Elucidating the Molecular Diversity Pattern of Major Bacterial Blight R-Genes and Identification of Candidate Bacterial Blight Resistance Genes in Rice

Thesis submitted to the University of Hyderabad
for the award of

Doctor of Philosophy (Ph.D.)

by

WAIKHOM BIMOLATA
(08LPPH08)

Supervisor

Dr. Irfan Ahmad Ghazi

Department of Plant Sciences
School of Life Sciences
University of Hyderabad
Hyderabad 500046, India
May 2014

University of Hyderabad
(A Central University established in 1974 by act of parliament)
HYDERABAD – 500 046, INDIA

"DECLARATION"

I, **Waikhom Bimolata** hereby declare that this thesis entitled **"Elucidating the molecular diversity pattern of major bacterial blight *R*-genes and identification of candidate bacterial blight resistance genes in rice"** submitted by me, was carried out in the Department of Plant Sciences, School of Life Sciences, University of Hyderabad, under the guidance and supervision of **Dr. Irfan Ahmad Ghazi** is an original and independent research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

Waikhom Bimolata
(Research Scholar)

University of Hyderabad
(A Central University established in 1974 by act of parliament)
HYDERABAD – 500 046, INDIA

CERTIFICATE

This is to certify that this thesis entitled "**Elucidating the molecular diversity pattern of major bacterial blight *R*-genes and identification of candidate bacterial blight resistance genes in rice**" is a record of bonafide work done by **Miss Waikhom Bimolata**, a research scholar for Ph.D. programme in Plant Sciences (Registration No. 08LPPH08), Department of Plant Sciences, School of Life Sciences, University of Hyderabad under my guidance and supervision. The thesis has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Dr. Irfan Ahmad Ghazi
(Supervisor)

(Head of the Department) **(Dean of the School)**

Preface

This thesis is submitted in partial fulfillment of the requirements for a degree in Doctor of Philosophy. This Ph.D. thesis contains the result of research undertaken at the Department of Plant Sciences, School of Life Sciences, University of Hyderabad, Hyderabad. I would have not been able to achieve this task without the financial support from the Department of Science and Technology, New Delhi, India under project grant no. SR/SO/PS-21/09. Certainly, I would have never reached the point of finishing my dissertation without the help and support of others. This thesis is based on research in the area of molecular diversity analysis of disease resistance genes coupled with identification of candidate genes for resistance against bacterial blight disease of rice. This dissertation is original and some part of it has been published in the journal PLANTA. The main aim of this study was to identify alternate form of resistance gene which will provide durable and sustainable resistance to elite susceptible rice varieties against diverse virulent *Xoo* strains. We addressed this question from the perspective of exploring and exploiting the available genetic resources for crop improvement. This dissertation is organized into five chapters and a general conclusion and summary. Each chapter is inter-related and linked to the problem addressed. The first chapter briefly introduces and addresses the background of the problem. Subsequently, we proposed the targeted means of going to the bottom for handling it. This chapter connects different objectives of the study. The second chapter provides brief review of the concerned problem and different levels of research being carried out in the area worldwide. The third chapter describes about different methodology used and applied to obtain the proposed solution. This chapter deals with the overall screening for identification of valuable genetic resource and the desired form of resistance genes. We investigated the natural variation in genes which are involved in bacterial blight resistance. The fourth chapter describes about the combine results obtained for the three different objectives. The last chapter discusses about the results obtained and different factors limiting the study. This chapter also explains how our results are informative, similar and different from other's report available in literatures. Finally, a brief summary and conclusion uniting the elements in the dissertation is presented. In the process of writing this thesis, I feel that I have learned a lot which will help me in future scientific endeavor.

CONTENTS

Page Nos.

Acknowledgement

I take this opportunity to thank and express my deep gratitude to my supervisor Dr. Irfan Ahmad Ghazi for his constant support, encouragement and guidance throughout my doctoral research. His critical assessments of my work help me in improving the quality of the work. I would also like to thank Dr. Insaf Ahmed Qureshi, Department of Biotechnology, who always extended his help and support throughout my research. I thank for his constructive suggestions and critical review of my doctoral work.

My most sincere thanks go to Dr. R. Meenakshi Sundaram and Dr. G. S. Laha, Senior Scientists, Directorate of Rice Research, Hyderabad who always extended their generous support from providing germplasms, Xoo culture and field for screening studies. Their key role in shaping my work with their valuable suggestions and guidance are deeply appreciated.

I express my gratitude to Prof. P. B. Kirti and Prof. Appa Rao Podile, my doctoral committee members for periodically evaluating my progress in research and also for their interest, advice, support and graciously extending their lab facilities. I thank the present and former Deans, school of Life Sciences, Prof. R. P. Sharma, Prof. Aparna Datta Gupta, Prof. M Ramanadham and Prof. A. S. Raghavendra, the present and former Heads, Department of Plant Sciences Prof. Ch. Venkataramana, Prof. A. R. Reddy and Prof. P. B. Kirti for allowing me to use the facilities of the School and the Department. I sincerely acknowledge the infrastructural support provided by UGC-SAP, DST-FIST, DBT-CREBB, to the Dept. of Plant Sciences and the CSIR, Govt. of India for the research fellowship. I also thank DST-SERB, DBT and UPE-II for research funding to the lab. I thank Dr. Ragiba Makander, Department of Plant Sciences for her constructive suggestions for my research work.

I would also like to thank Dr. N. Sarla, Principle Scientist, DRR, Project director, DRR Hyderabad, NBPGR New Delhi and CRRI Cuttuck for providing me with the rice germplasms. I also extend my gratitude to Mr. Usham Dijen Singh for providing seeds of MustiLaiphou and MohemPhou landraces.

While compiling this thesis, I am glad to acknowledge everyone who has contributed to the build-up of my scientific personality and therefore, to the completion of this thesis. My heartfelt thanks are also due to my colleagues Zahoor Gul, Anirudh Kumar and Yaseen Bhat for their timely help and co-operation throughout my research period. I thank the M.Sc. project students

Sai kiran Reddey, Ranjeet and Sashidhar Reddy for their patience and timely help.

Words are inadequate in offering my thanks to my dearest friends Chaitu, Debo, Rama, Jyoti, Narmada (Gannapa), Prasu (Latchi), Sashi, Venkat (the all round social worker), Kishore, Madhu and Sumit for their unconditional love, support and encouragement throughout my difficult times. I owe thanks to you guys for all those hard and tiring trips to hospitals in Hyderabad with me. All thanks to you for the unforgetful memories we shared in the campus. Special thanks to Debo for her timely help in my research work. I wish to thank Sapana, Mickey and Kalpana for the cheerful and memorable moments we shared in the hostel.

I would like to extend my gratitude to my friend Abhay Kumar and my seniors Dr. Dilip Kumar, Naveen Singh and Dr. Pawan Shukla for their constant encouragement, support and help throughout my doctoral study. I also thank Malati, Ashutosh, Shiven and other lab members of Dr. Niyaz Ahmed lab; Santhosh and Jyoti of Dr. Insaf Ahmed Qureshi's lab for their timely help. I am thankful to all my batch mates and other research scholars of the School of Life Sciences for their timely support.

I thank my lab attendees Raju and Ranjit for their assistance in the lab and in the green house and I highly acknowledge non-teaching staff of our school and department. My thanks also go to Dr. Prasad Babu, DRR for his timely help during my initial study and Mr. Ashok Reddy for all his co-operation, who took care of rice plants in green house and field. Helping hands from the technical staffs of DRR are highly acknowledged.

Finally, I would like to express my heartfelt thanks to my beloved parents W. Binodlal and W. Sorojini, my brothers, sis-in-laws, nieces and nephew for their blessings, wishes, support, patience and constant encouragement throughout my research carrier.
Last but not the least; I thank the Almighty for his blessings to accomplish this task successfully.

Bimolata

ABBREVIATIONS

2DE: two dimensional electrophoresis
π: nucleotide diversity
µg: microgram
µ: microlitre
µM: micromolar
°C: degree centigrade/degree Celsius
AFLP: amplified fragment length polymorphism
ACN: acetonitrile
BB: bacterial blight
BLAST: basic local alignment search tool
bp: base pair
BSA: bovine serum albumin
cDNA: complementary DNA
C-terminal: carboxy terminal
CBB: coomassie brilliant blue
CHCA: a-cyano-4-hydroxycinnamic acid
CTAB: cetyltrimethylammonium bromide, hexadecyltrimethylammonium bromide
DEPC: diethylpyrocarbonate
DMSO: dimethyl sulfoxide
DNA: deoxy ribonucleic acid
Dnase: deoxyribonuclease
dNTPs: deoxy nucleotide triphosphates
DTT: dithiothreitol
EDTA: ethylene diamine tetra acetic acid
ETI: effector triggered immunity
FAO: food and agriculture organization
g: gram
h: hour(s)
Hd: haplotype diversity
HR: hypersensitive response
HRGPs: hydroxyproline rich glycoproteins
InDel: insertion and deletion
IPTG: isopropyl β-D-thiogalactoside
IRRI: International Rice Research Institute
JM: juxtamembrane
K_a: rate of non synonymous substitution
K_s: rate of synonymous mutation
kb: kilobase pair
kDa: kilodalton
L: litre
LB: luria-bertani
LD: linkage disequillibrium

LRR: leucine rich repeats
M: molar
MAS: marker assisted selection
ME: maximum evolution
MQ: milli Q
MP: maximum parsimony
MR: major resistance gene
MS/Ms: tandem mass spectrometry
MALDI-TOF: matrix-assisted laser desorption/ionization time of flight
mg: milligram
min: minute
ml: milliliter
mM: millimolar
N-terminal: amino terminal
NBS: nucleotide binding sites
$NH_4HCO_{3:}$ ammonium bicarbonate
NIL: near isogenic lines
UTR: untranslated region
OD: optical density
ORF: open reading frame
PAGE: polyacrylamide gel electrophoresis
PAMP: pathogen associated molecular pattern
PCD: programmed cell death
PCR: polymerase chain reaction
PMSF: phenylmethylsulfonylfluoride
PR proteins: pathogenesis-related proteins
PRR: pathogen recognition receptor
PTI: PAMP triggered immunity
Pv: pathovar
QTL: quantitative trait loci
R genes: disease resistance genes
RNA: ribonucleic acid
RNase: ribonuclease
ROS: reactive oxygen species
rpm: revolutions per minute
RT-PCR: reverse transcriptase-polymerase chain reaction
SAR: systemic acquired resistance
SDS: sodium dodecyl sulphate
sec: seconds
SES: standard evaluation system
SP: signal peptide
SNP: single nucleotide polymorphism
TAL: transcription activator-like

TFA: trifluoroacetic acid
TE: tris-EDTA
TILLING: target induced local lesions in genomes
Tm: melting temperature
TSS: transcription start site
Tris: tris-(Hydroxymethyl) aminoethane
TTSS: type three secretion system
UN: United Nations
USDA: United States department of Agriculture
UPT: upregulated by transcription factor
V: volts
X-gal: 5-bromo-4-chloro-3-indolyl β-D- galactoside
Xoo: *Xanthomonas oryzae*

List of figures:

List of tables:

Chapter One

INTRODUCTION

INTRODUCTION

Rice is inevitably an important staple food crop supporting half of the world population. Rice cultivation has gained importance worldwide not only as a source of dietary constituents, but also play a major role in socioeconomic development, providing food security, poverty alleviation, employment, sustainable development and maintenance of cultural heritage, etc. (Nguyen and Ferrero, 2006). It is also a favorable model crop plant among cereals for genetic improvement studies and functional genomics due to its small genome size and availability of the full genome sequence. In the context of modernization and globalization, cultivable land areas are shrinking day by day, which in turn affects rice productivity. It's a big challenge to meet the demand of ever increasing population across the globe especially, in the Asian subcontinent. Many challenges lie ahead of rice cultivars to achieve the required target in coming years and rice diseases are among one of the major challenges.

Rice plants are frequently under attack from various insects, pest and pathogens. Most of the elite rice cultivars and varieties with favorable taste and yield, such as Basmati, Sambha Mashuri, Swarna, other aromatic rice, etc. are very much prone to diseases. Bacterial blight (BB) is one of the major diseases of rice affecting as much as 20-50% yield loss in severe cases (Reddy 1989; Mew 1993; Adhikari et al. 1994). *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), a gram negative bacterium causes bacterial leaf blight disease in rice. It is a vascular disease resulting in systemic infection to the plant. To control the disease, use of chemical pesticides, biological control agents and host mediated resistance have been widely adopted. However, spraying pesticides have tremendous long term ill effects to the environment and consumers. In order to provide durable, economical and sustainable disease resistance in plants, host resistance is considered the best strategy for crop protection. Resistance of rice plants towards *Xoo* is mediated both by qualitative and quantitative resistance. Pathogen associated molecular pattern (PAMP) triggered immunity (PTI), a quantitative resistance, also the first layer of innate immunity in plants leads to basal defense response. Through this, antimicrobial compounds like proteases, chitinases, phytoalexins, etc. are secreted to the site of infection. It also triggers reactive oxygen species (ROS) secretion, ethylene biosynthesis, salicylic acid accumulation and callose deposition for cell wall strengthening close to the site of infection (Scheel 1998; Alvarez 2000; Heath 2000; Dangl and Jones, 2001). This

process results in hypersensitive responses (HR) which stop biotrophic pathogens such as *Xoo* from invasion or proliferation in host plants. Resistance gene (*R* genes) or disease resistance gene mainly confers qualitative resistance. These *R* genes interact directly or indirectly with specific *Avr* genes present in the bacterial pathogen thereby inducing disease resistance in plants against the pathogen through signaling pathways (Flor 1971). Some of the important *R* genes and their natural alleles have been identified from wild species and landraces. Exploring and exploiting the wild genetic resources promises to be a valuable base for developing BB resistance varieties. Many of the popular *Xa* genes being susceptible to diverse *Xoo* strains had been reported (Ochiai et al. 2000; Ram et al. 2011). Hence, identifying new form and sources of resistance genes is quite essential due to the presence of vast pathogen diversity. Some wild species were found to be highly resistant to different virulent *Xoo* strains, however, their resistance factors were unknown and need to uncover.

Foremost, the main purpose of the study is to identify valuable genetic resources possessing novel and desired form of *R* gene alleles from a collection of wild species, cultivars, landraces and introgression lines, etc. secondly, to analyze the different levels of polymorphism among alleles and study molecular evolution undergoing at major BB resistance genes, which may influence their interaction pattern with the pathogen and phenotype of the plant. Finally, to find candidate disease resistance genes present in disease tolerant wild species of rice. In order to develop well defined strategies for controlling BB, a comprehensive understanding of underlying molecular mechanisms of host resistance to pathogens are quite essential. In recent past, considerable progress in rice research for enhancing host resistance to BLB has been made. Till date, around 38 *R* genes for BB resistance had been reported (Chen et al. 2008; Natrajkumar et al. 2012). In nature, there must be vast allelic variants of these genes.

To begin with There are overall 23 species in the genus *Oryza* originating from diverse environmental locations (Vaughan et al. 2003). The collection of wild relatives available at the genebank are the reservoir of new, novel genes related to crop improvement and also favors the study of genetic diversity. The wild relatives are proven

to have enormous tolerance and resistance to biotic and abiotic stress, which is contributed by pool of resistance genes present in them. The availability of rice genome sequence of Nipponbare and Indica 93-11 and cloning of some important BB resistant genes will enable the rapid analysis of the genetic diversity at these loci over a wide range of germplasms and subsequent identification of novel and superior alleles through allele mining.

In simple terms, searching for an allelic variation is nothing but allele mining. A known locus is targeted for finding valuable and unknown alleles. In this approach, the nucleotide sequence of one genotype is used to isolate functional and valuable alleles from related genotypes (Latha et al. 2004). Variation in gene sequence may then be correlated with the disease resistance performance of the accession, and may well identify the best alleles for future transgenic experiments and breeding program. Nucleotide sequence variation in the naturally occurring alleles may bring changes at the level of gene expression and protein function which may produce functional changes and also result in phenotypic variation of a trait (Doebley and Lukens, 1998; Buckler and Thornsberry, 2002). Recent studies have shown that the disease resistant or susceptible phenotype is not only due to nucleotide variation in the coding region, however nucleotide changes in the non coding and regulatory sites also played an important role in determining the phenotype (Gu et al. 2005; Romer et al. 2009). Less information are available about the molecular evolution and the evolutionary relationships of resistance or susceptible genes in plants. Through natural allelic variation, we can identify allelic differences which contribute to the phenotypic variation of disease symptoms (Iyer-Pascuzzi et al. 2007). In the coming future, naturally occurring allelic variants will be one of the valuable resources for functional genomics study.

In this study, we present the phenotypic screening of collection of rice species, natural polymorphism analysis of four major *Xa* genes, *Xa21, Xa26, Xa27* and *xa5* at the sequence level and identification of putative candidate genes for BB resistance in wild *Oryza*. Till date, there are scanty reports on the genetic diversity of any of the bacterial blight *R* gene. *Xa21* is a broad spectrum bacterial blight resistance gene and code for a LRR threonine rich receptor like kinase (Song et al. 1995). *Xa26* is another dominant *R*

gene conferring race specific resistance against Chinese, Japanese and Korean *Xoo* strains. Both *Xa21* and *Xa26* genes had been mapped on long arm of Chromosome 11 of rice encoded LRR receptor kinase-type protein (Sun et al. 2004). *Xa26* is the same gene as *Xa3* (Xiang et al. 2006), which is known to be an important bacterial blight resistance gene in the japonica cultivar breeding in China (Xu et al. 2007). *Xa27* is a dosage dependent, differentially expressed *R* gene which confers resistance against *Xoo* PXO99A. It also showed broad-spectrum resistance against diverse strains of *X. oryzae* obtained from different Asian countries (Gu et al. 2004). This gene was first identified in *Oryza minuta* and mapped on chromosome 6 of rice. *xa5* confer race specific resistance against Phillipines *Xoo* race 1 (PXO86). The *xa5* gene is a mutant form of rice transcription factor OsTFIIAγ5 (*Xa5*) where there was a substitution variant of single amino acid V39E (Iyer et al. 2004).

Our analysis which was conducted on wild species and domesticated rice may help in understanding the natural selection and evolution of the important BB resistance genes to sustain under ever evolving pathogen constraint. Studying the pattern and structure of genetic variation of BB *R* genes will help us in understanding the evolutionary history of these genes. It will also help in identifying good source of resistance which may be further utilized for developing disease resistant varieties. With these backgrounds, the following objectives have been framed for our study:

1. Screening of bacterial blight resistant germplasms for allele mining.
2. Allele mining of *Xa21, Xa26, Xa27* and *xa5* nucleotide diversity analysis of their alleles.
3. Study of differentially up-regulated genes at transcript and translational level during *Xoo* infection in resistant wild rice.


Chapter Two
Literature Review

The literature related to the present study on the allele mining of candidate disease resistance genes, their nucleotide diversity, expression analysis and proteomics has been reviewed under the following heads:

2.1 Major Yield Constraints
2.2 Major Diseases of Rice
2.3 Management of Bacterial Blight
2.4 Major Bacterial Blight Resistance genes
2.5 Utilization of Natural Diversity for Disease Resistance
2.6 Defense Responsive Genes for Disease Resistance
2.7 Transcriptome Analysis of Disease Resistance Genes
2.8 Proteomics Analysis of Disease Resistance Genes

Rice is an economically important staple food crop feeding more than half of the world population. Rice cultivation was believed to start in many countries over some 6500 years ago (Gnanamanickam 2009) and it is a major contributor of Indian economy with a GDP of 15% (Icrisat, 2008). Across the globe, rice consumption is highest in Asia. "The average per capita consumption of rice is higher than 80 kg/person per year" (Gnanamanickam, 2009). India is both a chief producer and consumer of rice. Among the top rice producing countries, India stands second in the world after China. The estimated annual production of rice in India is 131,274,000 tonnes (Global Rice Production, FAO 2011) (Fig.2.1). It is estimated that by the year 2035, an additional 116 million tons of rice will be required (USDA report, 2012). Statistics during 2011-12 showed that the global rice field under cultivation was 1, 60,000 HA and still 925 million people were starving in 2010 (UN, FAO) (Fig.2.2).

2.1 MAJOR YIELD CONSTRAINTS

Rice cultivation faces many biological and non biological constraints during the course of their production. The overall major constraints are biotic and abiotic stresses. Biotic stress includes insects, pests, rodents, birds, weeds, brown plant hopper and

diseases caused by bacteria, fungus, viruses and nematodes etc. An annual yield loss of rice due to insects and diseases was estimated up to 25%. Among insects, stem borer and common rice insects were most destructive while, rice blast caused by *Pyricularia oryzae* or *Magneporthe grisea* and bacterial blight caused by *Xanthomonas oryzae* are the two most widespread and serious diseases of rice (Nguyen and Ferrero, 2006). Abiotic stresses are drought, improper irrigation, submergence, salinity, cold, zinc deficiency, poor soil quality and its fertility. Repeated and long term cultivation of rice in the same field resulted in increased salinity in irrigated areas (McDonald 1994; Pingali and Rosegrant, 1996; Aguilar et al. 1997). Other factors contributing to yield constraints are socio-economic and technical constraints which include low input, inappropriate plant spacing, late sowing and selection of wrong cultivars (Jha et al. 2012). More upstream research is needed to address the yield loss caused by diseases and pests such as blast, sheath blight, bacterial blight, etc.

2.2 MAJOR DISEASES OF RICE

Rice cultivation comes under the scrutiny of many pests and pathogens affecting its yield and productivity. It is also one of the main factors which hamper in achieving the set target of production. The major diseases of rice are bacterial blight (BB) (bacterial disease), blast, sheath blight and sheath-rot (fungal diseases) and rice tungro disease (viral disease) (Fig. 2.3).

2.2.1 Bacterial Diseases of Rice

Major bacterial diseases of rice are bacterial leaf blight (caused by *Xanthomonas oryzae* pv. *oryzae,* Fig.2.3a) and bacterial leaf streak (caused by *Xanthomonas oryzae* pv. *oryzicola,* Fig.2.3b). Other bacterial diseases are foot rot (*Erwinia chrysanthemi*), strain rot (*Burkholderia glumae*) and sheath brown rot (*Pseudomonas fuscovaginae*). Bacterial blight is one of the oldest known and serious diseases of rice, which was first discovered by farmers of Japan in 1884 (Tagami and Mizukami, 1962). Considerable progress on the research of this disease had been made and reviewed worldwide (Mizukami and Wakimoto, 1969; Ou 1985; Mew 1987, 1989, 1993;

Gnanamanickam 1999; Subramoni et al. 2006). BB is a disease of vascular system where the pathogen enters through natural openings such as stomata, hydathodes and wounds (Mew 1987). After entering they multiply in the xylem vessels, thereby blocking the nutrients flow in the plants and spreading to other parts of the plant. The pathogen take up the nutrients from plants resulting in drying and wilting of rice leaves, thus reducing the net photosynthetic leaf area, 1000 grain weight, etc. and finally the yield (Mew, 1987). The severity of the disease is more in humid tropical and subtropical region where temperature ranges between 28ºC to 34ºC. The symptoms of this disease are observed mainly at the maximum tillering stage and there are two different phases of symptoms, the kresek phase and leaf blight phase (Fig.2.4). Kresek happen mostly at the seedling stage, the whole plant turn yellow and wilt. It is more destructive than blight phase and result in total crop failure as the whole plant die (Ou 1985; Mizukami and Wakimoto, 1969). Leaf blight phase has yellow lesion in the leaves which gradually turn brown and ultimately wilt. This phase is most commonly observed in the Indian fields. The manifestation of BB increases as the pathogen is spread in the form of bacterial ooze via rain, irrigation water, flood and wind. The total crop loss due to BB accounts from 10% to 50% (Ou 1985; Mew et al. 1993), while in susceptible cultivars it is estimated between 74-81% (Srinivasan and Gnanamanickam, 2005).

Unlike bacterial blight, the damage caused by leaf streak is minimal. The estimated loss is almost 32% in 1000 grain weight (Opina and Exconde, 1971; IRRI 1983; Nyvall 1999). Spreading of disease and pattern of infection is same as that of BB in case of bacterial leaf streak. Bacterial leaf streak is commonly found in Southern China, Taiwan, India, Southeast Asian countries, and West Africa but absent in temperate countries. Other bacterial diseases are less reported and minimal in nature. Scientific findings are comparatively very less to BB.

2.2.2 Fungal Diseases of Rice

There are numerous fungal pathogens causing different diseases in rice. Some of the diseases along with the causal organism are listed in Table 2.1. One of the major fungal diseases affecting rice productivity is rice blast caused by *Magneporthe grisea*

(Fig.2.3c). It causes severe damage during seedling stage. The annual yield loss worldwide due to blast disease is up to 50% (Scardaci et al. 1997). Yield losses between 50-85% were also reported in the Philippines (IRRI). This disease can be controlled by cultivation of resistant varieties and use of fungicides. Another widespread and destructive fungal pathogen is *Rhizoctonia solani* which causes Sheath blight (ShB) (Fig.2.3d) in rice. In tropical and temperate regions, an annual yield loss of 20–50% on average due to Sheath blight infection is reported (Miruta 1956; Boyatee and Lee 1979; Rajan 1987), while in susceptible cultivars the yield loss may go above 50% (Lee and Rush, 1983). In India, the annual yield loss due to the disease is up to 54.3% (Chahal et al. 2003). Presence of natural resistance gene is limited for this disease which limits its management. These diseases not only cause yield loss of rice, but also reduces the grain quality (Singh et al. 2011).

2.2.3 Viral and other Diseases of Rice

Study on rice viruses is quite less and however there has been considerable improvement in the recent past. Rice viruses are mostly transmitted through leafhoppers. Rice tungro virus (Fig.2.3e) is also one of the destructive diseases, which causes significant yield loss of rice in South East Asian countries (Gananamanickam 2009). Tungro disease is caused by rice tungro baciliform virus (RTBV) and rice tungro spherical virus (RTSV). Yield loss by this virus may be up to 100% when infection occurs at the early growth stage (IRRI, 2003). Other viral diseases are Rice stripe necrosis virus (RSNV), Rice crinkle disease and Rice yellow mottle virus (RYMV) (Fig.2.3f). These viral diseases may be controlled through integrated pest management strategies (Aboa and Fadhila, 2003). Other miscellaneous rice diseases are bronzing of leaf due to zinc deficiency, cold injury at low temperature, kernel spotting due to feeding injury by rice stink bug (*Oebalus pugnax*) etc.

2.3 MANAGEMENT OF BACTERIAL BLIGHT

Management and controlling of BB had been a major task lying in front of rice scientists across the globe. We need to formulate strategies in order to obtain sustainable

BB resistance which is quite a challenging task as new strains and isolates of the pathogen keep emerging due to natural mutation and evolution. Some of the control methods which have been widely adopted were use of chemical pesticides, bio control agents, breeding for resistance genes and exploitation of host mediated resistance. Among these, use of chemical pesticides is potent yet very harmful for the environment and ecology, so, the most preferable option is going for the bio-control agents, use of breeding lines with resistance genes and exploitation of host mediated resistance. In recent times, scientists have been working tremendously to achieve durable, sustainable and broad-spectrum resistance in rice.

2.3.1 Biological control agents

Biological control agents are generally the natural enemies of disease producing bacteria, fungi, viruses or any other pest. They are living organisms either feeding on the pathogen or inhibit the growth of pathogen population by producing toxins against it. Use of bio-control agents against BB was first reported by Mew and Rosales (1986). Later on, many groups reported on its application against the disease. Sakthivel and Mew, (1991) treated bacteriocin-producing bacteria with BB infected rice plants and observed reductions of bacterial blight incidence up to 31–99% in greenhouse tests. Sudhir et al. (2002) isolated 65 antagonistic bacteria out of 227 bacterial isolates from different locations in Indonesia. In India, many gram positive bacteria which were antagonistic against *Xoo* and produced bacteriocin were isolated by Gnanamanickam and group (2003, 2006 and 2009). *Pseudomonas fluorescens* and bacillus strains producing 2, 4-diacetylphioglucinol (DAPG) suppressed the growth of (*Xoo*) in IR24 at field trial and net house experiments (Vasudevan 2002; Velusamy and Gnanamanickam, 2004; Velusamy et al. 2006). Ji et al. (2008) isolated a novel bacterial strain of *Lysobacter antibioticus* 13-1 from the rice rhizosphere and was found to suppress bacterial blight (BB). Another natural bio-control agent reported to suppress the growth of (*Xoo*) is the epiphytic populations of *E. herbicola* present in rice leaf surfaces. It act as bio-control agent by producing an acid, which lowered the pH of rice leaves (Hsieh and Buddenhagen, 1975; Santhi et al. 1987). List of biocontrol agents for different rice diseases are given in Table 2.2.

2.3.2 Host mediated disease resistance

During the course of evolution plants had developed abilities to defend themselves against pathogens. There are two types of resistance by which plants confer resistance, one is qualitative and another is quantitative resistance. Qualitative resistance is governed by major disease resistance (MR) genes generally which function single handedly. This type of resistance is generally stronger and mostly pathogen and race specific in nature. On the other hand, quantitative resistance is conferred by multiple resistance genes also known as quantitative trait loci (QTL) which is less effective, but usually pathogen or race non specific (Kou and Wang, 2010). As quantitative resistance is provided by multiple genes, this type of resistance is more durable and broad spectrum. Pathogen invasion in plant is resisted by two types of innate immunity, pathogen associated molecular pattern (PAMP) triggered immunity (PTI) and effector-triggered immunity (ETI) (Jones and Dangl, 2006). Qualitative resistance is mediated by both PTI and ETI. PAMPs are recognized by plasma membrane localized plant pattern recognition receptors (PRRs) and initiate PTI. PTI is mostly weaker than ETI and it provides only the basal level of resistance which is race non specific (Thomma et al. 2011). ETI on the other side is conferred by the group of *R* genes which is pathogen-race specific through direct or indirect interaction with the pathogen's effector (Flor, 1942). Downstream of *R* genes and PRR genes mediated resistance lays the defense response genes in the signaling pathways where some of these defense responsive genes emerged as major contributors of quantitative disease resistance (Zhang et al. 2013).

2.3.3 Management of Bacterial Blight through breeding and transgenic approaches

Till date, breeding of cultivars with introgressed gene for bacterial blight resistance promises to be the best option for the control of bacterial blight disease. It is the most effective, economical, environmentally safe and ecologically sound method for providing sustainable disease resistance against *Xoo* (Khush et al. 1990, Singh et al. 2001, Singh et al. 2011). In recent times, breeding for resistance to bacterial blight is rapidly progressing with generation of numerous hybrids. Several near isogenic lines (NILs) carrying individual *Xa* genes had been developed by transferring resistance genes

from their natural sources (both cultivars and wild relatives) through breeding. These NILs served as donor lines of these genes to elite cultivars. In many instances, rice varieties carrying single resistant genes were found ineffective against various *Xoo* strains such as *Xa4* being susceptible to *Xoo* strains in many parts of Indonesia, India, China and Phillipines (Mew 1987, 1989; Mew et al. 1993) and IRBB21 being susceptible to virulent *Xoo* strains isolated from Punjab, other parts of India, Korea, Nepal and China (Goel et al. 1998; Lee et al. 1999; Adhikari et al. 1999; Marella et al. 2001; Zeng et al. 2002; Ji et al. 2003; Muralidharan et al. 2004). Afterwards, the method of integrating two to three resistance genes together in a single plant also known as gene pyramiding was introduced through marker assisted selection. Gene pyramiding strategy was first adopted by Yoshimura et al. (1995) combining *Xa4/xa5* and *xa5/Xa10* then followed by Huang et al. (1997) where four resistance genes *Xa4*, *xa5*, *xa13* and *Xa21* were put together and found wider spectrum and higher level of resistance than single gene possessing plants. This method had been successful in providing disease resistance against multiple *Xoo* strains in elite cultivars. Some of the successful examples are Singh et al. (2001) pyramided three BB resistance genes, *xa5*, *xa13* and *Xa21*, into cv. PR106, widely grown rice in Punjab, India after which successfully provided resistance to 17 *Xoo* isolates prevalent in Punjab. Four genes *Xa4*, *xa8*, *xa13* and *Xa21* were put together in popular aromatic high yielding rice variety Pusa Basmati 1. Through marker assisted selection BB resistant PB1 retaining basic Basmati quality and desirable agronomic traits were developed (Joseph et al. 2004). Two elite cultivars of southern India Samba Mahsuri and Triguna were made BB resistant successfully by pyramiding *xa5*, *xa13* and *Xa21* together (Sundaram et al. 2008, 2009). Recently Singh et al. (2011) developed improved PB1 through marker assisted backcross breeding, which combine three different genes (*xa13*, *Xa21* and *Pi54*) and one QTL (qSBR11-1) in one plant for BB, blast and ShB disease resistance.

Another approach is raising transgenic lines carrying resistance genes. For the first time *Xa21* carrying transgenic lines in japonica rice T-309 were developed by Wang et al. (1996). Later on, *Xa21* was introduced into several varieties such as IR72, MH63, and IR51500 (Datta, 2002; Tu et al., 1998; Tu et al. 2000). Generation of transgenic requires multiple backcrossing which is very tedious. Successful generation of transgenic

rice carrying *Xa21* gene was carried out by Wang et al. (1996). *Xa21* transformed into an elite restorer line has shown resistance to BB while retaining its original traits (Zhai et al. 2002). Swamy et al. (2006) also raised transgenic PB1 lines carrying *Xa21* and were found susceptible to six *Xoo* isolates, however transgenic lines carrying more than one *R* gene were resistant to different isolates. Datta (2004) carried out field trials of *Xa21* transgenic rice in India, Phillipines and China and found satisfactory result. Transgenic lines carrying one or the other *Xa* genes were developed for functional characterization of the gene; however, none has been introduced for commercial use as deregulation of these varieties are still pending. Major drawbacks of transgenic are, it requires enormous resources and acceptance by public is low. Hence, conventional breeding assisted by marker selection is still popular and an option of choice for developing resistant lines to control BB disease.

2.4 MAJOR BACTERIAL BLIGHT RESISTANCE GENES

For bacterial blight, most of the genes conferring resistance to the diseases are qualitative major resistance genes. Till date 38 *R* genes for resistance against BB have been identified from different wild species and cultivated rice (Chen et al. 2008; Natrajkumar et al. 2012). Out of this, 14 of them were found to be recessive in nature (Kou et al. 2013). Among these identified genes only six *R* genes have been cloned till date (*Xa1*, *Xa21*, *Xa3/Xa26*, *Xa27* and *xa5*, *xa13*) and isolated using map based cloning strategies (Yoshimura et al. 1998; Sun et al. 2004; Iyer and McCouch, 2004; Chu et al. 2006; Song et al. 1995; Gu et al. 2005). Among these genes, *Xa21*, *xa13* and *xa5* have been extensively used for the development of BB resistant cultivars and high yielding hybrids. Chromosomal locations of different genes along with their linked markers are presented in Figure 2.5.

2.4.1 *Xa21*

A major breakthrough in BB management research came with the cloning of *Xa21* gene. *Xa21* is the first BB *R* gene cloned with the help of map based positional cloning strategy by Song et al. (1995) which is 3.9 Kb in size. It was mapped on the long

arm of chromosome 11 of rice. It is a broad spectrum disease resistance gene, which was introgressed from wild species *O. longistaminata* into *O. sativa* (IR24) background (Khush et al. 1990). This gene codes for receptor like kinases with hydrophobic N terminal and four major domains, LRR, transmembrane, juxtamembrane (JM) and intracellular kinase domains. It confers race specific resistance against *Xoo* race 6 (Song et al. 1995*)*. Ronald et al. (1992) developed a PCR based STS marker pTA248, which was efficiently used for marker assisted selection of *Xa21* gene. *Xa21* belongs to multigene family designated as A1, A2, B (*Xa21*), C, D, E and F (Song et al. 1995 and Wang et al. 1996). Expression analysis of *Xa21* transcript shows that this gene is constitutively expressed at all time points, but not expressed in susceptible IR24 (Century et al. 1999). It further suggested that its expression is independent of developmental stages, wounding or infection and not correlated with the expression of *Xa21* disease resistance (Century et al. 1999). *Xa21* mediated disease resistance is negatively regulated by transcriptional regulator XB10/WRKY62 (Peng et al. 2008) and triggered by the binding of sulfated peptide axY^{S}22 from AX21 to the LRR domain of XA21 protein. AX21/AvrXa21 is a type I effector protein secreted by *Xoo* into rice plants through type I secretion system (Lee et al. 2006, 2009). Functionally, this gene is involved in pathogen recognition and response. Based on its characteristic features, XA21 is considered a PRR which provides high level of PTI, yet race specific resistance (Wang et al. 1996; Zhao et al. 2009). Due to its high resistance against various *Xoo* strains, *Xa21* had been widely adopted for development of many resistant varieties incorporation with other genes through gene pyramiding method.

2.4.2 *Xa26*

Xa26 is another dominant *R* gene conferring race specific resistance against Chinese, Japanese and Korean *Xoo* strains. *Xa26* gene was mapped on Chromosome 11 of rice within genomic region 20 kbp in length (Yang et al. 2003). The presence of *Xa26* in rice was first identified in Chinese *O. sativa* cultivar Minghui 63. This gene exhibited broad resistance both at seedling and adult stage. *Xa26* gene, conferring race-specific resistance to *Xoo*, also encodes LRR receptor kinase-type protein (Sun et al. 2004). Both *Xa26* and *Xa3* had been reported to be same form of gene (Xiang et al. 2006). This gene

is known to be an important bacterial blight resistance gene in the japonica cultivar breeding in China (Xu et al. 2007). *Xa26* belongs to a multigene family, including four members and encoding LRR RK family with all the members clustered in tandem on rice chromosome 11 (Sun et al. 2004, 2006). The gene family consists of MRKa, MRKb (known as *Xa26*), MRKc and MRKd. It has been predicted that MRKa and MRKc are intact genes, whereas MRKd is a pseudogene carrying one inframe stop codon, two frameshift sites, and two inserts. Among the three intact genes, *Xa3/Xa26*, MRKa and MRKc, only *Xa3/Xa26* mediated disease resistance to *Xoo* under their native promoters. The expression pattern of *Xa21* and *Xa26/Xa3* are functionally same as the level of resistance in plants increases as their expression increases. Their expression is developmentally regulated and maximum resistance in plants is obtained at their adult stage (Zhao et al. 2009; Century 1999; Li et al. 2012).

2.4.3 *Xa27*

Xa27 is a semi dominant *R* gene conferring resistance against *Xoo* PXO99A. It showed broad-spectrum resistance against multiple strains of *X. oryzae* prevalent in different Asian countries, including two Indian strains A3842 and A3857 (Gu et al. 2004). This gene had been mapped on chromosome 6 of rice and was identified from wild rice *Oryza minuta* (BBCC). The function and mechanism of action of this gene have been reported by Gu et al. (2005, 2009) and Romer et al. (2009) citing it to be differentially expressed gene. *Xa27* gene is activated by type III TAL effector AvrXa27 and hence, confers resistance against only those incompatible *Xoo* strains harboring avirulence gene *avrXa27*. The expression of this gene is induced by the binding of AvrXa27 effector to the 16 bp UPT box in the promoter region of *Xa27*. This interaction is very specific and any mutation in the UPT box leads to non expression of the gene. *Xa27* is an intronless gene, which encodes a protein of 113 amino acid residues. The recessive allele *xa27* identified from IR24 share same coding sequence, however due to deletion and substitution in the promoter, this allele is non expressed. The main polymorphic regions between these two alleles lay in the upstream of transcription start site (TSS) and TATA box. There is a deletion of three nucleotides GAA at -51 position

in *xa27* promoter (Gu et al. 2005). XA27 protein is unique and different from previously reported *R* genes. It is predicted to code for a trans-membrane protein (Gu et al. 2005).

2.4.4 *Xa1*

Another less explored dominant gene is *Xa1* which is the only NBS-LRR gene cloned for BB resistance. It conferred high level of resistance to the Japanese *Xoo* race 1 strain T7174 and was first reported and mapped on chromosome 4 by Sakaguchi (1967). It was tagged with an RFLP marker XNpb235. Later, Yoshimura et al. (1998) reported its positional cloning from *Xa1* donor lines IRBB1 and Kogyoku. It encoded an amino acid sequence of 1802 residues. Southern hybridization analysis suggests that *Xa1* is a single copy gene and its homologous sequence is present in susceptible isogenic line IR24 at the same locus. Its expression is induced both by wounding and infection, which increases the resistance towards pathogen (Yoshimura et al. 1998). The mechanism behind the induction of this gene is still unclear.

2.4.5 *xa5*

xa5 is a recessive *R* gene mapped on chromosome 5 which confer basal level of race specific resistance against Phillipines *Xoo* race 1 PXO86 (Blair et al. 2003; Iyer and McCouch, 2004). This gene is flanked by two microsatellite markers RM603 and RM607. The recessive form of the gene is widely present in Aus-Boro lines (Iyer and McCouch, 2004). The *xa5* gene is a mutant form of rice transcription factor *OsTFIIAγ5* (*Xa5*) where there was a substitution variant of single amino acid V39E (Iyer et al. 2004). The same form of mutation was detected in all Aus-Boro lines possessing the recessive *xa5* gene. This transcription factor is coded by 106 amino acids. Both the dominant and recessive genes are constitutively expressed in different tissues of the plant and not influenced by wounding, infection or compatible/incompatible interactions (Iyer and McCouch, 2004; Jiang et al. 2006). The amino acid polymorphism between susceptible and resistant allele also didn't influence its expression. A set of CAPs marker was also developed based on the SNP present in the coding sequence which can be used efficiently for easy identification of desired resistant alleles present in lines, cultivars and

accessions. Rice has another transcription factor in chromosome 1 known as *OsTFIIAγ1*. *OsTFIIAγ1* is also a candidate gene targeted by TAL effectors and involved in mediating host susceptibility like its homolog *OsTFIIAγ5* (Sugio et al. 2007; Iyer-Pascuzzi et al. 2008). *xa5* is highly homologous to the small subunit of TFIIA and its 3 dimensional structure was modeled using TFIIA small subunit as a template. The predicted 3-D structure of *xa5* after its superimposition with Xa5 and TFIIA shows that due to the substitution of V39E, a small change occurred in the third helix domain of the xa5 protein, a supposed conservative substitutable site (Jiang et al. 2006). Gu et al. (2009) also hypothesize the role of dominant *Xa5* in activating type III effector AvrXa27 for inducing the transcription of *Xa27* gene in *Xa27* possessing plants.

2.4.6 *xa13*

xa13 gene is another important fully recessive, constitutively expressed bacterial blight resistance gene consisting of five exons and encodes a plasma membrane protein of 307 amino acids (Chu et al. 2006). It confers resistance against Philippines race 6 (PXO99) of *Xoo*. This gene is located on the telomeric end of long arm of chromosome 8. It was first identified by Ogawa et al. (1987) from BJ1 group and isolated from isogenic line IRBB13 by Chu et al. (2006). Its dominant allele *Xa13* is pathogen responsive, susceptible gene (Yuan et al. 2009, 2010) and plays an important role in pollen development. *xa13* and *Xa13* genes encode proteins which are identical with 1–3 amino acid substitutions (Chu et al. 2006). Susceptible *Xa13* shows 68 percent similarity to the product of nodulin MtN3 (Chu et al. 2006) hence, it has been named as Os8N3 (Yang et al. 2006). This gene is transcritionally activated by type III TAL effector PthXo1 of *Xoo* whose corresponding cis regulating element, UPT_{PthXo1} lies in –80 to –56 of *Xa13* promoter, P_{Xa13} (Yang et al. 2006; Romer et al. 2010). Expression of this gene leads to the development of the disease after infection with incompatible strain, however mutation studies showed that suppressing its expression greatly enhanced the resistance of the plant against the strain. The difference in the promoter sequence between the dominant *Xa13* and recessive *xa13* is solely responsible for different response against the pathogen rather than the amino acid difference in the encoded protein. There is an insertion of repetitive sequences of 243bp near the TATA box of

xa13 in IRBB13 (Yang et al. 2006). This gene is unique among all *R* genes as it is being regarded as a mutant of pathogen induced expressional loss of function. *xa13* and all its alleles are unresponsive to TAL effector PthXo1 (Yang et al. 2006) which is due to the presence of mutated UPT_{PthXo1} (Yuan et al. 2010). Hence, *xa13* is able to resist PXO99A by not responding and not interacting with the pathogen.

In the present study, we have used the gene information on *Xa21, Xa26, Xa27* and *xa5* and reported their nucleotide diversity among species and varieties. Using these informations, we have also analyzed differential expression of few of the alleles of these genes at transcript level. For bacterial blight disease information on diversity analysis and allele mining of resistance gene is limited. Recently Utami et al. (2013) identified *Xa7* functional alleles from an Indonesian local rice population. Hittalmani et al. (2013) with the help of SSR markers could identify resistant and susceptible alleles of *Xa21* from local Indian landraces and traditional cultivars; however, there functional characterization has not been carried out.

2.5 UTILIZATION OF NATURAL DIVERSITY FOR DISEASE RESISTANCE

Genetic variability is the foundation for crop improvement. Exploring new resistant sources has become a thriving field to overcome newly evolving pathogenic strains of *Xoo*. Cultivation of resistant cultivars has been an efficient method of BB management through host resistance. However, due to the presence of strong selection pressure and rapid evolution of the pathogens, host resistance provided by *R* genes was overcome. Hence, exploring and exploiting the available pool of genetic variation in landraces, traditional varieties and wild relatives of cultivated rice will help in the development of effective and durable disease resistant varieties.

2.5.1 Genetic resources and natural variation

The rice genetic resource in international rice Genebank and different rice cultivating countries is the basis of world food security. Genus *Oryza* has 23 species, including two cultivated species *Oryza sativa* (Asian) and *Oryza glaberrima* (African)

(Vaughan et al. 2003). These 23 species have been classified into seven species group based on their genome content (Table 2.3). The first species group known as *O. sativa* complex comprised nine species, including *O. barthii, O. glaberrima, O. breviligulata, O. glumaepatula, O. longistaminata, O. meridionalis, O. nivara, O. rufipogon* and *O. sativa* which have AA genome (Morishima et al. 1992). The genetic relatedness among these species was confirmed by restriction fragment length polymorphisms (RFLPs) of nuclear DNA (Wang et al. 1989), isozymes (Second 1991), cpDNA (Dally and Second, 1990) and mtDNA (Second and Wang, 1992), and amplified fragment length polymorphism (AFLP) (Aggarwal et al. 1999). Genetically at the molecular level, *O. sativa* is less diverged than *O. rufipogon* while, at the phenotypic level *O. sativa* show vast diversities than *O. rufipogon* (Morishima 2001). Gene flow is believed to play important role in diversification of cultivated plants. In rice, gene flow mainly occurs from inbreeding cultivated races to out-breeding wild races. In wild the out-crossing rate ranged from 10% to 60% (Oka 1988). It has been found that many of the novel disease resistance genes were identified from wild species of *Oryza* and also Zhang and Ling (1994) and Zhang et al. (1998) reported *Oryza rufipogon* as a reservoir of many genes for bacterial blight resistance. Wild rice has been playing a key role in breeding for resistance and other yield enhancement traits in rice (Vaughan et al. 1994; Brar and Khush, 1997). Wild relatives are rich repositories of genes for crop improvement. Some of the bacterial blight resistance genes such as *Xa21* (Khush et al.1990), *Xa27* (Gu et al. 2004), *Xa30/Xa38* (Cheema et al. 2008), *Xa31t* (Wang et al. 2008), *Xa33* (Natrajkumar et al. 2012), *xa34t* (Chen et al. 2011) and *Xa35t* (Bin et al. 2010) have been identified from wild species *O. longistaminata*, *O. minuta, O. nivara, O. rufipogon* and *O.minuta,* respectively. Landraces maintained by traditional farmers too posses tremendous genetic variability and these resources have been underexploited (Ram et al. 2007). For instance, *Xa1*, *Xa4, xa5*, *xa13* and *Xa26* had been identified from landraces. Exploring the diversity of this underexplored wild rice and landraces genotypes for different alleles for these known resistance genes at the molecular level may lead to identification of superior/functional alleles which could be deployed for breeding resistance.

