Akira Kato. 1999. Cloning and sequencing of molecular markers Associated with Cold Sensitivity at seedling stage in rice. Intr. Plant and Animal Genome Conerence. Town and Country Hotel, San Diego, January 17-21, 1999, p. 328

Ko, H.L. and Cowan, D.C. 1994. Random amplified polymorphic DNA analysis of Australian rice (*Oryza sativa* L.) varieties. *Euphytica.* 80(3):179-189

Kumar, S.2005. Biochemical and molecular characterisation of *Njavara* types of Rice (*Oryza sativa* L.). Ph.D thesis, Kerala Agricultural University. Thrissur, 241p

Kumary, S.L. 2004. Genetic diversity, characterization and evaluation of medicinal rice in Kerala. Compendium of papers of Science - Society

interface on Medicinal and Aromatic Rices, August 20-21, 2004, KAU, Thrissur and MSSRF, Chennai at Regional Agricultural Research Station (KAU), Pattambi, Kerala, India, 123 p.

Kumary, S.L. 2007. Biodiversity of rice in Kerala. In: Muthunayagam, A.E.(ed.), *Proceeding of the nineteenth Kerala Science Congress*; 29-31 January, 2007; Kannur. Kerala state council for science, technology, and Environment, Government of Kerala ,pp 33-46

Kumary, S.L. Susamma, P.George., N., Dijee Bastian., K.C. Aipe and N. Rema Bai.2003.Aromatic Rices of Kerala.In.R.K.Singh and U.S. Singh.(eds).A treatise on the Scented Rices on India. Kalyani publishers, New Delhi.P.317-326

Kurata, N., Y. Nagamura, K. Yamamoto, Y. Harushima, N. Sue, J. Wu, B.A. Antonio, 1994. A300 kilobase interval genetic map of rice including 883 expressed sequences. *Nature Genet.* 8 :365-372

Lalitha, A. 2005. DNA marker technology in the improvement of Horticulture crops. In: ICAR National symposium on Biotechnology interventions for improvement of Horticulture crops: Issues and strategies.10-12 jan. 2005. COH, KAU, thrissur, Kerala .pp8

Lee, H. G., Kim, K. M. and Sohn, J. K. 2001. Identification of RAPD markers associated with grain weight in rice. *Korean J. Crop Sci.* 46 (4): 261-265

Liu, Q., Wang, L., Chen, S. Lin, F. and Pan, Q.A. 2005. Genetic and physical mapping of Pi 36(t), a novel rice blast resistance gene located on rice chromosome 8. *Mol. Genet. and Genomics.* 274 (4) :.394-401

Liu, W. and Yan, Y.S. 2007. Impart assessment of cadmium contamination on rice (*Oryza sativa* L.) seedlings at molecular and population levels using multiple biomarkers. *Geophysical Res. Abstrc.* .9 (07) :508

M.S. Swaminathan Research Foundation. 2005. *Navara.* Community Agro-biodiversity centre, 3p.

Mackill, Z. and Redona,E.D.1995. Microsatellite and RAPD markers polymorphism in rice. *USPA-ARS and Rpt. of Agron and Range Science*, Unt. of California

Mahmood, T., Hyder,M.Z.,Naqvi,S.M.S.2007.Cloning and sequence analysis of germin-like protein gene 2 promoter from *Oryza sativa* L. Spp. *indica. DNA sequence.*18 (1):26-32.

Manakkodan, K.K.V. 1949. *Pancha Karma adhava shodhan cikitsa* Vidyarambham Press, Alappuzha, Kerala.

Mane, S.P., Shashidar, H.E., Kanbas, A. and Hittalmani, S. 2003. Molecular marker analysis for root length in a diverse germplasm of rice (*Oryza sativa* L.). *Indian J. Genet.* 63(3):197-201

Martin, C., Juliano, A., Newbury, H. J., Lu, B. R., Jackson, M. T. and Lloyd, F. B. V. 1997. The use of RAPD markers to facilitate the identification of *Oryza* species within a germplasm collection. *Genet. Resour. Crop Evol.* 44 (2): 175-183

Masahiko, K.H. and Delin Hong. 2002. Characterization of CMS and maintains lines in indica rice (*Oryza sativa* L.) based on RAPD marker analysis. *Euphytica* 129: 249-252

Mc Cough, S.R., Chen, X., Panand, O., Temnykh, S., Xu, Y. Cho, Y.G., Huang, N. Ishii, T. and Blair, M. 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol. Biol.* 35: 89-99

Mc Cough, S.R., Teytelman, L., Keno, I. and Yano, M. 2002. Development of 2,240 new SSR markers for rice (*Oryza sativa* L.) *Genome Res.*

Menon, M.V. and Potty, N.N. 1996. Effects of different inputs on productivity and quality relations in Njavara (*Oryza sativa* L.). Ph.D. thesis, Kerala Agricultural University, Thrissur, India, 239p.

