QTL Mapping of Economically Important Traits and DNA based Detection and Quantification of Adulteration of Basmati rice (*Oryza sativa* L.)

Thesis submitted to

University of Hyderabad

For the award of the degree of

Doctor of Philosophy

By

Vemireddy Lakshmi Narayana Reddy

Laboratory of Molecular Genetics
Centre for DNA Fingerprinting and Diagnostics (CDFD)
June 2008

Centre for DNA Fingerprinting and Diagnostics
(An autonomous centre of the Department of Biotechnology,
Ministry of Science and Technology, Govt. of India)
Hyderabad 500 076, India

J. Nagaraju

Certificate

This is to certify that the research work in the thesis titled "QTL mapping of economically important traits and DNA based detection and quantification of adulteration of Basmati rice (*Oryza sativa* L.)" has been carried out by Mr. Vemireddy Lakshmi Narayana Reddy under my supervision. This work has not been submitted for other degree or diploma to any University or institute.

[J. Nagaraju]

Research supervisor

Group Leader, Laboratories of Molecular Genetics and DNA Fingerprinting
Team Leader, Centre of Excellence on Genetics and Genomics of Silkmoths
Scientific Consultant, APEDA-CDFD Centre for Basmati DNA Analysis
Email: jnagaraju@cdfd.org.in Phone:+9140 27171427
www.cdfd.org.in/jnagres.html

School of Life Sciences
University of Hyderabad
Hyderabad – 500 046, India

Certificate

This is to certify that Mr. Vemireddy Lakshmi Narayana Reddy has carried out the research embodied in the present thesis for a full period prescribed under the Ph.D. ordinance of this University. We recommend his thesis titled "QTL mapping of economically important traits and DNA based detection and quantification of adulteration of Basmati rice (*Oryza sativa* L.)" for the award of "Doctor of Philosophy" of the University.

Dr. J. Nagaraju, Research Supervisor
Laboratory of Molecular Genetics
Centre for DNA Fingerprinting and
Diagnostics (CDFD)

Prof. E. A. Siddiq, Co-Advisor
Distinguished Chair, CDFD
Adjunct Professor,
University of Hyderabad

Prof. P. Appa Rao, Head
Department of Plant Sciences
University of Hyderabad

Prof. A. S. Raghavendra, Dean
School of Life Sciences
University of Hyderabad

Declaration

I hereby declare that the thesis entitled “QTL mapping of economically important traits and DNA based detection and quantification of adulteration of Basmati rice (*Oryza sativa* L.)” embodies the original research work carried out by me at Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics, Hyderabad under the guidance of J. Nagaraju. This thesis has not been submitted for any other degree or diploma to any university or institute. Any assistance received is suitably acknowledged and original works referred to have been cited.

Vemireddy Lakshmi Narayana Reddy
ID.No: 01LPPH14

Acknowledgements

I am tempted to thank individually all who made this thesis possible. However, because the list might be too long and by fear of leaving someone out, I will simply say "thank you very much to you all".

I wish to express my profound gratitude and sincere thanks to Dr. J. Nagaraju, my thesis advisor. Although he is expertise in silkworm genetics, out of his own enthusiasm and curiosity on Basmati rice, albeit not having separate funds, he encouraged me to choose this topic. It is really pleasure to work with him. I am grateful for his guidance, useful suggestions throughout the period of this research, constructive criticisms and the correction of the thesis. Without his encouragement and constant guidance I could not have finished this thesis.

It is difficult to overstate my appreciation to Prof. E. A. Siddiq, my co-advisor, who first brought me into the world of research with whom I began to learn molecular techniques before joining PhD. He was also instrumental in bringing me to CDFD. Not only a great mentor, he has also been a cornerstone in my professional development. Only few scientists can spare their chambers for doing research work apart from office work. He is the one that kind who could spare his chamber for my crossing programme. In addition, his invaluable supervision and guidance helped in planning and pursuing of my Ph D education.

Besides my advisors, I would like to thank Drs. Syed E. Hasnain and J. Gowrishankar for making it possible for me to do research in CDFD. Grateful acknowledgement is also expressed to Drs. Shashidhara and Gayatri Ramakrishna for their valuable and timely assistance in providing chemicals.

I am grateful to Council of Scientific and Industrial Research (CSIR) which provided the fellowship for this degree.

I would also like to acknowledge ANGRAU for allowing me to do field experiments. Thanks to scientists of ARI, Drs. Anuradha, Srilakshmi and Hamidunnisa for their kind support during these years. I thank my colleagues Ramana, Sridhar, Jayaprada and Naresh in the Biotechnology Unit, ARI, for their help in the field operations and data collection. I express my sincere thanks to Dr. Shobarani, DRR for allowing me to perform quality analysis assays in her lab.

I would like to thank my colleagues in LMG - Satish, Ramana, Arun, Sunil, Mrinal, Jay, Subbu, Dharma, Archana, Jyoti, Asha, Muthu, Shoban, Padma priya, Deepa, Gayatri, Malathi, Satya, Vijay, Jana, Ravi, Murali, Sudhakar, Archana Tomer, Shekar and Bhaskar for their support, timely help and suggestions to finish my thesis work. My exceptional thanks to Basmati group members – Parimala, Pavani, Kaliappan, Swarna and Sudhakar for their support in the lab work. Thanks also to CDFD friends – Santhosh, Madhav, Subba rao, Aisha, KV, Nitin, Chandu, Ajay, Mehfooz, Pankaj, Uma, Charitha, Raghu, Srinu, Kalyani, Ramesh, Prasad, Gokul and Prashanti.

I am forever grateful to my parents, brothers, sister and their families for their support and encouragement throughout the tenure of Ph D. Special thanks to my parents-in-law and their family for their assistance in completing Ph D.

I cannot end without thanking my dear wife Srividya, despite pursuing Ph D, who endured the long evenings and weekends and provided constant encouragement and my loving son Siddharth who brought us enjoyment and liveliness in our lives. To them I dedicate this thesis.

[Vemireddy Lakshmi Narayana Reddy]

Contents

List of Figures

Contd….

List of Tables

Synopsis

Name of the Student	:	Vemireddy Lakshmi Narayana Reddy
Regd. No.	:	01LPPH14
Title of the thesis	:	QTL mapping of economically important traits and DNA based detection and quantification of adulteration of Basmati rice (*Oryza sativa* L.)
Degree	:	Doctor of Philosophy (Ph. D)
Department	:	Plant Sciences
Research institute	:	Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad
Research supervisor	:	Dr. J. Nagaraju
Co-Advisor	:	Prof. E. A. Siddiq
University	:	University of Hyderabad

Basmati rice (*Oryza sativa* L.*),* originated in the foothills of Himalayas, command premium price in the domestic and international markets on account of its unique quality traits characterized by long slender grain, extra lengthwise elongation and soft fluffy texture on cooking, and exquisite aroma. The varieties globally recognized as genuine traditional Basmati (TB) from India include Basmati370, Taraori Basmati and Dehradun Basmati. TB rice is not only in great demand in the domestic markets, but is also seen in the menu of connoisseurs worldwide. Basmati rice qualifies as a Geographical Indication (GI).

Earlier investigations on genetic analysis of different accessions of rice revealed that the Basmati rices form a distinct cluster different from *indica*, *japonica* and *javanica* varietal groups suggesting that Basmati rice falls in phylogenetically divergent group. As for inheritance and breeding behavior are concerned, all the indices of grain appearance and cooking/eating qualities with the exception of amylose content and to some extent aroma have been reported to be polygenically controlled. For instance, amylose content is reported to be governed by single gene located on chromosome 6 i.e., waxy locus in association with modifier complex. Some researchers are of the opinion, that *wx* gene also controls gelatinization temperature and gel consistency. In case of aroma, most researchers believe that it is under the control of single recessive gene. Recently, a gene for fragrance (*fgr*) was identified on chromosome 8 which has homology to the gene that encodes betain aldehyde dehydrogenase 2 (*bad2*). Thus, the complex genetic architecture of unique traits of Basmati and tedious screening methodologies

involved in quality testing have been serious constraints to breeding for quality of Basmati. Moreover, very limited information is available on the genetic basis of grain and cooking qualities of rice in general and basmati rice in particular.

In addition to the desired quality traits, Basmati rice also harbours many undesirable traits that include tall stature, low yield and photosensitivity. Attempts since 1970s to develop high yielding Basmati rice varieties by cross breeding, though, resulted in many high yielding Evolved Basmati (EB) varieties; they fell short of the quality standards of TB varieties. The evolved and non-Basmati rices looking similar to TB varieties in appearance have suited the interests of unscrupulous traders to use them as adulterants in Basmati trade. The practice of Basmati adulteration would seriously jeopardize our export trade of over Rs.2000 crores. Hence, identifying genuine Basmati variety from other look-alike long grain non-Basmati varieties is important from the viewpoint of sustainable trade. Traditionally employed morphological and biochemical assays for detecting adulteration in Basmati rice have not been found to be discriminative enough warranting more precise high-throughput techniques.
Keeping the foregoing in view, the present study was undertaken with the following objectives:

1. QTL mapping of economically important traits of Basmati rice.
2. Detection and quantification of adulteration in Basmati rice.

Objective 1 QTL mapping of economically important traits of Basmati rice

For mapping QTLs, an F_2 population of 181 plants was developed from a cross of a traditional Basmati variety Basmati370 and a semi dwarf variety Jaya. Phenotypic data of 18 agronomic and quality traits of the parents, hybrids and F_2 mapping population was recorded. DNA was isolated from the two parental varieties and screened for polymorphism using 502 rice microsatellite markers (RMs). Of these, 203 showed polymorphism (40.44%). The polymorphic markers were distributed on all the 12 rice chromosomes. These markers were used for screening the mapping population. A molecular linkage map comprising 134 microsatellite markers spanning the rice genome at 18.37 cM interval was constructed using MAPMAKER/EXP v 3.0 and MapDisto v 1.7 softwares. A total of 47 QTLs for 16 economically important traits of Basmati was identified employing interval mapping (IM) and composite interval mapping methods of the QTL

Cartographer v 2.5. These QTLs were found to be distributed on all the rice chromosomes except 7 and 11. Of the 47 QTLs, 17 contributed to more than 15% phenotypic variance. Interestingly, a single region on chromosome 5 at the marker interval of RM430-RM18600 was found to be controlling four important grain quality traits viz., grain length, grain breadth, length-breadth ratio and grain elongation ratio. The promising QTLs controlling important economic traits in basmati rice, identified in this study, could be used as candidates for future fine mapping and positional cloning.

Objective 2 Detection and quantification of adulteration of Basmati rice

In the present study, a capillary electrophoresis (CE) based microsatellite DNA profiling protocol was developed, which facilitates quick and accurate detection and quantification of the adulteration in basmati rice. The single-tube assay multiplexes eight microsatellite loci viz., RM1, RM55, RM44, RM348, RM241, RM202, RM72 and RM171 to generate variety-specific allele profiles that can detect adulteration from 1% upwards. Accuracy of quantification has been shown to be ±1.5%.

In addition, it was shown that CE is essential for microsatellite based protocol for detection and quantification of adulteration in basmati rice by comparing with slab-gel and agarose gel electrophoresis. Comparative analysis across eight microsatellite loci in twelve rice varieties demonstrated that CE method showed less error (±0.73bp) in estimation of allele sizes compared to slab-gel (±1.59bp) and agarose gel (±8.03bp) electrophoretic methods. When the average deviations were further scrutinized for the actual distribution of allele sizes, it was observed that nearly 58% of the alleles estimated by capillary electrophoresis showed exact values (zero deviation). In all, >90% of the capillary electrophoresis assays resulted in allele sizing within a base-pair bin, whereas the corresponding estimate was <50% for slab gel electrophoresis and a dismal 7 and 2%, respectively, for software-based and manually measured agarose gel electrophoresis. Moreover, CE method showed greater reproducibility (<0.5bp deviation) compared to slab-gel (1bp) and agarose (>3bp) based methods. CE method was significantly superior in quantification of the adulterant also, with mean error of ±3.91% in comparison to slab-gel (±6.09%). Lack of accuracy and consistency of the slab-gel and agarose electrophoretic methods warrants the employment of capillary electrophoresis for Basmati rice purity tests.

General Introduction

Rice, the second most widely grown cereal crop, is the staple food for more than one half of the world's population. Accounting for 23 percent of the world's supply of calories (Ashikari *et al* 2005), it is one of the most versatile crops grown worldwide under a wide range of agro-ecological conditions ranging from irrigated and rainfed low land and upland to deep water and tidal conditions. Asia accounts for 90% of the global production and consumption of rice. Although the world's rice production has more than doubled from 257 million tons in 1966 to 600 million tons in 2000, the increase has not kept up with the demand because of faster growth of human population than rice production during this time. It is estimated that rice production must increase by at least 50 per cent in the next 25 years to keep pace with the demand. This calls for development of rice varieties with higher yield potential combining tolerance to biotic and abiotic stresses as agriculture land is fast shrinking thus limiting the scope for horizontal expansion.

Rice is one of the very few crop species endowed with rich genetic diversity. More than one hundred thousand landraces and improved cultivars available in the gene banks world over largely constitute the diversity. Varietal diversification in rice has been by natural selection initially as a function of adaptation to diverse agro-climatic conditions and conscious and continuous selection subsequently by man for his diverse quality preferences. A unique varietal group that has gained wider acceptance as a speciality rice all over the world is **"Basmati rice"**.

Regarded as the "**king of the rices**", Basmati rice occupies a special place among all aromatic rice cultivars by virtue of its special quality characterized by extra long slender grain, lengthwise excessive elongation on cooking, soft and fluffy texture of the cooked rice, and pleasant distinct aroma. As a result, it commands premium price in the domestic as well as overseas markets unlike the non-aromatic rice varieties (Archak *et al* 2007; Nagaraju *et al* 2002; Vemireddy *et al* 2007). Amidst several grown and marketed Basmati varieties, only six in India are recognized as traditional Basmati (TB) varieties of which, Basmati370 Taraori Basmati and Dehradun Basmati are the most popular ones (Archak *et al* 2007). Unlike other aromatic rices, Basmati rice expresses its unique quality traits only when grown in the north-western foot hills of the Himalayan in the Indian sub-continent. And, because of its location specific quality performance Basmati is now a Geographical Indication (GI) belonging to a specific geographical area in the Indian subcontinent and it is like "champagne" among wines and "scotch" among

whisky. Basmati has attained "**heritage rice**" status as is considered "**farmers cultivar**" grown by farmers in India and Pakistan for years (Nene, 1998) stretching across the Punjab region of India and Pakistan and Haryana and Uttar Pradesh regions of India for more than 250 years (Figure A). The special qualities of Basmati are attributed to unique and complex combination of soil, water, and climate, cultural practices under which it is grown besides inherent genetics governing these features. Thus attempts to replicate the grain qualities by growing Basmati rice anywhere else in the world have not been successful.

Figure A Basmati rice growing areas of India and Pakistan

History of Basmati rice

Reference to existence of some aromatic rices can be traced to the documents of Susrutha (*circa* 400 BC), a great Indian pioneer in medicine and surgery. The first usage of the word "**basmati**" can be found in Punjabi classic ***Heer Ranjah*** (1766) by the Punjabi poet Waris Shah (Nene, 1998). The usage of names of rice cultivars such as *Basmati*, *Begumi*, *Satthi* (60-day rice) suggests that "Basmati" rices have been in cultivation since 1700 AD. Subsequent reference to the "Basmati"

can be found in the book "**A dictionary of the Economic Products of India**" compiled by George Watt and published in 1891. The word "Basmati" appears to have been derived from two Sanskrit words i.e, *vaas* (fragrance) and *matup* (possessing). *Vaasmati* pronounced now as Basmati appears to mean the one that "**possesses fragrance**." In North India "*va"* is often pronounced as "*ba"*, and that is how the word "*vasmati"* must have become "Basmati". Basmati rice was a preferred rice for the preparations of *Biryani* and *Pulao* on special occasions and was patronised by emperors while poets have been composing poems admiring its qualities (Nene, 1998). The credit for conserving and improving Basmati rices to suit the socioeconomic and agro-climatic should go to the farmers of geographical region in the Indian subcontinent. The names of Basmati varieties that exist now are quite often suggestive of the places where they are grown. For instance, Taraori Basmati is from the area known by Taraori in Haryana and Dehradun Basmati from Dehradun in Uttaranchal.

Diversity in Basmati rice

As mentioned earlier, rice is quite rich in its genetic diversity distributed worldwide. Out of 21 valid species of genus *Oryza*, only two viz., *Oryza sativa* (Asian cultivar) and *Oryza glaberrima* (African cultivar) are cultivated (Vaughan *et al* 2003). Since its origin and domestication, *O. sativa* is reported to have undergone differentiation into three eco-geographical races or sub-species, viz., *indica*, *japonica* (temperate) and *javanica* (tropical *japonica*) (Singh *et al* 2000; Vaughan *et al* 2003). The diversification of *O. sativa* is not confined only to these three sub species as it continued to further differentiate into different varieties as a result of natural and human selections over centuries under diverse agro-climates, cultural practices etc. According to traditional classification, Basmati is described under *indica* group because of its long and thin grains (Sweeney and McCouch 2007). However, earlier reports on diversity studies have suggested that Basmati varieties form a distinct group of their own separated from *indica* and *japonica*. For instance, isozyme analysis of Glaszmann, reveals Asian cultivars to have differentiated into six varietal groups viz., ***indica, aus, ashinas, rayadas, aromatic, and japonica*****.**

Majority of the aromatic rice varieties that included Basmati were clustered in a separate group (Group V), whereas typical *indica* and *japonica* varieties formed Group I and Group IV respectively (Glaszmann, 1987) (Figure B). Many of the varieties belonging to the Group V are characterized by elongated kernels after cooking. Although aromatic varieties separated quite

well from the other groups, they were comparatively more closer to the *japonica* group than to *indica* group (Garris *et al* 2005). Findings from a recent study, using 169 microsatellite loci and two chloroplast loci, are in complete agreement with the results of Glaszmann (Garris *et al* 2005) (Figure B). High percentage of hybrid sterility of crosses of Basmati with typical *indica* varieties further confirm that Basmati group is more distant genetically from *indica* (Nagaraju *et al* 2002).

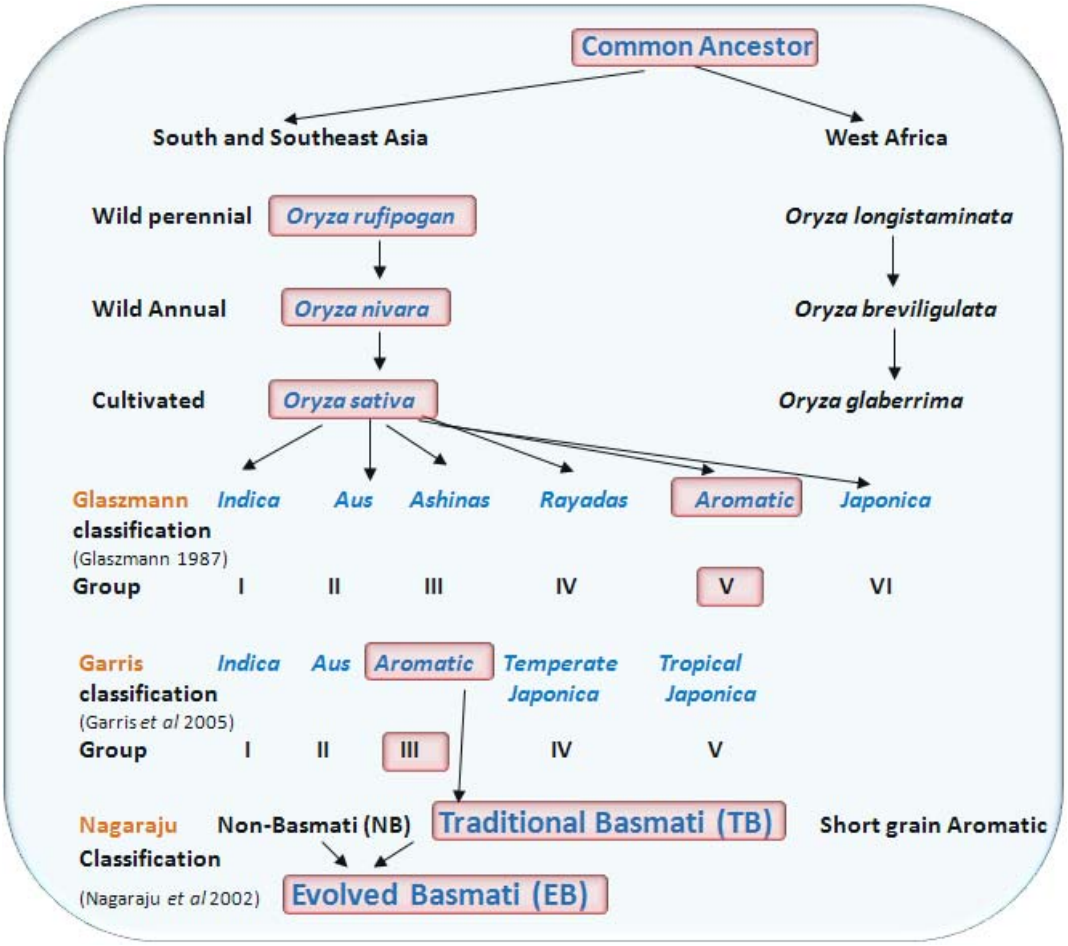


Figure B Evolutionary pathway of Basmati rice

Nagaraju *et al* (2002) studied genetic relationship of traditional basmati (TB), evolved basmati (EB), and non-basmati (NB) varieties using simple sequence repeats (SSR) loci and fluorescent labelled inter-SSR-PCR (FISSR) primers. The findings reported in the study indicate that in likelihood all the traditional basmati varieties might have evolved from a single landrace

and whatever little genetic variation observed among them could be as a result of selection practiced over years by farmers keeping in view the consumer practices as well as their own needs. Almost all the TB varieties known by different names are selections from one local variety, in all probability, Basmati 370. Isozyme patterns of 60 of the 65 Pakistani Basmati accessions and nine Indian Basmati varieties with the exception of Karnal Local matched that of Basmati 370 and Dehraduni Basmati (Nagaraju *et al* 2002). These results further support the view that all known TB varieties of India and Pakistan might have been the derivatives of either Basmati370-like or Dehraduni Basmati-like varieties.

Most of the agronomic and grain quality traits are controlled by many genes each of which with relatively a small effect on the overall phenotype. These traits do not show discrete phenotypes, consequently they are often measured and given a quantitative value and are referred to as quantitative traits. Understanding the genetic basis of quantitative traits is difficult as the observed phenotype is not exact of the genotype (Falconer and Mackay 1996). The expression of genes controlling quantitative traits is greatly influenced by environment. Consequently, improvement of polygenic traits by traditional breeding methods is time consuming and the gains are harder to realize. Breeders usually overcome this problem by multi-environmental evaluation of the traits in replicated trials to detain the effect of environment. Because of the intractable problem encountered while trying to improve quantitative traits by conventional means, breeders and geneticists have considered the potential use of molecular markers to identify chromosomal regions harboring genes that control quantitative traits.

Morphological markers and genetic markers were the first generation markers used for selection of quantitative traits. About 80 years ago Sax (1923) reported the association of quantitatively inherited seed size with simply inherited genetic markers that governed seed coat pigmentation and pattern in common bean, *Phaseolus vulgaris* L. In subsequent reports simply inherited and linked morphological mutations were used as markers for selection of quantitative traits. However, because of paucity of such markers and effect of them on quantitative traits often being larger than that of the linked QTLs, the extensive study of quantitatively inherited traits has been difficult (Tanksley *et al* 1998). Later, isozyme markers were used to identify QTLs in maize (Stuber and Edwards 1986). Isozymes have also been used for genetic analysis and to

study of linkage relationships in rice (Pham *et al* 1990). The available number of isozyme markers was too few to be of use for extensive study of quantitative traits.

Falconer and Mackay (1996) observed that if DNA markers linked to genes governing quantitative traits could be found indirect selection for such markers might improve the efficiency of breeding for the associated quantitative traits. This strategy is reliable because molecular markers are unaffected by environment and thus they can be used to identify the loci that affect respective quantitative trait(s). DNA markers have been used in a variety of crops to determine the number of genes controlling trait inheritance and for gene tagging (Pradeep *et al* 2005; Tanksley *et al* 1992).

The DNA markers are of two broad categories, *viz*. restriction enzyme based markers and DNA amplification or PCR based markers. The first category comprises restriction fragment length polymorphisms (RFLPs), which detect DNA polymorphisms through restriction endonuclease digestion followed by visualization via DNA blot hybridizations. Many of the RFLP markers are quite informative because of their co-dominant nature. They have been used for constructing saturated genome maps of rice (Causse *et al* 1994; Nagamura *et al* 1997). However, these markers require relatively large amounts of DNA for assay besides being time consuming and labor intensive and are not cost effective to construct high resolution maps (Liu *et al* 1994).

With the advent of polymerase chain reaction (PCR) coupled with technological advancements in DNA sequencing and genotyping high throughput methods were developed over the years for use in rice, including randomly amplified polymorphic DNA (RAPD) (Virk *et al* 1995), amplified fragment length polymorphism (AFLP) (Virk *et al* 2000) and microsatellites (McCouch *et al* 1997).

Of the several PCR-based markers employed for genetic mapping studies, microsatellite markers, also known as simple sequence repeats (SSR), are genetically informative and robust markers. Available in abundance in eukaryote genome including rice, SSR marker analysis requires only a small DNA template, technically simple to assay and are amenable to high throughput analysis (McCouch *et al* 1997).

The advantages of SSR markers have led to the development of high-density linkage maps of rice (Causse *et al* 1994; Harushima *et al* 1998). Such molecular maps have provided a powerful tool for elucidating genetic basis of several agriculturally important traits, many of which are quantitatively inherited.

In contrast to monogenic mutant traits, yield and its components show continuous phenotypic variation in segregating populations of crosses. The yield traits are controlled by polygenes often referred to as quantitative trait loci (QTLs). As many of the economically important traits are controlled by QTLs, identification of trait specific QTLs is important for crop improvement by marker associated improvement of the target traits. As QTL analysis enables finding genes that are otherwise difficult to identify, dissecting QTLs is quite informative and important to understand the genetic architecture of the traits in question. In the past, understanding the molecular basis of polygenic traits had been difficult in the absence of whole genome sequence information. Today, annotated whole genome sequence of rice is available using which we can design suitable molecular marker(s) for analyzing QTLs. Availability of high density marker linkage maps and powerful biometric methods has enabled considerable progress in QTL mapping.

Rice is an ideal model system for molecular genetic studies in monocots by virtue of :
(a) smallest genome size (390Mb) among cereals (maize-3,300Mb, barley-5,100Mb and wheat-16,000Mb, Ashikari and Matsuoka 2006), (b) diploid nature (2n=24), (c) availability of a large collection of cultivars and wild species (>1,20,000 accessions worldwide), (d) synteny with the genomes of other cereals, (e) amenability to regeneration of protoplasts, (f) availability of high density linkage maps (Causse *et al* 1994) and full length expressed tags (Sasaki *et al* 1994), and (g) relatively high degree of transformation efficiency (McCouch and Doerge 1995).

As of April 2008, 8646 QTLs have been identified and mapped in rice for different traits (Gramene V27; www.gramene.org). However, very few reports are available on QTL analysis of unique traits of Basmati rice. Hence, the present study was undertaken to identify and map QTLs pertaining to economically important agronomic and quality traits of Basmati rice. Well characterized Basmati rice specific molecular markers could serve as diagnostic marker tags for Basmati varieties. Markers tightly linked to any of the distinct traits of Basmati grain quality

could be used in marker assisted selection (MAS) programs for improvement of Basmati rices. Besides, the tightly linked markers could be used to identify candidate genes that confer/influence/control the trait(s) in question.

It is well known that TB varieties have many undesirable traits like tall stature hence proneness to lodging, low yield and photosensitivity. Attempts to correct these undesirable traits by cross breeding with high yielding semi-dwarf varieties have resulted in the proliferation of many evolved Basmati varieties but these varieties fell short of the quality of TB varieties (Nagaraju *et al* 2002). Hence attempts to replace the TB varieties with the high yielding EB varieties were not met with desired success. On the other hand, many "so called" high yielding of evolved and "look alike" non-Basmati rices have suited the interests of unscrupulous traders to use them as adulterants in Basmati trade. The practice of Basmati adulteration is reported to be rampant in Basmati trade and in the absence of intervention this would jeopardise export trade of over Rs.2000 crores. Hence, identifying genuine Basmati variety from other long grain look-alike non-Basmati varieties is important from the viewpoint of sustainable trade. Traditionally employed morphological and biochemical assays for detecting adulteration in Basmati rice have not been found to be discriminative enough warranting more precise high-throughput techniques. By make use of the DNA markers, we developed a robust DNA-based assay for detection and quantification of adulteration of Basmati varieties.

The salient features of the study are presented and discussed in this thesis.

References

Archak S, Lakshminarayanareddy V and Nagaraju J (2007) High-throughput multiplex microsatellite marker assay for detection and quantification of adulteration in Basmati rice (Oryza sativa). Electrophoresis 28: 2396-405.

Ashikari M and Matsuoka M (2006) Identification, isolation and pyramiding of quantitative trait loci for rice breeding. Trends Plant Sci 11: 344-50.

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H and Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science 309: 741-5.

Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SE and *et al* (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138: 1251-74.

Falconer DS and Mackay TFC (1996) Introduction to Quantitative Genetics. 4th edition,Longman,Harlow, UK.

Garris AJ, Tai TH, Coburn J, Kresovich S and McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. Genetics 169: 1631-8.

Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. Theor Appl Genet 74: 21-30.

Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, Yamamoto T, Lin SY, Antonio BA, Parco A, Kajiya H, Huang N, Yamamoto K, Nagamura Y, Kurata N, Khush GS and Sasaki T (1998) A high-density rice genetic linkage map with 2275 markers using a single F2 population. Genetics 148: 479-94.

Liu CJ, Witcombe JR, Pittaway TS, Nash CT, Busso CS and Gale MD (1994) An RFLP-based genetic map of pearl millet (*Pennisetum glaucum*). Thero Appl Genet 89.

McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T and Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Mol Biol 35: 89-99.

McCouch SR and Doerge RW (1995) QTL mapping in rice. Trends Genet 11: 482-7.

Nagamura Y, Antonio BA and Sasaki T (1997) Rice molecular genetic map using RFLPs and its applications Plant Mol Biol 35: 1-2: 79-87

Nagaraju J, Kathirvel M, Kumar RR, Siddiq EA and Hasnain SE (2002) Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. Proc Natl Acad Sci US 99: 5836-41.

Nene YL (1998) Basmati rice: A distinct variety (cultivar) of the Indian subcontinent. Asian Agri-History 2: 321-330.

Pham JL, Glaszmann JC, Sano R, Barbier P, Ghesquière A and Second G (1990) Isozyme markers in rice: genetic analysis and linkage relationships. Genome 33 348-359.

Pradeep M, Sarla N, Laxminarayana VR and Siddiq EA (2005) Identification and mapping of yield and yield related QTLs from an Indian accession of *Oryza rufipogon*. BMC Genetics 6: 33.

Sax K (1923) The Association of Size Differences with Seed-Coat Pattern and Pigmentation in *Phaseolus vulgaris*. Genetics 8: 552-60.

Singh RK, Singh US and Kush GS (2000) Aromatic rices. Oxford &IBH Publishing Co. Pvt. Ltd

Stuber CW and Edwards MD (1986) Genotypic selection for improvement of quantitative traits in maize using molecular marker loci In: Proceedings of the 41st Annual Maize and Sorghum Research Conference American Seed Trade Association, Chicago, Illinois, pp 40-83.

Sweeney M and McCouch S (2007) The complex history of the domestication of rice. Ann Bot (Lond) 100: 951-7.

Tanksley SD, Bernachi D and Beck-Bunn T (1998) Yield and quality evaluations on a pair of processing tomato lines nearly isogenic for the Tm2(a) gene for resistance to the tobacco mosaic virus. Euphytica 99 77-83.

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB and *et al* High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141-60.

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

Vemireddy LR, Archak S and Nagaraju J (2007) Capillary electrophoresis is essential for microsatellite marker based detection and quantification of adulteration of Basmati rice (*Oryza sativa*). J Agric Food Chem 55: 8112-7.

Virk PS, Ford-Lloyd BV, Jackson MT and Nwbury HJ (1995) Use of RAPD for the study of diversity within plant germplasm collections. Heredity 74: 170-179.

Virk PS, Newbury HJ, Jackson MT and Ford-Lloyd BV (2000) Are mapped markers more useful for assessing genetic diversity? Theor Appl Genet 100: 3-4: 607-613

Vos P (1998) AFLP fingerprinting of Arabidopsis. Methods Mol Biol 82: 147-55.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and *et al* (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-14.

Chapter I

QTL mapping of economically important traits of Basmati rice

Introduction

1.1 Introduction

Basmati rice, characterized by long slender grain, excessive elongation on cooking, exquisite aroma and soft and fluffy texture of cooked rice, is the culmination of centuries of cultivation and selection by farmers and patronisation by consumers in the north-western part of the Indian subcontinent The historical accounts and archaeological findings suggest that varieties with such unique combination of quality attributes were hardly present in the traditional rice growing areas anywhere in the world (Nagaraju *et al* 2002; Nene, 1998).

In the days of quality conscious consumers, no matter how expensive, demand for superior quality Basmati rice is on the increase in the domestic as well as international export markets. Grain quality attributes of rice may be divided into two groups: (i) grain appearance and (ii) cooking and eating qualities (Amarawathi *et al* 2008; Juliano and Villareal 1993; Wan *et al* 2004). Indices governing these attributes are determined by physico-chemical properties especially of starch and socio-cultural factors such as the history and traditions of regions where rice is grown. The grain appearance is determined by kernel length, breadth, length-breadth ratio and translucency (chalkiness). Accordingly the trade classifies them short bold, medium bold and long bold, medium slender, long slender and extra long slender types for purposes of fixing price. All these features are mainly confined to *indica* rices which cook dry and flaky because of their high amylose content (AC), high gelatinization temperature and high to medium gel consistency. In contrast, *Japonica* rices belong to only one category i.e., short and plumpy. Rice appearance quality traits are reported to be quantitatively inherited (Tan *et al* 2000).

The cooking and eating quality traits comprise kernel swelling/elongation, volume, glossiness, firmness/texture etc., determined by the physico-chemical properties of starch. Excessive elongation, aroma and taste are some of the traits of aesthetic value. As for their inheritance and breeding behavior, all the indices of grain appearance and cooking/eating qualities with the exception of amylose content have been reported to be polygenically controlled. Considering the fact that amylose content is not discrete and shows continuous variation in the segregating populations, Puri and Siddiq (1980) concluded it to be a polygenic trait, some still maintain it to be controlled by a single major gene in association with modifier complex (Bollich and Webb, 1973; He *et al* 1999; Kumar and Kush 1988; McKenzie and Rutger 1983). Pooni *et al* (1993) report AC to exhibit a complex genetics due to the triploid nature of

endosperm that results in additional cytoplasmic and epistatic effects. Tan *et al* (1999) reported AC to be governed by waxy locus (*wx*) and mapped to chromosome 6.

One of the prominent characters of basmati for which it has acquired its name and high price is aroma. The typical Basmati aroma is evident at all stages of crop growth in almost all aerial parts of the plant (Lorieux *et al* 1996; Sood and Siddiq 1978). Development and retention of aroma is known to be influenced by genetic and environmental factors. Aroma develops from a cocktail of more than 100 compounds that include hydrocarbons, alcohols, aldehydes and esters. However, one key fragrance causing chemical component reported in Basmati, Jasmine rice and Pandanus leaves is 2-acetyl-1-pyrroline (2-AP) (Buttery *et al* 1983). Many studies on the genetic control of the aroma trait in rice have been carried out (Ahn *et al* 1992; Bradbury *et al* 2005; Sood and Siddiq 1978), most of which suggest that rice fragrance is controlled by a recessive gene. Recently, a gene for fragrance (*fgr*) was identified on chromosome 8 which has homology to the gene that encodes betain aldehyde dehydrogenase 2 (*bad-2*) (Bradbury *et al* 2005). Accumulation of 2-AP in aromatic rices was explained by the presence of mutations caused either by 8 bp deletion and three SNPs in the exon 7 (Bradbury *et al* 2005) or 7 bp deletion in the exon 2 (Shi *et al* 2008) of *bad-2* on chromosome 8 resulting in a loss of function of the *fgr* gene.

Deviation of observed frequencies of individuals in a given genotypic class from their expected Mendelian frequencies within a segregating population has been defined as segregation distortion (Sandler and Novitski 1957; Taylor and Ingvarsson 2003). The phenomenon of segregation distortion has been frequently reported in wide crosses in many crop species (Kumar *et al* 2007; Lambridges *et al* 2004; Li *et al* 2007; Liu *et al* 2000). In the case of rices such deviations have been reported in inter sub-specific (Xu *et al* 1995) and inter-specific (Aluko *et al* 2004; Pradeep *et al* 2005) crosses. The genetic basis of segregation distortion in rice was first reported to be associated with the occurrence of a gametophytic gene (*ga-1*) on chromosome 6 (Iwata *et al* 1964) and of *ga-2* and *ga-3* on chromosome 3 by Nakagahra (1972). Lin *et al* (1992) attribute this phenomenon to abortion of male and female gametes or selective fertilization in certain inter-specific or inter-sub specific crosses. The work of Nakagahra (1972) has led to the suggestion that the gametophytic loci are responsible for the partial or total elimination of gametes carrying one of the parental alleles. Segregation distortion at a marker

locus would occur as a result of linkage between the marker and the gametophyte gene (*ga*) which would confer low pollinating ability.

Breeding for Basmati quality in high yield background requires sound knowledge of the genes governing the key traits such as quality and yield. Most of these traits as mentioned earlier are known to be polygenically controlled. Breeding/selection for desired one or more quality traits using conventional screening methods is a time consuming and difficult task for breeders and are often unreliable, warranting development and use of novel molecular marker tools. Though not extensive, limited information is available on molecular mapping of genes/QTLs governing grain characteristics of rice. Ahn *et al* (1992, 1993) identified linkage of the *fgr* gene with the RFLP marker RG28 on chromosome 8 at a genetic distance of 4.5 cM. In addition, one QTL has been identified on the same chromosome mapped at a distance of 14.6 cM using two RFLP markers viz., RZ323 and RZ562 for kernel elongation. Subsequently, Lorieux *et al* (1996) confirmed the linkage of the *fgr* gene with the marker RG28. And, they also identified additional two QTLs for aroma, one on chromosome 4 and the other on chromosome 12. The *fgr* gene was characterized recently and found to be homologous to the gene that encodes betain aldehyde dehydrogenase 2 (*bad*-2) (Bradbury *et al* 2005). Recently, Amarawathi *et al* (2008) have identified a few Basmati quality trait related QTLs in Pusa1121 that include three for grain length (L) on chromosomes 1 and 7, two each for grain breadth (B) and LB ratio on chromosome 7, three for aroma on chromosomes 3, 4 and 8 and one each for elongation ratio on chromosome 11, amylose content and alkali spreading value mapped on to chromosome 6.

The present study was initiated to identify, map and characterize the QTLs linked to economically important agronomic and quality traits of Basmati rice. Tightly linked molecular markers, if any, found for these traits could be used in marker assisted selection (MAS) programmes for further improvement of Basmati quality rices. Besides, the tightly linked markers can be used to characterize the genes controlling the traits by making use of the rice whole genome sequence.

Materials and Methods

1.2 Materials and Methods

Choice of parents

The traditional Basmati variety, Basmati370 and the semi dwarf non-Basmati variety, Jaya were chosen as parents for developing the mapping population for the following reasons.

- Traditional basmati varieties known by different names in the subcontinent, in all likelihood, are derivatives of the single local variety i.e., Basmati370 or Basmati370-like variety (Nagaraju *et al* 2002).
- Most of the basmati varieties released as elite basmati varieties since 1965 from India (12 of 19) and Pakistan (4 of 5) have Basmati370 as one of the donor strains in the breeding programmes.
- Genetic diversity study employing ISSRs and SSRs reveals that the high yielding variety Jaya to be genetically quite distinct from Basmati370 (Nagaraju *et al* 2002).
- Jaya and Basmati varieties possess distinct and contrasting physico-chemical characters.