2.5.2 Single Nucleotide Polymorphism (SNPs)

Single nucleotide polymorphism or commonly known as snips are genetic variation occurring at single base or nucleotide in a gene sequence, non coding region, coding region. SNPs are generally point mutation which may lead to synonymous; non synonymous or nonsense mutation. SNPs are predominantly present in almost all organisms. This form of polymorphism governs the main cause of diversity and allelic variation in nature. They are generally used for marker development, fine mapping and cloning of QTL, study of genetic variation and population structure of many crop plants (Bhattramakki et al. 2002; Jones et al. 2009; Esteras et al. 2013). SNP based markers have gained much importance in molecular breeding of crop plants due to their abundance in genomes (Mammadov et al. 2012). Some SNPs are responsible for differentiation between resistant and susceptible alleles of many disease resistance genes in plants. Development of rapid sequencing technologies and computational tools facilitate genome wide identification of SNPs. Validation and large scale genotyping of SNPs are essential to determine the functional significance of candidate genes for important traits such as biotic and abiotic stress resistance (Parida et al. 2012). Validated SNPs can be further used to develop functional markers. It has wide applications in genome wide association studies and marker assisted selection (MAS). They are being identified in whole genome and constantly used in identification and screening of resistance gene loci, QTL, etc. (Kim et al. 2010; Fang et al. 2010; Leonforte et al. 2013).

2.5.3 Approaches and applications of nucleotide diversity studies

Identifying the SNPs/InDels in the candidate gene, which contribute to the different alleles will help in finding functionally relevant phenotypes of interest. Most conveniently, this can be done with two different approaches: (1) EcoTILLING and (2) PCR and sequencing based allele mining (Ramkumar et al. 2010). EcoTILLING is technically similar to TILLING where instead of artificially induced mutation; here mutations are detected from naturally occurring alleles (Ramkumar et al. 2010). It also involves heteroduplex formation, enzymatic cleavage, running on LiCor gel for SNP detection and final confirmation through sequencing (Till et al. 2003). The availability of

enormous genome sequences of various important plant species has enhanced the adoption of sequencing based allele mining. Sequencing based allele mining involves amplification and sequencing of alleles in diverse accessions through PCR followed by identification of nucleotide variation by sequence analysis. This type of study helps to analyze individuals for haplotype structure and diversity to genetic association studies in plants. The first and the foremost step in the process of allele mining are germplasm collection and selection of desired resistant accessions. Some of the promising applications of allele mining are discovery of novel alleles/favorable alleles of important agronomic traits, genetic diversity analysis of available collections of genetic resources in the gene banks (Kaur et al. 2008), provide insight into molecular basis of novel trait variations, identification of new haplotypes with relation to the resulting phenotypic changes, development of allele specific marker for marker assisted selection, evolutionary study of alleles etc., (Ramkumar et al. 2010). With such promising aspects, we intend to utilize this approach and exploit the natural resources for identifying candidate disease resistance genes for BB disease. An overview for identification and functional validation of *R* genes through allele mining is illustrated in Fig. 2.6.

2.5.4 Allele mining of BLAST resistance genes

Blast is also one of the serious diseases of rice plants. Allele mining of blast resistance genes in different wild species and cultivated rice had been reported by various groups. Ramkumar et al. (2010) reported mining of two blast resistance genes *Pikh* and *Pita* known to confer broad-spectrum resistance to blast in India. Using sequence based strategy; naturally occurring allelic variation was identified from 27 landraces of northeastern India and 127 accessions of different wild *Oryza* species collected from IRRI. The plant materials were screened for disease resistance with *NLR-1,* most virulent pathogen race in southern India. In an inter and intraspecific DNA variation study on *Pita* gene in *O. rufipogon, O. meridionalis* and *O. officinalis* conducted by Yoshida and Miyashita (2009), resistance *Pita* allele had lower level of sequence variation than susceptible allele. Excess of non conservative amino acid replacement was observed between *O. rufipogon* and *O. officinalis Pita* leucine rich domain. The origin of *Pita* in cultivated rice (*indica* and *aus*) and weedy rice has been

suggested from a single source and introgression of susceptible *pita* from cultivated to weedy rice by outcrossing (Lee et al. 2011). Molecular diversity pattern of *Pita* in US weedy rice was found to be more similar to cultivated rice than *O. rufipogon*. Huang et al. (2008) also demonstrated low level of nucleotide polymorphism of *Pita* gene in *O. rufipogon*. Other than *Pita*, allele mining and polymorphism study of *Pi54* formerly known as *Pik*^{h} has been reported, which showed diverse nature of *Pi54* allele in land races and wild species of rice (Kumari et al. 2013). A similar kind of study was also conducted for rice blast resistance gene *Pid3* encoding a nucleotide-binding-site leucine-rich repeat (NBS-LRR) protein. They identified another resistant allele of this gene named as *Pid3-A4* from *O. rufipogon* (Lv et al. 2013). The availability of rice genome sequences of two cultivars of *O. sativa*, Nipponbare (IRGSP 2005) and *indica* cv. 93-11(BGI) and also introduction of fast and low cost sequencing technology has increased the provision of finding new alleles of various disease resistance genes from the vast genetic pool. Utilizing these resources, gives us an opportunity to explore and find superior, durable and broad-spectrum resistance alleles of the underexplored BB resistance genes. Based upon these information's, we have analyzed the sequence variation in the natural alleles of BB resistance genes *Xa21, Xa26, Xa27*and *xa5* in our present study.

2.5.5 Allele mining studies of *R* genes in other crop plants

Allele mining is the identification, of superior and favorable alleles of agronomic importance. Its foremost target is the identification of sequence polymorphisms for a given gene in naturally occurring populations. The subsistence of natural variation of various genes available in nature provides the platform to identify multiple allelic variations which result in different phenotypic patterns of resistance (Iyer-Pascuzzi et al. 2007). Alleles are also responsible for the trait and variation in other genotypes. In true sense, allele mining is not only identification of single nucleotide polymorphism (SNP) or insertion and deletion (InDel) at coding sequences of the gene but also include analysis of non-coding and coding regions of the genes in addition to detect sequence variations in the regulatory regions of agronomically valuable genes (Rangan et al. 1999; Latha et al. 2004, RamKumar et al. 2010). The method of allele

mining had been widely adopted for studying the diversity of different *R* genes of various crop plants. Seven new resistance alleles of powdery mildew resistance gene '*Pm3*' in wheat was identified by targeted approach for molecular utilization of germplasm collections available in gene bank. Rich resource of new functional alleles was found in the landraces (Bhullar et al. 2009). *Rpi-blb* homologues were identified from different species of *Solanum* and found to be present and highly conserved in both tuber bearing and non tuber bearing *Solanum* species (Wang et al. 2008). In natural diversity study of *Rps2* of *Arabidopsis thaliana* conferring resistance against *Pseudomonas syringae,* the site of pathogen recognition was the target of natural selection. The data of this study suggested that diversifying selection is required to maintain natural variation (Mauricio et al. 2003). Kover and Schaal (2002) also studied the genetic variation of different Arabidopsis accessions towards resistance against *Pseudomonas syringae*. The alleles of *Pto* gene in wild tomato showed excess polymorphism at non synonymous site, however the frequency of amino acid polymorphism, which negatively affect the function of *Pto* was much lower than silent polymorphism thus indicating the operation of both balancing and purifying selection in the evolution of this gene (Rose et al. 2007). Different alleles of *Sr* gene for resistance against stem rust in wheat was identified and the authors proposed that the allelic variation obtained for this gene can be exploited for marker assisted selection which will further provide durable resistance to stem rust etc (Tesfaye et al. 2013). Considering these findings, in a similar manner we also designed our study to identify new resistance alleles for BB disease in rice and analysis of the pattern of selection and evolution operating in cloned BB resistance *R* genes.

2.6 DEFENSE RESPONSIVE GENES FOR DISEASE RESISTANCE

Defense responses to protect plants from pathogens are activated in plants following pathogen recognition and signal transduction (Chen and Ronald, 2011). In general the defense responses are cell wall reinforcement, accumulation of antimicrobial secondary metabolites like phytoalexins, and expression of pathogenesis related (PR) proteins, defensins, peroxidases, proteinase inhibitors, free radical scavenging enzymes

etc., (Yang et al. 1997). After pathogen attack, a class of defense proteins known as hydroxyproline-rich glycoproteins (HRGPs) involved in cell wall biosynthesis accumulates at the site of infection making cell walls impermeable to pathogens. Extensin is an important example of HRGPs (Vidyasekharan 2004). Another class of enzymes playing important role in plants defense are β-glucosidase enzyme involved in cell wall metabolism, lignifications, synthesis of PR proteins and activation of defense compounds such as cyanide, isothiocyanates, or DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one) (Niemeyer 1988; Poulton 1990; Czjzek 2001; Barleban 2007; Morant et al. 2008). There are 17 groups of PR proteins (PR1-PR17). In rice, only a few pathogens induced PR genes, PR1, PR8 and PR10, have been reported (Park et al. 2004; Kim et al. 2008; Mitsuhara et al. 2008). Few phytohormones such as salicylic acid and jasmonic acid have also shown to play important roles in plant's defense response (Mei et al. 2006; De Vleesschauwer et al. 2010; Jiang et al. 2010). In plants, pathogen induced defense response have been intensively studied in *Arabidopsis* (Dong 2001; Glazebrook 2001), whereas in rice, the signal transduction pathways of this kind of defense response is still underway. Few reports are available in rice, claiming induction of defense response genes after infection with *Xoo* and *M. grisea* in IRBB13, Minghui 63 and Zhenshan 97. The response was both pathogen specific and pathogen non specific (Wen et al. 2003).

The whole genome sequence of both *Oryza sativa* and *Xanthomonas oryzae* are available which facilitate advance studies in functional genomics and proteomics. Advance techniques such as two-dimensional gel electrophoresis coupled with MS-MS/MALDI, microarray analysis, AFLP-cDNA analysis and RT-PCR will aid in identifying differentially expressed proteins and candidate genes after infection with *Xoo* in highly disease resistant wild rice species whose resistance factors are unknown.

2.7 TRANSCRIPTOME ANALYSIS OF PLANT PATHOGEN INTERACTION

During plant pathogen interaction, the expressions of different genes are regulated which mediates defense responses in the plants. Various techniques had been developed during the last decade in order to identify genes showing differential

expression at the transcript level. Some of the popular techniques used in transcriptome analysis are cDNA microarray (Schena et al. 1995), cDNA AFLP (Vuylsteke et al. 2007), suppression substractive hybridization (Diatchenko 1996), SAGE with the aid of NGS (Velculescu et al. 1995; Mardis 2008), DGE (Audic and Claverie, 1997), qPCR, RNA seq (Mortazavi et al. 2008) etc.

Transcriptome analysis helps in identifying the underlying signaling pathways and candidate disease resistance genes activated during plant pathogen interaction. Plant's defense in particular is carried out by *R* genes (through gene for gene interaction). Other defense genes are Pathogenesis Related (*PR*) genes, defensins, enzymes involved in phytoalexin biosynthesis, plant protectant enzymes and signal transduction components such as specific transcription factors. Induction of defense genes primarily occur at the transcription level and plant defense response is most importantly governed by sequential and spatial expression pattern of defense genes (Andersson et al. 2000).

Related to bacterial blight disease, a genome wide identification of defense response genes was performed in *xa13* gene mediated resistance plants where they identified 702 unique expressed sequences triggered by *xa13*. In addition, they could also identify genes putatively encoding NBS-LRR and LRR receptor kinase (Chu et al. 2004). Wen et al. (2003) also conducted an expression profile of 12 defense response genes where they observed constitutive expression of these genes, but significantly induced under the influence of *Xoo* and *Pyricularia grisea*. In an expression profiling study of rice genes conducted by Li et al. (2006) in rice-*Xoo* pathosystem (transgenic line TP309-Xa21 carrying *Xa21* and avirulent *Xoo* strain P6 and virulent strain K1), 498 and 454 differentially expressed genes were identified in compatible and incompatible interactions, respectively. Most of these genes were defense related and transcription factors. Transcriptional profiling study of *indica* rice cultivar Ajaya (IET8585), considered as a resistant variety to BB identified 1450 genes with increased expression after *Xoo* (Indian race Bxo1) infection (Kottapalli et al. 2007). Most of the genes were defense/stress related and many protein kinases and leurine rich repeats coding genes were up-regulated in IET8585 at all time points. Expression of pathogenesis related transcription factor, ERF (ethylene response element binding protein, EREBP) and

protein phosphatase 2C were induced in resistant IET8585. RNA sequencing technology was used to identify pathogen response genes in cotton (Xu et al. 2011), soybean (kim et al. 2011) and *Arabidopsis* (Zhu et al. 2012). DeepSAGE (Nielsen 2006), an improved version of SAGE was used to study the comparative transcriptome of compatible and incompatible interactions of potato infected with *Phytophthora infestans* (Gyetvai et al. 2012).

Microarray is an important tool to identify and quantify expression profile of host/pathogen genes during disease progression (Lodha and Basak, 2012). A comprehensive analysis of SAR genes in Arabidopsis was performed by Maleck et al. (2000) using microarray technology. A common promoter element called W box (WRKY factor binding site) which was co regulated with PR1 gene. Differential expressions of 705 mRNAs were observed in *Arabidopsis thaliana* after infection with *Alternaria brassicicola* or treatment with salicylic acid, jasmonic acid and ethylene (Schenk et al. 2003). The gene expression of biotrophic fungus *Blumeria graminis* was examined during barley infection where differential expression pattern of enzyme encoding genes for different primary metabolic pathways was observed (Both et al. 2005). Scheideler et al. (2002) using cDNA arrays identified new physiological processes during plant defense in *Arabidopsis* infected with *Pseudomonas syringae*. Major changes were observed in the enzymes involved in metabolic pathways. Two different groups Gjetting et al. (2004) and Catoni et al. (2009) also reported the genome wide expression profile of defense related genes and identification of candidate disease resistance genes after infection in barley and tomato plants, respectively. Till date, numerous reports are available on transcript analysis in the area of molecular plant microbe interaction though all the findings are not mentioned here. Similar results had been reported in almost all the findings with most of the defense related genes being induced. Using these informations, we present defense responsive transcript analysis of rice plants in response to *Xanthomonas oryzae* pv. *oryzae* infection and screening of resistance genes induced by the pathogen in wild relatives of rice, the wild relatives being better prospect for future BB research.

2.8 PROTEOMICS ANALYSIS OF DISEASE RESISTANCE GENES IN PLANTS

During the recent past, proteomics approach using 2DE coupled with MS/MS analysis for protein identification is fast becoming one of the popular methods to understand different stress response mechanisms and identify various proteins involved in the process. This method had been widely adopted to study the signaling cascade involved in plant pathogen interaction (Quirino et al. 2010). Proteomics has contributed to crop protection by helping in identification of fungal and bacterial effectors which trigger or induce immune response. It also helped in identification of host proteins and biomarkers of genes related to defense response (Yang et al. 2013). Comparative proteomics was used to monitor changes in protein expression pattern under biotic stress in rice (Xu et al. 2013). It helps to identify novel proteins; characterize particular and common changes in protein expression pattern post pathogen infection in plants. Comparative proteome profile was conducted between compatible and incompatible host pathogen interactions or between pathogen inoculated and mock treated in different plants-pathogen system. This approach has been used to detect differentially expressed defense response genes of rice leaf blades after *Xoo* inoculation in Java 14 (Mahmood et al. 2006). In this study, they could identify four different defense related proteins PBZ, PR5, SOD and peroxiredoxin. Geddes et al. (2008) identified 16 differentially expressed proteins related to defense response in Fusarium-barley interaction. Zhou et al. (2006) also conducted a similar kind of study for Fusarium-wheat interaction in fusarium head blight disease. In an experiment conducted to identify genes required for Cf-dependent hypersensitive cell death in tomato seedlings, it was found that 50% of the up-regulated proteins were defense related. The identified proteins were a set of pathogenesis related (PR) proteins, different class of glutathione S transferases, cysteine protease, phytopthora-inhibited protease 1 (Pip1), a trypsin-inhibitor family protein LeMir (Lycopersicon Esculentum Miraculin), chitinases and β-1, 3-glucanase. They identified that Pip1-like protein, a SIPK type MAP kinase, an ASN, and a LeMir-like protein are required for Cf-4/Avr4-dependent HCD (Xu et al. 2012). Generally in disease resistant plants, it was observed that most of the up-regulated proteins in comparison to susceptible plants were related to defense, energy, metabolism and protein synthesis. ROS-detoxification enzyme glutathione S-transferase 14 (GST14), APX7 and a PR

protein, Chia2a, was found to be differentially expressed between WT and spl5 mutant in rice (Chen et al. 2013). In a comparative 2DE and MS analysis between resistant and susceptible wheat plants against yellow rust disease by Maytalman et al. (2013), PR1, PR4, GST and Peroxiredoxin Q (Prx Q) were identified, which are well known defense related genes for resistance against yellow rust and other pathogen infection (Zeng et al. 2010; Coram et al. 2008). At the current scenario, proteome analysis on rice-*Xoo* interaction is still naïve. There are big scopes to explore in this area using wild relatives and landraces. Remarkable improvement in proteomics study coupled with availability of rice and *Xoo* whole genome sequences aided by bioinformatic tools will help in identifying novel elements for disease resistance and also enhance the understanding of molecular and cellular mechanisms involved in this popular host pathogen interaction. In our present study, we have exploited the above information and studied the influence of *Xoo* infection in protein profile of resistant wild relatives of rice. However, the full characterization of rice-*Xoo* proteome is quite a challenging task due to its complex experimental procedure, high cost and diverse protein dynamics.


Chapter Three
Materials & Methods

Kits, enzymes, chemicals and reagents:

All kits for PCR purification, gel extraction, PCR cloning, plasmid isolation, RNA isolation, cDNA synthesis and the enzymes required were procured from Simga-Aldrich (USA), MBI Fermentas (Germany), Genetix (India), Finnzymes, Thermo scientific (USA) and Invitrogen (USA). The required chemicals and reagents of analytical grade were obtained from Sigma-Aldrich (USA), GE health care (USA), Merck (Germany), Fermentas (Germany), Himedia and Qualigens fine chemicals (India).

3.1. IDENTIFICATION OF THE SUITABLE GENETIC RESOURCE FOR ALLELE MINING OF BACTERIAL BLIGHT RESISTANT GENES IN RICE

Before the beginning of natural sequence variation analysis, our target was to identify the genetic source which has high resistance potential against wide range of *Xanthomonas oryzae*. The selection of the individual rice accessions was based on the availability of germplasm with us as well as to represent *Oryza* diversity as much as possible. The materials and methods described in this section are according to Bimolata et al. (2013).

3.1.1 Plant Material and Bacterial Strains

In our study, we used 104 rice accessions (26 of *O. nivara*, 66 of *O. sativa*, two of *O. officinalis* and single accession of *O. alta*) to screen (Annexure 1) bacterial blight resistance. Nine near isogenic lines in the genetic background of the cultivar IR 24 possessing single BB *R* gene: IRBB1, IRBB3, IRBB4, IRBB5, IRBB7, IRBB10, IRBB11, IRBB13 and IRBB21 were included as control for the respective *R* genes. IR24, PB1 and TN1 were also used as bacterial blight susceptible checks. The rice genotypes were obtained from the Directorate of Rice Research, Hyderabad (Andhra Pradesh, India), Rajendra Agriculture University, PUSA (Bihar, India), CRRI, Cuttack (Odisha, India) and NBPGR (New Delhi, India). Rice seeds were germinated in dark at 28ºC in germination box layered with pre-soaked germination paper. After a week, the germinated seeds were shifted to 10 cm diameter pots containing soil. The seed beds

were uniformly watered and fertilized with a half-strength Hoagland nutrient solution. After 30 days, healthy seedlings were transplanted to 20 litres earthen pots. Five pots were kept for each genotype with three plants in each pot filled with a mixture of clay and peat (1:1, v/v).

3.1.2 Plant inoculation and disease Phenotyping

After 45 days of transplantation, each genotype was challenged with five Indian virulent isolates of *Xoo*. The isolates used for screening were DX011 (Pantnagar, Uttarakhand), DX127 (Cuttack, Odisha), DX020 (Hyderabad, Andhra Pradesh), DX015 (Aduthurai, Tamil Nadu) and DX133 (Raipur, Chhattishgarh). These strains were obtained from Department of Plant Pathology, DRR, Hyderabad and cultured on modified Wakimoto's medium or Hayward's medium (Annexure 2) at 28ºC for 72 hours. The bacteria were then scraped and suspended in sterile distilled water and the concentration was adjusted to 0.1-0.2 OD ($1x10^8$- $1x10^9$ CFU/ml). Using this bacterial suspension, five to six uppermost leaves of plants at the booting stage were inoculated following leaf clipping method (Kauffman et al. 1973). Three plants were used for each isolate. The control plants were clipped with scissors dipped in sterile water. After 15 days of inoculation, lesion lengths were measured and the ratios of lesion length to leaf length (RLL) were calculated. Scoring of disease was done following Standard Evaluation System for Rice (SES scale; IRRI, 2002). RLL less than 25% were scored as resistant or tolerant genotypes and above 25% were scored as susceptible (Fig. 3.1). During the process of screening for bacterial blight resistance, seedlings were uniformly irrigated and fertilized using Hoagland nutrient solution.

3.1.3 Genomic DNA isolation

Genomic DNA was isolated from fresh young leaves of 70 rice plants following CTAB method (Doyle and Doyle, 1990) with minor modifications (Rajendrakumar et al. 2007). 150 mg of rice leaves was crushed into fine powder using liquid nitrogen. Precautions were taken not to thaw the sample. In 2 ml microfuge tube, 1 ml CTAB buffer was added and finely crushed leaf sample powder was added to it. The sample

was mixed thoroughly with the buffer and incubated at 65ºC for 45 min. The solution was mixed in between to maintain uniformity. Into this, an equal volume of chloroform was added, mixed very slowly for 5 min and spinned down at 12,000 RPM for 10 min at 4ºC. The upper aqueous layer was then transferred into fresh 1.5 ml tube. Precaution was taken not to disturb the middle layer. The sample was then treated with 10 mg/ml RNase enzyme for 2 h at 37ºC. After RNase treatment, equal volume of absolute propane-2-ol was added and left at room temperature for 30 min to precipitate the genomic DNA. The precipitated DNA was pelleted down at 12000 RPM for 15 min and then washed with 70% ethanol for 5 min at 5000 RPM. A white pellet DNA was obtained which was air dried till the ethanol evaporates or till the DNA becomes transparent. Finally, DNA was dissolved in 50 µL of 1X Tris EDTA (10 mM Tris and 1 mM EDTA, pH 8.00) buffer and stored at -20ºC for PCR analysis.

The quality and quantity of DNA was determined by NanoDropND1000, Thermo scientific (USA) and electrophoretically on agarose gel. Standard known concentration of commercial vector DNA was loaded to check the concentration of DNA. The concentration of DNA was determined by measuring the absorbance at 260 nm. A value of OD 260 =1 corresponds to 50 ng/µL for DNA. A value of OD 260/280 = 1.8 was considered pure for DNA.

3.1.4 Screening of rice genotypes with functional marker

The genotypic screening with functional or gene linked marker was performed in order to identify individuals with identical alleles of the resistant form. PCR reactions were carried out for 10 µL reaction volume containing 50 ng of genomic DNA, 3 pmol of each primers, 2 mM of each dNTPs, 10 x PCR buffer (10 mM Tris-HCl, pH 8.0, 2.5 mM $MgCl_2$), and 0.5 unit (5U/ µL of *Taq* DNA polymerase (Sigma Aldrich). The PCR condition was set as 5 min initial denaturation at 95ºC, 32 cycles of amplification with 30 sec, DNA denaturation at 94ºC, 45 sec annealing at various temperatures for different genes. Elongation was kept at 72ºC with different time periods, depending on the product size. PCR analysis was repeated at least for 3 times. The amplified product was then

separated and resolved on 3% agarose gel. The primers reported by Iyer et al. (2006) and Hui et al. (2010), were used to screen the presence of resistant alleles.

3.2. ALLELE MINING OF *Xa21, Xa26, Xa27* AND *xa5* AND A COMPARATIVE NUCLEOTIDE DIVERSITY ANALYSIS OF ALLELES

The aim of this work was to evaluate the natural sequence variation and sequence evolution at the locus of four major BB resistant genes *Xa21, Xa26, Xa27* and *xa5* mainly in wild and cultivated species of *Oryza*. The materials and methods described in this section are according to Bimolata et al. (2013).

3.2.1 Primer designing and synthesis

Gene specific primers were designed for *Xa21, Xa26, Xa27* and *xa5* from GenBank accessions DQ426646.1 (cultivar Zhachanglong), U37133.1, AY986492 and Os05g0107700, respectively. In order to cover the complete gene with 5'UTR and 3'UTR, overlapping primers were designed 1 kb upstream of the start codon and 500bp downstream (Fig. 3.2). The number of primers depended on the size of the gene. Primer 3 online software (biotools.umassmed.edu/bioapps/primer3_www.cgi) (Rozen and Skaletsky, 2000) and Fast PCR was used to design the primers using optimal length 19 to 23 mers, GC content 40-55%, 2-3 GC clamps and Tm 54ºC to 65ºC. Some of the primers were designed manually. Self complimentary and potential hair pin loop formation was checked on Oligocalc (http://www.basic.northwestern.edu). The name of primers used in the whole work is listed in Table 3.1. The primers were synthesized from IDT and Sigma (Germany). The primer tubes were spinned down before opening and dissolved in 1X TE in the concentration of 100 µM as a stock. Working concentration of 10 µM was diluted from the stock with sterile MQ water.

3.2.2 Plant material and genomic DNA isolation for allele mining

The ORF including 5' and 3' UTR for the four genes were amplified from different wild species and *O. sativa.* Genomic DNA was isolated following Doyle and

Doyle (1990) same as described in section 3.1.3. For *Xa27* gene, 27 alleles were amplified from *O. nivara* (OS 28, ON 38-1, M 209 A, OS 23-2, 21028, 81832, 100963, 102166, 106133, 210223, ON132 ,ON15, ON24, ON 34 and ON44), *O. sativa* (IR20, TN1, PAU933, R2, R9, R23, MohemPhou, MustiLaiphou, PAU201, R4), *O. alta* (IC384116), *O. officinalis* (IC203740); for *Xa26*, 14 alleles were amplified from *O. nivara* (ON 38-1, M 209 A, 81832, 106133, 210223, ON132, *O. sativa* (TN1, PB1, MohemPhou, MustiLaiphou, R47), *O. alta* (IC384116), *O. officinalis* (IC203740); 19 alleles of *Xa21* was isolated from *O. nivara* (M209A, 81832, 81852, 100963, 102166, 106133, 210223, ON132, ON15), *O. sativa* (IR20, TN1, PB1, PAU933, R47, MohemPhou, MustiLaiphou, R4), *O. longistaminata* (OL1 and OLD) and 15 alleles of *xa5* was amplified from (M209A, 81832, 81852, 100963, 106133, ON132 ,ON15), *O. sativa* (TN1, PB1, MohemPhou, MustiLaiphou, Kalamekri, AC32753), *O. longistaminata* (OL1 and OLD).

3.2.3 PCR isolation of alleles

The alleles of *Xa21, Xa26, Xa27* and *xa5* were isolated from selected accessions using gene specific primers. PCR reactions were carried out for 30 µl reaction system containing 50 ng of genomic DNA, 9 pmole of each primer, 2 mM of each dNTPs, 5 x PCR buffer (10 mM Tris-HCl, pH 8.0, 1.5 mM $MgCl_2$), and 0.6 unit (2U/µL) of High Fidelity phusion polymerase (Finnzymes). The PCR condition was set as 3 min initial denaturation at 98ºC, 35 cycles of amplification with 30 sec, DNA denaturation at 98ºC, 30 sec annealing at different temperatures for different primers. Elongation was kept at 72ºC with different time periods, depending on the product size. There was slight variation in the PCR amplification pattern from one primer to another. The amplified product was then separated and resolved on 0.8% agarose gel and purified using gel extraction kit from Sigma Aldrich (USA).

3.2.4 Purification of PCR product from agarose gel

The amplified band was excised and weighed. To this three volumes of gel solublization buffer were added and kept at dry bath or incubator at 50-60ºC for 5-10

min with regular shaking. After solublization of the gel it was transferred into mini spin column placed in collection tube provided in the kit. The samples were then spinned at 14,000-16,000 RPM for 30 sec and discarded the flow through. Into this 700 µL of wash buffer was added in mini spin column and centrifuge at 14,000-16,000 RPM for 30 sec. The above step was repeated without adding wash buffer to remove traces of wash buffer. Finally, 50 µL of elution buffer were added to elute the PCR product. The concentration of the purified DNA was checked in NanoDropND1000 spectrophotometer.

3.2.5 Polyadenylation of blunt end PCR product

Hi–Fi phusion polymerase produce blunt end PCR product. Before proceeding for ligation into T- vector, the amplicons were polyadenylated using *Taq* polymerase. For 20 µL PCR product, 25 µL polyadenylation reactions were kept. The reaction mixture consists of 2.5 µL 10X PCR buffer, 2 µL dATP (10 mM) and 0.5 µL (5U/µL) Taq polymerase. The reaction was carried out at 72ºC for 30 min.

3.2.6 Cloning and sequencing of alleles

3.2.6.1 Ligation

DNA ligation involves creating a phosphodiester bond between the 3' hydroxyl of one nucleotide and the 5' phosphate of another. The polyadenylated product was ligated into cloning vector pTZ57R/T (2886 bp in length). The ligation reaction was carried out in a total reaction volume of 15 µL, containing 50 ng of pTZ57R/T vector DNA, appropriate amount of PCR product in the ratio of 3:1 (PCR product: vector), 1.5 µL of 10X T_4 DNA ligase buffer and 1 µL of T_4 DNA ligase (5 U/µL). The reaction was carried out at 22°C for 2 hrs.

3.2.6.2 Preparation of competent cells by $CaCl_2$ method

Competent cells for transformation were prepared following Sambrook et al. (1989). Primary culture of *E coli* DH5α cells was inoculated from freshly prepared mother plate. This culture was incubated overnight/12 hr at 37ºC. This primary culture was transferred into 50 ml of LB broth and incubated at 37ºC until the OD_{600} reaches 0.4. The cells were then stored at 4ºC for 10 min and the cells were pelleted down at 5000 RPM for 4 min at 4ºC. The pellet was resuspended in 10 ml of ice cold 0.1 M $CaCl_2$ and stored at ice for 30 min. Cells were recovered by centrifugation at 4ºC for 4000 RPM for 5 min. The pellets were resuspended slowly in 3 ml of 0.1 M of $CaCl_2$ and 3 ml of 50% glycerol (final concentration of glycerol is 25%). The competent cells were then aliquoted 150 µL each into 1.5 ml of centrifuge tube and frozen in liquid N_2. These cells were then stored at -70ºC for long term use.

3.2.6.3 Blue white screening

The ligated product was added to the competent cells DH5α maintaining sterile condition. Cells were incubated on ice for 30 minutes and subjected to heat shock at 42°C for 90 sec and cold shock for 5 min. To the transformed cells 1 ml of LB broth was added and then incubated at 37°C (shaker) for 1 hr. The LB agar plates were then prepared with ampicillin (100 µg/ml), 40 µL of X-gal (25 mg/ml) and 5 µL of IPTG (25 mg/ml). The cell pellets were then resuspended and spreaded onto the agar plates. LB plates were kept for incubation at 37°C overnight. After successful transformation we get isolated blue and white color colonies on the plate. Transformed colonies with desired insert appear white in color while colonies without insert are blue in color. White colonies were then selected for further confirmation.

3.2.6.4 Colony PCR and sequencing

The presence of positive clones were detected by colony PCR followed by plasmid PCR and restriction digestion. In a normal PCR reaction, colonies were used as template and the amplified colonies were inoculated in LB broth overnight at 37ºC. Plasmid DNA was isolated by conventional alkaline lysis method or using Plasmid Miniprep kit from Fermentas (USA). Further confirmation of cloning was done by

plasmid PCR in which 0.3-0.5 µL of plasmid was used as template with *Taq* 10X buffer system. PCR was performed in a total volume of 10 µL containing 3 pmole of each forward and reverse primer, 0.2 mM of each of the dNTPs, 0.75 units of *Taq* DNA polymerase, 10X *Taq* PCR buffer and 10-20 ng/µL concentration of plasmid DNA. The PCR condition was set as 5 min initial denaturation at 95ºC, 32 cycles of amplification with 30 sec DNA denaturation at 94ºC, 30 sec annealing at different annealing temperatures for different primers. Elongation was kept at 72ºC with different times depending on the product size. The amplified product were separated and resolved on 1% agarose gel. Restriction digestion of the plasmid to confirm the presence of insert was performed with specific two restriction enzymes at two sites outside the insert. Double restriction digestion was performed by addition of recommended buffer compatible for the both enzymes, plasmid DNA up to 1 µg, 1µL each of the enzymes and incubating at appropriate temperature (usually 37ºC) and incubation period following manufacturer's instruction. The release of insert was checked on 1% agarose gel. The cloned plasmids with the insert DNA of interest were sequenced by SciGenom. The quality of the sequences was analyzed on the chromatogram visually.

3.2.7 Analysis of allelic diversity and sequence data

The analysis of nucleotide sequence variation helps in identifying the selection pressures acting upon different genes. The raw sequence data obtained after sequencing were edited, aligned and the overlapping nucleotide sequences were removed. Then the sequenced products of different primer combinations were arranged into complete sequence. Overlapping sequences between the products of two adjacent primers was identified using align two sequences using online tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi?) and deleted from one of the sequenced product. The coding regions and the UTR regions were determined from the reference sequences. The reference sequences used for different genes were U37133.1, DQ426646.1 (cultivar Zhachanglong), AY986492 and Os05g0107700 for *Xa21, Xa26, Xa27* and *xa5*, respectively. The identified alleles were then compared to their respective susceptible and resistant alleles. The alleles obtained, respectively, for *Xa21, Xa26, Xa27* and *xa5* were compared with that of IRBB21, IRBB3/Minghui63, IRBB27, IRBB5 (all resistant

allele) and IR24 (susceptible allele). Different software utilities available online and offline were used to study the allelic and nucleotide diversity. The percentage sequence identity was determined by pair wise alignment using NCBI BLAST (http://blast.ncbi.nlm.nih.gov/; Altschul et al. 1990). The sequence variation among alleles was compared using standard summary statistics π (average pairwise difference or average no. of nucleotide diversity per site, Nei and Li, 1979) and θ (population mutation parameter or number of segregating sites, Watterson 1975). Intra and interspecific polymorphisms of the alleles were analyzed using an offline software DnaSP program version 5.1 (Librado and Rozas, 2009). Sequences were first aligned using ClustalX (Thompson et al. 1997). The output aligned file was saved as Fasta format and used as an input file for analysis in DnaSP (http://www.ub.edu/dnasp). The numbers of polymorphic sites, including SNPs, InDel polymorphism, synonymous and non - synonymous substitutions in coding and non-coding region (Nei and Gojobori, 1986) were estimated. All the analysis was done in sliding window for 50 bp per window. Alignment gaps were excluded from comparative analysis. Each allele was also compared to their respective resistant and susceptible allele.

Different test selection analysis, such as Tajima's *D* test (Tajima 1989) and Fu and Li's *D* test (Fu and Li, 1993) to determine departures from neutrality were conducted using DnaSp program. The calculation of Tajima's D test is based on the value of Θ_w and π, where under standard model that value of θ_W and π are nearly equal. Presence of excess of rare polymorphisms is indicated by negative D value while, a positive Tajima's D signifies an excess of intermediate frequency polymorphisms. The 'θ' value for Fu and Li's D test, is determined by the number of singleton mutations 'ηe' and the total number of mutations 'η'. Linkage disequilibrium (LD) was estimated based on the parameter of R^2 between all SNPs using DnaSP v5.10. Significance of LD was measured statistically using Fisher's exact test and Chi-square test. Decay in LD was determined by plotting graphs of R^2 versus pairwise distance (bp).

3.2.8 Constructing a phylogenetic tree

The phylogenetic trees of different alleles were constructed using MEGA 4 (Tamura et al. 2007). The alleles were aligned in ClustalX. After alignment, an unrooted linear Neighbour joining (NJ) tree was constructed with an option of 10,000 bootstrap values. In order to determine the authenticity of the NJ plot, phylogenetic distances were calculated. An example of calculating the phylogenetic distance: For two alleles the sequence similarity was determined in EZTaxon server 2.1, and then the distances of the alleles were calculated from their respective nodes. Both the distances from their nodes were added and multiplied by the scale given in the tree. The obtained value was then divided by the distance of the allele from its node (we should choose the shorter one). This final value was subtracted from 100; and if the obtained value is equivalent to the percentage sequence similarity between the two alleles, the phylogenetic tree could be defined as authentic.

3.2.9 *In silico* structural analysis of deduced proteins of the alleles/ functional prediction of polymorphic alleles

Amino acid sequences were deduced for the identified alleles using ORF finder or by translation of the CDS region. The sequences were then aligned in ClustalX and amino acid changes or non synonymous mutations were identified. Change in the protein function due to amino acid substitution was predicted through SIFT (Kumar et al. 2009) and PROVEAN (Choi et al. 2012) software tools. SignalP (Petersen et al. 2011) and TOPCONS (Bernsel et al. 2009) software were used to predict the presence of signal peptide and trans membrane region.

3.2.10 Expression analysis of *Xa21* and *xa5* allele

3.2.10.1 Plant material and infection

Two genes (*Xa21* and *xa5*) and four alleles for each gene were selected for expression analysis. The selected alleles were IRBB21 (*Xa21*) and IRBB5 (*xa5*) as

resistant control and IR24 as a susceptible control for both the gene. The *Xa21* test alleles were from two accessions of *Oryza longistaminata* and the *xa5* test alleles were from Kalamekri and AC32753 (*O. sativa*). The rice plants were grown in glass house condition as described in section 3.1.1. The plants were then inoculated with *Xoo* strain DX011 in triplicates by syringe infiltration method. The control plants were treated in a similar manner with sterile water instead of *Xoo* inoculums. The day and night temperature of the glass house was maintained at 28ºC. Leaf samples for RNA isolation were collected at different time intervals such as 0 h and 48 h.

3.2.10.2 Isolation of total RNA and cDNA synthesis

Approximately 100 mg of frozen leaf material was ground with mortar and pestle to a fine power with liquid nitrogen. Total RNA from the fine powder was extracted using an RNA isolation kit (Genetix) according to the manufacturers' instructions. The total RNA was treated with DNase to remove genomic DNA impurities. Total RNA was isolated from triplicate leaf samples. Care was taken to avoid degradation of RNA by RNase. After determining the quality and quantity of total RNA, aliquots were made and stored at - 80ºC for further use. cDNA was synthesized on the same day of RNA isolation to avoid degradation of RNA due to freeze thaw process. Equal amount of RNA (1-2 µg) was taken and cDNA was synthesized using Superscript III cDNA synthesis kit (Invitrogen, USA) following manufacturer's instruction.

3.2.10.3 RT-PCR and amplification

Reverse Transcription PCR was performed with Superscript cDNA synthesis kit to generate first strand cDNA. For a 20 µL reaction, 5X VILO Reaction Mix, 4 µL 10X SuperScript Enzyme Mix, 2 µL RNA (up to 2.5 µg) and DEPC water were added to make up the final volume to 20 µL in sterile RNase free tubes. The contents were mixed gently and incubated at 25°C for 10 min followed by incubation at 42°C for 60 min. The reaction was terminated by increasing the temperature to 85°C for 5 min. The quality and quantity of the cDNA was checked on agarose gel and spectrophotometrically. Semi-quantitative PCR for each gene was done as a preliminary expression profile study.

3.2.10.4 Real time PCR

Primers for Q-RT PCR were designed in such a way that the 3' end contains dA and the last five nucleotides contain no more than two 'G' or 'C' to avoid primer dimer formation. Primers for *xa5* and *Xa21* were designed from the exonic junction (Table 3.2). A master mix containing 2X Sybr select master mix (Invitrogen, USA), 150- 400 nM forward and reverse primers, 100 ng of cDNA was used as a template. No template control was used to identify PCR contamination. In a 96 well optical plate, 20 µL each of reaction mixture was loaded and sealed tightly with optical adhesive cover (Applied Biosystem). The PCR program was set as UDG activation at 50ºC for 2 min, Ampli*Taq* DNA Polymerase, UP activation at 95ºC for 2 min and denaturation at 95ºC for 15 sec. Annealing of primer was kept at melting temperature of the primer for 15 sec and extension at 72ºC for 1 min. The cycle was repeated for 40 times from step 3.

3.2.10.5 Data analysis

The real time data were analyzed by relative quantification method. The average Ct value of two nearest readings for each sample of different time points and different primers were calculated. Each reaction was kept in triplicate and a standard error was calculated. The expression profile of different genes at different time point was normalized by zero hour samples and three internal control genes. The fold increase in expression was calculated using the following formula:

$\Delta\Delta C_T = (\Delta C_{T,\ Target} - \Delta C_{T,\ Reference})_{\ Time\ x} - (\Delta C_{T,\ Target} - \Delta C_{T,\ Reference})_{\ Time\ 0}$

$\Delta C_{T,\ Target} = C_{T,\ Control} - C_{T,\ Treatment,}\ \Delta CT_{\ Reference} = C_{T,\ Control} - C_{T,\ Treatment}$

Fold change in expression = $2^{-\Delta\Delta C_T}$

3.3. STUDY OF DIFFERENTIALLY UP-REGULATED GENES AT TRANSCRIPT AND TRANSLATIONAL LEVEL DURING *XOO* INFECTION IN RESISTANT WILD RICE

3.3.1 Plant material and infection for transcript and translational analysis

Oryza nivara Acc. No. 81832, *Oryza longistaminata* (OL1) and *Oryza sativa* IR24 (Fig. 3.3) were grown and infected with *Xoo* strain DX011 as described in section 3.2.10.1. The same procedure was followed for infection and collecting the leaf samples too.

3.3.2. Isolation of total RNA and cDNA synthesis

Approximately 100 mg of frozen leaf material was ground with mortar and pestle to a fine power with liquid nitrogen. Total RNA from the fine powder was extracted using RNA isolation kit according to the manufacturers' instructions, Genetix (India). The total RNA was treated with DNase to remove genomic DNA impurities. Total RNA was isolated from triplicate leaf samples. Care was taken to avoid degradation of RNA by RNase. After determining the quality and quantity of total RNA, aliquots were made and stored at - 80ºC for further use. The concentration of RNA was determined by measuring the absorbance at 280 nm. A value of OD 260 = 1 corresponds to 40 ng/uL for RNA. A value of OD 260/280 = 1.9 to 2.0 was considered pure for RNA. cDNA was synthesized on the same day of RNA isolation to avoid degradation of RNA due to freeze thaw process. Equal amount of RNA (1-2 µg) was taken and cDNA was synthesized using Superscript III cDNA synthesis kit of Invitrogen (USA) following manufacturer's instruction.

3.3.3. Expression analysis for defense response genes

Reverse Transcription PCR was performed with Superscript cDNA synthesis kit to generate first strand cDNA. For a 20 µL reaction, 5X VILO Reaction Mix, 4 µL 10X SuperScript Enzyme Mix, 2 µL RNA (up to 2.5 µg) and DEPC water were added to make up the final volume to 20 µL in sterile RNase free tubes. The contents were mixed gently and incubated at 25°C for 10 min followed by incubation at 42°C for 60 min. The reaction was terminated by increasing the temperature to 85°C for 5 min. Semi-quantitative PCR for each gene was done as a preliminary expression profile study. The list of genes and their primers used in this study is listed in Table 3.3. The genes were

selected based on the previous reports stating their role and up regulation during infection in rice. The samples used were uninfected or zero time point, mock inoculated and bacterial infected at 24 h and 48 h time points. Primers designing and real time PCR was performed in similar manner as described above in section 3.2.10.4.

3.3.4. Data analysis

The real time data were analyzed by relative quantification method. The average C_T value of two nearest readings for each samples of different time points and different primers were calculated. Each reaction was kept in triplicate and a standard error was calculated. The expression profile of different genes at different time point was normalized by zero hour samples and three internal control genes. The fold increase in expression was calculated using the following formula:

$$\Delta\Delta C_T = (\Delta C_{T,\ Target} - \Delta C_{T,\ Reference})_{Time\ x} - (\Delta C_{T,\ Target} - \Delta C_{T,\ Reference})_{Time\ 0}$$

$$\Delta C_{T,\ Target} = C_{T,\ Control} - C_{T,\ Treatment,}\ \Delta CT_{\ Reference} = C_{T,\ Control} - C_{T,\ Treatment}$$

$$\text{Fold change in expression} = 2^{-\Delta\Delta C_T}$$

3.3.5. Total protein isolation from leaves of rice

Total plant protein was isolated from control and infected plants of *O. nivara* acc. No. 81832 at 3DPI. One gram leaf samples were collected from infected and control plants and stored at -80ºC for further use. Total leaf proteins were isolated following Sengupta et al. (2011) method. One gram leaf tissue was ground to fine powder in prechilled mortar and pestle using liquid nitrogen and suspended in 4 ml pre-cooled extraction buffer [0.5 M Tris–HCl (pH 7.5), 0.7 M sucrose, 0.1 M KCl, 50 mM EDTA, 2% 2-mercaptoethanol and 1 mM PMSF]. To this equal volume of Tris–HCl (pH 7.5) saturated phenol was added, mixed thoroughly for 30-45 min at 4ºC or keeping on ice. The homogenates were centrifuged at 6,000 g for 20 min at 4ºC. The upper phenolic phase was collected slowly and transferred into fresh 2 ml tubes. An equal volume of extraction buffer was added to it. The above step was repeated and the upper phenolic phase was re-extracted. Four volumes of 0.1 M ammonium acetate in methanol was

added to the collected phenolic phase and kept 1-3 hours or overnight at -20ºC to precipitate the protein. The samples were then recentrifuged at 10,000 g at 4ºC for 30 min and the protein precipitate was washed thrice in ice cold methanol and in ice cold acetone. The pellet was air dried for 2-3 mins or until the acetone evaporated completely and finally resolved in 200 µL of rehydration buffer [8 M (w/v) urea, 2 M (w/v) thiourea, 4% (w/v) CHAPS, 30 mM DTT, 0.8% (v/v) IPG buffer, pH range 4–7 (GE Healthcare, Uppsala, Sweden)].