Monna, L., Miyao, A., Inoue, T., Fukuoka, S., Yamazaki, M., Zhong, H. S., Tanaka, T., Shibayama, S. and Sasaki, T.1994. Rice RAPD markers and their conversion into STSs. *Rice Genome.*3 (2): 2-4

Moos, A.N.E.T. 2004. The internal uses of *Njavara* and other medicinal rice varieties.In paper presented at Science-Society Interface on Medicinal and Aromatic Rices. Regional Agricultural Research station, pattambi, Kerala, India, August 20-21,2004.

Mundy, J. and N.H. Chua, 1988. Abscisic acid and water stress induced expression of a novel rice gene. *EMBO J.* 7: 2279-2286.

Nadarajan, N., Vivekanandan, P. and Henry, R. J. 1999. Use of molecular markers for diversity analysis in rice. *Indian J. Genet. Plant Breed.* 59 (3): 247-259

Newbury, H. J. and Ford, L. B. V. 1993. The use of RAPD for assessing variation in plants. *Plant Growth Regulation.* 12 (1-2): 43-51

Olsen, K.M., Caicedo.A.L. Polato, N, McClung, A., McCouch, S. and Purugganan, M.D.2006. Selection Under Domestication: Evidence for a Sweep in the Rice *Waxy* genomic region. *Genetics* .173(2): 975–983.

Parsons, B. J., Newbury, H. J., Jackson, M. T. and Ford, L. B. V. 1997 Contrasting genetic diversity relationships are revealed in rice (*Oryza sativa* L.) using different marker types. *Molec. Breed.* 3 (2): 115-125

Pordie,L.2006.The Institutionalisation of Therapeutic Practices in India. Social and legal perspectives.7-8 DEC 2006.French inst of pondichery (FIP) india. http://www.ifpindia.org/societies-and medicines –in south-asia. html.

Prabhakaran, 2004, June 07."*Njavara* facing extinction".*The Hindu*

Presting, G.G., Budiman, N.A., Wood, T. and Yu, Yeisoo. 2000. A frame work for sequencing the rice genome. Proceedings of an international conference, Los Banos, Laguna, Philippines, 27-29 March 2000

Qian, W., Ge, S. and Hong, D. Y. 2001 Genetic variation within and among populations of a wild rice *Oryza granulata* from China detected by RAPD and ISSR markers. *Theor. Appl. Genet.* 102 (3): 440-449

Qu, L.J., Chen,J., Liu, M. and Pan, N.2003. Molecular cloning and functional analysis of a novel type of Bowman-Birk inhibitor gene family in rice. *Plant Physio.*. 133:560-570.

Rafalski, J. A., Tingey, S. V. and Williams, J. G. K. 1991. RAPD markers - a new technology for genetic mapping and plant breeding. *Ag Biotech News and Information.* 3 (4): 645-648

Raghunathachari, P., Khanna, V. K., Singh, U. S. and Singh, N. K. 2000. RAPD analysis of genetic variability in Indian scented rice germplasm (*Oryza sativa* L.). *Curr. Sci.* 79 (7): 994-998

Ragot, M. and Hoisington, D. A. 1993. Molecular markers for plant breeding: comparisons of RFLP and RAPD genotyping costs. *Theor. Appl. Genet.* 86 (8): 975-984

Rana, P., Naina, W.H.S., Sindhu, A., Jain, R.K. and Chowdhury, J.B. 1999. RAPD based assessment of diversity in Indica rice genotypes. *Indian J. of Exp. Biol.* 37 :.1209-1212

Randhawa,G., Firke, K.P. 2006. Genetic analysis of aromatic rice (*Oryza sativa* L.) using sequence-tagged microsatellite markers. 2nd Int. Rice Congress 2006. Oct. 9-13, 2006, New Delhi, pp.161-162

Ravi 2003. Molecular Marker based Genetic Diversity Analysis in Rice (*Oryza sativa* L.) using RAPS and SSR markers. *Euphytica* 133:243-252

Reddy, G.S. 2000. Characterisation and evaluation of the rice (*Oryza sativa* L.) cultivar *Njavara*. M.Sc. (Ag.) thesis, Kerala Agricultural University, Thrissur, India, 137p.