Development of mapping population

The crosses of Basmati370 and Jaya were made during *Rabi,* 2002. One hundred F_1 seeds were used to raise F_2 generation during *Kharif,* 2003 (Figure 1.1). The F_2 population was grown along with F_1s and the parents in wetland farm of the Agricultural Research Institute (ARI), Rajendranagar, Hyderabad. Hybrid nature of F_1s was confirmed using microsatellite markers that are polymorphic between parents (Figure 1.2). Out of 10,000 F_2 plants, 181 were randomly chosen as mapping population. The phenotypic measurements of the nine agronomic and nine quality traits were carried out for the 181 F_2 offspring. The standard procedure followed for recording the observations of agronomic traits is as under:

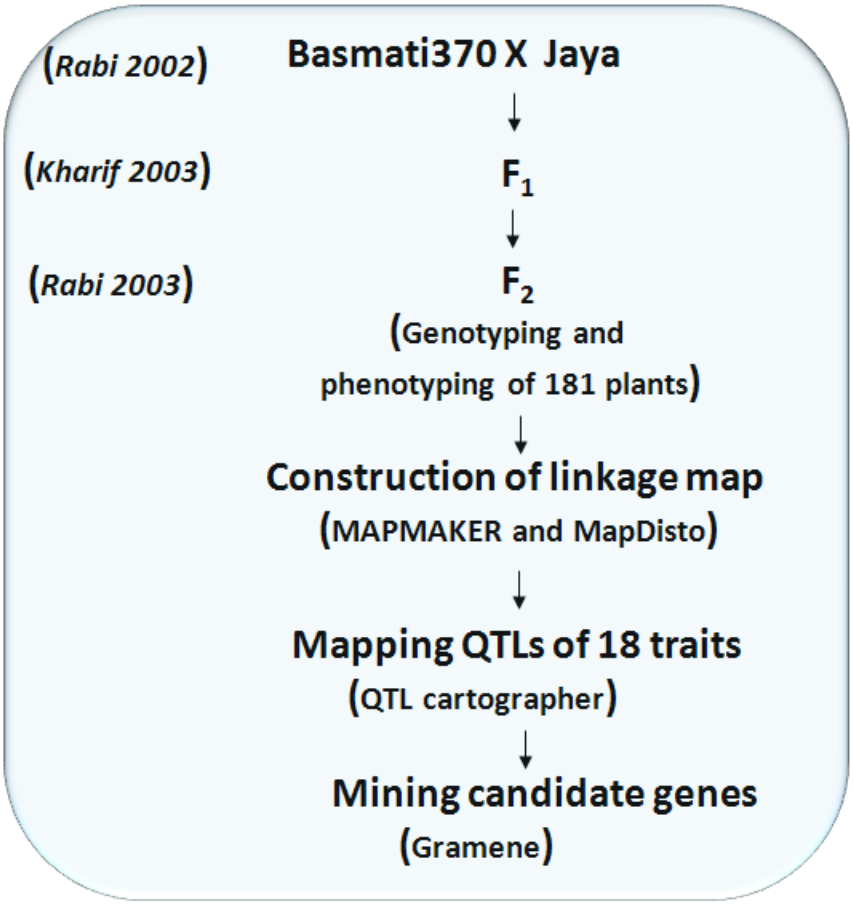
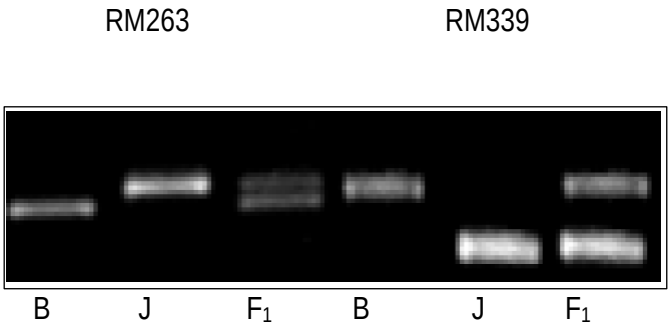


Figure 1.1 Development of the mapping population and QTL mapping



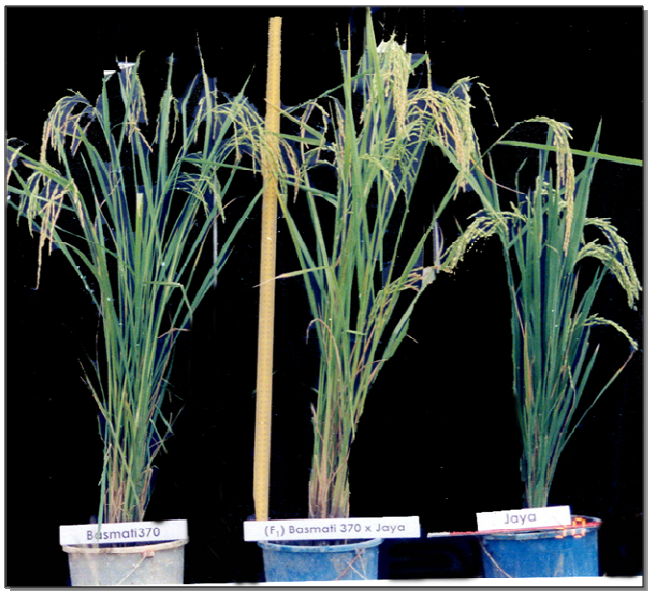
B - Basmati370; J - Jaya and F_1 – Hybrid

Figure 1.2 Confirmation of heterozygosity of the F_1 hybrid of Basmati370/Jaya. RM263 and RM339 are informative or polymorphic microsatellite markers between Basmati370 and Jaya

Plant height (PH) **-** Length of the tallest tiller from ground level to the tip of the panicle (Figure 1.3)

Number of panicles (NP) **-** Number of ear bearing tillers per plant

Panicle length (PL) **-** Length in cm from neck to the tip of the panicle excluding awn (Figure 1.4)



Basamti370 F_1 (Basmati370 x Jaya) Jaya

Figure 1.3 Plant types of Basmati370, F_1 and Jaya

Jaya F_1 Basmati370

Figure 1.4 Panicle lengths of Jaya, F_1 and Basmati370

Spikelet number (SN) **-** Number of spikelets including empty and filled ones per panicle averaged over 4-5 panicles

Filled grains (FG) **-** Number of filled spikelets per panicle averaged over 4-5 panicles

Chaffy grains (CG) **-** Number of sterile spikelets or chaffy grains per panicle averaged over 4-5 panicles

Spikelet fertility (SF) **-** Ratio of filled spikelets to the total number of filled and chaffy spikelets per panicle, expressed in percentage

Grain weight (GW) **-** Weight in grams of 1000 filled spikelets

Single plant yield (SPY) **-** Weight in grams of total filled grains per plant

After maturation, the grains were harvested and stored at room temperature for at least 3 months before processing. The analysis of quality traits was carried out at Directorate of Rice Research (DRR), Hyderabad. Hulls were removed from 50 g of rough rice from each of F2 plants using a Model TH035A Satake Huller (Houston, TX) to obtain brown rice. Embryos and the bran layers were removed (polished) from brown rice using McGill miller, model #1, (Phillip Rahm International). The standard procedures were followed for recording data of quality traits as given under:

Grain length (GL) and grain breadth (GB)

The length and breadth was measured using grain shape tester or dial micrometer for a minimum of 10 full rice grains with both the tips are intact from each of the F_2 plants (Figure 1.5).

Grain length-grain breadth (LB ratio)

The grain length-grain breadth ratio (LB ratio) was calculated as the grain length divided by grain breadth.

Chalkiness

Ten whole grains from each of the F_2 plants were placed on light box for scoring chalkiness. Grains with 50% or more chalk content were identified (Figure 1.6). Degree of chalkiness was determined by adopting the following scale according to increasing intensity:

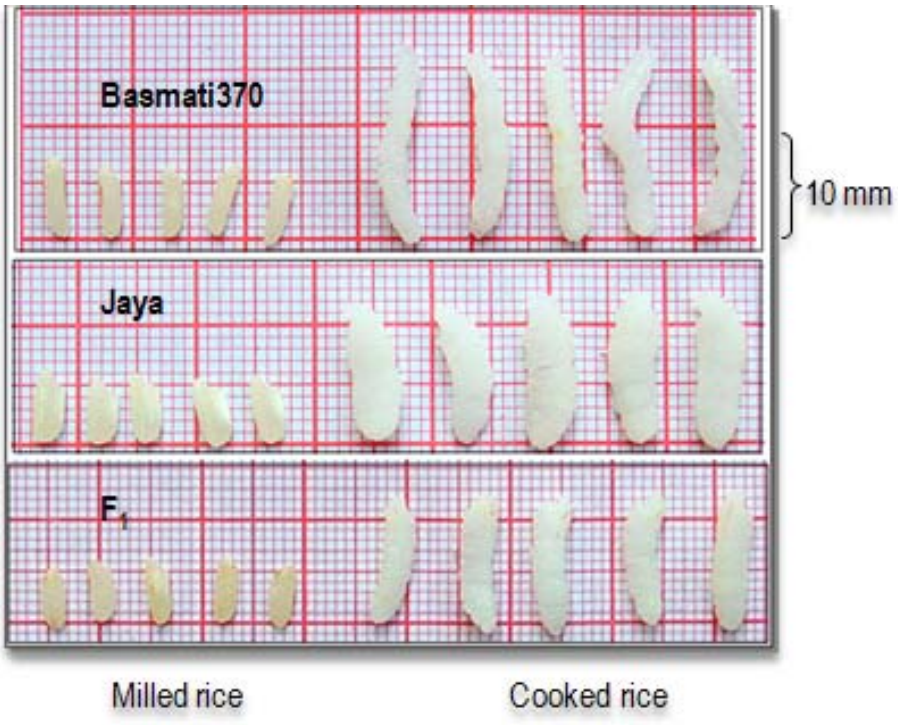


Figure 1.5 Grain appearance traits of Basmati370, Jaya and F_1 before and after cooking

Figure 1.6 Grain chalkiness of Basmati370, Jaya and F_1

Scale	Score	Percent of grain area
Absent (A)	0	0
Very occasionally present (VOC)	1	<10
Occasionally present (OC)	3	10-30
Present (P)	5	>30%
Fully present (FP)	9	100%

Grain length after elongation (GLAC) and elongation ratio (ER)

Kernels of rice varieties expand either breadthwise or lengthwise upon cooking (Figure 1.5). Lengthwise expansion with least increase in girth is considered a desirable trait from the angle of consumers' preference. For instance, Basmati rices of India and Pakistan, Bahra of Afghanistan, Domsiah of Iran, Bashful of Bangladesh, and D25-4 of Myanmar elongate 100% upon cooking. The F_2 offspring were measured for the degree of elongation using the method of Azeez and Shafi (1966). The elongation test consisted of soaking of 25 whole milled kernels in 20 mL of distilled water and subsequently placing them in water bath at 98°C for 10 min. The cooked rice was then transferred to a petri dish lined with filter paper. Ten cooked whole grains were selected and length was measured by placing them on graph paper. The elongation, measured as the ratio of the average length of cooked rice kernels to the average length of uncooked rice kernels, was scored.

Aroma

The presence of aroma from the rice leaf was evaluated by following the method developed by Sood and Siddiq (1978). The method consisted of adding ~ 10mL of 1.7% KOH solution to small petri dishes containing about 2g finely minced sample of green leaf or stem. The petri dishes were covered immediately after addition of alkali and left at room temperature for about 10 min. The plates were then opened and immediately smelled. The samples are scored as strongly aromatic, moderately aromatic and non-aromatic depending on the intensity of aroma. A strongly scented variety, Basmati370 and a non-scented variety Jaya were used as checks for scoring for aroma in F_2 individuals.

Alkali Spreading Value (ASV)/Gelatinization temperature (GT)

The method of Little *et al* (1958) was used for conducting the alkali spreading test. A duplicate set of six whole-milled grains without cracks was selected and placed in a plastic box (5cm x 5cm x 1.9cm) containing 1.7 % KOH solution at 29^{o}C for 23 hrs. Grains were carefully separated using forceps, and the spreading value of the grains was scored by visual assessment as given below:

Score	Spreading scale	Clearing scale
1	Kernel not affected	Kernel chalky
2	Kernel swollen	Kernel chalky, collar powdery
3	Kernel swollen, collar incomplete narrow	Kernel chalky, collar cottony or cloudy
4	Kernel swollen, collar complete and wide	Centre cottony, collar cloudy
5	Kernel split or segmented, collar complete and wide	Centre cottony, collar clearing
6	Kernel dispersed, merging with collar	Centre cottony, collar, cleared
7	All kernel dispersed and inter mingled	Centre and collar cleared

Gelatinization temperature was estimated by the extent of alkali spreading value as given below:

Classification	Alkali Spreading value	Gelatinization temperature
1-2	Low	High>74^{o}C
3	Low, Intermediate	High, Intermediate
4-5	Intermediate	Intermediate (70-74 ^{o}C)
6-7	High	Low (55-69 ^{o}C)

Amylose content (AC)

Many of the cooking and eating characteristics of milled rice are influenced by the ratio of two fractions of starch viz., amylose and amylopectin, in the rice grain. Rice varieties are grouped on the basis of their AC into the following four types (Kumar and Kush 1988):

Classification	AC (%)
waxy	0–2
very low	3–9
Low	10– 19
intermediate	20–25
high	>25

The simplified procedure of Juliano (1971) detailed as under was used for estimation of AC. Twenty whole-milled rice kernels were ground in an Udy cyclone mill (sieve mesh size 60), 100 mg of rice powder taken in a 100 mL volumetric flask was added with 1mL of 95% ethanol and 9 mL of 1 M NaOH. The content was heated over a boiling water bath to gelatinize the starch. After cooling for 1 hr distilled water was added and the contents were thoroughly mixed. For each set of samples run, low-, intermediate-, and high-amylose standard varieties were included to serve as checks. Five mL of the starch solution was taken in a 100 mL volumetric flask. One mL of 1 M acetic acid and 2 mL of iodine solution (0.2 g iodine and 2.0 g potassium iodide in 100 mL aqueous solution) were added and the volume was made up with distilled water. The contents were shaken well and left to stand still for 20 min. Absorbance of the solution was measured at 620 nm with a spectrophotometer (Bausch and Lomb Spectronic 20). Amylose content was determined by using a conversion factor and the results were expressed on a dry weight basis. Moisture content of the samples was not determined as the relative humidity and temperature of the laboratory was controlled. For making standard curve, 40 mg of potato amylose (Sigma Chemical Co. or Stein Hall Co., Inc.) of known moisture content was wetted with 1 mL of ethanol and 9 mL of 1 M NaOH, heated for 5–10 min over a boiling water bath, cooled, and made up to 100 mL with distilled water. Solution (1, 2, 3, 4, 5 mL) was placed with a pipette in 100 mL volumetric flasks. The solution was acidified with 1 M acetic acid (0.2, 0.4, 0.6, 0.8, and 1.0 mL, respectively) and treated as above. The absorbance values were plotted at 620 nm against the concentration of anhydrous amylose (mg) and the conversion factor was determined. The dilution factor of 20 for the samples was included in the conversion factor.

Phenotypic data analysis of parents, F_1 and F_2 individuals

Correlations between character pairs and test for normal distribution were computed at $p < 0.05$ and $p < 0.01$ in Microsoft-Excel (2007). Heterosis, heterobeltiosis and inbreeding depression were calculated using the following formulae.

Heterosis= [(F_1-MP)/MP] x 100

Heterobeltiosis= [(F_1-BP)/BP] x 100

Inbreeding depression= [(F_1- F_2)/ F_1] x 100

Where, MP- Mid parent and BP- Better parent

Test of significances among parents, F_1 and mid parental value were calculated employing StatPlus v 4.6 software (www.analystsoft.com/en).

Genomic DNA extraction and quantification

DNA from the two parents, F_1 and 181 F_2 individual plants was extracted from leaves by using the modified CTAB method described as under:

Steps followed:

- Approximately 1 g of fresh leaf tissue was ground to fine powder in liquid nitrogen using a pestle and mortar.
- The powder was quickly added with 20 mL of pre-warmed (65^{0}C) 2X extraction buffer (2% CTAB, 200 mm Tris–HCl p^H 8.0, 20 mM EDTA; 1.4M NaCl; 1% PVP and distilled water) in a capped polypropylene or oakridge tubes. The clumps were suspended with a spatula, and incubated for 1 hr with frequent stirring in a water bath at 65^{o}C.
- 20 mL of Chloroform: Iso-amyl alcohol (24:1) mixture was added and kept for shaking for10 min.
- It was centrifuged at 13000 rpm for 15 min and supernatant was collected into a new polypropylene tube. This step was repeated till a clear supernatant was obtained.
- DNA was precipitated by addition of 0.7 volume of ice-cold 100% Isopropanol followed by gentle mixing.
- It was then incubated for 1 hr at -20 ^{0}C followed by centrifugation at 13000 rpm for 15 min at 4^{0}C. The resulted supernatant was discarded and the DNA pellet was air dried. The pellet was washed with 5 mL of 70% ethanol and spun again at 13000 rpm for 5 min at 4 ^{0}C and was air dried again.
- The DNA pellet was dissolved in 200 µL of TE buffer (10 MM Tri-HCl, 1mM EDTA p^H 8.0) over night at room temperature followed by incubation for 1 hr at 65 ^{0}C.
- The DNA was transferred from polypropylene tube to eppendorf tube using wide-mouthed glass pipette or half cut tip (to avoid DNA shearing) 5 µL RNAase (20mg/mL) (pre-boiled) was added to the DNA solution and incubated at 37^0 C for an hour.

- The DNA was reprecipitated by adding 0.1 volume of 5M ammonium acetate and two volumes of 100% ethanol followed by gentle mixing and then stored at -20^0 C for an hour.
- Samples were centrifuged at 13000 rpm for 10 min and the supernatant was drained out.
- 100 μL of 70% ethanol was added to the DNA pellet and centrifuged at 13000 rpm for 10 min. Supernatant was drained carefully and the DNA pellet was air dried over night.
- The DNA was dissolved in 200 μL of TE buffer at 65^0 C for one hour.
- The DNA was quantified at 260 nm of UV absorbance in spectrophotometer.

An aliquot of DNA sample was diluted in 1x TE buffer (10 mm Tri-HCl, 1 mM EDTA p^H 8.0) in a ratio of 1:100 in a micro cuvette.

PCR amplification

PCR amplification was performed in a 10 μl volume containing 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.5 unit of Taq polymerase, 50 μM of dNTPs, and 0.1 μM of each primer with 10 ng of genomic DNA on a Thermal Cycler (PE9700) with a Ramp speed of 9700 (Applied Biosystems,USA). The PCR products were run on a 3% agarose gel (Sigma-routine agarose) or 3% metaphor agarose (Sigma-high resolution) mixed with ethidium bromide or in 0.5X tris-acetic acid-EDTA (TAE) buffer (242 g Tris base, 57.1 mL Acetic acid 100mL 0.5M EDTA mixed and made up the volume to 1 litre with double distilled water and pH adjusted to 8.5). The resolved PCR bands were detected using Bio-Rad Molecular Imager Gel Doc XR System.

Agarose gel electrophoresis

PCR samples were mixed with bromo-phenol blue (0.25% bromo phenol blue and 40% (w/v) sucrose mixed in water) and run on a 3% agarose gel (Sigma) containing ethidium bromide (10mg/mL) along with the marker Lambda Hind III digest (MBI Fermentas) for genomic DNA or 50 bp ladder (MBI Fermentas) for PCR product at 5.3 V/cm (Bio-Rad PowerPac 300) for an hour. Gels were photographed using Bio-Rad Molecular Imager Gel Doc XR System.

Parental polymorphism, linkage map construction and QTL detection

A set of 502 SSR (simple sequence repeat) markers spanning all the 12 rice chromosomes was screened between Basmati370 and Jaya strains (Figure 1.7). 134 markers that are polymorphic between parents were used for screening 181 F_2 individuals (Figure 1.8) and tested the $\chi 2$ goodness of fit against 1:2:1 segregation ratio using MapDisto software (Lorieux 2006).

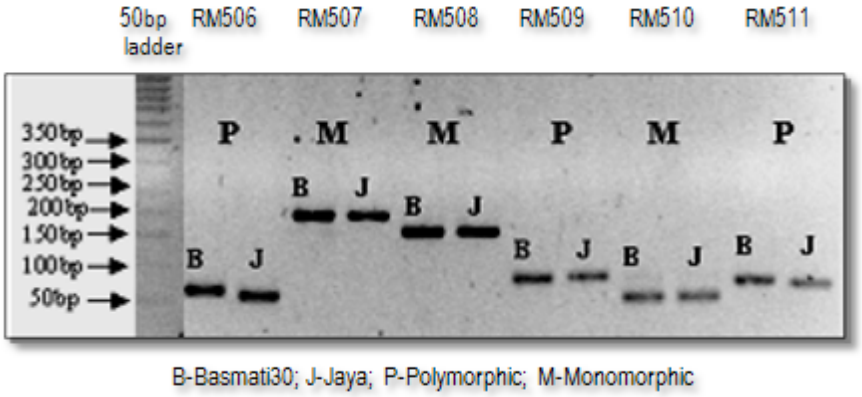


Figure 1.7 Polymorphism between Basmati370 and Jaya. Parents were amplified with rice microsatellite markers and run on a 3% agarose gel for screening parental polymorphism.

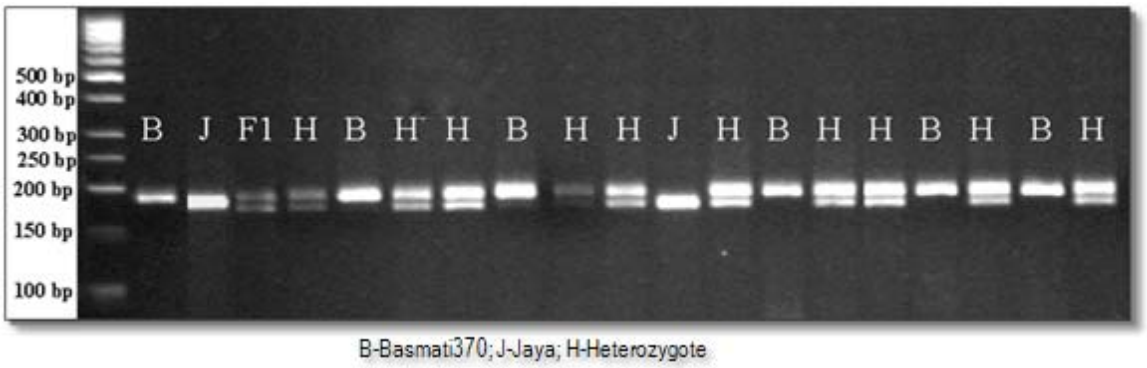


Figure 1.8 Segregation pattern of RM17 locus in F_2 population of cross between Basmati370 and Jaya

Linkage map was constructed using the 134 markers that are polymorphic between parents were used for screening 181 F_2 individuals (Figure 1.8) and tested the χ2 goodness of fit against 1:2:1 segregation ratio using MapDisto software (Lorieux 2006). Linkage map was constructed using the MAPMAKER version 3.0 (Lincoln *et al* 1993) following Kosambi Function. Linkage groups were determined using 'group' command with LOD (Logarithm of odds ratio) score of 3.0 and a recombination fraction of 0.4.

Order of the markers for each group was determined using 'order' and 'ripple' commands. Assignment of linkage groups to the respective chromosomes was done based on the rice maps developed at Cornell University (McCouch *et al* 2002). QTLs were detected by interval (Lander and Botstein 1989) and composite interval mapping procedures of Windows QTL Cartographer v.2.5 software (Wang *et al* 2006b). Composite interval mapping (Zeng 1994) was conducted using the default settings (e.g., Model 6, five cofactors selected automatically by forward regression with a 10-cM window) (http://statgen.ncsu.edu/qtlcart/cartographer.html). Permutation tests (Churchill and Doerge 1994) were used to determine chromosome-wise significance thresholds (α=0.05) for QTL detection using QTL Cartographer with 1000 shuffles of the trait data. Quantitative trait loci peaks with significant likelihood ratio test statistics under H3/H0 were recorded.

These two hypotheses are:

H0: $a = 0$, $d = 0$

H3: $a \neq 0$, $d \neq 0$

Where a = additive effect and d = dominance effect.

Conversion factor: LOD = 0.217 x LR (likelihood ratio)

Results

1.3 Results

Phenotypic trait measurements and frequency distribution

The parents used for mapping, Basmati370 and Jaya differed significantly in respect of all the traits studied except for panicle length, chaffy grains, spikelet fertility and single plant yield (Table 1.1). The mean of the F_1 hybrids was intermediate for panicle length, 1000 seed weight, grain length (L), grain breadth (B), and LB ratio, alkali spreading value, amylose content and aroma. In all the cases, F_2 mean was intermediate to the parental values except in 1000 seed weight, chaffy grains, spikelet number and elongation ratio for which the mean was more than that of parents. In case of single plant yield, spikelet fertility and chalkiness, the means were less than the parents. Plant height ranging from 84.98cm to 114.79cm with a mean of 112.85cm was well above the mid parent value of 99.89cm. The same was the case with characters such as panicle length, chaffy grains, spikelet number, 1000 seed weight, grain breadth, grain length after cooking and elongation ratio. On the contrary, F_2 mean values for number of panicles, filled grains, spikelet fertility, single plant yield, grain length, length-breadth ratio, alkali spreading value, amylose content, aroma and chalkiness were well below their respective mid parental values.

Except aroma, all the scored agronomic and quality traits showed transgressive segregation ranging between 3 and 100%. Such segregants would be of value for improvement of these traits (Table 1.2 and Figures 1.9 and 1.10). As for aroma measured on 1-9 scale with the parents scoring the extremes of the scales it was not possible to get transgressive segregants for this trait. However, in case of spikelet fertility, the mean value of F_2 (71.55%), was quite low as compared to that of parents and F_1 and as a result all the F_2 plants fell below the parental average leading to 100% transgressive segregants. Transgressive segregants observed for traits such as panicle length, filled grains, spikelet number, spikelet fertility, single plant yield and grain length significantly exceeded both the parents. However, in the case of plant height, grain length, elongation ratio, alkali spreading value and amylose content transgressive segregants exceeded only Basmati370 whereas number of panicles, chaffy grains and seed weight exceeded Jaya parent. However, the number of transgressive segregants in respect of grain breadth, length-breadth ratio and chalkiness did not significantly exceed both the parents.

Table 1.1 Test of significance among parents, F_1 and mid parental values for 18 traits studied

S.No.	Trait	Code	Basmati370 (B) (n=10)	Jaya (J) (n=10)	F_1 (n=10)	F_2 (n=181)	MP (n=10)	B/J (n=10)	MP/F_1 (n=10)	B/F_1 (n=10)	J/F_1 (n=10)
1	Plant height (cm)	*PH*	114.79 ± 0.39	84.98 ± 4.65	120.25 ± 2.06	112.85	99.85	**	*	**	**
2	No. of panicles	*NP*	12.57 ± 3.64	8 ± 1.10	15 ± 2.94	7.36	10.285	*	**	NS	**
3	Panicle length (cm)	*PL*	25.29 ± 2.66	23.33 ± 4.02	24.88 ± 1.03	24.53	24.31	NS	NS	NS	NS
4	Filled grains (no.)	*FG*	75.50 ± 4.12	109.25 ± 4.65	167 ± 4.24	90.43	92.38	**	**	**	**
5	Chaffy grains (no.)	*CG*	4.86 ± 1.68	7.67 ± 4.50	20.50 ± 3.54	37.65	6.27	NS	**	**	*
6	Spikelet number	*SN*	80.25 ± 4.79	116.75 ± 0.50	187.5 ± 0.71	128.11	98.5	**	**	**	**
7	Spikelet fertility (%)	*SF*	94.13 ± 2.70	93.58 ± 4.09	89.06 ± 1.93	71.55	93.9	NS	NS	NS	NS
8	1000 Seed weight (g)	*SW*	18.2 ± 2.27	23.65 ± 1.25	22.53 ± 1.49	24.9	20.93	**	NS	*	NS
9	Single plant yield (g)	*SPY*	14.19 ± 4.78	17.10 ± 1.10	27.96 ± 1.41	13.58	15.64	NS	**	**	**
10	Grain length (mm)	*GL*	6.49 ± 0.27	5.95 ± 0.37	6.24 ± 0.18	6.09	6.22	**	NS	*	*
11	Grain breadth (mm)	*GB*	1.82 ± 0.05	2.53 ± 0.11	2.20 ± 0.05	2.22	2.12	**	NS	**	**
12	Length-Breadth ratio	*LB*	3.57 ± 0.17	2.36 ± 0.18	2.84 ± 0.07	2.76	2.97	**	**	**	**
13	Grain length after cooking (mm)	*GLAC*	15.1 ± 0.57	9.88 ± 0.83	15.6 ± 0.84	14.48	12.5	**	**	NS	**
14	Elongation ratio	*ER*	2.33 ± 0.17	1.68 ± 0.17	2.5 ± 0.15	2.38	2	**	**	*	**
15	Alkali spreading value	*ASV*	5± 0.00	7 ± 0.00	6.0 ± 1.05	5.75	6	**	NS	**	NS
16	Amylose content (%)	*AC*	21.03 ± 0.37	26.79 ± 0.29	22.8 ± 1.25	22.52	23.9	**	NS	NS	**
17	Aroma	*ARM*	9± 0.00	1± 0.00	2.00 ± 1.05	4.66	5	**	**	**	*
18	Chalkiness	*CHK*	1.80 ± 1.03	3 ± 1.63	1.60 ± 0.97	1.21	2	**	NS	NS	**

P<0.05 (*); p<0.001 (**); NS = Non-significant; MP- Mid parental value

Table 1.2 Transgressive segregants, heterosis, heterobeltiosis and inbreeding depression for 18 traits in the F_2 population

S.No.	Trait	Transgressive segregants (%)	Heterosis (%)	Heterobeltiosis (%)	Inbreeding depression (%)
1	Plant height (cm)	49	20.39	4.76	6.15
2	No. of panicles	72	45.84	19.33	50.93
3	Panicle length (cm)	77	2.34	-1.62	1.41
4	Filled grains (no.)	65	80.78	52.86	45.85
5	Chaffy grains (no.)	93	227.21	167.28	-83.66
6	Spikelet number*	69	90.36	60.60	31.67
7	Spikelet fertility* (%)	100	-5.11	-5.39	19.66
8	1000 Seed weight (g)	68	7.67	-4.74	-10.52
9	Single plant yield (g)	87	78.72	63.51	51.43
10	Grain length (Calin *et al.*)	44	0.32	-3.85	2.40
11	Grain breadth (Calin *et al.*)	3	1.15	-13.04	-0.91
12	Length-Breadth ratio*	9	-4.22	-20.45	2.82
13	Grain length after cooking (Calin *et al.*)	33	24.9	3.31	7.18
14	Elongation ratio*	62	24.69	7.30	4.80
15	Alkali spreading value	28	0	-14.29	4.17
16	Amylose content (%)	15	-4.64	-14.89	1.23
17	Aroma	0	0	-44.44	6.80
18	Chalkiness	6	50	0.00	59.67

* Derived characters

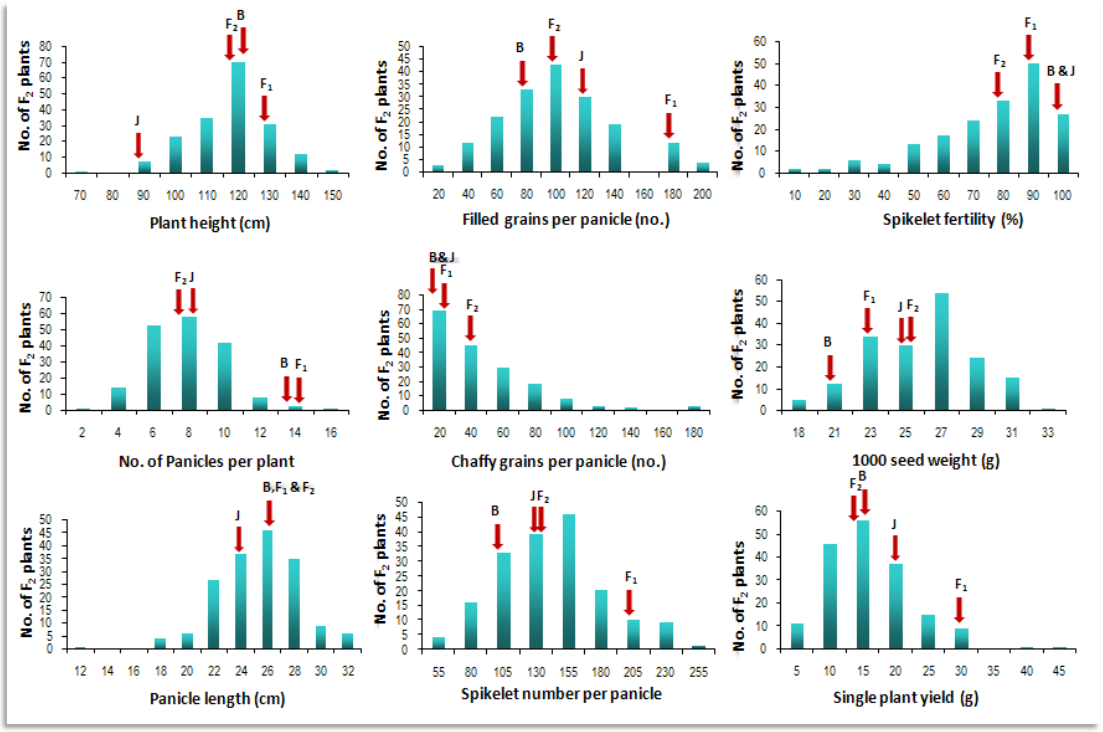


Figure 1.9 Phenotypic distributions of agronomic traits in 181 F_2 offspring

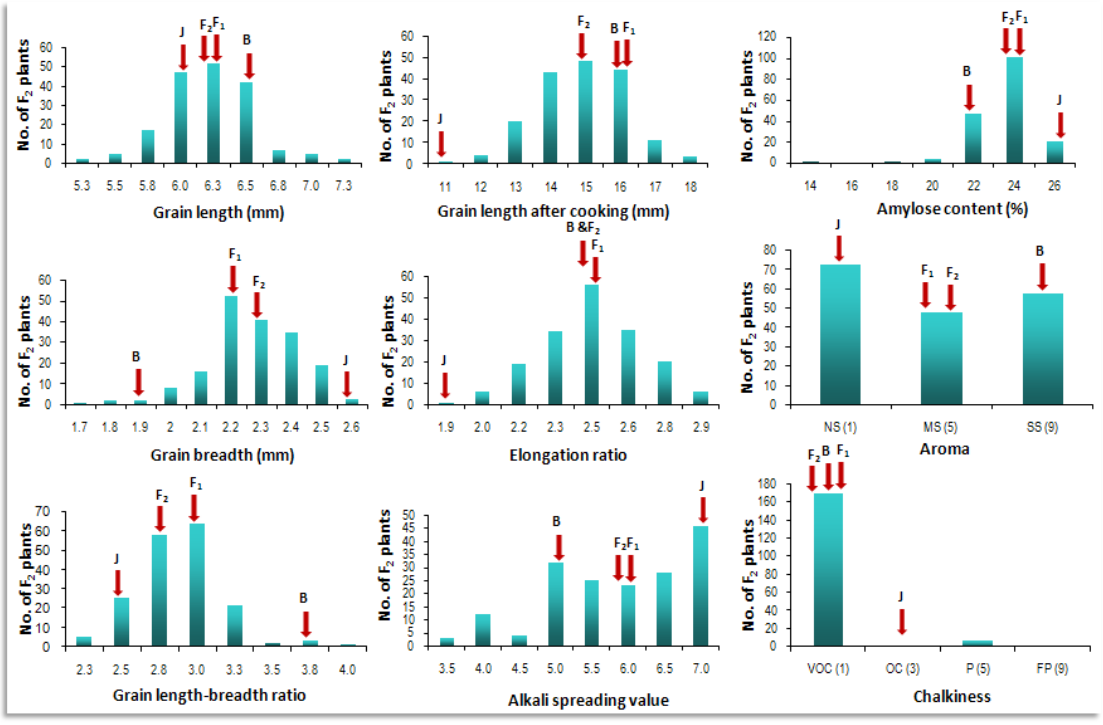


Figure 1.10 Phenotypic distributions of quality traits in 181 F_2 offspring

All the traits, except chaffy grains, spikelet fertility, alkali spreading value, aroma, chalkiness and amylose content showed normal distribution in F_2 (Figures 1.9 and 1.10). As expected, chaffy grains and spikelet fertility skewed towards the lowest and highest values respectively. Amylose content and chalkiness showed unimodal distribution suggesting these traits may be controlled by few genes.

F_1 hybrids, in all cases, exhibited a wide range of heterosis the range being from -5.11% to 227.21%. However, significant, positive and beneficial heterosis was observed only for a few traits such as spikelet number, filled grains, single plant yield, number of panicles, grain length after cooking, elongation ratio and plant height (Table 1.2).

Trait correlations

All the agronomic traits except chaffy grains and 1000 seed weight showed significant correlation with single plant yield. Plant height also showed highly significant positive correlations with all the traits except number of panicles, chaffy grains and 1000 seed weight (Table 1.3). Panicle length showed highly significant and positive correlation with filled grains and spikelet number. Filled grains also showed positive correlation with spikelet number and spikelet fertility. As expected, chaffy grains and spikelet fertility showed highly significant negative correlation. Interestingly, in the present study, 1000 seed weight, aroma, alkali spreading value and amylose content were found to be correlated with none of the traits studied. Chalkiness significantly and positively correlated with grain breadth whereas negatively correlated with length-breadth ratio. Surprisingly, grain length after cooking showed significant and positive correlation with filled grains and grain length while less significant correlation was observed with spikelet number, seed weight, single plant yield and length- breadth ratio. As expected, grain length showed negative correlation with grain breadth and significant and positive correlation with grain length after cooking and length-breadth ratio (Table 1.3).

Parental polymorphism

A total of 552 microsatellite markers was used for screening of parents i.e., Basmati370 and Jaya. Of 552 markers, 482 gave clearly scorable amplification products (87.32%), 30 (5.34%) did not amplify (null alleles) either Basmati370 or Jaya, 15 (2.72%) could not be scored on agarose

Table 1.3 Correlation coefficients among 18 traits of F_2 population derived from the Basmati370 and Jaya cross

Trait	PH	NP	PL	FG	CG	SN	SF	SW	SPY	GL	GB	LB	GLAC	ER	ASV	AC	ARM	CHK
PH	1.000																	
NP	0.075	1.000																
PL	0.454**	-0.026	1.000															
FG	0.395**	-0.006	0.557**	1.000														
CG	-0.046	0.054	0.090	-0.313**	1.000													
SN	0.316**	0.038	0.579**	0.643**	0.524**	1.000												
SF	0.208*	-0.072	0.111	0.617**	-0.874**	-0.151	1.000											
SW	0.143	-0.059	0.158	0.066	0.132	0.166	-0.062	1.000										
SPY	0.345**	0.629**	0.267**	0.502**	-0.069	0.394**	0.237*	0.131	1.000									
GL	0.065	0.081	0.100	-0.033	0.032	-0.002	-0.029	0.223*	0.129	1.000								
GB	0.063	0.020	0.051	0.038	0.177	0.172	-0.081	0.380**	0.110	-0.266**	1.000							
LB	-0.003	0.031	0.017	-0.024	-0.110	-0.106	0.052	-0.149	0.006	0.714**	-0.856**	1.000						
GLAC	0.039	0.165	0.145	0.261**	0.008	0.238*	0.065	0.227*	0.245*	0.402**	-0.028	0.236*	1.000					
ER	0.004	0.113	0.088	0.315**	-0.014	0.268**	0.092	0.097	0.174	-0.205**	0.146	-0.203*	0.811**	1.000				
ASV	-0.008	-0.114	0.181	0.009	0.009	0.014	0.019	0.029	-0.067	-0.074	-0.063	-0.008	-0.118	-0.082	1.000			
AC	0.123	-0.056	0.038	0.102	-0.188	-0.051	0.170	0.000	0.051	-0.030	-0.016	0.001	-0.076	-0.051	0.149	1.000		
ARM	0.130	0.115	0.063	-0.100	0.028	-0.07	-0.080	-0.082	0.100	0.099	-0.070	0.0979	0.034	-0.020	-0.030	-0.080	1.000	
CHK	-0.00	0.041	-0.040	0.057	0.020	0.066	-0.000	0.1844	0.119	-0.070	0.341**	-0.27**	0.129	0.173	-0.210	-0.060	-0.140	1.000

** Significant at p=0.01 * Significant at p=0.05 ; For trait codes refer Table 1.1

gel and 25 (4.53%) exhibited heterozygosity in either Basmati370 or Jaya. A total of 203 markers showed polymorphism (42.12%) between the parents indicating the possibilities of constructing a linkage map of Basmati370/Jaya. Polymorphic markers were distributed on all the 12 chromosomes (Table 1.4). Chromosome 1 showed maximum number (35) of polymorphic markers.

Table 1.4 Chromosome-wise distribution of microsatellite markers polymorphic between Basmat370 and Jaya

Chromosome	Polymorphic markers	Chromosome	Polymorphic markers
1	35	7	11
2	28	8	17
3	21	9	18
4	13	10	9
5	16	11	8
6	19	12	8

Linkage map construction

Segregation of marker loci

Of 203 polymorphic loci, 60 markers which could not be scored were excluded from screening F_2 population. Nine markers were found to be unlinked. The remaining 134 markers were used for construction of genetic linkage map. These 134 markers included 129 rice microsatellite markers (www.gramene.org), two markers from the waxy gene (MX4 and WXSSR), two markers flanking the major QTL for grain length (RM353w and JL14) and one gene (*fgr*) specific STS (sequence tagged site) marker. Of these, 98 (73.13%) showed varying degrees of segregation distortion on all the 12 chromosomes suggesting that the distortion was random and not confined to any specific part of the rice genome (SI Table 1.1). These results are in agreement with earlier studies carried out by Amarawathi *et al* (2008) wherein segregation distorted loci are reported to have distributed over eight chromosomes viz., 2, 3, 4, 6, 7, 8, 9 and 10. Majority of the markers represented heterozygotes while very few (~9%) show Basmati370 alleles (SI Table 1.1). The highest number of markers showing distorted segregation were mapped to chromosome 8 (12), whereas the lowest number (1) to chromosome 12.

Marker interval and map length

Total length of the Basmati genetic map was estimated to be 2443.6cM (Figure 1.11). The average genetic distance between adjacent markers for 133 intervals on 12 chromosomes was 18.37 cM. However, there were five large genetic gaps of 55-72 cM on chromosomes 1, 2, 8, 9 and 12 where map distances could not be determined accurately. Excluding these gaps, the average distance of remaining intervals was 16.41 cM.