3.3.6. Rehydration and isoelectric focusing

The concentration of the isolated total leaf protein was quantified by Bradford's method. BSA protein was used as a standard to determine the concentration of protein sample. A total of 1 mg protein made up to a final volume of 320-340 µL with rehydration buffer. Into this 3 µL of 1% bromophenol blue tracking dye was added. The protein samples were mixed uniformly with a pipette. Then, the protein samples were loaded on immobilized linear gradient gel strips (pH 4-7, 18 cm; GE Healthcare) in a loading tray avoiding air bubbles. The strips were then over layered with mineral oil and placed on to Ettan IPGphor II (GE Healthcare) isoelectric focusing unit. Active rehydration and focusing of protein in first dimension was carried out at 20ºC in three steps: 30 min at 500 V, 3 h to increase from 500 to 10,000 V and 6 h at 10,000 V (a total of 60,000 Vh). This program was set in the system connected to the isoelectric focusing unit.

3.3.7. Two dimensional gel electrophoresis (2DE)

After Isoelectric focusing, strips were stored at -80ºC until the 2DE set up was ready. The strips were equilibrated with equilibration buffer I and II for 20 min each with gentle rocking at room temperature (25 ± 2ºC). The first equilibration buffer contains 6 M urea, 50 mM Tris–HCl buffer (pH 8.8), 30% (w/v) glycerol, 2% (w/v) SDS, 2% DTT and 0.002% (w/v) bromophenol blue. The second equilibration buffer contains 2.5% (w/v) iodoacetamide instead of DTT. The proteins were separated on 12% second dimension SDS–PAGE containing 12% (v/v) acrylamide, (30% (w/v) separation

acrylamide, and 0.135% (w/v) bisacrylamide), 1×(v/v) separation buffer (1 M Tris, pH 8.8 containing 0.27% (w/v) SDS), 0.07% APS, and 0.1% TEMED] and MQ water. The gels were then stained with coomassie brilliant blue G250. Destaining of the gels was done with destaining solution containing 45% methanol and 10% acetic acid. Destaining solution was changed after every 15 mins for first one hour and for every 45 mins till complete destaining of the background. The gels were then stored at 10% acetic acid at 4ºC. Protein patterns in the gels were recorded as digitized images using a calibrated densitometric scanner (GE Healthcare).

3.3.8. Image and data analysis

Analysis of the scanned image of the gels (100 dpi, TIFF file) was performed using ImageMaster 2D Platinum imaging software ver.6.0 (ImageMaster software; GE Healthcare). Spot detection and normalization was done using the software. Protein abundance was expressed as relative volume to compensate the differences in sample loading, gel staining, and destaining from gel to gel. All the gels were run together at a time to minimize technical error. The spots volume between control and treated was compared. The mean volume of treated was divided by control and the value > 1 was considered up-regulated proteins and < 1 as down-regulated proteins. Protein spots up-regulated more than two folds were selected for MALDI MS/MS TOF analysis and protein identification.

3.3.9. Enzymatic trypsin digestion of protein and protein identification

The Protein spots which showed more than 2 fold up-regulations were excised from the CBB stained gels. Protein spots of low intensity were pulled from multiple gels and transferred to sterile dust free 1.5 ml microfuge tubes. The protein spots were incubated in 200 µL of 50% acetonitrile (ACN) and 50 mM ammonium bicarbonate (NH_4HCO_3) for 20 min or until the gel was destained. Thereafter, the excised gel pieces were treated with 10 mM DTT in 50 mM NH_4HCO_3 and incubated at 56ºC for 1 h. This was followed by treatment with 55 mM iodoacetamide in 50 mM NH_4HCO_3 for 45 min in dark at room temperature (25 ± 2ºC). The gel pieces were then washed with 25 mM

NH_4HCO_3 and ACN, dehydrated in speed vac at ambient temperature. Gel pieces were rehydrated and swollen in digestion buffer containing 15 µL of 25 mM NH_4HCO_3 solution containing 25 ng μL^{-1} trypsin at 4ºC for 10 min and then digested at 37ºC for overnight (Promega, USA). After incubation, a short spin was given and the supernatant was collected in a fresh microfuge tube. The left gel pieces were further sonicated for 10 min followed by frequent vortexing for 5 min in 10 µL of 0.1% trifluoroacetic acid (TFA) and 100% ACN (1:1), to extract the remaining peptides. This extraction step was repeated twice to improve the extraction yield. The supernatants were pooled together, dried in a speed vac and the obtained peptides were resolved in 5 µL of 1:1 100% ACN and 0.1% TFA. The peptide suspension was mixed in the ratio of 1:1 with a-cyano-4-hydroxycinnamic acid (CHCA) matrix in 50% ACN and 1% TFA (1:1). The sample were spotted onto a MALDI plate and dried at room temperature for mass spectrophotometry.

MS/MS experiment for peptide identification was done by Matrix-assisted laser desorption/ionization time of flight mass spectrometric (MALDI-TOF MS) analysis using MALDI-TOF/TOF mass spectrometer (Bruker Autoflex III smartbeam, Bruker Daltonics, Bremen, Germany) according to the protocol of Shevchenko et al. (1996) with minor modifications. Mass data acquisitions were piloted by FlexControl 3.0 (Build 100) software using batched-processing and automatic switching between MS and MS/MS modes. All MS survey scans were acquired over the mass range of 800-3500 m/z in the reflectron positive-ion mode and accumulated from an average of 2500 laser shots with acceleration of 19 kV. Peptide precursor ions corresponding to contaminants including keratins and trypsins autolytic products were excluded in a mass tolerance of ±0.5 Da. The filtered precursor ions with a user defined threshold were selected for the MS/MS scan. Fragmentation of precursor ions was performed using MS-MS 1KV positive mode.

The MALDI-TOF/TOF data were loaded into the MASCOT program (http://www.matrixscience.com) employing Biotools software (Bruker Daltonics) and protein identification was performed against NCBInr and Swiss Prot database using combined MS (peptide mass fingerprint approach) with MS/MS. All lift and TOF files were combined for the MS/MS analysis. The taxonomic category was set to *Oryza sativa*

L. The other search parameters were: monoisotopic peptide mass (MH^+); 0-2 missed cleavage per peptide; enzyme, trypsin; precursor-ion mass tolerance in the range of 50-800 ppm; MS/MS fragment-ion mass tolerance, 1-2 Da; variable modifications, carbamidomethylation (C) for cysteine and oxidation for methionine (M) were allowed. Contaminating ions like trypsin and keratins were excluded from the peak lists before database searching. Top hit for each protein with highest score greater than the significant value ($P < 0.05$) and with a minimum of two peptides matched were considered to be the identified protein. If a protein spot matched multiple protein under different accession numbers, the candidate protein with the maximum Mascot score with highest peptides matched were selected. The nearest experimental MW (molecular weight) and PI values to the theoretical values if having the same Mascot score were given equal weightage in spot selection. The identified proteins were named according to the corresponding annotations in NCBI. Proteins which do not have functional annotation in the database, homologous proteins were taken and search against NCBI database (BLASTP: http://blast.ncbi.nlm.nih.gov/) to annotate the function of such protein.


Chapter Four
Results

4.1. ASSESMENT OF DISEASE PHENOTYPE AFTER *XOO* INFECTION

4.1.1 Wild species were highly resistant to tested Indian *Xoo* isolates

The search for new source of BB resistance in wild rice is an underexplored area to fight the disease in a country like India with diverse population of *Xoo* pathogens. We screened 104 rice accessions for their resistance against five different *Xoo* isolates DX011, DX127, DX020, DX015, and DX133 (Fig.4.1.1; Annexure I). Out of 104 accessions screened, a total of 23 accessions (14 accessions of *O. nivara*, one each of *O. alta* and *O. officinalis*, two introgression lines - PAU201 and R2; two landraces – Musti Laiphou (ML) and Mohem Phou (MP); and three isolines - IRBB13, IRBB21 and IR59) showed resistance against all the five Indian *Xoo* isolates while remaining accessions showed susceptibility to one or more of the isolate (Fig.4.1.2). Some of the wild species were highly resistant to the disease showing the presence of strong *R* gene(s), which can be transferred to elite cultivars through breeding strategy. The resistance score of the tested lines showed that only IRBB21 (*Xa21*) and IRBB13 (*xa13*) could resist against all the five isolates, while IRBB5 (*xa5*) possessed moderate resistance to DX011, DX127 and DX015 but susceptible to DX020 and DX133 isolate. Remaining lines containing single *R* genes IRBB1 (*Xa1*), IRBB3 (*Xa3/Xa26*), IRBB4 (*Xa4*), IRBB7 (*Xa7*), IRBB10 (*Xa10*) and IRBB11 (*Xa11*) were found to be defeated and susceptible (Fig.4.1.3). Most of the introgression lines and cultivars were found susceptible while majority of the wild species conferred high resistance against these five Indian *Xoo* isolates. Among the screened accessions, 52% of the wild species were found to be highly resistant and only three accessions, *O. nivara* (42446 and 21990) and *O. alta* (IC384076) were completely susceptible to all the isolates. Remaining 10 accessions were resistant to one or two *Xoo* isolates (Fig.4.1.4). Among the cultivated rice and landraces, only ML and MP were resistant. Those cultivars which were previously reported to be resistant to BB were found susceptible and could resist moderately to one or two isolates (Fig.4.1.5). Out of the 42 introgression lines of *O. nivara* x Swarna, only four lines exhibited moderate resistance to the disease. Our result also indicated that introgression lines containing three pyramided *Xa* genes *xa5*, *xa13* and *Xa21* (IR59) was highly resistant to the isolates.

4.1.2 Genotypic screening

Genotypic screening using gene linked markers were done for three genes *xa5, Xa21* and *xa13*. For other genes, appropriate markers were not available which distinguished clearly between susceptible and resistance allele. All the accessions which showed resistance to all five *Xoo* isolates as well as one or two isolates were screened using linked markers. Susceptibility check lines such as TN1, PB1, IR24 etc., and IRBB lines possessing specific genes were also included as control. For *xa5,* all the screened accessions except two accessions of *O. sativa* showed the presence of dominant form of the gene. Among the genotypes screened, recessive *xa5* were present only in two *O. sativa* accessions AC32753 and Kalamekri identified using CAPs marker reported by Iyer et al. (2006) (Fig.4.1.6). Presence of *Xa21* gene was screened across the genotypes using PTA248 marker (Ronald et al. 1992) and found the presence of DNA band size equal to that of IRBB21 in OL1 and OLD only (Fig. 4.1.7) while remaining accessions prominently showed the presence of recessive (susceptible) form of *Xa21*. Similarly, *xa13* linked functional marker designed from the polymorphic site between the promoters of recessive and dominant form of the gene was used to investigate the existence of recessive (resistance) allele in the natural population (Fig. 4.1.8). The recessive allele was found to be absent in all the tested wild and cultivated accessions.

4.2. MOLECULAR EVOLUTION OF *R* GENES *Xa21, Xa26, Xa27* AND *xa5*

4.2.1 Polymorphism at *xa21* locus

The total ORF comprising 3921 nucleotides of *Xa21* gene was sequenced from 18 accessions including one intron of 843 base pairs. Allelic sequences ranged between 3828 to 3921 nucleotides and the length of the aligned sequence was 3997 bp. The percentage identity of the nucleotide sequences of the alleles ranged between 95-99 % in comparison to IRBB21. There were a total of 196 mutations, 191 segregating sites, 113 parsimony informative sites, 64 InDel events and 78 singleton variable sites (rare kind of polymorphism, which is present once in the sample) within the entire sequenced region. The average number of nucleotide difference (K) was found to be 41.977. There were a

total of 136 single nucleotide polymorphism in exonic regions and 40 SNPs in the intron. The frequency of SNP was one SNP per 22.7 bp for the entire sequence, one SNP in 23 bp for the exonic region and one SNP in 21.5 bp for the intron. The SNP frequency was lower in *O. sativa* with one SNP/51.9 bp and one SNP/55.25 bp in *O. nivara*. The average frequency of InDel polymorphism was one InDel site for every 62.4 bp. Maximum InDel was observed in the first exon and least in last exon. The frequency of InDel was highest in intron followed by exon 1 and least in exon 2. Nucleotide substitution was observed throughout *Xa21* but highest nucleotide diversity was seen between first 500 bp of the coding region. The positions of substitution can be seen in Fig.4.2.1. The nucleotide diversity pi of the total sequence was 0.01112 and the intronic region had higher nucleotide diversity (π = 0.01328; Θ_w = 0.01500) than the coding region (π = 0.01055; Θ_w = 0.01433). Within the coding region, exon 1 (π = 0.01083) had substantially higher diversity than exon 2 (π = 0.00874). In intra specific polymorphism analysis, *O. nivara* exhibited higher sequence variation (π = 0.00782; Θ_w = 0.00846) than *O. sativa* (π = 0.00688; Θ_w = 0.00782). Nucleotide diversity was higher than InDel diversity (0.00238). All the alleles except OLD and OL1 showed an equal level of polymorphism (Fig.4.2.2). Frequency of A to G transition was more than other form of transitions and overall transition bias of the alleles was R = 1.221, where R = [A*G*k_1 + T*C*$K2$]/ [(A+G)*(T+C)]. Among four *R* genes, the transition bias observed in *Xa21* alleles was minimal.

Divergence at the synonymous site (π_{syn} = 0.01439) was higher than the non synonymous site (π_{non} = 0.00876). Most of the polymorphic sites in coding region resulted in silent substitution, both conservative and non conservative amino acid changes along with internal stop codons in all 18 alleles. K_a/K_s were determined by sliding window and following Jukes and Cantor model. We observed that value of K_s was greater than K_a leading to the ratio of $K_a/K_s < 1$, indicating the alleles were under the process of purifying selection with respect to IRBB21 (Fig. 4.2.3).

Linkage Disequilibrium was significant for both *O. nivara* ($R^2 > 0.2$, Fig.4.2.4A) and *O. sativa* ($R^2 > 0.35$, Fig.4.2.4B) ($0.001 < P < 0.01$). The LD plot of R^2 value, as a function of pairwise distance between polymorphic sites revealed slight decay

of the analyzed loci within 3500 bp in *O. sativa* and 4000 bp in *O. nivara*. While determining the recombination sites of this locus, a minimum of seven Rm events were detected considering only the coding region.

The Tajima's test showed no significant difference between π and Θ, thus, showing consistency with the neutral theory (Haseneyer et al. 2008). The Tajima's D was estimated to be -1.05476 ($P > 0.10$, not significant). Similarly, separate Tajima's D test for coding, non coding and intraspecies also showed negative and nonsignificant departure from neutrality. Summary statistics of polymorphism analysis of *Xa21* alleles is summarized in Table 4.1.

4.2.1.1 *Sequence polymorphism between resistant and susceptible alleles*

Among the alleles, we considered allele from TN1 and PB1 as susceptible allele since they were found to be susceptible to all the isolates and *Xa21* from IRBB21 as resistant allele. We analyzed these sequences together to identify unique polymorphisms which distinguished the two. We identified 95 fixed differences between IRBB21 and TN1; 80 differences between IRBB21 and PB1. Some of these polymorphisms were common with other alleles too. These changes resulted in both silent and mis sense mutations along with premature termination of proteins. At the sequence level other than the alleles from accession OL1 and OLD, remaining alleles were more similar to susceptible alleles.

4.2.1.2 *Molecular evolution of Xa21 gene*

Nucleotide sequences were used to deduce neighbor joining (NJ) and parsimony trees. The alignment for the phylogeny analysis was same as that of diversity analysis. The presence of two major clades I and II was supported by high bootstrap value. Clade I again comprised of two minor clades. IRBB21 and two alleles from *O. longistaminata* were clustered separately in clade II, showing that these alleles were highly distinct from the remaining alleles of different species (Fig.4.2.5). These three alleles emerged together in NJ, minimum evolution (ME) and maximum parsimony (MP). Between clade

I and II, there were 33 fixed differences across the entire sequence. Clade II showed slightly higher within clade diversity (π = 0.00946) than those of clade I (π = 0.00772). There were 55 singletons, zero parsimony sites in clade II and 61 singleton and 60 parsimony informative sites among alleles in clade I. Highest divergence among the alleles was observed between IRBB21 related and remaining alleles. Intraspecific divergence was quite low compared to interspecific divergence, however the divergence between alleles of *O. sativa* and *O. nivara* individuals was minimal (Fig. 4.2.6).

4.2.1.3 *Allelic expression analysis of Xa21*

In order to confirm the functional significance of alleles identical to resistant *Xa21* (IRBB21), we checked the expression of *Xa21* alleles in two accessions of *O. longistaminata* (OL1 and OLD). OL1 and OLD were highly resistant to BB as shown in Fig. 4.2.7. These two accessions were selected for expression analysis based on the sequence analysis and genotypic screening analysis with pTA248 marker. IRBB21 and IR24 were used as positive and negative control. Unfortunately, these two tested alleles didn't show any expression after the infection or before infection. We could see the expression of this gene only in control and infected IRBB21, however, slightly up regulated in infected plant as shown by Semi-Q RT PCR (Fig. 4.2.8).

4.2.2 Sequence analysis of *Xa26* alleles

Alleles of *Xa26* were isolated from individuals of 14 rice accessions of *O. sativa* and *O. nivara.* Complete sequence analysis was performed with the alleles from IR24, IRBB3, Minghui 63 and Zhachanglong (*O. sativa*). Sequence polymorphism was detected across 3875 bp of sequence covering 3315 bp exonic region, 106 bp intronic region and 458 bp 3'UTR. The total alignment length was 3987 bp. The percentage sequence similarity among alleles ranged between 95-99%. IRBB3 showed least sequence similarity with IC203740 and IR24 of 97 and 96%, respectively.

Nucleotide sequence polymorphism was analyzed across 18 accessions of different species at *Xa26* locus. There were a total of 304 polymorphic (segregating sites)

and 312 mutations with an estimated average nucleotide divergence (K) of 63.072. Among them 182 sites belonged to singleton variable and 122 parsimony informative site. There were 16 haplotypes among the 18 alleles and the haplotype diversity, Hd = 0.98 ± 0.028. The overall nucleotide diversity was equal to 0.01958 pi and the Watterson estimator 'Θ_w'= 0.02744. The highest nucleotide diversity was observed in the 3' UTR followed by coding region (π = 0.01684) (Fig.4.2.9). As many as 61 InDel events and 315 nucleotide substitutions were present. Frequency of SNPs was more than the InDels being one SNP in every 24.91 bp and one InDel/65.36 bp (Fig.4.2.10). The average InDel length was longer in the coding region (8.2271) than the entire gene (2.15) and its InDel diversity (π_i = 0.0012) was least different from that of coding region (π_i = 0.00091). The longest InDel was found in two *O. nivara* accessions (ON132 and 210223). Among allelic comparison highest polymorphism was found between IR24 and IRBB3. Many of the SNPs and InDels occurred in the exons which resulted in stop codons and premature termination of coding frame. Pattern of nucleotide substitution was estimated by maximum composite likelihood method where the overall transition bias 'R' = 1.421, the highest among the four genes**.** Here also, we could see 'G' to 'A' transition being most favored among other form of transitions. The transition and transversion rate was almost same in all alleles when compared individually with IRBB3 (Fig.4.2.11).

Similarly, as seen for *Xa21* alleles, the divergence at silent site (π_{syn} = 0.02299) was higher than the non synonymous site (π_a= 0.01507). When individual alleles were compared with IRBB3/Minghui 63, total number of synonymous substitution was more than the non synonymous changes resulting in $K_a/K_s < 0$ which illustrate that the alleles are undergoing purifying selection with reference to *Xa26* of IRBB3. In test of neutrality, the Θ value was greater than π giving negative Tajima's D (-1.297, $p > 0.10$). The difference was statistically not significant, consistent with the neutral theory (Moriyama and Powell, 1996; Heseneyer et al. 2008) and also illustrating absence of significant selection of this gene in the selected population of our study. Out of the 14 alleles isolated, only two alleles (M209A and 106133) showed intact ORF giving full length protein, whereas internal stop codons were found in the sequences of the remaining alleles. Amino acid substitution was found both in LRR and kinase domains.

Similar to *Xa21* alleles, interspecific nucleotide diversity was slightly higher in *O. nivara* (π = 0.02216) than *O. sativa* (π = 0.02139). However intraspecifically, *O. sativa* (254) had more number of mutations than *O. nivara* (191). Both the species showed negative Tajima's D with nonsignificant departure from neutrality. The linkage disequilibrium was found to be significant for both *O. sativa* ($R^2 \approx 0.54$) and *O. nivara* ($R^2 = 0.7$). Significant pairwise comparison was performed with Fisher's exact test and Chi-square test. Decay in LD was also observed in both the species when R^2 was plotted against pairwise distance between polymorphic sites (Fig.4.2.12). Summary statistics of different parameters are illustrated in Table 4.2.

4.2.2.1 *Molecular evolution tree of Xa26*

A dendrogram depicting the relationship and divergence of 18 *Xa26* alleles was plotted using Maximum Parsimony and NJ plot. Two major clades supported by strong bootstrap value were formed. Clade II had only four accessions, IR24 (*O. sativa*), 81832, 106133 (*O. nivara*) and IC203740 (*O. officinalis*) while resistant alleles from IRBB3, Minghui63 and Zhachanglong were clustered in clade I (Fig.4.2.13). The cluster presenting remaining alleles from *O. nivara*, *O. sativa* and *O. officinalis* were not distinguished at the species level. Intermixing of genetic component was observed which may be due to out-crossing.

4.2.3 Polymorphism at *Xa27* locus

The entire sequence of *Xa27* gene was amplified from 27 rice accessions using overlapping gene specific primers. The length of the sequenced alleles ranged between 2056 to 2733 bp. Overall sequence similarity of the entire gene ranged from 96 - 99%. We observed a deletion of three base pairs (TAC) in the exonic region, which lead to deletion of threonine (T) in OS 23-2, 81832, 210223, ON15, PAU933, R9 and ON132, hence, these alleles coded only 112 amino acids (Fig.4.2.14). There was also deletion of single nucleotide 'G' in M209A at 90th position resulting to a frameshift mutation in the protein sequence. The remaining alleles encoded protein product of 113 amino acids. For

the public interest the entire allelic sequences were deposited in the NCBI GenBank database and obtained the accession numbers as listed in Table 4.3.

The allelic sequences, including IR24 and IRBB27 were aligned and the multiple alignment length ranged between 2056 bp to 2780 bp. The alignment length included promoter, 5' UTR, 3'UTR and also alignment gaps. Among the 29 alleles, excluding sites with gaps, there were 103 polymorphic sites (in total, 1,999 available sites), 74 singleton variable sites, 40 parsimony informative sites and 32 InDel events ranging from 1- 300 bp in length. InDel polymorphism was also detected predominantly in the promoter region with fewer cases in exonic part. Some of the InDel sites were consistent between *O. nivara* and *O. sativa*. Deletion of ≈ 400 bp from the upstream region of *Xa27* alleles was noticed in ON132, 21028, 81832, PAU933, ON15, and 210223. The average number of nucleotide difference, 'K' of the entire gene was estimated to be 15.892. The average genetic diversity, 'π' was 0.00795 and 'Θ_w' equal to 0.01312 (Table 4.4). Among wild and landrace accessions, Mustilaiphou was found to be highly diverse from IRBB27 (π = 0.02291) than IR24 (π = 0.0163). The highest sequence variability was observed in the 5' flanking region of the alleles with 'π' value 0.00916 and least variability in the 3'UTR. The test of neutrality failed to give significant Tajima's and Fu & Li's D (P > 0.01) in *Oryza nivara* and *Oryza sativa* species, eventhough, Θ_w value was greater than π resulting in negative Tajima's D. Fixed number of nucleotide differences were detected among *O. sativa* and *O. nivara* alleles with few cases of variation within the species too. Intraspecific polymorphism was lower in *O. nivara* (π = 0.00790) than *O. sativa* (π = 0.00845). The most crucial polymorphism governing the function of *Xa27* was observed in the promoter where the new alleles shared similar UPT box to that of IR24 with deletion of three nucleotides 'AGA' at -51 position (from transcription start site) and substitutions from 'C' to 'A' (Fig. 4.2.15) which is crucial for binding of AVRXa27 that in turn activate the transcription of the gene (Romer et al. 2009). These three nucleotides were proved to be part of the UPT (UpRegulated by Transcription Activator) box, a stretch of 16 nucleotides just after the TATA box which actually binds the AVRXa27 protein (Romer et al. 2009; Bogdanove et al. 2010).

At the protein level highest number of amino acid polymorphism was observed in *O. nivara*. Nine of the alleles from species *O. nivara* had single amino acid replacement, whereas 21028 (2 changes), OS 28 and ON44 (3 changes) had more than one. Single amino acid replacement was seen only in three accessions of *O. sativa*. Both conservative and non conservative amino acid polymorphisms were present in the alleles, out of which, two conservative and three non conservative amino acid polymorphism lie in the first trans-membrane domain, however these mutations didn't make any major changes in the trans-membrane structure. SIFT analysis also predicted the tolerable nature of these amino acid mutations in the alleles. In M209A, TOPCONS prediction showed the absence of trans-membrane (TM) region and N-terminal signal peptide whereas, in the remaining alleles, presence of TM was predicted at two regions from 32-52 residues and 69-89 residues (Fig.4.2.16).

4.2.3.1 *Molecular Evolution of Xa27*

An unrooted NJ plot based on the nucleotide variations in *Xa27* locus was constructed to analyze the phylogenetic relationships among the alleles. An unrooted tree has no evolutionary direction since the history of evolution and origin of these alleles cannot be traced solely from the available sequence data. The entire coding and non coding part of the sequences were considered to generate the phylogenetic tree. When all the 29 sequences were considered for constructing the tree, due to low polymorphism among *O. sativa* accessions, a monophyletic clade was formed and was difficult to get a resolved tree with higher bootstrap value. From the phylogenetic tree, we observed two major groups of sequence type where ML was the lone member in the second group (Fig.4.2.17). This observation may be the result of insertion of ≈ 350 bp in the upstream region of ML. The separation shown were mainly at the species level as *O. sativa* and *O. nivara* formed separate clusters with the exception of one or two accessions, however, the alleles were not clustered on the basis of resistance/susceptibility genotype. The nucleotide as well as the protein sequence similarity of OS 28 (His-17, Phe-33, Val-40), ON24 (His-17) and ON44 (His-17, Phe-33, Val-40) found in the same clad suggest their similar origin. It was also observed that all the susceptible alleles from IR24, IR20 and TN1 etc. belonged to same group. Furthermore, *Xa27* gene in IRBB27 which was

originally derived from *O. minuta* was found nearer to *O. nivara* than IR24 (*O. sativa*) from evolutionary point of view. The resistant *Xa27* allele and those of other wild accessions evolved around the same period.

4.2.3.2 *Selection at the Xa27 Locus*

The K_a/K_s ratio was calculated for each allele with resistant (IRBB27) and susceptible (IR24) allele to study their divergence rate from the resistant and susceptible one. Comparison of each allele individually with IRBB27 and IR24 had shown that K_a/K_s was zero for 10 accessions namely IR20, IC203740, M209A, 102166, Musti laiphou, Mohem phou, R2, R4, PAU201 and ON34 which means only synonymous changes were present. It suggests that the selection was neutral for these accessions. Even though we found a lot of sequence diversity in and around the promoter and 5'UTR of Musti laiphou, non synonymous change was not detected. High rate of amino acid replacement was also observed in 50% of alleles where rate of non-synonymous changes by rate of synonymous changes (K_a/K_s) was >1. Alleles of TN1 and R23 when compared to IR24 showed very high $K_a/K_s \approx 8.00$, where the rate of amino acid replacement was occurring at a very fast rate than the synonymous changes which suggest rapid adaptive evolution (Bergelson et al. 2001). However, when compared with IRBB27, the rate of amino acid changes was substantially lower with $K_a/K_s = 0.788$. Along with positive selection, we could also see purifying selection in some of the alleles, ON24, 100963, IC384116, 106133 and 21028, where $K_a/K_s<1$.

4.2.4 Polymorphism at *xa5* locus

xa5 alleles were isolated from 15 accessions, out of which two *O. sativa* accessions were similar to recessive *xa5* and remaining alleles were similar to dominant *Xa5*. The GenBank accessions of the alleles are listed in Table 4.5. Comparative analysis of the alleles with respect to IRBB5 showed an average sequence identity in the range of 97-99%. The total alignment length was 6395 sites and size of the alleles ranged between 6140-6264 bp. The comparative sequence analysis showed 310 polymorphic sites (out of 316 total mutations), 105 singleton variable sites and 205 parsimony informative sites.

There were a total of 16 haplotypes (Hd = 0.993 ± 0.023), 87 InDel events and minimum of 21 recombination events. In general, frequency of substitutions and deletions were observed more in the intronic and non coding region than the exonic region. The mean frequency of InDel was one InDel for every 73.5 bp and that of SNP was one SNP per 27.56 bp. As observed in previous studies, transition/transversion bias was present with R = 1.066, where G → A and C → T transition were most prevalent. In the exonic region, we observed significant SNPs, both synonymous and non synonymous substitutions. Other than the type of substitution reported by Iyer and McCouch (2004), we also found 9 other SNPs in the coding region. Out of the nine SNPs, seven of them resulted in amino acid substitution (Fig.4.2.18) and among them S → A substitution in AC32753 and Q → L in PB1 were non conservative amino acid substitution. The frequency of SNP was higher in *O. nivara* accessions with one SNP/63.31 bp than *O. sativa* accessions (one SNP/70.27 bp). We also observed 39 fixed differences between *O. longistaminata* and remaining accessions. The summary statistics of diversity analysis showed a mean nucleotide diversity of π = 0.01324, Θ_w = 0.01539 and an average nucleotide difference of K = 78.904 (Table 4.6). Maximum diversity was observed in the second intron (Fig.4.2.19). Intraspecific polymorphism was found to be higher in *O. nivara* (π = 0.01106, Θ_w = 0.01029) than *O. sativa* (π = 0.00733, Θ_w = 0.00644). The rate of divergence, K_a/K_s were determined by sliding window and following Jukes and Cantor model. K_a/K_s value for each allele were determined in comparison to *xa5* (IRBB5) sequence. The value of K_a was more than that of K_s for PB1, TN1, 81832 and 81852; hence K_a/K_s were greater than one (Fig.4.2.20). Selection was neutral for AC32753 and Kalamekri while, remaining alleles were undergoing purifying selection with respect to IRBB5. We observed a similar trend in all the genes where the Tajima's test showed no significant difference between π and Θ, thus consistent with the neutral theory (Haseneyer et al. 2008). Though, some individual alleles showed $K_a/K_s>1$, the average Tajima's D was estimated to be -0.67191 ($P > 0.10$, not significant). Similarly, separate Tajima's D test for intraspecies also showed negative and nonsignificant departure from neutrality.

Linkage Disequilibrium was significant for both *O. nivara* ($R^2 > 0.75$, Fig.4.2.21A) and *O. sativa* ($R^2 > 0.55$, Fig.4.2.21B) ($0.001< P < 0.01$). LD plot of R^2

value as a function of pairwise distance between polymorphic sites revealed slight decay of the analyzed loci within 4000bp in *O. nivara*. The pairwise comparison was significant as determined by Fisher's exact test and Chi-square test. While determining the recombination sites of this locus, zero Rm event was detected considering only the coding region.

4.2.4.1 *xa5 gene tree and molecular evolution*

An unrooted phylogenetic tree (NJ plot) depicting genetic relatedness among the alleles was plotted using Mega 4. The entire sequence comprising 5'UTR, coding, non coding and 3'UTR were considered in plotting the tree. Two major clades supported with high bootstrap values were present in the plot. Resistant alleles were found to be more closer to alleles from *O. nivara* accessions than the susceptible alleles from *O. sativa* accessions. IRBB5 possessing resistant allele was found clustered in clade I along with other *O. nivara* accessions (Fig.4.2.22). Similar pattern of clade formation was also observed with Minimum Evolution and Maximum Parsimny tree. Within clade diversity was found to be higher in clade II ($\pi = 0.01288$) than clade I ($\pi = 0.01040$). There were 66 singleton variable sites, 109 parsimony informative sites in clade I and 76 SVS, 114 PIS in clade II. In disparity index analysis, the highest composite distance was found between 81852 (*O. nivara*) and OLD (*O. longistaminata*) (Fig.4.2.23).

4.2.4.2 *Allelic expression analysis of xa5 gene*

Expression analysis of *xa5* alleles identified from accessions AC32753 and Kalamekri were tested using IRBB5 and IR24 as resistant and susceptible, respectively. Phenotypically, AC32753 and Kalamekri showed a moderate level of resistance to *Xoo* strain DX011 (Fig.4.2.24). We were interested to see if any of the substitutions in the alleles other than the previously reported one had any influence in its expression or not. Only the basal level of expression of these alleles was observed in all the tested samples (Fig.4.2.25) suggesting that the level of expression of this gene is not related to its resistivity towards the pathogen.

4.3. STUDY OF DIFFERENTIALLY UP-REGULATED GENES AT TRANSCRIPT AND TRANSLATIONAL LEVEL DURING *XOO* INFECTION IN RESISTANT WILD RICE

Disease resistances in plants are governed by different resistance factors. As a plant's defense response during pathogen invasion, many host encoded proteins are differentially expressed. Proteomics and real time expression analysis are useful tools to identify and understand the functions of different genes and proteins involved in plant pathogen interaction. It also gives an opportunity to identify PAMP triggered immunity (PTI) and effector triggered immunity (ETI) related disease resistance genes. In this study we intend to compare differences between compatible and incompatible interactions as well as mock and pathogen inoculated plants.

The great potential of the wild accessions to resist BB was well observed in the screening study. Under this objective to identify the resistance factors, we focused on defense responsive genes which are race non specific instead of *R* genes which are race specific for the pathogen. To study rice – *Xoo* interaction and to identify novel defense related genes in BB resistance wild rice, 45 days old *O. nivara*, 81832, *O. longistaminata* (OL1) and *O. sativa* (IR24, susceptible control) plants were inoculated with DX011 by multiple leaf pricking method to cover the whole leaf area. The wild rice 81832 and OL1 showed symptomless resistance till six days after inoculation (Fig.4.3.1).

4.3.1 Transcript analysis of specific defense response genes after *Xoo* inoculation

To study the defense responses of rice to the pathogen, expression profiling was performed for a set of 13 defense related genes reported by different groups (Zhou et al. 2002). This was performed using cDNA of rice cultivars susceptible to the disease (*O. sativa* IR24) and wild relatives resistance to the disease (*O. nivara*, 81832 and *O. longistaminata*, OL1) in order to assess the role of different genes involved in *Xoo*-rice interaction in the tested wild rice. Out of the 13 genes only seven genes were found to be expressed during infection. They were *OsBAK1*, *RPLD1*, *RPLD2*, *BI26N5*, *EI1K8*, *EI10P9* and *EI35I3*. These genes showed constitutive expression, however various

degrees of increased level of expression were observed in response to pathogen infection. *BI26N5*, *EI1K8*, *EI10P9* and *EI35I3* are cDNA sequences which were identified by cDNA arrays and showed elevated expression during pathogen infection (Zhou et al. 2002). These genes were pathogen non specific and showed higher level of expression after infection both with *Xoo* and *P. grisea* during incompatible interaction (Wen et al. 2003). Only a basal level of expression was observed for *BI26N5* in all the samples except for 81832 (48h treated) and OL1 (48h treated). This gene is homologous to translation initiation factor (GOS2) and had been mapped on chromosome 7 (Wen et al. 2003). *EI1K8* and *EI10P9*, the single copy genes mapped on chromosome 1 and 6 respectively, showed constitutive basal level of expression in susceptible lines while differential expression was observed in resistant 81832 and OL1 infected (Fig.4.3.2). Both these genes were mapped to chromosomal locations previously identified as QTL for blast resistance (Wang et al. 1994; Chen et al. 2003). *EI10P9* is a homologue of Polyubiquitin (Rubq1). The relative change in expression was twofold higher in 81832 at 24 h time point than IR24 plants. *EI35I3* is also a single copy gene mapped on chromosome 7. It is homologous to thiamine biosynthetic enzyme. Its expression was increased more than 5 fold only in OL1 at 24 h control and infected and 48 h control while infected sample at 48h time point showed negligible expression. Susceptible IR24 and resistant 81832 showed least expression. *OsBAK1* is a pathogen recognition receptor (PRR) generally involved in PTI. It is a homologue of Arabidopsis BAK1, brassinosteroid insensitive associated kinase 1 (BAK1/SERK3). BAK1 in general is regarded as a key regulator of plant immunity (Chinchilla et al. 2007; Heese et al. 2007). Enhanced susceptibility of *OsBAK1* RNAi transgenic rice to *M. oryzae* was reported thus indicating its role in resistance against rice blast fungus (Park et al, 2011). In our study, enhanced level of expression of this gene was observed in resistant wild rice 81832 at 48 h sample. The fold change in expression was 1.5 fold higher than susceptible IR24. The expression level of this gene was slightly lower than 81832 in OL1. *RPLD*s are rice phospholipase D isoforms which are similar to PLD subfamilies of Arabidopsis. *RPLD1* is located on chromosome 1 and *RPLD2* on chromosome 5. The involvement of *RPLD*s in plant defense has been reported by Young et al. (1996). More than three-fold expression for *RPLD1* was observed only for OL1 at 48 h post infection while *RPLD2* showed differential expression for resistant lines (Fig.4.3.3).

4.3.2 Identification of proteins induced after *Xoo* inoculation by 2DE and MALDI analysis

Two dimensional gel electrophoresis was performed for 81832 leaf samples of three days post infection with DX011 and water infiltrated in order to identify proteins induced during the host pathogen interaction process. IPG strips with pH gradient 4 to 7 were rehydrated with protein samples and then placed on isoelectric focusing machine (IEF). The proteins were then separated in a pH gradient manner based on their difference in isoelectric point (pI). The IEF separated proteins were then subjected to 2DE and we observed approximately around 200 proteins being separated (Fig.4.3.4). Using Image Master 2D platinum 5 software, the differentially expressed proteins were grouped based on their relative expression as up regulated or down regulated proteins. Though the resistance symptoms appear later after six days post infection, we could observe changes internally at the molecular level much earlier. To understand the function of differentially expressed proteins, each protein was identified by MALDI-TOF MS analysis and protein sequencing (Fig.4.3.5). The proteins were identified from the hits of protein sequence databases available in NCBI GenBank and Swissport.

Among the expressed proteins 61 protein spots were differentially expressed or up regulated while 47 spots were up regulated above 1.5 fold. Among this, 30 spots were given for analysis in which only 18 of them were successfully identified and 29 spots were undetermined protein. The sequence and identity of these proteins could not be determined either due to low quantity for MS or the N-terminus of the protein was blocked. The identified proteins were classified into five categories, they were energy/metabolism (6/47, 13%), defense (5/47, 10%), transcription factor (2/47, 4%), hypothetical (5/47, 11%) and undetermined protein (29/47, 62%) (Fig.4.3.6; Table 4.7).

Earlier reports tell that the proteins related to energy and metabolisms were commonly expressed during stress responses such as RuBisCo LSU (spot no. 1127). There was fivefold increase in expression of this protein in infected plant. Other proteins related to energy and metabolisms were triose phosphate isomerase (spot no. 1293), phosphoribulose kinase (spot no. 1196), fructose-bis phosphate aldolase (spot no. 1226),

Ribose-phosphate pyrophosphokinase 2 (spot no. 1216) and esterase (spot no. 1295). Any new proteins related to defense response unique from previous reports were not identified in this study as majority of the proteins could not be identified and analyzed. The identified putative defense related proteins were methyl transferase (spot no. 1190), cysteine synthase (spot no. 1457), PHD finger protein (spot no. 1302), germin like protein (spot no. 1317), putative thiazole synthase (spot no. 1218) and serine threonine protein kinase (spot no. 1455). The level of expression of these proteins was elevated more than two fold in pathogen inoculated plant than the control. During disease resistance response, methyltransferase helps in reinforcement of the plant cell wall (Pakusch et al. 1989) and biosynthesis of phytoalexins. In response to pathogen attack, Germin like proteins (GLP) are involved in conversion of active oxygen species (AOS) to H_2O_2 and hence help in its accumulation in plants (Bolwell and Wojtaszek 1997), thereby sending downstream signals for systemic acquired resistance and helping in cell wall strengthening. PHD finger proteins are involved in DNA recognition, RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding. Thiazole synthase may have roles in adaptation to various stress conditions and in DNA damage tolerance. Serine threonine kinases are proteins generally involved in downstream signaling of defense response and also many of the identified disease resistance genes are serine threonine kinases. Remaining proteins were two transcription factor proteins or putative stress related (spot no. 1290 and 1302) and five non determined or hypothetical proteins. Their level of expression was twofold increased in inoculated plant except spot number 1334 with only 1.5 fold changes.


Chapter Five
Discussion

5.1 FINDING NEW SOURCE OF RESISTANCE

To locate potential sources of resistance which will provide durable and sustainable resistance against bacterial leaf blight, we screened both cultivated and wild species of rice with five highly virulent *Xoo* strains. We could see variation in pathogen recognition and resistivity at different *R* gene loci which may be the result of genetic diversity both at the host and pathogen level. This genetic and natural diversity promised to be a foundation for developing stronger and durable disease resistance crops. Genotypically, the resistant form of the alleles was found to be less frequent in nature.

5.1.1 *Xa21* and *xa13*: Potent genes against virulent *Xoo* strains

Based on the screening analysis, *Xa21* and *xa13* showed strong resistance against five virulent *Xoo* isolates DX011, DX127, DX020, DX015 and DX133 indicating the presence of corresponding *avr* gene in these isolates. *Xa21* codes for the LRR receptor like kinase and act as a plant pathogen recognition receptor (PRR). The resistance of *Xa21* to these five isolates suggests their production of AVRXa21. AVRXa21 is secreted into the plants through type I secretion system (Lee et al. 2006). *xa13* confer resistance in recessive form because the dominant *Xa13* also known as *Os8N3* which is a susceptibility gene. The strong resistance by this gene to all the isolate means that *Os8N3* specific type III TAL effector PthXo1 is also present in those isolates as *xa13* is un-induced by this effector and plants with *xa13* confer resistance to those strains whose virulence depend only on PthXo1 (Yang et al. 2006). Among all the reported *Xa* genes, *Xa21 and xa13* were found to be broadly resistant to most of the Indian *Xoo* isolates. *xa5* exhibited partial resistance to three strains except DX133 and DX020. *xa5* provide race specific resistance and may attenuate TAL effector activation of susceptibility genes in the host plant (Iyer-Pascuzzi et al. 2008). The combination of three genes *xa5 + Xa21 + xa13* in IR59 provides strong resistance against the pathogen. The advantage of this three gene combination in popular elite varieties had also been reported (Joseph et al. 2004; Sundaram et al. 2008). While the remaining genes *Xa1, Xa3/Xa26, Xa4, Xa7, Xa10* and *Xa11* were susceptible to all the five isolates which was used in this study. These genes were completely defeated to the tested isolates. The

breakdown of some of these genes was reported due to large scale and long term cultivation of plants carrying these genes mainly because of the expansion of the virulence factor of the pathogens during the course of evolution (Mew et al. 1993).

5.1.2 Wild species: The hub of resistance genes

In this study, 52% of the wild *Oryza* species were found to be highly resistant to all five *Xoo* isolates suggesting the presence of strong *R* genes. Only three accessions, *O. nivara* (42446 and 21990) and *O. alta* (IC384076) were completely susceptible to all isolates. Remaining 10 accessions showed resistance to one or two *Xoo* isolates. Plant researchers have started realizing the importance of wild genetic resources and efforts have been made to explore the wealth of wild resources for many agronomical traits. More and more resistance genes have been identified from wild species. Some of the widely used and highly resistant *Xa* genes had been identified from wild rice such as *Xa21* (from *O. logistaminata*), *Xa27* (*O. minuta*) and few newly identified genes *Xa30* (Kuljit et al. 2008) and *Xa33* (Natrajkumar et al. 2012) from *O. nivara*. Other than these two genes from *O. nivara*, we could find out the presence of strong *R* genes in this species based on our study too. Different accessions of *O. nivara* which provided high resistance against BB promises to be the reservoir of future BB resistance genes. There are also evidences of other *Oryza* species conferring wide spectrum resistance to *Xoo*. Two different groups from India (Srinivasan and Gnanamanickam, 2005; Ram et al. 2011) identified the presence of strong resistant like genes in *O. rufipogon* accessions. The accessions used by these groups were highly resistant to diverse multiple Indian and Philippines *Xoo* isolates.

Among the cultivated rice and landraces, only ML and MP were resistant. The cultivars which were previously reported to be resistant to BB were mostly found susceptible and could resist moderately to one or two isolates. Out of the 42 introgression lines of *O. nivara* x Swarna, only four lines exhibited moderate resistance to the disease. We could observe limited scope of finding resistance gene from cultivars and landraces. Both wild and cultivated accessions showing wide spectrum disease resistance may be further deployed for introgression into elite cultivars.

5.2 MOLECULAR EVOLUTION AND DIVERSITY OF *XA* GENES

"Genetic variations are the foundation of crop improvement."

(Zhang et al. 2013)

Natural variation(s) in the alleles of *R* genes due to different form of polymorphisms play an important role in providing phenotypic diversity or variability with respect to disease resistance. Considering the existence and emergence of virulent pathotypes of the bacterial blight pathogen, there is a need to exploit the potential of allele mining by identification of novel alleles of *R* genes. Study on molecular evolution and understanding genetic diversity of known *R* gene(s) will also help us in discovering important region(s) controlling disease resistance. Host-pathogen co-evolution resulted in maintenance of allelic diversity at *R* gene in natural population (May and Anderson, 1983). Pathogens are believed to play an important role as a selective agent during the course of *R* gene evolution (Kover and Schaal, 2002). In this study, we have analyzed the sequence polymorphism of four major bacterial blight resistance genes *Xa21, Xa26, Xa27* and *xa5* in order to identify favorable alleles and also to uncover genetic variants of these genes in wild and cultivated species. This analysis is the first of its kind for BB related genes and may also help in molecular evolution study and mining of useful alleles for crop improvement. In our study, we could observe similar phenomenon in their pattern of sequence variation and divergence. Increasing the number of natural population with the use of pathogen possessing specific *Avr* will help in better understanding of the co evolution of host resistance and pathogen.