Reece, K. S., McElroy, D. and Wu, R. 1990. Genomic nucleotide sequence of four rice (*Oryza sativa*) actin genes. *Plant Mol. Biol.* 14: 621-624.

Rohlf, F. J. 1998. NTSYSpc. *Numerical Taxonomy and Multivariate Analysis System*. Applied Biostatistics, Inc., New York, USA, 583p.

Saker, M.M. 2005.Genetic analysis of some Egyptian rice genotypes using RAPD,SSR and AFLP.*Afr. J. of Biotech.*4(9):882-890.

Sandhu, S. S., Bastos, C. R., Azini, L. E., Neto, A. T. and Colombo, C. 2002. RAPD analysis of herbicide-resistant Brasilian rice lines produced via mutagenesis. *Genet. Mol. Res.* 1(4): 359 - 370

Santhy, V., Mohapatra, T., Dadlani, M., Sharma, S. P. and Sharma, R. P. 2000. DNA markers for testing distinctness of rice (*Oryza sativa* L.) varieties. *Plant Varieties Seeds.*13 (3): 141-148

Sawey, M.J. 2001.Role of Gap-Junctional Communication in Breast Cancer Progression and Chemoprevention. *J. of nutrition.*131:167-169.

Shirasawa, K., Kishitani, S. and Nishio, T.. 2005. Conversion of AFLP markers to sequence - specific markers for closely related lines in rice by use of the rice genome sequence. *Mol. Breeding.*14 (3): 283-292

Siddiq, E.A. 2004. Aromatic and medicinal rices: Urgency for validation and conservation for sustainable utilization towards value added rice farming. In Keynote address at Science - Society Interface on Medicinal and Aromatic Rices (Regional Agrl. Res. Station, Pattambi, Kerala, India, August 20-21, 2004)

Singh, N.K., Saurabh Raghuvanshi and Srivastava, S.K. 2004. Sequence analysis of the long arm of rice chromosome II for rice wheat synteny. Functional and Integrative Genomics. 4(2):102-117

Song, G.E., Giancarlo, C. 1999. RAPD variation within a between atural populations of the wild rice *Oryza rufipagon* from China and Brazil. Laboratory of systemic and Evolutionary Botany, Inst. of Botany, Chinese Academic of Sciences, Beijing, China Heridity. 82:638-644

Sreejayan, Kumar, V.R. and Thomas, G. 2003. Collections and morphological evaluation of Njavara, a traditional medicinal rice (*Oryza sativa* L.) in Kerala, India. *Plant Genet. Resour. Newsl.* 135:12-17

St.Clair, W.H., st.Clair, D.K. Effects of the Bowman Birk protease inhibitor on the expression of oncogenes in the irradiated mouse colon. *Cancer Res.*1991;51:4539-4543.

Subudhi, P. K., Borkakati, R. P., Virmani, S. S. and Huang, N. 1997. Molecular mapping of a thermosensitive genetic male sterility gene in rice using bulked segregant analysis. *Genome.* 40 (2): 188-194

Sujith Kumar, C.K.1999. *'Krishi Malayalam'* Akshara samskrithi, Samskrithi publications, Kannur, Kerala.

Sushruthacharya. 2500 BC. *Sushrutha Samhitha* (Translated by Bhishagratna and Kunjalal, K., 1991) Fourth edition. Chowkhamba Sanskrit Series Office, Varanasi, India, 2671p.

Tashiro, M., Hashinok, Shiozaki, M. Ibuki, F., Maki, Z. 1987. The complete amino acid sequence of rice bran trypsin inhibitor. *J. Biochem.* 102: 297-306

Thomas, G. 2006. *Njavara*, the single largest used medicinal rice in the world: Genetic structure, varietal discrimination and phytogenetic affinity.In: 2nd International Rice Congress 2006 , 9-13 october 2006. New Delhi, Abstract: 618

Thomas, G., Sreejayan, Nair, P. 2004. Rationalisation of genetic resource collection of *Njavara*, the traditional medicinal rice in Kerala, India using morphological traits and AFLP markers. *In:Compendium of papers of Science – Society Interface on Medicinal and Aromatic Rices*, 20 – 21 August 2004, Regional Agricultural Research station (RARS), KAU, Pattambi, Kerala, India.