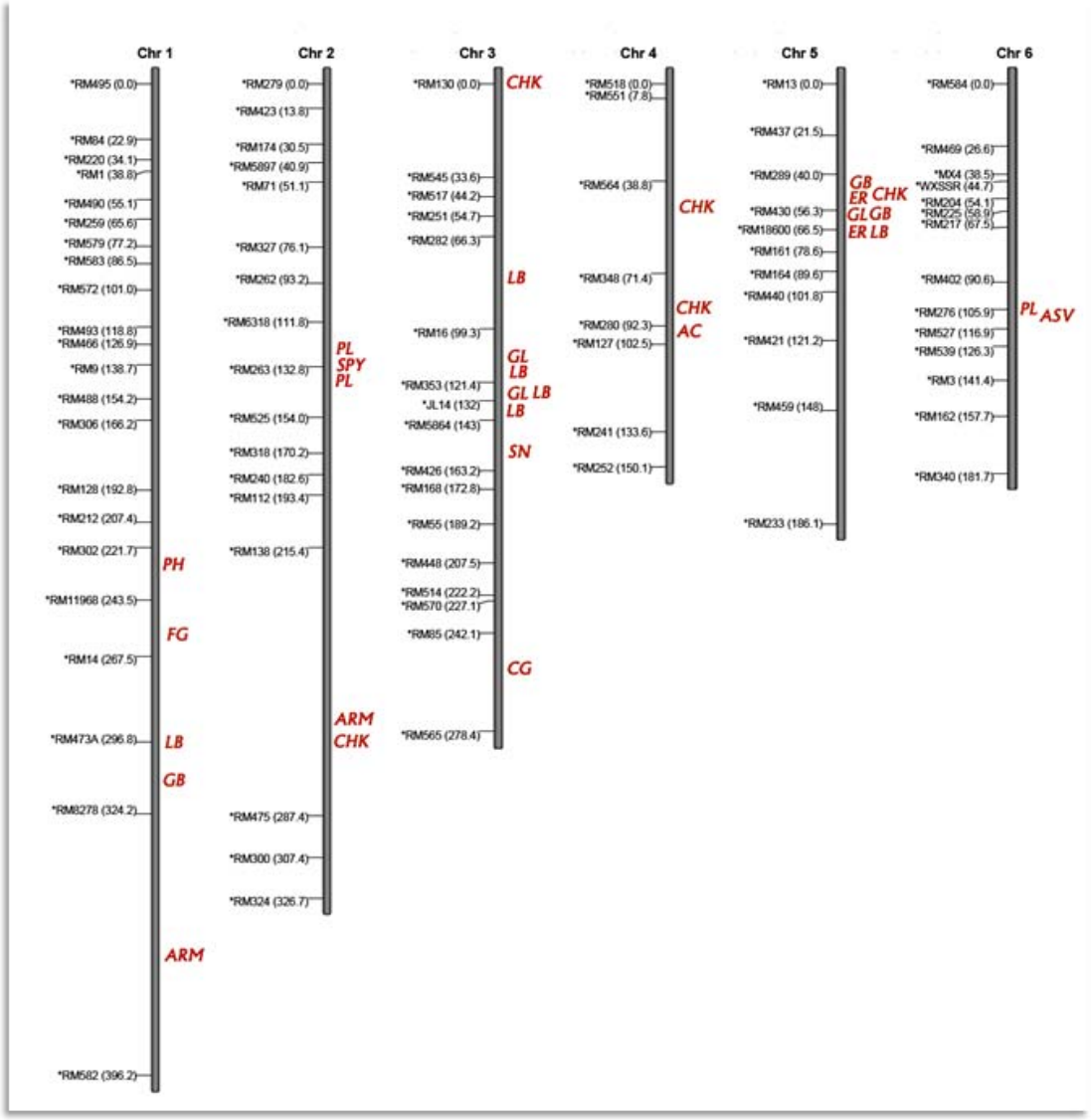
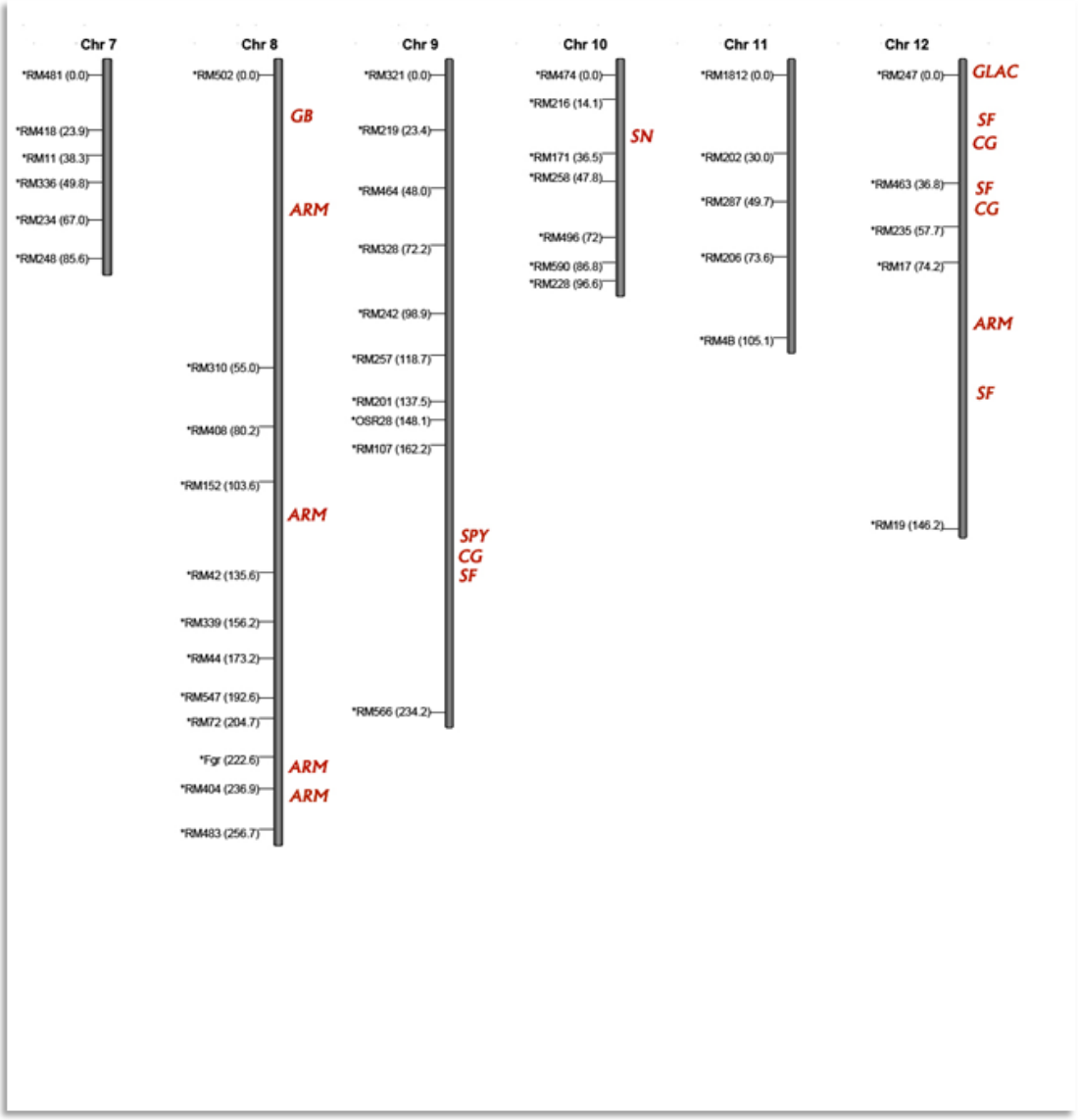


Figure 1.11 Distribution of QTLs for 16 traits in the molecular linkage map of Basmati. The QTLs are indicated in bold (red colour) and right side of the linkage group. For codes of the traits refer Table 1.1. Names of the markers are represented left side of the linkage group. Numbers in parenthesis are genetic distance between markers in centi morgans (cM).

Contd..



For details refer Figure 1.11

Genome composition

Average percent heterozygosity per plant in the F_2 population was estimated to be 59.7% the range being from 33 to 90% indicating that the heterozygosity was close to the expected (50%) and exceeded the range 3 SD (standard deviation) on each side. The average percent homozygosity was 19.1% (range: 3.7 - 49.6%) for B/B genotypes (Basmati370) and 21.2% (range: 4.7-43.6%) for J/J genotypes (Jaya). Deviations from expected homozygotes ratios were observed with over abundance of heterozygotes in most of the cases. Of the 134 markers, on an average 131 markers per F_2 individual were scored with a range of 113 to 134.

Mapping of QTLs for agronomic and quality traits

Each of the scored traits was subjected to QTL mapping using rice microsatellite markers. Many QTLs were identified using interval mapping (IM) and composite interval mapping (CIM) methods of QTL Cartographer v 2.5 software at 2.5 as LOD threshold for detecting a QTL. By IM, 77 QTLs relating to 17 traits were identified except 1000 seed weight (SI Table 1.2) whereas by CIM 82 QTLs were identified for 17 traits except number of panicles (SI Table 1.3). In all, 47 QTLs for all the agronomic and quality traits of Basmati were identified using both the methods (IM and CIM) except for number of panicles and 1000 seed weight (Table 1.5). These QTLs were distributed on all the rice chromosomes except 7 and 11 (Figure 1.11). On an average 2.9 QTLs were detected per character and 3.9 QTLs per chromosome (Figure 1.11). Percent phenotypic variation contributed by each QTL for each of the traits ranged from zero as in the case of spikelet number and grain breadth to 82.16% as in single plant yield. Of 47 QTLs, 17 explained more than 15% phenotypic variance.

Plant height

Only one QTL, designated as *qPH1.1,* was identified for plant height trait on chromosome 1 at an interval of RM302-RM11968 and it accounted for 15.42% phenotypic variance. Allele from Basmati370 was associated with increased plant height.

Panicle length

Of the three QTLs for panicle length, two were on chromosome 2 (*qPL2.1, qPL2.2*) and one was on chromosome 6 (*qPL6.1*). These three QTLs together contributed to 2.04% total phenotypic variance. Increased quantitative effect came from Basmati370 suggesting a major part of the variation in panicle length still remained unexplained due to environmental influence on this trait.

Filled grains

A single QTL designated as *qFG1.1* was identified on chromosome 1 in the marker interval of RM11968-RM14. It explained 22.68% of the phenotypic variance between the parents indicating the possible involvement of a major gene governing the trait. Its increasing effect resulted from the Basmati parent.

Table 1.5 Quantitative trait loci (QTLs) detected in Basmati370/Jaya F_2 population

S.No.	Trait	QTL	Chr.	Marker interval	LFM	RFM	LOD	Additive	Dominance	PVE
1	Plant height (cm)	*qPH1.1**	1	RM302-RM11968	16	10.4	5.138	7.908	-0.858	15.418
2	Panicle length (cm)	*qPL2.1*	2	RM6318-RM263	16	9.28	3.039	0.456	1.636	0.925
3		*qPL2.2*	2	RM263-RM525	10	15.55	2.818	0.267	1.920	0.296
4		*qPL6.1*	6	RM276-RM527	2	10.22	3.413	0.408	-1.773	0.819
5	Filled grains (no.)	*qFG1.1*	1	RM11968-RM14	10	19.55	3.244	31.165	-28.073	22.677
6	Chaffy grains (no.)	*qCG3.1*	3	RM85-RM565	20	30.2	4.284	-3.532	-13.439	0.460
7		*qCG9.1*	9	RM107-RM566	34	80	3.021	-2.710	-10.480	0.328
8		*qCG12.1*	12	RM247-RM463	34	15.23	5.211	-7.804	-15.738	2.458
9		*qCG12.2*	12	RM463-RM235	18	7.15	10.403	-9.381	-8.383	2.335
10	Spikelet number (no.)	*qSN3.1*	3	RM5864-RM426	14	10.16	2.788	666.000	-1.593	0.000
11		*qSN10.1*	10	RM216-RM171	26	1.29	2.885	-19.354	-6.115	6.661
12	Spikelet fertility (%)	*qSF9.1*	9	RM107-RM566	56	58	2.562	4.208	5.350	2.202
13		*qSF12.1*	12	RM247-RM463	34	15.23	4.062	4.390	9.115	2.150
14		*qSF12.2*	12	RM463-RM235	14	11.15	7.255	7.155	1.973	4.249
15		*qSF12.3*	12	RM17-RM19	48	66	3.441	5.987	-4.491	4.472
16	Single plant yield (g)	*qSPY2.1*	2	RM263-RM525	0	25.55	3.720	-2.258	3.979	4.060
17		*qSPY9.1*	9	RM107-RM566	48	66	3.154	8.397	-4.769	82.155
18	Grain length (mm)	*qGL3.1*	3	RM16-RM353	6	20.91	3.663	0.179	-0.215	10.901
19		*qGL3.2**	3	RM353-JL14	10	1.7	9.217	0.362	-0.125	46.065
20		*qGL5.1*	5	RM430-RM18600	6	5.2	6.603	0.217	0.031	17.468
21	Grain breadth (mm)	*qGB1.1*	1	RM473A-RM8278	0	34.52	6.714	-0.038	0.119	1.649
22		*qGB5.1*	5	RM289-RM430	2	16.88	3.345	-0.096	0.061	13.078
23		*qGB5.2**	5	RM430-RM18600	4	7.2	3.333	-0.106	0.052	17.149
24		*qGB8.1*	8	RM502-RM310	16	64.66	3.454	666.000	0.015	0.000
25	Length-Breadth ratio	*qLB1.1*	1	RM473A-RM8278	0	34.52	5.063	0.116	-0.208	3.928
26		*qLB3.1*	3	RM282-RM16	20	23.13	3.323	0.433	-0.194	74.353
27		*qLB3.2*	3	RM16-RM353	8	18.91	3.491	0.169	-0.211	12.326
28		*qLB3.3**	3	RM353-JL14	8	3.7	4.358	0.220	-0.129	22.342
29		*qLB3.4*	3	JL14-RM5864	2	10.21	4.770	0.229	-0.111	24.958
30		*qLB5.1**	5	RM430-RM18600	8	3.2	4.650	0.405	-0.070	46.531
31	Grain length after cooking (mm)	*qGLAC12.1*	12	RM247-RM463	0	49.23	3.512	0.312	0.396	2.680
32	Elongation ratio	*qER5.1*	5	RM289-RM430	14	4.88	3.706	-0.143	0.075	19.931
33		*qER5.2*	5	RM430-RM18600	4	7.2	3.711	-0.136	0.067	18.931
34	Alkali spreading value	*qASV6.1**	6	RM276-RM527	4	8.22	26.746	-1.257	0.264	71.735
35	Amylose content (%)	*qAC4.1*	4	RM280-RM127	0	11.15	4.077	-0.970	0.315	14.249
36	Aroma	*qARM1.1*	1	RM8278-RM582	74	40	6.735	0.654	5.284	1.859
37		*qARM2.1*	2	RM138-RM475	80	32.06	7.590	-0.178	-5.332	0.133
38		*qARM8.1*	8	RM502-RM310	36	44.66	6.976	-0.230	-5.309	0.218
39		*qARM8.2*	8	RM152-RM42	18	23.56	6.132	0.968	-5.312	3.116
40		*qARM8.3**	8	Fgr-RM404	10	6.36	4.910	1.946	1.161	10.990
41		*qARM8.4*	8	RM404-RM483	8	16	4.998	2.476	0.511	20.226
42		*qARM12.1*	12	RM17-RM19	30	84	7.556	-0.589	-5.334	1.512
43	Chalkiness	*qCHK2.1*	2	RM138-RM475	56	56.06	2.696	1.587	-0.468	69.576
44		*qCHK3.1*	3	RM130-RM545	0	44.07	3.234	0.915	-0.575	15.083
45		*qCHK4.1*	4	RM564-RM348	14	28.62	3.138	2.107	-0.142	63.795
46		*qCHK4.2*	4	RM348-RM280	22	3.15	5.896	1.087	-0.754	25.092
47		*qCHK5.1*	5	RM289-RM430	6	12.88	3.835	-0.809	0.359	14.533

* Stable QTLs; PVE- Phenotypic variance explained by each QTL; Left (LFM) and right (RFM) flanking marker distance from the QTL (cM);Positive and negative values of additive effect indicates the increasing effect coming from the alleles of Basmati370 and Jaya respectively.

Chaffy grains

A total of four QTLs influencing chaffy grains designated as *qCG3.1 qCG9.1 qCG12.1 qCG12.2* was identified one each on chromosomes 1 and 9 and two on chromosome 12. Together they explained 5.6% phenotypic variation. It was Jaya parent that increased chaffy grains at all the loci.

Spikelet number

Two regions were found to be associated with QTLs for spikelet number viz., *qSN3.1* and *qSN10.1* on chromosome 3 and 10 respectively. Of the two QTLs, the QTL *qSN3.1* explained zero percent phenotypic variation of the trait suggesting genes within this QTL region might be having opposite effects whereas *qSN10.1* accounted for 6.7% of the phenotypic variation with the allele coming from the parent Jaya.

Spikelet fertility

Four QTLs, one on chromosomes 9 (*qSF9.1*) and remaining three on chromosome 12 (*qSF12.1 qSF12.2* and *qSF12.3*) affecting spikelet fertility were identified. Together they accounted for 12.89% of the phenotypic variance. In all the loci Basmati parent contributed to spikelet fertility.

Single plant yield

Two QTLs, *qSPY2.1* and *qSPY9.1* were identified for single plant yield on chromosomes 2 and 9 respectively. The QTL *qSPY9.1* on chromosome 9 explained 82.1% phenotypic variance. The other QTL, *qSPY2.1* accounted for only 4.06% of the phenotypic variance. The allele for increased grain yield came from Basmati370 for *qSPY9.1* while from Jaya for *qSPY2.1.*

Grain length

A total of three QTLs viz., *qGL3.1*, *qGL3.2* and *qGL5.1* two on chromosome 3 and one on chromosome 5 were detected. For all the three QTLs, the increasing effect was associated with Basmati370 allele. These three QTLs explained phenotypic variance of 10.9, 46.01 and 17.47% respectively.

Grain breadth

Four QTLs, *qGB1.1*, *qGB5.1*, *qGB5.2* and *qGB8.1* were found to be responsible for grain breadth. Of the four QTLs, two QTLs, *qGB5.1*, *qGB5.2* on chromosome 5 had major effect explaining 13.08 and 17.15% phenotypic variance respectively and one QTL *qGB1.1* on chromosome 1 had relatively minor effect explaining 3.9% phenotypic variance. In all these cases, increased QTL effect came from the parent Jaya. For the QTL *qGB8.1*, Basmati370 and Jaya alleles have opposite effects resulting in zero percent variance in phenotype.

Length-Breadth ratio (LB) / Grain size

A total of six QTLs influencing this trait was identified. In all the QTLs, alleles from Basmati370 contributed to increase in LB ratio. The QTLs, *qLB3.2*, *qLB3.3* on chromosome 3 and *qLB5.1* on chromosome 5 explained 12.33, 22.34 and 46.53% phenotypic variation respectively. Further, these loci were located in the vicinities of *qGL3.1* and *qGL3.2* controlling grain length and *qGB5.2* controlling grain breadth traits. Such association is not surprising because LB ratio is a derived trait obtained by dividing grain length by grain breadth. The other three QTLs viz., *qLB1.1*, *qLB3.1* and *qLB3.4* explained 3.93, 74.35, 24.96% phenotypic variance respectively.

Grain length after cooking (GLAC)

A QTL associated with GLAC, *qGLAC12.1* contributing 2.68% phenotypic variance was located on chromosome 12. Basmati allele was associated with an increase of GLAC as was the case in grain length.

Elongation ratio (ER)

Two QTLs, *qER5.1* and *qER5.2* were identified for this trait on chromosome 5 explaining 19.9 and 18.9% phenotypic variance respectively. The allele from Jaya variety at these regions increased the elongation ratio.

Alkali spreading value (ASV)/ Gelatinization temperature (GT)

One major QTL on chromosome 6, *qASV6.1* with the highest LOD value of 26.75, was identified. It explained a maximum of 71.74% phenotypic variance (Figure 1.12). The allele from Jaya had a strong positive effect on ASV.

Amylose content (AC)

In the present study, one QTL *qAC4.1* explaining 14.25% phenotypic variance on chromosome 4 was detected. The Jaya allele had increasing effect on this trait. A QTL for amylose content viz., *qAC6.1* was also identified on chromosome 6 at an interval of RM469- MX4 employing CIM method of QTL Cartographer (SI Table 1.3). It explained 15% phenotypic variation. As expected the increased effect of this QTL was associated with high amylose Jaya parent.

Aroma (ARM)

Seven QTLs designated as *qARM1.1, qARM2.1, qARM8.1, qARM8.2, qARM8.3, qARM8.4* and *qARM12.1* influencing aroma were identified. Of these, four QTLs *qARM8.1, qARM8.2, qARM8.3* and *qARM8.4* were located on chromosome 8 explaining 0.22, 3.12, 10.9 and 20.23% phenotypic variance, respectively. The other three QTLs *qARM1.1, qARM2.1* and *qARM12.1* located on chromosomes 1, 2 and 12 respectively together contributed 3.51% phenotypic variance. These QTLs are novel ones and are specific to Basmati varieties.

Chalkiness (CHK)

A total of five QTLs, *qCHK2.1*, *qCHK3.1, qCHK4.1 qCHK4.2* and *qCHK5.1* were identified on chromosomes 2, 3, 4 and 5 respectively. Most of the QTLs explained >14% phenotypic variation. In all, except QTL *qCHK5.1* on chromosome 5, the increased effect was from the Basmati parent.

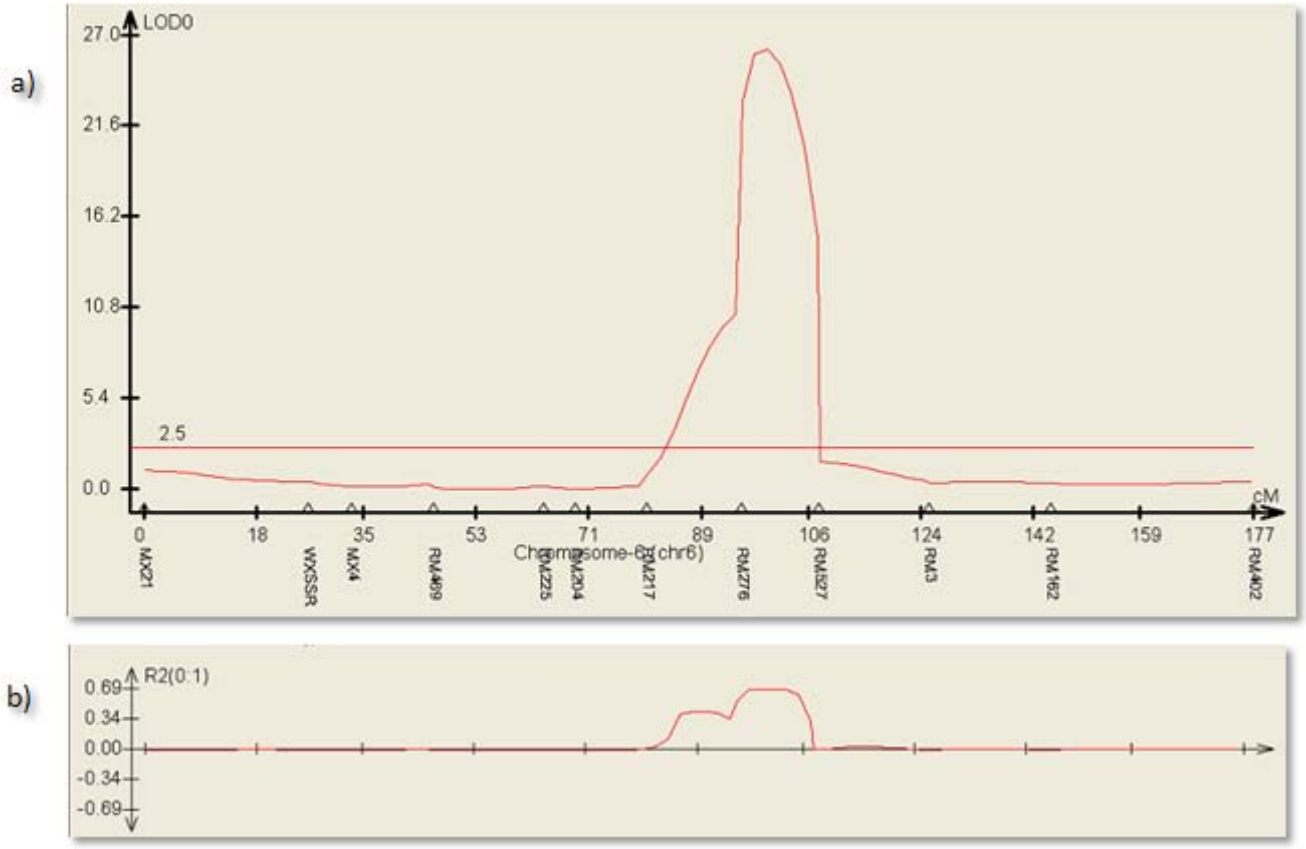


Figure 1.12 QTL cartographer LOD peak for alkali spreading value

a) Markers and their genetic distances are given in X-axis and LOD values in Y-axis
b) Phenotypic variance explained by the alkali spreading value QTL

Discussion

1.4 Discussion

Basmati rices of the Indian sub continent: genetics of quality traits

Basmati rices of the Indian subcontinent are highly distinctive rices because of their unique traits. With the advent of high yielding varieties ensuring higher farm returns serious threat to Basmati rices was perceived by the breeders pushing them to resort to breeding for varieties of Basmati quality in high yielding background. But for reasons that are beginning to be understood, no variety ideally matching the traditional Basmati could be evolved. Genetic investigations have revealed that all except one or two are controlled quantitatively and selections based on phenotype are not reliable enough (Puri and Siddiq 1980; Amarawati *et al* 2008). The present study undertaken with the objective of identifying quantitative trait loci (QTLs) governing the key characters of Basmati rice using the mapping population (F_2) derived from a cross of traditional Basmati variety, Basmati370 and high yielding non-Basmati variety Jaya. To our knowledge, the current study is the first one to carry out genome-wide mapping to dissect the genetic basis of important agronomic and quality traits of Basmati rice.

Phenotypic variation in the mapping population

All the Basmati traits, except chaffy grains, spikelet fertility, alkali spreading value, aroma, chalkiness and amylose content show continuous variation suggesting their polygenic nature (Figures 1.9 and 1.10). Transgressive segregants found in both the directions that means beyond both the parents for all the agronomic traits and a few quality traits viz., grain length, grain breadth and amylose content suggest that neither of the parents carry all the positive or the negative alleles. Occurrence of such transgressions is possibly due to accumulation of complementary alleles from both the parents at multiple loci in certain F_2 population (Tanksley, 1993) and unmasking of recessive deleterious alleles due to inbreeding (Rick and Smith 1953).

According to Singh and Narayan (1997), high heterosis followed by inbreeding depression in next generation for a trait is an indication of non-additive gene action (dominance and epistasis) while similar performance of a trait in both F_1 and F_2 or negative performance in F_1 and enhanced performance in F_2 is suggestive of additive gene action. In the present study, all

the traits except number of panicles, panicle length and chalkiness seemed to show non-additive gene action (Table 1.2).

In general, in the present study, correlations were not strong among quality traits suggesting that very few genes are controlling these traits. Interestingly, 1000 seed weight, aroma, alkali spreading value and amylose content were found to be correlated with none of the traits studied implying that these traits are controlled by few genes. In contrast, earlier reports have suggested positive and significant correlation of amylose content with alkali spreading value (Bao *et al* 2002) and cooked rice elongation (Tian *et al* 2005).

Construction of genetic map

For mapping QTLs, a genetic map was constructed employing 181 F_2 offspring and 134 markers. High parental polymorphism (42.12%) observed in the present study gives further evidence to the divergence and distinctness of Basmati rices from the other *indica* and *japonica* rice groups as revealed by an earlier study (Nagaraju *et al* 2002). The percent polymorphism detected in the present study is higher than the previously reported 28.9% (Amarawati *et al* 2008) and lower than 63.95% reported by Govindaraj *et al* (2005). The former had used evolved Basmati i.e, Pusa1121 as one of the parents and in the latter study a traditional Basmati i.e. Basmati370 was used as parent. The linkage map (LOD-score $\geq$3.0) places 134 markers on 12 linkage groups spanning a total map length of 2443.6 cM with an average distance of 18.37 cM between adjacent marker loci. But there are five large genetic gaps of 55-72 cM on chromosomes 1,2,8,9 and 12. Excluding these genetic gaps the average interval of remaining markers is 16.41 cM. According to Lander and Botstein (1989), the average interval being less than 20 cM the linkage map is suitable for QTL mapping. The map reported in the present study is relatively denser with higher percentage of markers on chromosome 1 as compared to the earlier reported molecular maps (Redona and MacKill 1998; Zhuang *et al* 1997). The Basmati genetic map (2443.6 cM) is much longer than some of the notable maps constructed using inter-sub specific populations, (1794.7 cM, Temnykh *et al* (2001); 1521.6 cM, Harushima *et al* (1998) and shorter (2455.7 cM) than the map reported by Qi-Jun *et al* (2006) (Table 1.6). Higher genetic distance observed between some markers in the present study could be attributed to: (a) employment of more number of segregation distorted markers (98), (b) stretching effect of markers on chromosomes caused by small population size contribute to increased map length (Subudhi and Huang 1999),

and (c) map expansion due to excess heterozygosity in segregating markers. This is in agreement with the earlier study of Knox and Ellis (2002). Total map length increase due to stretching effect is known in many crops including rice (Subudhi and Huang 1999), sorghum (Biovin *et al* 1998) and barley (Becker *et al* 1995). Lorieux *et al* (1996) showed that linkage map length could be reduced by 27% if the frequency of distorted alleles is corrected.

The linear order of microsatellite markers reported in the present genetic map is not in total agreement with that of the high-density linkage map of Temnykh *et al* (2001). The two maps differ in genetic distances between microsatellite markers (Table 1.6). Such observed discrepancies in the map distances is most likely to occur on account of different parental strains, number of markers, differences in the size and type of mapping populations and levels of polymorphism. Nevertheless, for exploratory mapping, resolution and genome coverage of the present linkage map may be adequate at least for some of the chromosomes to detect QTLs relating to economically important traits.

Out of 134 markers screened, 98 (73.13%) have shown distorted segregation and this phenomenon has been reported earlier in many mapping populations (Causse *et al* 1994; Xiao *et al* 1996b; Xu *et al* 1995; Redona and Mackill, 1998). The gametophytic gene *ga-1* on chromosome 6 is reported to be associated with segregation distortion (Iwata *et al* 1964). Nakagahra (1972) further reported the presence of *ga-2* and *ga-3* on chromosome 3. Thus distortion in segregation does not seem to be confined to any specific chromosome as opined by Xu *et al* (1995). The genetic basis of such skewed segregation may be due to either selective abortion of male or female gametes, or selective gametic mating. The few markers that showed segregation distortion in the present study viz., RM71, RM85, RM241, RM348, RM280 and RM42 were also reported to exhibit segregation distortion by Aluko *et al* (2004). Likewise, the distorted markers in the current study i.e., RM279 and RM248 were also found to show segregation distortion in an earlier study (Amarawati *et al* 2008). The order of distorted markers on our linkage map was altered in position when compared to the previously published maps. Generally, segregation distortion is associated with the crosses of rice sub-species and species (Aluko *et al* 2002 and Pradeep *et al* 2005). The high percent of distortion observed in the present study gives further support to the contention that of Basmati rice is genetically quite distinct from *indica* and *japonica* varieties.

Table 1.6 Comparison of Basmati genetic map with previously published rice genetic maps (Based on the whole genome length of the haploid rice i.e., 4.3 x 10^8, Arumuganathan and Earle, 1991)

	Current study	Qi-Jun *et al* (2006)	Temnykh *et al* (2001)	Harushima *et al* (1998)
Parents	Basmati370/Jaya	Nipponbare/93-11	IR64/Azucena	Nipponbare/Kasalath
Type of the population	F_2	F_2	DH	F_2
Size of the population	181	90	96	186
Type of the markers	SSR	SSR	SSR & RFLP	RFLP
Number of the markers	134	152	>500 SSRs & 145 RFLPs	2275
Map length (cM)	2443.6	2455.7	1794.7	1521.6
Genetic distance between markers (cM)	18.23	16.16	2.78	<2
Physical distance between markers (kb)	3208.9	2828.9	666.7	189.01

Mapping of QTLs of agronomic and quality traits

In all, 47 QTLs were identified for 16 agronomic and grain quality traits. Of these, majority of the alleles of increasing effect was found to be contributed from Basmati parent. Of 47 QTLs, 17 QTLs explained more than 15% phenotypic variation between parents. The large amount of variation explained by these QTLs seems to suggest that inheritance of these traits is controlled by gene(s) with large effect and possible involvement of a few major genes. Though agronomic traits especially yield and yield components are known to be controlled by polygenes, contrary to the expected higher QTLs, very few QTLs (1 to 2) were identified in the present study for plant height, number of filled grains, spikelet number and single plant yield. Such an unexpected result may be due to genetically distant population used in the study, minor QTLs not detected and environmental effects.

Plant height

Interestingly, near the plant height QTL *qPH1.1* identified in the current study, semi dwarf gene, *sd1* which encodes a gibberellic 20-oxidase (*OsGA20ox-1*), an enzyme that catalyzes the conversion of GA_{53} to GA_{20} in gibberellic acid biosynthesis in rice (Monna *et al* 2002; Sasaki *et al* 2002; Speilmeyer *et al* 2002) was also located on chromosome 1. In another study, Ishimaru *et al* (2004) identified sucrose phosphate synthase gene controlling plant height on different region of the same chromosome.

Panicle length

Recently, *aberrant panicle organization-1 (APO-1)* gene encoding an F-box protein on chromosome 6 has been reported. The mutant strain of this gene results in reduced panicle length and thereby accommodates less number of spikelets per panicles (Ikeda *et al* 2007). Further experiments are required to confirm the gene underlying the QTL for panicle length on chromosome 6 (*qPL6.1)* also encodes F-box protein.

Filled grains

Like plant height QTL, the major QTL *qFG1.1* for filled grains was also close to *sd1* gene (~80kb). Recently, Ashikari *et al* (2005) have identified a gene underlying the major QTL controlling grain number on chromosome 1 i.e., *Gn1a* encoding for cytokinin oxidase/dehydrogenase (*OsCKX2*),

an enzyme that degrades phytohormone cytokinin. And further concluded that the reduced expression of *OsCKX2* causes cytokinin accumulation in inflorescence meristems and increases the number of reproductive organs resulting in enhanced grain yield. Pyramiding of the genes for grain number and plant height into the same rice background resulted in 26% more grain number per panicle and 18% short in plant height (Ashikari *et al* 2005). The QTL detected in the present study and *Gna1* are not the same suggesting that *qFG1.1* seemed to harbour other candidate genes that control the grain number through mechanism(s) that remain to be elucidated.

Single plant yield

The QTL *qSPY9.1* explained highest (82.1%) phenotypic variance suggesting involvement of a major gene controlling this trait and this QTL could be a potential candidate for fine mapping and map based cloning projects. Recently, He *et al* (2006) fine mapped an yield improving QTL *GY2-1* on chromosome 2 which was different from the *qSPY2.1* identified in the current study to a 102.9 kb region and concluded that it has a haplotype of leucine rich repeat receptor kinase gene cluster which showed an extensive allelic variation between parents i.e., Dongxiong (a wild rice, *Oryza rufupogan* Griff.) and Guichao2 (*Oryza sativa ssp indica*).

Grain length

As the QTL on chromosome 3 i.e., *qGL3.2* was tightly linked to the right flanking marker i.e., JL14 with a genetic distance of 1.7cM, it can be right away used in marker-assisted selection for improvement of grain length. Moreover, this region coinciding with a major effect QTL for grain size was consistently detected around the centromeric region of chromosome 3 in numerous studies across different environments and genetic backgrounds (Table 1.7) (Fan *et al* 2006; Wan *et al* 2006; Wang *et al* 2006; Zhou *et al* 2006). Therefore, the present study tends to support the general conclusion made by Tanksley (1993) that a substantial proportion of QTL affecting a trait particularly those having major effects can be identified under different environments. Recently this QTL *GS3* was dissected further into a gene, encoding a putative PEPB (phosphatidyl ethanolamine-binding protein) -like domain, a transmembrane region, a putative TNFR (tumor necrosis factor receptor) /NGFR (nerve growth factor receptor) family cystein rich domain and a VWFC (von willebrand factor type C) module. Comparative sequence analysis identified a non-sense mutation in the second exon of the putative *GS3* gene in all long-grain varieties when

compared to small grain varieties. This mutation causes a 178 amino acid truncation in the C-terminal region of the predicted protein suggesting that *GS3* may function as a negative regulator for grain size (Fan *et al* 2006). The other two QTLs viz., *qGL3.1* and *qGL5.1* also explained >10% phenotypic variance suggesting that these QTLs are good candidates for fine mapping and cloning projects. Taken together with the results of previous studies it can be inferred that major QTL for grain length detected in the study on chromosome 3 is likely to be the same locus as the one reported by earlier studies (Tan *et al* 2000; Wan *et al* 2005). In contrast, Redona and Mackill (1998) detected a QTL for grain length around RG476 on chromosome 4 which was not detected in the present study. It is also interesting to note that the chromososme region of maize flanked by umc164c and umc157 on chromosome 1 harbouring a QTL for kernel length is homeologous to the short arm of the rice chromosome 3 suggesting the possibility of orthology between rice and maize genes governing kernel length in this region (Li *et al* 2004).

Grain breadth

The QTLs, *qGB5.1* and *qGB5.2* have been consistently reported in the earlier studies across different environments and genetic backgrounds (Table 1.7). For instance, Wan *et al* (2006) reported that the QTL *qGW-5.2* explains 44% phenotypic variation across the eight environments. The flanking markers of this QTL may be suitable for marker assisted selection after validation (Tan *et al* 2000; Wan *et al* 2005). On the other hand, Song *et al* (2007) identified a major QTL for grain width i.e., *GW2* on chromosome 2 which encodes a previously unknown RING type protein with E3 ubiquitin ligase activity and is known to function in the degradation by the ubiquitin-proteasome pathway. Further, loss of *GW2* function increased cell numbers resulting in a larger or wider spikelet hull and accelerated the grain milk filling rate which consequently enhances grain width, weight and yield. Surprisingly, similar kind of enzyme exists in the *qGW-5.2* region i.e., interval of RM430 and RM18600 on chromosome 5 with a physical distance of 9.2kb from the RM18600 marker (Figure 1.13) (Table 1.8). This suggests that RING E3 ligase encoding gene may be the major gene underlying the QTL *qGB-5* on chromosome 5. However, fine mapping of this QTL region and map-based cloning of this gene is required to initiate functional analysis of the candidate gene.

Table 1.7 Previously reported QTL information on grain appearance traits of rice on chromosomes 3 and 5

Trait	QTL*	Chr.	Marker interval	LOD	PVE	Population	Parents	Cross	Ref.
Grain length (L)	qGL3	3	R19-C1677	6.46	32.77	71 RILs	Asominori/IR24	J/I	Wan *et al* (2005)
	gl3	3	RM251-RM554	6.32	19.3	192 F_2	Domsephid/Gerdeh	I/J	Rabiei *et al* (2004)
	gl5	5	RM437-RM289	3.61	13.5	192 F_2	Domsephid/Gerdeh	I/J	Rabiei *et al* (2004)
	Kl3.1	3	RZ251	22.2	25	308 BC_3F_1	*O.sativa/O. glaberrima*	interspecific	Li *et al* (2004)
	gl3	3	RM251-RM338	3.8	12.5	312 DH	*O.sativa/O. glaberrima*	interspecific	Aluko *et al* (2004)
	GL3	3	RG393-C1087	32.7	59	241 F_2	Zhenshan 97/Minghui 63	I/J	Tan *et al* (2000)
	QGl3	3	RD301-RD3.2	15.31	17.3	292 RILs	Lemont/Teqing	J/I	Xu *et al* (2004)
	QGl5	5	RM163-RM161	8.79	10	292 RILs	Lemont/Teqing	J/I	Xu *et al* (2004)
Grain breath (B)	qGB5	5	R3166-R569	6.3	26.99	71 RILs	Asominori/IR24	J/I	Wan *et al* (2005)
	gb3	3	RM7-RM251	15.02	34.1	192 F_2	Domsephid/Gerdeh	I/J	Rabiei *et al* (2004)
	gb5	5	RM437-RM289	6.2	14.9	192 F_2	Domsephid/Gerdeh	I/J	Rabiei *et al* (2004)
	GW	5	MRG0002-RM218	7.8	18.5	190 DH	Zhenshan 97/WYJ-2	I/J	Jiang *et al* (2005)
	GB	5	RG360-C37349	19.9	52.3	241 F_2	Zhenshan 97/Minghui 63	I/J	Tan *et al* (2000)
	QGw5	5	R569a-RM249	12.87	17.6	292 RILs	Lemont/Teqing	J/I	Xu *et al* (2004)
	Qwij-5	5	RM13-RM289	3.43	26.1	143 BC_1F_1	IR36/NK2//IR36	I/J//I	Abdelkhalik *et al* (2005)
L-B ratio/Grain size	qLWR-3	3	R19-C1677	3.93	20.6	71 RILs	Asominori/IR24	J/I	Wan *et al* (2005)
	qLWR-5	5	R3166-R569	6.6	26.4	71 RILs	Asominori/IR24	J/I	Wan *et al* (2005)
	gs3	3	RM251-RM554	11.29	27.1	192 F_2	Domsephid/Gerdeh	I/J	Rabiei *et al* (2004)
	gs5	5	RM437-RM289	6.03	14.5	192 F_2	Domsephid/Gerdeh	I/J	Rabiei *et al* (2004)
	lwr3.1	3	RZ251	7	8	308 BC_3F_1	*O.sativa/O. glaberrima*	interspecific	Li *et al* (2004)
	GS	3	RM36-MRG0002	8.9	17.1	190 DH	Zhenshan 97/WYJ-2	I/J	Jiang *et al* (2005)
	LWR	3	C1087-R2403	12.8	29.4	241 F_2	Zhenshan 97/Minghui 63	I/J	Tan *et al* (2000)
	LWR	5	RG360-C7349	9.9	31.3	241 F_2	Zhenshan 97/Minghui 63	I/J	Tan *et al* (2000)
	QGs5a	5	R569a-RM249	5.67	7.8	292 RILs	Lemont/Teqing	J/I	Xu *et al* (2004)
	QGs5b	5	RM163-RM161	3.10	3.8	292 RILs	Lemont/Teqing	J/I	Xu *et al* (2004)

I-indica; J- japonica; *Note: Codes used in previous studies for grain length – GL or KL; grain breadth – GB or GW; grain size – LWR or GS

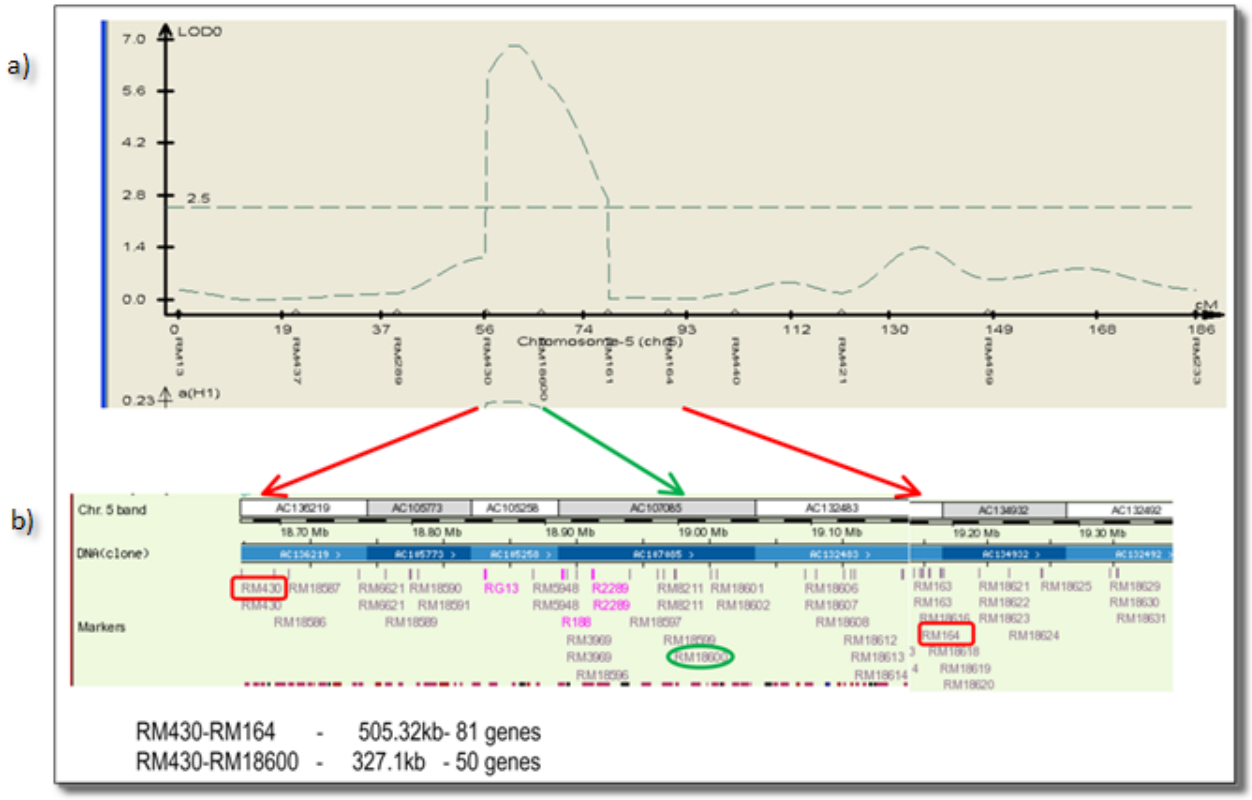


Figure 1.13 a) QTL cartographer LOD peak for grain breadth and its flanking microsatellite markers. LOD 2.5 taken as threshold for detecting a QTL. X-axis shows the markers and their genetic distances in cM. Y-axis shows the LOD values.

b) Corresponding markers for grain breadth QTL on chromosome 5 of *Japonica* genome in gramene web site (www.gramene.org). Markers highlighted with red boxes are the ones flanking grain breadth QTL. After fine mapping, this QTL has been narrow down to the interval of RM430 and RM18600.