5.2.1 Natural diversity in *Xa21* alleles

Xa21 gene is one of the most commercially exploited and studied BB resistance genes. This gene has been reported to confer wide spectrum resistance to diverse *Xoo* strains across world's rice growing countries. Yet, there are reports on the susceptibility of *Xa21* to few pathotypes of Indian and South Asian isolates, (Bustamam et al. 1996; Lavanya et al. 1998; Goel et al. 1998; Ochiai et al. 2000; DRR 2002) which may be due to mutation and evolution in the isolates. Consequently, finding a candidate alternate

form of the gene with the scope of better performance will help in combating with new virulent strains of the pathogen.

Among alleles, OL1 and OLD showed maximum homology with IRBB21 which defined their relatedness and similar origin as *Xa21* in IRBB21 was derived from *O. longistaminata*. The presence of two divergent groups was evident from the divergence table and the phylogram (Fig.4.2.6). *Xa21* alleles showed higher nucleotide diversity (pi = 0.01112) than *Xa27* and other *R* genes such as *Pita* (0.00235, Huang et al. 2008) from rice. The nucleotide diversity of rice genes was comparatively less than those of *Arabidopsis* and other crop plants. At the species level, we found a little higher diversity in *O. nivara* than the cultivated *O. sativa*. The level of diversity may increase if we include more number of natural population and species. The decrease in diversity among *O. sativa* may be due to the domestication bottleneck (Eyre-Walker et al. 1998; Buckler et al. 2001; Zhu et al. 2007). As intron evolves rapidly than exon (Small and Wendel, 2000), sequence polymorphism was higher in the intron than coding region, which is often seen in most of the genes studied. As Haldane (1949) stated that high polymorphism is expected at the locus involved in pathogen recognition, we observed elevated level of nucleotide and amino acid polymorphism at *Xa21* locus. Numerous substitutions and InDels resulted in amino acid polymorphism however, making the alleles non functional including two alleles from *O. longistaminata* due to premature termination of codon, which was also evident from expression analysis. These two alleles failed to give any expression even after infection with pathogen while the positive control IRBB21 was found expressing (Fig.4.2.8). This result hints us that only marker analysis is not sufficient to find resistance allele. Looking for *Xa21* alleles in more number of *O. longistaminata* accessions and other wild species may help in finding the desired or novel form of the allele.

Universally transition bias is prevalent across different Kingdom, genera and species. It was observed to be a common phenomenon during the course of evolution. The alleles showed more number of transition from ‘G’→‘A’ and ‘C’→’T’ than vice versa. Similar feature has been recorded in plants (Olmstead et al. 1998), nematodes (Blouin et al. 1998), Drosophila and mammals (Petrov and Hartl, 1999) and prokaryotes

(Kowalczuk et al. 2001). High rate of 'C' to 'T' transition may be due to methylation of cytosine, which increases the probability of this form of substitution (Petrov and Hartl, 1999).

Statistically, Tajima's D value suggest the absence of natural selection at this locus, however there are also possibility of natural selection as we can see lots of polymorphism among the alleles leading to amino acid substitution and immature termination of protein. There were also two divergent groups in the phylogenetic tree. Some of the *R* genes studied for diversity analysis failed to show statistically significant Tajima's D indicating absence of selection, however, the likelihood of selection was indicated by negative Tajima's D value which shows excess of rare variants as in the case of *RPP13* (Rose et al. 2004) and *Pto* (Rose et al. 2007). There may be relaxed selection pressure or the gene might be evolving neutrally during which the gene is depleted from the population through deletion, frameshift or nonsense mutation (Rose et al. 2004).

Linkage disequillibrium was measured by plotting R^2 as a function of pairwise distance between the SNPs. The LD decay was more for *O. sativa* accessions. The LD measured was on an average the same as that of other plants reported. Locus specific selection increases LD. Higher R^2 was observed in *O. sativa* than *O. nivara* and the extent of LD was more in *O. nivara*. Recombination and cross pollination influences the decay of LD. There were a minimum of seven recombination events and as rice is a self pollinated plant, there were only slight decay of LD. Rapid decay of LD is because of high recombination rate, which mainly happens in the cross pollinated plants.

Though OL1 and OLD showed the presence of identical allele of *Xa21*, functionally these alleles failed to show any expression after pathogen infection. Phylogenetically and sequentially, these two alleles were nearest to IRBB21 (resistant allele) however, non functional due to nonsense mutation in the coding region.

5.2.2 Natural variation in *Xa26* alleles

Xa26 also code for LRR receptor kinase just like *Xa21*. IRBB3 was susceptible to the five virulent isolates used in our study, but we were interested in understanding its molecular diversity statistics hence this was included in our study. For the diversity analysis only the coding region with the short length intron was considered. Among the four genes, the highest sequence polymorphism was found among *Xa26* alleles. Within the selected accessions of different species, maximum variation was observed between IRBB3 (reference resistance allele originally derived from *O. sativa* cv. Minghui 63) and IR24 (reference susceptible allele). Hence, the highest sequence variation among *Xa26* alleles was found at intraspecies level between resistant and susceptible allele. This shows that natural selection favored polymorphism between resistant and susceptible alleles rather among species. In fact, IRBB3 was found susceptible to all five *Xoo* isolates, hence, this allele may be considered susceptible. Yet, it is complicated to distinguish the alleles from resistant accessions to be susceptible or resistance as phenotypically they were resistant to the disease. In nature, there are many known and unknown *R* genes responsible for resistance to *Xoo* but we don't know yet the resistance factor present in those wild and land races. The functionality of these alleles also couldn't be determined as *Xa26* is a constitutively expressed gene.

Most of the studied features of selection and polymorphism analysis, such as K_a/K_s, transition bias, Tajima's D, linkage disequilibrium, inter and intra species diversity, presence of internal stop codons etc., at *Xa26* locus is similar to that of *Xa21*. Higher K_s value indicates purifying selection, however non significant Tajima's D value again suggests influence of selection at this locus. The probable reasons for this observation have been discussed in previous section of *Xa21* alleles. The phylogenetic tree also depicts the intermixing of genetic component during the course of evolution, perhaps due to out-crossing between species. As *Xa26* is predominantly found in the Chinese cultivars, including more number of accessions from this region would enrich the information on diversity analysis and will be able to identify more resistant form of allele of the gene.

5.2.3 *Xa27* is less frequent in natural population

Unique SNPs were present in the 5'UTR of IC203740, OS 23-2, Musti laiphou, PAU933, PB1, M209A, R9 and nearby region of UPT boxes in Mohem phou, M209A, R9, OS 28, ON24, PAU933, Pusa Basmati1, R4 and 106133 (*O. nivara*). Fixed polymorphisms were also observed in wild rice accessions ON44, ON24 and OS 28 both in the protein coding as well as the UPT box region and these accessions were clustered in the same clade at the phylogenetic tree. At the species level, we observed a set of SNPs and InDels unique to *O. nivara* and *O. sativa*, respectively, in the genomic region encompassing *Xa27* but some polymorphic traits were common in both. We could see less intra-specific polymorphism in both *O. sativa* and *O. nivara* which may be because of small sample size. Moderate inter-specific polymorphism was detected in the promoter region of *Xa27* locus. A lesser amino acid polymorphism at the intra-specific level may be due to purifying selection resulting in conservation of amino acid sequence as reported in the case of *R* gene loci *RPS5*, *RPM1*, *RPS2* and *RPP13* in *Arabidopsis thaliana* (Caicedo et al. 1999; Mauricio et al. 2003). Low level of polymorphism was also reported for other *R* genes in rice such as most widely studied *Pita* conferring resistance against *M. grisea* (Huang et al. 2008; Yoshida et al. 2009; Lee et al. 2011). At the genome level, IC384116 (CCDD genome), IC203740 (CC) and IRBB27 didn't show much sequence divergence from the AA genome (*O. sativa* and *O. nivara*); earlier studies have also shown close relatedness and colinearity between AA and CC genome (Vaughan 1994; Zou et al. 2008; Feng et al. 2009). Highly conserved colinearity between AA and CCDD genome was also reported by Huang et al. (1994).

Xa27 alleles showed an average polymorphism of π_s = 0.00795. This low variation is comparable with that of *O. sativa*, which has a nucleotide diversity (π_{silent}) in the range of 0.0011–0.0035 (Tang et al. 2006; Zhu et al. 2007) and also '*Pita*' locus of *O. rufipogon* with π = 0.00235 (Huang et al. 2008). Nucleotide divergence for *O. nivara* accessions were also much lower with 'π'= 0.0079. Yoshida et al. (2009) also reported very low diversity at silent site of '*Pita*', π_s = 0.0101. However, we may get higher allelic variability if we include more number of species such as *O. rufipogon*, *O. longistaminata* or *O. minuta* etc. in the study. We may also conduct this study in

different sets of *Oryza* species for better understanding of *Xa27* molecular evolution. The nucleotide polymorphism in 5' non coding region (π = 0.00916) was shown to be much higher than that of the entire region. This may be due to the low polymorphisms of the 3'UTR and exclusion of sites with gaps from the analysis. We found that the genetic diversity of *Xa27* alleles was in and around the range of those reported defense related genes. Nucleotide diversity of other defense related genes such as *Adh3* locus in wild barley has π value 0.0219 (Lin et al. 2001); *Adh* loci of allogamous species depicted Pi values from 0.00204 to 0.01742 (Cummings and Clegg, 1998). In *A. thaliana*, a chitinase-encoding gene has π = 0.0104 (Kawabe et al. 1997) and *Pto*, disease resistance gene of wild tomato (*Lycopersicon*) was reported with π = 0.012 (Rose et al. 2007). Bakker et al. (2008) also reported that seven genes encoding pathogen-related proteins, had π_s = 0.00183 and π_a = 0.00126.

The ratios of divergence rate of most of the *O. nivara* accessions from IRBB27 were high with $K_a/K_s > 1$ which suggest the occurrence of positive selection or adaptive evolution (Bergelson et al. 2001). Most of the *O. sativa* accessions exhibited neutral selection with $K_a/K_s = 0$, while few *O. nivara* and introgression lines (ILs) have $K_a/K_s < 1$ illustrating the presence of selection constraint. Even though, M209A (*O. nivara*) showed zero divergence rate from IRBB27, at the translation level, it encoded a non functional protein due to frame shift mutation. This miscalculation arose due to the exclusion of single bp deletion from the analysis by the software. We also observed similar pattern of divergence rate between the susceptible allele (IR24) and the new alleles. A strong positive selection was shown in TN1, R23 and 210223 (*O. nivara*) against IR24. This observation was due to very low polymorphism in silent sites (π_s) of the said alleles. In general, adaptive evolution is most commonly observed in regions involved in host pathogen recognition (Endo et al. 1996) such as in the case of *Xa21* and *Xa21D*, occurrence of non-synonymous mutation was higher than synonymous mutation in LRR domain (Wang et al. 1998). For thorough analysis of adaptive and positive selection of *R* genes, we should consider only those sites which are proven to play an important role in recognition (Wang et al. 1998). "Amino acids evolve at a faster rate in functionally important regions of R-proteins than the corresponding rate of synonymous changes" (Meyers et al. 1998; Bergelson et al. 2001). Similar pattern was observed in

this case with majority of the amino acid replacement taking place in the TM domain despite the fact that this substitution didn't cause major changes in the pattern of the TM. Probably, *Xa27* were evolving under relaxed constraint. The role of these TMs in pathogen recognition is yet to be explored in order to understand the resistance mechanism of *Xa27* completely.

When Tajima's D was calculated for *O. nivara* accessions and *O. sativa* accessions, the difference between π and Θ was very less which failed to give a significant 'D' value. The non-significant Tajima's D and Fu and Li's F may be due to small sample size and random sampling. Simultaneously, we cannot overrule the possibility of selection occurring at this locus. Firstly, as the samples under study were selected randomly from a natural population and not from an interbreeding population, secondly, *Xa27* locus interacts directly with the pathogen where the possibility of co-evolution of both plant and pathogen exist. So, a population level study of *R* gene locus will yield accurate measurement of selection (Caicedo et al. 1999). Other than looking only at the coding region for selection, we found high level of nucleotide polymorphism at the promoter and 5'UTR among the accessions which possessed many regulatory sites. The roles of these polymorphisms in the gene function are yet to be explored and this region may also be under natural selection. More importantly, the deletion and substitution in the $UPT_{AvrXa27}$ box in all alleles except IRBB27 will make these alleles unable to signal AvrXa27 recognition (Romer et al. 2009) which suggest the selective maintenance of dominant *Xa27* or we can say gain of function of this allele during the course of evolution among the natural population. This might also be the possible case of an ''evolutionary arms race" or a negative frequency dependent selection. A more detailed study with more number of natural populations may provide crucial insights to understand the exact evolutionary dynamics of *Xa27* locus.

Twenty seven alleles of *Xa27* showing numerous sequence diversity have been identified. Other than having few changes in the coded protein sequences, the UPT box was similar to that of IR24. We presume that the binding affinity of AvrXa27 to the UPT box of these alleles will be weak due to presence of defective UPT box and hence binding of *avrXa7* will be unable to induce the expression of the alleles. It may be

interesting to study the different alleles where amino acid changes had occurred, after fusing these newly found coding sequences with a functional $UPT_{Avrxa27}$ by promoter engineering as described by Romer et al. (2009) and Hummel et al. (2012). Clear understanding of functional difference among alleles will help in deciphering its structure and function (Bergelson et al. 2001) as well as studying the molecular diversity of *avrxa27* will clear the co-evolutionary process of this *R-avr* gene. The available sequence data of different alleles and the level of diversity found at the *Xa27* locus may provide great opportunity for studying the evolution of plant–pathogen interactions as well as a chance to study both evolutionary and genetic aspects of this unique *R* gene.

5.2.4 Sequence variation at the naturally occurring *xa5* alleles

The summary of statistics describe that *xa5* also show common and similar trend for most of the features of polymorphism analysis to those of previously discussed genes as the samples selected for analysis were randomly selected from natural population and require increasing the number of samples. In the entire genomic region, as intron evolves rapidly than exon (Small and Wendel, 2000), sequence polymorphism and frequency of SNPs and InDels were higher in the intron and noncoding region than coding region which is often seen in most of the genes studied. Recent studies have shown that the disease resistance or susceptible phenotype is not only due to nucleotide variation in the coding region, however, nucleotide changes in the non coding and regulatory sites also played an important role in determining the phenotype (Gu et al. 2005; Romer et al. 2009). This variation in the naturally occurring alleles may bring changes at the level of gene expression and protein function which may produce functional alteration and also result in phenotypic variation of a trait (Doebley and Lukens, 1998; Buckler and Thornsberry, 2002). We also observed decrease in diversity among *O. sativa* which may be due to the domestication bottleneck (Eyre-Walker et al. 1998; Buckler et al. 2001; Zhu et al. 2007).

Functionally, resistant IRBB5, susceptible IR24 and the remaining tested alleles AC32753 and Kalamekri showed almost equal level of expression both under infected and non infected state indicating that the level of expression of the alleles might not

influence its function, however, may be their ability/non ability to bind DNA polymerase and manipulate the induction of susceptibility genes determine their role in resistance. Functionally, AC32753 and Kalamekri were resistant to the disease and these two alleles may be employed for gene pyramiding into elite cultivars. The functional significance of amino acid substitutions in AC32753 (*O. sativa*), 81832 and 81852 (*O. nivara*) may be elucidated however, structurally amino acid substitutions in 81832, 81852 (*O. nivara*) and AC32753 occurred in the loop region, hence functional changes in these alleles due to these mutations are least expected.

5.2.5 Limitations in distinguishing between susceptible and resistant alleles

As per observation in this study, many of the wild accessions showed strong resistance to some of the highly virulent *Xoo* strains. However, it was difficult to distinguish whether the identified allele was responsible for resistance or susceptible as there was possibility of presence of many strong resistance genes which confer resistance to the plants. We could differentiate between susceptible and resistant allele based only on reference sequence information available. One approach to identify the disease resistance factors present in the wild genetic resource is through RNAi construct and developing transgenic lines. However, this approach is time consuming and tedious. It will be difficult to validate the function of all alleles with this method. Yet, it would be of great interest to identify different unknown resistance genes present in various wild species of rice. It may be achieved by screening wide collections of germplasms with more number of markers reported for identified *Xa* genes. Crossing the wild species and landraces harboring resistance genes with a cultivated susceptible variety, followed by phenotypic screening and background cleaning by marker assisted backcross selection would facilitate the identification of new BB resistance genes present in the wild rice species. Another method is by comparing the whole transcriptome and proteome profile of infected and non infected plants of both resistant and susceptible rice. It will help us in identifying some key genes and proteins up-regulated during the infection process.

5.3 STUDY OF DIFFERENTIALLY UP-REGULATED GENES AT TRANSCRIPT AND TRANSLATIONAL LEVEL DURING *XOO* INFECTION IN RESISTANT WILD RICE

Plants counteract pathogens by activating innate immune response (Jones and Dangl, 2006). Many defense response genes also play important roles in keeping plants protected during infection. They are mostly pathogen and race non specific. There are several key components in the signal transduction pathways or secondary pathways as intensively reported for Arabidopsis. Here, we discussed the transcript and proteomic analysis of rice defense response induced by *Xanthomonas oryzae* pv. *oryzae* infection. Real time PCR of previously reported defense response genes in rice was performed for *O. nivara*, 81832 and *O. longistaminata*, OL1 and two-dimensional gel electrophoresis (2DE) was utilized to identify protein spots up-regulated by *Xoo* infection in highly resistant wild species *O. nivara*, 81832.

5.3.1 Transcript profile of defense response genes

Transcript profile of wild species of rice plants after *Xoo* infection will give insight into diverse resistance factors activated during the interaction. Out of 13 genes analyzed in incompatible and compatible interaction only seven of them were differentially expressed at 24HPI and 48HPI. The remaining genes were found unexpressed at this time point may be due to early response of those genes during infection process. Both compatible and incompatible interactions showed constitutive expression. Elevated expression of *EI1K8* more than 3 fold changes was observed in resistant *O. nivara* 81832 and *O. longistaminata* OL1. However, expression of this gene in susceptible IR24 and water infiltrated samples were minimal at basal level. Similar pattern was observed in IRBB13, Minghui 63, C101LAC, C101A51 and IRBB4 after inoculation with different strains of *Xoo* or different isolates of *P. grisea* (Wen et al. 2003). *EI1K8* showed a high degree of sequence similarity to a light-regulated gene (Reimmann and Dudler, 1993). A little elevated level of expression was observed in resistant wild rice for *BI26N5* while *EI10P9* showed only basal level of expression in all susceptible and resistant rice except *O. nivara* 81832 at 24 h post infection. The *EI35I3*

is a single copy gene mapped to chromosome 7 (Wen et al. 2003). In our case, its significant expression was observed only in resistant *O. longistaminata* OL1. Wen and group (2003) had reported its elevated expression in IRBB13 and Minghui 63 after infection both with *Xoo* and *P. grisea*. Two isoforms of another rice gene phospholipase D (*RPLD1* and *RPLD2*) also showed significantly higher level of expression in resistant wild rice than susceptible IR24. Rice phospholipase D is an enzyme involved in fatty acid metabolism. This gene is reported to be an important enzyme involved in signal transduction, stress responses, protein trafficking, and membrane metabolism (McGee et al. 2003). During bacterial blight stress, increase level of expression was observed for both the isoforms in resistant wild rice than susceptible IR24. *RPLD2* showed higher fold changes in expression than *RPLD1*. Up-regulation of this gene during *Xoo* infection had also been reported by McGee et al. (2003). The last but not the least gene *OsBAK1* is the rice homolog of Arabidopsis *BAK1* (*AtBAK1*) and both of them were highly conserved. Over expression of this gene in rice resulted in phenotypic changes such as plant height, leaf erectness, grain morphology and disease resistance responses (Li et al. 2009). With respect to disease resistance, *OsBAK1* possess leucine rich repeat (LRR) domain and is termed as pathogen recognition receptor (PRR) (Roux et al. 2011). PRRs are generally involved in PAMP triggered immunity. During bacterial blight stress, we observed up-regulation of *OsBAK1* in incompatible interaction with *O. nivara* and *O. longistaminata.* The level of expression was higher in *O. nivara*, 81832 than *O. longistaminata,* OL1. The highest fold change in expression was observed at 48h post infection. Among the induced genes, transcript levels of some of the genes were higher in either of the two resistant wild species or both. At this point, we need further confirmation for these genes to be targeted for crop improvement. These genes may also be playing important role in making the plants resistant against BB, however, we need further experiment like breeding with elite susceptible variety in order to identify the major resistance factor present in these wild rice.

5.3.2 Protein profile after *Xoo* infection in *O. nivara*

The approach of proteomics for plant pathogen interaction is a late blossoming field and had been adopted to compare protein profile between compatible and

incompatible interactions between Java 14 and *Xoo* strain (Mahmood et al. 2006). Until now there are very few reports on proteome analysis of defense signaling pathways involving plant pathogen interaction in rice especially the wild rice. In the proteome study, due to interference by polysaccharides and other metabolites, the total number of spots or protein separated on the gel were comparatively less than previous reports. The proteome analysis was performed only in *Xoo* infected and water infiltrated resistant wild rice and only the up-regulated proteins were identified. Comparative analysis between compatible (susceptible to BB) and incompatible (resistant to BB) was not performed in this study. Previous study by our group has observed the increase in expression of metabolic and energy related proteins, however, none of the stress and defense related proteins were up-regulated in compatible interaction. Much of the studies related to stress have shown increase in expression of most of the metabolism and energy related proteins specially the Rubisco large subunit (Kumar et al. unpublished; Xu et al. 2013; Yang et al. 2013). It could be possible that due to decrease in the total photosynthetic leaf area in the plant during infection, the functions of these proteins were enhanced to obtain the normal energy state. The Rubisco large subunit and other metabolic proteins were found to be commonly up-regulated or differentially expressed during incompatible interaction, after Jasmonic acid and probenazole treatment (Mahmood et al. 2006). Among the identified proteins only five proteins were related to defense. As a major drawback of this kind of analysis, it was unsuccessful to identify most of the proteins either due to low intensity/low quantity or blocking of the N-terminus of the protein. The fold changes in expression of the proteins were not so high as the comparison was done between control and infected resistant plant but not between susceptible and resistant plant. Most of the identified defense related proteins were commonly expressed during stress while any major resistance genes could not be identified in the wild rice species as majority of the spots were undetermined. Further confirmation of function with expression analysis will help in validating this proteome analysis. At this stage it is difficult to establish the functional network of proteome involved during the host pathogen interaction as some proteins of interest could not be identified may be due to low abundance or condition-dependent expression added with technical shortcomings and high cost.


Summary & Conclusion

SUMMARY AND CONCLUSION

In this thesis, we investigated the natural variation in genes which are involved in bacterial blight resistance. The aim of this study was to identify alternate source and form of resistance which will make elite susceptible rice varieties stand strongly against virulent *Xoo* strains. We addressed this question from the perspective of exploring and exploiting the available genetic resources for crop improvement.

Studies of evolutionary dynamics of disease resistance genes started in the last two decades, mainly in Arabidopsis and tomato plants. Lately, reports on molecular evolution of disease resistance genes in rice came up for fungal blast disease caused by *Magneporthe grisae*, however, the study on the natural sequence diversity and allelic polymorphism of bacterial blight resistance genes are limited. The pathogen imposed selection of *R* gene or the co evolution of host and the pathogen in plants is quite complex and different from other source of natural selection. This co evolutionary process resulted in maintenance of allelic variation of *R* genes, hence ensuing from the balancing selection (May and Anderson, 1983).

Phenotypic screening of 104 accessions including wild species, cultivars, landraces and introgression lines for resistance against five different *Xoo* strains DX011, DX127, DX020, DX015, and DX133 revealed that the wild accessions were highly resistant towards BB than the cultivated species. Other than phenotypic screening, we also tested for the presence of resistance alleles of *Xa* genes in the resistant plants. Here, Only *Xa21* (IRBB21) and *xa13* (IRBB13) were resistant to all the five isolates, while *xa5* (IRBB5) could resist moderately to DX011, DX127 and DX015 isolates but susceptible to DX020 and DX133 isolate. Through marker analysis two resistant forms of *xa5* alleles were identified from *O. sativa* cultivar which may be useful for developing BB resistant lines. This study suggests that the wild species of *Oryza* possess potential genes which may be targeted for developing BB resistant varieties. Other than the currently known BB resistant *R* genes, many more are yet to identify from the unexplored wild genetic resource. Phenotypic screening suggests the presence of *R* gene(s) or QTL in nature which confers stronger resistance than the known genes.

SUMMARY AND CONCLUSION

Polymorphisms of four BB resistance *R* genes *Xa21, Xa26, Xa27* and *xa5* were determined as quantified by average pair wise difference across all sites denoted by 'π'. As a reference sequence, we have considered the previously cloned and characterized *Xa21, Xa27, Xa26* and *xa5* genes for analysis. The average nucleotide diversities in these genes were elevated than the reported *R* genes of tomato, Arabidopsis and blast resistance genes of rice. Along the entire gene sequence, it was commonly observed that intron and non coding region exhibited higher diversity for the entire four locus. The allelic diversity was higher in *O. nivara* than *O. sativa*. In natural habitat there are more chances of intermixing genetic components in wild species by out crossing which increases the intensity of allelic diversity. Transition bias > 1 was observed in all the genes which suggest that transition pattern of substitution is favored over transversion during the course of evolution. It's difficult to draw a concrete conclusion regarding the nature of selection at the four loci due to random sampling with diverse species. It was observed that in nature the presence of favorable forms of *R* gene is limited to geographic location, related species and cultivars. *Xa26* alleles had been identified from China and recessive *xa5* alleles from rice cultivars of Indian Subcontinent. It would be more beneficial to target species specific and cultivar specific individuals of a particular location to identify novel and favorable alleles of these genes. Though insignificant negative Tajima's D value indicate the absence of selection constraint on these genes, the value of K_a/K_s less than 1 suggest the operation of purifying selection. High level of amino acid polymorphism in XA5 and XA27 is indicative of balanced polymorphism, where both resistant and susceptible forms of alleles are prevalent in nature. However, the primary confirmation was not changed in XA5 and XA27 due to amino acid substitution; elevated non-synonymous substitution rates in these domains represent relaxed selective constraint and identification of single resistant allele for *Xa21, Xa26* and *Xa27* suggest negative frequency dependent selection. Species specific analysis with larger sample size would yield more significant result of this selection study.

During *Xoo* infection in wild rice *O. nivara* 81832 and *O. longistaminata* OL1, out of 13 genes analyzed, only seven genes were differentially expressed. These genes were constitutively expressed in both resistant and susceptible plants. *OsBAK1, RPLD* isoforms, *EI10P8* and *EI1K8* were found to induce and may be targeted for developing

BB resistant lines. The remaining genes were found unexpressed at the selected time points which may be due to early response of those genes during infection process. These genes may also be playing role in making the plants resistant against BB, however, we need further experiment like breeding with elite susceptible variety in order to identify the major resistance factor present in these wild rice.

Proteins identified by MALDI TOF- TOF analysis were found to be involved in defense and stress, energy metabolism, photosynthesis and regulatory proteins during BB stress condition. In the proteome study due to interference by polysaccharides and other metabolites, the total number of spots or protein separated on the gel were comparatively less. The proteome analysis was performed only in *Xoo* infected and water infiltrated resistant wild rice and only the up-regulated proteins were identified. Some of the protein identified in this study might play important roles in plant disease resistance, although further studies are required to verify the detailed localization and regulation of these proteins.

Some of the future prospects of the above study are:

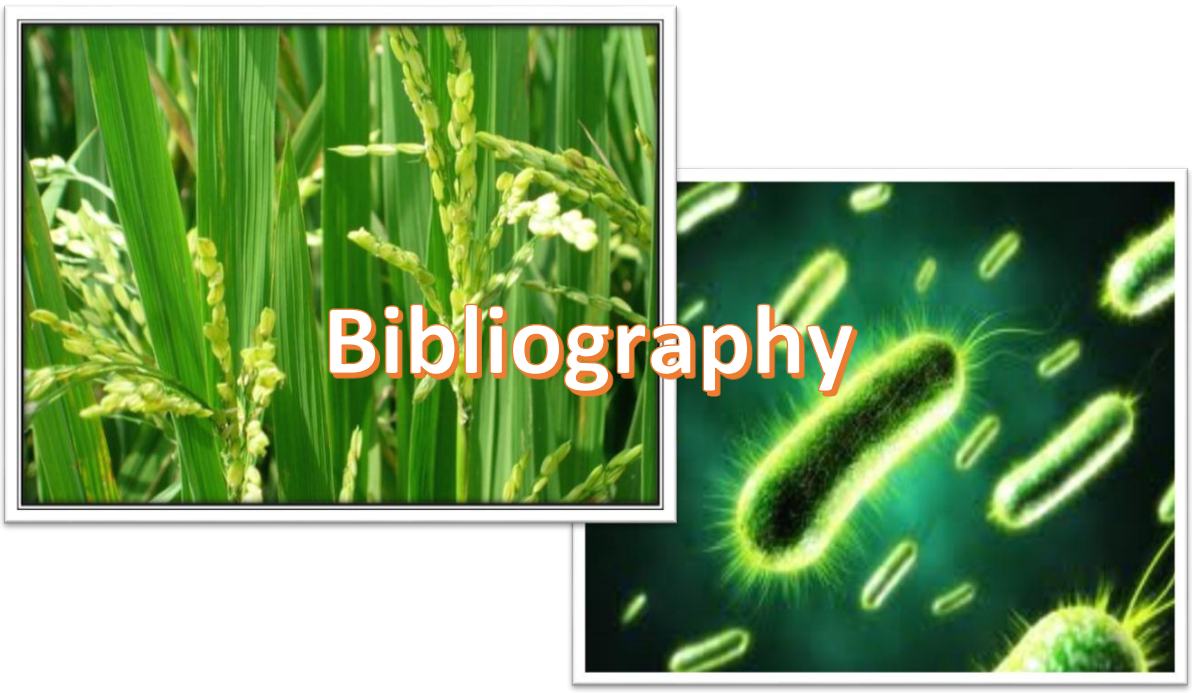
1. Conducting the diversity analysis with more number of accessions of same species from different geographical locations or population level study of *R* gene locus may yield accurate measurement of selection.

2. Functional study of allelic variants of *Xa27* by fusing with a functional $UPT_{Avrxa27}$ using promoter engineering will be of great interest.

3. Increasing number of targeted *R* genes is required for allele mining in order to dig out the responsible resistance genes in many of the wild rice.

4. Studying the molecular diversity of *avr* gene of the pathogens will be an interesting feature and will give a light on the co-evolutionary process of this *R-avr* gene interaction.

5. More intense analysis on transcript and translational level study in wild resistant rice will help in identification of novel resistance genes for bacterial blight.


Bibliography

Aboa ME and Fadhilab HA (2003) Epidemiology and management of rice viruses and virus diseases. *Plant virology in sub-Saharan Africa,*112-127.

Adhikari TB, Mew TW and Teng PS (1994) Progress of bacterial blight on rice cultivars carrying different *Xa*-genes for resistance in the field. *Plant Disease*, 78:73–77.

Adhikari TB, Mew TW, Leach JE (1999) Genotypic and pathotypic diversity in *Xanthomonas oryzae* pv. *oryzae* in Nepal. *Phytopathology*, 89:687–694.

Aggarwal RK, Brar DS, Nandi S, Huang N and Khush GS (1999) Phylogenetic relationships among Oryza species revealed by AFLP markers. *Theoretical Applied Genetics*, 98:1320-1328.

Aguilar MD, Grau M and Contreras JM (1997) Effect of pre-seeding nitrogen fertilization on rice yield components under water high salinity conditions in Southern Spain. In: Paper presented at the international symposium on rice quality. Nottinghan, UK, 24–27 November 1997.

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215:403–410.

Alvarez ME (2000) Salicylic acid in the machinery of hypersensitive cell death and disease resistance. *Plant Molecular Biology*, 44:429-442.

Andersson CR, Foley RC, Oñate-Sánchez L and Singh KB (2000) Transcriptional Control of Plant Defence Gene Expression. *Australian Biochemist*, 31:8-11.

Audic S and Claverie JM (1997) The significance of digital gene expression profiles. *Genome Research,* 7:986-95.

Bakker EG, Traw MB, Toomajian C, Kreitman M and Bergelson J (2008) Low levels of polymorphism in genes that control the activation of defence response in *Arabidopsis thaliana*. *Genetics*, 178:2031-2043.

Barleban L, Panjikar S, Ruppert M, Koepke J and Stöckigt J (2007) Molecular architecture of strictosidine glucosidase: The gateway to the biosynthesis of the monoterpenoid indole alkaloid family. *Plant Cell*, 19:2886–2897.

Bergelson J, Kreitman M, Stahl EA and Tian D (2001) Evolutionary dynamics of plant *R* genes. *Science*, 292:2281-2285.

Bernsel A, Viklund H, Hennerdal A and Elofsson A (2009) TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Research*, 37: (Webserver issue) W465-8

Bhattramakki D, Dolan M, Hanafey M et al. (2002) Insertion-deletion polymorphisms in 3′ regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Molecular Biology*, 48:539–547.

Bhullar NK, Street K, Mackay M, Yahiaoui N and Keller B (2009) Unlocking wheat genetic resources for the molecular identification of previously undescribed functional alleles at the *Pm3* resistance locus. *Proceedings of National Academy of Scince* USA, 106:9519–9524.

Bimolata W, Kumar A, Sundaram RM, et al. (2013) Analysis of nucleotide diversity among alleles of the major bacterial blight resistance gene *Xa27* in cultivars of rice (*Oryza sativa*) and its wild relatives. Planta 238:293-305.

Bin GS, Pin ZD and Hua LX (2010) Identification and mapping of a novel bacterial blight resistance gene *Xa35* (t) originated from *Oryza minuta*. *SciAgr Sinica*, 43:2611-2618.

Blair MW, Garris AJ, Iyer AS, Chapman B, Kresovich S and McCouch SR (2003) High resolution genetic mapping and candidate gene identification at the xa5 locus for bacterial blight resistance in rice (*Oryza sativa* L.), *Theoretical and Applied Genetics*, 107:62–73.

Blouin MS, Yowell CA, Courtney CH and Dame JB (1998) Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. *Molecular Biology and Evolution*, 15:1719–1727.

Bogdanove AJ, Schornack S and Lahaye T (2010) TAL effectors: finding plant genes for disease and defense. *Current Opinion in Plant Biology*, 13:394-401.

Bolwell GP and Wojtaszek P (1997) Mechanism for the generation of reactive oxygen species in plant defense - a broad perspective. *Physiology Molecular Plant Pathology*, 51:347–366.

Both M, Eckert SE, Csukai M, Mueller E, Dimopoulos G and Spanu PD (2005) Transcript profiles of *Blumeria graminis* development during infection reveal a cluster of genes that are potential virulence determinants. *Mol. Plant Microbe Interact*, 18**:**125–133.

Boyatee CD and Lee FN (19790 Reduction in rice yield caused by sheath blight. Arkans for Res, 28:5.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Annals of Biochemistry*, 72:248-254.

Brar DS and Khush GS (1997): Alien introgression in rice. *Plant Molecular Biology*, 35:35-47.

Buckler ESI, Thornsberry JM and Kresovich S (2001) Molecular diversity, structure and domestication of grasses. *Genetic Research*, 77:213-218.

Buckler ES and Thornsberry JM (2002) Plant molecular diversity and applications to genomics. *Current Opinion Plant Biology*, 5:107-111.

Bustamam MM, Yunus HR, Hifni M, Baroidan EY, Ardales and Nelson RJ (1996) Population structure of the rice bacterial pathogen, *Xanthomonas oryzae* pv. *oryzae* in Java and Bali,

Indonesia. In: Pages 107, Proc. Of the 3rd Asia-Pacific Conf. on Agricultural Biotechnology: Issues and Choices, Nov 10-15, Prachuapkhirikhan, Thailand.

Cao Y, Duan L, Li H, Sun X, Zhao Y, Xu C, Li X and Wang S (2007) Functional analysis of Xa3/Xa26 family members in rice resistance to *Xanthomonas oryzae* pv. *oryzae. Theoretical Applied Genetics*, 115:887–895.

Caicedo AL, Schaal BA and Kunkel BN (1999) Diversity and molecular evolution of the *RPS2* resistance gene in *Arabidopsis thaliana*. *PNAS*, 96:302–306.

Catoni M, Miozzi L, Fiorilli V, Lanfranco L and Accotto GP (2009) Comparative analysis of expression profiles in shoots and roots of tomato systemically infected by Tomato spotted wilt virus reveals organ-specific transcriptional responses. *Molecular Plant Microbe Interaction*, 22:1504–1513.

Century KS, Lagman RA, Adkisson M, Morlan J, Tobias R, Schwartz K, Smith A, Love J, Ronald PC and Whalen MC (1999) Short communication: developmental control of *Xa21*-mediated disease resistance in rice. *Plant Journal*, 20:231-236.

Chahal KKS, Sokhi SS and Rattan GS (2003) Investigation on sheath blight of rice in Punjab. *Indian Phytopathology*, 56:22-26.

Cheema KK, Grewal NK, Vikal Y, Sharma R, Lore JS, Das A, Bhatia D, Mahajan R, Gupta V, Bharaj TS and Singh K (2008) A novel bacterial blight resistance gene from *Oryza nivara* mapped to 38 kb region on chromosome 4 and transferred to *Oryza sativa* L. *Genet Res*, 90:1-11.

Chen H, Wang S, Xing Y, Xu C, Hayes PM and Zhang Q (2003) Comparative analyses of genomic locations and race specificities of loci for quantitative resistance to *Pyricularia grisea* in rice and barley. *Proceedings in National Academy of Science*, USA 100:2544–2549.

Chen S, Huang ZH, Zeng LX, Yang JY, Liu QG, et al. (2008) High resolution mapping and gene prediction of *Xanthomonas oryzae* pv. *oryzae* resistance gene *Xa7*. *Mol Breeding*, 22:433-441.

Chen S, Liu X, Zeng L, Ouyang D, Yang J and Zhu X (2011) Genetic analysis and molecular mapping of a novel recessive gene *xa34*(t) for resistance against *Xanthomonas oryzae* pv. *oryzae. Theoretical Applied Genetics*, 122:1331-1338.

Chen X and Ronald PC (2011) Innate immunity in rice. *Trends in Plant Science*, 16:451-458.

Chen X, Fu S, Zhang P, Gu Z, Liu J, Qian Q and Ma B (2013) Proteomic analysis of a disease resistance-enhanced lesion mimic mutant spotted leaf 5 in rice. Rice 6:1.

Chinchilla D, Zipfel C, Robatzek S, Kemmerling B, Nurnberger T, Jones JD, FelixG and Boller T (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature*, 448:497–500.

Choi Y, Sims GE, Murphy S, Miller JR and Chan AP (2012) Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE* , 7: e46688.

Chu Z, Ouyang Y, Zhang J, Yang H and Wang S (2004) Genome-wide analysis of defense responsive genes in bacterial blight resistance of rice mediated by a recessive R gene, *xa13*. *Mol Gen Genomics,* 271:111-120.

Chu Z, Fu B, Yang H, Xu C, Li Z, Sanchez A, Park YJ, Bennetzen JL, Zhang Q and Wang S (2006) Targeting *xa13*, a recessive gene for bacterial blight resistance in rice. *Theoretical and Applied Genetics*, 112:455–461.

Coram TE, Wang M and Chen XM (2008) Transcriptome analysis of the wheat–Puccinia striiformis f. sp. tritici interaction. *Molecular Plant Pathology*, 9:157-169.

Cummings MP and Clegg MT (1998) Nucleotide sequence diversity at the *alcohol dehydrogenase* 1 locus in wild barley (*Hordeum vulgare* ssp. spontaneum): an evaluation of the background selection hypothesis. *Proceedings of National Academy of Science*, 95:5637-5642.

Czjzek M, Cicek M, Zamboni V, Burmeister WP, Bevan DR, Henrissat B and Esen A (2001) Crystal structure of a monocotyledon (maize ZMGlu1) betaglucosidase and a model of its complex with p-nitrophenyl beta-D-thioglucoside. *Biochem. Journal*, 354:37–46.

Dally AM and Second G (1990) Chloroplast DNA diversity in wild and cultivated species of rice (genus Oryza, section Oryza): cladistic-mutation and genetic-distance analysis. *Theoretical Applied Genetics,* 80:209-222.

Dangl JL and Jones JDG (2001) Plant pathogens and integrated defense responses to infection. *Nature*, 411:826-833.

Datta SK (2002) Transgenic plants for the management of plant diseases: Rice a case study. In S.S. Gnanamanickam (Ed.), *Biological control of crop diseases* (pp. 31–62). New York: Dekker.

Datta SK (2004) Rice biotechnology: A need for developing countries. *AgBioForum*, 7:31–35.

De Vleesschauwer, D. et al. (2010) Abscisic acid-induced resistance against the brown spot pathogen Cochliobolus miyabeanus in rice involves MAP kinase-mediated repression of ethylene signaling. *Plant Physiology*, 152:2036–2052.

Diatchenko L (1996) "Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries". *Proceedings of the National Academy of Sciences*, 93:6025–6030.

Dong X (2001) Genetic dissection of systemic acquired resistance. *Current Opinion in Plant Biology*, 4:309–314.

Doebley J and Lukens L (1998) Transcriptional regulators and the evolution of plant form. *Plant Cell,* 10:1075–1082.

Doyle JJ and Doyle JL (1989) Isolation of plant DNA from fresh tissue. *Focus*, 12:13-15.

DRR (2002) Directorate of Rice Research Progress Report, 2001, Vol. 2. Entomology and Pathology, All India Coordinated Rice Improvement Programme (ICAR), Directorate of Rice Research, Rajendranagar, Hyderabad-500 030, A.P., India.

Endo T, Ikeo K and Gojobori T (1996) Large-scale search for genes on which positive selection may operate. *Molecular Biology and Evolution,* 13:685-90.

Esteras C, Formisano G, Roig C, Dı´az A, Blanca J, Garcia-Mas J et al. (2013) SNP genotyping in melons: genetic variation, population structure, and linkage disequilibrium. *Theoretical Applied Genetics*, 126:1285–1303.

Eyre-Walker A, Gaut RL, Hilton H, Feldman DL and Gaut BS (1998) Investigation of the bottleneck leading to the domestication of maize. *Proceedings of National Academy of Sciences* USA, 95:4441-4446.

Fang DD, Xiao J, Canci PC and Cantrell RG (2010) A new SNP haplotype associated with blue disease resistance gene in cotton (*Gossypium hirsutum* L.) *Theoretical Applied Genetics*, 120:943–953.

Feng Q, Huang T, Zhao Q, Zhu J, Lin Z and Han B (2009) Analysis of collinear regions of Oryza AA and CC genomes. *Journal of Genetics and Genomics*, 36:667-77.

Flor HH (1942) Inheritance of pathogenicity in Melampsora lini. *Phytopathology*, 32:653-669.

Flor HH (1971) Current status of the gene-for-gene concept." *Annu Rev Phytopathol*, 9:275–296.

Fu YX and Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, 133:693-709.

Geddes J, Eudes F, Laroche A and Selinger LB (2008). Differential expression of proteins in response to the interaction between the pathogen *Fusarium graminearum* and its host, Hordeumvulgare. *Proteomics,* 8:545–554.

Gjetting T, Carver TL, Skot L and Lyngkjaer MF (2004) Differential gene expression in individual papilla resistant and powdery mildew infected barley epidermal cells. *Mol Plant Microbe Interact*, 17:729–738.

Glazebrook J (2001) Genes controlling expression of defense responses in Arabidopsis – 2001 status. *Current Opinion in Plant Biology*, 4:301–308.

Global Rice Production, FAO Statistical Database on crops, 2011, http://faostat.fao.org.

Gnanamanickam S, Brindha Priyadarisini V, Narayanan N, Vasudevan P and Kavitha S (1999) An overview of bacterial blight disease of rice and strategies for its management. *Current Science*, 77:1435–1443.

Gnanamanickam SS (2009) Biological Control of Rice Diseases, Progress in Biological Control 8, DOI 10.1007/978-90-481-2465-7 5,

Gnanamanickam S, Vasudevan P and Velusamy P (2004) Biological control of bacterial blight of rice in India with rice-associated bacteria: Insights into mechanisms. Proceedings of the 1st International Conference on Bacterial Blight of Rice. Tsukuba, Japan.

Goel RK, Kaur L and Saini RG (1998) Effectiveness of different *Xa* genes against *Xanthomonas oryzae* pv. *oryzae* populations causing bacterial blight of rice in Punjab (India). *Rice Genetic Newsletter*, 15:131–132.

Gu K, Tian D, Yang F, Wu L, Sreekala C, Wang D, Wang GL and Yin Z (2004) High-resolution genetic mapping of *Xa27*(t), a new bacterial blight resistance gene in rice, *Oryza sativa* L. *Theoretical Applied Genetics*, 108:800-807.

Gu K, Yang B, Tian D, Wang D, Wu L, Sreekala C, Yang F, Chu Z, Wang GL, White FF and Yin Z (2005) *R*-gene expression induced by a type-III effector triggers disease resistance in rice. *Nature*, 435:1122-1125.

Gu K,Tian D and Qiu C,Yin Z (2009) Transcription activator-like type III effector AvrXa27 depends on OsTFIIAg5 for the activation of *Xa27* transcription in rice that triggers disease resistance to *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Pathology*, 10:829-835

Gyetvai G, Sønderkaer M, Göbel U, Basekow R, Ballvora A, Imhoff M, et al. (2012) The transcriptome of compatible and incompatible interactions of potato (*Solanum tuberosum*) with Phytophthora infestans revealed by DeepSAGE analysis. *PLoS ONE*, 7:e31526.10.1371/journal.pone.0031526.

Haldane JBS (1949) Disease and evolution. *Ric. Sci. Suppl*. 19:68–76 [often requoted, *e.g.*, in Dronamraju, K. (Editor), 1990, *Selected Genetic Papers of J.B.S. Haldane*, Garland Publishing, New York/London]

Haseneyer G, Ravel C, Dardevet M, Balfourier F, Sourdille P, Charmet G, Brunel D, Sauer S, Geiger HH, Graner A and Stracke S (2008) High level of conservation between genes coding for the GAMYB transcription factor in barley (*Hordeum vulgare* L.) and bread wheat (*Triticum aestivum* L.) collections. *Theoretical Applied Genetics*, 117:321–331.

Heath MC (2000) Nonhost resistance and nonspecific plant defenses. *Current Opinion in Plant Biology*, 3:315-319.

Heese A, Hann DR, Gimenez-Ibanez S, Jones AM, He K, Li J, Schroeder JI, Peck SC and Rathjen JP (2007) The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *Proceedings in National Academy of Sciences* USA, 104:12217–12222.