Troll,W.,Wiesner,B.,Shellabarger,C.J.,Holtzman,S.,Stone,J.S.1980.Soybean diet lowers breast tumor incidence in irradiated rats.*Carcinogenesis*.1:469-472.

Ullman, A., Jacob, F. and Monod,J.1967.Characterization by *in vitro* complementation of a peptide corresponding to an operator-proximal segment of the ß-galactosidase structural gene of *Escherichia Coli*.*J.Mol.Biol.*24:339-341

Vagbhata. 500 BC. *Ashtanga Hridayam* (Translated by Murthy, K.S.R., 1991). Krishnadas Academy, Varanasi, India, 2153p.

Virk, P. S., Ford, L. B. V., Jackson, M. T., Pooni, H. S., Clemeno, T. P. and Newbury, H. J.1996. Predicting quantitative variation within rice germplasm using molecular markers. *Heredity*. 76 (3): 296-304

Virmani, S.S. 1999. Exploitation of heterosis for shifting the yield frontier in rice. In: J.G. Coors and S. Pandey (Eds.), The Genetics and Exploitation of heterosis in crops, CIMMYT, ASA, CSSA, Madison, NI. pp.423-438

Wang Zongyang, Gu Ming Hong, Liu Qiao Quch, Chen Xiu Hua, Hsinkan. 2004. CDNA cloning and sequence analysis of rice sbe 1 and sbe 3 genes. *Rice Sci.* 2 (3): 81-85

Wang, Z. Y. and Tanksley,S.D.1989. Restriction fragment length polymorphism in *Oryza sativa* L. *Genome.* 32: 1113-1118.

Waugh, R. and Powell, W. 1992. Using RAPD markers for crop improvement. *Trends in Biotech.* 10: 186-191

Welsh, J. and McClelland, M. 1990. Finger printing genomes using PCR with arbitrary primers. *Nucleic Acids Res.* 8:7213-7218

Williams, J.G.K., Kubelk, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18:6531-6535

Wu, C. J., Cheng, Z. Q., Huang, X. Q., Yin, S. H., Cao, K. M. and Sun, C. R. 2004. Genetic diversity among and within populations of *Oryza granulate* from Yunnan of China revealed by RAPD and ISSR markers: implications for conservation of the endangered species. *Plant Sci.* 167: 35 - 42

Wu, K.S. and S.D. Tanksley, 1993. Abundance, Polumorphism and genetic mapping of rie satellites in rice *Mol. Gen.* 241: 225-35

Wu, Y.J. and Chen, Y. 2006. RAPD analysis of jasmine rice - specific genomic structure. *Genome* . 49(6): 716-719

Xiang Taihe, Wang, Lilin and Pang Jilians. 2005. Clonging and characterization of a full length cab gene encoding the light - harvesting chromophyll a/b - binding proteins in rice (*Oryza sativa* L.). *Acta Agronomica Sinica,* 31(9): 1227-1232

Yamaguchi,S. K., J. Mundy and N.H. Chua, 1989. Four tightly linked *rab* genes are differentially expressed in rice. *Plant Mol. Biol.* 14: 29-39.

Yang, G.P., M.A.S. Maroof, C.G. Xu, Q. Zhang and R.M. Biyasher, 1994. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. *Mol. Gen.* 245:187-94

Yang,Z., Zongyang, W., Minghong, G., Liu, Q., Chen, X. and Wu, H.2004.cDNA cloning and sequence analysis of rice sbe1 and sbe3 genes. *Rice Sci.* 11(3):81-85.

Yavelow, J., Collins, M., Birk, y., Troll, W. and Kennedy, A.R. 1985. *Proc. Natl. Acad. Sci.* USA,82:5395-5399

Yu, L.X. and Ngugen, H.T. 1994. Genetic variation detected with RAPD markers among upland and low land rice cultivars (*Oryza sativa* L.). *Theor. Applied Genet.* 87(6):668-672

Zabcaum, M. and P. Vos. 1993. Selective restriction fragment amplification: a general method for DNA finger printing. European patent 0534858

Zeng, L., Kivon, T.R., Liu, X. and Wilson, C. 2004. Genetic diversity analysed by microsatellite markers among rice (*Oryza sativa* L.) genotypes with different adaptations to saline soils. Plant Science, April 20, 2004, 66:1275-1285