Table 1.8 List of genes present in the marker interval of RM430-RM18600 on chromosome 5*

	RM430			
S.No.	Gene Name	Start	End	Function
1	Os05g32170.1	18674175	18677127	crinkly4-like protein, putative, expressed
2	Os05g32180.1	18680509	18684040	N-acetyltransferase, putative, expressed
3	Os05g32190.1	18685899	18688970	tungus CG8253-PA, putative, expressed
4	Os05g32200.1	18690921	18691766	hypothetical protein
5	Os05g32210.1	18696550	18703677	NAD kinase 3, putative, expressed
6	Os05g32220.1	18706496	18710630	50S ribosomal protein L1, putative, expressed
7	Os05g32230.1	18713284	18718949	expressed protein
8	Os05g32240.1	18719305	18719631	hypothetical protein
9	Os05g32250.1	18721123	18724096	hypothetical protein
10	Os05g32260.1	18725711	18725992	hypothetical protein
11	Os05g32270.2	18731187	18734545	AP2 domain transcription factor, putative, expressed
12	Os05g32270.1	18729277	18734545	AP2 domain transcription factor, putative, expressed
13	Os05g32280.1	18736757	18737946	conserved hypothetical protein
14	Os05g32290.1	18739991	18745654	retrotransposon protein, putative, Ty1-copia subclass
15	Os05g32300.1	18747126	18747926	hypothetical protein
16	Os05g32310.1	18755190	18757678	SAD1, putative, expressed
17	Os05g32320.1	18759608	18762691	transporter, folate-biopterin transporter, putative, expressed
18	Os05g32330.1	18762799	18766799	hydrolase, putative, expressed
19	Os05g32330.2	18762799	18766799	hydrolase, putative, expressed
20	Os05g32340.1	18771955	18776279	retrotransposon, putative, centromere-specific, expressed
21	Os05g32350.1	18779494	18784141	protein binding protein, putative, expressed
22	Os05g32360.2	18786041	18789028	serine/threonine-protein kinase bur-1, putative, expressed
23	Os05g32360.1	18785338	18789091	serine/threonine-protein kinase bur-1, putative, expressed
24	Os05g32370.1	18789174	18792756	pre-mRNA-splicing factor ATP-dependent RNA helicase DHX16, putative, expressed
25	Os05g32380.1	18803258	18804753	triacylglycerol lipase, putative, expressed
26	Os05g32390.1	18808838	18817716	dynamin family protein, expressed
27	Os05g32390.2	18808857	18817716	dynamin family protein, expressed
28	Os05g32400.1	18820276	18826921	retrotransposon protein, putative, unclassified
29	Os05g32410.1	18831511	18831999	conserved hypothetical protein
30	Os05g32420.1	18836176	18837358	expressed protein
31	Os05g32430.1	18845063	18850166	CPL3, putative, expressed
32	Os05g32430.2	18845063	18849231	CPL3, putative, expressed
33	Os05g32440.1	18858318	18859761	expressed protein
34	Os05g32450.1	18860547	18863614	hypothetical protein
35	Os05g32460.1	18875578	18877004	VQ motif family protein, expressed
36	Os05g32470.1	18880590	18884566	retrotransposon protein, putative, unclassified
37	Os05g32474.1	18886077	18887583	expressed protein
38	Os05g32480.1	18887197	18893017	transposon protein, putative, Mutator sub-class
39	Os05g32490.1	18893843	18895043	hypothetical protein
40	Os05g32500.1	18910173	18914617	expressed protein
41	Os05g32500.2	18910449	18914617	expressed protein
42	Os05g32510.1	18916164	18917713	retrotransposon protein, putative, unclassified
43	Os05g32520.1	18918331	18921033	conserved hypothetical protein
44	Os05g32530.1	18927734	18936628	mannosylglycoprotein endo-beta-mannosidase, putative, expressed
45	Os05g32544.2	18939120	18947702	glycosyltransferase, putative, expressed
46	Os05g32544.1	18938548	18947702	glycosyltransferase, putative, expressed
47	Os05g32560.1	18957264	18962941	transposon protein, putative, unclassified
48	Os05g32570.1	18972424	18979195	RING E3 ligase protein, putative, expressed
49	Os05g32580.1	18988309	18995149	expressed protein
50	Os05g32580.2	18988309	18995149	expressed protein
	RM18600			

-Shaded genes are probable candidate genes involved in grain breadth

*Source: www.gramene.org;www.plantgdb.org/OsGDB/

Length-Breadth ratio (LB) / Grain size

Like QTLs for grain length and grain breadth, which we detected on chromosomes 3 and 5 respectively, the QTLs for length-breadth ratio (LB) *qLB3.2*, *qLB3.3* and *qLB5.1* were detected on the same chromosomes (Table 1.7). These results are consistent with previous reports obtained across different environments and genetic backgrounds (Tan *et al* 2000; Wan *et al* 2005) suggesting that these QTLs are either controlled by a single gene with pleiotropic effects or might be tightly linked genes. Hence, these QTLs are potential candidates for future fine mapping and cloning projects.

Grain length after cooking (GLAC)

In contrast to single QTL identified in the present study on chromosome 12, Li *et al* (2004) reported a single QTL on chromosome 3 and Tian *et al* (2005) two QTLs each on chromosomes 2 and 6 for this trait. Ahn *et al* (1993) identified a QTL on chromosome 8 associated with cooked kernel elongation and concluded that this QTL was loosely linked to the fragrance gene.

Elongation ratio (ER)

In a recent study, a QTL for ER, *elr11-1* was identified on chromosome 11 which explained 6.8% phenotypic variance and the increased effect coming from an aromatic variety i.e, Pusa1121 (Basmati like parent) (Amarawati *et al* 2008). On the contrary, in the present study, on chromosome 5 where QTLs for grain appearance traits located, two QTLs for ER were detected.

Alkali spreading value (ASV)/ Gelatinization temperature (GT)

Interestingly, the QTL *qASV6.1* mapped along with *alk* gene which encodes soluble starch synthase IIa (SSIIa) is associated with gelatinization temperature (GT). These results are in agreement with the previous reports which concluded that GT is primarily controlled by *alk* gene (He *et al* 1999; Lanceras *et al* 2000). Contrary to these results, Tan *et al* (1999) and Zheng *et al* (2008) reported that GT was controlled by waxy gene. This result shows that the genetic factors beyond the *alk* gene are probably involved in altering the GT variation leading to the proposal that *alk* is a major but not the sole player in GT variation. Previous reports suggested that the SSIIa was one of the important biosynthetic enzymes determining the starch structure and its properties (Nakamura 2002; Umemoto *et al* 2002). The SSIIa enzyme seems to have a role in

elongation of A and B1 amylopectin chains, and also actively determine the ratio of two chain lengths i.e., L- type (present in *indica* rices) and S-type (present in japonica rices) (Nakamura 2002; Umemoto *et al* 2002). However, in Basmati rice, being a separate group from *indica* and *japonica* rice, it is interesting to understand the role as well as the structure of SSIIa.

Amylose content (AC)

In many previous reports, the major QTL or gene controlling AC was identified as *waxy* (*wx*) gene located on chromosome 6. Interestingly, the *qAC6.1* QTL identified in the present study is coinciding with the *waxy* gene. This gene reported to be encodes granule bound starch synthase (GBSS), the last enzyme catalyzing the formation of amylose (Vandeputte and Delcour 2004).

In fact, the rice amylose content is determined by the activity level of GBSSI, which in turn depends on the three functional alleles at the *wx* locus: i) the wx^a allele, ii) the wx^b allele derived from the wx^a allele during domestication of rice, and iii) the wx^{op} (opaque) allele, also derived from the wx^a allele independently of the wx^b gene. Unlike wx^a rice, wx^{op} contains ~10% amylose (Hirano *et al* 1998). The wx^a and wx^b alleles regulate the level of GBSS. Rice plants with the wx^a alleles (*indica* varieties) produce more GBSS than wx^b allele (*japonica* varieties) resulting in higher amylose content than in the japonica rices. The mutation in wx^b is a single base substitution of GBSSI, changing the sequence at the splice site in the first intron of GBSSI from AGGTATA to AGTTATA (Olsen and Purugganan 2002). Recently it is reported that a CT microsatellite length in the 5'untranslated region (UTR) of exon 1 of the GBSS gene discriminates between different types of amylase (Jayamani *et al* 2007; Ayres *et al* 1995; Bligh *et al* 1995):

i) rices with 18-19 CT repeats are low amylose temperate *Japonica* rices (wx^b)
ii) rices with 14-20 CT repeats are intermediate amylose tropical *Japonica* rices (wx^a)
iii) rices with 8,10 or 11 CT repeats are high amylose *indica* rices (wx^a)
iv) rices with 17 CT repeats have evolved in several phylogenetic categories including waxy and carry either the wx^a or wx^b allele.

However, there are no published reports on characterization of GBSSI in Basmati rice. Based on the published literature from non-Basmati rice one can deduce that Basmati might be having wx^a allele of waxy gene as well as 14-20 CT repeats in waxy region resulting in intermediate amylose content.

Aroma (ARM)

The QTLs, *qARM8.3* and *qARM8.4* were mapped on to the same region of chromosome 8 as that reported by Ahn *et al* (1992) and Lorieux *et al* (1996). Of seven QTLs for aroma four from Basmati370 and three from Jaya explained the increased effect suggesting that the environment where the experiment was conducted seemed to be unsuitable for expression of all aroma QTLs in Basmati370. Moreover, Basmati needs cool temperatures during flowering for expression of pleasant aroma. In Hyderabad where experiment was conducted the temperatures were higher compared to the Himalayan region which is the native habit of Basmati and hence it is likely that some QTLs or major QTLs would have gone undetected. Recently, a major QTL for aroma on chromosome 8 was fine mapped and a gene which encodes a betain aldehyde dehydrogenase (*bad-2*) was identified (Bradbury *et al* 2005). However, the exact role of *bad-2* enzyme in aroma development remains to be established. Higher level of 2-AP in aromatic rices was attributed to the presence of mutations caused either by 8 bp deletion and three SNPs in the exon 7 (Bradbury *et al* 2005) or 7 bp deletion in the exon 2 (Shi *et al* 2008) of *bad-2* on chromosome 8 resulting in a loss of function of the *fgr* gene.

Chalkiness (CHK)

The co-localization of QTLs for grain breadth and chalkiness as well as positive and significant correlation between these two traits observed in the present study suggests that breeders can simultaneously improve these two traits. These results are consistent with the earlier study by Tan *et al* (2000) who also mapped the QTLs for grain width and chalkiness at a marker interval of RG360 and C37349 on same region of the chromosome 5.

Stable QTLs of promise

QTLs, which are consistently detected over a range of environments are considered to be "stable QTLs", and are preferred targets in crop improvement. Though the present study carried out in single environment, together with the results of previous studies, seven QTLs viz., *qPH1.1* (Sasaki *et al* 2002; Speilmeyer *et al* 2002; Monna *et al* 2002)*, qGL3.2, qGB5.2, qLB3.3, qLB5.1* (Wan *et al* 2005; Fan *et al* 2006)*, qASV6.1* (Shu *et al* 2006) and *qARM8.3* (Bradbury *et al* 2005) that are associated with six traits of Basmati could be considered as stable QTLs. According to Wan *et al* (2005) QTLs with major effects are more likely to behave as stable QTLs across

multiple environments. These QTLs, apart from their suitability in improvement of the traits concerned, can also serve as potential candidates for fine mapping. These stable QTLs also facilitate development of near-isogenic lines and advanced breeding lines. Further, several QTLs, each with different environment specificity, can be introgressed into a single genotype to develop phenotypes stable over a range of environments. In fact, in conventional plant breeding selections are made in target environment and testing is done in multiple diverse environments. This exercise is cumbersome and time consuming. However, use of stable QTLs based selection can accelerate the pace of selection process in rice breeding programme.

Genetic dissection of putative QTLs identified in the study

Contradictory hypotheses have been invoked over the past few years about the major genes corresponding to the detected QTLs in cereals in general and rice in particular. For instance, Yano *et al* (1996) reported that three of the five putative QTLs for heading date were mapped to the same region where genes for photoperiod sensitivity in rice are located. Beavis *et al* (1991) reported that map location of a QTL for dwarfing in maize coincident with that of QTL for plant height. Conversely, Xiao *et al* (1996) observed that none of the QTLs for plant height was mapped to regions of semi dwarf or dwarfing genes. In the present study, 17 promising QTLs were identified. Out of these, six QTLs are consistent with the ones identified in previous studies (Table 1.9). The present study demonstrates that it is possible to detect major QTLs employing primary mapping populations like F_2 derived from a cross of potential parents.

Potential markers for marker assisted selection (MAS)

It is difficult and time consuming for breeders to improve desired rice grain appearance by conventional breeding, especially before crop maturity. Molecular markers, tightly linked to these target traits, could help in reliable selection within shorter duration of time. For instance, recently, Liu *et al* (2006) employed CAPS (cleaved amplified polymorphic sequence) markers for introgressing waxy locus to reduce amylose content in two elite Chinese cultivars. Prerequisite for identification of markers for MAS is the tight association of marker with the target QTL. According to Wan *et al* (2006) marker-QTL association should be as low as possible (< 2.6cM) to avoid linkage drag during introgression of the QTL(s) of interest. Wan *et al* (2006) have identified seven candidate markers for marker assisted introgression of the gene *gl-3* governing grain

Table 1.9 Correspondence between known major genes and promising QTLs identified in the present study

Trait	Chr.	Marker interval	PVE*	Gene	References
Plant height (cm)	1	RM302-RM11968	15.418	Gibberellin- 20 oxidase 2	Monna *et al*, Speilmeyer *et al* (2002)
Filled grains (no.)	1	RM11968-RM14	22.677	Gibberellin- 20 oxidase 2	Monna *et al*, Speilmeyer *et al* (2002)
Single plant yield (g)	9	RM107-RM566	82.155		
Grain length (mm)	3	RM353-JL14	46.065	Putative transmembrane protein	Fan *et al* (2006)
	5	RM430-RM18600	17.468		
Grain breadth (mm)	5	RM430-RM18600	17.149		
Length-Breadth ratio	3	RM282-RM16	74.353		
	3	RM353-JL14	22.342	Putative transmembrane protein	Fan *et al* (2006)
	3	JL14-RM5864	24.958		
	5	RM430-RM18600	46.531		
Elongation ratio	5	RM289-RM430	19.931		
	5	RM430-RM18600	18.931		
Alkali spreading value	6	RM276-RM527	71.735	soluble starch synthase II-3	He *et al* (1999); Lanceras *et al* (2000)
Aroma	8	RM404-RM483	20.226	Betain aldehyde dehydrogenase-2	Bradbury *et al* (2005)
Chalkiness	3	RM130-RM545	15.083		
	4	RM564-RM348	63.795		
	4	RM348-RM280	25.092		

*PVE-Phenotypic variance explained by the QTL

Table 1.10 List of microsatellite markers identified in the present study for possible marker assisted selection (MAS) for different traits of Basmati rice

S.No.	Trait	Marker	Marker-QTL distance (cM)	PVE* (%)
1	Panicle length	RM276	2	0.82
2	Single plant yield	RM263	0	4.06
3	Spikelet number	RM171	1.29	6.67
4	Grain length	JL14	1.7	46.07
5	Grain breadth	RM473A	0	1.65
6	Length breadth ratio	RM473A	0	3.93
7	Length breadth ratio	JL14	2	24.96
8	Grain length after cooking	RM247	0	2.68
9	Chalkiness	RM130	0	15.08

* PVE- Phenotypic variance explained by the QTL

length into leading rice cultivars. In the present study nine candidate markers for eight different traits have been identified based on their association with the concerned QTLs (Table 1.10).

Of the nine markers, JL14 and RM130 seem to be more useful. Apart from being reasonably tightly linked (1.7 cM and 0 cM, respectively), they contribute higher phenotypic variance for the traits of grain length, length-breadth ratio and chalkiness. In an earlier study Li *et al* (2004) have reported close association (< 2cM) of the marker JL14 to the grain length on chromosome 3. These findings as well as earlier reports suggest that the markers that are tightly linked to the grain length are on chromosome 3, grain breadth on chromosome 5, amylose content and alkali spreading value on chromosome 6 and these markers could be employed for marker assisted selection to improve the grain traits of Basmati rice (Table 1.10).

Candidate genes at QTL regions

One of the ultimate goals of genetic mapping of complex traits is to isolate candidate genes at QTL regions. In the present study, one QTL region i.e., at marker interval of RM430 and RM18600 on chromosome 5 spanning a physical distance of 327.1kb was chosen for identification of candidate genes as it controls grain breadth, grain length, length-breadth ratio and elongation ratio. Based on the rice genome sequence information 50 genes are present at this region (Table 1.8). Of the 50 genes, two predicted genes viz., AP2 transcription factor (Ohto *et al* 2005; Jofuku *et al* 2005) and RING E3 ligase (Song *et al* 2007) have been reported to be involved in controlling the seed size and weight. Hence, these two genes have been chosen for further studies. The details of the two genes are discussed below.

AP2 (*APETALA2)* transcription factor

AP2 belongs to a large family of DNA-binding proteins and is specific to plants. Each AP2 domain consists of approximately 68 amino acids. AP2 encoding genes are divided into three different classes based on the number of AP2 domains they harbor (Kim et al 2006).

- First class of genes encodes a protein containing two AP2 domains (AP2-like) eg., *AINTEGUMENTA* (*ANT*), and *Glossy15 (GL15)* (Jofuku *et al* 1994).
- Second class of genes encodes a protein with only one AP2 domain (*ERF* (Ethylene-responsive element-binding factor) like) eg., *ERFs, TINY, AtEBP and AB14.*

- Third class of genes, *RAV1* [Related to ABI3/VP1 (ABSCISIC ACID INSENSITIVE3/ VIVIPAROUS1)] and *RAV2*, encodes proteins that have two different DNA-binding domains, AP2 and B3.

Genetic studies have shown that AP2 has at least three major roles during development. AP2 is required for the specification of floral meristem identity in flower development (Irish and Sussex, 1990; Okamuro *et al* 1997) and has role in determining the seed size, seed weight and the accumulation of seed oil and protein. In addition, AP2 is also required for defining floral organ identity and proper specification of the outer two whorl floral organs, sepals and petals, and as a negative regulator of the floral homeotic MADS box gene *AGAMOUS* (Bowman *et al* 1989; Drews *et al* 1991). When mutated, ap2 flowers display distinct homeotic alterations where sepals are replaced by carpels and petals by stamens. Lastly, *AP2* expression is required for normal ovule and seed development (Jofuku *et al* 1994; Western *et al* 2001).

RING E3 ligase

The RING (Really Interesting New Gene)-type E3 ligase is characterized by the presence of a cysteine-rich domain that coordinates two zinc atoms. Unlike AP2 gene, RING type proteins are widely present in animals, plants, yeast, and viruses and are predicted to contain a large diversity of RING domains. The RING- type proteins function as E3 ubiquitin ligases *in vitro*, targeting proteins for ubiquitin dependant degradation by the 26S proteasome (Song *et al* 2007). Such RING proteins are involved in the regulation of numerous cellular processes including transcription, signal transduction, recombination and cell cycle progression. Previous studies in *A.thaliana* have shown that the RING-type (C3H2C3) protein *BIG BROTHER*, which has E3 ligase activity, function as a central negative regulator of floral organ size, most likely by marking cellular proteins for degradation (Song *et al* 2007).

Control of seed size in Basmati rice

Seed/grain size is one of the important quality traits of Basmati rice and played an important role in the evolution of cereal crops. Since the dawn of agriculture, food grains have been subjected to intense selection and breeding for seed size as well as for other qualities traits, and most of the grains consumed today have seeds far larger than their wild relatives.

Domestication has greatly enhanced the diversity of grain shape and size in addition to other changes. Although the preference for rice grain characters varies with consumer groups, long and slender grain rice is preferred by the majority of the consumers in Asia and USA (Fan *et al* 2006). According to the notification standards issued by the Ministry of Commerce, Govt. of India, the minimum length is 7mm and length-width ratio is 3.5 for A-grade Basmati rice (Amarawati *et al* 2008). Hence, elucidating the genetic mechanism by which seed size is controlled is important for rice improvement programs. Based on the published literature and current study we have proposed a model to explain genetic control of grain size in Basmati rice (Figure 1.14).

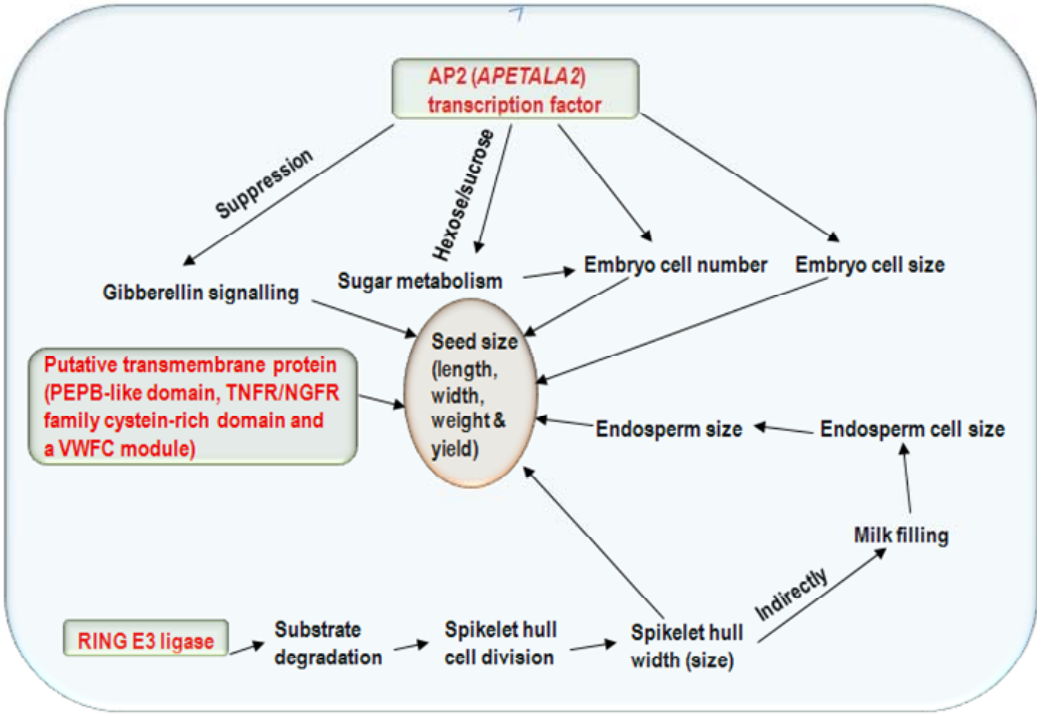


Figure 1.14 Proposed model for control of seed size in Basmati rice.
Candidate genes are highlighted in boxes with red colour text.

In angiosperms (flowering plants), seed development begins with double fertilization of the egg cell (n) and the central cell (2n) of the haploid embryo sac (female gametophyte) with pollen grains (n) (male gametophytes) within the maternal tissues of the ovule. This event results in the formation of the diploid embryo (2n) and the triploid endosperm (3n), respectively. Seed development is marked by rapid growth of endosperm and embryo, until

seed maturation. Simultaneously, maternal ovule (integument) also undergoes regulated growth to accommodate growing embryo and endosperm and integuments of ovule ultimately comprise mature seed coat. Hence, seed size is the consequence of three different growth programs: those of the diploid embryo, the triploid endosperm, and the diploid maternal ovule. The control and coordination of these growth programs are under genetic regulation. In Basmati rice, seed size seems to be controlled by at least three major genes and a few modifiers. Based on the speculation that AP2 and RING E3 ligase on chromosome 5 and putative transmembrane protein encoding gene on chromosome 3 are the candidate genes controlling the grain size a model for the control of Basmati grain size is proposed as shown in Figure 1.14.

AP2 is reported to act maternally to control seed mass by regulating embryo cell size and cell number (Jofuku *et al* 2005). Also, it controls seed mass through its effects on sugar metabolism by changing the ratio of hexose to sucrose during seed development (Weber *et al* 1997). In addition, genetic and physiological studies suggest that AP2 acts in part by suppressing gibberellin signaling to ensure uniformity in flower development against normal physiological fluxes (Okamuro *et al* 1997). RING type protein with E3 ubiquitin ligase, on the other hand, acts as a novel negative regulator of cell division that regulates rice seed development by targeting its substrate(s) to the 26S proteasome regulating proteolysis. Loss of the gene controlling this protein causes increased endosperm cell number and size, resulting in a larger (wider) spikelet hull and accelerate the grain milk filling rate, resulting in enhanced grain width, weight and yield (Song *et al* 2007). Recently, a gene encoding a putative transmembrane protein having PEPB-like domain, a putative TNFR/NGFR family cystein-rich domain and a VWFC module, has been identified. Loss of function of this protein results in long grains. This gene is also reported to govern grain width, grain weight and thickness (Fan *et al* 2006). However, further investigations are required to unravel the mechanism by which seed size is controlled in Basmati rice.

QTL clusters for grain appearance quality traits

Grain length, grain breadth/width, length/breadth (LB) ratio and chalkiness are important attributes determining the grain appearance quality of Basmati rice. In the present study three QTL clusters for grain appearance traits of Basmati have been identified on different chromosomes (Figure 1.11). Two QTL clusters on chromosome 5, at the marker interval of RM430 and RM18600 comprising *qGL5.1, qGB5.2, qLB5.1* and *qER5.2* and the second one in the

marker interval of RM289 and RM430 consists of *qGB5.1, qER5.1* and *qCHK5.1.* The third cluster comprises of *qGB3.1, qLB3.1* is on chromosome 3. Similar results were also obtained by Wan *et al* (2005) except that they identified one additional cluster on chromosome 1 for grain length, length width ratio and chalkiness. Recently, Zhang *et al* (2006) reported a QTL cluster on chromosome 8 for four traits viz., spikelets per panicle, grains per panicle, heading date and plant height. In the current study, however, no QTL clusters have been identified for the agronomic traits.

Several earlier studies have demonstrated that QTLs for correlated traits often map to the same chromosome regions (Paterson *et al* 1991, Lin *et al* 1996, Xiao *et al* 1996, Tan *et al* 1999). For instance, QTLs related to highly correlated traits like grain breadth, grain length, and length-breadth ratio have found to be located on the same genomic region of chromosome 5 (Wan *et al* 2005; Tan *et al* 2000) (Table 1.7). However, this trend was not seen for other significantly correlated traits such as plant height, panicle length, filled grains and single plant yield. The QTLs relating to these traits have been mapped on to different chromosomes implying that these traits are possibly controlled by independent and unrelated genes. Classical quantitative genetics assumes that trait correlation can be attributed to the effect of pleiotropy or to the tight linkage of causative genes. If pleiotropism is the major reason, coincidence of both the location of QTL for related traits as well as the directions of their genetic effects can be expected. If tight physical linkage of the genes is the major reason, the directions of the genetic effect of QTL for different traits may be different although the coincidence of the locations of QTL can still be expected (Wan *et al* 2005). In the present study, one QTL region at a marker interval of RM289 and RM430 on chromosome 5 harbors QTLs for chalkiness and grain breadth and also exhibits similar increasing effect from Jaya indicating pleiotropism rather than close linkage between these QTLs as the reason for correlation of these two traits. However, in the region of RM430 and RM18600 effects of three QTLs for grain breadth, grain length, and length-breadth ratio are in different directions suggesting involvement of tightly linked genes as the cause for the correlation of these traits.

Physical grain quality of rice assumes importance to a great extent in trade and milling. However, cooking and eating qualities, being the most important determinants from consumers' angle, are widely studied. Parental choice in crop breeding research is done keeping in view the

quality needs/preference of targeted users. Also, it is important to have knowledge of genetics of various quality indices for directed improvement or molecular breeding approaches. Majority of the indices, both grain and cooking/eating qualities are polygenically controlled. Often one finds contradictory reports on the genetic architecture of such indices. It is generally believed that amylose content, for instance, a simply inherited Mendelian trait is associated with several modifiers (Bollich and Webb 1973; McKenzie and Rutger 1983; Kumar and Khush 1988; Tan *et al* 1999). There are also reports that suggest it to be a quantitative trait (Puri and Siddiq (1980), Lanceras (2000)). A QTL for amylose content that accounted for 91 percent of phenotypic variation was found to reside on chromosome 6 (He *et al* 1999). A QTL identified on Chromosome 4 and continuous variation in the mapping population as observed in the present study, suggests that, in all probability, the amylose content a quantitative trait. The classical genetic (McKenzie and Rutger 1983; Umemoto *et al* 2002) and molecular mapping (Bao *et al* 2004; He *et al* 1999; Tian *et al* 2005) approaches mapped a major gene for alkali spreading value on chromosome 6. Nevertheless, contradictory results have been reported about the mode of inheritance of this trait in rice. For instance, Ghosh and Govindswamy (1972) reported that ASV was quantitatively inherited whereas McKenzie and Rutger (1983) observed that the segregation pattern in crosses did not conform to any identifiable genetic model. No matter whether a trait is Mendelian or polygenically inherited, detection of promising markers could be of value for directed improvement of the trait in question. Failure to detect a QTL for any given trait or bias in QTL mapping may be attributed to any one or many of the following reasons:

(i) two parental inbreds may have identical alleles at a linked QTL, which may thus escape detection,

(ii) the parents may have different alleles at the QTL, having equivalent expression for the particular trait,

(iii) type and size of mapping populations,

(iv) type/number of molecular markers used in the construction of genetic map,

(v) the statistical threshold for detecting putative loci,

(vi) heritability of target traits (Yano and Sasaki 1997)

(vii) low level of polymorphism between parents, and

(viii) properties of the genome such as variation in recombination rates and gene density among different regions of the genome as opined by Norman 2001.

Based on the results of the current study as well as of the previous reports the following rice chromosomes can be termed as the ones that harbour the QTLs/genes controlling important agronomic and quality traits.

- Chromosome 1 is termed as "**architecture chromosome**" as it harbours gene(s) for plant height and grain number.
- Chromosomes 3 and 5 include QTLs/genes for grain length, grain breadth, length-breadth ratio/grain size and chalkiness and can be named as "**grain appearance chromosomes**"
- Chromosome 6 is referred to as "**quality chromosome**" as it harbours genes controlling most of the important quality traits such as amylose content, alkali spreading value, gel consistency, water absorption, volume expansion and elongation of cooked kernel.
- Chromosome 8 is named as "**fragrant chromosome**" as it is having gene(s) governing aroma/fragrance.

Conclusions

1.5 Conclusions

Markers derived from the stable and well behaved QTLs can be candidates in application of marker-assisted selection (MAS) of the corresponding traits in rice breeding. QTLs of promise, identified in the present study, for quantitative and quality traits of Basmati can be transferred to high yielding varieties and parents of heterotic hybrids by recombination breeding using the associated marker(s). Besides, the promising QTLs can also serve as candidates for further fine mapping and map-based cloning projects in future. Rapid advances being made in the field of rice genetics and genomics would help in identification and characterization of the candidate genes underlying QTLs, which can be utilized in transgenic rice development. Being a model cereal crop with all the available genetic and genomic resources, understanding of rice QTLs in other cereal crops would facilitate their positional cloning. By pyramiding the genes from "**quality chromosome**" (chromosome 6) and "**architecture chromosome**" (chromosome 1) into a single variety it may be possible to develop a high yielding superior quality rice variety so that it can be available to common man who dream to taste quality rices like Basmati.

References

1.6 References

Abdelkhalik AF, Shishido R, Nomura K and Ikehashi H (2005) QTL-Based Analysis of Heterosis for Grain Shape Traits and Seedling Characteristics in an *indica-japonica* Hybrid in Rice (*Oryza sativa* L.). Breed. Sci. 55: 41-48.

Ahn SN, Bollich CN, Tanksley SD (1992) RFLP tagging of a gene for aroma in rice. Theor Appl Genet 84:825–828

Ahn SN, Bollich CN, McClung AM, Tanksley SD (1993) RFLP analysis of genomic regions associated with cooked kernel elongation in rice. Theor Appl Genet 87:27–32

Aluko G, Martinez C, Tohme J, Castano C, Bergman, C and Oard JH (2004) QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* x *O. glaberrima*. Theor Appl Genet 109: 630-9.

Amarawathi Y, Singh R, Singh A, Singh V, Mohapatra T, Sharma T and Singh N (2008) Mapping of quantitative trait loci for basmati quality traits in rice (*Oryza sativa* L.). Molecular Breeding 21: 49-65.

Archak S, Lakshminarayanareddy V and Nagaraju J (2007) High-throughput multiplex microsatellite marker assay for detection and quantification of adulteration in Basmati rice (Oryza sativa). Electrophoresis 28: 2396-405.

Arumuganathan K and Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. 9:208-218

Ashikari M and Matsuoka M (2006) Identification, isolation and pyramiding of quantitative trait loci for rice breeding. Trends Plant Sci 11: 344-50.

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H and Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science 309: 741-5.

Ayres NM, McClung AM, Larkin PD, Bligh HFJ, Jones CA. and Park WD (1997) Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. Theor. Appl. Genet. 94, pp. 773–781

Bao JS, Sun M and Corke H (2002) Analysis of the genetic behavior of some starch properties in indica rice (*Oryza sativa* L.): thermal properties, gel texture, swelling volume. Theor Appl Genet 104: 408-413.

Bao JS, Sun M and Corke H (2004) Analysis of the quantitative trait loci of some starch properties in indica rice (*Oryza sativa* L.): thermal properties, gel texture, swelling volume. Theor Appl Genet 39: 379-385.

Beavis WD, Gram D, Albertson M, Fincher R (1991) Quantitative trait loci. for plant height in four maize populations and their association with qualitative genetic loci Theor Appl Genet 83:141-145

Becker J, Vos P, Kuiper M, Salamini F and Heun M (1995) Combined mapping of AFLP and RFLP markers in barley. Mol Gen Genet 249: 65-73.

Biovin K, Deu M, Trouche G, and Hamon P (1998) Integration of AFLP markers into an RFLP linkage map of sorghum. In: Abstracts of Plant and Animal Genome VI. International conference on the status of plant and animal genome research, January 18-22, 1998, San Diego, CA, USA, pp130.

Bligh HFL, Till RI and Jones CA (1995) A microsatellite sequence closely linked to the wx gene of Oryza sativa. Euphytica 86 pp. 83–85

Bollich CN and Webb BD (1973) Inheritance of amylose in two hybrid populations of rice. Cereal Chem 50: 631-636.

Bowman JL, Smyth DR, and Meyerowitz EM (1989) Genes directing flower development in *Arabidopsis*. Plant Cell 1, 37–52

Bradbury LM, Fitzgerald TL, Henry RJ, Jin Q and Waters DL (2005) The gene for fragrance in rice. Plant Biotechnol J 3: 363-70.

Buttery RG, Ling LC, Juliano BO and G, Turnbaugh JG (1983) Cooked rice aroma and 2-acetyl-1-pyrroline. J. Agric. Food Chem 31: 823-826.

Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SE and *et al* (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138: 1251-74.

Churchill GA and Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138: 963-971.

Drews GN, Bowman JL, Meyerowitz EM. (1991) Negative regulation of the *Arabidopsis* homeotic gene *AGAMOUS* by the APETALA2 product. Cell 65, 991–1002

Falconer DS and Mackay TFC (1996) Introduction to Quantitative Genetics. 4th edition, Longman, Harlow, UK.

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X and Zhang Q (2006) GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet 112: 1164-71.

Garris AJ, Tai TH, Coburn J, Kresovich S and McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. Genetics 169: 1631-8.

Ghosh AK and Govindaswamy S (1972) Inheritance of starch iodine blue and alkali digestion value in rice and their genetics association. II Riso 21:123-132

Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. Theor Appl Genet 74: 21-30.

Govindaraj P, Arumugachamy S and Maheswaran M (2005) Bulked segregant analysis to detect main effect QTL associated with grain quality parameters in Basmati 370/ASD 16 cross in rice (*Oryza sativa* L) using SSR markers. Euphytica 144: 61-68.

Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, Yamamoto T, Lin SY, Antonio BA, Parco A, Kajiya H, Huang N, Yamamoto K, Nagamura Y, Kurata N, Khush GS and Sasaki T (1998) A high-density rice genetic linkage map with 2275 markers using a single F2 population. Genetics 148: 479-94.

He P, Li SG, Qian Q, Ma YQ, Li JZ, Wang WM, Chen Y and Zhu LH (1999) Genetic analysis of rice grain quality. Theor Appl Genet 98: 502-508.

He G, Luo X, Tian F, Li K, Zhu Z, Su W, Qian X, Fu Y, Wang X, Sun CQ and Yang J (2006) Haplotype variation in structure and expression of a gene cluster associated with a quantitative trait locus for improved yield in rice. Genome Res. 16: 618-626.

Hirano HY, Eiguchi M and Sano Y (1998) A single base change altered the regulation of the Waxy gene at the posttranscriptional level during the domestication of rice. Mol Biol Evol 15: 978-87.

Ikeda K, Ito M, Nagasawa N, Kajunko K and Yasu O (2007) Rice aberrant panicle organization 1, encoding an f-box protein, regulates meristem fate. The plant journal: 51(6):1030-40

Irish VF, and Sussex IM. (1990) Function of the *apetala-1* gene during *Arabidopsis* floral development. Plant Cell **2,** 741–753

Ishimaru K, Ono K and Kashiwagi T (2004) Identification of a new gene controlling plant height in rice using the candidate-gene strategy. Planta 218: 388-95.

Iwata T, Nagamatsu T and Omura T (1964) Abnormal segregation of waxy and apiculus coloration by a gametophytic gene belonging to the first linkage group in rice. Jpn J Breed 14: 33-39.

Jain S, Jain RK and McCouch SR (2004) Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently-labeled microsatellite markers. Theor Appl Genet 109: 965-77.

Jayakody L and Hoover R (2002) The effect of lintnerization on cereal starch granules. Food Research International 35: 665-680.

Jayamani P[a, b], Negrão S[a], Brites[c] C and Oliveira MM (2007) Potential of *Waxy* gene microsatellite and single-nucleotide polymorphisms to develop *japonica* varieties with desired amylose levels in rice (*Oryza sativa* L.) Journal of Cereal Science 46, 2, 178-186

Jofuku KD, Omidyar PK, Gee Z & Okamuro JK (2005) Control of seed mass and seed yield by the floral homeotic gene APETALA2 Proc. Natl. Acad. Sci. USA 102, 3117–3122.

Jofuku KD, den Boer BGW, van Montagu M, Okamuro JK (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2*. Plant Cell 6, 1211–1225.

Juliano BO and Villareal CP (1993) Grain quality evaluation of world rices. International Rice Research Institute, Manila, Philippines.

Juliano BOA (1971) Simplified assay for milled-rice amylose Cereal Science Today 16: 334-340.

Kim S, Soltis PS, Wall K and Soltis DE (2006) Phylogeny and Domain Evolution in the APETALA2-like Gene Family Mol. Biol. Evol. 23(1):107–120.

Knox MR and Ellis THN (2002) Excess Heterozygosity Contributes to Genetic Map Expansion in Pea Recombinant Inbred Populations Genetics 162: 861-873

Kumar I and Khush GS (1988) Inheritance of amylose content in rice (*Oryza sativa* L.). Euphytica 38: 261-269.

Kumar S, Gill BS and Faris JD (2007) Identification and characterization of segregation distortion loci along chromosome 5B in tetraploid wheat. Mol Genet Genomics 278: 187-96.

Lambrides CJ, Godwin ID, Lawn RJ and Imrie BC (2004) Segregation distortion for seed testa color in Mungbean (*Vigna radiata* L. Wilcek). J Hered 95: 532-5.

Lanceras JC, Huang ZL, Naivikul O, Vanavichit A, Ruanjaichon V and Tragoonrung S (2000) Mapping of genes for cooking and eating qualities in Thai jasmine rice (KDML105). DNA Res 7: 93-101.

Lander ES and Botstein D (1989) Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. Genetics 121: 185-199.

Li J, Xiao J, Grandillo S, Jiang L, Wan Y, Deng Q, Yuan L and McCouch SR (2004) QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*O. sativa* L.) and African (*O. glaberrima* S.) rice. Genome 47: 697-704.

Li W, Lin Z and Zhang X (2007) A novel segregation distortion in intraspecific population of Asian cotton (*Gossypium arboretum* L.) detected by molecular markers. J Genet Genomics 34: 634-40.

Lin SY, Ikehashi H, Yanagihara S, and Kawashima A (1992) Segregation distortion via male gametes in hybrids between Indica and Japonica or wide-compability varieties of rice (*Oryza sativa* L.). Theor Appl Genet 84:812-818

Lin HX, Qian HR, Zhuang JY, Lu J, Min SK, Xiong ZM, Huang N, and Zheng KL (1996) RFLP mapping of QTLs for yield and yield related characters in rice (*Oryza sativa* L.). Theor Appl Genet, 92: 920-927.

Lincoln S, Daly M, and Lander E, (1993) Mapping genes controlling quantitative traits with MAPMAKER/QTL 1.1: a tutorial and reference manual. 2nd edn. Whitehead Institute Technical Report, Cambridge, UK.

Little RR, Hilder GB and Dawson EH (1958) Differential effect of dilute alkali on 25 varieties of milled white rice. Cereal Chem 35: 111-126.

Liu, CJ, Witcombe JR, Pittaway TS, Nash CT, Busso CS and Gale MD (1994) An RFLP-based genetic map of pearl millet (*Pennisetum glaucum*). Thero Appl Genet 89.

Liu F, Wu XL and Chen SY (2000) Segregation distortion of molecular markers in recombinant inbred populations in soybean (*G. max*). Yi Chuan Xue Bao 27: 883-7.