Hill WG, Weir BS (1988) Variances and covariances of squared linkage disequilibria in finite populations. Theor Pop Biol 33:54–78.

Hittalmani S, Kahani F, Dhanagond SM and Rao AM (2013) DNA marker characterization for allele mining of blast and bacterial leaf blight resistant genes and evaluation for grain yield. *African Journal of Biotechnology*, 12:2331-2340.

Hsieh SPY and Buddenhagen IW (1974) Suppressing effects of *Erwinia herbicola* on infection by *Xanthomonas oryzae* and on symptom development in rice. *Phytopathology*, 64:1182-1185.

Huang H and Kochert G (1994) Comparative RFLP mapping of an allotetraploid wild rice species (*Oryza latifolia*) and cultivated rice (*O. sativa*). *Plant Molecular Biology*, 25:633-648.

Huang CL, Hwang SY, Chiang YC and Lin TP (2008) Molecular Evolution of the *Pi-ta* Gene Resistant to Rice Blast in Wild Rice (*Oryza rufipogon*). *Genetics*, 179:1527-1538.

Huang N, Angeles ER, Domingo J, Magpantay G, Singh S, Zhang G, Kumaravadivel N, Bennett J and Khush GS (1997) Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and PCR. *Theoretical Applied Genetics*, 95:313-320.

Hui XZ, Fei H, Li-fen G et al. (2010) Application of Functional Markers to Identify Genes for Bacterial Blight Resistance in *Oryza rufipogon*. *Rice Science*, 17:73-76.

Hummel AW, Doyle EL and Bogdanove AJ (2012) Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytology*, 195:883-93.

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature*, 436: 793-800.

International Rice Research Institute (IRRI) (1983) Field problems of tropical rice. Manila (Philippines): IRRI. 172 p.

IRRI (2003) *Rice doctor*, Philippines: The International Rice Research Institute.

Iyer AS and McCouch SR (2004) The rice bacterial blight resistance gene xa5 encodes a novel form of disease resistance. *Molecular Plant Microbe Interaction*, 17:1348–1354.

Iyer-Pascuzzi AS, Sweeney MT, Sarla N and McCouch SR (2007) Use of naturally occurring alleles for crop improvement. In: Upadhyaya NM (ed), Rice functional genomics- Challenges, progress and prospects. Springer, pp 113-143.

Iyer-Pascuzzi AS and McCouch SR (2007) Functional markers for *xa5*-mediated resistance in rice. *Molecular Breeding*, 19:291-296.

Iyer-Pascuzzi AS, Jiang H, Huang L and McCouch SR (2008) Genetic and functional characterization of the rice bacterial blight disease resistance gene *xa5*. *Phytopathology*, 98:289-295.

Jha S, Singh AK, Krishna M, Meena MS and Singh RK P (2012) Constraints of Rainfed Rice Production in Eastern India: An Overview. doi.org/10.2139/ssrn.2061953.

Ji GH, Zhang SG and Qian J (2003) Preliminary analysis on races of *Xanthomonas oryzae* pv. *oryzae* in Yunnan Province. *Plant Prot*., 29:19-21.

Ji GH, Wei LF, He YQ, Wu YP and Bai XH (2008) Biological control of rice bacterial blight by *Lysobacter antibioticus* strain 13–1. *Biological Control*, 45:288–296.

Jiang GH, Xia ZH, Zhou YL, Wan J, Li DY, Chen RS, Zhai WX and Zhu LH (2006) Testifying the rice bacterial blight resistance gene *xa5* by genetic complementation and further analyzing *xa5* (*Xa5)* in comparison with its homolog *TFIIAc1*. *Molecular Genetics and Genomics,* 275:354–366.

Jiang CJ et al. (2010) Abscisic acid interacts antagonistically with salicylic acid signaling pathway in rice-*Magnaporthe grisea* interaction. *Molecular Plant Microbe Interaction,* 23:791–798.

Jones JD and Dangl JL (2006) The plant immune system. *Nature*, 444:323-329.

Jones E, Chu WC, Ayele M, Ho J, Ed Bruggeman, Yourstone K, Rafalski A et al. (2009) Development of single nucleotide polymorphism (SNP) markers for use in commercial maize (*Zea mays* L.) germplasm. *Molecular Breeding,* 24:165–176.

Joseph M, Gopalakrishnan S, Sharma RK, Singh AK, Singh VP, Singh NK and Mohapatra T (2004) Combining bacterial blight resistance and Basmati quality characteristics by phenotypic and molecular marker assisted selection in rice. *Molecular Breeding*. 13:377–387.

Kalendar R, Lee D and Schulman AH (2009) FastPCR Software for PCR Primer and Probe Design and Repeat Search. *Genes, Genomes and Genomics*, 3:1-14.

Kauffman HE, Reddy APK, Hsieh SPY and Merca SD (1973) An improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. *Plant Disease Report*, 57:537-541.

Kaur K, Street K, Mackay M, Yahiaoui N and Keller B (2008) Allele mining and sequence diversity at the wheat powdery mildew resistance locus *Pm3*. In: Appels R, Eastwood R, Lagudah E, Langridge P, Mackay M, McIntyre L, editors. Proceedings of the 11th International Wheat Genetics Symposium. Sydney, Australia: Sydney University Press.

Kawabe A, Innan H, Terauchi R and Miyashita NT (1997) Nucleotide polymorphism in the acidic chitinase locus (ChiA) region of the wild plant *Arabidopsis thaliana*. *Molecular Biology and Evolution,* 14:1303-1315.

Khush GS, Bacalangco E and Ogawa T (1990) A new gene for resistance to bacterial blight from *O. longistaminata*. *Rice Genetic Newsletter,* 7:121-122

Kim ST et al. (2008) The rice pathogen-related protein 10 (JIOsPR10) is induced by abiotic and biotic stresses and exhibits ribonuclease activity. *Plant Cell Report*, 27:593–603.

Kim KH, Kang YJ, and Lee SH (2011) RNA-Seq Analysis of a Soybean Near-Isogenic Line Carrying Bacterial Leaf Pustule-Resistant and Susceptible Alleles. *DNA Research*, 18:483–497.

Kim JS, Ahn SN, Kim CK and Shim CK (2010) Screening of Rice Blast Resistance Genes from Aromatic Rice Germplasms with SNP Markers. *Plant Pathology Journal*, 26:70-79.

Kottapalli KR, Rakwal R, Satoh K, Shibato J, Kottapalli P, Iwahashi H, Kikuch S (2007) Transcriptional profiling of *indica* rice cultivar IET8585 (Ajaya) infected with bacterial leaf blight pathogen *Xanthomonas oryzae* pv *oryzae*. *Plant Physiology and Biochemistry*, 45:834-850.

Kou Y and Wang S (2010) Broad-spectrum and durability: understanding of quantitative disease resistance. *Current Opinion in Plant Biology*, 13:181-185.

Kou Y and Wang S (2013) Bacterial blight resistance in rice. In Genomics Applications in Plant Breeding. Edited by Varshney R, Tuberosa R. Wiley-Blackwell Press; in press.

Kover PX and Schaal BA (2002) Genetic variation for disease resistance and tolerance among *Arabidopsis thaliana* accessions. *Proceedings in National Academy of Sciences* USA, 99:11270–11274.

Kowalczuk M, Mackiewicz P, Mackiewicz D, Nowicka A, Dudkiewicz M, Didek MR, Cebrat S (2001) DNA asymmetry and the replicational mutational pressure. *Journal of Applied Genetics*, 42:553–577.

Kuljit K, Cheema, Navjit K, Grewal, Vikal Y, Sharma R et al (2008) A novel bacterial blight resistance gene from *Oryza nivara* mapped to 38 kb region on chromosome 4L and transferred to *Oryza sativa* L. *Genetic Research Camb,*. 90:397–407.

Kumar P, Henikoff S and Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocol*, 4:1073-1081.

Kumar PN, Sujatha K, Laha GS, Rao KS, Mishra B, Viraktamath BC, Hari Y, Reddy CS, Balachandran SM, Ram T, Sheshu MM, Rani NS, Neeraja CN, Reddy GA, Shaik H, Sundaram RM (2012) Identification and fine-mapping of *Xa33*, a novel gene for resistance to *Xanthomonas oryzae* pv. *oryzae*. *Phytopathology*, 102:222-228.

Kumari A, Das A, Devanna BN, Thakur S, Singh P K, Upreti, H C, Singh N K, Sharma T R (2013) Mining of Rice Blast Resistance Gene *Pi54* shows effect of single nucleotide polymorphism on phenotypic expression of alleles. *Europan Journal of Plant Pathology***.** Doi 10.1007/s10658_013_0216-5.

Latha R, Rubia L, Bennett J, and Swaminathan MS (2004) Allele Mining for Stress Tolerance Genes in *Oryza* Species and Related Germplasm. *Molecular Biotechnology*, 27:101–108.

Lavanya B, Priyadarisini VB, Leenakmari S, Gnanamanickam SS and Levy M (1998) Lineage-exclusion resistance breeding: pyramiding of blast and bacterial blight resistance genes for rice disease management in India. In: Page 21 Proc. 7th Natl. Rice Biotechnology Network Meeting, Oct 25-29, 1998, UAS-NCBS, Bangalore, India.

Lee FN and Rush MC (1983) Rice sheath blight: a major rice disease. *Plant Disease*, 67:29-32.

Lee SW, Choi SH, Hon SS, Lee DG and Lee BY (1999) Distribution of *Xanthomonas oryzae* pv. o*ryzae* strain virulent to *Xa21* in Korea. *Phytopathology*, 89:928–933.

Lee SW, Han SW, Bartley LE and Ronald PC (2006) Unique characteristics of *Xanthomonas oryzae* pv. *oryzae* AvrXa21 and implications for plant innate immunity. *Proceedings in National Academy of Sciences* USA, 103:18395–18400.

Lee SW, Han SW, Sririyanum M, Park CJ, Seo YS and Ronald PC (2009) A type I-secreted, sulfated peptide triggers XA21-mediated innate immunity. *Science*, 326:850-853.

Lee S, Jia Y, Jia M, Gealy DR, Olsen KM et al. (2011) Molecular Evolution of the Rice Blast resistance Gene *Pi-ta* in Invasive Weedy Rice in the USA. *PLoS ONE*, 6: e26260. doi:10.1371/journal.pone.0026260.

Leonforte A, Sudheesh S, Cogan NO, Salisbury PA, Nicolas ME, Materne M, Forster JW and Kaur S (2013) SNP marker discovery, linkage map construction and identification of QTLs for enhanced salinity tolerance in field pea (*Pisum sativum* L.). *BMC Plant Biology*, 13:161.

Li Q, Chen F, Sun L, Zhang Z, Yang Y and He Z (2006) Expression profiling of rice genes in early defense responses to blast and bacterial blight pathogens using cDNA microarray. *Physiological and Molecular Plant Pathology*, 68:51–60.

Li D, Wang L, Wang M, Xu YY, Luo W, Liu YJ, Xu ZH, Li J, Chong K (2009). Engineering *OsBAK1* gene as a molecular tool to improve rice architecture for high yield. *Plant Biotechnology Journal*, 7:791–806.

Li H, Li X, Xiao J, Wing RA and Wang S (2012) Ortholog alleles at *Xa3/Xa26* locus confers conserved race-specific resistance against *Xanthomonas oryzae* in rice. *Molecular Plant*, 5:281-290.

Librado P and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25:1451-1452.

Lin JZ, Brown AHD and Clegg MT (2001) Heterogeneous geographic patterns of nucleotide sequence diversity between two alcohol dehydrogenase genes in wild barley (*Hordeum vulgare* subspecies spontaneum). *Proceedings of National Academy of Sciences* USA, 98:531-536.

Lodha TD and Basak J (2012) Plant-pathogen interactions: what microarray tells about it? *Molecular Biotechnology*, 50:87-97.

Lv Q , X Xu, J Shang, G Jiang, Z Pang, Z Zhou, J Wang et al. (2013) Functional Analysis of *Pid3-A4*, an Ortholog of Rice Blast Resistance Gene *Pid3* Revealed by Allele Mining in Common Wild Rice. *Phytopathology*, 103:594-599.

Mahmood T, Jan A, Kakishima M and Komatsu M (2006) Proteomic analysis of bacterial-blight defense - responsive proteins in rice leaf blades. *Proteomics*, 6:6053–6065.

Maleck K, Levine A, Eulgem T, Morgan A, Schmid J, Lawton KA, Dangl JL, Dietrich RA (2000) The transcriptome of Arabidopsis during systemic acquired resistance. *Nature Genetics*, 26:403-410.

Mammadov J, Aggarwal R, Buyyarapu R and Kumpatla S (2012) SNP Markers and Their Impact on Plant Breeding. *International Journal of Plant Genomics*, http://dx.doi.org/10.1155/2012/728398.

Mardis ER (2008) Next-Generation DNA Sequencing Methods. *Annual Review of Genomics and Human Genetics*, 9:387-402.

Marella LS, Vera CM, George LC, Bernardo MA, Nelson RJ, Lenug H, Reddy JN and Sridhar R (2001). Idenification of resistance genes effective against rice bacterial bight pathogen in Eastern India. *Plant Disease*, 85:506-512.

Mauricio R, Stahl EA, Korves T, Tian D, Kreitman M et al. (2003) Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana*. *Genetics*, 163:735–746.

Maytalman D, Mert Z, Baykal AT, Inan C, Günel A, Hasançebi S (2013) Proteomic analysis of early responsive resistance proteins of wheat (*Triticum aestivum*) to yellow rust (*Puccinia striiformis* f. sp. *tritici*) using ProteomeLab PF2D**.** *Plant Omics*, 6:24-35.

May RM and Anderson RM (1983) Parasite-host coevolution, in Coevolution, edited by D. Futuyama and M. Slatkin, Sinauer Associates, Sunderland, MA. pp. 186–206.

McDonald DJ (1994) Temperate rice technology for the 21st century-the New South Wales example. In: Proceedings of the first temperate rice conference, Yanco, Australia, pp 1–12.

McGee JD, Hamer JE and Hodges TK (2001) Characterization of a PR-10 pathogenesis-related gene family induced in rice during infection with *Magnaporthe grisea*. *Molecular Plant Microbe Interaction*, 14:877–886.

Mei C et al. (2006) Inducible overexpression of a rice allene oxide synthase gene increases the endogenous jasmonic acid level, PR gene expression, and host resistance to fungal infection. *Molecular Plant Microbe Interaction*, 19:1127–1137.

Mew TW and Rosales AM (1986) Bacterization of rice plants for control of sheath blight caused by *Rhizoctonia solani*. *Phytopathology*, 76:1260-1264.

Mew TW (1987) Current status and future prospects of research on bacterial blight of rice. *Annual Review of Phytopathology*, 25, 359–382.

Mew TW (1989) An overview of the world bacterial bight situation. In: Bacterial Blight of Rice, pp. 7–12. International Rice Research Institute, Los Banos, Laguna, Philippines.

Mew TW, Alvarez AM, Leach JE and Swings J (1993) Focus on bacterial blight of rice. *Plant Disease*, 77:5–12.

Meyers BC, Shen KA, Rohani R, Gaut BS and Michelmore RW *(*1998*)* Receptor-like Genes in the Major Resistance Locus of Lettuce Are Subject to Divergent Selection. *Plant Cell*, 10*:*1833-1846.

Mitsuhara I et al. (2008) Characteristic expression of twelve rice PR1 family genes in response to pathogen infection, wounding and defense-related signal compounds. *Molecular Genetics and Genomics*, 279:415–427.

Miruta H (1956) On the relation between yield and inoculation times of sheath blight, Corticium sasakii in the earlier planted paddy rice. Association of Plant Protection Kyushu, 2:100-102.

Mizukami, T. and Wakimoto, S. (1969) Epidemiology and control of bacterial leaf blight of rice. *Annual Review of Phytopathology*, 7:51–72.

Morant AV, Jorgensen K, Jorgensen C, Paquette SM, Sanchez-Perez R, Moller BL and Bak S (2008) beta-glucosidases as detonators of plant chemical defense. *Phytochemistry*, 69:1795–1813.

Morishima (2001) Molecular marker, genetic diversity and evolution. Book chapter Rice genetics IV G.S. Khush, D.S. Brar and B. Hardy, IRRI.

Morishima H, Sano Y and Oka HI (1992) Evolutionary studies in cultivated rice and its wild relatives. Oxford Surv. *Evolutionary Biology*, 8:135-184.

Mortazavi A, et al. (2008) Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature Methods*, 5:621-628.

Moriyama EN and Powell JR (1996) Intraspecific nuclear DNA variation in Drosophila. *Molecular Biology and Evolution*, 13:261–277.

Muralidharan K, Khush GS, Krishnaveni D, Reddy CS, Reddy JN, Sridhar R. 2004. Field performance of genotypes by molecular-assisted breeding to bacterial blight in India. *Rice Genetic Newsletter*, 21:29-31.

Natrajkumar P, Sujatha K, Laha GS, et al. (2012) Identification and fine-mapping of *Xa33*, a novel gene for resistance to *Xanthomonas oryzae* pv. *oryzae*. *Phytopathology*, 102:222-228.

Nei M and Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3:418-426.

Nei M and Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of National Academy of Sciences,* 76: 5296–5273.

Nguyen NV and Ferrero A (2006) Meeting the challenges of global rice production. *Paddy Water Environ*, 4: 1–9.

Nielsen KL, Høgh AL and Emmersen J (2006) DeepSAGE-digital transcriptomics with high sensitivity, simple experimental protocol and multiplexing of samples. *Nucleic Acids Research*, 34: e133.

Niemeyer HM (1988) Hydroxamic acids (4-hydroxy-1,4-Benzoxazin-3-Ones), defense chemicals in the gramineae. *Phytochemistry*, 27, 3349–3358.

Nyvall RF (1999) Field crop diseases. Iowa State University Press, USA. 1,021 p.

Ogawa T, Lin L, Tabien RE and Khush GS (1987) A new recessive gene for resistance to bacterial blight of rice. *Rice Genetic Newsletter,* 4:98–100.

Oka HI (1988) Origin of cultivated rice. Tokyo (Japan): Japanese Science Society Press. 254 p.

Ochiai H, Horino O, Miyajima K and Kaku H (2000) Genetic diversity of *Xanthomonas oryzae* pv. *oryzae* strains from Sri Lanka. *Phytopathology*, 90:415-421.

Olmstead RG, Reeves PA and Yen AC (1998) Patterns of sequence evolution and implications for parsimony analysis of chloroplast DNA. In: Soltis DE, Soltis PS, Doyle JJ (eds) Molecular systematics of plants II. Kluwer Academic, London, pp 164–187

Opina DS and Exconde BR (1971) Assessment of yield loss due to bacterial leaf streak of rice. *Philippine Phytopathologyl*, 7:35-39.

Ou SH (1985) Rice diseases. Commonwealth Mycological Institute, Kew, Surrey, United Kingdom.

Pakusch AE, Kneusel RE and Matern U (1989) *S*-adenosyl-l-methionine:*trans*-caffeoyl-coenzyme A 3-*O*-methyltransferase from elicitor-treated parsley cell suspension cultures. *Archieves of Biochemistry and Biophysics*, 271:488–494.

Parida SK, Mukerji M, Singh AK, Singh NK and Mohapatra T (2012) SNPs in stress responsive rice genes: validation, genotyping, functional relevance and population structure. *BMC Genomics*, 13:426.

Park CJ, Kim KJ, Shin R, Park JM, Shin YC and Paek KH (2004) Pathogenesis-related protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. *The Plant Journal*, 37:186–198.

Park HS, Ryu HY, Kim BH, Kim SY, Yoon IS and Nam KH (2011) A subset of *OsSERK* genes, including *OsBAK1*, affects normal growth and leaf development of rice. *Molecular Cells*, 32:561–569.

Petrov DA and Hartl DL (1999) Patterns of nucleotide substitution in Drosophila and mammalian genomes. *Proceedings of National Academy of Sciences* USA, 96:1475–1479.

Peng Y, Bartley LE, Chen X, Dardick C, Chern M, Raun R, Canlas PE and Ronald PC (2008) OsWRKY62 is a negative regulator of basal and *Xa21*- mediated defense against *Xanthomonas oryzae* pv. *oryzae* in rice. *Molecular Plant*, 1:446-458.

Petersen TN, Brunak S, Heijne GV and Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8:785-786.

Pingali PL and Rosegrant MW (1996) Confronting the environmental consequences of the Green Revolution. In: Proceedings of the 18th session of the international rice commission. Rome, Italy, pp. 59–69.

Poulton JE (1990) Cyanogenesis in plants. *Plant Physiology*, 94:401–405.

Quirino BF, Candido ES, Campos PF, Francoa OL and Krüger RH (2010) Proteomic approaches to study plant–pathogen interactions. *Phytochemistry*, 71:351–362.

Rajan CPD (1987) Estimation of yield losses due to sheath blight of rice. *Indian Phytopathology*, 40:174-177.

Ram SG, Thiruvengadam V and Vinod KK (2007) Genetic diversity among cultivars, landraces and wild relatives of rice as revealed by microsatellite markers. *Journal of Applied Genetics*, 48:337–345.

Ram Kumar G, Sakthivel K, Sundaram RM, Neeraja CN, Balachandran SM, Shobha Rani N, Viraktamath BC and Madhav MS (2010) Allele mining in crops: prospects and potentials. *Biotechnology Advances*, 28:451–461.

Ram T, Laha GS, Deen R, Ramos JM, VeraCruz CM and Brar DS (2011) *Oryza rufipogon*, a valuable source for resistance to bacterial blight of rice. *Plant Breeding*, 130:715–718.

Rangan L, Constantino S, Khush GS, Swaminathan MS and Bennett J (1999) The feasibility of PCR based allele mining for stress tolerance genes in rice and related germplasm. *Rice Genetic Newsletter*, 16:43–7.

Rajendrakumar P, Sujatha K, Rao KS, Natarajkumar P, Viraktamath BC, Balachandran SM, Biswal AK and Sundaram RM (2007) A protocol for isolation of DNA suitable for rapid seed and grain purity assessments in rice. *Rice Genetics Newsletter*, 23:92-95.

Reddy APK (1989) Bacterial blight: crop loss assessment and disease management. In Proceeding of the International Workshop on Bacterial Blight of Rice. International Rice Research Institute, Manila, the Philippines. p. 79–88.

Reimmann C and Dudler R (1993) Circadian rhythmicity in the expression of a novel light-regulated rice gene. *Plant Molecular Biology*, 22:165–170.

Romer P, Recht S and Lahaye T (2009) A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens. *Proceedings of National Academy of Sciences* USA, 106:20526-20531.

Römer P, Fecht S, Straub T, Elsasser J, Schornack S, Boch J, Wang S and Lahaye T (2010) Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist*, 187:1048-1057.

Ronald PC, Albano B, Tabien R, Abenes L, Wu K, McCouch S, Tanksley SD (1992) Genetic and physical analysis of the rice bacterial blight disease resistance locus, *Xa21*. *Mol Gen Genet*, 236:113-120.

Roux M, Schwessinger B, Albrecht C, Chinchilla D, Jones A, Holton N, Malinovsky FG, Tor M, de Vries S, and Zipfel C (2011) The *Arabidopsis* leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. *Plant Cell*, 23:2440–2455.

Rose LE, Bittner-Eddy PD, Langley CH, Holub EB, Michelmore RW and Beynon JL (2004) The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. *Genetics*, 166:1517-1527.

Rose LE, Michelmore RW and Langley CH (2007) Natural Variation in the *Pto* Disease Resistance Gene Within Species of Wild Tomato (Lycopersicon). II. Population Genetics of *Pto*. *Genetics*, 175:1307–1319.

Rozen S and Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386.

Sakthivel N and Mew TW (1991) Efficacy of bacteriocinogenic *Xanthomonas oryzae* pv. *oryzae* on the incidence of bacterial blight disease of rice (Oryza sativa, L). *Canadian Journal of Microbiology*, 37:764–768.

Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning: a laboratory manual. Second ed. NY, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press.

Santhi DP, Unnamalai N and Gnanamanickam SS (1987) Epiphytic association of Erwinia herbicola with rice leaves infected by *Xanthomonas campestris* pv. *oryzae* and its interaction with the pathogen. *Indian Phytopathology*, 40:327-332.

Scardaci SC, Webster RK, Greer CA, Hill JE, William JF, Mutters RG, Brandon DM, McKenzie KS and Oster JJ (1997) Rice blast: a new disease in California. Agronomy Fact Sheet Series. 1997–2. Department of Agronomy and Range Science, University of California, Davis.

Scheel D (1998) Resistance response physiology and signal transduction. Current Opinion in Plant Biology, 1:305–310.

Schena M, Shalon D, Davis RW and Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270:467-70.

Schenk PM, Kazan K, Manners JM, Anderson JP, Simpson RS, Wilson IW, Somerville SC and Maclean DJ (2003) Systemic gene expression in Arabidopsis during an incompatible interaction with *Alternaria brassicicola*. *Plant Physiology*, 132:999–1010.

Scheideler M, Schlaich NL, Fellenberg K, Beissbarth T, Hauser NC, Vingron M, Slusarenko AJ and Hoheisel JD (2002) Monitoring the switch from housekeeping to pathogen defense metabolism in *Arabidopsis thaliana* using cDNA arrays. *Journal of Biological Chemistry*, 277:10555–10561.

Second G (1991) Molecular markers in rice systematics and the evaluation of genetic resources. In: Bajaj YPS, editor. Biotechnology in agriculture and forestry. Vol 14. Rice. New York (USA): Springer. p 468-494.

Second G and Wang ZY (1992) Mitochondrial DNA RFLP in genus Oryza and cultivated rice. *Genetic Research and Crop Evolution*, 39:125-140.

SES (Standard Evaluation System for Rice) (2002) The Philippines: International Rice Research Institute, IRRI, Manila.

Sengupta D, Kannan M and Reddy AR (2011) A root proteomics-based insight reveals dynamic regulation of root proteins under progressive drought stress and recovery in *Vigna radiata* (L.) Wilczek. *Planta*, 233:1111–1127.

Singh S, Sidhu JS, Huang N, Vikal Y, Li Z, Brar DS, Dhaliwal HS and Khush GS (2001) Pyramiding three bacterial blight resistance genes (*xa5, xa13* and *Xa21*) using marker-assisted selection into *indica* rice cultivar PR106. *Theoretical Applied Genetics*, 102:1011–1015.

Singh AK, Gopala Krishnan S, Singh VP, Prabhu KV, Mohapatra T, Singh NK et al. (2011) Marker assisted selection: a paradigm shift in Basmati breeding. *Indian Journal of Genetics and Plant Breeding*, 71:1–9.

Small RL and Wendel JF (2000) Copy number lability and evolutionary dynamics of the *Adh* gene family in diploid and tetraploid cotton (Gossypium). *Genetics*, 155: 1913–1926.

Song W, Wang G, Chen L, Kim H, Pi L, Holsten T, Gardner J, Wang B, Zhai W, Zhu L, et al (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science*, 270:1804–1806.

Sreenivasaprasad S and Johnson R (2001) Eds XIII, Major Fungal Diseases of Rice: Recent Advances. 458 p.

Srinivasan B and Gnanamanickam S (2005) Identification of a new source of resistance in wild rice, *Oryza rufipogon* to bacterial blight of rice caused by Indian strains of *Xanthomonas oryzae* pv.*Oryzae. Current Science*, 88:1229-1231.

Sundaram RM, Vishnupriya MR, Biradar SK, Laha GS, Reddy GA, ShobhaRani N, Sarma NP and Sonti RV (2008) Marker assisted introgression of bacterial blight resistance in Samba Mahsuri, an elite*indica* rice variety. *Euphytica*, 160:411–422.

Sundaram RM, Vishnupriya MR, Laha GS, ShobhaRani N, SrinivasRao P, Balachandaran SM, Reddy GA, Sarma NP and Shonti RV (2009). Introduction of bacterial blight resistance into Triguna, a high yielding, mid-early duration rice variety. *Biotechnology Journal*, 4:400–407.

Subramoni S, Jha G and Sonti RV (2006) Virulence functions of xanthomonads. In SS Gnnnanamanickam (Ed.), Plant-associated bacteria, pp. 535–571.

Sugio A, Yang B, Zhu T and White FF (2007) Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OsTFIIAγ1* and *OsTFX1* during bacterial blight of rice. *Proceedings of the National Academy of Sciences*, USA, 104:10720-10725.

Sun X, Cao Y, Yang Z, Xu C, Li X, Wang S and Zhang Q (2004) *Xa26*, a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein. *Plant Journal*, 37:517–527.

Sun X, Cao Y and Wang S (2006) Point mutations with positive selection were a major force during the evolution of a receptor-kinase resistance gene family of rice. *Plant Physiology*, 140:998–1008.

Swamy P, Panchbhai AN, Dodiya P, Naik V, Panchbhai SD, Zehr UB, Azhakanandam K and Char BR (2006) Evaluation of bacterial blight resistance in rice lines carrying multiple resistance genes and *Xa21* transgenic lines. *Current Science*, 90.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123:585-595.

Tamura K, Dudley J, Nei M and Kumar S (2007) *MEGA4*: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24:1596-1599.

Tang T, Lu J, Huang JZ, He JH, McCouch SR and Shen Y (2006) Genomic variation in rice: genesis of highly polymorphic linkage blocks during domestication. *PLoS Genetics*, 2:1824–1833.

Tesfaye L, M Marco, Badebo A, Ammar K, Ricci A, Crossa J and Tuberosa R (2013) Searching for novel sources of field resistance to Ug99 and Ethiopian stem rust races in durum wheat via association mapping. *Theoretical Applied Genetics*, 126:1237–1256.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higguns DG (1997) The CLUSTAL_X Windows Interface: Flexible stratgies for Multiple Sequene Alignment Aided by Quality Analysis Tools. *Nucleic Acids Research*, 24:4876-4882.

Tagami Y and Mizukami T (1962) Historical review of the researches on bacterial blight of rice caused by *Xanthomonas oryzae* (Uyede and Ishiyama) Dowson. *Special Report of Plant Disease and Insect Pests Forecasting Service*, 10:1–112.

Thomma BP, Nurnberger T and Joosten MH (2011) Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *Plant Cell*, 23:4-15.

Till BJ, Reynolds SH, Green EA, Codomo CA, Enns LC, Johnson JE, et al. (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Research*, 13:524–30.

Tu J, Ona I, Zhang Q, Mew TW, Khush G S and Datta SK (1998) Transgenic rice variety IR72 with *Xa21* is resistant to bacterial blight. *Theoretical and Applied Genetics*, 97:31–36.

Tu J, Datta K, Khush GS, Zhang Q and Datta SK (2000) Field performance of *Xa21* transgenic *indica* rice. *Theoretical and Applied Genetics*, 101:15–20.

Utami DW, Lestari P and Koerniati S (2013) A Relative Expression of *Xa7* Gene Controlling Bacterial Leaf Blight Resistance in Indonesian Local Rice Population (*Oryza sativa* L.). *Journal of Crop Science and Biotechnology*, 16:1-7.

Vasudevan P (2002) Isolation and characterization of Bacillus spp. from the rice rhizosphere and their role in biological control of bacterial blight of rice caused by *Xanthomonas oryzae* pv. *oryzae*. Ph. D dissertation, University of Madras, India.

Vaughan DA (1994) Wild relatives of rice: Genetic resources handbook. (International Rice Research Institute, Manila, Philippines).

Vaughan DA, Morishima H and Kadowaki K (2003) Diversity in the genus *Oryza*. *Current Opinion in Plant Biology*, 6:139-146.

Velusamy P and Gnanamanickam SS (2003) Identification of 2,4-diacetylphloroglucinol (DAPG) production by plant-associated bacteria and its role in suppression of rice bacterial blight in India. *Current Science*, 85:1270–1273.

Velusamy P, Immanuel EJ, Gnanamanickam SS and Thomashow LS (2006) Biological control of rice bacterial blight by plant-associated bacteria producing 2,4-diacetylphloroglucinol. *Canadian Journal of Microbiology*, 52:56–65.

Vidhyasekaran P (2004) Consise Encyclopedia of Plant Pathology; Haworth Press: Philadelphia, PA, USA, pp. 511-517.

Velculescu VE, Zhang L, Vogelstein B and Kinzler KW (1995) Serial analysis of gene expression. *Science*, 270:484–487.

Vossen JH, Jo KR and Vosman B (2014) Mining the Genus *Solanum* for Increasing Disease Resistance R. Tuberosa et al. (eds.), Crop productivity, food security and nutritional quality. *Genomics of Plant Genetic Resources*, 27 DOI 10.1007/978-94-007-7575-6_2, 2:515.

Vuylsteke M, Peleman JD and van Eijk MJT (2007) AFLP-based transcript profiling (cDNA-AFLP) for genome-wide expression analysis. *Nature Protocol*, 2:1399–1413.

Wang ZY and Tanksley SD (1989) Restriction fragment length polymorphism in *Oryza sativa* L. *Genome*, 32:1113-1118.

Wang GL, Mackill DJ, Bonman M, McCouch SR, Champoux MC and Nelson RJ (1994) RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. *Genetics*, 136:1421–1434.

Wang GL, Song WY, Ruan DL, Sideris S and Ronald PC (1996) The cloned gene *Xa21* confers resistance to multiple *Xanthomonas oryzae* pv. *oryzae* isolates in transgenic plants. *Molecular Plant-Microbe Interactions*, 9:850–855.

Wang M, Allefs S, Berg RG, Vleeshouwers VGAA, Vossen EAG and Vosman B (2008) Allele mining in Solanum: conserved homologues of *Rpi-blb1* are identied in *Solanum stoloniferum. Theoretical Applied Genetics*, 116:933–943.

Wang C, Wen G, Lin X, Liu X and Zhang D (2008) Identification and fine mapping of the new bacterial blight resistance gene, *Xa31*(t), in rice. *European Journal of Plant Pathology*, 123:235-240.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7:256–276.

Wen N, Chu Z and Wang S (2003) Three types of defense-responsive genes are involved in resistance to bacterial blight and fungal blast diseases in rice. *Molecular Genetics and Genomics*, 269:331–339.

Xia C, Chen H and Zhu X (2012) Identification, Mapping, Isolation of the Genes Resisting to Bacterial Blight and Breeding Application in Rice. *Molecular Plant Breeding*, 3:120-130.

Xing Y, Frei U, Schejbel B, Asp T and Lübberstedt T (2007) Nucleotide diversity and linkage disequilibrium in 11 expressed resistance candidate genes in *Lolium perenne*. *BMC Plant Biology*, 7:43.

Xu QF, Cheng WS, Li SS, Li W, Zhang ZX, Xu YP, Zhou XP and Cai XZ (2012) Identification of genes required for Cf-dependent hypersensitive cell death by combined proteomic and RNA interfering analyses. *Journal of Experimental Botany*, 63:2421–2435.

Xiang Y, Cao Y, Xu C, Li X and Wang S (2006) *Xa3*, conferring resistance for rice bacterial blight and encoding a receptor kinase-like protein, is the same as *Xa26*. *Theoretical Applied Genetics*, 113:1347–1355.

Xu S, Cao Y, Li X and Wang S (2007) Expressional and biochemical characterization of rice disease resistance gene *Xa3/Xa26* family. *Journal of Integrative Plant Biology*, 49:852–862.

Xu Q, Ni H, Chen Q, Sun F, Zhou T, et al. (2013) Comparative Proteomic Analysis Reveals the Cross-Talk between the Responses Induced by H_2O_2 and by Long-Term Rice-Black Streaked Dwarf Virus Infection in Rice. *PLoS ONE*, 8: e81640.

Yang Y et al. (1997) Signal perception and transduction in plant defense responses. *Genes Development*, 11:1621–1639.

Yang Z, Sun X, Wang S and Zhang Q (2003) Genetic and physical mapping of a new gene for bacterial blight resistance in rice. *Theoretical Applied Genetics*, 106:1467-1472.

Yang B, Sugio A and White FF (2006) Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of National Academy of Sciences* USA, 103:10503–10508.

Yang F, Jacobsen S, Jørgensen HJL, Collinge DB, Svensson B and Finnie C (2013) Fusarium graminearum and its interactions with cereal heads: studies in the proteomics era. *Frontiers in Plant Science*, 4:1-8.

Yoshida K and Miyashita NT (2009) DNA polymorphism in the blast resistance gene *Pita* of wild rice *Oryza rufipogon* and its related species. *Genes Genetics and Systemetics*, 84:121-126.

Yoshimura S, Yoshimura A, Iwata N, McCouch SR, Abenes ML, Baraoidan MR, Mew TW and Nelson RJ (1995) Tagging and combining bacterial blight resistance genes in rice using RAPD and RFLP markers. *Molecular Breeding*, 1:375-387.

Yoshimura S, Yamanouchi U, Katayose Y, Toki S, Wang ZX, Kono I, Kuruta N, Yano M, Iwata N and Sasaki T (1998) Expression of *Xa1*, a bacterial blight resistance gene in rice, is induced by bacterial inoculation. *Proceedings of National Academy of Science* USA, 95:1663–1668.

Young SA, Wang X and Leach JE (1996) Changes in the plasma membrane distribution of rice phospholipase D during resistant interactions with *Xanthomonas oryzae* pv. *oryzae*. *Plant Cell*, 8:1079–1090.

Yuan M, Chu Z, Li X, Xu C and Wang S (2009) Pathogen-induced expressional loss of function is the key factor in race-specific bacterial resistance conferred by a recessive R gene *xa13* in rice. *Plant Cell Physiology*, 50*:*947-955.

Yuan M, Chu Z, Li X, Xu C and Wang S (2010) The bacterial pathogen *Xanthomonas oryzae* overcomes rice defenses by regulating host copper redistribution. *Plant Cell*, 22*:*3164-2176.

Zeng LX, Huang SH, Wu SZ (2002) The resistance of IRBB21 (*Xa21*) against 5 races of Guangdong Province. *Phytopathology*, 29:97-100.

Zeng X, Wang , Ali M, Zhang H, Liu X, Li W and Ji W (2010) Profiling gene expression patterns of stripe rust (*Puccinia striiformis* f. sp. tritici) resistance gene in new wheat germplasm. *Pak J Bot*, 42:425-4266.

Zhang Q and Ling SC (1994) Evaluation of resistance to bacterial blight (*Xanthomonas oryzae pv. Oryzae)* in wild rice species. *Scientia Agricultura Sinica*, 27:1-9.

Zhang Q, Lin SC, Zhao BY, Wang CL, Yang WC, Zhou YL, Li DY, Chen CB and Zhu LH (1998) Identification and Tagging a new gene for resistance to bacterial blight (*Xanthomonas oryzae pv oryzae*) from *Oryza rufipogon. Rice Genetic Newsletter*, 15:138-142.

Zhang W, Fraiture M, Kolb D, Löffelhardt B, Desaki Y, et al. (2013) Arabidopsis RECEPTOR-LIKE PROTEIN and receptor-like kinase SUPPRESSOR OF BIR1-1/EVERSHED mediate innate immunity to necrotrophic fungi. *Plant Cell*, 25:4227-4241.

Zhai WX, Wang WM, Zhou YL, Li XB, Zheng XW, Zhang Q, et al. (2002) Breeding bacterial blight-resistant hybrid rice with the cloned bacterial blight resistance gene *Xa21*. *Molecular Breeding*, 8:285–293.

Zhao J, Fu J, Li X, Xu C and Wang S (2009) Dissection of the factors affecting development controlled and race specific disease resistance conferred by leucine rich repeat receptor kinase type *R* genes in rice. *Theoretical and Applied Genetics*, 119:231-239.

Zhu Q, Zheng X, Luo J, Gaut BS and Ge S (2007) Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: Severe bottleneck during domestication of rice. *Molecular Biology and Evolution*, 24:875-882.

Zhu QH et al. (2012) Characterization of the defense transcriptome responsive to *Fusarium oxysporum*-infection in Arabidopsis using RNA-seq. *Gene*, http://dx.doi.org/10.1016/j.

Zhou B, Peng K, Chu Z, Wang S, Zhang Q (2002) The defense-responsive genes showing enhanced and repressed expression after pathogen infection in rice (*Oryza sativa* L.). *Sci China C Life Sci*, 45:449–467.

Zhou W, Eudes F and Laroche A (2006) Identification of differentially regulated proteins in response to a compatible interaction between the pathogen *Fusarium graminearum* and its host, *Triticum aestivum*. *Proteomics*, 6:4599–4609.

Zou XH, Zhang FM, Zhang JG, Zang LL, Tang L, Wang J, Sang T and Ge S (2008) Analysis of 142 genes resolves the rapid diversification of the rice genus. *Genome Biology*, 9:113.

ANNEXURE 1

Accessions of different *Oryza* species used in our study and their disease reaction against five different Xoo strains.

S.No.	Code Name	IRGC/ Accession No.	Scientific name	Disease response against five *Xoo* strains									
				DX015		DX011		DX133		DX127		DX020	
				S†	P	S†	P	S†	P	S†	P	S†	P
1.	WR1*	OS 28	*O. nivara*	5	R	7	S	7	S	5	R	5	R
2.	WR9	ON 38	*O. nivara*	7	S	7	S	5	R	5	R	3	R
3.	WR10*	ON 38-1	*O. nivara*	3	R	3	R	1	R	3	R	3	R
4.	WR11*	M 209 A	*O. nivara*	1	R	3	R	5	R	5	R	3	R
5.	WR14*	OS 23-2	*O. nivara*	1	R	7	S	7	S	3	R	5	R
6.	WR17	42446	*O. nivara*	7	S	7	S	7	S	7	S	7	S
7.	WR38	45991	*O. nivara*	7	S	7	S	7	S	1	R	5	R
8.	WR52	101514	*O. nivara*	1	R	1	R	5	R	3	R	3	R
9.	WR56	21028	*O. nivara*	1	R	3	R	5	R	1	R	1	R
10.	WR69	21990	*O. nivara*	5	R	7	S	9	S	7	S	7	S
11.	WR82*	81832	*O. nivara*	1	R	1	R	1	R	1	R	1	R
12.	WR84	81852	*O. nivara*	1	R	1	R	1	R	1	R	3	R
13.	WR85*	100963	*O. nivara*	3	R	3	R	5	R	3	R	3	R
14.	WR86*	102166	*O. nivara*	7	S	5	R	3	S	5	S	5	R
15.	WR102*	106133	*O. nivara*	1	R	1	R	1	R	1	R	1	R
16.	WR110*	210223	*O. nivara*	3	R	3	R	5	R	3	R	3	R
17.	WR117	Unknown	*O. nivara*	5	R	5	R	5	R	7	S	5	R
18.	WR132*	ON132	*O. nivara*	1	R	1	R	3	R	1	R	3	R
19.	SL1	ON01	*O. nivara*	9	S	5	R	5	R	3	R	7	S
20.	SL15*	ON15	*O. nivara*	1	R	1	R	1	R	1	R	1	R
21.	SL17	ON17	*O. nivara*	9	S	7	S	9	S	7	S	5	R
22.	SL19	ON19	*O. nivara*	5	R	7	S	7	S	7	S	3	R
23.	SL24*	ON24	*O. nivara*	3	R	5	R	7	S	5	R	3	R
24.	SL33	ON33	*O. nivara*	1	R	5	R	3	R	5	R	3	R
25.	SL34*	ON34	*O. nivara*	9	S	5	R	5	R	5	R	5	R
26.	SL44*	ON44	*O. nivara*	1	R	1	R	1	R	3	R	1	R
27.	NB16*	IC384116	*O. alta*	3	R	3	R	1	R	3	R	1	R
28.	NB40*	IC203740	*O. officinalis*	5	R	3	R	5	R	1	R	3	R
29.	N76	IC384076	*O. officinalis*	7	S	5	R	5	R	5	R	5	R
30.	CEMPOCELAK	Not available	*O. sativa*	7	S	7	S	9	S	9	S	9	S
31.	DV85	Not available	*O. sativa*	7	S	7	S	7	S	9	S	9	S
32.	IR20*	Not available	*O. sativa*	9	S	7	S	7	S	5	R	7	S

33.	ZENITH	Not available	*O. sativa*	7	S	7	S	9	S	7	S	9	S
34.	TETEP	Not available	*O. sativa*	7	S	5	R	7	S	7	S	7	S
35.	CAS209	Not available	*O. sativa*	7	S	7	S	9	S	9	S	9	S
36.	IET8320	Not available	*O. sativa*	7	S	7	S	7	S	9	S	9	S
37.	AJAYA	Not available	*O. sativa*	5	R	7	S	9	S	7	S	9	S
38.	JANA14	Not available	*O. sativa*	7	S	9	S	7	S	5	R	7	S
39.	IR59	Not available	*O. sativa*	3	R	3	R	3	R	1	R	3	R
40.	PB1	Not available	*O. sativa*	9	S	7	S	9	S	9	S	9	S
41.	IR24	Not available	*O. sativa*	9	S	9	S	9	S	9	S	9	S
42.	TN1*	Not available	*O. sativa*	9	S	9	S	9	S	9	S	9	S
43.	RADHA	Not available	*O. sativa*	9	S	7	S	9	S	9	S	9	S
44.	R.BHAGAWATI	Not available	*O. sativa*	9	S	9	S	7	S	7	S	9	S
45.	SAROJ	Not available	*O. sativa*	7	S	7	S	7	S	7	S	7	S
46.	SATYAM	Not available	*O. sativa*	7	S	7	S	9	S	7	S	7	S
47.	SANTOSH	Not available	*O. sativa*	9	S	7	S	7	S	7	S	5	R
48.	DHANALAXMI	Not available	*O. sativa*	7	S	9	S	9	S	9	S	7	S
49.	PAU201	Not available	*O. sativa*	1	R	5	R	5	R	3	R	3	R
50.	PAU915	Not available	*O. sativa*	7	S	7	S	5	R	7	S	5	R
51.	PAU933*	Not available	*O. sativa*	5	R	5	R	5	R	7	S	3	R
52.	R1	Not available	*O. sativa*	7	S	7	S	9	S	9	S	7	S
53.	R2*	Not available	*O. sativa*	5	R	5	R	3	R	3	R	5	R
54.	R3	Not available	*O. sativa*	7	S	7	S	7	S	5	R	7	S
55.	R4*	Not available	*O. sativa*	5	R	7	S	3	R	3	R	5	R
56.	R5	Not available	*O. sativa*	7	S	7	S	7	S	7	S	7	S
57.	R6	Not available	*O. sativa*	7	S	7	S	9	S	7	S	9	S
58.	R8	Not available	*O. sativa*	5	R	9	S	9	S	7	S	7	S
59.	R9*	Not available	*O. sativa*	5	R	7	S	5	R	5	R	5	R
60.	R10	Not available	*O. sativa*	5	R	7	S	9	S	9	S	9	S
61.	R12	Not available	*O. sativa*	7	S	7	S	7	S	7	S	7	S
62.	R14	Not available	*O. sativa*	7	S	9	S	9	S	9	S	9	S
63.	R15	Not available	*O. sativa*	7	S	7	S	7	S	9	S	9	S
64.	R17	Not available	*O. sativa*	7	S	9	S	9	S	7	S	9	S
65.	R18	Not available	*O. sativa*	7	S	7	S	7	S	9	S	9	S
66.	R19	Not available	*O. sativa*	7	S	9	S	9	S	7	S	7	S
67.	R20	Not available	*O. sativa*	9	S	9	S	9	S	9	S	7	S
68.	R21	Not available	*O. sativa*	9	S	9	S	9	S	9	S	9	S
69.	R22	Not available	*O. sativa*	7	S	7	S	7	S	7	S	7	S
70.	R23*	Not available	*O. sativa*	5	R	5	R	5	R	7	S	5	R
71.	R24	Not available	*O. sativa*	9	S	7	S	9	S	7	S	9	S
72.	R25	Not available	*O. sativa*	7	S	7	S	7	S	5	R	5	R

73.	R28	Not available	*O. sativa*	7	S	7	S	9	S	9	S	9	S
74.	R29	Not available	*O. sativa*	7	S	7	S	7	S	7	S	7	S
75.	R30	Not available	*O. sativa*	9	S	7	S	9	S	7	S	7	S
76.	R31	Not available	*O. sativa*	7	S	7	S	7	S	7	S	7	S
77.	R33	Not available	*O. sativa*	7	S	7	S	9	S	9	S	9	S
78.	R35	Not available	*O. sativa*	7	S	9	S	9	S	9	S	9	S
79.	R36	Not available	*O. sativa*	7	S	7	S	9	S	9	S	9	S
80.	R37	Not available	*O. sativa*	9	S	9	S	9	S	9	S	9	S
81.	R38	Not available	*O. sativa*	7	S	7	S	9	S	7	S	9	S
82.	R39	Not available	*O. sativa*	7	S	7	S	7	S	7	S	7	S
83.	R40	Not available	*O. sativa*	7	S	9	S	9	S	9	S	9	S
84.	R41	Not available	*O. sativa*	9	S	7	S	9	S	9	S	9	S
85.	R42	Not available	*O. sativa*	7	S	7	S	9	S	7	S	7	S
86.	R43	Not available	*O. sativa*	9	S	7	S	7	S	7	S	7	S
87.	R44	Not available	*O. sativa*	7	S	9	S	9	S	7	S	7	S
88.	R45	Not available	*O. sativa*	7	S	7	S	7	S	9	S	7	S
89.	R46	Not available	*O. sativa*	7	S	9	S	9	S	9	S	7	S
90.	R47*	Not available	*O. sativa*	7	S	5	R	5	R	5	R	5	R
91.	R48	Not available	*O. sativa*	7	S	7	S	9	S	9	S	7	S
92.	R49	Not available	*O. sativa*	7	S	9	S	9	S	9	S	9	S
93.	R27	Not available	*O. sativa*	7	S	7	S	7	S	7	S	7	S
94.	MP*	Not available	*O. sativa*	5	R	5	R	5	R	5	R	3	R
95.	ML*	Not available	*O. sativa*	5	R	5	R	3	R	3	R	3	R
96.	IRBB1	115095	*O. sativa*	9	S	9	S	9	S	7	S	9	S
97.	IRBB3	115100	*O. sativa*	7	S	9	S	9	S	7	S	7	S
98.	IRBB4	115101	*O. sativa*	7	S	9	S	9	S	9	S	9	S
99.	IRBB5	115102	*O. sativa*	5	R	5	R	7	S	5	R	9	S
100.	IRBB7	115119	*O. sativa*	9	S	9	S	9	S	7	S	7	S
101.	IRBB10	115606	*O. sativa*	9	S	9	S	9	S	7	S	7	S
102.	IRBB11	115096	*O. sativa*	7	S	7	S	7	S	7	S	7	S
103	IRBB13	115097	*O. sativa*	3	R	3	R	5	R	1	R	3	R
104.	IRBB21	115099	*O. sativa*	1	R	1	R	3	R	5	R	5	R

S†= Disease Score, P = phenotype, R = Resistant, S = Susceptible

ANNEXURE 2

Media preparation

1. Modified Wakimoto medium

$Ca(NO_3)_2 . 4H_2O$.....0.5 g
$Na_2HPO_4 . 12H_2O$.....2.0 g
Peptone.....5.0 g
Sucrose.....15.0 g
$FeSO_4 . 7H_2O$ .....0.5 g
Agar.....15.0 g
Distilled water.....1.0 L

2. Luria Bertani medium

Tryptone……..10.0 g
Yeast Extract……...5.0 g
Nacl……..10.0 g
Distilled water.....1.0 L

ANNEXURE 3

Preparation of Stock Solutions

1 M Tris HCl: 30.275 g of 1 M Trizma base (Tris hydroxy methyl amino methane-MW 121.1) was dissolved in about 180 ml distilled water. Then, the pH was adjusted to 8 using 1 N HCl and finally the volume was made up to 250 ml. This stock solution was autoclaved and stored at room temperature.