Zhao, X.P. and G. Kochut, 1993. Phytogenetic distribution and genetic mapping of (GGC) (a) micro satellite from rice (*Oryza sativa* L.). *Plant Mol. Biol.* 21: 607-614

Zhuang, J. Y., Ma, W. B., Peng, Y. C., Wang, H. C., Zhang, H. L. and Zheng, K. L. 1998. Screening of RAPD markers for the parentage identification of three-line hybrid rice. *Chinese Rice Res. Newsl.* 6: 1- 2

MOLECULAR DOCUMENTATION OF *NJAVARA* TYPES OF RICE (*Oryza sativa* L.) FOR CULTIVAR IDENTIFICATION

By

SHAREESH. N

ABSTRACT OF THE THESIS

Submitted in partial fulfilment of the
requirement for the degree of

Master of Science in Agriculture

Faculty of Agriculture
Kerala Agricultural University

Department of Plant Breeding and Genetics

COLLEGE OF HORTICULTURE
VELLANIKKARA, THRISSUR - 680 656
KERALA, INDIA

2007

ABSTRACT

Characterisation and evaluation of *Njavara* types of rice (*Oryza sativa* L.) was under taken in the Department of Plant Breeding and Genetics and Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara during 2005 - 2007 with the objective of molecular characterization and gene sequencing in *Njavara* for developing suitable molecular markers for cultivar identification.

Njavara genotypes exhibited high variability with respect to lemma and palea colour with two major classes viz., yellow (straw colour) and black types. Detailed characterization revealed that yellow type could be further grouped into gold furrows on straw (N3 & N5) and brown furrows on straw (N4 & N7). Three genotypes viz., N1, N2 and N6 exhibited a lemma and palea colour dominated by black. N1 and N2 exhibited variations in black colour for lemma and palea as pure black, black furrows/black patches on straw background whereas N6 exhibited light shade of black. Variations in seed coat colour as red, light brown and brown were also observed among the genotypes.

The method suggested by Dellaporta *et al.* (1983) with slight modifications was found to be effective in isolating good quality genomic DNA from *Njavara.* Good amplifications were observed when RAPD analysis was performed with sequences OPA 1, OPA 4, OPA 6, OPA7 , OPA 9, OPN 6, OPN 18, OPP 6, OPP 11 and OPE 6. OPA 1 and OPP 11 were found to be promising in the amplification of rice genomic DNA with maximum amplification. Amplification of *Njavara* DNA with primer OPE 6 exhibited unique bands (1.375 kb , 1.29 kb and 0.44 kb) for *Njavara* genotypes and hence are valuable as DNA markers for the identification of this unique cultivar. The dendrogram with RAPD markers showed distinct clusters for *Njavara.*

Cloning and sequencing of the unique molecular band with M 13 primer gave the sequence data of a gene segment of size 762 bp.The homology search of this sequence with BLASTN showed that it has maximum identity with genes from *Oryza sativa* (*japonica* cultivar-group) mitochondrial gene for tRNA-Asn, complete sequence, *O. sativa* mRNA for chilling-inducible protein, *O.sativa rbbi2-5* gene for

putative Bowman Birk tryspin inhibitor, *O.sativa rbbi2-3* gene for putative Bowman Birk trypsin inhibitor, *O.sativa rbbi2-4* gene for putative Bowman Birk trypsin inhibitor and *O. sativa* (*japonica* cultivar-group) mRNA for chilling tolerance related protein.The homology of cloned DNA fragments of *Njavara* (N5) with BBI genes (*rbbi 2-3, rbbi 2-4* and *rbbi 2-5*) is a preliminary indication of the medicinal property (anticarcinogenic) of this unique medicinal cultivar of Kerala and also its thermostable nature.

Sequence analysis revealed the presence of five ORF's . The longest open reading frame had 180 bases encoding 59 amino acids in the predicted coding region. Among the amino acids, serine was occurring more frequently than other amino acids. ORF in +3 reading frame was found to be of a residue length of 102 bases, encoding 33 amino acids. Genscan tool determined two internal exons from the clone with a length of 91 and 129 residues.

Sequence analysis of the data with VecScreen showed strong match to vector sequences in the database eventhough the sequences were not matching with pSCA (vector used in the present study) vector. Alignment of sequences through CLUSTAL W programme revealed poor homology with query sequence and vector sequence used for cloning. Homology was shown between the sequences when BLAST 2 SEQUENCES programme was used. These results are to be confirmed through further studies.