Liu QQ, Li QF, Cai XL, Wang HM, Tang SZ, Yu HX, Wang ZY, Gu MH (2006) Molecular Marker-Assisted election for Improved Cooking and Eating Quality of Two Elite Parents of Hybrid Rice Crop Sci 46: 2354-23

Lorieux M, Petrov M, Huang N, Guiderdoni E and GhesquiÃ¨re A (1996) Aroma in rice: genetic analysis of a quantitative trait. Theor. Appl. Genet 93: 1145-1151.

Lorieux MJ (2006) Mapdisto, A Tool For Easy Mapping of Genetic Markers. Plant & Animal Genomes XIV Conference January 14-18, Town & Country Convention Center San Diego, CA

McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T and Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Mol Biol 35: 89-99.

McCouch SR and Doerge RW (1995) QTL mapping in rice. Trends Genet 11: 482-7.

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom R, DeClerck G, Schneider D, Cartinhour S, Ware D and Stein L (2002) Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). DNA Res 9: 199-207.

McKenzie KS and Rutger JN (1983) Genetic analysis of amylose content, alkali spreading score, and grain dimensions in rice. Crop Sci 23: 306-311.

Monna L, Kitazawa N, Yoshino R, Suzuki J, Masuda H, Maehara Y, Tanji M, Sato M, Nasu S and Minobe Y (2002) Positional cloning of rice semidwarfing gene, sd-1: rice "green revolution gene" encodes a mutant enzyme involved in gibberellin synthesis. DNA Res 9: 11-7.

Nagamura Y, Antonio BA and Sasaki T (1997) Rice molecular genetic map using RFLPs and its applications Plant Molecular Biology Volume 35, Numbers 1-2,79-87

Nagaraju J, Kathirvel M, Kumar RR, Siddiq EA and Hasnain SE (2002) Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. Proc Natl Acad Sci U S A 99: 5836-41.

Nakagahra M, Omura T, Iwata N (1972) Gametophyte genes and their loci on the eleventh linkage group of cultivated rice. Japan J Breed, 22 (6): 305-312.

Nakamura Y (2002) Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue. Plant Cell Physiol 43: 718-25.

Nene YL (1998) Basmati rice: A distinct variety (cultivar) of the Indian subcontinent. Asian Agri-History 2: 321-330.

Ohto M, Fischer RL, Goldberg RB, Nakamura K and Harada JJ (2005) Control of seed mass by APETALA2 Proc. Natl. Acad. Sci. USA 102, 3123–3128.

Olsen KM and Purugganan MD (2002) Molecular evidence on the origin and evolution of glutinous rice. Genetics 162: 941-50.

Okamuro JK, Caster B, Villarroel R, van Montagu M, Jofuku KD (1997) The AP2 domain of *APETALA2* defines a large new family of DNA binding proteins in *Arabidopsis*. Proc. Natl. Acad. Sci. USA 94, 7076–7081

Qi-Jun Z, Shao-Ping YE, Jie-Qin LI, Bing Z, Yong-Shu L, Yong P and Ping LI (2006) Construction of a Microsatellite Linkage Map with Two Sequenced Rice Varieties Acta Genetica Sinica 33: 152-1.

Paterson A, Hewitt J. Peterson S, Lincoln S, Lander, E and Tanksley S. (1991) Mendelian Factors Underlying Quantitative Traits: Comparison across Species, Generations, and Environments. Genetics 127: 169-179.

Pham JL, Glaszmann, JC, Sano, R, Barbier, P, Ghesquière, A and Second, G (1990) Isozyme markers in rice: genetic analysis and linkage relationships. Genome 33 348-359.

Pooni HS, Kumar I and Khush GS (1993) Genetical control of amylose content in selected crosses of indica rice. Heredity 70: 269-280.

Pradeep M, Sarla N, Laxminarayana VR and Siddiq EA (2005) Identification and mapping of yield and yield related QTLs from an Indian accession of *Oryza rufipogon*. BMC Genetics 6: 33.

Puri RP and Siddiq EA (1980) Inheritance of gelatinization temperature in rice. Indian J Genet Plant B reed.40:4 60-62

Redona ED and Mackill DJ (1998) Quantitative trait locus analysis for rice panicle and grain characteristics. Theor Appl Genet 96 957-963.

Rick CM and Smith PG (1953) Novel variation in tomato species hybrids. Am. Nat. 88.

Sandler L and Novitski E (1957) Meiotic drive as an evolutionary force. Am. Nat. 41: 105-110.

Sasaki T, Song J, Koga-Ban Y, Matsui E, Fang F, Higo H, Nagasaki H, Hori M, Miya M, Murayama-Kayano E, Takiguchi T, Takasuga A, Niki T, Ishimaru K, Ikeda H, Yamamoto Y, Mukai Y, Ohta A, Miyadera N, Havukkala I, Minobe Y (1994) Toward cataloguing all rice genes: large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. Plant J 6: 615-624

Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS, Kitano H and Matsuoka M (2002) Green revolution: a mutant gibberellin-synthesis gene in rice. Nature 416: 701-2.

Sax K (1923) The Association of Size Differences with Seed-Coat Pattern and Pigmentation in *Phaseolus vulgaris*. Genetics 8: 552-.

Shi W Yang Y Chen and Xu M (2008) Discovery of a new fragrance allele and the development of functional markers for the breeding of fragrant rice varieties Mol Breeding DOI 10.1007/s11032-008-9165-7

Shu XL, Shen SQ, Bao JS, Wu DX, Nakamura Y and Shu QY (2006) Molecular and biochemical analysis of the gelatinization temperature characteristics of rice (*Oryza sativa* L.) Starch granules, J Cereal Sci. 44, 1:40-48

Singh P and Narayan SS (1997) Biometrical techniques in Plant breeding Kalyani publishers

Singh RK, Singh US and Kush GS (2000) Aromatic rices. Oxford &IBH Publishing Co. Pvt. Ltd

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase Nature Genetics 39;623-630

Sood BC and Siddiq EA (1978) A rapid technique for scent determination in rice. Indian J. Genet. Plant Breed 38:268-271.

Spielmeyer W, Ellis MH and Chandler PM (2002) Semidwarf (*sd-1*), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. Proc Natl Acad Sci U S A 99: 9043-9048.

Stuber CW and Edwards MD (1986) Genotypic selection for improvement of quantitative traits in maize using molecular marker loci In: Proceedings of the 41st Annual Maize and Sorghum Research Conference American Seed Trade Association, Chicago, Illinois, pp 40-83.

Subudhi PK, and Huang N (1999) RAPD mapping in a doubled haploid population of rice (*Oryza sativa* L.) Hereditas 130, 41-49.

Sweeney M and McCouch S (2007) The complex history of the domestication of rice. Ann Bot (Lond) 100: 951-7.

Tan YF, Li JX, Yu SB, Xing YZ, Xu CG and Zhang QF (1999) The three important traits for cooking and eating qualities of rice grain are controlled by a single locus. Theor Appl Genet 99:642-648.

Tan YF, Xing YZ Li JX, Yu SB, Xu CG and Zhang QF (2000) Genetic bases of appearance quality of rice grains in Shanyou 63, an elite rice hybrid. Theor Appl Genet 101: 823-829.

Tanksley SD, Bernachi D and Beck-Bunn T (1998) Yield and quality evaluations on a pair of processing tomato lines nearly isogenic for the Tm2(a) gene for resistance to the tobacco mosaic virus. Euphytica 99 77-83.

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB and *et al* (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141-60.

Taylor DR and Ingvarsson PK (2003) Common features of segregation distortion in plants and animals. Genetica 117: 27-35.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S and McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11: 1441-52.

Tian R, Jiang GH, Shen LH, Wang LQ and He YQ (2005) Mapping quantitative trait loci underlying the cooking and eating quality of rice using a DH population. Molecular Breeding 15: 117-124.

Umemoto T, Yano M, Satoh H, Shomura A and Nakamura Y (2002) Mapping of a gene responsible for the difference in amylopectin structure between japonica-type and indica-type rice varieties. Theor Appl Genet 104: 1-8.

Vandeputte GE and Delcour JA (2004) From sucrose to starch granule to starch physical behaviour: a focus on rice starch. Carbohydrate Polymers 58: 245-266.

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

Vemireddy LR, Archak S and Nagaraju J (2007) Capillary electrophoresis is essential for microsatellite marker based detection and quantification of adulteration of Basmati rice (*Oryza sativa*). J Agric Food Chem 55: 8112-7.

Virk PS, Ford-Lloyd BV, Jackson MT and Nwbury HJ (1995) Use of RAPD for the study of diversity within plant germplasm collections. Heredity 74: 170-179.

Virk PS, Newbury HJ, Jackson MT and Ford-Lloyd BV (2000) Are mapped markers more useful for assessing genetic diversity? Theor Appl Genet 100.607-613

Vos P (1998) AFLP fingerprinting of Arabidopsis. Methods Mol Biol 82: 147-55.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and *et al* (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-14.

Wan XY, Wan JM, Jiang L, Wang JK, Zhai HQ, Weng JF, Wang HL, Lei CL, Wang JL, Zhang X, Cheng ZJ and Guo XP (2006) QTL analysis for rice grain length and fine mapping of an identified QTL with stable and major effects. Theor Appl Genet 112: 1258-70.

Wan XY, Wan JM, Su CC, Wang CM, Shen WB, Li JM, Wang HL, Jiang L, Liu SJ, Chen LM, Yasui H and Yoshimura A (2004) QTL detection for eating quality of cooked rice in a population of chromosome segment substitution lines. Theor Appl Genet 110: 71-9.

Wan XY, Wan JM, Weng JF, Jiang L, Bi JC, Wang CM and Zhai HQ (2005) Stability of QTLs for rice grain dimension and endosperm chalkiness characteristics across eight environments. Theor Appl Genet 110: 1334-46.

Wang J, Wan X, Crossa J, Crouch J, Weng J, Zhai H and Wan J (2006a) QTL mapping of grain length in rice (*Oryza sativa* L.) using chromosome segment substitution lines. Genet Res 88: 93-104.

Wang S, Basten CJ and Zeng ZB (2006b) Windows QTL Cartographer 2.5. North Carolina State University, Raleigh, NC (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm).

Weber H, Borisjuk L and Wobus U (1996) Controlling seed development and seed size in *Vicia faba*: a role for seed coat-associated invertases and carbohydrate state. Plant J., 10(5):823-834

Weber H, Borisjuk L and Wobus U (1997) Sugar import and metabolism during seed development Trends in Plant Science, Volume 2, 5:169-174

Western TL, Burn J, Tan WL, Skinner DJ, Martin-McCaffrey L, Moffatt BA, Haughn GW (2001) Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in *Arabidopsis*. Plant Physiology 127, 998–1011.

Xiao J, Li J, Yuan L and Tanksley SD (1996) Identification of QTLs affecting traits of agronomic importance in a recombinant inbred population derived from a subspecific rice cross. Theor Appl Genet 92:230-244.

Xu Y, Shen Z, Chen Y and Zhu LH (1995) Distorted segregations of RFLP markers and their distribution on chromosomes in an indica-japonica F2 population of rice (*Oryza sativa* L.). Acta Bot Sin 37: 91-96.

Xu JL, Yu SB, Luo L J, Zhong DB, Me HW and Li ZK (2004) Molecular dissection of the primary sink size and its related traits in rice Plant Breeding 123,43-50

Yano M and Sasaki T (1997) Genetic and molecular dissection of quantitative traits in rice Plant Molecular Biology 35, 1-2, 145-153

Yano M, Harushima Y, Nagamura Y, Kurata N, Sasaki T and Minobe Y (1996) QTL analysis as an aid to tagging genes conferring heading date in rice. Proceeding of 3rd International Rice Genetics Symposium, International Rice Research Institute, Manila

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136: 1457-1468.

Zheng X, Wu JU, Lou XY, Xu HM and Shi CH (2008) The QTL analysis on maternal and endosperm genome and their environmental interactions for charcters of cooking quality in rice (*Oryza sativa* L.) Theor Appl Genet 116, 3; 335-342

Zhou LQ, Wang YP and Li SG (2006) Genetic analysis and physical mapping of Lk-4(t), a major gene controlling grain length in rice, with a BC2F2 population. Yi Chuan Xue Bao 33: 72-9.

Zhuang JY, Lin HX, Lu J, Qian HR, Hittalmani S, Huang N and Zheng KL (1997) Analysis of QTL X environment interaction for yield components and plant height in rice. Theor Appl Genet 95: 799-808.

Supporting information

1.7 Supporting information (SI)

SI Table 1.1 Chi square values of microsatellite markers showing segregation distortion among F_2 population of Basmati370/Jaya

S.No.	Marker	Chromosome	Chi square	Probability	Skewness
1	RM84	1	1159	0.00304	Heterozygote
2	RM490	1	12.08	0.00238	Basmati370
3	RM259	1	22.63	0.00001	Heterozygote
4	RM583	1	6.38	0.04121	Basmati370
5	RM579	1	13.24	0.00134	Heterozygote
6	RM9	1	10.06	0.00655	Heterozygote
7	RM306	1	16.81	0.00022	Jaya
8	RM473A	1	32.31	0.00000	Heterozygote
9	RM128	1	18.91	0.00008	Heterozygote
10	RM302	1	37.27	0.00000	Basmati370
11	RM212	1	18.93	0.00008	Heterozygote
12	RM11968	1	22.00	0.00002	Heterozygote
13	RM8278	1	12.75	0.00171	Heterozygote
14	RM14	1	19.90	0.00005	Heterozygote
15	RM5897	2	1185	0.00267	Heterozygote
16	RM71	2	10.57	0.00506	Heterozygote
17	RM327	2	12.42	0.00201	Basmati370
18	RM300	2	20.75	0.00003	Heterozygote
19	RM475	2	13.71	0.00106	Heterozygote
20	RM6318	2	15.74	0.00038	Jaya
21	RM263	2	1151	0.00316	Heterozygote
22	RM525	2	7.12	0.02841	Jaya
23	RM318	2	14.02	0.00090	Jaya
24	RM240	2	17.99	0.00012	Heterozygote
25	RM112	2	1114	0.00382	Heterozygote
26	RM138	2	6.64	0.03615	Jaya
27	RM545	3	8.98	0.01124	Heterozygote
28	RM517	3	12.45	0.00198	Heterozygote
29	RM251	3	20.87	0.00003	Heterozygote
30	RM282	3	17.27	0.00018	Jaya
31	JL14	3	15.04	0.00054	Jaya
32	RM353	3	18.00	0.00012	Heterozygote
33	RM5864	3	19.73	0.00005	Heterozygote
34	RM426	3	8.60	0.01360	Jaya
35	RM168	3	17.17	0.00019	Heterozygote
36	RM570	3	12.32	0.00211	Heterozygote
37	RM565	3	26.24	0.00000	Heterozygote
38	RM514	3	7.93	0.01901	Heterozygote
39	RM85	3	10.76	0.00461	Basmati370
40	RM551	4	19.19	0.00007	Heterozygote
41	RM518	4	13.24	0.00134	Heterozygote
42	RM564	4	28.62	0.00000	Heterozygote
43	RM252	4	8.41	0.01491	Heterozygote
44	RM241	4	13.31	0.00129	Jaya
45	RM127	4	6.64	0.03608	Jaya
46	RM280	4	1139	0.00336	Heterozygote
47	RM437	5	9.80	0.00746	Jaya
48	RM289	5	2121	0.00002	Heterozygote

S.No.	Marker	Chromosome	Chi square	Probability	Skewness
49	RM430	5	7.93	0.01901	Jaya
50	RM18600	5	10.49	0.00526	Jaya
51	RM161	5	6.18	0.04544	Jaya
52	RM233	5	1184	0.00268	Jaya
53	RM421	5	27.55	0.00000	Heterozygote
54	MX4	6	10.38	0.00557	Heterozygote
55	WXSSR	6	13.02	0.00149	Heterozygote
56	RM204	6	20.21	0.00004	Heterozygote
57	RM225	6	13.21	0.00136	Heterozygote
58	RM584	6	8.95	0.01139	Heterozygote
59	RM539	6	8.22	0.01643	Basmati370
60	RM3	6	23.13	0.00001	Heterozygote
61	RM162	6	7.16	0.02792	Jaya
62	RM340	6	7.61	0.02222	Jaya
63	RM481	7	8.11	0.01733	Jaya
64	RM11	7	9.89	0.00712	Jaya
65	RM336	7	16.30	0.00029	Heterozygote
66	RM248	7	15.26	0.00049	Heterozygote
67	RM408	8	26.69	0.00000	Heterozygote
68	RM152	8	10.82	0.00447	Heterozygote
69	RM310	8	3141	0.00000	Heterozygote
70	RM547	8	32.86	0.00000	Heterozygote
71	Fgr	8	22.13	0.00002	Heterozygote
72	RM72	8	1180	0.00273	Jaya
73	RM44	8	15.10	0.00053	Jaya
74	RM404	8	12.43	0.00200	Jaya
75	RM483	8	7.12	0.02839	Basmati370
76	RM339	8	17.71	0.00014	Jaya
77	RM42	8	27.74	0.00000	Heterozygote
78	RM502	8	9.18	0.01013	Jaya
79	RM464	9	53.70	0.00000	Heterozygote
80	RM321	9	10.41	0.00550	Basmati370
81	RM566	9	7.58	0.02258	Jaya
82	RM257	9	15.78	0.00038	Heterozygote
83	RM242	9	16.50	0.00026	Basmati370
84	RM201	9	25.91	0.00000	Basmati370
85	RM107	9	15.72	0.00039	Basmati370
86	OSR28	9	2178	0.00002	Basmati370
87	RM474	10	14.93	0.00057	Jaya
88	RM216	10	13.54	0.00115	Jaya
89	RM258	10	9.89	0.00712	Heterozygote
90	RM171	10	27.45	0.00000	Heterozygote
91	RM228	10	6.65	0.03600	Heterozygote
92	RM496	10	13.15	0.00140	Jaya
93	RM590	10	20.16	0.00004	Heterozygote
94	RM4B	11	23.58	0.00001	Jaya
95	RM1812	11	9.77	0.00755	Heterozygote
96	RM202	11	13.99	0.00091	Heterozygote
97	RM287	11	18.09	0.00012	Heterozygote
98	RM235	12	56.35	0.00000	Heterozygote

SI Table 1.2 QTLs identified in Basmati370/Jaya F_2 population employing interval mapping (IM) method

S.No.	Trait	Chr.	Marker interval	LFM	RFM	LOD	Additive	Dominance	PVE
1	Plant height	1	RM128-RM212	6	6.65	3.080	8.402	-3.899	15.595
2		1	RM212-RM302	8	6.27	3.398	7.430	-6.409	13.589
3		1	RM302-RM11968	12	9.79	4.865	8.366	-3.108	17.091
4		1	RM11968-RM14	0	24.01	3.818	5.115	1.294	5.560
5	No.of panicles	10	RM216-RM171	8	14.43	2.603	-1.683	1.287	19.774
6	Panicle length	2	RM6318-RM263	12	8.99	3.061	0.544	1.502	1.390
7		2	RM263-RM525	6	15.19	2.681	0.342	1.603	0.482
8		6	RM276-RM527	0	11.02	2.788	0.349	-1.557	0.607
9	Filled grains	1	RM11968-RM14	8	16.01	3.501	32.435	-25.570	23.994
10	Chaffy grains	1	RM8278-RM582	28	44	3.339	11.455	-12.330	5.349
11		3	RM282-RM16	12	20.96	4.770	42.569	-9.364	61.129
12		3	RM353-JL14	6	4.59	3.210	25.909	-10.741	22.173
13		3	JL14-RM5864	4	7.01	2.621	24.724	-7.656	20.473
14		3	RM5864-RM426	10	10.18	6.792	52.963	7.035	86.338
15		3	RM426-RM168	0	9.54	3.947	13.467	8.288	7.458
16		3	RM85-RM565	16	22	3.819	-2.750	-6.352	0.259
17		4	RM348-RM280	14	17.25	3.043	14.540	-9.953	6.513
18		9	RM107-RM566	30	42	2.901	-3.929	-3.078	0.432
19		10	RM216-RM171	12	10.43	2.508	-11.063	-8.865	3.678
20		12	RM247-RM463	24	12	6.640	-7.049	-18.096	1.967
21		12	RM463-RM235	12	8.9	10.453	-7.639	-17.807	1.706
22		12	RM17-RM19	36	36	2.814	-7.961	11.241	2.762
23	Spikelet number	3	RM5864-RM426	12	8.18	2.511	28.505	0.462	19.036
24		3	RM570-RM85	0	15	2.627	20.565	-21.140	8.961
25		10	RM216-RM171	20	2.43	2.719	-19.348	-8.364	6.858
26	Spikelet fertility	1	RM8278-RM582	32	40	3.022	666.000	2.399	0.000
27		3	RM426-RM168	0	9.54	3.566	-5.948	-6.707	4.013
28		7	RM418-RM11	8	6.4	2.784	-11.811	0.971	14.609
29		9	RM107-RM566	36	36	4.177	1.861	16.824	0.412
30		12	RM247-RM463	24	12	4.359	4.290	9.621	2.007
31		12	RM463-RM235	12	8.9	6.736	6.629	2.472	3.540
32		12	RM17-RM19	34	38	4.725	6.632	-9.461	5.305
33	Single plant yield	2	RM263-RM525	0	21.19	2.795	-1.068	3.273	1.090
34		9	RM107-RM566	34	38	2.848	3.922	-3.890	17.646
35	Grain length	3	RM251-RM282	6	5.59	3.057	0.200	0.069	13.904
36		3	RM16-RM353	12	10.16	6.303	0.335	-0.239	41.275
37		3	RM353-JL14	8	2.59	8.724	0.407	-0.169	56.033
38		5	RM430-RM18600	6	4.17	4.944	0.206	0.034	16.131
39	Grain breadth	1	RM473A-RM8278	2	25.37	5.772	-0.033	0.126	1.361
40		5	RM289-RM430	4	12.28	4.280	-0.122	0.074	21.665
41		5	RM430-RM18600	6	4.17	4.104	-0.115	0.055	22.488
42		8	RM502-RM310	12	43.03	2.594	-0.126	0.009	28.904
43		9	RM107-RM566	70	2	2.831	0.014	0.074	0.357
44	Length-Breadth ratio	1	RM473A-RM8278	0	27.37	6.068	0.069	-0.221	1.704
45		3	RM251-RM282	6	5.59	2.739	0.196	0.000	17.805
46		3	RM282-RM16	16	16.96	4.172	0.306	-0.175	50.119
47		3	RM16-RM353	12	10.16	4.943	0.229	-0.230	25.882
48		3	RM353-JL14	6	4.59	5.363	0.282	-0.162	34.792
49		3	JL14-RM5864	4	7.01	5.806	0.295	-0.147	38.972
50		3	RM5864-RM426	2	18.18	2.891	0.210	-0.107	19.154
51		5	RM430-RM18600	6	4.17	6.296	0.253	-0.066	30.951

S.No.	Trait	Chr.	Marker interval	LFM	RFM	LOD	Additive	Dominance	PVE
52	Grain length after cooking	3	RM282-RM16	20	12.96	2.593	0.823	-0.877	19.463
53		12	RM247-RM463	0	36	2.689	0.254	0.394	1.804
54	Elongation ratio	5	RM289-RM430	10	6.28	3.462	-0.148	0.103	20.651
55		5	RM430-RM18600	4	6.17	3.135	-0.128	0.075	17.130
56		12	RM247-RM463	0	36	2.633	0.030	0.071	1.022
57	Alkali spreading value	1	RM8278-RM582	34	38	3.218	-0.145	-1.593	0.949
58		4	RM551-RM564	20	11.25	3.531	-0.998	-0.553	31.461
59		6	RM217-RM402	12	11.05	10.646	-1.032	-0.007	43.130
60		6	RM402-RM276	12	3.3	25.412	-1.305	0.261	72.908
61		6	RM276-RM527	4	7.02	26.133	-1.254	0.254	70.732
62		6	RM527-RM539	4	5.44	19.137	-1.159	0.141	57.308
63		6	RM539-RM3	4	11.1	15.521	-1.299	0.365	59.567
64		12	RM17-RM19	44	28	3.926	0.260	0.184	3.118
65	Amylose content	4	RM280-RM127	2	8.13	3.339	-0.943	0.284	14.562
66		4	RM127-RM241	16	15.13	4.412	-3.376	0.335	73.199
67	Aroma	1	RM8278-RM582	40	32	9.228	0.831	5.766	2.879
68		2	RM138-RM475	50	21.99	9.439	0.998	-5.709	3.897
69		8	RM502-RM310	26	29.03	8.677	-0.935	-5.708	3.345
70		8	RM152-RM42	14	17.96	6.348	1.698	-5.774	9.068
71		8	RM339-RM44	4	12.97	3.661	2.493	0.367	19.168
72		8	RM72-Fgr	12	5.86	4.121	2.580	0.705	20.419
73		8	Fgr-RM404	10	4.33	6.538	2.265	1.358	16.497
74		8	RM404-RM483	6	14	6.368	2.541	0.907	22.993
75		12	RM17-RM19	26	46	9.402	-1.306	-5.702	7.269
76	Chalkiness	2	RM138-RM475	32	39.99	6.719	0.005	-0.209	0.001
77		3	RM130-RM545	0	33.56	4.047	0.757	-0.525	17.767
78		4	RM564-RM348	10	22.65	3.740	1.818	0.024	66.076
79		4	RM348-RM280	18	2.9	5.037	1.072	-0.763	25.368
80		4	RM127-RM241	10	21.13	3.133	1.187	-0.498	32.071
81		5	RM289-RM430	0	16.28	2.578	-0.745	0.435	10.992
82		8	Fgr-RM404	4	10.33	3.008	-0.226	-0.653	1.033

PVE- Phenotypic variance explained by each QTL; LFM- Left flanking marker; RFM- Right flanking marker

Sl Table 1.3 QTLs identified in Basmati370/Jaya F_2 population employing composite interval mapping (CIM)

S.No.	Trait	Chr.	Marker interval	LFM	RFM	LOD	Additive	Dominance	PVE
1	Plant height	1	RM302-RM11968	16	10.4	5.138	7.908	-0.858	15.418
2	Panicle length	2	RM6318-RM263	16	9.28	3.039	0.456	1.636	0.925
3		2	RM263-RM525	10	15.55	2.818	0.267	1.920	0.296
4		6	RM276-RM527	2	10.22	3.413	0.408	-1.773	0.819
5	Filled grains	1	RM11968-RM14	10	19.55	3.244	31.165	-28.073	22.677
6		10	RM258-RM496	0	29.91	2.521	11.257	-19.112	3.501
7	Chaffy grains	3	RM85-RM565	20	30.2	4.284	-3.532	-13.439	0.460
8		9	RM464-RM328	10	19.74	3.219	-5.241	-3.966	0.792
9		9	RM107-RM566	34	80	3.021	-2.710	-10.480	0.328
10		11	RM206-RM4B	22	20	2.721	666.000	15.575	0.000
11		12	RM247-RM463	34	15.23	5.211	-7.804	-15.738	2.458
12		12	RM463-RM235	18	7.15	10.403	-9.381	-8.383	2.335
13	Spikelet number	1	RM1-RM490	14	4.91	2.709	2.945	-20.888	0.214
14		1	RM490-RM259	6	5.59	2.813	7.033	-22.117	1.001
15		1	RM11968-RM14	18	11.55	3.568	45.936	-27.808	37.881
16		1	RM8278-RM582	70	44	3.369	-10.655	-47.484	3.328
17		2	RM138-RM475	52	60.06	3.087	-9.778	45.190	2.806
18		3	RM5864-RM426	14	10.16	2.788	666.000	-1.593	0.000
19		9	RM107-RM566	58	56	3.766	-12.298	46.761	4.434
20		10	RM216-RM171	26	1.29	2.885	-19.354	-6.115	6.661
21		12	RM17-RM19	46	68	3.962	666.000	47.732	0.000
23	Spikelet fertility	3	RM5864-RM426	24	0.16	2.981	-4.261	-6.624	1.968
24		3	RM426-RM168	0	10.44	2.989	-4.211	-6.596	1.922
25		3	RM85-RM565	20	30	2.531	2.958	5.927	0.891
26		9	RM107-RM566	56	58	2.562	4.208	5.350	2.202
27		12	RM247-RM463	34	15.23	4.062	4.390	9.115	2.150
28		12	RM463-RM235	14	11.15	7.255	7.155	1.973	4.249
29		12	RM17-RM19	48	66	3.441	5.987	-4.491	4.472
30	1000 Seed weight	7	RM234-RM248	20	2	3.530	-0.184	0.044	13.730
31	Single plant yield	2	RM263-RM525	0	25.55	3.720	-2.258	3.979	4.060
32		9	RM107-RM566	48	66	3.154	8.397	-4.769	82.155
33	Grain length	3	RM16-RM353	6	20.91	3.663	0.179	-0.215	10.901
34		3	RM353-JL14	10	1.7	9.217	0.362	-0.125	46.065
35		5	RM430-RM18600	6	5.2	6.603	0.217	0.031	17.468
36		9	OSR28-RM107	4	12.09	2.542	-0.098	0.028	3.441
37	Grain breadth	1	RM473A-RM8278	0	34.52	6.714	-0.038	0.119	1.649
38		3	JL14-RM5864	10	2.21	2.921	-0.098	0.068	14.157
39		3	RM55-RM448	6	15.6	3.122	0.049	0.038	4.220
40		5	RM289-RM430	2	16.88	3.345	-0.096	0.061	13.078
41		5	RM430-RM18600	4	7.2	3.333	-0.106	0.052	17.149
42		8	RM502-RM310	16	64.66	3.454	666.000	0.015	0.000
43	Length-Breadth ratio	1	RM306-RM128	2	31.27	2.560	0.026	-0.147	0.390
44		1	RM473A-RM8278	0	34.52	5.063	0.116	-0.208	3.928
45		3	RM282-RM16	20	23.13	3.323	0.433	-0.194	74.353
46		3	RM16-RM353	8	18.91	3.491	0.169	-0.211	12.326
47		3	RM353-JL14	8	3.7	4.358	0.220	-0.129	22.342
48		3	JL14-RM5864	2	10.21	4.770	0.229	-0.111	24.958
49		5	RM289-RM430	10	8.88	4.124	0.361	-0.129	41.529
50		5	RM430-RM18600	8	3.2	4.650	0.405	-0.070	46.531

Contd..

S.No.	Trait	Chr.	Marker interval	LFM	RFM	LOD	Additive	Dominance	PVE
51		11	RM206-RM4B	24	18	2.870	0.443	-0.046	69.477
52	Grain length after cooking	2	RM318-RM240	8	5.95	2.725	0.138	0.631	0.473
53		6	RM539-RM3	0	17.34	2.981	-0.241	0.664	1.401
54		12	RM247-RM463	0	49.23	3.512	0.312	0.396	2.680
55	Elongation ratio	5	RM289-RM430	14	4.88	3.706	-0.143	0.075	19.931
56		5	RM430-RM18600	4	7.2	3.711	-0.136	0.067	18.931
57		6	RM539-RM3	0	17.34	3.292	-0.035	0.111	1.198
58	Alkali spreading value	6	RM276-RM527	4	8.22	26.746	-1.257	0.264	71.735
59	Amylose content	4	RM280-RM127	0	11.15	4.077	-0.970	0.315	14.249
60		6	RM469-MX4	0	13.23	3.743	-0.886	0.265	14.989
61	Aroma	1	RM8278-RM582	74	40	6.735	0.654	5.284	1.859
62		2	RM138-RM475	80	32.06	7.590	-0.178	-5.332	0.133
63		8	RM502-RM310	36	44.66	6.976	-0.230	-5.309	0.218
64		8	RM152-RM42	18	23.56	6.132	0.968	-5.312	3.116
65		8	Fgr-RM404	10	6.36	4.910	1.946	1.161	10.990
66		8	RM404-RM483	8	16	4.998	2.476	0.511	20.226
67		9	RM107-RM566	84	30	5.969	0.038	-1.029	0.006
68		12	RM17-RM19	30	84	7.556	-0.589	-5.334	1.512
69	Chalkiness	1	RM8278-RM582	48	66	2.968	1.582	0.310	69.727
70		2	RM138-RM475	56	56.06	2.696	1.587	-0.468	69.576
71		3	RM130-RM545	0	44.07	3.234	0.915	-0.575	15.083
72		4	RM551-RM564	24	16.45	3.497	2.264	-0.606	53.905
73		4	RM564-RM348	14	28.62	3.138	2.107	-0.142	63.795
74		4	RM348-RM280	22	3.15	5.896	1.087	-0.754	25.092
75		5	RM289-RM430	6	12.88	3.835	-0.809	0.359	14.533
76		11	RM1812-RM202	22	16.54	2.747	1.919	-0.981	59.057
77		12	RM17-RM19	58	56	2.613	1.591	-0.418	69.014

PVE- Phenotypic variance explained by each QTL; LFM- Left flanking marker; RFM- Right flanking marker

Chapter II

DNA based detection and quantification of adulteration of Basmati rice

Introduction

2.1 Introduction

Basmati rice commands a special place among all aromatic rice cultivars on account of extra long slender grain, soft and fluffy texture of the cooked rice and pleasant distinct aroma. Traditional Basmati (TB) rice cultivars are not only in great demand in the domestic markets, but also seen in the menu of connoisseurs worldwide. However, unsuitability of TB for intensive cultivation prompted the breeders to develop evolved Basmati (EB) varieties (that fall short of the quality traits of TB), leading to the presence of both TB and EB varieties in the market in addition to a relatively inferior non-aromatic non-Basmati long-grain rice varieties (Archak *et al* 2007; Nagaraju *et al* 2002a; Vemireddy *et al* 2007). TB label not only brings high premium but also qualifies for duty reduction in some European markets (European Commission, Commission-Regulation: EC-No.2294/2003, 2003, pp. 12–15). Basmati370, Dehradun Basmati (Type-3), and Taraori Basmati of India, recognized as TB varieties, command higher price (U.S. $850/MT) as compared to EB varieties such as Pusa Basmati, Mahi Sugandha (U.S. $480/MT) from India and Super Basmati (U.S. $500–520/MT) from Pakistan (www.basmati.com, www.oryza.com), whereas NB varieties such as PR106 command only U.S. $167/MT. Because of difficulty in distinguishing TB, EB, and NB rice grains on the basis of visual observation or physicochemical tests, unscrupulous traders adulterate TB with EB or NB grains to earn higher profits (Figure 2.1). Survey conducted by the Food Standards Agency of the United Kingdom (www.food.gov.uk/science/ surveillance/ fsis2004branch/fsis4704 Basmati) of Basmati rice consignments from India and Pakistan report that such practices are not uncommon. By mixing 30% of EB rice with TB rice, for instance, traders can earn profit of almost $111/ MT, whereas similar degree of adulteration with NB varieties can earn a profit of $204/MT. Globally, annual Basmati rice export market alone is valued at US$1 billion and it is reported to be on the rise (Agricultural and Processed Food Products Export Development Authority (India), www.apeda.com; Trade Development Authority of Pakistan, www.epb.gov.pk) (Figure 2.2). When the subcontinent is in such an enviable growing export of this high quality rice but also the country's credibility in the international market. Therefore, distinguishing TB varieties from EB varieties and look-alike long-grain NB varieties and certifying to the genuineness of TB is an important task to protect the Basmati trade as well as interests of consumers. Although TB varieties can be distinguished from long grain Basmati-like rice varieties based on the intrinsic

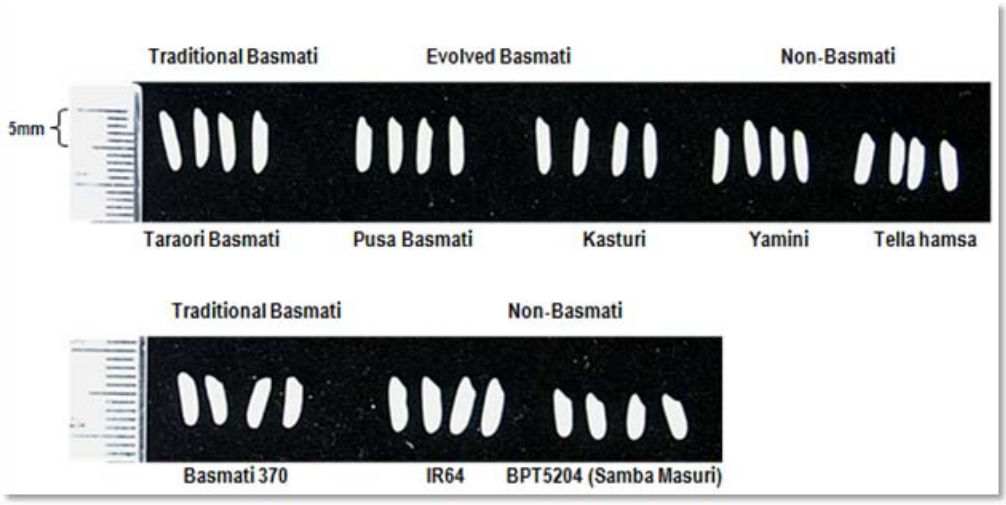


Figure 2.1 Comparison of grain shape among traditional and evolved Basmati and non-basmati varieties

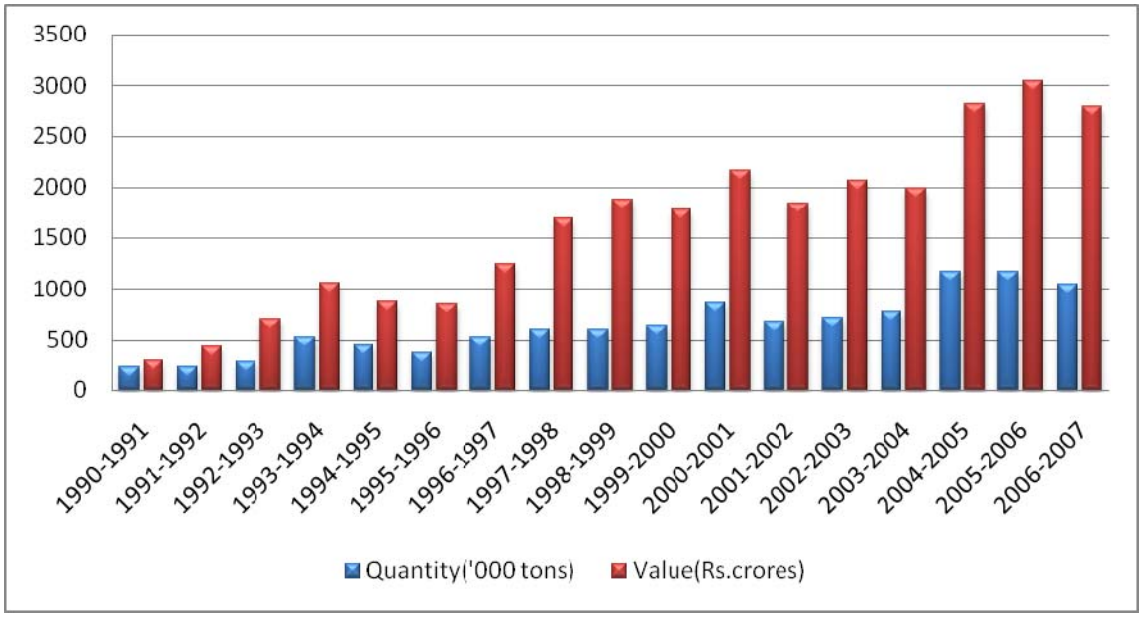


Figure 2.2 Export of Basmati rice from India

Data obtained from http;//commerce.nic.in

properties of the grain such as presence or absence of aroma, physico-chemical properties of starch, and grain elongation ratio on cooking, their reliability and applicability for routine and large scale applications are yet to be demonstrated (Figure 2.3 and 2.4). These drawbacks necessitate to look for DNA-based approaches.

A similar but more stringent traceability requirement in case of genetically modified (GM) food is being routinely and comprehensively met by use of real-time PCR and array-based protocols (Baeumler *et al* 2006; Leimanis *et al* 2006). In the case of GM food, detection and quantification of transgenes is relatively straightforward because they are distinct "foreign" DNA elements in the host genome. The foreign sequences, therefore, can be specifically amplified with high sensitivity (detection) and accuracy (quantification) using transgene-specific primers by real-time PCR. In case of Basmati,

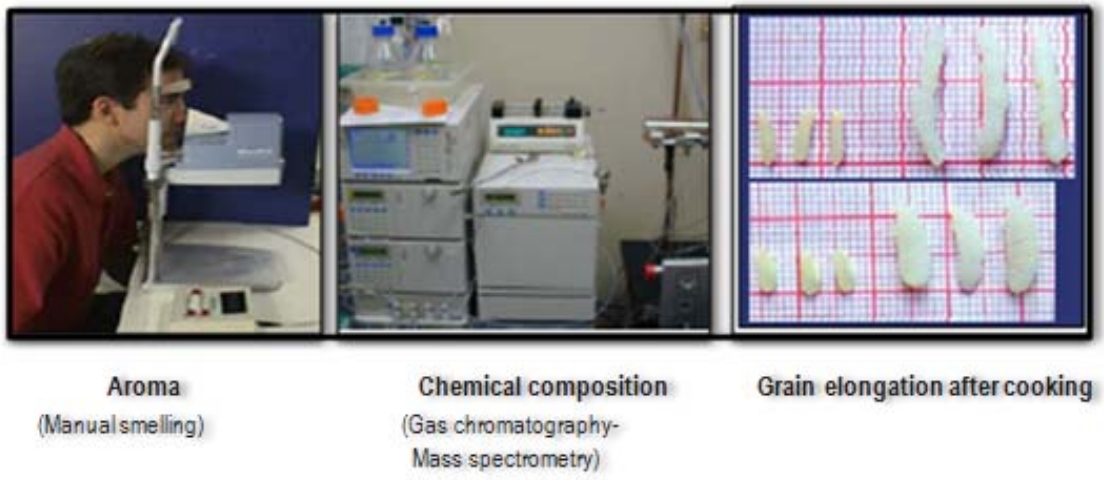


Figure 2.3 Methods for detection of the adulterant

the so-called "adulterant" is another long-grain rice variety with some of the traits of Basmati. Thus far, no specific sequence (genic or otherwise), upon amplification can generate private alleles specific to various Basmati cultivars, has been identified. The need for identifying Basmati-specific sequence/alleles has made researchers to look for polymorphisms using microsatellite markers.