0.5 M EDTA: 46.53 g of Ethyl diamine tetra acetic acid (EDTA-MW 372.24) was dissolved in the 180 ml of distilled water. Then, the pH of the solution was adjusted to 8.0 with NaOH and volume was made to 250 ml. This stock solution was autoclaved and stored at room temperature.

TE buffer (pH 8.0): The buffer was prepared with following concentrations.

S.No.	Components	Stock conc.	Final conc.	Amount of stock per 100 ml
1	Tris-Hcl (pH 8)	1M	10 mM	1 ml
2	EDTA (pH 8)	0.5 M	1 mM	0.2 ml
3	Distilled water	-	-	98.8 ml

2.5 M NaCl: 36.5 g of NaCl (MW 58.44) was dissolved in 250 ml of distilled water. The solution was autoclaved and stored at room temperature.

DNA extraction buffer: It was prepared by mixing the following components.

S.No.	Components	Stock conc.	Amount of stock per 100 ml
1	Tris-Hcl (pH 8)	1.0 M	10 ml
2	EDTA (pH 8)	0.5 M	4 ml
3	NaCl	5.0 M	28 ml
4	CTAB	2 %	2 g
5	PVP	2 %	2 g
6	Distilled water	-	58 ml

Phenol : Chloroform : Isoamyl alcohol (25:24:1): For preparation of 250 ml of stock solution, 125 ml phenol, 120 ml chloroform and 5 ml isoamyl alcohol were mixed well to get the required concentration.

TAE Buffer preparation: One litre of (50X) stock solution of TAE was prepared as follows: 242 g Tris base (FW = 121.14; Sigma, USA) was dissolved in approximately 750 ml deionized water. Carefully 57.1 ml of glacial acid and 100 ml of 0.5 M EDTA (pH 8.0) was added, mixed well by stirring. The volume was finally made up to 1000 ml using double distilled water. The solution was filter sterilized, thoroughly mixed, autoclaved and stored at room temperature. This stock solution can be stored at room temperature. The working solution of 1x TAE buffer was made by simply diluting the stock solution by 50X in deionized water. The buffer is now ready for running an agarose gel.

10X TBE Buffer Preparation: Mix following components to make 1 liter 10X TBE buffer

S.No.	Components	Amount	10X Stock Conc.	Final 1X Conc.
1	Tris Base	108 g	890 mM	89 mM
2	Boric Acid	55 g	890 mM	89 mM
3	EDTA (pH 8.0)	40 ml	20 mM	2 mM

ANNEXURE 4

Preparation protein extraction buffer

S.No.	Components	Final concentration	Amount (For 100 ml)
1	2MTris-HCl, pH 7.5	0.5 M	25 ml
2	Sucrose	0.7 M	23.96 g
3	Kcl	0.1M	0.745 g
4	EDTA	50 mM	10 ml
5	β-mercaptoethanol	2 %	*
6	PMSF	1mM	*

* β -mercaptoethanol and PMSF was added after grinding in Eppendorf tube directly.

Rehydration buffer

S.No.	Components	Final concentration	Amount (For 10 ml)
1	Urea (MW-60)	8 M	4.8 g
2	Thiourea	2 M	1.52 g
3	CHAPS	4 %	0.4 g
4	DTT	20 mM	0.015 g
5	IPG Buffer	0.8 %	80 µl

SDS equilibration buffer

(50 mM Tris-HCl, pH 8.8, 6 M urea, 30% glycerol, 2% SDS, bromophenol blue, 200 ml)

S.No.	Components	Final concentration	Amount (for 200 ml)
1	Tris-HCl, pH 8.8	50 mM	10.0 ml
2	Urea (FW 60.06)	6 M	72.07 g
3	Glycerol (87% v/v)	30% (v/v)	69 ml
4	SDS (FW 288.38)	2% (w/v)	4.0 g
5	Bromophenol blue	0.002% (w/v)	400 µl of 1% solution
6	Double distilled H2O		to 200 ml

This is a stock solution. Prior to use DTT or iodoacetamide are added. Store at -20 °C.

12% SDS Polyacrylamide Gel:

S.No.	Components	Stock conc.	Amount (for 100 ml)
1	Acrylamide	30 %	40.0 ml
2	Tris, pH 8.8	1.5 M	25 ml
3	SDS	10%	1 ml
4	APS	10%	1 ml *
5	TEMED		40 µl *
6	Double distilled H2O		33 ml

***APS and TEMED are added just before poring.**

Agarose Sealing Mixture:

S.No.	Components	Amount
1	1X TAE Buffer	50 ml
2	Agarose	0.5 g
3	1% Bromophenol Blue	100 µl

Microwave to boiling point to dissolve the Agarose.

SDS PAGE Running Buffer (10x):

S.No.	Components	Amount (1 litre)
1	Glycine	141.1 g
2	Tris base	30.3 g
3	SDS	10 g
4	Double distilled H2O	to 1000 ml

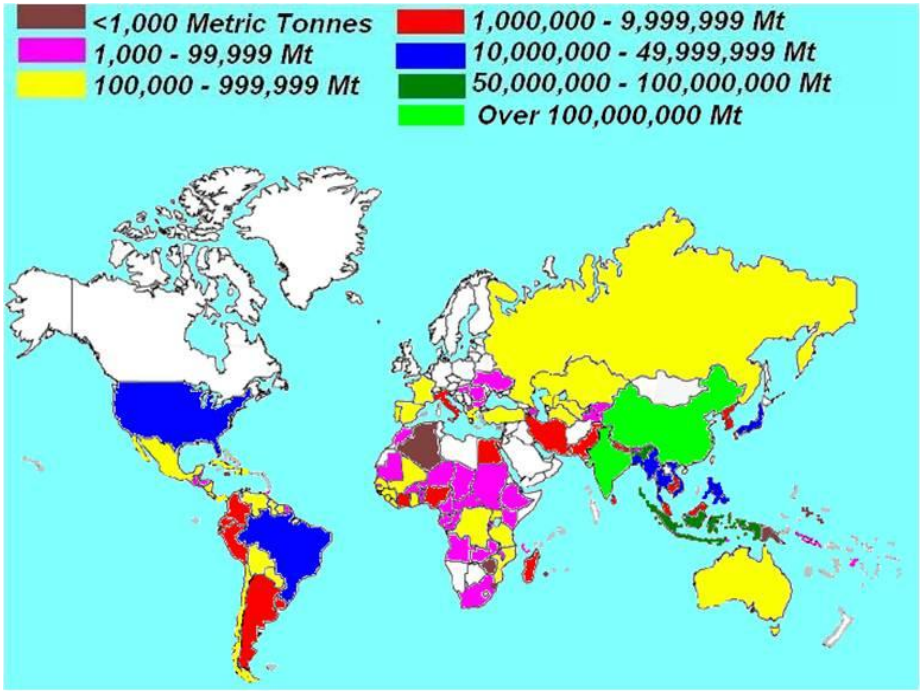


Fig. 2.1: A world map showing rice production in metric tonnes across the world in 2005. (Source: gramene.org)

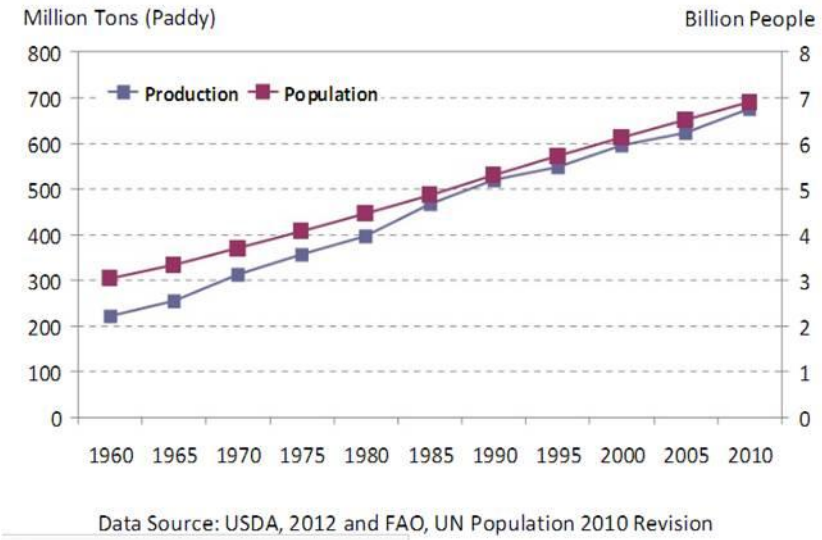
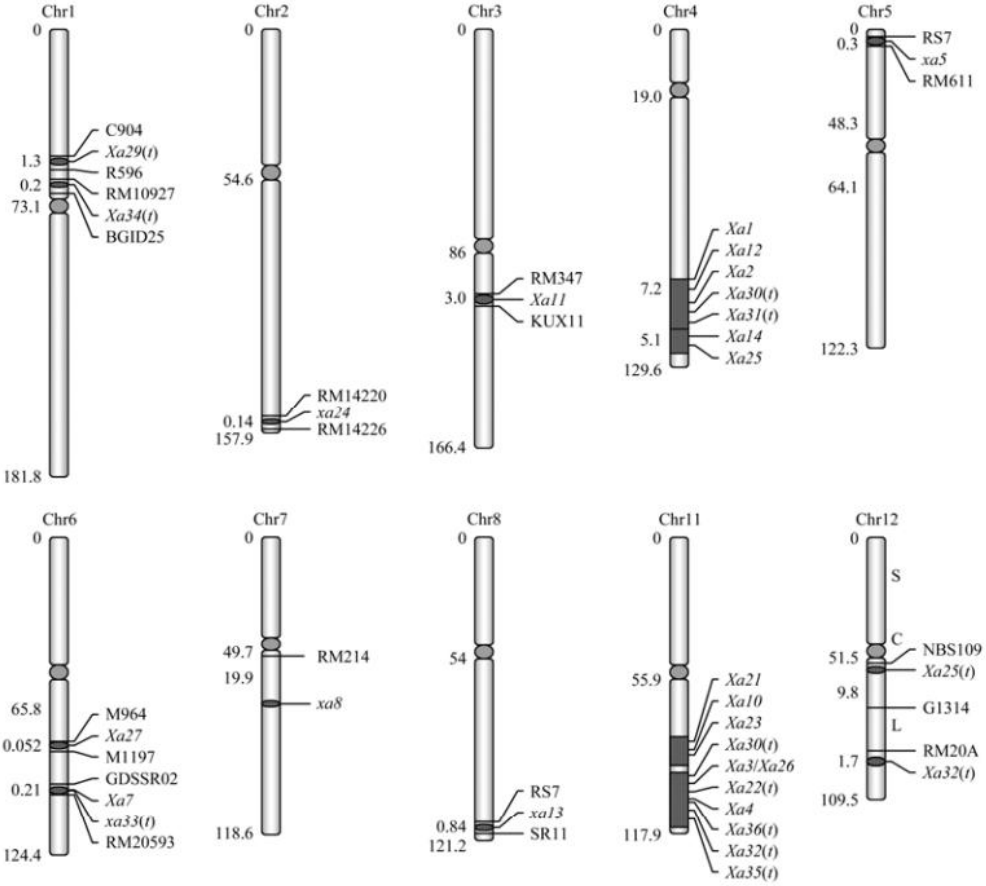


Fig.2.2: A graph showing global rice production versus population growth. From 1960 onwards we can see an exponential growth in population and it exceeds the global rice production.

Fig. 2.3: Major diseases of rice affecting its yield and productivity. (a) and (b) bacterial diseases, (c) and (d) fungal diseases and (e) and (f) viral diseases of rice. (Source: internet)

(a) (b)

Fig. 2.4: Two phases of bacterial blight diseases has been shown (a) Kresek and (b) Leaf blight in the picture. The severity of the disease in field condition is shown in the picture. (Source: internet)



Source: Xia et al. 2012

Fig. 2.5: Chromosomal location of different mapped *R* genes for BB resistance is shown on the different chromosomes of rice with their linked markers. Chromosome 1 possess *Xa29(t)* and *Xa34(t),* Chromosome 2 *(xa24),* Chromosome 3 *(Xa11),* Chromosome 4 *(Xa1, Xa12, Xa2, Xa30t, Xa31t, Xa14 and Xa25),* Chromosome 5 *(xa5),* Chromosome 6*(Xa27, Xa7* and *xa33t),* Chromosome 7 *(xa8),* Chromosome 8 *(xa13),* Chromosome 11 *(Xa21, Xa10, Xa23, Xa30t, Xa3/Xa26, Xa22t, Xa4, Xa36t, Xa32t and Xa33t)* and Chromosome 12 *(Xa25t, Xa32t).*

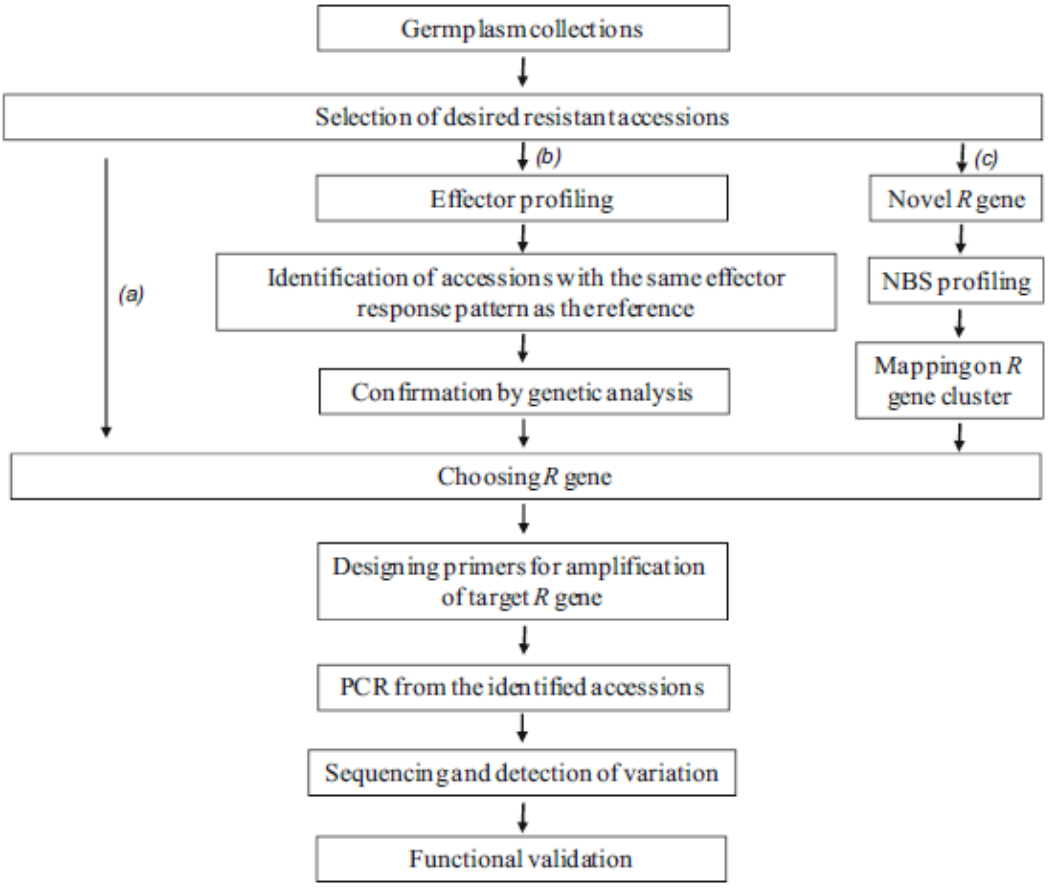


Figure Courtesy: Vossen et al. 2014

Figure 2.6 A flowchart showing the pipeline of allele mining *R* genes. (a) Allele mining process where effector molecules are unknown for pathogen under study, (b) combine effector profiling and allele mining approach and (c) discovery of novel *R* genes in combination of NBS profiling and allele mining

Fig.3.1: Screening of rice plants for resistance against DX011, DX127, DX020, DX015 and DX133 *Xoo* strains in (A, B, C) green house and (D) field

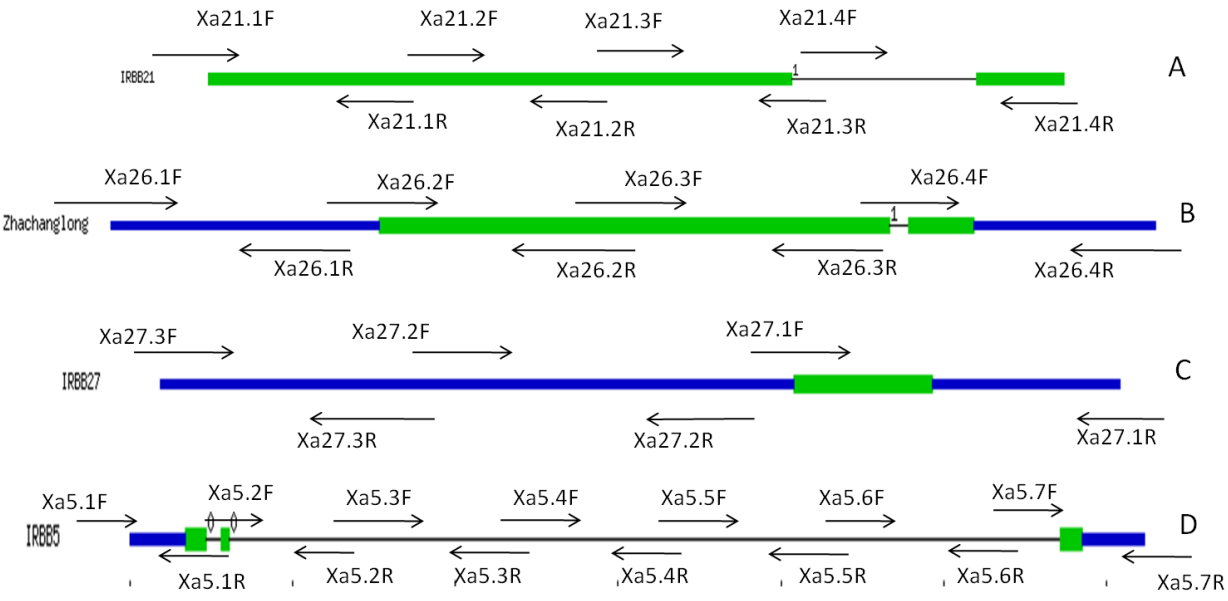


Fig. 3.2: Strategy of designing gene specific primers for allelic diversity analysis and allele mining. Primers were designed for (A), *Xa21* (B), *Xa26* (C) *xa5* and (D) *Xa27* in which one reverse primer overlapped with another forward primer. Blue regions indicate non coding and green color indicate exonic region.

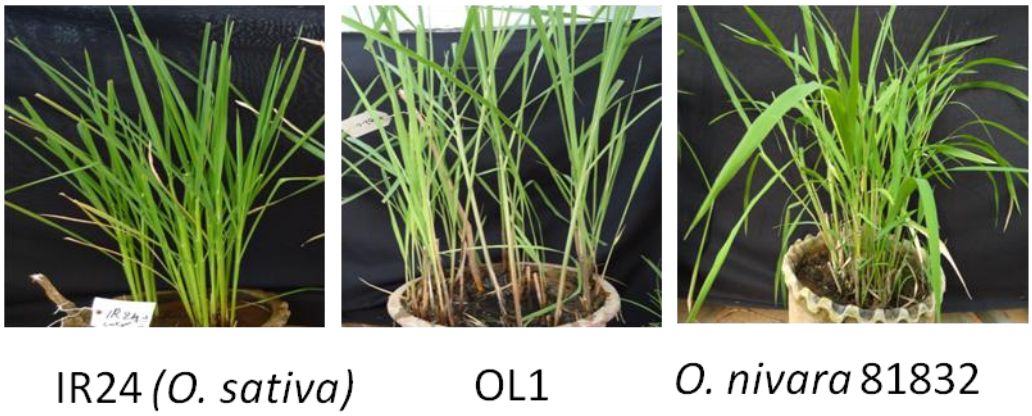


Fig. 3.3: Plant material used for transcript and translation profile. Three genotypes IR24 (susceptible control), *Oryza longistaminata* OL1 (resistant) and *Oryza nivara* 81832 were infected with Xoo isolate DX011 by syringe infiltration and leaf clipping method. Transcript analysis was carried out for IR24, OL1 and 81832 while proteome study was done only for infected and non infected 81832.

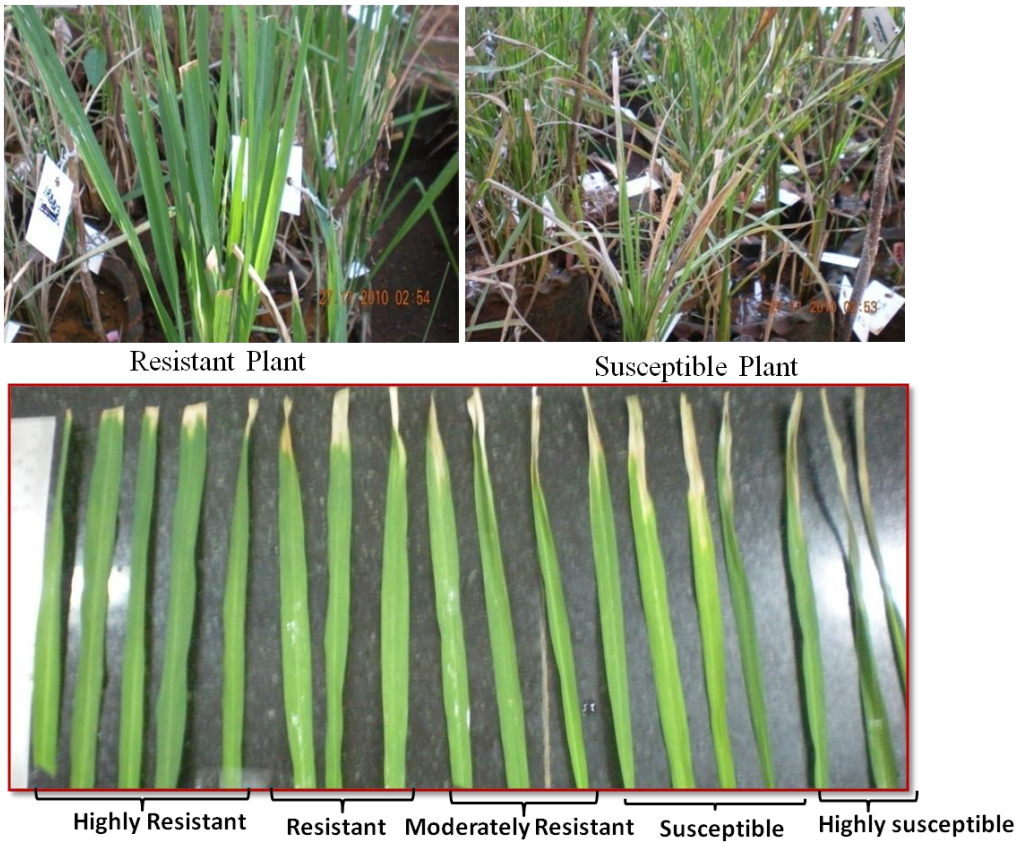


Fig.4.1.1: Green house screened plants are shown in the picture. Lesion length of different wild and cultivated rice species after leaf clipping inoculation are shown. Picture on the top indicate resistant and susceptible plants. Resistance plants were healthy while the susceptible plants were wilted. The leaf samples with more lesion length indicate more susceptibility to BB while leafs with least lesion length indicate strong resistance to the disease.

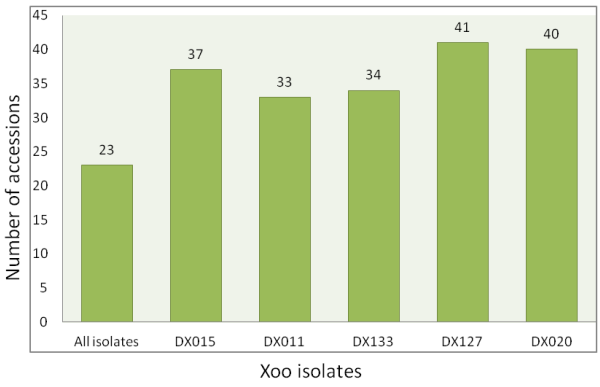


Fig.4.1.2: A bar graph representing the number of accessions showing resistance to different *Xoo* isolates. X axis represent name of isolates and Y axis indicate total number of accessions.

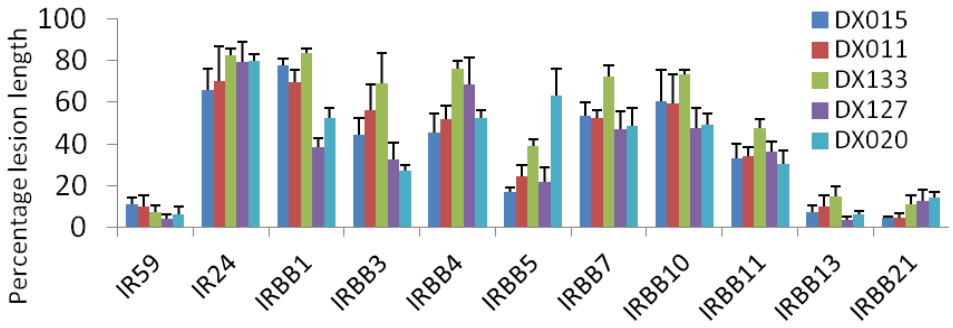


Fig.4.1.3: Level of resistance of different NILs towards five *Xoo* isolates are shown in this graph. X-axis represent NILs and Y- axis the disease score. Lesser the disease score, more the resistance.

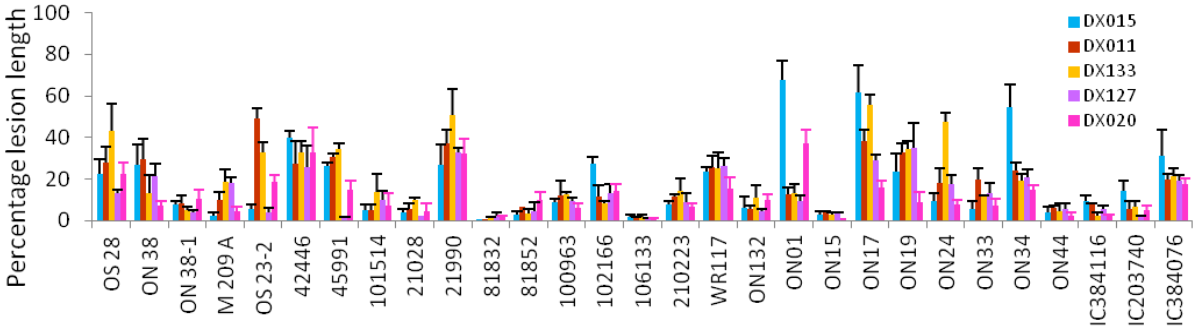


Fig.4.1.4: Level of resistance of different wild species towards five *Xoo* isolates. The names of the isolates are mentioned on the right side of the graph.

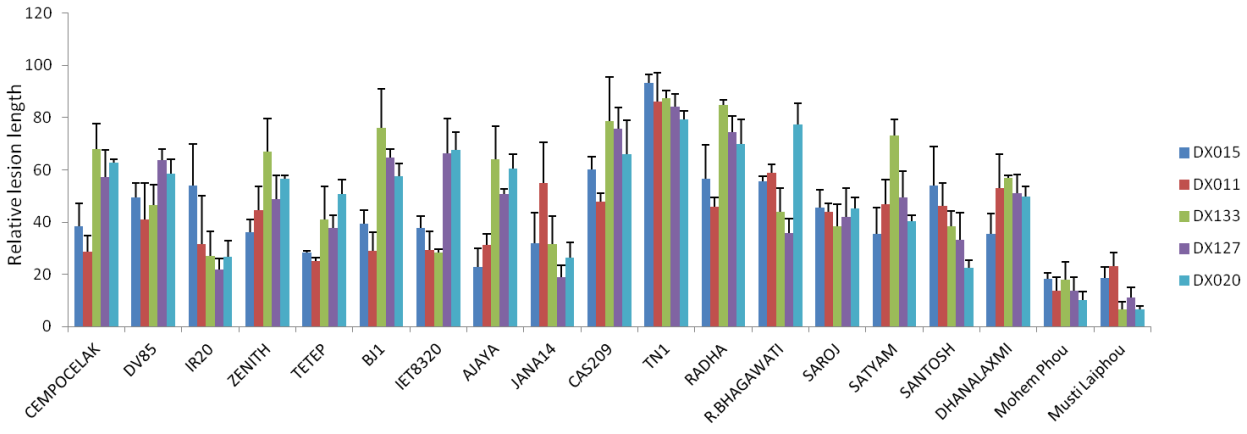


Fig. 4.1.5: Level of resistance of different cultivars towards five *Xoo* isolates. The names of the isolates are mentioned on the right side of the graph.

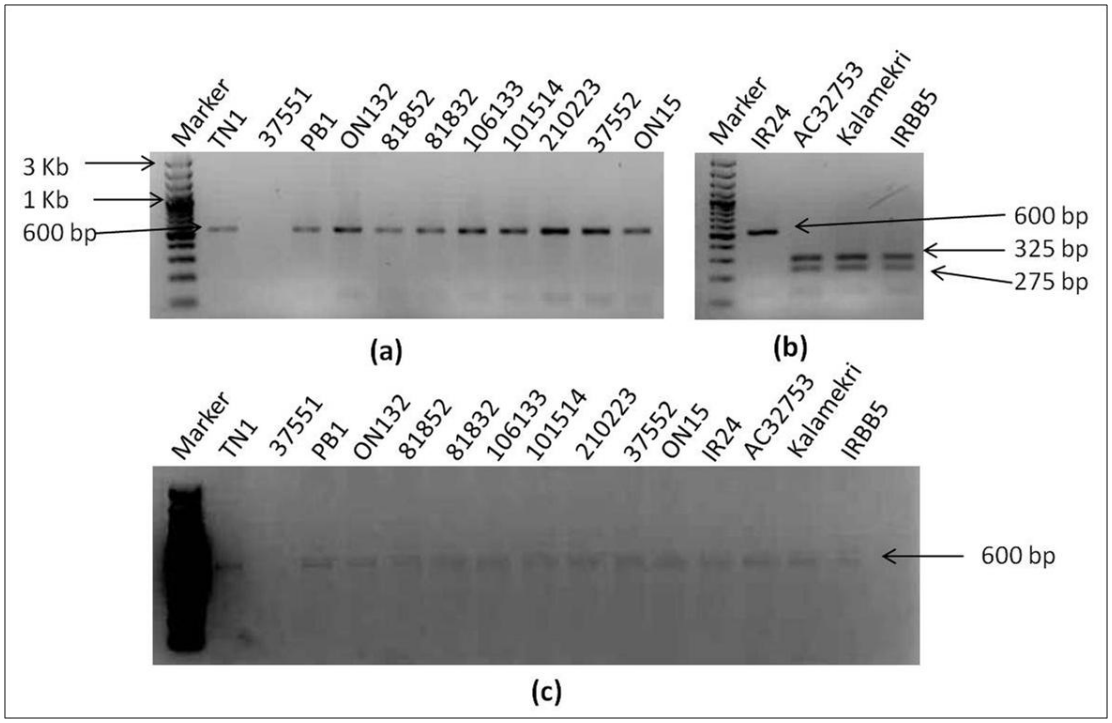


Fig. 4.1.6: Schematic presentation of genotypic screening using functional CAPs marker reported by (Iyer et al. 2006). (a) Gel picture showing PCR product after restriction digestion with BsrI enzyme resolved on 3% gel. The primer was designed at the polymorphic site between dominant (*Xa5*) and recessive (*xa5*). (b) Only the recessive alleles are digested by BsrI. Band size around 600 bp is dominant and susceptible allele while two bands around 325 bp and 275 bp indicated the presence of recessive and resistant allele. (c) PCR product before restriction digestion with BsrI enzyme. Accession numbers of the genotypes are mentioned above the respective wells. The DNA marker used is 100 bp plus (100 bp – 3000 bp).

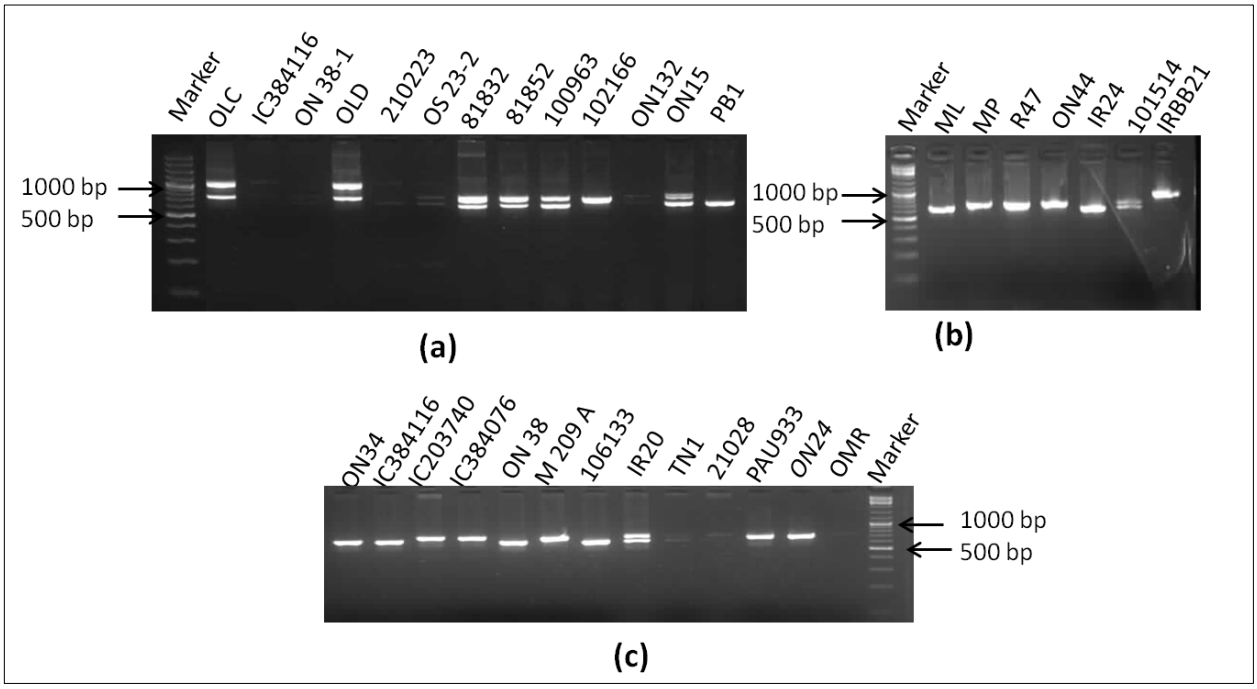


Fig. 4.1.7: Schematic presentation of genotypic screening using PTA248 marker (Ronald et al. 1996). (a), (b) & (c) Gel picture showing PCR product after resolving on 2% agarose gel. The primer was designed at the polymorphic site between dominant (*Xa21*) and recessive (*xa21*). Dominant alleles showed a band size around 1 Kb. Accession numbers of the genotypes and the cultivars are mentioned above the respective wells. The DNA marker used is 100 bp plus (100 bp – 3000 bp).

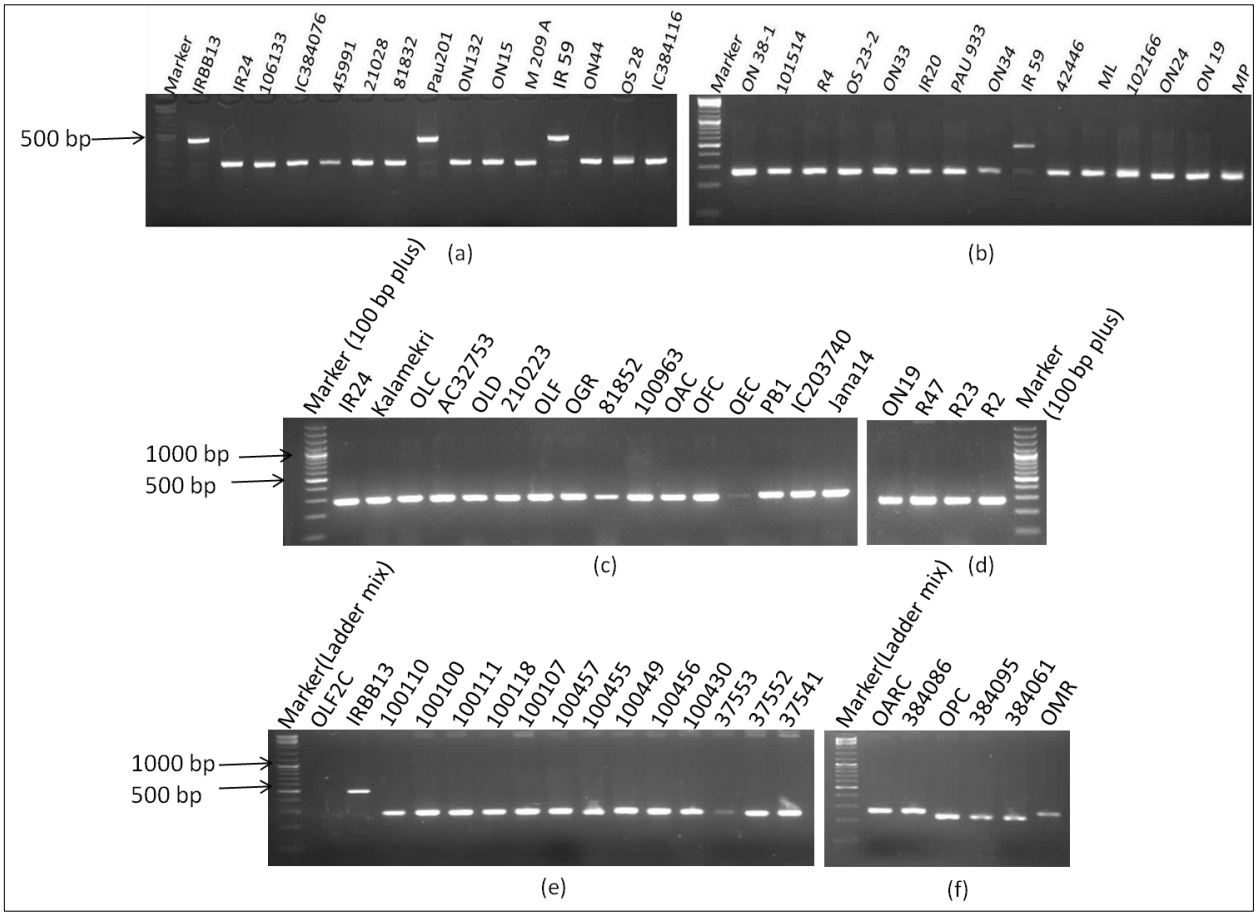


Fig. 4.1.8: Schematic presentation of screening of polymorphic bands of *xa13* gene (a), (b), (c), (d), (e) and (f). PCR was performed using functional marker FM13 (Hui et al. 2010). Band size around 500 bp indicates the presence of recessive and resistant *xa13* allele. Accession numbers of the genotypes and the cultivars are mentioned above the respective wells. The DNA marker used is 100 bp plus (100 bp – 3000 bp).

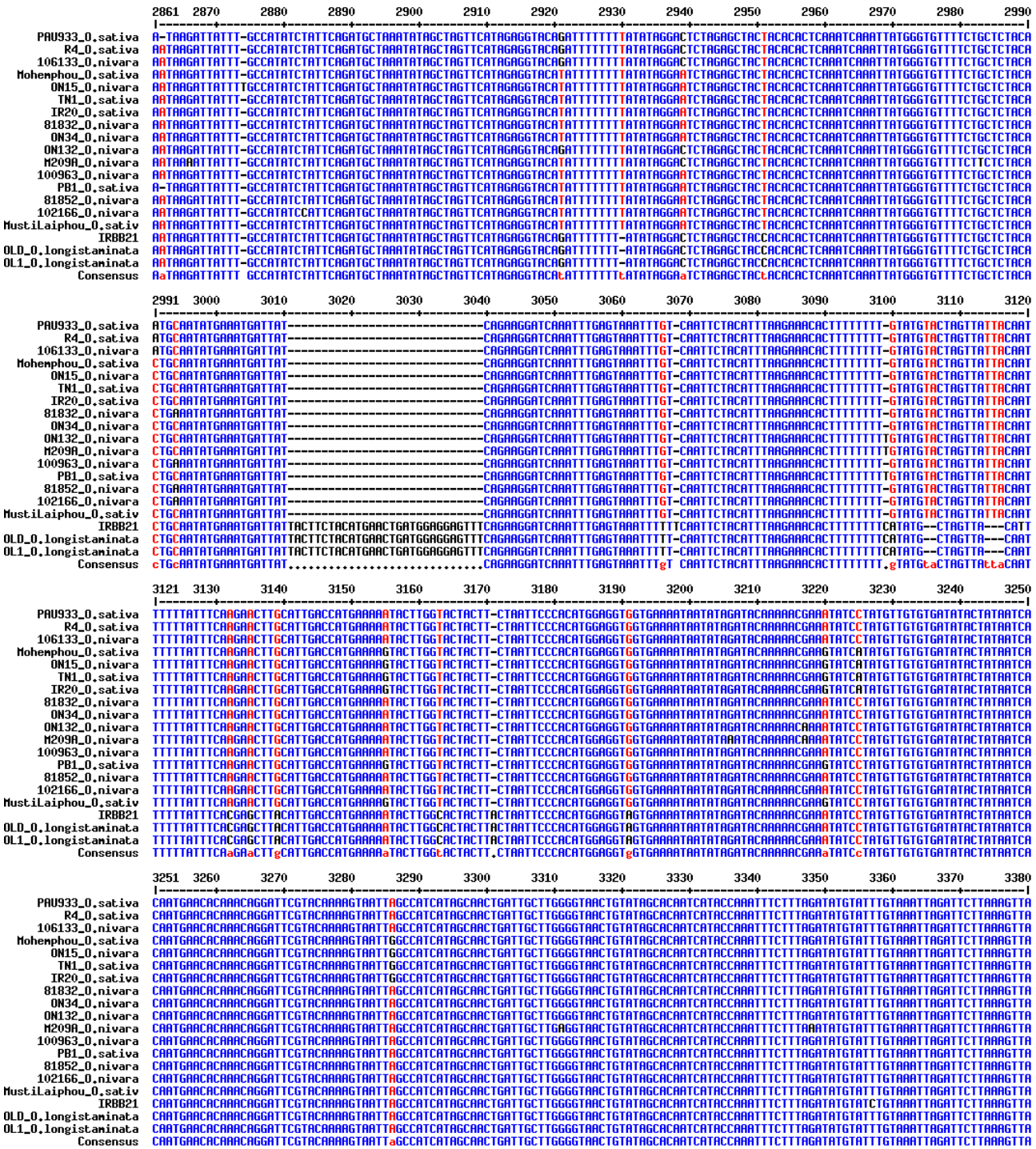


Fig. 4.2.1: Multiple sequence alignment of *Xa21* alleles showing substitutions and InDels in different regions. Deletions are shown in dotted lines while red and black color nucleotides indicate substitution. The rice genotypes are indicated in left column. The numbers on the top of the sequences indicate the position of nucleotides. Sequence similarity is observed among three alleles of *Xa21* (IRBB21, OLD and OL1).

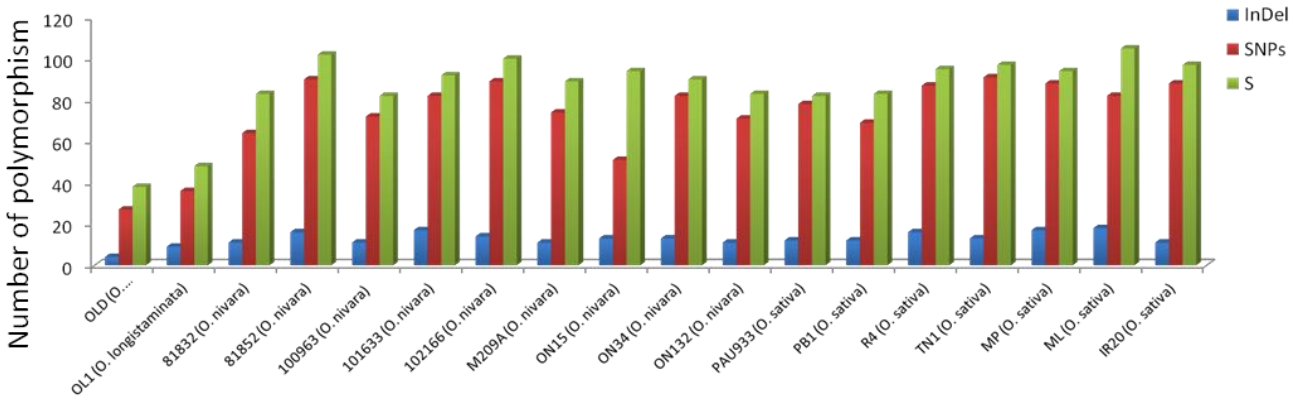


Fig.4.2.2: This graph depicts total number of different types of polymorphisms found at *Xa21* locus among different accessions. Blue, red and green colors indicate InDels, SNPs and total number of polymorphic sites, respectively. Different accessions are shown in X-axis and number of polymorphisms on Y- axis.

Seq 1	Seq 2	SilentDif	SilentPos	Ks	NSynDif	NSynPos	Ka
IRBB3(O_sa	IC203740(O	14.00	759.50	0.0187	36.00	2432.50	0.0149
IRBB3(O_sa	IRGC81832(	12.00	759.58	0.0160	40.00	2432.42	0.0166
IRBB3(O_sa	IRGC106133	11.00	759.58	0.0146	41.00	2432.42	0.0170
IRBB3(O_sa	Zhachanglo	0.00	762.00	0.0000	0.00	2430.00	0.0000
IRBB3(O_sa	Minghui63(	0.00	762.00	0.0000	0.00	2430.00	0.0000
IRBB3(O_sa	TN1(O_sati	7.00	762.08	0.0092	2.00	2429.92	0.0008
IRBB3(O_sa	ON132(Oryz	3.00	761.67	0.0039	6.00	2430.33	0.0025
IRBB3(O_sa	IRGC210223	2.00	762.00	0.0026	7.00	2430.00	0.0029
IRBB3(O_sa	MustiLaiph	7.00	760.83	0.0093	19.00	2431.17	0.0079
IRBB3(O_sa	M209AOryza	10.00	763.50	0.0132	16.00	2428.50	0.0066
IRBB3(O_sa	IRGC81852(	12.00	760.50	0.0159	19.00	2431.50	0.0079
IRBB3(O_sa	PusaBasmat	17.00	762.58	0.0226	39.00	2429.42	0.0162
IRBB3(O_sa	MohemPhoul	7.50	759.83	0.0099	20.50	2432.17	0.0085
IRBB3(O_sa	ON38-1(Ory	14.00	760.42	0.0186	30.00	2431.58	0.0124
IRBB3(O_sa	IC384116Or	12.00	761.42	0.0159	22.00	2430.58	0.0091
IRBB3(O_sa	R47(O_sati	23.00	761.42	0.0308	39.00	2430.58	0.0162
IRBB3(O_sa	IR24(O_sat	33.50	759.58	0.0455	61.50	2432.42	0.0257

Fig. 4.2.3: DnaSP output for estimation of rate of synonymous and non synonymous mutation. K_a is rate of non synonymous changes and K_s is rate of synonymous changes. K_a and K_s values are estimated comparing Seq1 and Seq2. Seq1 is single allele IRBB3 which is compared to Seq2. Seq2 has lists of different alleles.

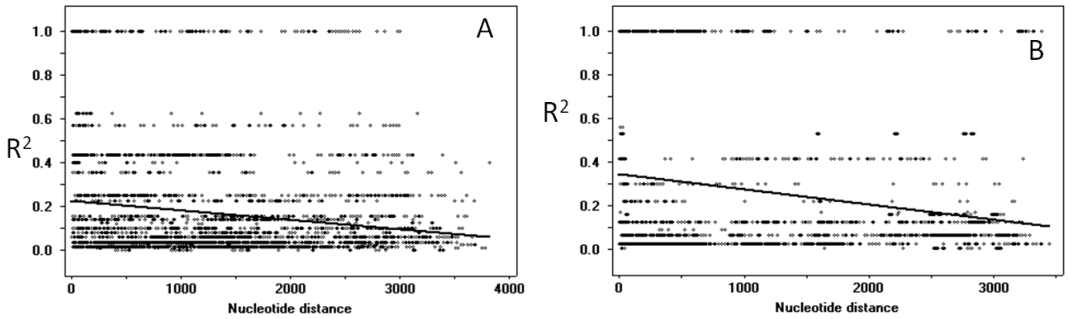


Fig. 4.2.4: Pattern of Linkage disequilibrium among *Xa21* alleles (A) *O. nivara* (B) *O. sativa.* Decay of LD 'R^2' as a function of distance between pairs of polymorphic sites in *Xa21* alleles. The Black line depicts the expected decline of LD against distance based on the equation given by HILL and WEIR (1988).

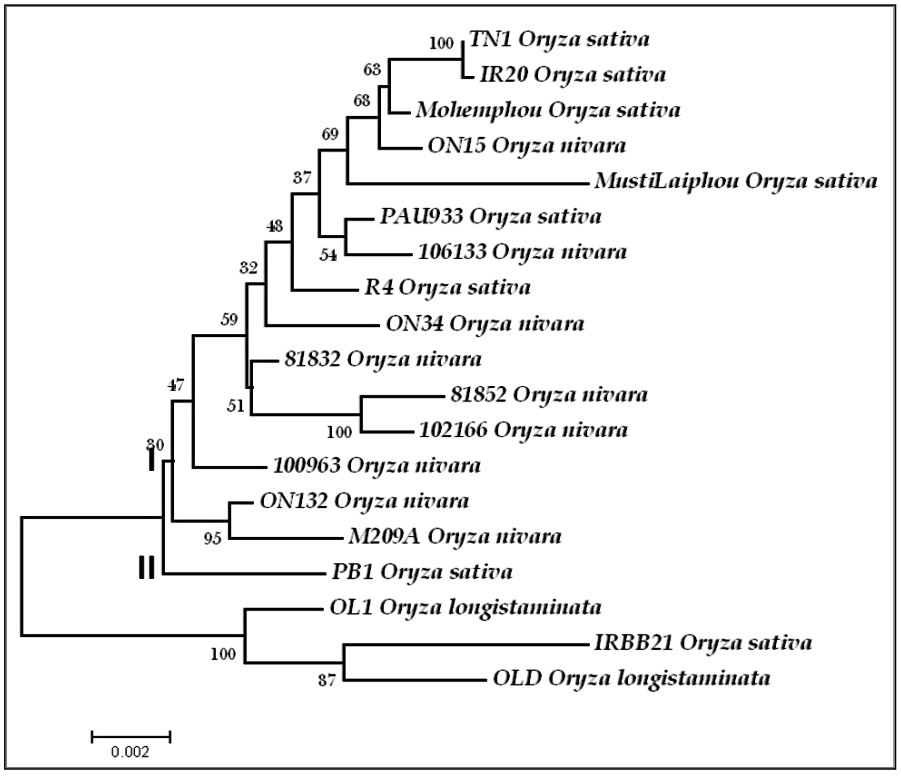


Fig. 4.2.5: The evolutionary relationship of *Xa21* alleles was inferred using the Neighbor-Joining method. The bootstrap consensus tree which is inferred from 10000 replicates is taken to represent the evolutionary history of the alleles analyzed. The tree is drawn to scale, with branch lengths. There are two major clades, clade I and II. The IRBB21 related alleles were clustered together in clade II. Each allele is indicated with the accession name followed by species name. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. IRBB21																		
2. ON132 O.n	0.022																	
3. M209A O.n	0.023	0.003																
4. 100963 O.n	0.022	0.004	0.006															
5. PB1 O.sativ	0.022	0.007	0.008	0.007														
6. R4 O.sativa	0.022	0.006	0.009	0.006	0.010													
7. ON34 O.niv	0.023	0.007	0.009	0.007	0.010	0.005												
8. PAU933 O.	0.021	0.006	0.009	0.007	0.007	0.003	0.007											
9. TN1 O.satir	0.025	0.009	0.012	0.009	0.009	0.006	0.008	0.006										
10. IR20 O.sa	0.025	0.010	0.012	0.009	0.009	0.007	0.008	0.006	0.000									
11. ON15 O.n	0.024	0.009	0.011	0.008	0.007	0.006	0.006	0.005	0.003	0.003								
12. 106133 O.	0.023	0.008	0.011	0.009	0.009	0.004	0.006	0.002	0.005	0.006	0.003							
13. Mohemph	0.023	0.007	0.010	0.007	0.007	0.006	0.007	0.004	0.002	0.003	0.002	0.005						
14. MustiLaip	0.028	0.013	0.015	0.012	0.012	0.010	0.011	0.009	0.009	0.009	0.008	0.010	0.007					
15. 81852 O.n	0.026	0.010	0.012	0.007	0.013	0.008	0.008	0.009	0.010	0.011	0.010	0.010	0.010	0.013				
16. 102166 O.	0.026	0.009	0.011	0.006	0.013	0.007	0.007	0.008	0.010	0.010	0.009	0.009	0.009	0.013	0.003			
17. 81832 O.n	0.022	0.005	0.007	0.003	0.008	0.004	0.004	0.004	0.006	0.007	0.006	0.006	0.005	0.009	0.006	0.005		
18. OLD O.lor	0.010	0.018	0.020	0.018	0.020	0.020	0.022	0.019	0.023	0.023	0.022	0.021	0.021	0.026	0.024	0.024	0.019	
19. OL1 O.lon	0.013	0.014	0.016	0.014	0.016	0.016	0.017	0.015	0.018	0.019	0.017	0.017	0.016	0.021	0.020	0.019	0.015	0.006

Fig. 4.2.6: Estimates of base composition bias difference between sequences of *Xa21* alleles. It depicts the divergence plot among different alleles of *Xa21*. The difference in base composition bias per site is shown in each column. Left column indicate the alleles and the numbering columns indicate the divergence rate corresponding to the numbered rows.

Fig. 4.2.7: Phenotypic response of the genotypes after *Xoo* infection. Leaves with lesion length more than 5 cm were susceptible. IR24 and IRBB21 was used as susceptible and resistance control, respectively. Control (C), Treated/Infected (T). Lesion length in IR24T was more than 10 cm.

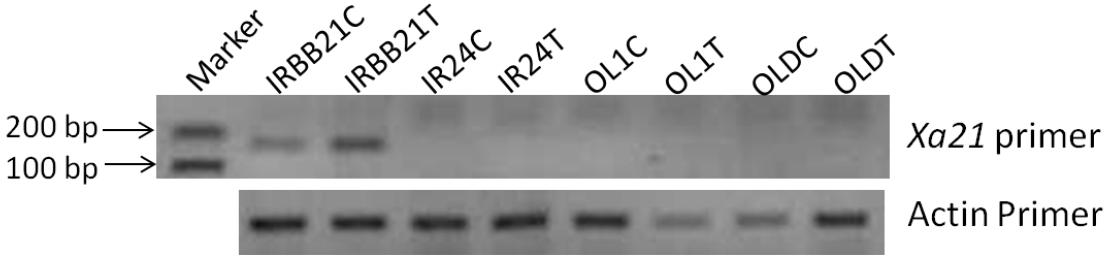


Fig. 4.2.8: Semi-quantitative PCR analysis of *Xa21* alleles in control and treated samples of IR24, IRBB21, OL1 and OLD. Primer X21RT was used for expression analysis. Band size between 100-200 bp is observed in IRBB21 control and infected with *Xoo* showing expression of *Xa21* gene, while *Xa21* expression is absent in remaining samples. Expression pattern of internal control in the samples is shown with actin primer.

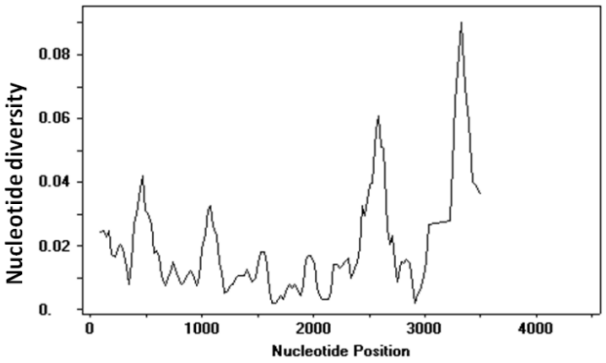


Fig. 4.2.9: Nucleotide diversity graph of *Xa26* alleles showing nucleotide positions in base pairs on X-axis and nucleotide diversity, Pi on Y-axis. Nucleotide diversity is a measure of DNA sequence variation based on the average pairwise distance between all sequences in the sample.

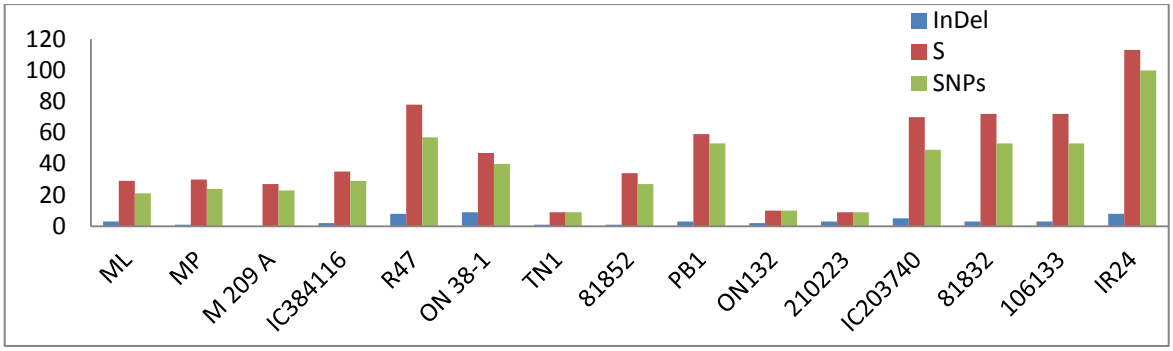


Fig. 4.2.10: Graph showing different forms of polymorphism in each accession in comparison to IRBB3 (*Xa26*) reference sequence 3709 bp. InDel – Insertion/Deletion; SNPs – Single Nucleotide polymorphism; S – Total no. of polymorphic of segregating sites.

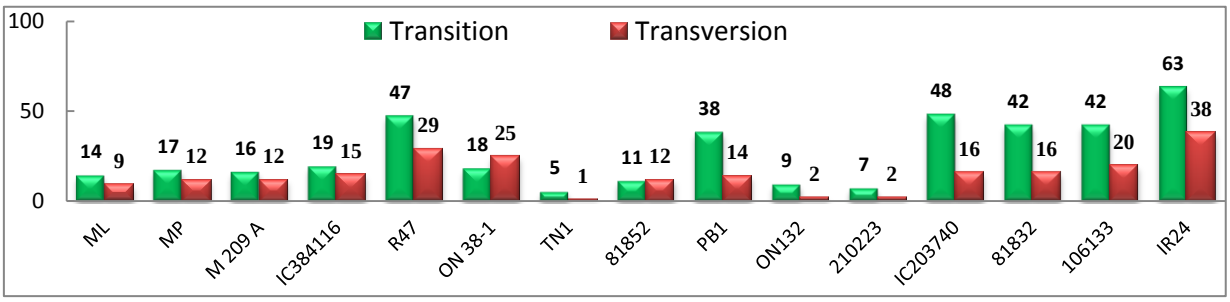


Fig. 4.2.11: Graph showing levels of transition and transversion substitution in each allele of *Xa26* (3709 bp).

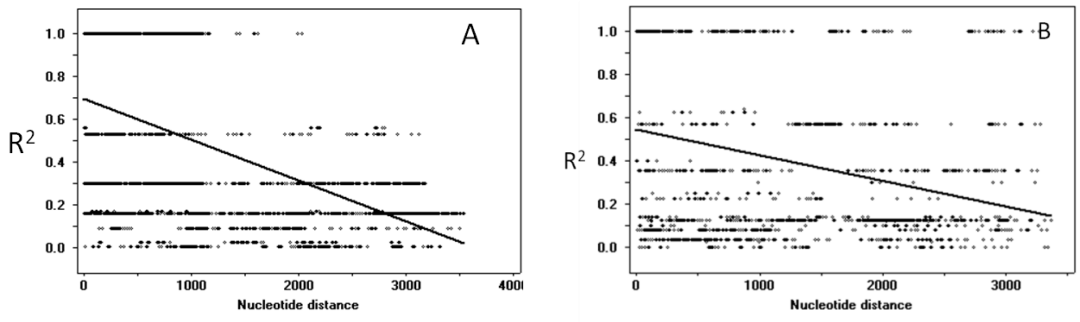


Fig. 4.2.12: Pattern of Linkage disequilibrium among *Xa26* alleles (A) *O. nivara* (B) *O. sativa.* Decay of LD 'R^2' as a function of distance between pairs of polymorphic sites in *Xa26* alleles.

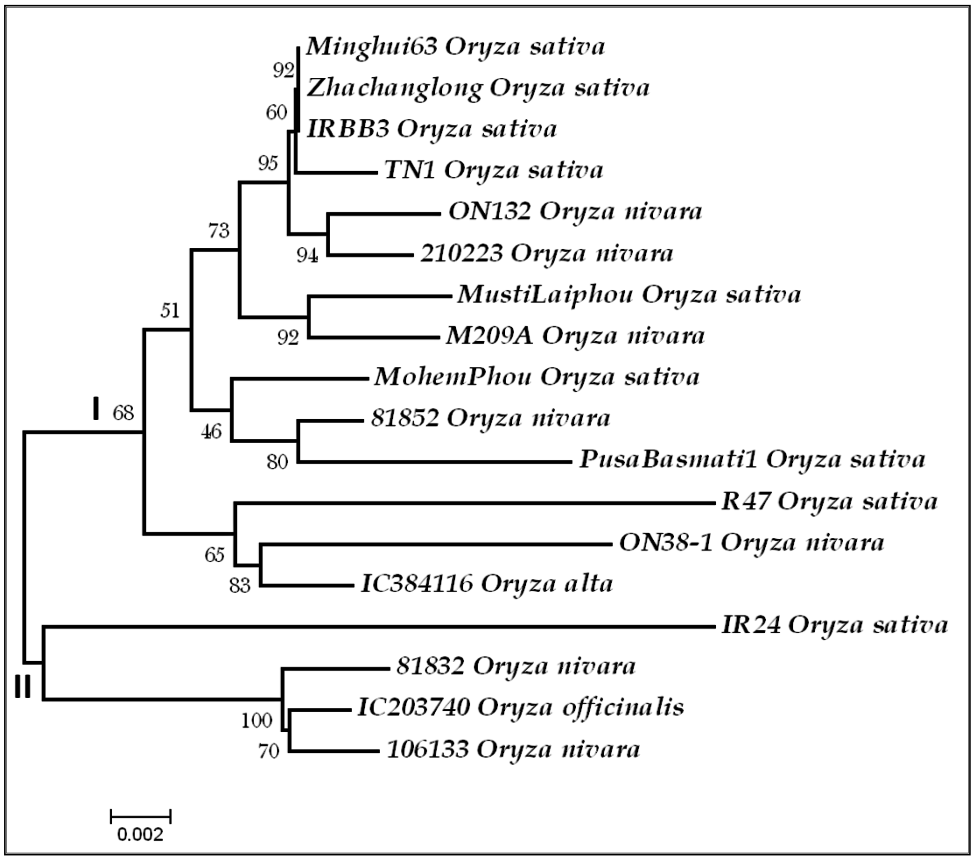


Fig. 4.2.13: Neighbour joining gene tree depicting the relationship of all studied individuals at *Xa26* locus. The tree was inferred from 10000 replicates bootstrap values to represent the evolutionary history of the alleles. The highest bootstrap value of each branch is indicated correspondingly. The tree is divided into two clades I and II. The scale of branch length is indicated at the bottom. Each allele is indicated with the accession name followed by species name. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007).

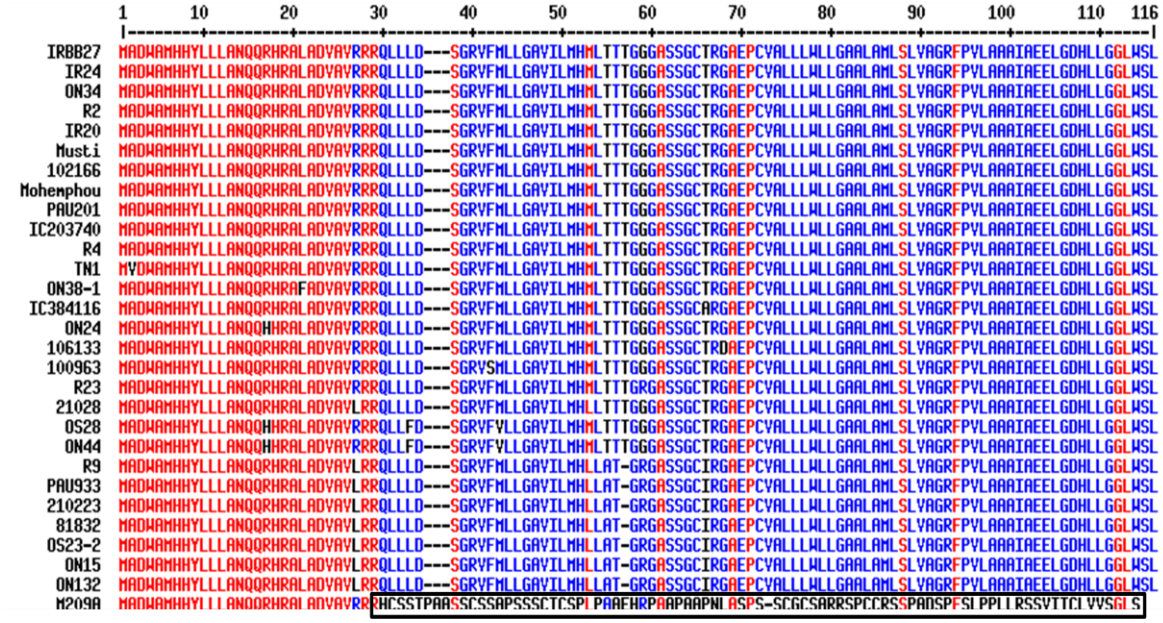


Fig. 4.2.14: Output of multiple alignment for the predicted amino acid sequences of *Xa27* alleles. The rice genotypes are indicated in left column. Alignment was performed using Multalin program. The numbers on the top of the sequences indicate the position of amino acids. Insertions and deletions are shown with gaps. Different colors in amino acid shows non synonymous changes. Frameshift mutation in M209A is also shown in the bottom.

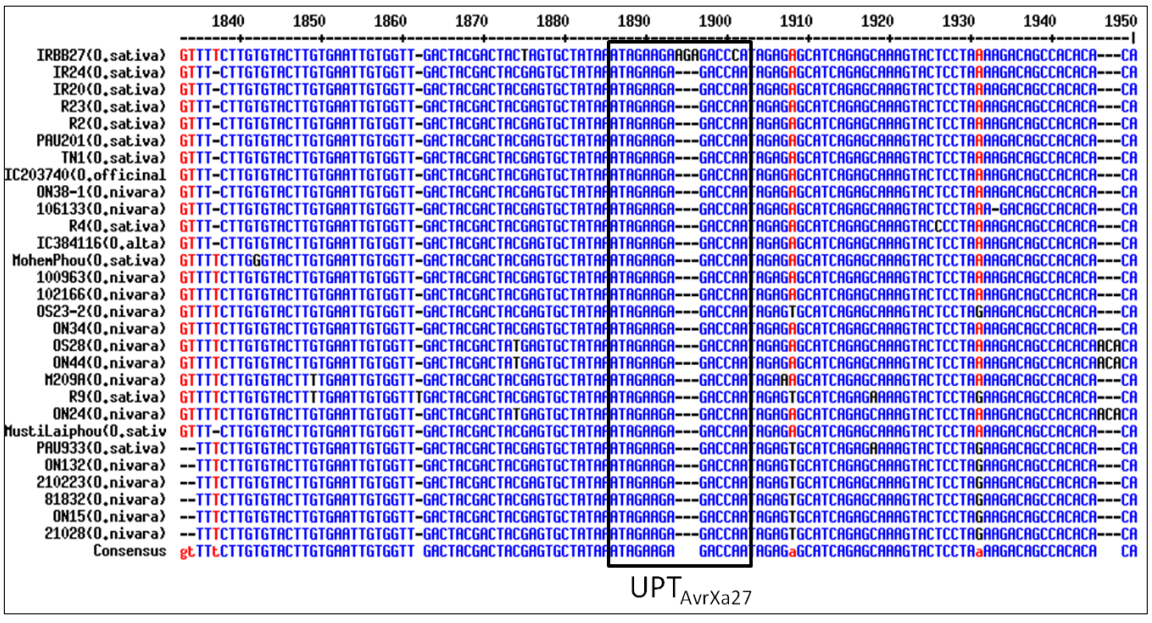


Fig. 4.2.15: Multiple alignment of the promoter region of *Xa27* covering the predicted $UPT_{AvrXa27}$. The rice genotypes are indicated in left column. Multiple alignment was constructed in Multalin. Deletion of three bp 'AGA' and substitution of 'C' by 'A' in the UPT box is indicated by black bar. It is at -51 position of the gene (from TSS).

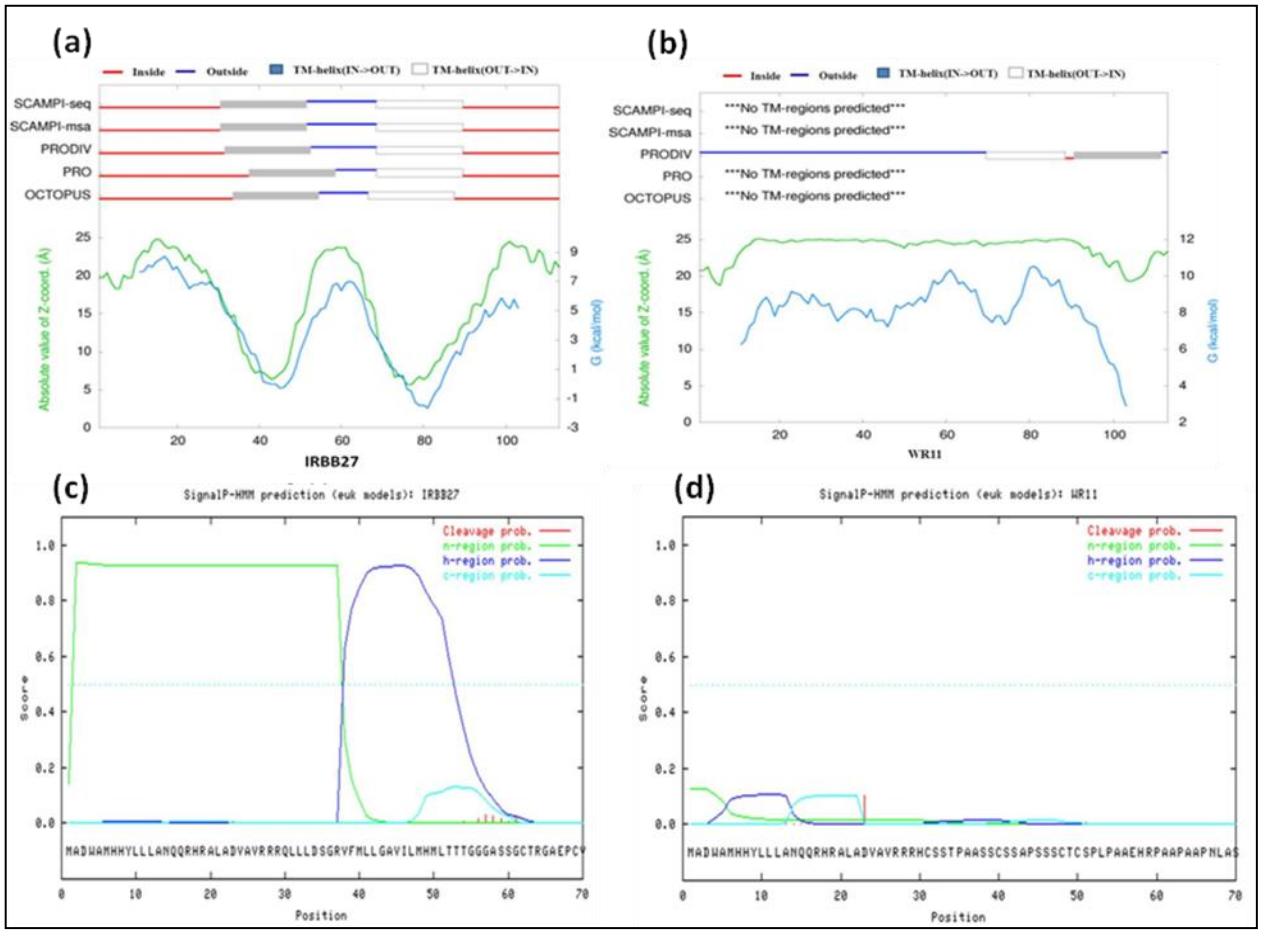


Fig. 4.2.16: TOPCONS and SignalP prediction for presence of trans-membrane (TM) domain and N- terminal signal peptide. Output of TOPCONS prediction (a) IRBB27 and (b) M209A. Depression in peaks at two regions around 32 to 52 residues and 69 to 89 residue (X-axis) in IRBB27 shows the presence of two TM domain whereas it is absent in M209A. Position of predicted TM is indicated by white and grey bar. SignalP output (c) IRBB27 and (d) M209A. Presence of N - terminal signal peptide is clearly indicated with green and blue color peaks in IRBB27 while it is absent in M209A.

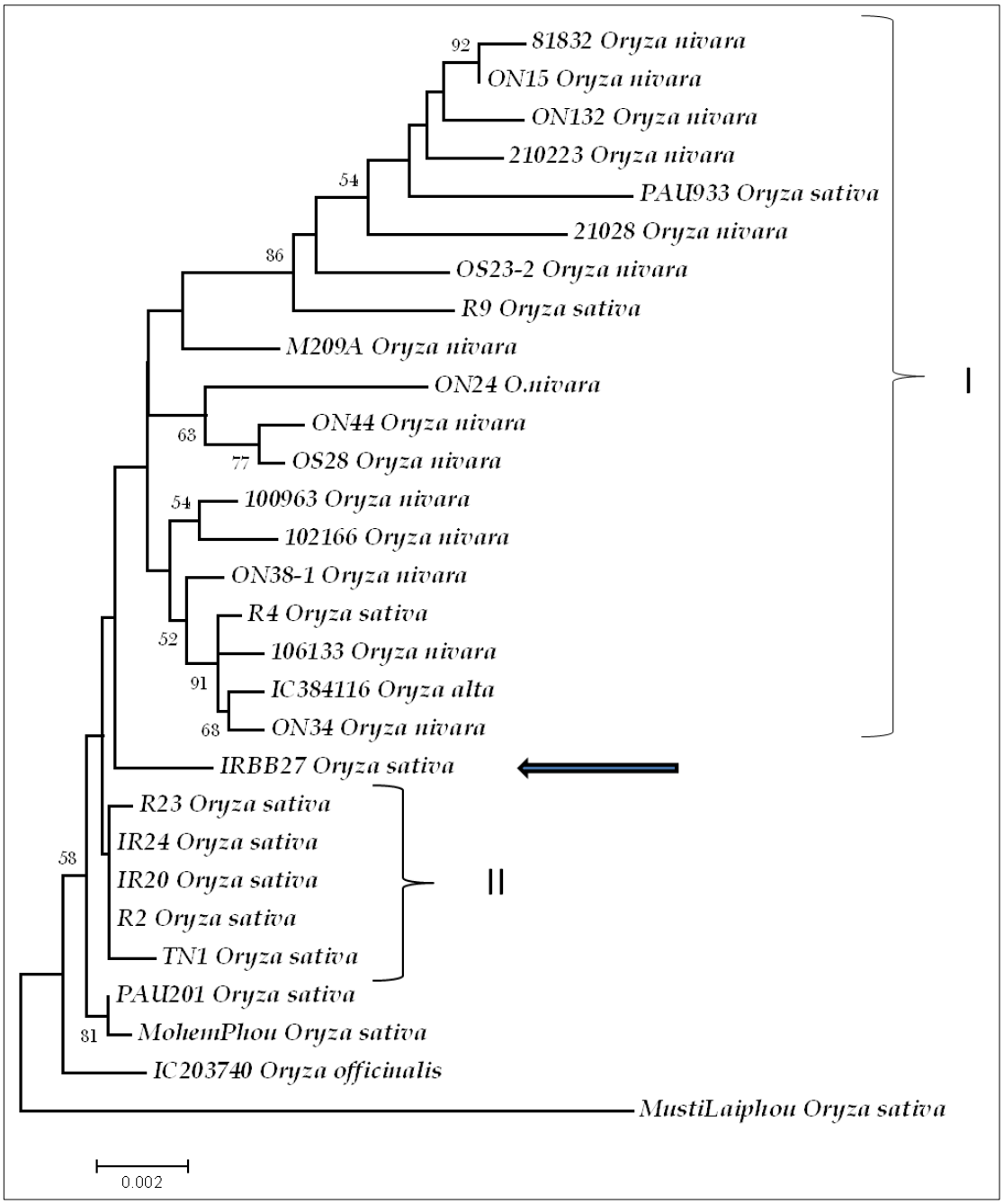
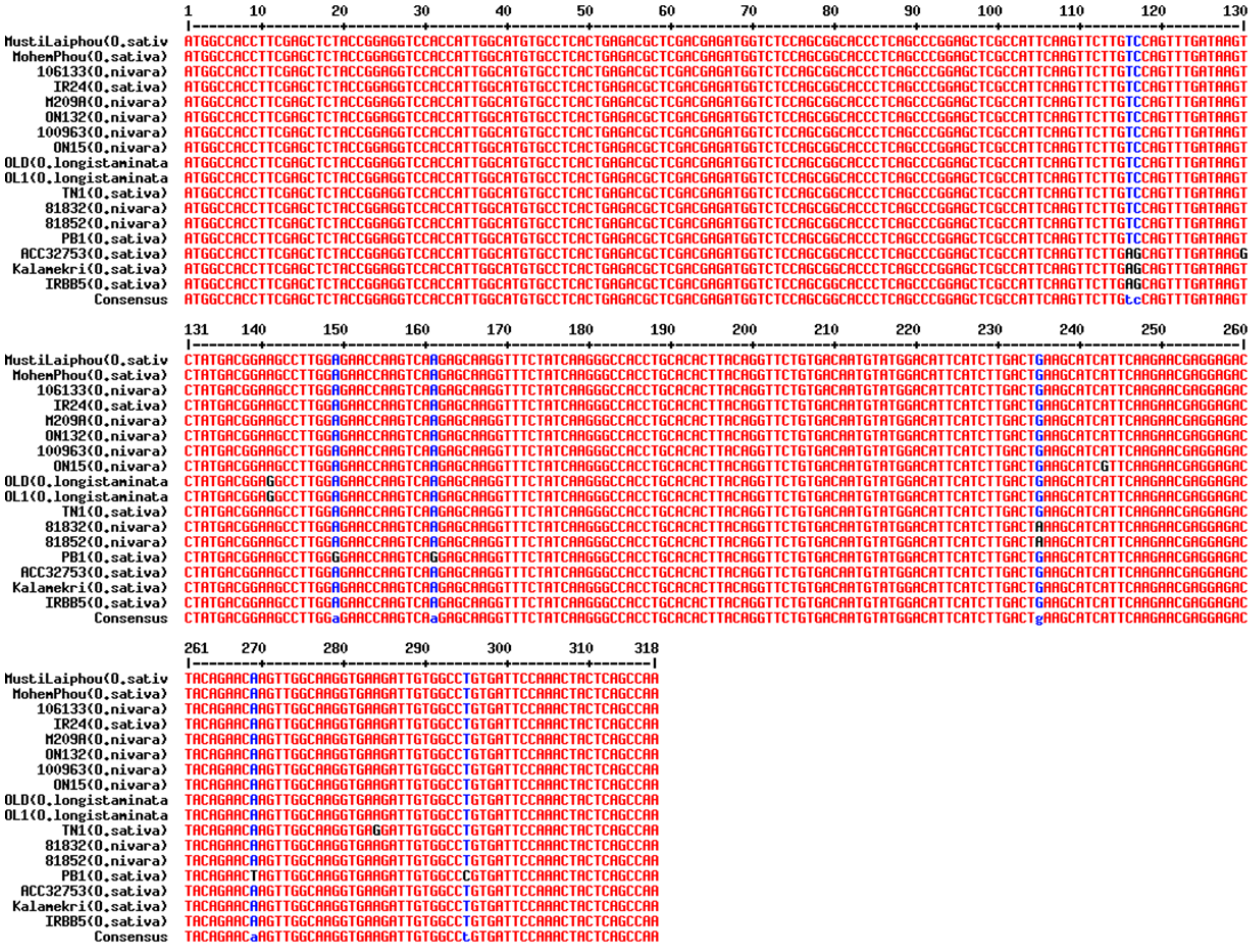
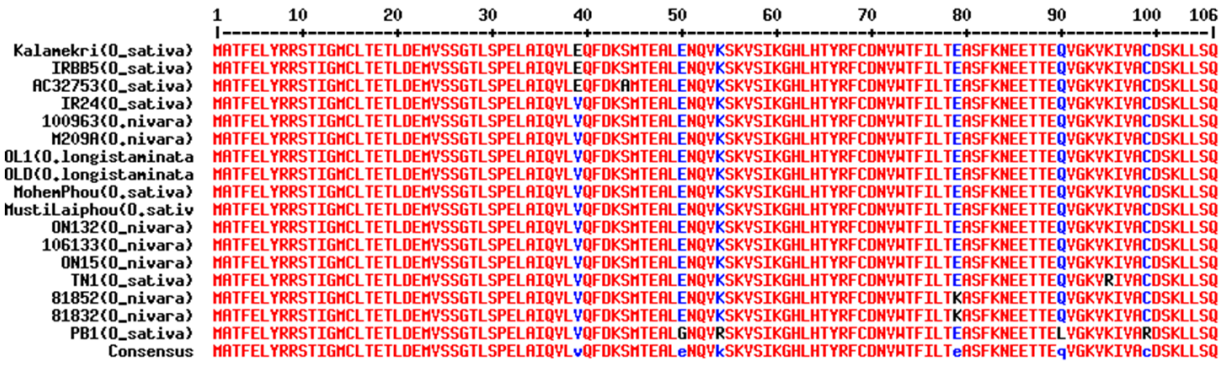


Fig. 4.2.17: The evolutionary relationship of *Xa27* alleles was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.07944762 is shown. The bootstrap consensus tree is inferred from 10000 replicates. The percentage of replicate trees in which the alleles clustered together in the bootstrap test (10000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007). Group I alleles are closer to wild relatives and group II alleles are related to IR24. IRBB27 indicated by arrow lies between these two groups.



A



B

Fig. 4.2.18: (A) Multiple alignment of coding region of *xa5* alleles. Substitutions in the CDS are represented with different colors. (B) Output of multiple alignment for the predicted amino acid sequences of *xa5* alleles. The rice genotypes are indicated in left column. The numbers on the top of the sequences indicate the position of amino acids. Different colors in amino acid shows non synonymous changes. Alignment was performed using Multalin program.

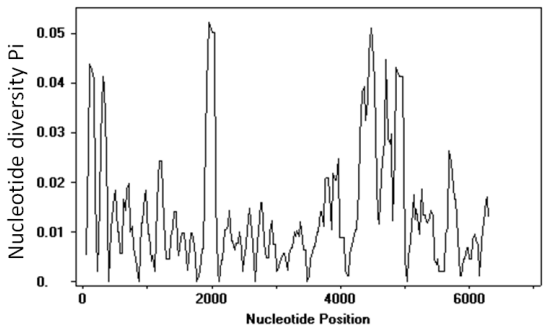


Fig. 4.2.19: Nucleotide diversity graph of *xa5* alleles showing nucleotide positions in base pairs on X-axis and nucleotide diversity, Pi on Y-axis. Nucleotide diversity is a measure of DNA sequence variation based on the average pairwise distance between all sequences in the sample.

Seq 1	Seq 2	SynDif	SynPos	Ks	NSynDif	NSynPos	Ka
IRBB5(O.sa	PB1(O.sati	0.50	71.08	0.0071	5.50	246.92	0.0226
IRBB5(O.sa	ACC32753(O	0.00	69.50	0.0000	1.00	248.50	0.0040
IRBB5(O.sa	Kalamekri[	0.00	69.50	0.0000	0.00	248.50	0.0000
IRBB5(O.sa	ON15(O.niv	1.50	69.83	0.0218	1.50	248.17	0.0061
IRBB5(O.sa	TN1(O.sati	0.50	70.00	0.0072	2.50	248.00	0.0101
IRBB5(O.sa	81832(O.ni	0.50	69.83	0.0072	2.50	248.17	0.0101
IRBB5(O.sa	81852(O.ni	0.50	69.83	0.0072	2.50	248.17	0.0101
IRBB5(O.sa	MustiLaiph	0.50	69.83	0.0072	1.50	248.17	0.0061
IRBB5(O.sa	MohemPhoul	0.50	69.83	0.0072	1.50	248.17	0.0061
IRBB5(O.sa	M209A(O.ni	0.50	69.83	0.0072	1.50	248.17	0.0061
IRBB5(O.sa	100963(O.n	0.50	69.83	0.0072	1.50	248.17	0.0061
IRBB5(O.sa	ON132(O.ni	0.50	69.83	0.0072	1.50	248.17	0.0061
IRBB5(O.sa	IR24(O.sat	0.50	69.83	0.0072	1.50	248.17	0.0061
IRBB5(O.sa	106133(O.n	0.50	69.83	0.0072	1.50	248.17	0.0061
IRBB5(O.sa	OLD(O.long	1.50	69.83	0.0218	1.50	248.17	0.0061
IRBB5(O.sa	OL1(O.long	1.50	69.83	0.0218	1.50	248.17	0.0061

Fig. 4.2.20: DnaSP output for estimation of rate of synonymous and non synonymous mutation of *xa5* alleles. K_a is rate of non synonymous changes and K_s is rate of synonymous changes. K_a and K_s values are estimated comparing Seq1 and Seq2. Seq1 is single allele IRBB3 which is compared to Seq2. Seq2 has lists of different alleles.

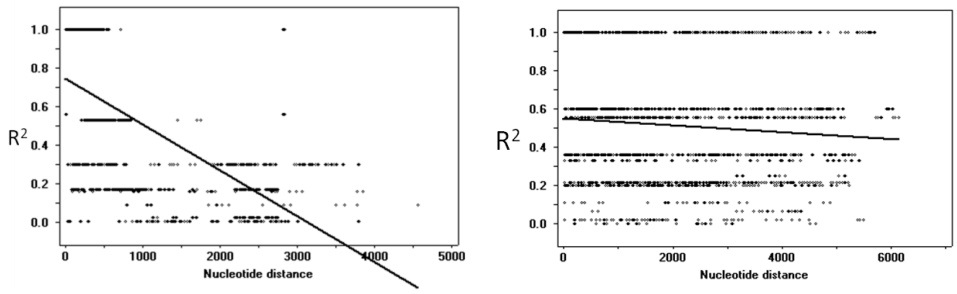


Fig. 4.2.21: Pattern of Linkage disequilibrium among *xa5* alleles (A) *O. nivara* (B) *O. sativa.* Decay of LD 'R^2' as a function of distance between pairs of polymorphic sites in *Xa26* alleles.