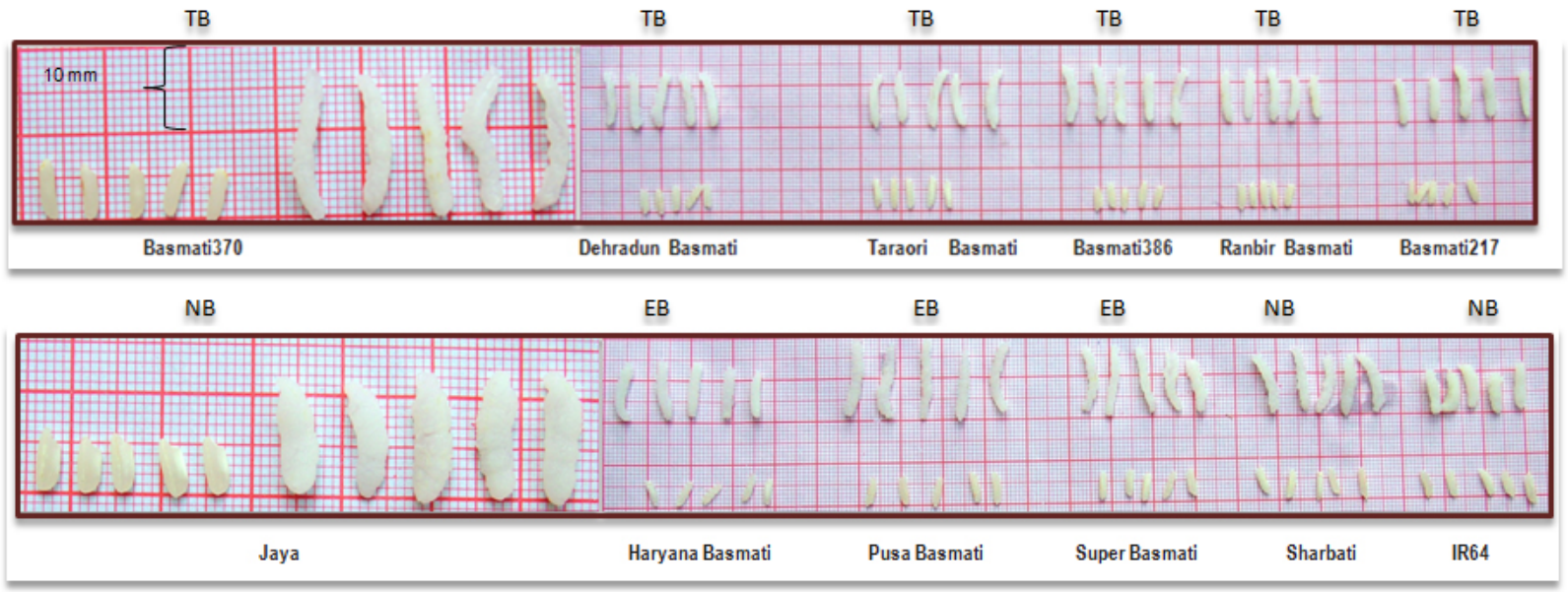


Figure 2.4 Comparison of grain length before and after cooking of traditional (TB) and evolved Basmati (EB) and non-basmati (NB) varieties

Microsatellite DNA markers have been extensively used in mapping and tagging genes, study of diversity, phylogeny and population genetics, disease diagnosis, and forensic investigations, (Nagaraju *et al* 2002a; Garris *et al* 2005; Jain *et al* 2004; Saini *et al* 2004). However, there is no instance of genotyping application akin to forensic DNA fingerprinting except for an initial report on the use microsatellites for detection of adulteration in Basmati rice (Bligh 2000), and subsequent demonstration of the potential of specific microsatellite profiles for unambiguous identification of TB rice varieties (Nagaraju *et al* 2002a). The multiplex panel of microsatellites has been employed for study of intra- and inter-subspecific variability in rice (Blair *et al* 2002; Coburn *et al* 2002). Recently, indel (sequences with an insertion or deletion) markers have been used for distinguishing Basmati varieties from other fragrant rice varieties (Steele *et al* 2008).

The initial use of microsatellite markers was based on agarose gel assays, and even today construction of linkage maps and diversity analysis in plants are usually based on agarose gel assays. However, forensic genetic analysis requires higher resolution fragment separation and greater accuracy in allele sizing as diagnostic markers often possess alleles with as few as 2 bp size difference. Thus, polyacrylamide gel electrophoresis was employed to achieve better fragment separation, and agarose gel electrophoresis was replaced by sophisticated and semi-automated PAGE platforms such as ABI373 and ABI377 (Applied Biosystems) and ALF express DNA sequencer (Pharmacia Biotech). Fluorescent dye chemistry (e.g., TAMARA, JOE, FAM, etc.) amenable to laser-based detection has enabled further improvement of resolution and accuracy of electrophoresis (Bandelj *et al* 2004; Nagaraju *et al* 2002b). A major upward shift in the accuracy of microsatellite marker based forensic analysis, however, took place with the introduction of capillary electrophoresis technology by various manufacturers, which includes Applied Biosystems (ABI310, ABI3100, ABI3700, ABI3730), Amersham Biosciences (MegaBASE500, 1000, 4000), Beckman Coulter (CEQ8800), and SpectruMedix Corporation (SCE2410, 9610, 19210) (Butler *et al* 2004). It was the capillary electrophoresis that enabled accurate and consistent allele sizing with minimal manual intervention, reducing thereby the sizing error (Koumi *et al* 2004). Furthermore, the higher throughput of capillary electrophoresis was found to be unmatched by any of the previous methods. Application of capillary-based microsatellite analysis was initially confined to human forensic DNA fingerprinting (Tagliaro and

Bortolotti 2006). The technique was subsequently extended to animal (Altet *et al* 2001) and plant (Tang *et al* 2004) variety identification and high-density genetic mapping projects.

Availability of different DNA detection and fragment separation methods have enabled Basmati importers and exporters to move toward a uniform and strict regime of DNA-based testing of export lots. A certificate of purity based on DNA test is considered to be value addition to a consignment of Basmati import into the European Union. Although there has been a consensus on employing a panel of microsatellite markers to solve the identity problem, the protocol accredited across laboratories to generate profiles has not been finalized as yet. Laboratories working on mapping and diversity studies of Basmati have been using different fragment separation techniques and have reported contradictory allele profiles of various Basmati varieties. For instance, Jain et al (2004) used fluorescent-labeled primers coupled with slab gel, whereas Saini et al (2004) and Singh et al (2004) used PAGE and agarose to study the genetic relationship of different Basmati cultivars. These published results have generated ambiguity among stakeholders in Basmati trade resulting in delay in establishing a "**Basmati Code of Practice"** describing diagnostic DNA markers for Basmati purity test. Differences in skill, cost, and infrastructure requirement among various techniques are so enormous that a simple decision is far from possible about the necessity and sufficiency criteria of fragment separation and detection techniques. The Centre for Basmati DNA Analysis in the Centre for DNA Fingerprinting and Diagnostics (CDFD) commissioned by the Government of India agencies, since its establishment is in the process of developing standards for fragment separation to be employed for purity testing of Basmati rice. Consumer preference brings higher returns for TB varieties, leading to generation of brand equity. Therefore, exporting firms and countries would be careful to maintain the purity of the consignment while importing countries would be keen on protecting consumer interests. Consequently, testing the samples for purity becomes important. Considering the trade volume, a high-throughput and accurate method for testing the purity of Basmati rice consignments is necessary. In fact, some Basmati importers particularly the European Union insist that all Basmati imports should carry mandatory certificate of purity based on DNA test (European Commission, Commission-Regulation: (EC)-No.2294/2003, 2003, pp. 12–15).

Keeping in view the foregoing developments relating to Basmati trade the present study was initiated with the following two specific objectives:

1. development of a high-throughput “single-tube assay” method based on multiplexing microsatellite markers for unambiguous identification of Basmati cultivars and detection and quantification of adulteration, and
2. demonstration of capillary electrophoresis as the accurate and reliable method for detection and quantification of adulteration of Basmati rice by comparing of various fragment separation techniques.

Materials and Methods

2.2 Materials and methods

Basmati rice varieties

Although the term "Basmati" actually denotes only a small number of aromatic rice cultivars of unique quality, vernacular names used to identify them in the areas of their cultivation have made the list longer. Recent breeding efforts to develop Basmati varieties in high yield background have also resulted in a number of evolved Basmati-like varieties. However, government agencies in India and Pakistan have recognized and notified 10 (5 TB and 5 EB) and 3 (1 TB and 2 EB) varieties, respectively as Basmati. Basmati370 and Super Basmati are recognized by the both the countries. In the present study, we included six notified TB varieties and nine officially released EB varieties (Figure 2.5), along with two long-grain non-aromatic varieties and a short-grain variety as listed in Table 2.1 to establish the utility of the protocol.

Sampling procedure and DNA extraction

Each rice sample consisted of at least 50 g of powdered grain, from which sub samples of 1 g each were drawn. Three sub samples were randomly selected and bulked, from which 100 mg of grain powder was collected for DNA extraction. DNA was isolated using Qiagen DNeasy plant mini kit (Japan) and DNA concentration was fluorometrically estimated with VersaFluor (Bio-Rad,USA) using Hoechst 33258 (Sigma,USA) dye with 300–490 excitation/460–320 emission filters. Average value of two readings was taken for each sample. Quality of the DNA (260/280 absorbance) was verified using UV1 UV-Vis spectrophotometer (Thermo Fisher Scientific Inc.,USA).

Microsatellite markers

Three hundred and fifty microsatellite loci (sequence source: www.gramene.org) were screened using six TB (Basmati370, Dehradun Basmati, Taraori Basmati, Basmati386, Ranbir Basmati, Basmati217), three EB (Pusa Basmati, Super Basmati, Haryana Basmati), two NB long grain rice varieties (Sharbati, IR64), and a short grain variety, Jaya.

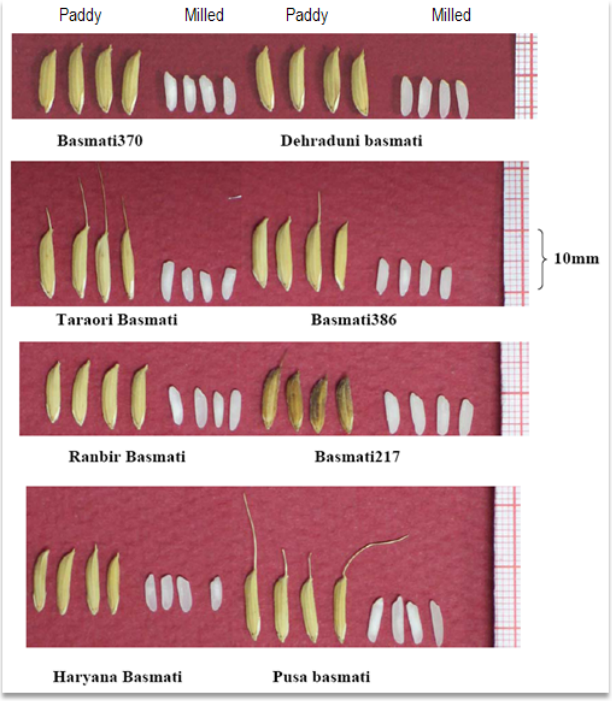


Figure 2.5 Traditional (TB) and Evolved Basmati (EB) varieties in paddy and milled form

Table 2.1 Rice varieties employed in the study

Name of the variety	Source	Type	Pedigree	Traits
Basmati370	I, P	TB	Selection	Aromatic, long grain
Dehradun Basmati	I	TB	Selection (Type-3)	Aromatic, long grain
Ranbir Basmati	I	TB	Selection (IET11348)	Aromatic, long grain
Taraori Basmati	I	TB	Selection (Karnal Local, HBC-19)	Aromatic, long grain
Basmati386	I	TB	Selection	Aromatic, long grain
Basmati217	I	TB	Selection	Aromatic, long grain
Kernel Basmati	P	EB	CM7/Basmati370, 1968	Aromatic, long grain
Basmati385	P	EB	TN1/Basmati370, 1985	Aromatic, long grain
Super Basmati	P, I	EB	Basmati320/IR661, 1996	Aromatic, long grain
Basmati198	P	EB	Basmati370/TN1, 1972	Aromatic, long grain
Pusa Basmati	I	EB	Pusa150/Karnal Local, 1989	Aromatic, long grain
Punjab Basmati	I	EB	Sona/Basmati370, 1982	Aromatic, long grain
Kasturi	I	EB	Basmati370/CR88-17-1-5, 1989	Aromatic, long grain
Mahi Sugandha	I	EB	BK79/Basmati370, 1995	Aromatic, long grain
Haryana Basmati	I	EB	Sona/Basmati370, 1991	Aromatic, long grain
Sharbati	–	NB	Selection from a UP landrace	Nonaromatic, long grain
IR64	–	NB	IR5657-33-2-1/IR2061-465-1-5-5, 1985	Nonaromatic, long grain
Jaya	–	NB	TN1/T141, 1967	Nonaromatic, medium grain

I, Basmati variety approved originally by India; P, Basmati variety approved originally by Pakistan.

PCR amplification

The PCR mixture (10 uL) contained 10 ng of DNA template, 80 uM dNTPs, 2 mM MgCl2, 0.5 U Amplitaq Gold DNA polymerase (Applied Biosystems,USA), and 0.1 uM each of forward and reverse primers. The 5'-end of forward primers (Sigma,USA) was labeled with any one of the following fluorescent fluorophores: carboxytetramethyl rhodamine (TAMRA), 6-carboxy-40-50-dichloro-20-70-dimethoxyfluorescein (JOE), or 50-carboxyfluorescein (50-FAM) (Table 2.2). Thermal cycling conditions for fluorescent labeled primers were: after an initial denaturation at 94^0C for 10 min, the PCR mix was cycled 25 times at 94, 55, and 72^0C for 45, 90, and 60 s, respectively (GeneAmp PCR System 9700, Applied Biosystems,USA). This was followed by a final extension step at 60 ^{0}C for 30 min. Thermal cycling conditions for non-fluorescent labeled primers were: initial denaturation of 15 min at 94 ^{0}C, 35 cycles of 45 s at 94 ^{0}C, 45 s at 55 ^{0}C, and 1 min at 72 ^{0}C, and final extension of 10 min at 72 ^{0}C. The PCR was performed on a GeneAmp PCR System 9700 (Applied Biosystems,USA). Amplification was verified on a 1.5% agarose gel.

Table 2.2 Microsatellite marker panel employed for multiplex assay

Locus	Repeat motif	Chr	Forward primer (5′ to 3′)	Reverse primer (5′ to 3′)	Allele pool (in base pairs)	Fluorophore
RM1	$(GA)_{26}$	1	GCGAAAACACAATGCAAAAA	GCGTTGGTTGGACCTGAC	73, 100, 106, 108	5′-FAM
RM72	$(TAT)_5C(ATT)_{15}$	8	CCGGCGATAAAACAATGAG	GCATCGGTCCTAACTAAGGG	148, 158, 164, 173	5′-FAM
RM171	$(GATG)_5$	10	AACGCGAGGACACGTACTTAC	ACGAGATACGTACGCCTTTG	322, 336, 344, 348	5′-FAM
RM202	$(GA)_{30}$	11	CAGATTGGAGATGAAGTCCTCC	CCAGCAAGCATGTCAATGTA	160, 182, 186	JOE
RM241	$(GA)_{31}$	4	GAGCCAAATAAGATCGCTGA	TGCAAGCAGCAGATTTAGTG	128, 140, 143	JOE
RM44	$(GA)_{16}$	8	ACGGGCAATCCGAACAACC	TCGGGAAAACCTACCCTACC	103, 109, 113	TAMRA
RM55	$(GA)_{17}$	3	CCGTCGCCGTAGTAGAGAAG	TCCCGGTTATTTTAAGGCG	220, 230, 235, 237	TAMRA
RM348	$(CAG)_7$	4	CCGCTACTAATAGCAGAGAG	GGAGCTTTGTTCTTGCGAAC	131, 140	TAMRA

Agarose gel electrophoresis

PCR samples were mixed with bromo-phenol blue and loaded on to a 3% agarose gel (Pharmacia model EPS-200, Sweden) containing ethidium bromide and run at 5.3 V/cm (Bio-Rad Power Pac 300) for an hour. Gels were photographed using a gel documentation system (Bio-Rad Molecular Imager Gel Doc XR System). Allele sizes were estimated in comparison with a 50 bp ladder (MBI Fermentas, Canada). Three researchers from the laboratory, who were carrying out PCR, agarose gel electrophoresis, and genotyping work on a routine basis, were asked to size the

alleles independently. Average allele sizes were determined on the basis of triplicate measurements obtained manually as well as by using software (Quantity One version 4.1.1). As it is difficult to quantify the amplified products manually in agarose gels, it was not carried out.

Slab gel electrophoresis

One micro liter of PCR product was mixed with 1.5 µL of 6x loading buffer (1:4 mixture of loading buffer and formamide; Sigma,USA), and 0.3 µL of GeneScan-500 ROX-labeled molecular weight standard (Applied Biosystems,USA) was included in the loading samples. The samples were denatured at 95 ^{0}C for 3 min before loading onto an ABI 377 automated sequencer (Applied Biosystems,USA) and electrophoresed on 5% polyacrylamide gel (Long Ranger, FMC,USA) containing 7 M urea, in 1x TBE buffer (90 mM Tris borate, pH 8.3, and 2 mM EDTA). Allele sizes were estimated using GeneScan 3.1 and Genotyper 2.1 software (Applied Biosystems) following the manufacturer's instructions.

Capillary electrophoresis (CE)

The fluorescently labeled PCR products were mixed with 0.3 µL of GeneScan-500 ROX (carboxy-x-rhodamine) size standard (Applied Biosystems,USA) and 12 µL of Hi-Di Formamide (Applied Biosystems,USA), electrophoresed by capillary electrophoresis on an ABI PRISM 3100 genetic analyzer (Applied Biosystems,USA) according to the manufacturer's instructions. Subsequently, fluorescent DNA fragments were resolved using the GeneScan version 3.7, and allele size and peak-area of the true peaks were determined by Genotyper version 3.7 (Applied Biosystems,USA). Each experiment was replicated at least three times to verify the reproducibility of markers.

Sequencing

Bidirectional sequencing of PCR products was carried out on an ABI PRISM 3100 (Applied Biosystems,USA) sequencer according to the manufacturer's instructions. The sequencing was repeated three times to obtain accurate sequences of the repeat regions. Comparison of allele sequences among Basmati varieties was carried out for each of the loci by multiple sequence alignment using ClustalX 1.8 program (Thompson *et al* 1997) and manual editing using GeneDoc Version 2.6.002 (Nicholas *et al* 1997).

Calculation of accuracy and consistency in estimation of allele sizes

Allele sizes estimated on the basis of sequencing results were considered to be *actual allele sizes* and were compared with *observed allele sizes* or *determined allele sizes* (in case of agarose) or *called allele sizes* in case of agarose, slab gel, and capillary electrophoresis instruments. Although different versions of the software were used to estimate the allele sizes in slab gel and capillary electrophoresis, as mentioned above, hardly any noticeable size differences were found as the basic algorithm is common and had no influence on our analysis. Allele size differences were averaged over all 12 rice varieties to obtain accuracy at each locus. The accuracy of allele size estimation was expressed as the difference (in base pairs) between the actual allele size and the observed allele size. Consistency in allele size estimation was computed from standard deviations obtained in triplicated assays involving electrophoresis and size determination of 12 rice varieties at each of the eight microsatellite loci.

Calculation of accuracy in quantification of adulteration

In slab gel and capillary electrophoresis methods, the quantity of an amplified PCR product is represented by peak area (measured in relative fluorescent units, rfu). Accuracy in quantification of adulteration was calculated by mixing of Basmati370 grains with a common adulterant, Haryana Basmati, in progressive proportions from 10 to 90% with a 10% interval at two informative loci, RM72 and RM348. Adulteration was expressed as a percent fraction:

$$\text{adulteration} = \left[\frac{A}{(A+B)}\right] \times 100$$

Where,

A is the rfu of the adulterant and

B is the rfu of the main variety.

Accuracy of quantification was expressed as the difference (in percent adulterant) between the actual adulteration and the observed adulteration.

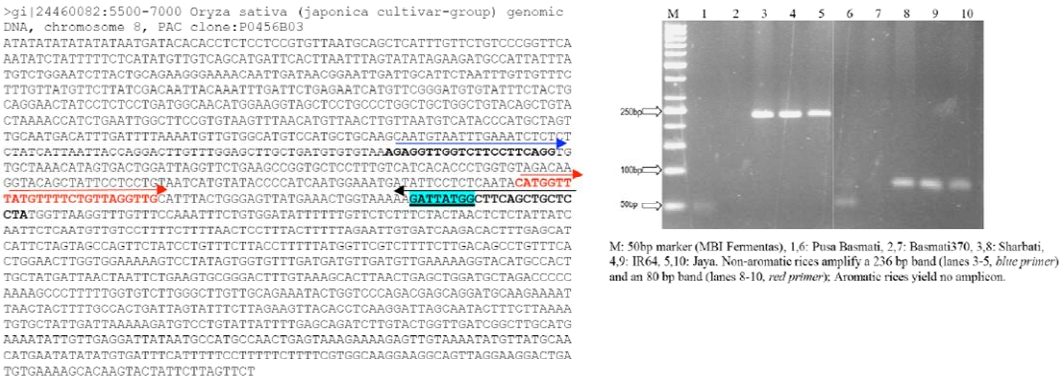
Preparation of standard samples

To verify the sensitivity and accuracy levels of the methodology in detection and quantification of adulteration, standard samples of Basmati370 were prepared by mixing the grains of a

common adulterant, Sharbati, at progressive ratio of 1, 3, 5, 7, 10, 15, 30, 50, and 60% to generate data at nine score points. Subsequent to genotyping, peak heights (relative fluorescence units, rfu) were recorded at each score point for evaluation. The peak areas (rfu) were determined and were plotted against the percent adulterant to develop a standard curve based on logistic model ($y = a/1\ 1\ be2cx$) by using CurveExpert version 1.37 (http://curveexpert.webhop.biz).

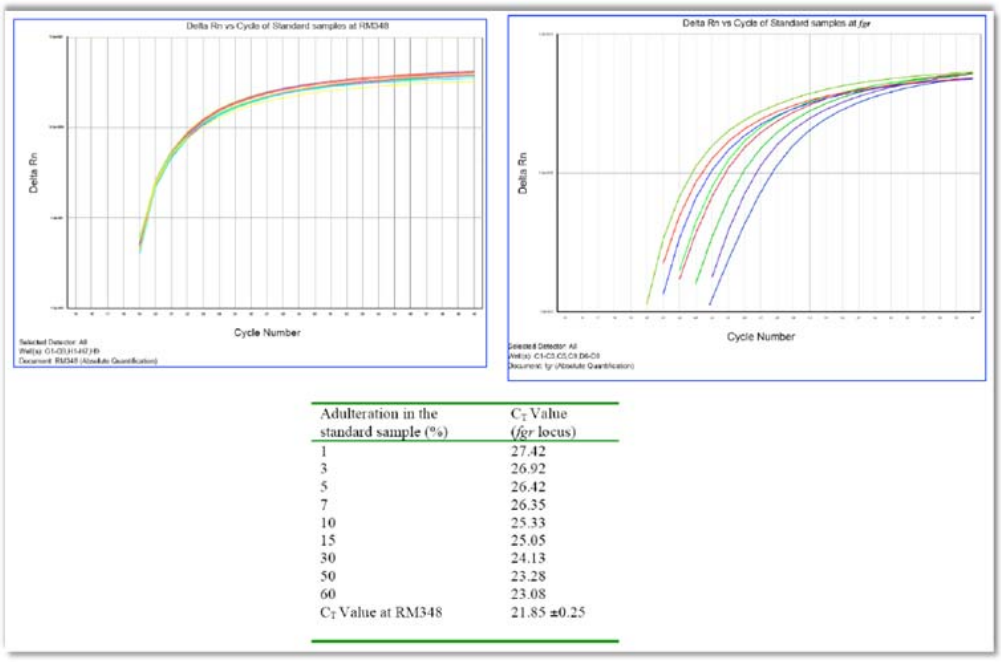
Comparison of capillary electrophoresis with real time PCR

Real time PCR is acknowledged as an accurate protocol for DNA quantification and is routinely employed in GM detection assays. An 8 bp deletion in the betaine-aldehyde dehydrogenase-2 gene *(bad-2)* specific to aromatic rices (Bradbury *et al* 2005) was exploited to amplify an 80 bp product exclusively in a non-Basmati variety (i.e., adulterant) in a mixture (Figure 2.6). Real time PCR and capillary electrophoresis (at two loci, RM348 and RM72) were carried out on standard samples prepared by mixing grains of Basmati370 with Sharbati in different proportions of 1, 5, 7, 10, 15, 30, 50 and 60% (Figure 2.7). PCR amplification and capillary electrophoresis were performed as described earlier. Real time PCR was carried out by employing qPCR Mastermix Plus for SYBR Green I kit (Eurogentec). A 50 µL reaction contained 2x reaction buffer containing dNTP, HotGoldStar, MgCl2, 0.5 µM primers (forward CATGGTTTATGTTTTCTGTTAGGTTG and reverse TAGGAG-CAGCTGAAGCCATAAT), SYBR Green I, stabilizers, and ROX passive reference. The PCR conditions included an initial DNA denaturation step at 95 ^{0}C for 10 min followed by 40 cycles of PCR with DNA denaturation at 95 ^{0}C for 15 s, primer annealing at 60 ^{0}C for 1 min, and extension at 72 ^{0}C for 1 min. Relative fluorescence units at various points of the amplification cycle were recorded with an ABI Prism 7000 sequence detection system (Applied Biosystems, USA) and Ct values were obtained using Sequence Detection Software 1.2.3. Ct values (reflecting the changed adulteration content) were determined from averaged Ct values of the replicate runs. Template DNA (expressed as percent quantity of adulterant in the sample) was computed by taking 7% score-point as the reference. These estimates were then compared to the adulteration quantified by the capillary electrophoresis (averaged over two loci, RM348 and RM72). Optimization of real-time PCR was carried out by comparing amplifications at *fgr* and RM348 loci (Figure 2.7).



M: 50bp marker (MBI Fermentas), 1,6: Pusa Basmati, 2,7: Basmati370, 3,8: Sharbati, 4,9: IR64, 5,10: Jaya. Non-aromatic rices amplify a 236 bp band (lanes 3-5, *blue primer*) and an 80 bp band (lanes 8-10, *red primer*); Aromatic rices yield no amplicon.

Figure 2.6 Real-time PCR amplification of *(fgr)* fragrance gene. Amplicon obtained from *blue* forward primer gives a 236 bp amplicon and *red* forward primer amplifies an 80 bp band. Reverse primer is common, and binds in the region where deleted sequence (GATTATGG) in the aromatic rices thus, amplifying only in non-aromatic rices.



Adulteration in the standard sample (%)	C_T Value (*fgr* locus)
1	27.42
3	26.92
5	26.42
7	26.35
10	25.33
15	25.05
30	24.13
50	23.28
60	23.08
C_T Value at RM348	21.85 ±0.25

Figure 2.7 Representative real-time PCR runs of RM348 and *fgr* loci

Results and Discussion

2.3 Results and Discussion

High-throughput multiplex microsatellite marker assay for detection and quantification of adulteration of Basmati rice

Detection of adulteration

Microsatellite marker panel for adulteration detection

Earlier reports on DNA-based detection of adulteration in food/feed material had employed specific DNA sequences to amplify fragment length polymorphisms to reveal the existence of an adulterant. The specific loci included 12S rRNA (Bellagamba *et al* 2003), 5S rRNA (Dhiman and Singh 2003), trnT-trnL spacer (Mooney *et al* 2006), and chloroplastic-SNP (Spaniolas *et al* 2006). However, in case of Basmati rice, genuine TB samples are adulterated with either EB or NB that are quite similar to TB and hence offer no leeway to employ routinely used genomic sequences for differentiating them (Figure 2.1). Under such circumstances, microsatellite loci that exhibit greater allelic polymorphism can be employed for the detection of adulterants. Primary step in the process of detection of an adulterant in Basmati samples is generation of variety-specific microsatellite profiles. This was achieved by screening for polymorphic microsatellite loci so that upon PCR, a genuine TB variety, at a given microsatellite locus yields a single allele of known size. Any admixture of TB with either EB or NB would be detected readily because of distinct allele sizes (Figure 2.8). Eight out of 350 microsatellite loci consistently amplified single allele per variety and distinguished, in combination, all the Basmati varieties notified by India and Pakistan among themselves and from some putative adulterants (Table 2.2). The panel comprised of microsatellite loci with five di, one tri, one tetra nucleotide repeats and one compound repeat. The panel generated an average of 3.3 alleles per locus in a size range of 73–348 bp. Success of a microsatellite-based genotyping application depends upon the informativeness of the loci. This means that although some of the microsatellite loci (dinucleotide repeats) exhibited great discriminatory power (*e.g.*, RM252, 22 alleles and allele frequency 11.6%, (Jain *et al* 2004)), they produced a series of nonspecific amplification products even under the best of conditions, precluding their inclusion in the genotyping assay. To avoid erroneous allele sizing, which is not acceptable for sensitive assays such as genotyping and determination of adulterants, the allele sizes in 12 varieties were confirmed by sequencing the alleles and counting the number of repeat units in each allele at all the loci (Figure 2.9)

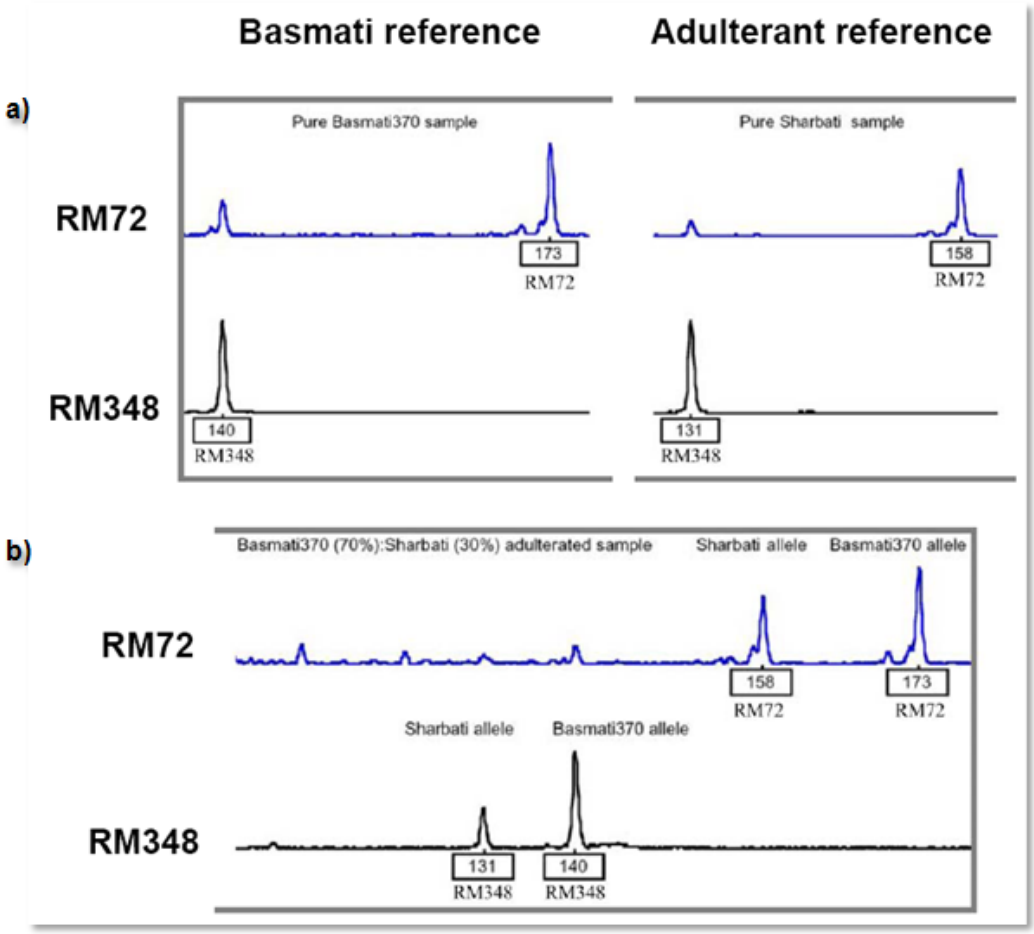


Figure 2.8 Detection of the adulterant. a) Electropherogram of amplification of pure Basmati (Basmati370) and pure adulterant (Sharbati) using RM72 and RM348 b) Electropherogram of amplification of adulterated Basmati (Basmati370 adulterated with 30% Sharbati) using RM72 and RM348

(SI Figure 2.2). Further, diversity analysis of Basmati varieties (McCouch *et al* 2002) showed that the microsatellites included in our panel such as RM171 (34.8%), RM55 (55%), RM44 (39%), and RM1 (50%) exhibited moderate allele frequency. This, as evident from the present study, brings adequate polymorphism to differentiate between TB, EB, and NB varieties. In the wake of the threat to the sustenance of Basmati trade by adulteration with any or a combination of several long grain rice varieties, the methodology reported here would help the trade by unambiguously identifying the authentic Basmati varieties and distinguishing them from trace adulterants based on deviations from the designated allele profile (Figure 2.10).

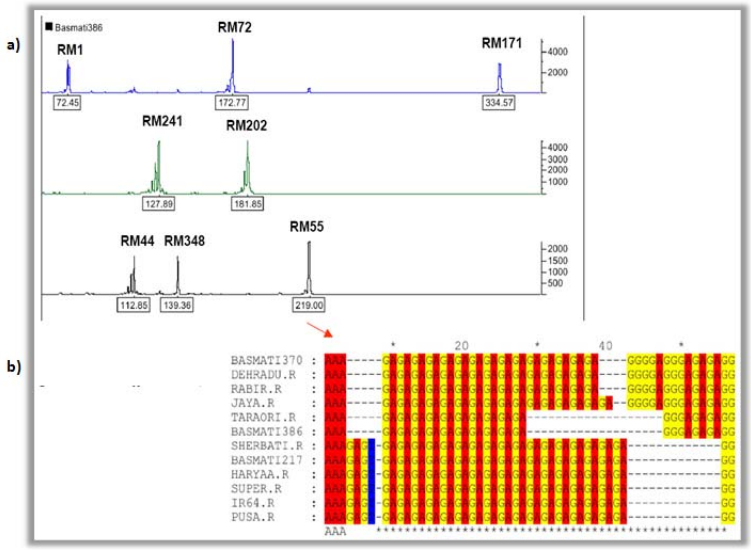


Figure 2.9 a) Genotyping of Basmati386 using 8 multiplexed microsatellite loci
b) Sequence alignment of alleles of RM55 showing variation in microsatellite repeat length

A. Identity codes

Variety	RM171	RM55	RM202	RM72	RM348	RM241	RM44	RM1
	Allele code							
Basmati370	B	C	B	D	B	B	B	A
Dehradun Basmati	B	C	B	D	B	B	B	A
Ranbir Basmati	B	C	B	D	B	C	B	A
Taraori Basmati	B	A	B	D	B	A	C	A
Basmati386	B	A	B	D	B	A	C	A
Basmati217	C	B	B	B	A	A	A	B
Kernel Basmati	C	A	B	D	B	B	B	A
Basmati385	B	C	A	D	B	B	C	A
Super Basmati	C	A	B	D	B	B	B	C
Basmati198	B	C	A	D	B	B	B	C
Pusa Basmati	C	B	B	B	A	A	C	A
Punjab Basmati	C	A	A	A	B	B	A	A
Kasturi	C	B	B	B	A	A	A	A
Mahi Sugandha	C	B	A	A	B	B	A	B
Haryana Basmati	C	B	A	B	A	A	A	D
Sharbati	A	B	A	B	A	A	A	C
IR64	D	B	C	C	A	A	A	C
Jaya	A	D	A	B	A	A	A	C

B. Code key

Allele code	RM171	RM55	RM202	RM72	RM348	RM241	RM44	RM1
	Allele size in base pairs							
A	322	220	160	148	131	128	103	73
B	336	230	182	158	140	140	109	100
C	344	235	186	164		143	113	106
D	348	237		173				108

Figure 2.10 A. Specific allele profiles of Basmati rice varieties and putative adulterants at eight microsatellite loci.
B. Based on the allele size of each locus code was assigned.

Stability of marker profile

Rice is highly homozygous due to absolute autogamy. If at all any variation is observed in rice, it is expected between populations. Unlike NB varieties that are cultivated as a staple crop, Basmati rice, grown as a commercial crop, is being maintained with little variability across populations Further, SSR- and FISSR-based analyses showing no significant difference among the notified TB varieties support the view that TB varieties are possibly the descendants/derived/selected from a single land race source, and minor differences seen among them are result of human selection *vis a vis* preference of farmers and consumers over centuries (Nagaraju *et al* 2002). Varietal homogeneity in TB was further evident from an earlier study that demonstrated absence of within variety polymorphism in some of the Basmati varieties (Singh *et al* 2004). In the present study, Basmati varieties were sampled from multiple sources and specificity of the designated alleles in a given Basmati variety was confirmed by genotyping at least ten individual plants from each of the varieties. There was no intravarietal polymorphism among populations (SI Table 2.1). Varietal homozygosity and lack of interpopulation variation among Basmati varieties coupled with moderate allele frequency of the selected microsatellite loci ensure stability of the genotyping assay described in this report.

Multiplex assay and variety specific profiles

Multiplexing all the informative microsatellite loci into a single-tube assay facilitates high throughput and reduces the cost of assay (Figure 2.9). Eight microsatellite loci of the panel were multiplexed into three groups after confirming that their allele sizes did not overlap (Table 2.2) and that all the primers showed compatibility when cross-checked against one another (Table 2.3). The primers were multiplexed by using three different fluorophores viz., 50FAM (RM1, RM171, RM72), JOE (RM202, RM241), and TAMRA (RM55, RM44, RM348). Microsatellite allele profiles of the varieties are given in Figure 2.1 of Supporting Information and the identity codes of the varieties are presented in Figure 2.10. All the TB varieties except Basmati217 generated identical allele profiles at each of the five loci – RM1, RM72, RM171, RM202, and RM348, whereas EB and NB varieties showed a diverse allele pattern. The TB varieties were distinguished readily and indisputably from EB and NB varieties using a combination of the five loci. Three loci, RM55, RM241, and RM44, were useful in detecting differences among TB varieties. Taraori Basmati/Basmati386 and Basmati370/Dehradun Basmati pairs had the same allele profiles and could not be distinguished. Otherwise, the eight-letter code was different for

Table 2.3 Inter-locus peak balance values for the multiplex loci

Locus	RM1	RM72	RM44	RM55	RM202	RM241	RM171	RM348
RM1	-							
RM72	0.91							
RM44	6.2	7.05						
RM55	5.92	5.69	0.69					
RM202	1.14	1.55	0.09	0.25				
RM241	1.77	2.64	0.14	0.31	0.7			
RM171	1.11	1.18	0.19	0.2	0.52	0.5		
RM348	1.29	1.28	0.24	0.1	0.94	0.77	1.01	-

all the varieties analyzed in this study and specificity of the assay was beyond doubt (Figure 2.10). Most common adulterants such as Basmati385 and Sharbati as well as EB varieties were assigned unique codes of identity.

Validation of the methodology and test of accuracy

Validation of a newly developed protocol is an integral process before it is adopted by other laboratories or for commercial purpose. DNA was extracted from aged rice grains, since traded grains are usually aged by a year. Care was taken to extract pure DNA since inhibitory components could be co-extracted from rice grains that are particularly rich in starch and also contain small amount of lipids. Four procedures – CTAB method (Saghai-Maroof *et al* 1984), modified CTAB method (Kang *et al* 1998), Nucleon Phytopure DNA extraction kit (Amersham Biosciences), and Qiagen DNeasy plant mini kit – were tested for DNA extraction. The Qiagen kit was found to yield good quality DNA (260/280.1.7) and exhibited the least sample-to-sample variation in PCR amplification (Cankar *et al* 2006). Experiments were carried out (i) to determine the optimum primer concentrations and the minimum number of PCR cycles, (ii) to assess the LOD (limit of detection), and (iii) to confirm the accuracy of quantification with real-time PCR.

Primer concentration

Varying the concentration of primers of one locus may affect the amplification at the other seven loci in the multiplex. Differences in the DNA sequences of the loci can lead to variation in the efficiency of primer binding. Therefore, it becomes essential to tune the primer concentrations to obtain not only best signal intensities but also balanced peak heights in a multiplex reaction. Equal concentrations of the primers resulted in intense allelic products at

RM1 and RM348, leading to dwarfing of other peaks. Optimal primer concentrations for the multiplex were determined to be 0.1uM (RM171, RM55, RM44), 25 nM (RM202, RM72, RM241), and 6.25 nM (RM1, RM348).

Number of PCR cycles

Improvement in sensitivity of PCR assay can be achieved by increasing the number of cycles. Forensic geneticists routinely use increased PCR cycle numbers (38–43) to obtain amplification from ancient DNA (Gill *et al* 1994). With increased template DNA, however, effect of cycle number is not that apparent; instead it could lead to artifact production and compromise on peak balance. Therefore, increased number of PCR cycles would adversely affect quantification of the adulterant. This is effectively achieved only in the linear phase of the PCR, where differences in the template DNA concentrations are best represented (Bubner and Baldwin 2004). To optimize the cycle number keeping sensitivity as well as quantification of the adulterant in mind, template DNA was amplified at 20, 22, 24, 25, 26, 28, and 30 cycles. Allele dropout in 1–5% adulterated samples was observed at 20, 22, and 24 cycles frequently for the adulterant allele. Multiplex loci gave good signal intensities from 25 cycles onwards. Twenty five cycles was chosen as a minimum number of cycles for quantification of adulterant using multiplex assay.

Peak balance and quantification

Amplification of polymorphic alleles may be imbalanced due to stochastic effects in the PCR reaction. Additionally, one allele can be preferentially amplified over the other due to unequal sampling of polymorphic alleles during the early stages of the PCR reaction (Findlay *et al* 1997). Since an important requirement for accurate quantification is to produce balanced allele peaks, peak balance ratio was calculated at various ratios of adulteration. Only polymorphic loci were included in the calculations. Complete dropout of one allele was assigned a zero peak balance ratio. Good intralocus balance (≥0.7 ratio) was obtained at all the multiplexed loci except RM1 (Table 2.3) (SI Figure 2.1). RM72 and RM348 exhibited excellent peak balance (≥0.80), confirmed over two sets of adulterated standard samples, Basmati370: Sharbati and Basmati370:IR64 (the latter is given as Figure 2.3 of Supporting Information). Therefore, RM348, RM72, and RM171 were used for the quantification of the adulterant in that order of preference (Figure 2.11).