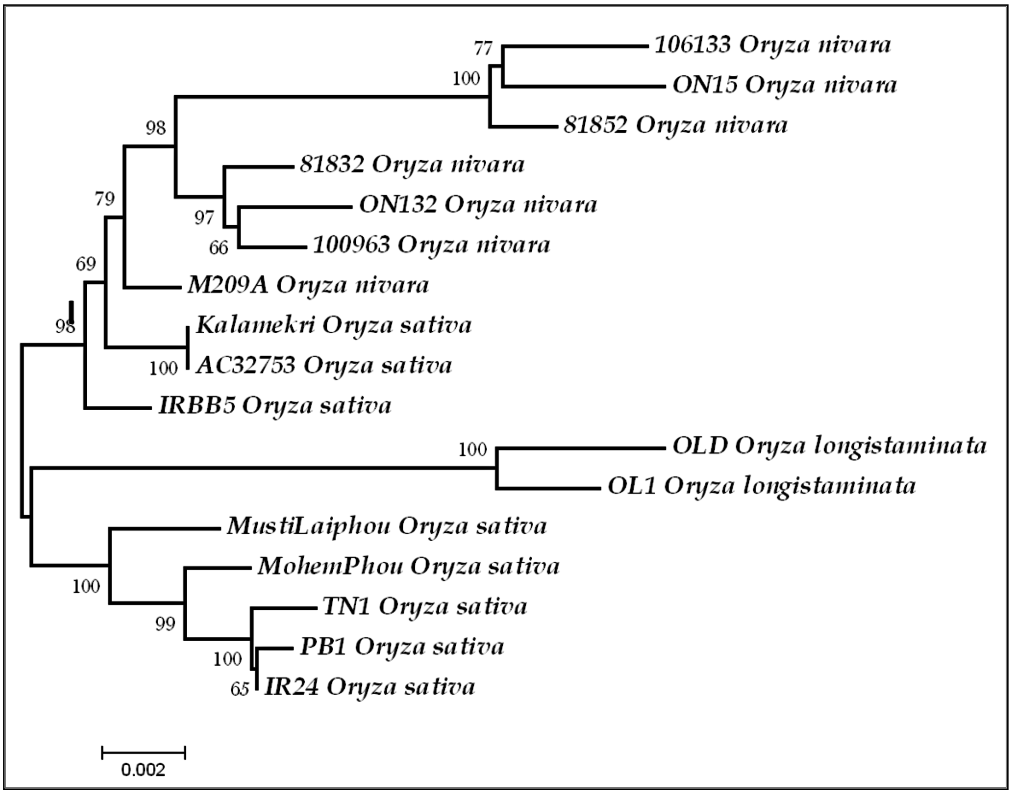


Fig. 4.2.22: Neighbour joining gene tree of all studied individuals at *xa5* locus. The tree was constructed with 10,000 bootstrap values. The highest bootstrap value of each branch is indicated correspondingly. The scale of branch length is indicated at the bottom. Each allele is indicated with the accession name followed by species name. The alleles are clustered separately at the species level, while the *xa5* alleles (resistant and recessive form) IRBB5, AC32753 and Kalamekri show more relatedness to *O. nivara*.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. 106133(O.nivara)																	
2. ON132(O.nivara)	0.012																
3. 100963(O.nivara)	0.011	0.002															
4. 81832(O.nivara)	0.008	0.001	0.004														
5. OLD(O.longistaminat	0.020	0.013	0.021	0.008													
6. 81852(O.nivara)	0.009	0.010	0.005	0.013	0.033												
7. ON15(O.nivara)	0.004	0.011	0.007	0.011	0.026	0.002											
8. MohemPhou(O.sativa	0.007	0.002	0.004	0.000	0.007	0.011	0.009										
9. PB1(O.sativa)	0.004	0.004	0.007	0.002	0.011	0.013	0.009	0.002									
10. IR24(O.sativa)	0.006	0.001	0.003	0.000	0.010	0.009	0.007	0.000	0.002								
11. TN1(O.sativa)	0.009	0.011	0.018	0.005	0.006	0.027	0.019	0.006	0.003	0.007							
12. OL1(O.longistamina	0.014	0.009	0.015	0.005	0.001	0.024	0.017	0.004	0.007	0.006	0.005						
13. MustiLaiphou(O.sati	0.023	0.002	0.006	0.005	0.015	0.020	0.023	0.006	0.010	0.006	0.016	0.013					
14. M209A(O.nivara)	0.006	0.002	0.001	0.003	0.017	0.004	0.005	0.002	0.004	0.001	0.012	0.011	0.008				
15. Kalamekri(O.sativa)	0.014	0.001	0.002	0.003	0.011	0.010	0.011	0.002	0.007	0.002	0.014	0.007	0.004	0.003			
16. AC32753(O.sativa)	0.014	0.001	0.002	0.003	0.011	0.010	0.011	0.002	0.007	0.002	0.014	0.007	0.004	0.003	0.000		
17. IRBB5(O.sativa)	0.001	0.014	0.015	0.009	0.019	0.015	0.007	0.008	0.004	0.008	0.006	0.014	0.025	0.009	0.017	0.017	

Fig. 4.2.23: Divergence plot among different alleles of *xa5*. Left column indicate the alleles and the numbering columns indicate the divergence rate corresponding to the numbered rows.

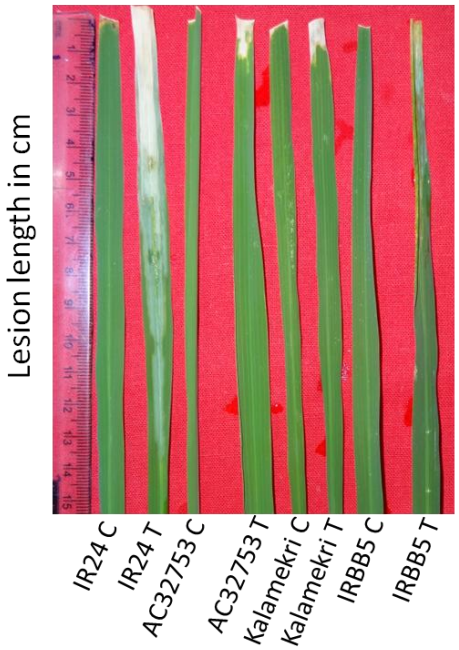


Fig. 4.2.24: Lesion length of Rice samples used for allelic expression analysis of *xa5* gene. Leafs with lesion length more than 5 cm are susceptible. IR24 was used as susceptible control. Control (C) ; Treated/Infected (T).

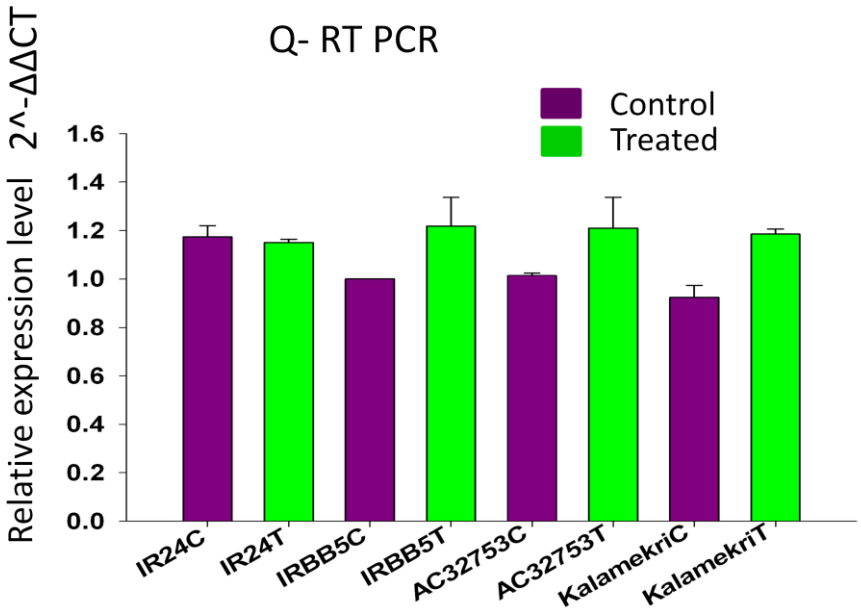


Fig. 4.2.25: This graph depicts the expression pattern of *xa5* alleles in control and treated samples of IR24, IRBB5, AC32753 and Kalamekri. 'C' indicates control and 'T' means treated. Fold change in expression is expressed in terms of $2^{-\Delta\Delta CT}$. The relative expression was almost at the basal level in all samples.

Fig. 4.3.1: Phenotypic response of the genotypes after *Xoo* infection. Leafs with lesion length more than 5 cm are susceptible. IR24 was used as susceptible control, respectively. Control (C), Treated/Infected (T)

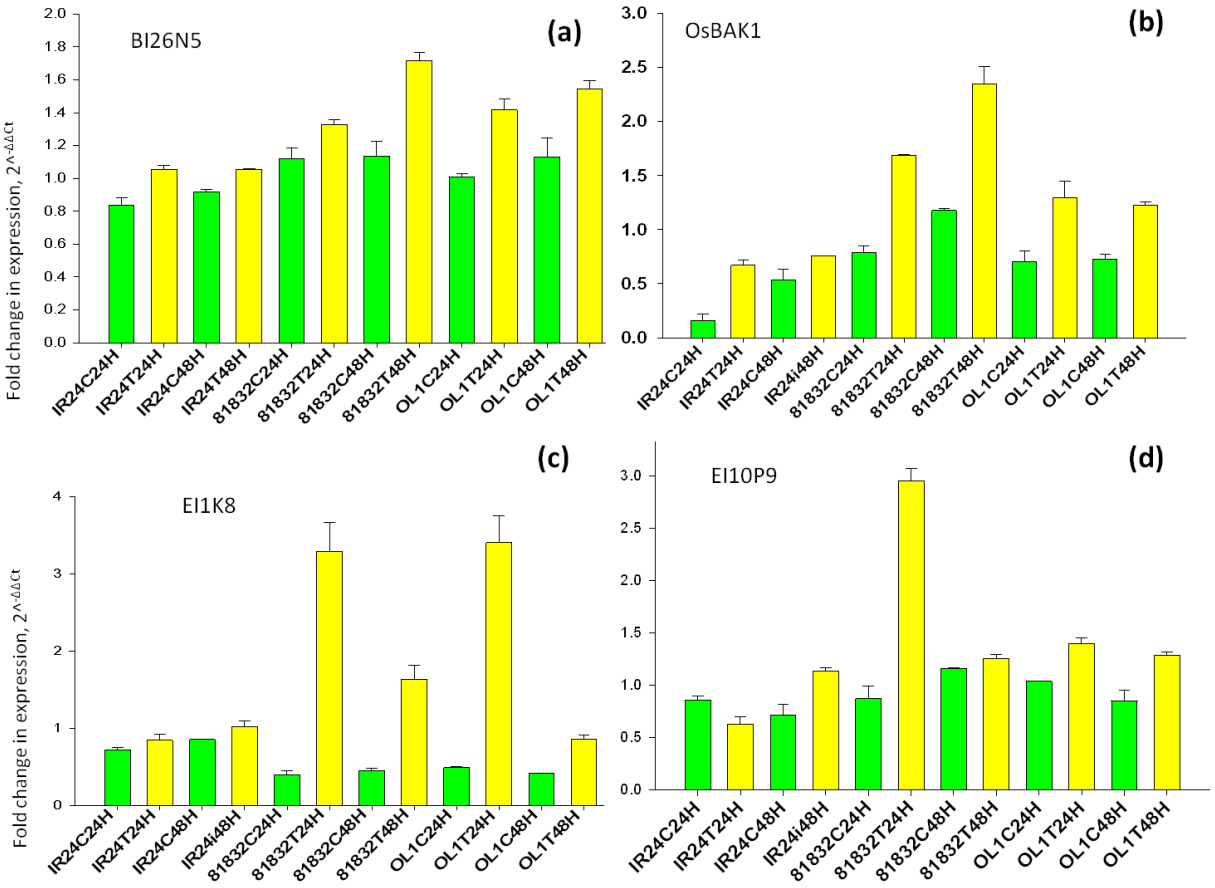


Fig. 4.3.2: Expression analysis of different genes (a) *BI26N5*, (b) *OsBAK1*, (c) *EI1K8* and (d) *EI10P9* by Q-RT PCR. X-axis represents the name of the samples and the level of expression in fold changes (relative change in expression, $2^{\wedge -\Delta\Delta Ct}$) is represented by Y-axis.

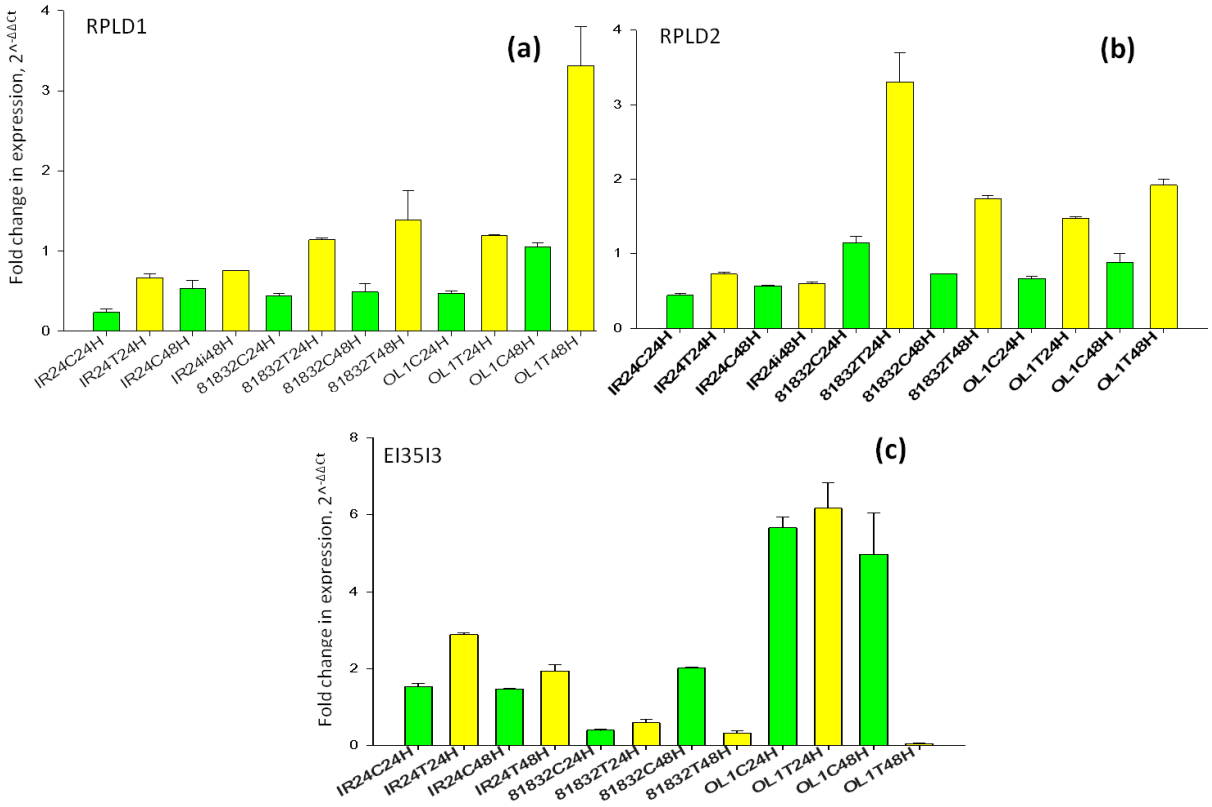


Fig. 4.3.3: Expression analysis of genes (a) *RPLD1*, (b) *RPLD2* and (c) *EI35I3*by Q-RT PCR. X-axis represents the name of the samples and the level of expression in fold changes (relative change in expression, $2^{\wedge -\Delta\Delta Ct}$) is represented by Y-axis.

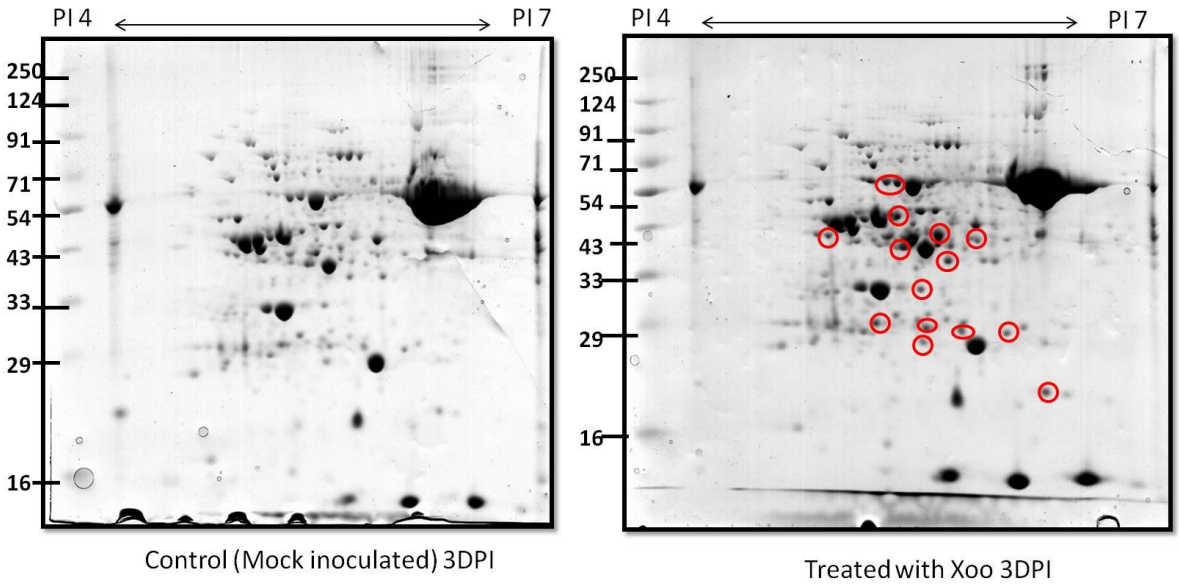
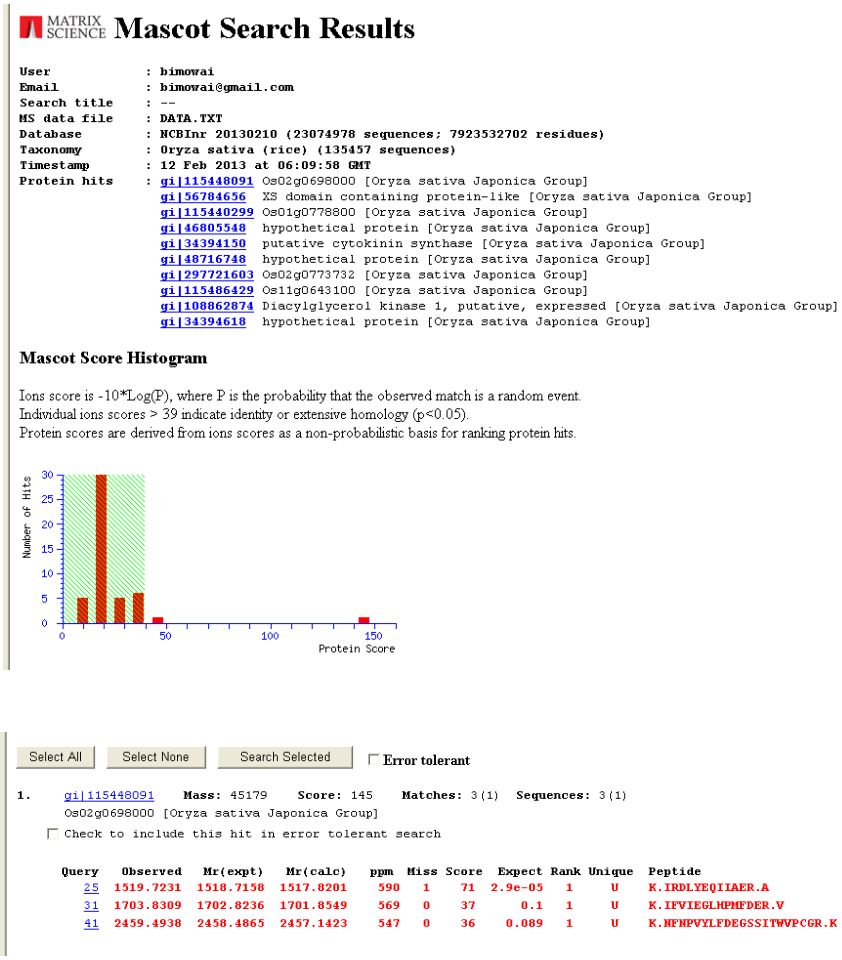


Fig. 4.3.4: 2DE of *Xoo* inoculated and mock treated (sterile water) of resistant wild rice *Oryza nivara* 81832. The protein samples were separated on IPG strips pH 4-7, 12% SDS PAGE. The size of the protein marker is labeled in the left side of the gel.



MATRIX SCIENCE **Mascot Search Results**

```
User            : bimowai
Email           : bimowai@gmail.com
Search title    : --
MS data file    : DATA.TXT
Database        : NCBInr 20130210 (23074978 sequences; 7923532702 residues)
Taxonomy        : Oryza sativa (rice) (135457 sequences)
Timestamp       : 12 Feb 2013 at 06:09:58 GMT
Protein hits    : gi|115448091 Os02g0698000 [Oryza sativa Japonica Group]
                  gi|56784656  XS domain containing protein-like [Oryza sativa Japonica Group]
                  gi|115440299 Os01g0778800 [Oryza sativa Japonica Group]
                  gi|46805548  hypothetical protein [Oryza sativa Japonica Group]
                  gi|34394150  putative cytokinin synthase [Oryza sativa Japonica Group]
                  gi|48716748  hypothetical protein [Oryza sativa Japonica Group]
                  gi|297721603 Os02g0773732 [Oryza sativa Japonica Group]
                  gi|115486429 Os11g0643100 [Oryza sativa Japonica Group]
                  gi|108862874 Diacylglycerol kinase 1, putative, expressed [Oryza sativa Japonica Group]
                  gi|34394618  hypothetical protein [Oryza sativa Japonica Group]
```

Mascot Score Histogram

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 39 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

Select All | Select None | Search Selected | Error tolerant

1. gi|115448091 **Mass:** 45179 **Score:** 145 **Matches:** 3(1) **Sequences:** 3(1)
   Os02g0698000 [Oryza sativa Japonica Group]
   Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
25	1519.7231	1518.7158	1517.8201	590	1	71	2.9e-05	1	U	K.IRDLYEQIIAER.A
31	1703.8309	1702.8236	1701.8549	569	0	37	0.1	1	U	K.IFVIEGLHPMFDER.V
41	2459.4938	2458.4865	2457.1423	547	0	36	0.089	1	U	K.MFNPVYLFDEGSSITWVPCGR.K

Search Parameters

```
Type of search          : MS/MS Ion Search
Enzyme                  : Trypsin
Fixed modifications     : Carbamidomethyl (C)
Variable modifications  : Oxidation (M)
Mass values             : Monoisotopic
Protein Mass            : Unrestricted
Peptide Mass Tolerance  : ± 650 ppm
Fragment Mass Tolerance: ± 1.5 Da
Max Missed Cleavages    : 1
Instrument type         : MALDI-TOF-TOF
Number of queries       : 48
```

Fig. 4.3.5: Schematic presentation of the output of the MS/MS ion search analysis using Mascot search in www.matrixscience.com. The first 10 protein hits in NBCI or SwissProt database is given on top with their accession numbers. Protein hits in red bar outside the green box with maximum protein score is considered. The number of peptide match with the peptide sequence is shown in red color. The parameters used for searching the proteins in databases are also indicated below the output.

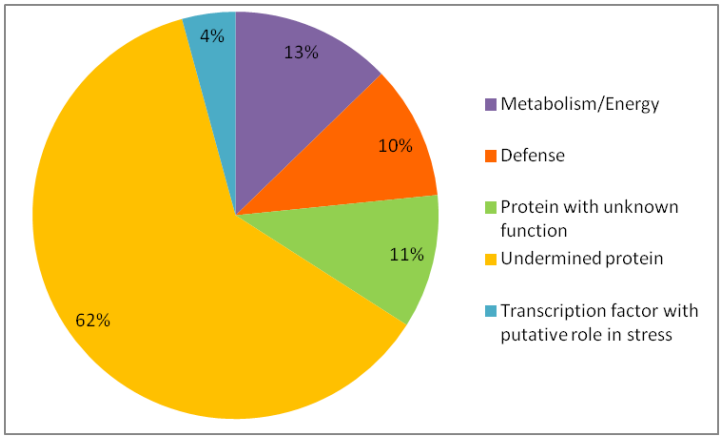


Fig. 4.3.6 (a): Percentagewise presentation of different protein detected in 2DE.

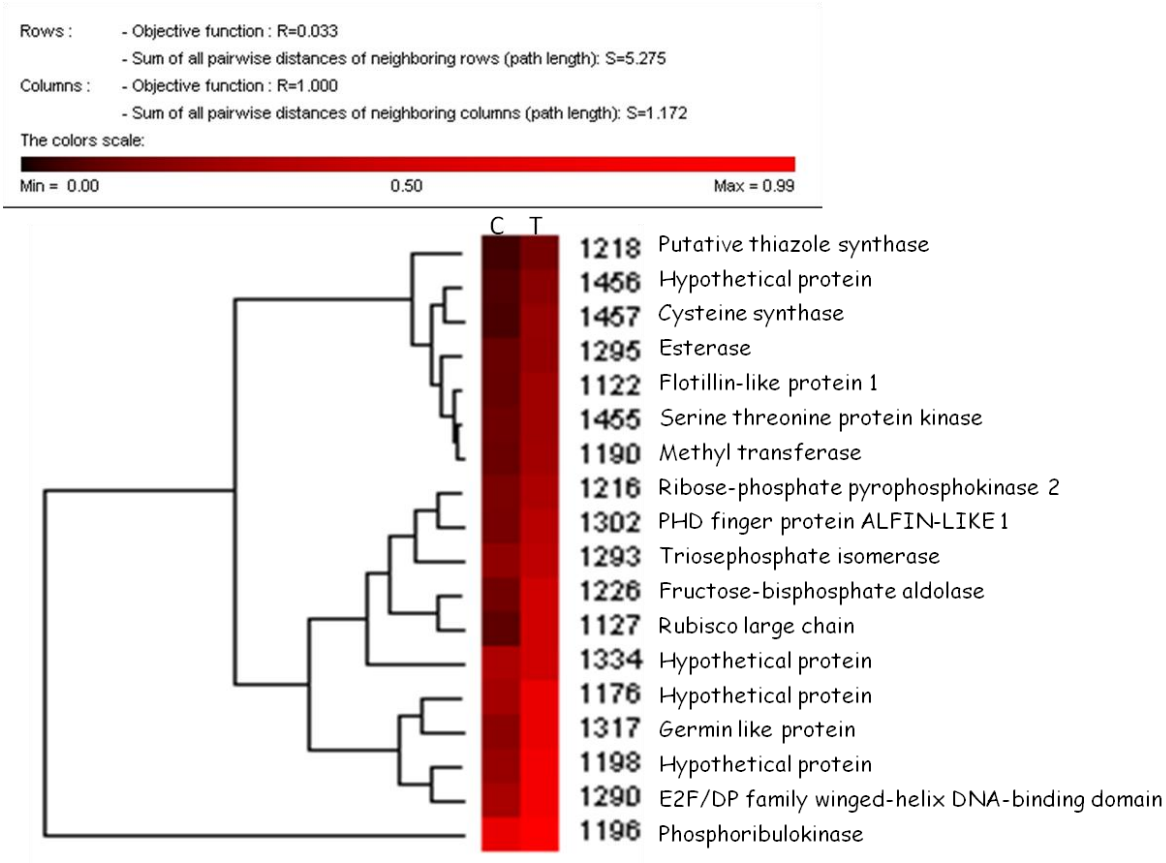


Fig. 4.3.6 (b): Heat map showing identified differentially expressed proteins after challenged inoculation with *Xoo* in *O. nivara* 81832 and mock inoculated plant. C- Control, T- Treated. The intensity of the spot increases from top to bottom. Difference in the level of expression can be clearly seen between control and treated.

Table 2.1 List of rice fungal diseases and their causal organisms (Source: Srinivasaprasad and Johnson, 2001)

Diseases	Causal Organism
Blast (leaf, neck [rotten neck], nodal and collar)	*Pyricularia grisea = Pyricularia oryzae* *Magnaporthe grisea* [teleomorph]
Brown spot	*Cochliobolus miyabeanus* *Bipolaris oryzae* [anamorph]
Crown sheath rot	*Gaeumannomyces graminis*
Downy mildew	*Sclerophthora macrospora*
Eyespot	*Drechslera gigantea*
False smut	*Ustilaginoidea virens*
Kernel smut	*Tilletia barclayana = Neovossia horrida*
Laf smut	*Entyloma oryzae*
Leaf scald	*Microdochium oryzae = Rhynchosporium oryzae*
Narrow brown leaf spot	*Cercospora janseana = Cercospora oryzae* *Sphaerulina oryzina* [teleomorph]
	Damage by many fungi including *Cochliobolus miyabeanus* *Curvularia* spp.
Pecky rice (kernel spotting)	*Fusarium spp.* *Microdochium oryzae* *Sarocladium oryzae* and other fungi.
Root rots	*Fusarium* spp. *Pythium* spp. *Pythium dissotocum* *Pythium spinosum*
	Cochliobolus miyabeanus *Curvularia* spp. *Fusarium* spp.
Seedling blight	*Rhizoctonia solani* *Sclerotium rolfsii* *Athelia rolfsii* [teleomorph] and other pathogenic fungi.
Sheath blight	*Thanatephorus cucumeris* *Rhizoctonia solani* [anamorph]
Sheath rot	*Sarocladium oryzae = Acrocylindrium oryzae*
Sheath spot	*Rhizoctonia oryzae*
Stackburn (Alternaria leaf spot)	*Alternaria padwickii*
Stem rot	*Magnaporthe salvinii* *Sclerotium oryzae* [synanamorph]
Water-mold (seed-rot and seedling disease)	*Achlya conspicua, Achlya klebsiana,* *Fusarium* spp., *Pythium* spp., *Pythium dissotocum,* *Pythium spinosum*

Table 2.2 Major Rice diseases and their biocontrol agents (Source: Vasudevan et al. 2002)

Disease	**Causal organism**	**Biocontrol agent**
Blast	*Pyricularia grisea*	*Pseudomonas fluorescens*
Brown spot	*Bipolaris oryzae*	*Pseudomonas sp.*
		P. aeruginosa
		Bacillus sp.
		B. subtilis
Bacterial blight	*Xanthomonas oryzae* pv. *oryzae*	*Bacillus* sp.
Sheath blight	*Rhizoctonia solani* Kuhn	*P. fluorescens*
		P. putida
		Bacillus sp.
		B. subtilis
		B. laterosporus
		B. pumilus
		Serratia marcescens
		Pseudomonas sp.
		P. aeruginosa
Sheath rot	*Sarocladium oryzae*	*P. fluorescens*
		B. subtilis
		P. aeruginosa
		Pseudomonas sp.
Stem rot	*Sclerotium oryzae*	*P. fluorescens*
		P. aeruginosa
		B. subtilis
		B. pumilus
Tungro	Rice tungro virus Vector - *Nephotettix* spp.	*P. fluorescens* (forvector)

Table 2.3 The species and genome groups of *Oryza* (Source: Vaughan et al. 2003)

Species	2n	Genome	Potentially useful traits	Distribution
Sect. *Oryza*				
Ser. *Sativae*				
O. barthii	24	AA		Sub-Saharan Africa
O. glaberrima	24	AA	Cultigen	West Africa
O. glumaepatula	24	AA	Elongation ability, CMS	South, Central America
O. longistaminata	24	AA	Resistance to BB and drought	Sub-Saharan Africa
O. meridionalis	24	AA	Elongation ability, drought tolerance	Tropical Australia
O. breviligulata	24	AA	Resistance to GLH, BB, drought	Africa
O. nivara	24	AA	Resistance to grassy stunt, virus, blast, BB and drought	Tropical, Subtropical Asia
O. rufipogon	24	AA	Elongation ability, CMS, resistance to BB	Tropical, Subtropical Asia, Tropical Australia
O. sativa	24	AA	Cultigen	Worldwide
Ser. *Latifoliae*				
O. alta	48	CCDD	Resistance to striped stem borer, high biomass production	South, Central America
O. eichingeri	24	CC	Resistance to yellow mottle virus, BPH, WBPH, GLH	South Asia, East Africa
O. grandiglumis	48	CCDD	High biomass production	South, Central America
O. latifolia	48	CCDD	Resistance to BPH, high biomass production	South, Central America
O. minuta	48	BBCC	Resistance to BPH, Sheath blight, BB, GLH	Philippines, Papua New Guinea
O. officinalis	24, 48	CC, BBCC	Resistance to thrips, BPH,GLH, WBPH	Tropical, Subtropical Asia
O. punctata	24, 48	BB, BBCC	Resistance to BPH, zigzag leafhopper	Sub-Saharan Africa
O. rhizomatis	24	CC	Drought tolerance	Sri Lanka
Ser. *Australienses*				
O. australiensis	24	EE	Drought tolerance, resistance to BPH	Tropical Australia
Sect. Brachyantha				
Ser. *Brachyanthae*				
O. brachyantha	24	FF	Resistance to yellow stem borer, leaf folder, whorl maggot, tolerance to laterite soil	Sub-Saharan Africa
Sect. Padia				
Ser. *Meyerianae*				
O. granulata	24	GG	Shade tolerance,	South, Southeast Asia
O. meyeriana	24	GG	Shade tolerance	Southeast Asia

Species	2n	Genome	Potentially useful traits	Distribution
O. neocaledonica	24	??	No report	New Caledonia
Ser. *Ridleyanae*				
O. longiglumis	48	HHJJ	Resistance to blast, BB	Indonesia (Irian Jaya), Papua New Guinea
O. ridleyi	48	HHJJ	Resistance to blast, BB, whorl maggot, stemborrer	Southeast Asia
Ser. *Schlechterianae*				
O. schlechteri	48	HHKK	Stoloniferous	Papua New Guinea

Table 3.1 List of gene specific primers for allele mining

Gene	Gene specific Primer	5' to 3' Sequence	Tm (◦C)	Product size (≈) bp
Xa21	Xa21.1F	GGGGTGTTTACATCCATAGG/ ATGATATCACTCCCATTATTGC	60	1700
	Xa21.1R	CGCTCAAGTTGTTTTCGTAGG		
	Xa21.2F	GACCTCGGCGACAACTACCT	60	1000
	X21.2R	GGCAATCACCAAGCGTGTTAG		
	Xa21EF	GACCTCGGCGACAACTACCT	61	3000
	Xa21ER	TTCAGCCAGTTCTCAGAATCC		
	Xa21.3F	CAGATACCCACATCCTTAGC	58	1000
	Xa21.3R	CTCCATCAGTTTCACTAGAGG		
	Xa21.4PF	GCAATGTGCTGTTAGATTCTG	60	1600
	Xa21.4PR	CACACACGTCAACTAGGAAGC/ TCAGAATTCAAGGCTCCCACC		
Xa26	Xa26.1F	CAGGGGTTGTGATGTACTGG	60	1400
	Xa26.1R	GTCGCTGCTGTTACTCTTGC		
	Xa26.2F	GCATACTCTTGCTGCCAATGC	60	1300
	Xa26.2R	GATCCATCCAGCAAGTTTCC		
	Xa26.3F	AGATTTGACGACGTGCAACC	60	1300
	Xa26.3R	GTCAGGCTTACCAGCAGAAG		
	Xa26.4F	ATGTGGTGCTGCCCGTTTAGG	62	1300
	Xa26.4R	CATGTCAAGCTACCAGTGAAGC		
Xa27	Xa27.1F	CGTCTCGAATCCCGTCCCTA	62	1200
	Xa27.1R	TAGCCAACCAGCAACGCCAC		
	Xa27.2F	AGGAGGGCGACGCAAGGTTC	62	1200
	Xa27.2R	TCCTTTCTTCAGCGGGCACC		
	Xa27.3F	TTTCAGAGTTCCCCATCACC	60	850
	Xa27.3R	CCATTTCCGTTTGGTACAGC		
xa5	Xa5.1F	AGCAGCATTTCCAAGAGTGG	54	1030
	Xa5.1R	TAC GTG TCG AAC GTG AAT GG		
	Xa5.2F	CTTTTCCGTACACACCTTGC	54	1234
	Xa5.2R	ACA ACC GAG TGA CCC AAA GG		
	Xa5.3F	CGAAATTCTGTGTGGGTTGC	53	1130
	Xa5.3R	TTA GCT CAG CAG GGA GAA GC		
	Xa5.4F	TGCTACCGGAGACTAGAAAGG	62	1200
	Xa5.4R	AGG TGT GCC GTA TTT AGA AGC		
	Xa5.5F	ACGGCACACCTTTATCATCC	54	690
	Xa5.5R	CCA GTT CTC CTG CAT TTT CG		
	Xa5.6F	CGAAAATGCAGGAGAACTGG	54	950
	Xa5.6R	GCC CTG AGA AAA GAA CAT GC		
	Xa5.7F	GACGCAATGTCGAACAGACC	60	750
	Xa5.7R	GCC AAA TGT TGC TAG GGT TG		

N.B. ≈ - approximate

Table 3.2 List of primers for expression analysis of *xa5* and *Xa21* genes

Gene	Primer name	5' to 3' sequence	Tm (◦C)	Product size (≈) bp
Xa21	X21RTF	GCTCGATGGGATTTATAGGG	60	150
	X21RTR	GCTTCCCGGTTACTATTTCC	60	
xa5	X5RTF1	GAGCAAGGTTTCTATCAAGG	60	150
	X5RTR1	CCAACTTGTTCTGTAGTCTCC	60	

N.B. ≈ - approximate

Table 3.3 List of primers used in RT-PCR for defense response genes

Gene	Primer name	5' to 3' sequence	Tm (◦C)	Product size (≈) bp
OsBAK1	OsBAK1F	TTGCAATAACGACAACAGTG	58	150
	OsBAK1R	ACTAGGTATCGTTCCGCTTA		
SA	EI12I1F	GTGCAGTAGCTCCAAGGGGT	62	150
	EI12I1R	GGCACTCGGAAGTTCGCCA		
	BI26N5F	AACTTCAGGGTGATCAGAGG	60	150
	BI26N5R	CCCCAACCTATAGTCTTTGC		
	EI35I3F	GCGTCCAGCCTCCTCAAGAC	62	150
	EI35I3R	GTCATCTCGCGGGACACGAT		
	EI10P9F	AAATCTCCCCTCGAAGCGA	62	150
	EI10P9R	CAATGGTGTCCGAGGACTC		
	EI1K8F	TGTTTCCGATGGAGGCGTG	62	150
	EI1K8R	GCAGAACTCTCCGCCCAGAT		
	EI5P11F	CGAACACAAGTACCTTGACG	60	150
	EI5P11R	GAGCCCAATTTAGCACTCG		
	OsPR1#074F	GTATGCTATGCTACGTGTTTATGC	60	150
	OsPR1#074R	GCAAATACGGCTGACAGTACAG		
	OsPR1#011F	ACGCCTTCACGGTCCATAC	58	150
	OsPR1#011R	CAGAAAGAAACAGAGGGAGTAC		
	OsPR1#012F	CGCTGTGTGTTTGTGTTATGTC	62	150
	OsPR1#012R	CGTGGTTTTGTCTTTATTTCAATCC		
RPLD1	RPLD1F	CTGAGATCACCTTGGTTAGG	58	150
	RPLD1R	CAACTGAAGTCCTGTCATCC		
RPLD2	RPLD2F	GAAGAGGATCAATGCTGACC	60	150
	RPLD2R	TAGCATAGTCTGTGCCATCC		
OsAOS	OsAOSF	CATCGTGGACACACTATCAC	58	150
	OsAOSR	AGAGAAATAACGAGGAACCG		
Reference gene	ActinF	GAGTATGATGAGTCGGGTCCAG	58-60	150
	ActinR	ACACCAACAATCCCAAACAGAG		
	NBPF	GGAATGTGGACGGTGACACT	60	150
	NBPR	TCAAAATAGAGTCCAGTAGATTTGTCA		
	ExPF	AGGAACATGGAGAAGAACAAGG	60	150
	ExPR	CAGAGGTGGTGCAGATGAAA		

N.B. ≈ - approximate

Table 4.1 Polymorphism and neutral test of different regions of the *Xa21* gene

	Total sites (excluding sites with gaps)	S	π	Θ_w	Fu and Li's D	Tajima's D	K	Parsimony informative site	Singleton variable sites
Entire gene	3775	191	0.0168	0.0233	-	-1.0547*	54.55	113	78
Coding	2994	150	0.0105	0.0143		-1.1837*	31.59	84	66
O. nivara seq	3828	88	0.0078	0.0084	-0.2982	-0.3874*	29.9	45	43
O. sativa seq	3812	73	0.0068	0.0078	-0.7547	-0.7641*	26.23	23	50

Note: * P > 0.01 (not significant), S = No. of polymorphic sites, K = Average nucleotide difference, π = Nucleotide diversity, Θ_w = No. of segregating sites.

Table 4.2 Polymorphism and neutral test of different regions of the *Xa26* gene

	Total sites (excluding sites with gaps)	S	π	Θ_w	Fu and Li's D	Tajima's D	K	Parsimony informative site	Singleton variable sites
Entire gene	3221	304	0.0195	0.0274	-	-1.2974*	63.07	122	182
Coding	3240	260	0.0168	0.0233	1.56665	-1.2754*	54.55	111	148
O. nivara seq	3408	190	0.0221	0.0227	-0.01961	-0.1833*	75.5	98	92
O. sativa seq	3377	248	0.0213	0.027	-1.11	-1.18*	72.22	79	169

Note: * P > 0.01 (not significant), S = No. of polymorphic sites, K = Average nucleotide difference, π = Nucleotide diversity, Θ_w = No. of segregating sites.

Table 4.3 GenBank accession numbers of *Xa27* alleles

S. No.	Genotype	Accession no.	Sequence length	Percentage homology with reference sequence
1.	ON24	JF304301	2,341	97
2.	102166	JF304302	2,390	97
3.	ON34	JF304303	2,380	97
4.	IC384116	HQ888852	2,406	97
5.	106133	HQ888853	2,406	97
6.	R4	HQ888854	2,407	97
7.	R2	HQ888855	2,392	97
8.	MP	HQ888856	2,395	97
9.	ML	HQ888857	2,733	97
10.	M209A	JN016505	2,371	96
11.	100963	JN016506	2,390	98
12.	R9	JN016507	2,364	96
13.	PAU933	JN016508	2,058	96
14.	210223	JN016509	2,057	98
15.	R23	JN016510	2,393	98
16.	81832	JN016511	2,056	96
17.	21028	JN016512	2,060	96
18.	ON38-1	JN016513	2,405	97
19.	Os-28	JN016514	2,369	97
20.	ON44	JN016515	2,383	97
21.	OS 23-2	JN016516	2,387	97
22.	ON15	JN016517	2,058	96
23.	IC203740	JN016518	2,393	98
24.	IR20	JN016519	2,393	98
25.	ON132	JN016520	2,057	97
26.	PAU201	JN016521	2,394	97
27.	TN1	JN601064	2,395	97

Table 4.4 Polymorphism and neutral test of different regions of the *Xa27* gene

	Total sites (excluding sites with gaps)	S	π	Θ_w	Fu and Li's D	Tajima's D	K
Entire gene	1999	103	0.00795	0.01312	-	-1.56*	15.892
Coding	338	17	0.01202	0.01281	-	-0.2107*	4.064
Promoter & 5'UTR	1198	84	0.00916	0.01785	-	-	10.975
O. nivara seq	2004	57	0.00790	0.00875	-0.82225	-0.41969*	15.829
O. sativa seq	2039	73	0.00845	0.01186	-1.48507	-1.6476*	17.22

Note: * $P > 0.01$ (not significant), S = No. of polymorphic sites, K = Average nucleotide difference, π = Nucleotide diversity, Θ_w = No. of segregating sites.

Table 4.5 GenBank accession numbers of *xa5* alleles

S. No.	Genotype	Accession no.	Sequence length	Percentage homology with reference sequence
1.	106133	KF689667	6204	97
2.	ON132	KF689663	6212	99
3.	81832	KF689666	6209	99
4.	81852	KF689664	6269	97
5.	ON15	KF689665	6208	98
6.	AC32753	KF689660	6365	99
7.	Kalamekri	KF689659	6227	99
8.	TN1	KF689661	6252	99
9.	PB1	KF689662	6253	98
10.	100963	KF689658	6190	99
11.	M209A	KF689657	6209	99
12.	OL1	KF689656	6201	99
13.	OLD	KF689655	6242	99
14.	MP	KF689654	6225	99
15.	ML	KF689653	6243	99

Table 4.6 Polymorphism and neutral test of different regions of the *xa5* gene

	Total sites (excluding sites with gaps)	S	π	Θ_w	Fu and Li's D	Tajima's D	K
Entire gene	5958	310	0.01324	0.01539	-	-0.6719*	78.9
Coding	318	11	0.00592	0.01023	-	-1.5639*	1.882
O. nivara seq	6108	154	0.01106	0.01029	0.00425	0.20337*	67.57
O. sativa seq	6108	102	0.00733	0.00644	0.48035	0.74876*	44.75

Note: * $P > 0.01$ (not significant), S = No. of polymorphic sites, K = Average nucleotide difference, π = Nucleotide diversity, Θ_w = No. of segregating sites.

Table 4.7 List of different proteins identified in 2DE and MALDI MS/MS search analysis

Spot No.	Protein identification	Peptide sequence	Accession	Score (%)	Peptide coverage (%)	Fold increase in expression
		Metabolism/Energy				
1127	Rubisco large chain	TFQGPPHGIQVER GGLDFTKDDENVNSQPFMR	gi|11466795	88		5
1293	Triosephosphate isomerase	**FFVGGNWK** **VIACVGETLEQR** **VATPDQAQEVHDGLRK**	gi|553107	132	14	1.8
1196	Phosphoribulokinase (PRK)	**IFVIEGLHPMFDER** **NFNPVYLFDEGSSITWVPCGR** **IRDLYEQIIAER**	gi|115448091	145	11	1.52
1216	Ribose-phosphate pyrophosphokinase 2	LSANLLTEAGSDR	KPRS2_ORYSJ	29	3	1.95
1226	Fructose-bisphosphate aldolase	GLVPLAGSNNESWCQGLDGLASR YAAISQDNGLVPIVEPEILLDGEHGIDR	ALFC_ORYSJ	47	13	3.4
1295	Esterase	**ADSSSTEEMQRWMVAMSPGTDVEE** **IAGADHAVMNSKPR**	PIR7B_ORYSI	35	14	1.97
		Undetermine/Hypothetical				
1176	Hypothetical	**MASNKQGGGK** **KQQQLQEEEEATK** **EEEEEAAAK**	gi|125553543	70	13	2.4
1198	Hypothetical protein	RSTLESGEGSSESVASDTGTAK	gi|218201674	27	5	3.2
1122	Flotillin-like protein 1	GVIEGETRVLAASMTMEEIFQGTK	FLOT1_ORYSJ	19	4	2.37
1334	Hypothetical protein	**SSEMISTPGSTTKTNTSISSR**	gi:125535351	7	9	1.53
1456	Hypothetical protein	D**LASMVALAVESPGAAAGR**	gi|115471157	84	4	2.1
		Defense related				
1218	Putative thiazole synthase	KAAAAAAAHPELILASK	gi|125558603	12	4	2.9
1190	Methyl transferase	LMLDVDEDNLSTWHQM AAAVVSGGPSAFER	METL_ORYSJ	17	8	2.3
1317	Putative germin-like protein	AASNFFLLTAFIALVATQA MASDPSPLQDFCVADK	GL92_ORYSJ	13	15	3.3
1455	Serine threonine protein kinase		NEK5_ORYSJ			2.1
1457	cysteine synthase	IQGIGAGFVPR LVVVVFPSFGER	gi|115442595	56	5	3.6
		Transcription factor/putative stress related				
1290	E2F/DP family winged-helix DNA-binding domain					2.68
1302	PHD finger protein ALFIN-LIKE 1	SGMDNSSR	C3H36_ORYSJ	8	3	2.4