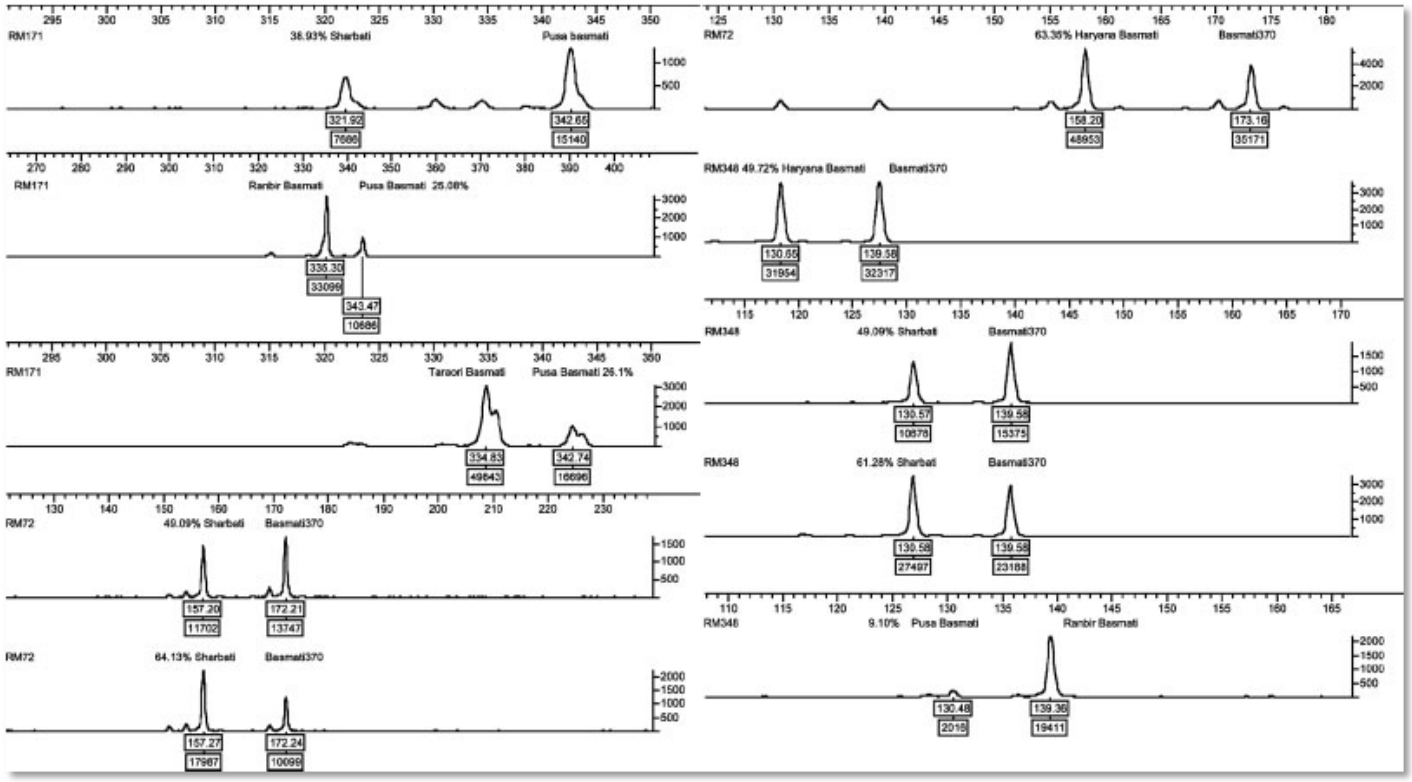


Figure 2.11 Quantification of the adulterant(s) using three microsatellites. Each electropherogram represents a quantification assay of various Basmati rice samples at a given locus (any of RM348, RM72, and RM171). Peaks are annotated by allele size (top box) and peak area (bottom box). Main variety and adulterant are resolved, identified and adulterant is quantified (expressed as per cent) peak area ratios. Alleles are identified within one base pair bin compared to the sizes given in Figure 2.9

Limit of detection (LOD) and Limit of quantification (LOQ)

LOD and LOQ were determined by an empirical approach consisting of measuring progressively more dilute concentrations of the adulterant. The LOQ (the adulteration levels at which quantitative results can be reported with a high degree of confidence) was also empirically determined. At 25 cycles PCR, all the eight loci were not equally efficient in detecting 1% adulteration level. Multiple runs show increasing SD, in that order, for RM348, RM72 (trinucleotide microsatellite repeat loci), RM171 (tetranucleotide microsatellite locus) and followed by dinucleotide repeat loci (RM202, RM55, RM1, RM44 and RM241) in detecting 1% adulterant in the otherwise pure Basmati sample. LOQ was also found to be 1% in our analyses particularly for RM348 and RM72. Number of PCR cycles influenced both LOD and accuracy of quantification. Detection efficiencies were drastically improved at 30 cycles for all the eight loci in terms of consistent detection of 1% adulteration, and reduction of allele dropouts. Conversely, the error limits of quantification expressed as SD, calculated over replicated experiments substantially increased from +1.11% (25 cycles) to ±3.3% (30 cycles) for adulteration up to 15%.

Concordance studies with real-time PCR analysis

Quantification based on competitive-quantitative PCR coupled with the use of standard curve, essentially relies upon the accuracy of measuring the template DNA concentration in the PCR mixtures. Hence, it was considered critical to confirm the template DNA concentrations by employing real-time PCR. However, the challenge was to find a specific sequence (genic or otherwise) that can, upon amplification, generate private alleles specific to Basmati varieties and adulterants. An amplicon of 80 bp size that is specific to nonaromatic rices was identified (Figure 2.6). Using this sequence, standard samples of Basmati370 containing Sharbati (nonaromatic) as adulterant were analyzed by real-time PCR to quantify the Sharbati DNA. *C*t values of standard samples were converted to the actual adulteration levels and the values were then compared to the adulteration quantified by the CE-based microsatellite assay. The values indicated reasonable agreement between quantification based on real-time and electrophoresis methods (Figure 2.12), implying that the standard samples were reliable enough to construct standard curves and subsequently employ those standard curves for quantification of adulteration.

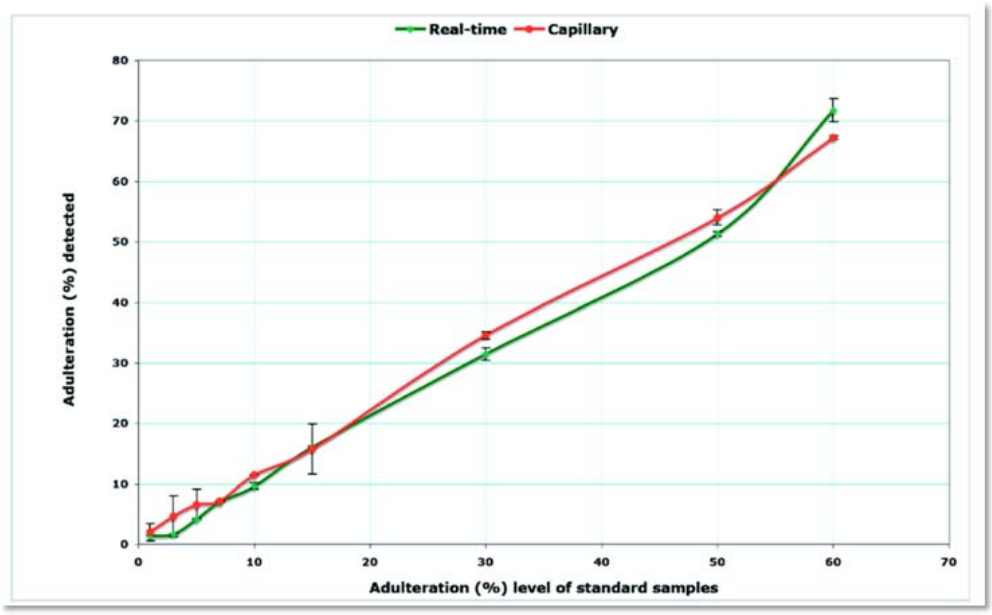


Figure 2.12 Corroboration of quantification of adulterant based on microsatellite marker assay by *fgr gene-based* real-time PCR assay

Quantification of adulteration

It is possible that some Basmati rice samples contain adventitious mixture because of inadvertent mixing in the field during harvest/storage. If we can measure the actual amount of the adulterant, such samples having admixture within limits allowed by the importing countries could be certified as practically genuine. For instance, Basmati Code of Practice recommends 7% as the ceiling for inadvertent mixtures in Basmati imports in the UK (www.riceassociation.org.uk**)**. Hence, it is essential that the methodology incorporate accurate quantification of adulteration. The protocol assumes (i) equivalent contribution of DNA from the grain mixtures and (ii) faithful representation of the samples. Therefore, adulterant is quantified based on the relative quantities of the amplified allelic products at a common locus between competing DNA templates (those of the Basmati and the adulterant) in the PCR mixture (Gilliland *et al* 1990) (Figure 2.13). In this way, adulteration was computed in the standard samples (premixed grains of the genuine Basmati and a putative adulterant in a progressive ratio). The results were represented by plotting percentages of adulterant on the abscissa and the corresponding peak-area ratios (expressed in percentage) on the ordinate. A smooth curve was obtained using logistic model (Figure 2.14).

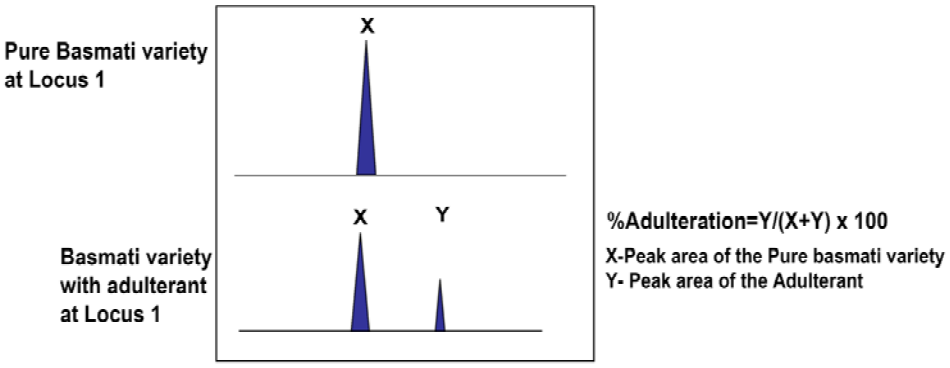


Figure 2.13 Quantification of the adulterant. The principle underlying the quantification of the adulterant is that ratio of quantities of the amplicons (ratio of peak area of pure variety and adulterant) is directly proportional to the quantities of competing DNA template in a PCR reaction.

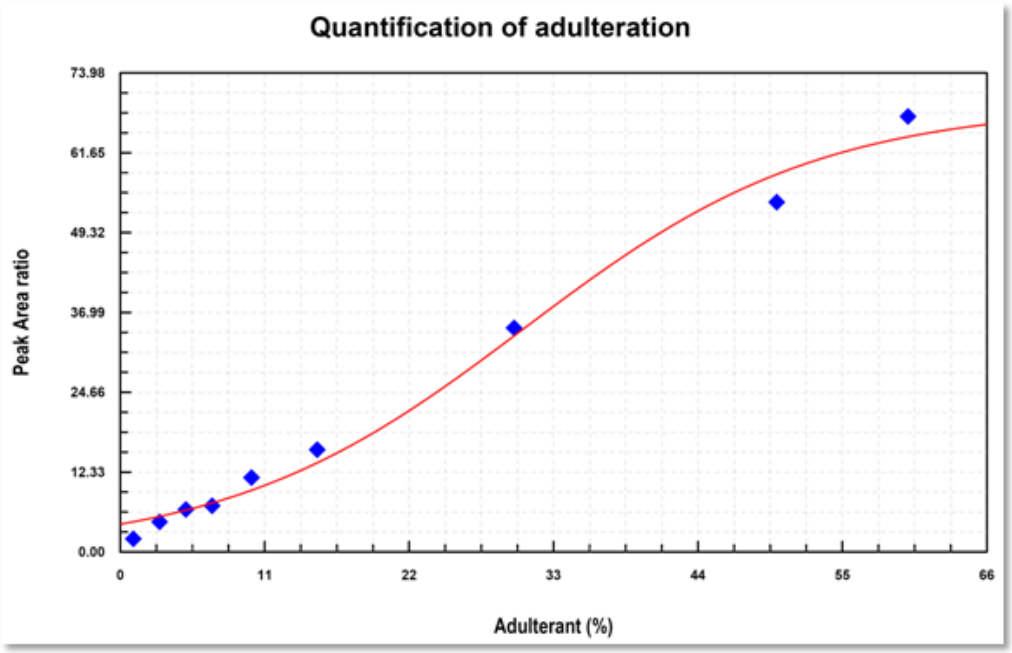


Figure 2.14 Standard curve generated based on Basmati370: Sharbati standard samples. Pure Basmati370 was mixed with Sharbati in progressive proportions (1, 3, 5, 7, 10, 15, 30, 50, and 60%). Peak area ratios obtained from microsatellite-based CE at RM348 locus were plotted against per cent adulteration at nine score points. The standard curve is developed based on logistic model.

Using both RM72 and RM348-based standard curves; adulterants were marginally overestimated at lower levels of adulteration (up to 15%) and were generally overestimated at higher levels of adulteration. The estimation of adulterant for adulteration up to 15% was within +1.11% of the actual adulteration. However, accuracy of estimation was lower for adulteration beyond 15%, which was +2.5% (Figure 2.14). Post DNA-test regime, legal ceiling of adventitious adulteration is proposed to be 7% for the lowest grade Basmati (www.food.gov.uk/multimedia/pdfs/ fsis4704Basmati.pdf). Therefore, samples having higher adulteration could be rejected, and hence accuracy of estimation beyond 15% of the adulteration becomes practically redundant. Detection of adulteration is relatively error-free once the representative samples are ensured. However, quantification (and subsequent acceptance/rejection of a sample) could be susceptible to minor experimental errors. When three blind samples with 4, 8 and 12% adulteration were genotyped and the peak-area ratios were plotted on the standard curves, the adulteration was estimated with an error of ±1.94, ±0.17 and ±0.93%, respectively. Possible sources of experimental error in the protocol were identified to be (i) DNA extraction, (ii) amplification of the microsatellite alleles by PCR, and (iii) CE. An experiment was laid down to quantify any possible deviation from actual adulteration value. Each of these steps was triplicated in series. For instance, 15% adulterated Basmati370 samples were assayed at RM72 indicated that deviation from the 15% was obtained with an overall error of ±1.23%. The same value at another locus RM348 was ±1.45%. This consistency level is adequate for all practical purposes of Basmati certification.

Limitations of the multiplex microsatellite marker assay

Taraori/Basmati386 and Basmati370/Dehradun Basmati pairs had the same allele profiles in the 8-loci multiplex and could not be distinguished. Nonetheless, this has no effect on the effectiveness of the protocol, as these varietal pairs are treated as practically the same variety in the market. Further, in our ongoing search for better loci, a few more potential candidate loci (RM440, RM525, JL14, RM16, and RM85) were found that are being developed as reserve set markers in case new adulterants in future challenge the existing multiplex. Among them, it was discovered that JL14 could distinguish between Taraori/Basmati386. The gap between real-time PCR in GM detection (LOQ 0.06%, www.food.gov.uk/multimedia/pdfs/fsis0506.pdf) and CE-based microsatellite marker assay in the estimation of Basmati adulteration (LOQ 1%, present study) in terms of error is significant and probably reflects the efficiency of the techniques

rather than the problem in hand or standardization of protocol. For instance, quantification of pork adulteration using SINE-specific sequences could be achieved at 1% limit (Calvo *et al* 2002) and a much-sophisticated approach (lab-on chip capillary-based electrophoresis) could also detect at a limit of 5% contamination on Arabica coffee samples (Spaniolas *et al* 2006). Opportunely, the stringency levels involved in detecting GM (legal limit of 0.9%) and detecting adulteration in case of Basmati trade (proposed legal limit of 7%) are as much disparate, and that allows us to employ the protocol presented here effectively for all practical purposes even if errors in the estimation of adulteration ranges at ±1.5%.

Capillary electrophoresis is essential for microsatellite marker based detection and quantification of adulteration of Basmati rice

There are three essential aspects for an electrophoresis technique to be successfully employed in genotyping assays: (i) accuracy (allele sizing), (ii) consistency (reproducibility), and (iii) sensitivity (quantification). Because the demand on accuracy in genotyping assays having legal (as in forensic DNA fingerprinting) or commercial (as in Basmati purity tests) connotations is much more stringent than required for typical electrophoretic applications, it is imperative that the best available technique is employed for the purpose. Multiple electrophoretic techniques are employed for Basmati genotyping, although it is apparent that capillary-based genotyping methodology is a superior choice (Bligh 2000; Jain *et al* 2004; Saini *et al* 2004; Singh *et al* 2004). In the study reported here, it was demonstrated, in statistical terms, that capillary electrophoresis shows greater accuracy and replicate concordance. Furthermore, it was also shown that capillary electrophoresis is up to the mark for quantitative assay by comparison with real time PCR analysis.

Accuracy and consistency of adulterant detection

To evaluate the accuracies of agarose gel, slab gel, and capillary electrophoresis, a total of 864 independent microsatellite-based genotyping assays was carried out (SI Table 2.2). To exclude the variations that may arise due to diverse DNA sources measured at different genomic loci, each technique employed the same 12 rice varieties (Table 2.1) genotyped at 8 microsatellite loci (Table 2.2). Each assay was run in triplicate, and average values of allele sizes and deviations were computed across all varieties at each microsatellite locus. Accuracy was expressed as the

difference (in base pairs) between observed allele size (obtained by genotyping assay) and the actual allele size (obtained by sequencing). Allele sizing is carried out with co-electrophoresed or internal fragments of known sizes. Capillary electrophoresis returned the most precise allele sizes, twice as accurately as slab gel electrophoresis could produce and more than an order of magnitude greater than agarose gel electrophoresis could achieve (Table 2.4).

Table 2.4 Allele size estimation accuracies of agarose, slab-gel and capillary electrophoresis methods

Locus	Size range (bp)	Mean difference (±bp)			
		Agarose manual	Agarose software	Slab-gel electrophoresis	Capillary electrophoresis
RM171	323-347	6.50	6.88	0.08	1.80
RM55	218-235	5.42	8.08	3.05	0.44
RM202	161-182	4.22	2.59	0.40	0.45
RM72	148-173	14.61	13.02	2.15	0.08
RM348	130-139	6.03	9.64	1.30	0.58
RM241	128-144	12.53	9.15	2.40	0.28
RM44	103-113	5.92	7.32	0.03	0.21
RM1	73-108	9.00	8.69	3.31	2.00
Average		8.03	8.17	1.59	0.73

The observations on averaged mean differences suggested that if the actual size of an allele were 74 bp, capillary electrophoresis, with a mean deviation of 0.73 bp, would estimate the allele size to be in the range of 73.27–74.73 bp. This would effectively create an error bin of 3 bp (73–75 bp). Slab gel electrophoresis, on the other hand, would create a 5 bp error bin (72–76 bp), and the agarose gel electrophoresis technique would have a 17 bp range (66–82 bp). The latter values illustrate why previous studies on Basmati genotyping based on agarose gel electrophoresis technique generated contradictory results (Bligh 2000; Jain *et al* 2004; Saini *et al* 2004; Singh *et al* 2004). When the average deviations were further scrutinized for the actual distribution of allele sizes, it was observed that nearly 58% of the alleles estimated by capillary electrophoresis showed exact values (zero deviation, Figure 2.15). In all, >90% of the capillary electrophoresis assays resulted in allele sizing within a base-pair bin, whereas the corresponding estimate was <50% for slab gel electrophoresis and a dismal 7 and 2%, respectively, for

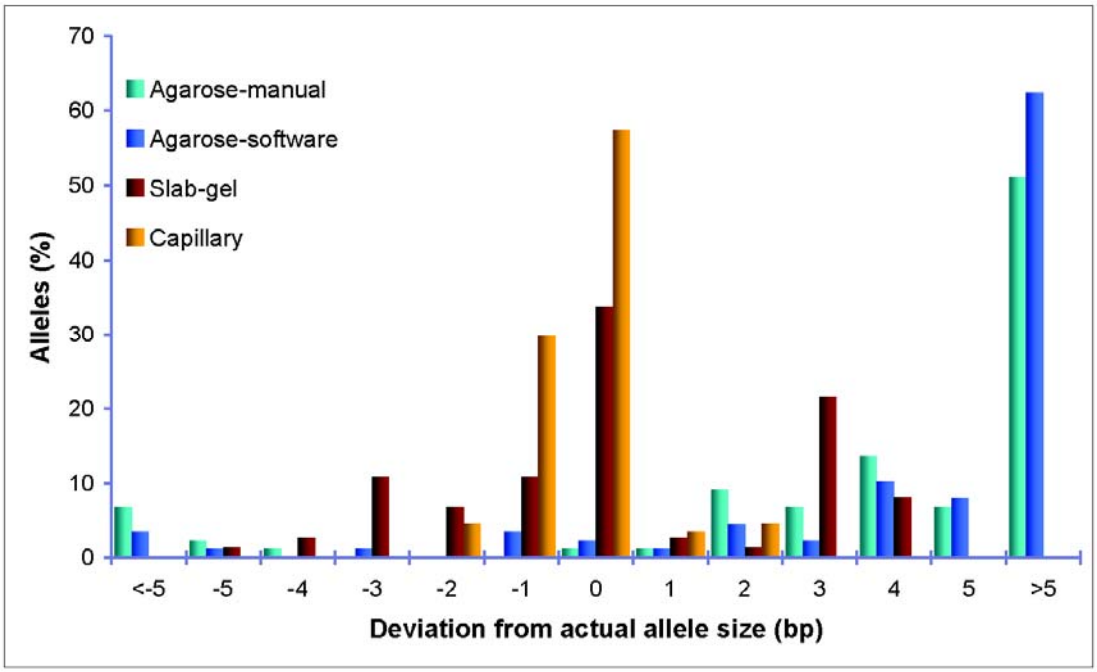


Figure 2.15 Allele size estimation accuracies of agarose, slab-gel and capillary electrophoresis methods. Y-axis presents fraction of total alleles (expressed in percentage) showing deviations from the actual allele sizes. X-axis shows the deviation (in bp) of the observed allele sizes from the actual allele sizes obtained by bidirectional sequencing.

software-based and manually measured agarose gel electrophoresis (Figure 2.15). It was observed that agarose gel electrophoresis returned overestimated allele sizes, whereas departure from the actual allele sizes in the case of slab gel electrophoresis spread like a normal distribution (Figure 2.15). Although similar algorithms were used for allele calling and allele sizing, characteristics of slab gel and capillary electrophoresis peaks differed noticeably (Figure 2.16). The broader peaks observed in slab gel meant greater chance of errors in sizing. Only microsatellite loci having trinucleotide repeats (RM348 and RM72) appear to produce sharp peaks as good as capillary electrophoresis. Broader peaks at other loci (dinucleotide repeats) restricted the bin size and ended up in differential allele sizing over repeated runs and across template DNA samples. The reproducibility problem was clearly reflected in 1 bp deviation observed in the case of slab gel electrophoresis (Table 2.5). On the other hand, capillary electrophoresis showed <0.5 bp. Agarose-based methods obviously exhibited far greater inconsistencies (~6 and ~3 bp in manual and software methods, respectively).

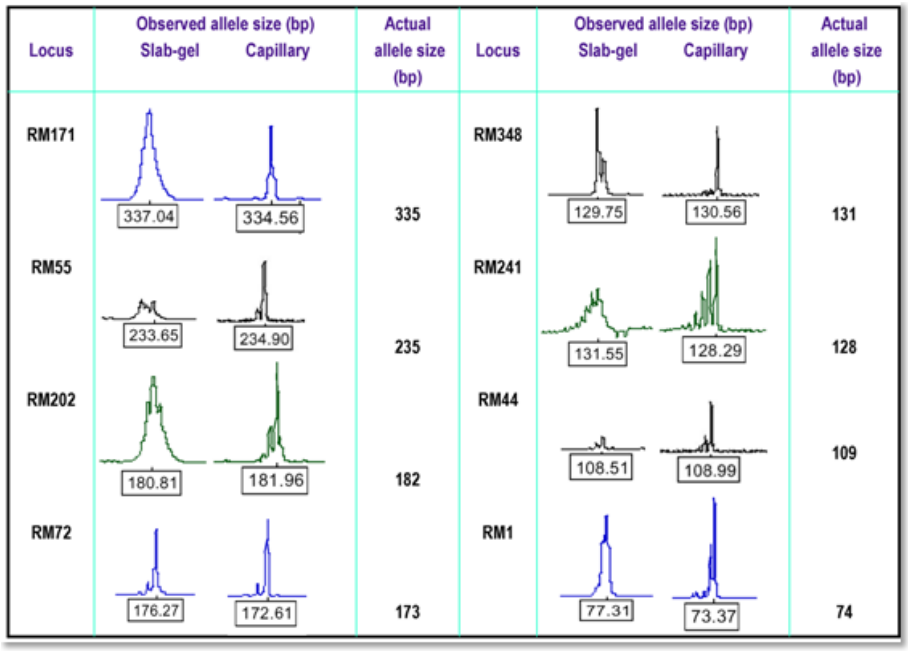


Figure 2.16 Extent of allele size differences between slab-gel and capillary electrophoresis at eight rice microsatellite loci. Each peak represents a single allele, whose size (in bp) is given below the peak. Please see Table 2 for allele size a range; only a representative profile is shown here. The actual allele sizes, determined by bidirectional sequencing, are also provided in the column for comparison.

Table 2.5 Consistency in estimation of allele sizes

Locus	Standard deviations (±bp)			
	Agarose manual	Agarose software	Slab-gel electrophoresis	Capillary electrophoresis
RM171	10.68	3.28	0.16	0.26
RM55	7.15	0.78	2.76	0.47
RM202	5.52	1.64	0.67	0.22
RM72	7.74	3.52	0.68	0.26
RM348	4.74	4.43	0.47	0.22
RM241	2.15	3.69	1.38	0.66
RM44	4.47	3.61	0.16	0.11
RM1	5.77	2.03	1.13	0.41
Average SD	6.03	2.87	0.93	0.33

Interestingly though, deviations across all of the loci did not show any significant relationship with allele size range in all three separation methods (Table 2.4). This is in contrast to locus-specific sizing bias observed in capillary electrophoresis based genotyping reported earlier (Mansfield *et al* 1996).

Accuracy of adulterant quantification

It is common for Basmati rice, a farm product, to contain adventitious mixtures because of inadvertent mixing during postharvest operations in the field or during storage. Taking this practical problem into consideration, each importing country allows a certain limit of adulteration (e.g., up to 15%) in Basmati imports, beyond which consignments are rejected. This calls for accurate quantification of adulterants in the Basmati export samples. Adulteration in Basmati lots was accurately quantified on the basis of peak areas of additional peaks in relation to the peak area corresponding to Basmati in the electropherogram. The mean difference in quantification of adulteration estimated from slab gel and capillary electrophoretic instruments assayed using premixed standard samples revealed that slab gel shows high mean difference (6.09%) compared to the capillary electrophoresis method (3.91) (Figure 2.17). Although neither assay exhibited a smooth curve, the slab gel based assay showed a greater amplitude of variations. More significantly, in the range of adulteration at which samples are accepted or rejected, slab gel electrophoresis showed unacceptably erroneous results (Figure 2.17). On the other hand, capillary electrophoresis showed greater accuracy in that critical zone (<1%). The quality of the peaks, as in detection accuracy, was also observed to affect the sensitivity of the assay to quantify the adulterant. Figure 2.18 illustrates comparison of the quantification accuracy between slab gel and capillary methods at two loci, RM348 and RM72, in a mixture of Basmati370 and Sharbati. It was obvious that the difference in the accuracies is mainly due to the peak quality, sharper peaks of capillary electrophoresis being more accurate. These observations strongly point out the indispensability of capillary electrophoresis for quantification assay to protect the interests not only of consumers but also of farmers and traders.

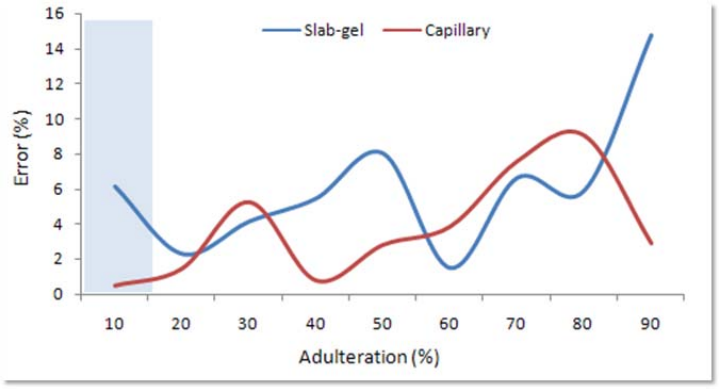


Figure 2.17 Accuracy in quantification of adulterant(s) by Slab-gel and Capillary electrophoresis. Quantification assays were carried out in standard samples of Basmati370 mixed with Haryana Basmati, in progressive proportions from 10% to 90% with a 10% interval at two informative loci, RM72 and RM348. The range of adulteration where samples are accepted or rejected (<15%) is depicted in light blue.

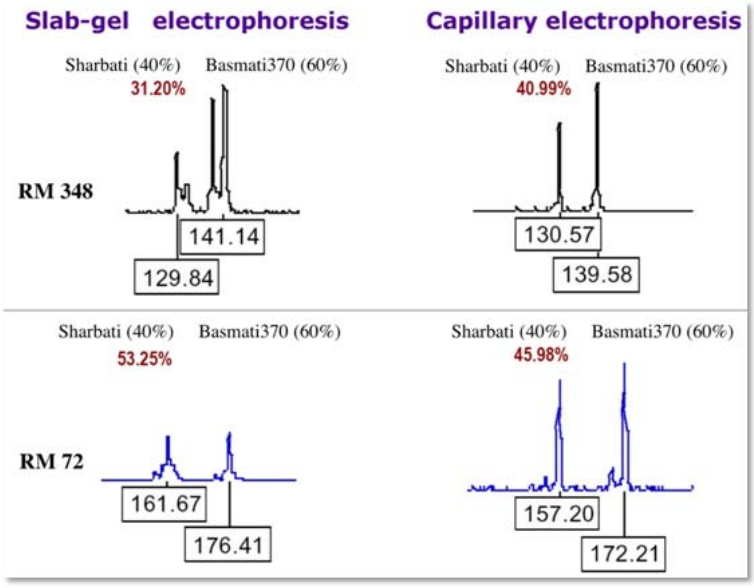


Figure 2.18 Accuracy of slab-gel and capillary electrophoresis techniques in quantification of Basmati rice adulteration. A representative profile of Basmati370 mixed with Sharbati (a common adulterant) in a 60:40 ratio is shown. Adulteration was quantified at two microsatellite loci, RM348 and RM72. Sharbati is detected by the shorter amplicons (left). Peaks are labeled with allele sizes (bp) as determined by Genotyper software; estimated adulteration values (%) are given in bold face.

Concordance studies of capillary electrophoresis and real time PCR

Comparison with other electrophoresis methods showed that capillary electrophoresis is far superior in accuracy and consistency. However, accuracy of quantification needed further confirmation using a sensitive assay. A similar but more stringent quantification assay of genetically modified (GM) food samples has been comprehensively met by the use of real time PCR based protocols (Baeumler *et al* 2006; Leimanis *et al* 2006). However, in the case of GM food, detection and quantification of transgenes have been relatively straightforward because trans- genes are distinct "foreign" DNA elements in the host genome, which can be specifically amplified with high accuracy (detection) and sensitivity (quantification) by real time PCR. On the other hand, in the case of Basmati adulteration, the so-called "adulterant" is another long-grain rice (often a recently bred Basmati derivative). So far, no Basmati-specific sequence (genic or otherwise) has been identified that, upon amplification, can generate varietal-specific private alleles. For want of signature sequences, real time PCR has not been useful in adulteration detection. Although it was possible to distinguish genuine Basmati (Basmati370) from non-Basmati (Sharbati) by real time PCR at the betaine-aldehyde dehydrogenase-2 (*bad-2*) gene (Bradbury *et al* 2005), this assay is not suitable for practical applications as many adulterants also carry the same genuine Basmati allele at this locus. In the present study, however, it was employed only to assess the quantification accuracy of capillary electrophoresis. Capillary electrophoresis showed a mean deviation of 1.76% compared to real time assay (Table 2.6), indicating that the estimation of adulteration based on the capillary electrophoresis method is realistically accurate for practical applications.

Table 2.6 Comparison of quantification of adulteration by capillary electrophoresis with real-time PCR

	Adulteration quantified by		
% of Actual adulteration	Real-Time PCR (RT)	Capillary (CE)	Deviation of RT and CE
1	1.82	1.9	0.08
5	5.18	8.89	3.71
7	7	9.37	2.37
10	11.48	13.68	2.2
15	15.9	15.98	0.08
30	30.9	33.5	2.6
50	51.45	50.15	1.3
			1.76

Conclusions

2.4 Conclusions

Traceability is increasingly becoming standard across agri-produce and agri-food market, largely driven by the experience with GM food as well as availability of a range of DNA-based concepts and technologies adapted to suit different industry needs. Traceability in the form of detection and quantification of adulteration forms an essential component of Basmati trade management as well as post-marketing surveillance. Given the unfeasibility of traceability through otherwise accurate and high-throughput technology like real time PCR at the moment, a microsatellite-based multiplex assay has been designed and tested for greater accuracy and practicability. Basmati trade is moving toward a uniform code of practice with a long-term goal to peg the adulteration of pure Basmati samples at 7%. The CE-based microsatellite multiplex assay presented here has high throughput and can effectively identify and quantify the level of adulteration with sensitivity and accuracy levels adequate for Basmati trade. The methodology is amenable to the development of a ready-to-use kit that can facilitate adherence to a common certification protocol by Basmati exporters, importers and regulators. The protocol has been authorized by the export development and regulatory agencies of the Government of India for issuance of certificate of purity for Basmati export samples. Based on the multiplex assay, 525 referral samples of Basmati rice destined to international market have so far been certified. Patents are pending for the protocol reported here (USPTO 10/357, 488 and 11/406, 257; PCT/IN06/00254). The know-how on ready-to use kit has been transferred to M/S Labindia (Indian partners of Applied Biosystems) under license transfer agreement.

In another study, the necessity of capillary electrophoresis for a microsatellite marker based assay for accurate and consistent detection and quantification of adulteration of Basmati samples has been demonstrated. Even if routinely used agarose and polyacrylamide gel electrophoresis methods can distinguish Basmati rice varieties (provided microsatellite makers have huge differences among allele sizes, at least 10–15 bp), they are not suitable for accurate quantification of adulterants. Although advanced versions of PAGE or slab gel electrophoresis coupled with fluorescent labeled primers allow the detection and quantification of adulteration in Basmati rice, they often give inaccurate estimation of allele sizes. Capillary electrophoresis, on the other hand, produces accurate and consistent results. Hence, it is necessary that a microsatellite-based genotyping assay to detect and quantify adulteration in Basmati rice employs capillary electrophoresis.

References

2.5 References

Altet L, Francino O and Sanchez A (2001) Microsatellite polymorphism in closely related dogs. J Hered 92: 276-9.

Amarawathi Y, Singh R, Singh A, Singh V, Mohapatra T, Sharma T and Singh N (2008) Mapping of quantitative trait loci for basmati quality traits in rice (*Oryza sativa* L.). Molecular Breeding 21: 49-65.

Archak S, Lakshminarayanareddy V and Nagaraju J (2007) High-throughput multiplex microsatellite marker assay for detection and quantification of adulteration in Basmati rice (*Oryza sativa*). Electrophoresis 28: 2396-405.

Baeumler S, Wulff D, Tagliani L and Song P (2006) A real-time quantitative PCR detection method specific to widestrike transgenic cotton (event 281-24-236/3006-210-23). J Agric Food Chem 54: 6527-34.

Bandelj D, Jakše J and Javornik B (2004) Amplification of fluorescent-labelled microsatellite markers in olives by a novel, economic method. Acta agriculturae slovenica 83 323-329.

Bellagamba F, Valfrè F, Panseri S and Moretti VM (2003) Polymerase chain reaction–based analysis to detect terrestrial animal protein in fish meal. Journal of Food Protection 66: 682-685.

Blair M, Hedetale V and McCouch S (2002) Fluorescent-labeled microsatellite panels useful for detecting allelic diversity in cultivated rice (*Oryza sativa* L.) Theoretical and Applied Genetics 105: 449-457.

Bligh HFJ (2000) Detection of adulteration of Basmati rice with non-premium long-grain rice. International Journal of Food Science & Technology 35: 257.

Bradbury LM, Fitzgerald TL, Henry RJ, Jin Q and Waters DL (2005) The gene for fragrance in rice. Plant Biotechnol J 3: 363-70.

Bubner B and Baldwin IT (2004) Use of real-time PCR for determining copy number and zygosity in transgenic plants. Plant Cell Rep 23: 263-71.

Butler JM, Buel E, Crivellente, F and McCord BR (2004) Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. Electrophoresis 25: 1397–1412.

Buttery RG, Ling LC, Juliano BO and G, Turnbaugh JG (1983) Cooked rice aroma and 2-acetyl-1-pyrroline. J. Agric. Food Chem 31: 823-826.

Calvo JH, Osta R and Zaragoza P (2002) Quantitative PCR detection of pork in raw and heated ground beef and pate. J Agric Food Chem 50: 5265-7.

Cankar K, Stebih D, Dreo T, Zel J and Gruden K (2006) Critical points of DNA quantification by real-time PCR--effects of DNA extraction method and sample matrix on quantification of genetically modified organisms. BMC Biotechnol 6: 37.

Coburn JR, Temnykh S, Paul E and McCouch S (2002) Design and Application of Microsatellite Marker Panels for Semiautomated Genotyping of Rice (*Oryza sativa* L.) Crop Sci. 42: 2092–2099.

Dhiman B and Singh M (2003) Molecular Detection of Cashew Husk (*Anacardium occidentale*) Adulteration in Market Samples of Dry Tea (*Camellia sinensis*). Planta Med 69.

Findlay I, Taylor A, Quirke P, Frazier R and Urquhart A (1997) DNA fingerprinting from single cells. Nature 389: 555-6.

Garris AJ, Tai TH, Coburn J, Kresovich S and McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. Genetics 169: 1631-8.

Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E and Sullivan K (1994) Identification of the remains of the Romanov family by DNA analysis. Nat Genet 6: 130-5.

Gilliland G, Perrin S, Blanchard K and Bunn HF (1990) Analysis of cytokine mRNA and DNA: detection and quantitation by competitive polymerase chain reaction. Proc Natl Acad Sci U S A 87: 2725-9.

Jain S, Jain RK and McCouch SR (2004) Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently-labeled microsatellite markers. Theor Appl Genet 109: 965-77.

Kang H, Cho YG, Yoon U and Eun MY (1998) A Rapid DNA Extraction Method for RFLP and PCR Analysis from a Single Dry Seed. Plant Molecular Biology Reporter 16: 1-9.

Koumi P, Green H, Hartley S, Jordan D, Lahec S, Livett R, Tsang K and Ward D (2004) Evaluation and validation of the ABI 3700, ABI3100, and the MegaBACE 1000 capillary array electrophoresis instruments for use with short tandem repeat microsatellite typing in a forensic environment Electrophoresis 25: 2227-2241.

Leimanis S, Hernandez M, Fernandez S, Boyer F, Burns M, Bruderer S, Glouden T, Harris N, Kaeppeli O, Philipp P, Pla M, Puigdomenech P, Vaitilingom M, Bertheau Y and Remacle J (2006) A microarray-based detection system for genetically modified (GM) food ingredients. Plant Mol Biol 61: 123-39.

Mansfield ES, Vainer M, Enad S, Barker DL, Harris D, Rappaport E and Fortina P (1996) Sensitivity, reproducibility, and accuracy in short tandem repeat genotyping using capillary array electrophoresis. Genome Res 6: 893-903.

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li, Z, Xing, Y, Zhang, Q, Kono, I, Yano, M, Fjellstrom, R, DeClerck, G, Schneider, D, Cartinhour, S, Ware, D and Stein, L (2002) Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). DNA Res 9: 199-207.

Mooney R, Chappell L and Knight A (2006) Evaluation of a polymerase chain reaction-based heteroduplex assay for detecting the adulteration of processed orange juice with mandarin juice. Journal - Association of Official Analytical Chemists 89: 1052-60.

Nagaraju J, Kathirvel M, Kumar RR, Siddiq EA and Hasnain SE (2002a) Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. Proc Natl Acad Sci U S A 99: 5836-41.

Nagaraju J, Kathirvel M, Subbaiah EV, Muthulakshmi M and Kumar LD (2002b) FISSR-PCR: a simple and sensitive assay for highthroughput genotyping and genetic mapping. Mol Cell Probes 16: 67-72.

Nicholas KB, Nicholas HBJ and Deerfield DWI (1997) GeneDoc: Analysis and Visualization of Genetic Variation. EMBNEW.NEWS 4: 14.

Saghai-Maroof MA, Soliman KM, Jorgensen RA and Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci U S A 81: 8014-8.

Saini N, Jain N, Jain S and Jain RK (2004) Assessment of genetic diversity within and among Basmati and non-Basmati rice varieties using AFLP, ISSR and SSR markers. Euphytica 140: 133-146.

Singh RK, Sharma RK, Singh AK, Singh VP, Singh NK, Tiwari SP and Mohapatra T (2004) Suitability of mapped sequence tagged microsatellite site markers for establishing distinctness, uniformity and stability in aromatic rice. Euphytica 135: 135-143.

Spaniolas S, Bennett MJ, Tucker GA and May ST (2006) Authentication of Coffee by Means of PCR-RFLP Analysis and Lab-on-a-Chip Capillary Electrophoresis. J. Agric. Food Chem. 54: 7466-7470.

Steele KA, Ogden R, McEwing R, Briggs H, and Gorham J (2008) InDel markers distinguish Basmatis from other fragrant rice varieties. Field Crops Research 105:81–87

Tagliaro, F and Bortolotti, F (2006) Recent advances in the applications of CE to forensic sciences (2001-2004) Electrophoresis. 27: 231-43.

Tang T, Huang J, Zhong Y and Shi S (2004) High-throughput S-SAP by fluorescent multiplex PCR and capillary electrophoresis in plants. Journal of Biotechnology 114: 59-68.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876-82.

Vemireddy LR, Archak S and Nagaraju J (2007) Capillary electrophoresis is essential for microsatellite marker based detection and quantification of adulteration of Basmati rice (*Oryza sativa*). J Agric Food Chem 55: 8112-7.

Supporting information

2.6 Supporting Information (SI)

SI Table 2.1 Population study

Bulked samples of one traditional and one evolved Basmati cultivars from various places across growing area were collected and allele profiles at eight microsatellite loci were generated using multiplex assay. There was absolute homogeneity in the inter-population allele profiles. Populations are labeled after the place of sampling. Suffix 'B' indicates that it is one of the multiple samples from a particular place.

1. Taraori Basmati

Population	RM72	RM202	RM348	RM44	RM55	RM171	RM241	RM1
Alewa B	173	182	140	113	220	336	128	73
Ambala	173	182	140	113	220	336	128	73
Amritsar	173	182	140	113	220	336	128	73
Babain B	173	182	140	113	220	336	128	73
Bahadurgarh B	173	182	140	113	220	336	128	73
Bapauli B	173	182	140	113	220	336	128	73
Barara B	173	182	140	113	220	336	128	73
Bhikhiwindi	173	182	140	113	220	336	128	73
Devigarh	173	182	140	113	220	336	128	73
Gannoua B	173	182	140	113	220	336	128	73
Gharunda B	173	182	140	113	220	336	128	73
Gohana	173	182	140	113	220	336	128	73
Guhla B	173	182	140	113	220	336	128	73
Hansi	173	182	140	113	220	336	128	73
Indri	173	182	140	113	220	336	128	73
Israna B	173	182	140	113	220	336	128	73
Jandialaguru	173	182	140	113	220	336	128	73
Kaithal B	173	182	140	113	220	336	128	73
Kalayat B	173	182	140	113	220	336	128	73
Kharkhoda B	173	182	140	113	220	336	128	73
Madlouda B	173	182	140	113	220	336	128	73
Nariangarh B	173	182	140	113	220	336	128	73
Narwana	173	182	140	113	220	336	128	73
Nilokheri	173	182	140	113	220	336	128	73
Panipat	173	182	140	113	220	336	128	73
Patran	173	182	140	113	220	336	128	73
Pundri B	173	182	140	113	220	336	128	73
Radour B	173	182	140	113	220	336	128	73
Rajaund	173	182	140	113	220	336	128	73
Rajpura	173	182	140	113	220	336	128	73
Ratia B	173	182	140	113	220	336	128	73
Rohtak	173	182	140	113	220	336	128	73
Safidon B	173	182	140	113	220	336	128	73
Shahbad	173	182	140	113	220	336	128	73
Sirsa	173	182	140	113	220	336	128	73
Taran Taran	173	182	140	113	220	336	128	73
Thanesar B	173	182	140	113	220	336	128	73
Tohana B	173	182	140	113	220	336	128	73
Uchanakalan	173	182	140	113	220	336	128	73

2.Pusa Basmati

Population	RM72	RM202	RM348	RM44	RM55	RM171	RM241	RM1
Ambala	158	182	131	113	230	344	128	73
Assandh	158	182	131	113	230	344	128	73
Babain	158	182	131	113	230	344	128	73
Bapoli	158	182	131	113	230	344	128	73
Barara	158	182	131	113	230	344	128	73
Bhuna	158	182	131	113	230	344	128	73
Bilaspur	158	182	131	113	230	344	128	73
Chhachroli	158	182	131	113	230	344	128	73
Ellnabad	158	182	131	113	230	344	128	73
Fatiabad	158	182	131	113	230	344	128	73
Gannour	158	182	131	113	230	344	128	73
Gharonda	158	182	131	113	230	344	128	73
Gulha	158	182	131	113	230	344	128	73
Indri	158	182	131	113	230	344	128	73
Israna	158	182	131	113	230	344	128	73
Jagadhari	158	182	131	113	230	344	128	73
Kaithal	158	182	131	113	230	344	128	73
Kalayat	158	182	131	113	230	344	128	73
Karnal	158	182	131	113	230	344	128	73
Kathura	158	182	131	113	230	344	128	73
Kharkhoda	158	182	131	113	230	344	128	73
Ladwa	158	182	131	113	230	344	128	73
Madlouda	158	182	131	113	230	344	128	73
Mundlana	158	182	131	113	230	344	128	73
Narayangarh	158	182	131	113	230	344	128	73
Nissing	158	182	131	113	230	344	128	73
Panipat	158	182	131	113	230	344	128	73
Pehowa	158	182	131	113	230	344	128	73
Pundri	158	182	131	113	230	344	128	73
Rai	158	182	131	113	230	344	128	73
Rajond	158	182	131	113	230	344	128	73
Rania	158	182	131	113	230	344	128	73
Ratia	158	182	131	113	230	344	128	73
Sirsa	158	182	131	113	230	344	128	73
Sonepat	158	182	131	113	230	344	128	73
Thanesar	158	182	131	113	230	344	128	73

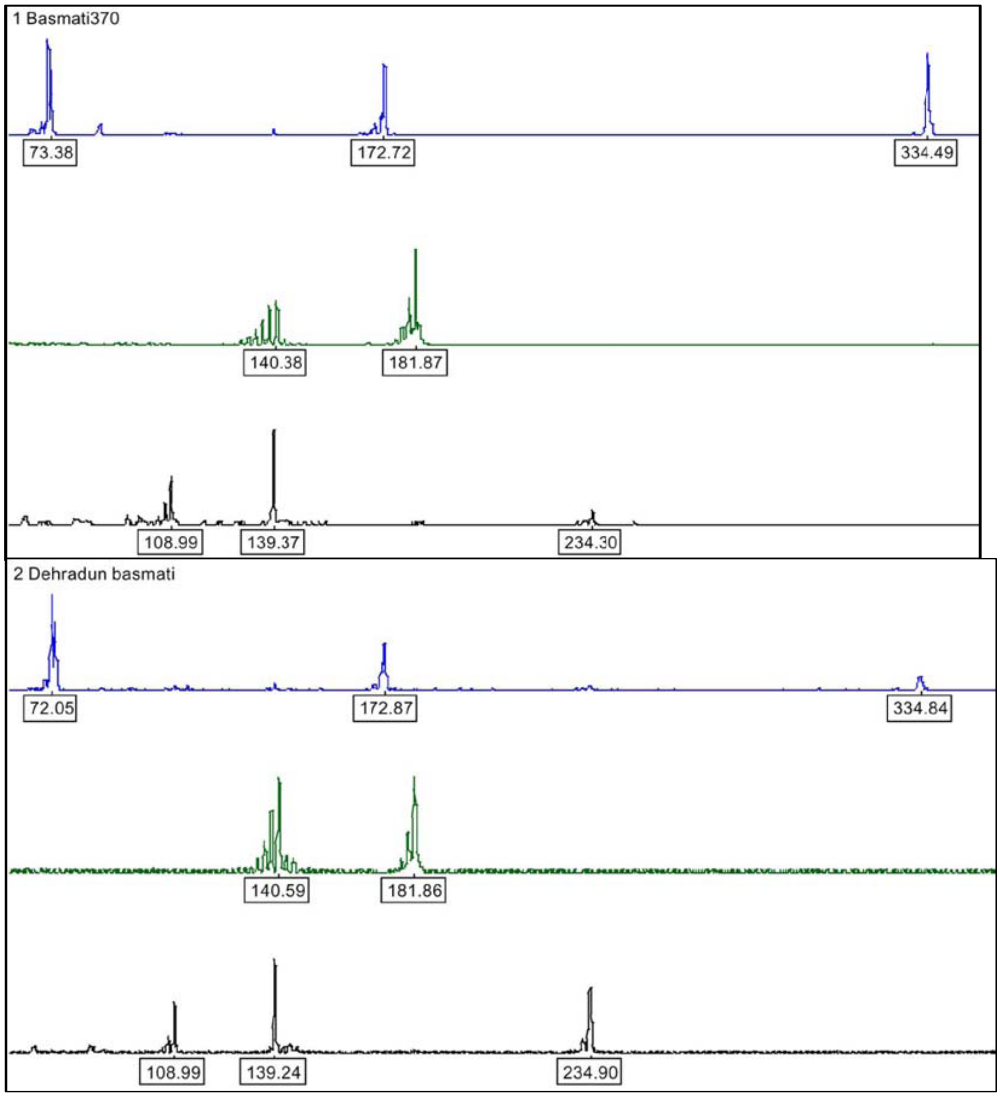
SI Table 2.2 Allele sizes estimated using 8 microsatellite markers across 12 rice varieties in different electrophoresis methods

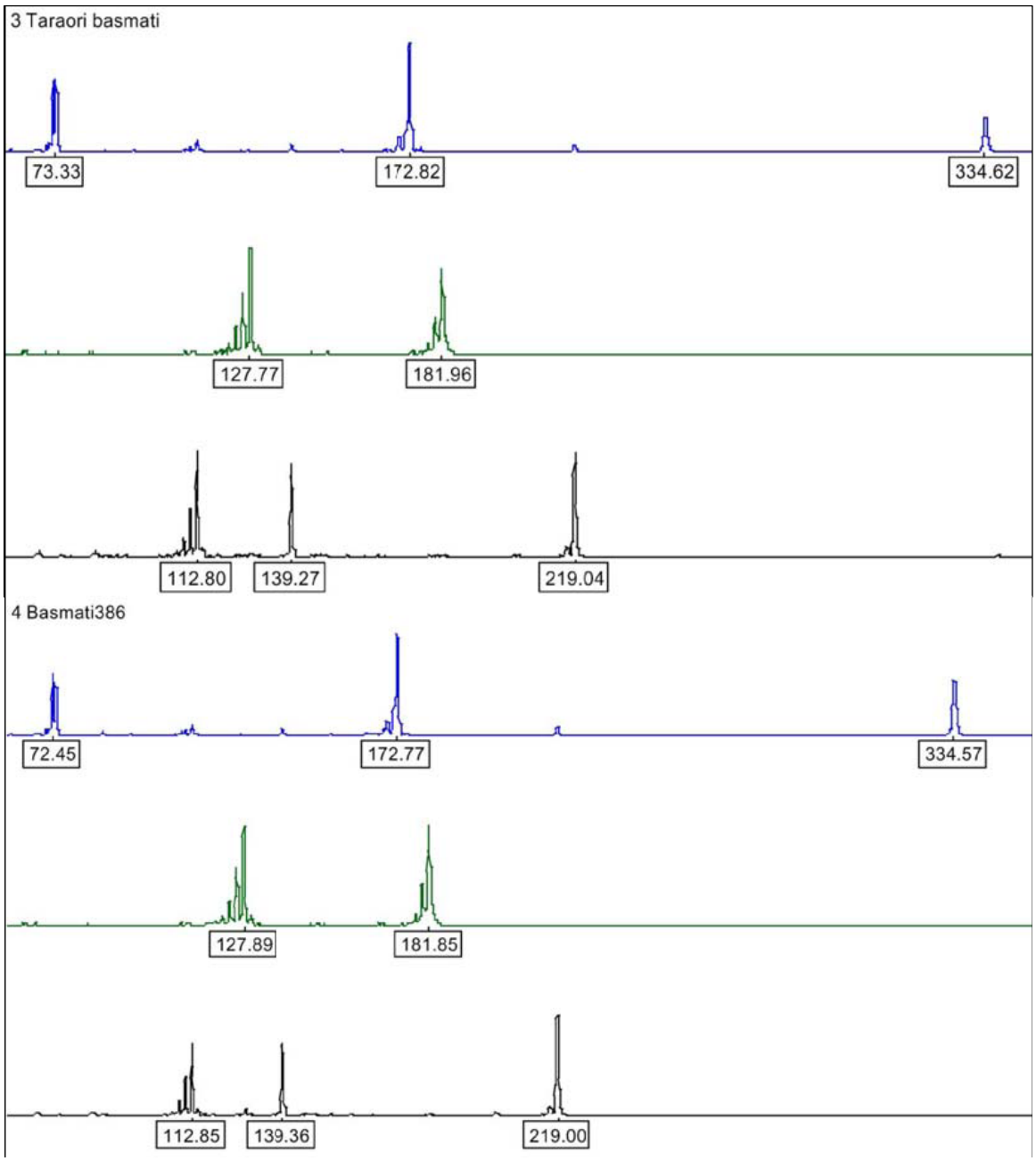
S.No.	Name of the variety	Agarose (Manual)	Agarose(Software)	Slabgel	Capillary	Sequencing
	RM1					
1	Basmati370	85.00	79.87	78.00	72.83	74.00
2	Dehradun Basmati	85.00	81.38	78.00	72.63	74.00
3	Taraori Basmati	85.00	80.73	78.00	73.07	74.00
4	Basmati386	84.67	79.23	78.00	72.99	74.00
5	Ranbir Basmati	85.00	79.23	78.00	73.51	74.00
6	Basmati217	103.33	102.53	103.00	99.27	108.00
7	Haryana Basmati	107.67	106.80	109.00	107.93	108.00
8	Pusa Basmati	NA	NA	NA	NA	NA
9	Super Basmati	79.00	68.89	108.50	106.08	108.00
10	Sharbati	99.67	94.86	108.50	106.16	108.00
11	Jaya	97.33	91.36	NA	105.78	108.00
12	IR64	96.67	87.99	NA	105.82	108.00
	RM241					
1	Basmati370	147.67	148.97	143.33	138.97	139.00
2	Dehradun Basmati	150.33	154.04	141.33	140.05	139.00
3	Taraori Basmati	140.67	140.60	130.67	127.96	128.00
4	Basmati386	140.67	138.99	131.00	127.97	128.00
5	Ranbir Basmati	152.00	154.04	146.00	144.58	143.00
6	Basmati217	140.67	137.39	131.00	127.99	128.00
7	Haryana Basmati	140.67	135.81	130.50	128.00	128.00
8	Pusa Basmati	140.67	133.36	130.50	128.05	128.00
9	Super Basmati	148.67	146.26	131.00	127.91	128.00
10	Sharbati	140.67	131.84	128.50	125.91	126.00
11	Jaya	140.67	128.83	NA	127.89	128.00
12	IR64	140.00	127.35	NA	129.67	130.00
	RM44					
1	Basmati370	111.67	113.29	109.00	109.51	109.00
2	Dehradun Basmati	113.33	115.47	109.00	108.96	109.00
3	Taraori Basmati	116.00	120.10	113.00	112.81	113.00
4	Basmati386	116.67	121.68	113.00	112.85	113.00
5	Ranbir Basmati	113.67	119.26	109.00	108.92	109.00
6	Basmati217	110.33	112.49	102.67	103.25	103.00
7	Haryana Basmati	110.33	111.03	103.00	103.19	103.00
8	Pusa Basmati	117.00	122.43	113.00	112.88	113.00
9	Super Basmati	115.33	119.26	103.00	103.25	103.00
10	Sharbati	110.67	108.16	103.00	103.28	103.00
11	Jaya	110.33	104.61	103.00	103.20	103.00
12	IR64	109.67	101.90	NA	103.24	103.00
	RM202					
1	Basmati370	185.33	186.40	182.00	182.10	182.00
2	Dehradun Basmati	186.00	185.54	182.33	181.98	182.00
3	Taraori Basmati	186.00	185.54	182.00	182.00	182.00
4	Basmati386	186.00	184.09	181.50	182.02	182.00
5	Ranbir Basmati	186.67	184.09	181.50	182.03	182.00
6	Basmati217	186.00	180.96	181.50	182.11	182.00
7	Haryana Basmati	168.67	160.47	157.00	161.67	160.00
8	Pusa Basmati	186.00	181.80	182.00	182.17	182.00
9	Super Basmati	NA	NA	NA	NA	NA
10	Sharbati	NA	NA	NA	NA	NA
11	Jaya	169.33	166.60	NA	161.65	160.00
12	IR64	188.67	191.11	NA	185.63	184.00

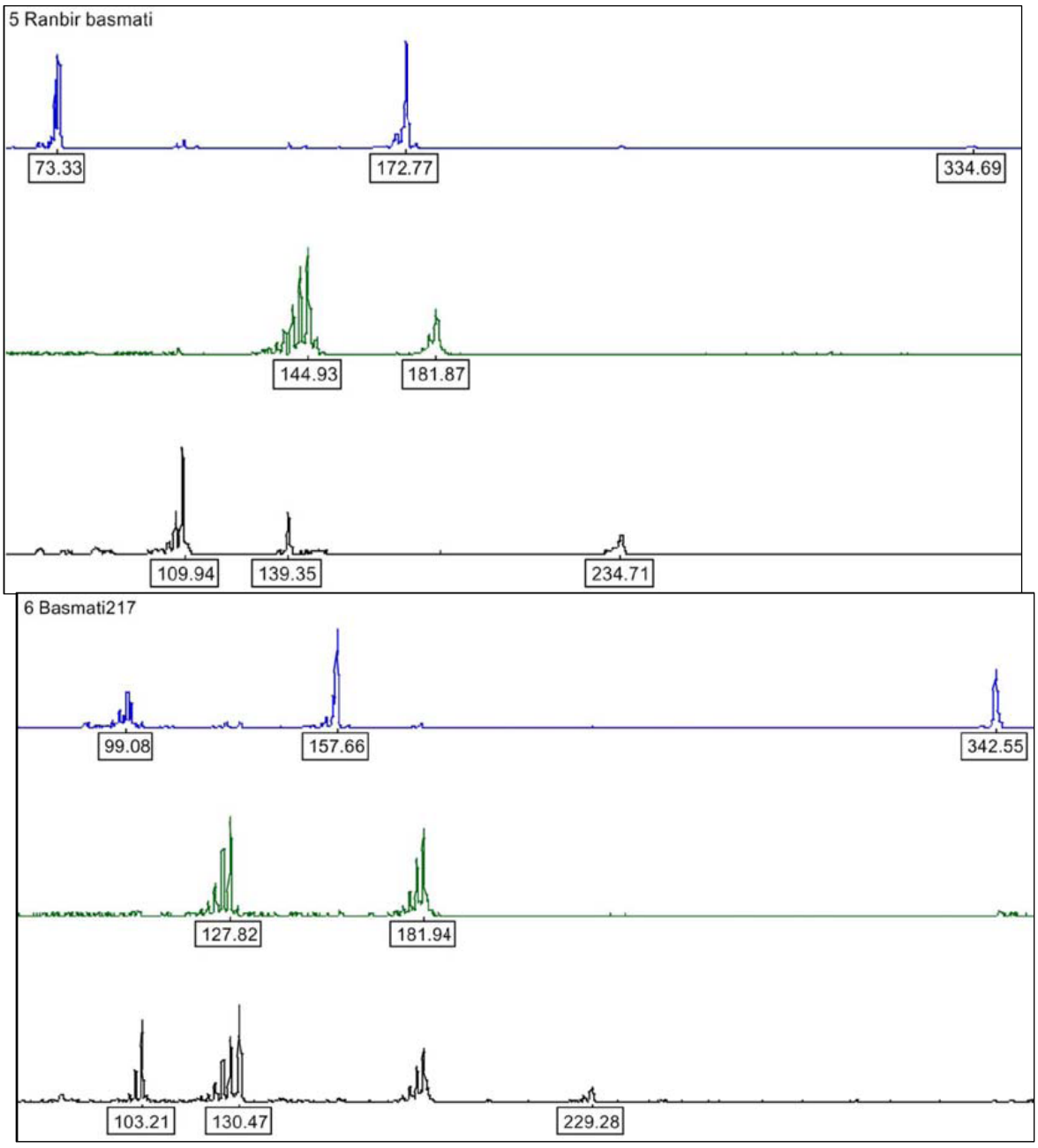
S.No.	Name of the variety	Agarose (Manual)	Agarose(Software)	Slabgel	Capillary	Sequencing
	RM171					
1	Basmati370	340.00	351.61	336.00	334.91	336.00
2	Dehradun Basmati	339.33	354.82	336.00	335.03	336.00
3	Taraori Basmati	338.33	364.46	336.00	334.83	336.00
4	Basmati386	338.33	360.13	336.00	334.95	336.00
5	Ranbir Basmati	336.67	364.46	336.00	334.93	336.00
6	Basmati217	340.33	382.32	344.00	342.76	344.00
7	Haryana Basmati	338.67	390.52	344.00	343.18	344.00
8	Pusa Basmati	338.00	396.29	344.00	342.77	344.00
9	Super Basmati	337.00	406.45	344.00	342.71	344.00
10	Sharbati	NA	NA	NA	NA	NA
11	Jaya	311.33	390.52	NA	322.02	332.00
12	IR64	325.33	438.49	NA	346.37	348.00
	RM348					
1	Basmati370	141.67	144.23	138.50	139.40	140.00
2	Dehradun Basmati	141.67	150.74	138.50	139.36	140.00
3	Taraori Basmati	141.67	153.26	138.50	139.32	140.00
4	Basmati386	142.33	154.75	138.50	139.39	140.00
5	Ranbir Basmati	142.33	155.48	138.00	139.38	140.00
6	Basmati217	137.33	147.09	130.00	130.57	131.00
7	Haryana Basmati	137.33	145.98	130.00	130.54	131.00
8	Pusa Basmati	137.33	148.73	130.00	130.53	131.00
9	Super Basmati	144.33	158.67	130.00	130.50	131.00
10	Sharbati	140.00	146.43	130.00	130.20	131.00
11	Jaya	142.00	147.25	NA	130.49	131.00
12	IR64	141.33	144.23	NA	131.58	131.00
	RM72					
1	Basmati370	183.33	181.04	175.50	173.18	173.00
2	Dehradun Basmati	184.33	181.55	175.50	173.06	173.00
3	Taraori Basmati	186.33	182.05	175.50	173.01	173.00
4	Basmati386	187.00	182.57	175.50	172.97	173.00
5	Ranbir Basmati	187.33	181.04	175.50	173.08	173.00
6	Basmati217	NA	NA	NA	NA	NA
7	Haryana Basmati	179.33	174.78	161.00	157.99	158.00
8	Pusa Basmati	179.00	181.04	161.00	158.01	158.00
9	Super Basmati	191.33	191.04	161.00	157.75	158.00
10	Sharbati	NA	NA	NA	NA	NA
11	Jaya	180.00	181.04	NA	157.82	158.00
12	IR64	178.33	181.04	NA	163.89	164.00
	RM55					
1	Basmati370	239.00	250.00	232.00	234.74	235.00
2	Dehradun Basmati	239.00	241.00	232.00	234.73	235.00
3	Taraori Basmati	230.00	226.08	216.50	219.15	220.00
4	Basmati386	230.00	228.50	216.50	219.32	220.00
5	Ranbir Basmati	239.00	243.58	230.00	234.17	235.00
6	Basmati217	235.00	235.92	227.50	229.74	230.00
7	Haryana Basmati	235.33	234.78	227.50	229.81	230.00
8	Pusa Basmati	235.33	235.92	227.50	229.02	230.00
9	Super Basmati	239.00	241.00	227.50	229.77	230.00
10	Sharbati	233.33	234.74	227.50	229.66	230.00
11	Jaya	239.00	245.02	NA	236.85	237.00
12	IR64	233.00	242.42	NA	229.83	230.00

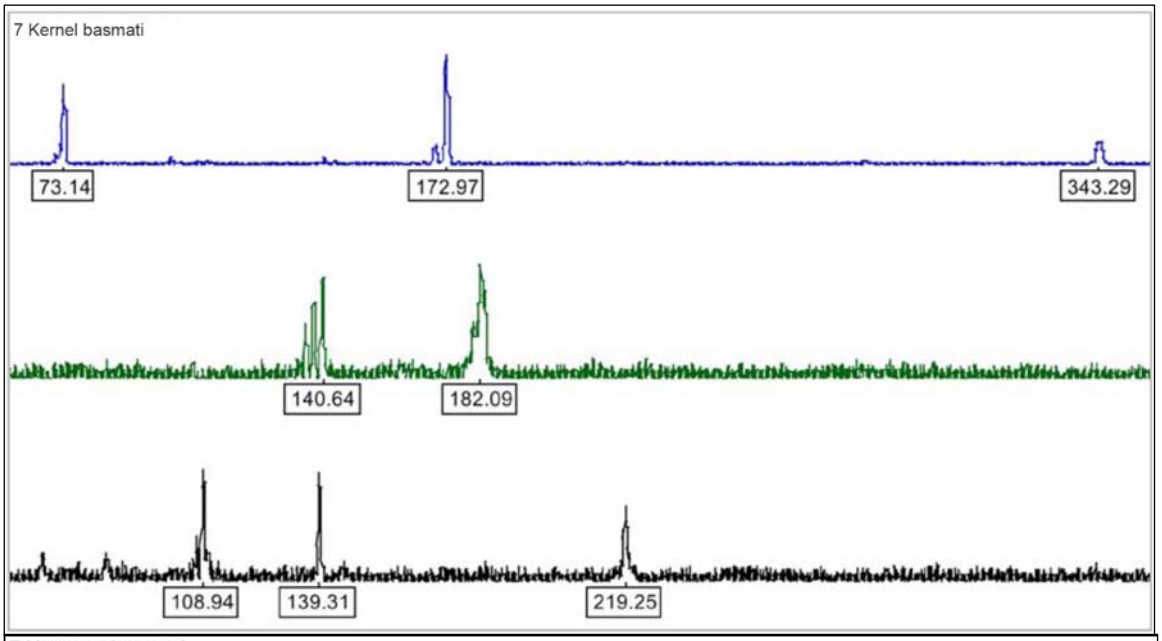
NA-Not available

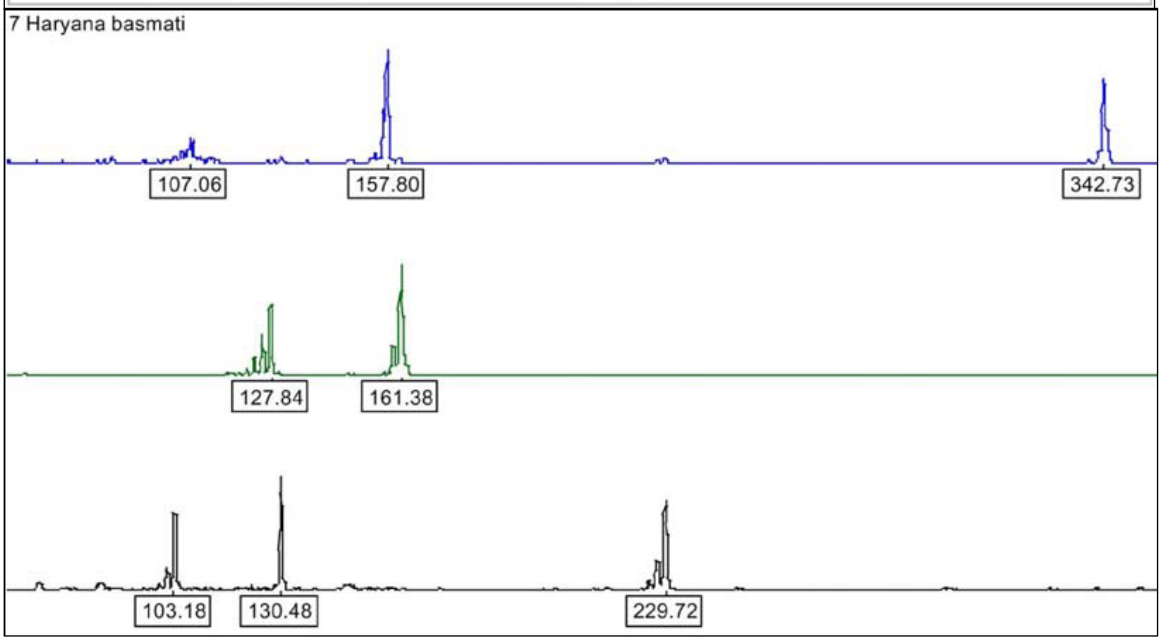
SI Figure 2.1 Representative Multiplex allele profiles of traditional, evolved, and non-Basmati rice varieties Each peak of the multiplex profile represents the variety specific allele at that locus, with allele size annotated in the box. The loci (order: from left to right) are RM1, RM72 and RM171 (in blue); RM241 and RM202 (in green); RM44, RM348 and RM55 (in black). Please see Table 1 in the main document for details regarding the rice varieties and Table 2 for details regarding the loci.

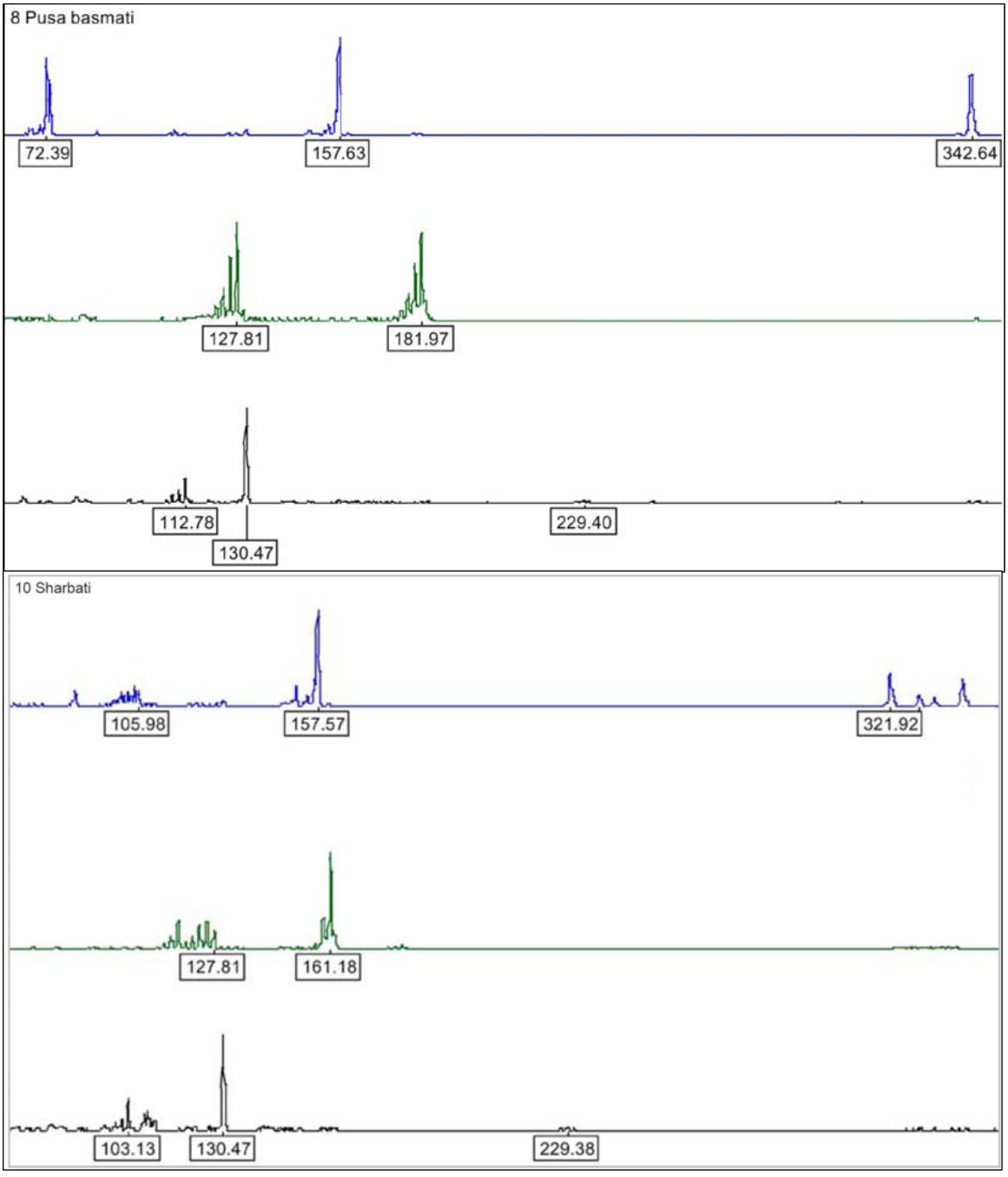


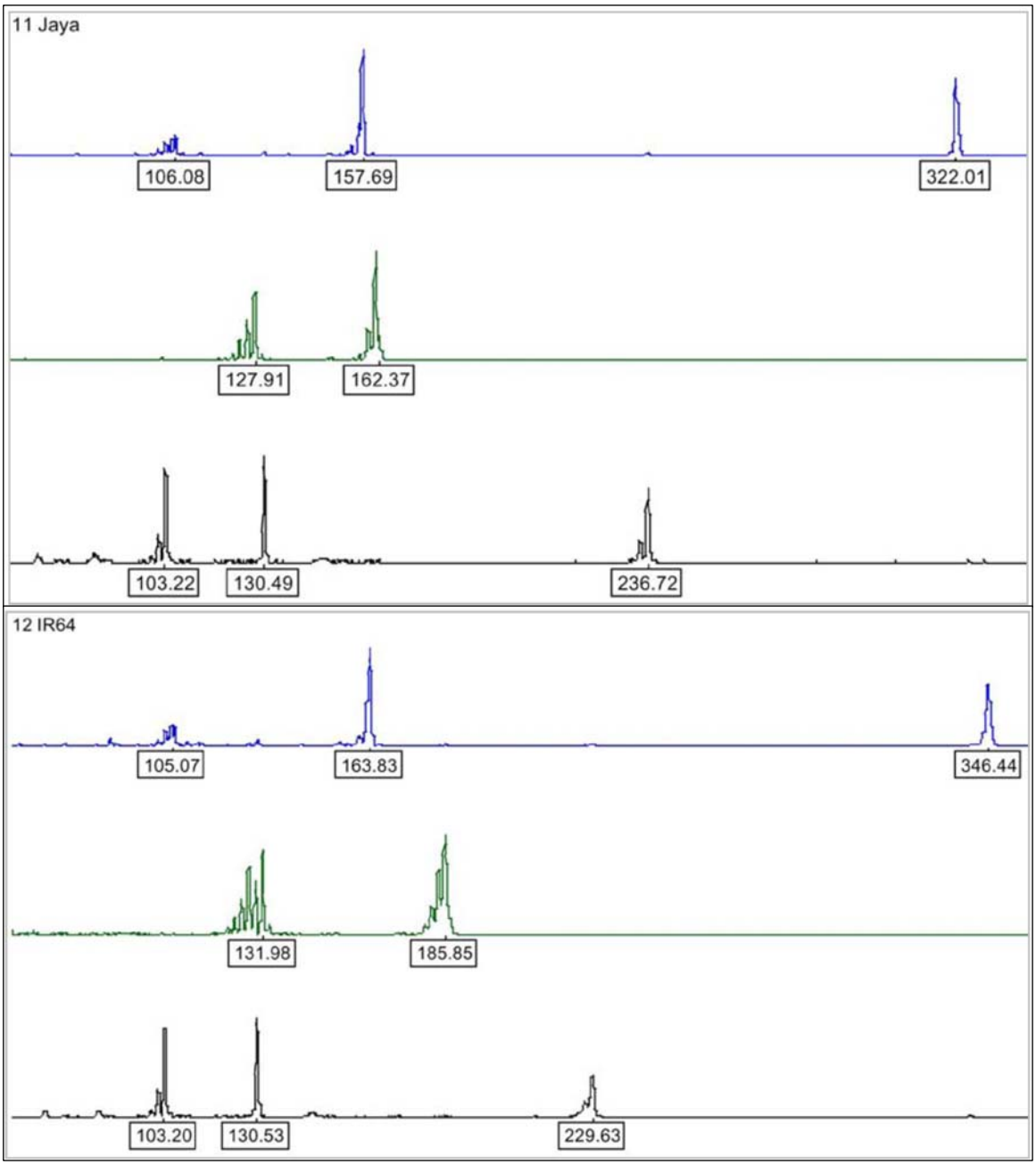

3 Taraori basmati
73.33
172.82
334.62
127.77
181.96
112.80
139.27
219.04
4 Basmati386
72.45
172.77
334.57
127.89
181.85
112.85
139.36
219.00


5 Ranbir basmati
73.33
172.77
334.69
144.93
181.87
109.94
139.35
234.71
6 Basmati217
99.08
157.66
342.55
127.82
181.94
103.21
130.47
229.28

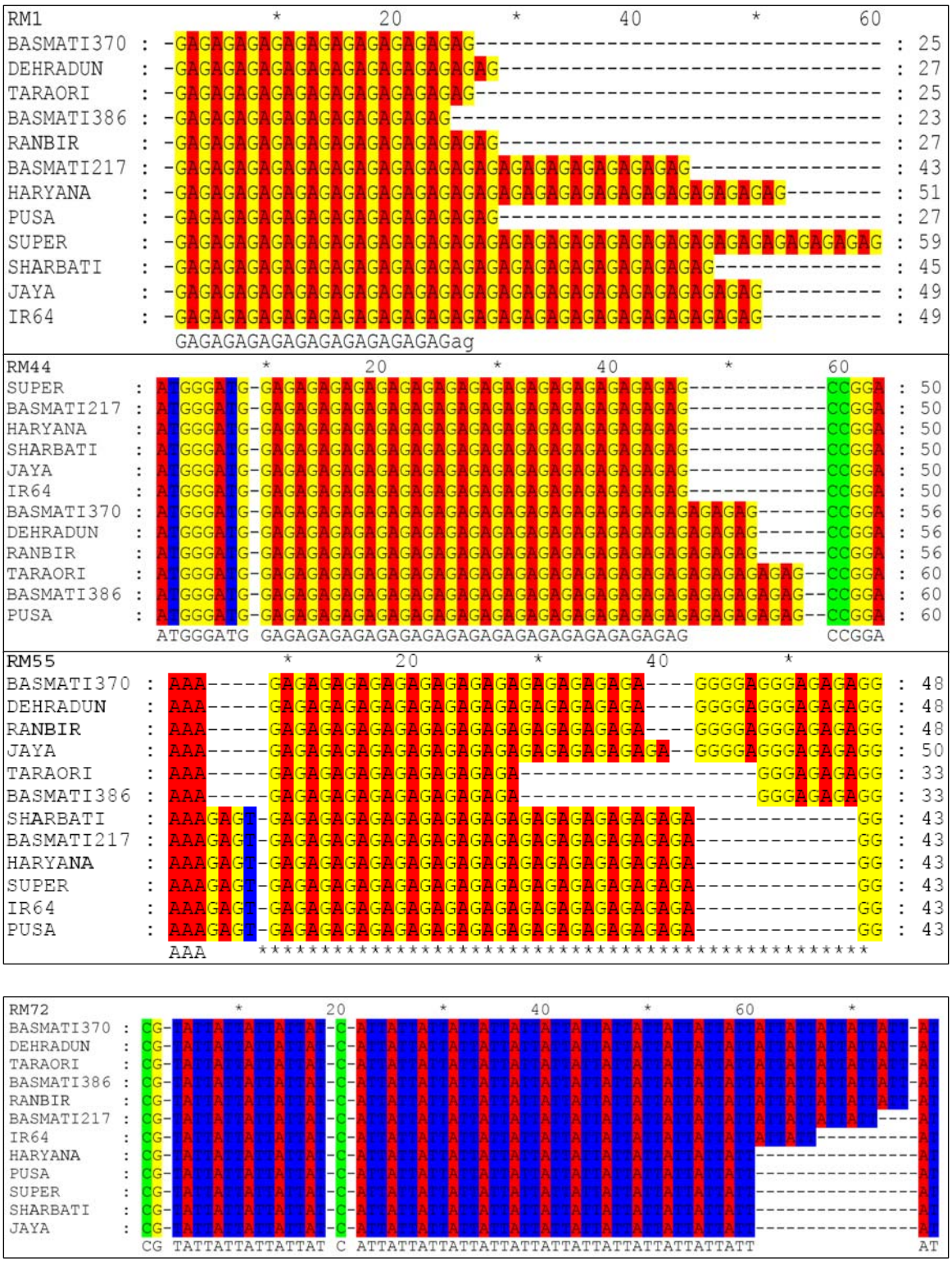

7 Kernel basmati
73.14
172.97
343.29
140.64
182.09
108.94
139.31
219.25

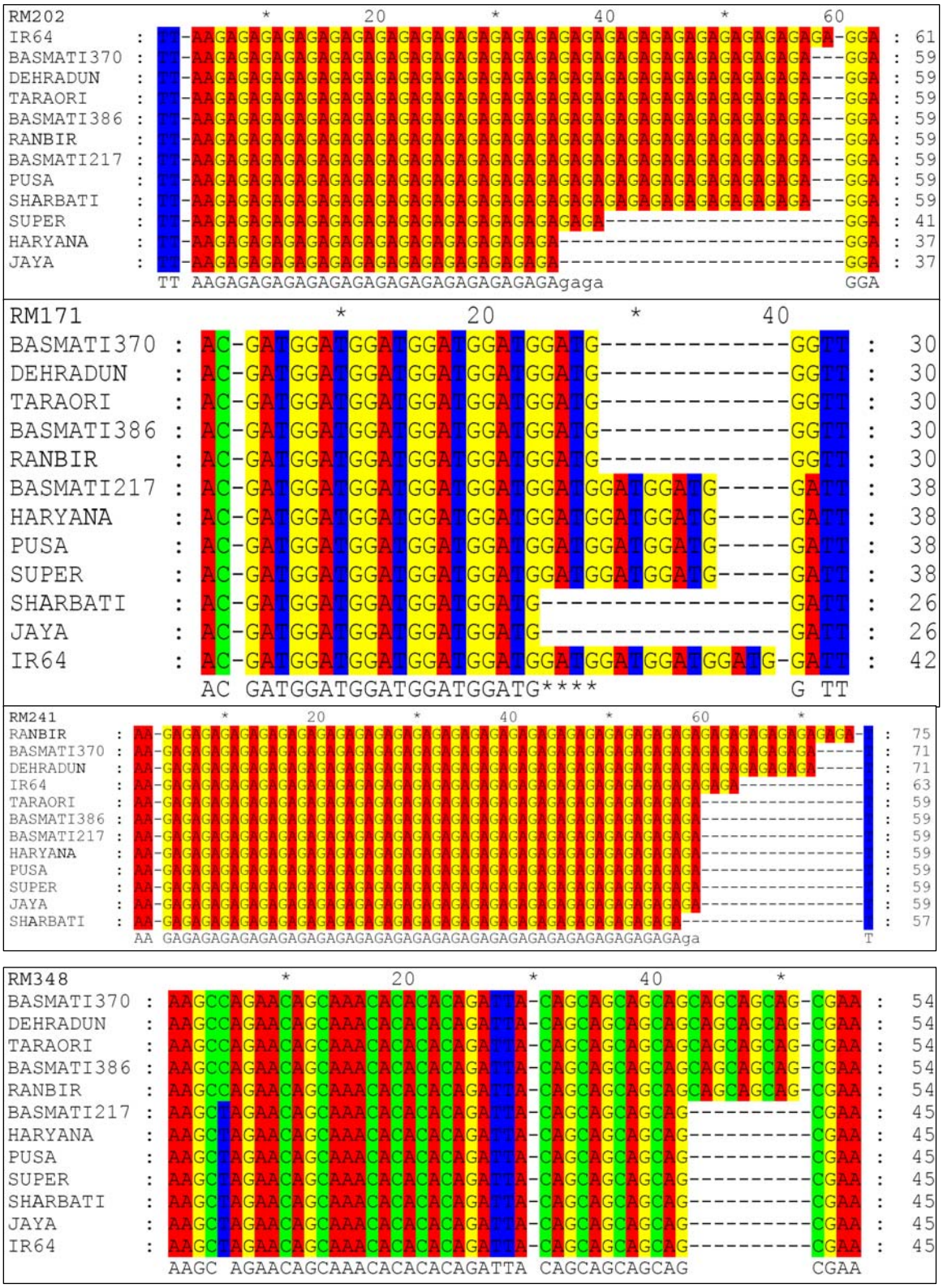

7 Haryana basmati
107.06
157.80
342.73
127.84
161.38
103.18
130.48
229.72


8 Pusa basmati
72.39
157.63
342.64
127.81
181.97
112.78
130.47
229.40
10 Sharbati
105.98
157.57
321.92
127.81
161.18
103.13
130.47
229.38

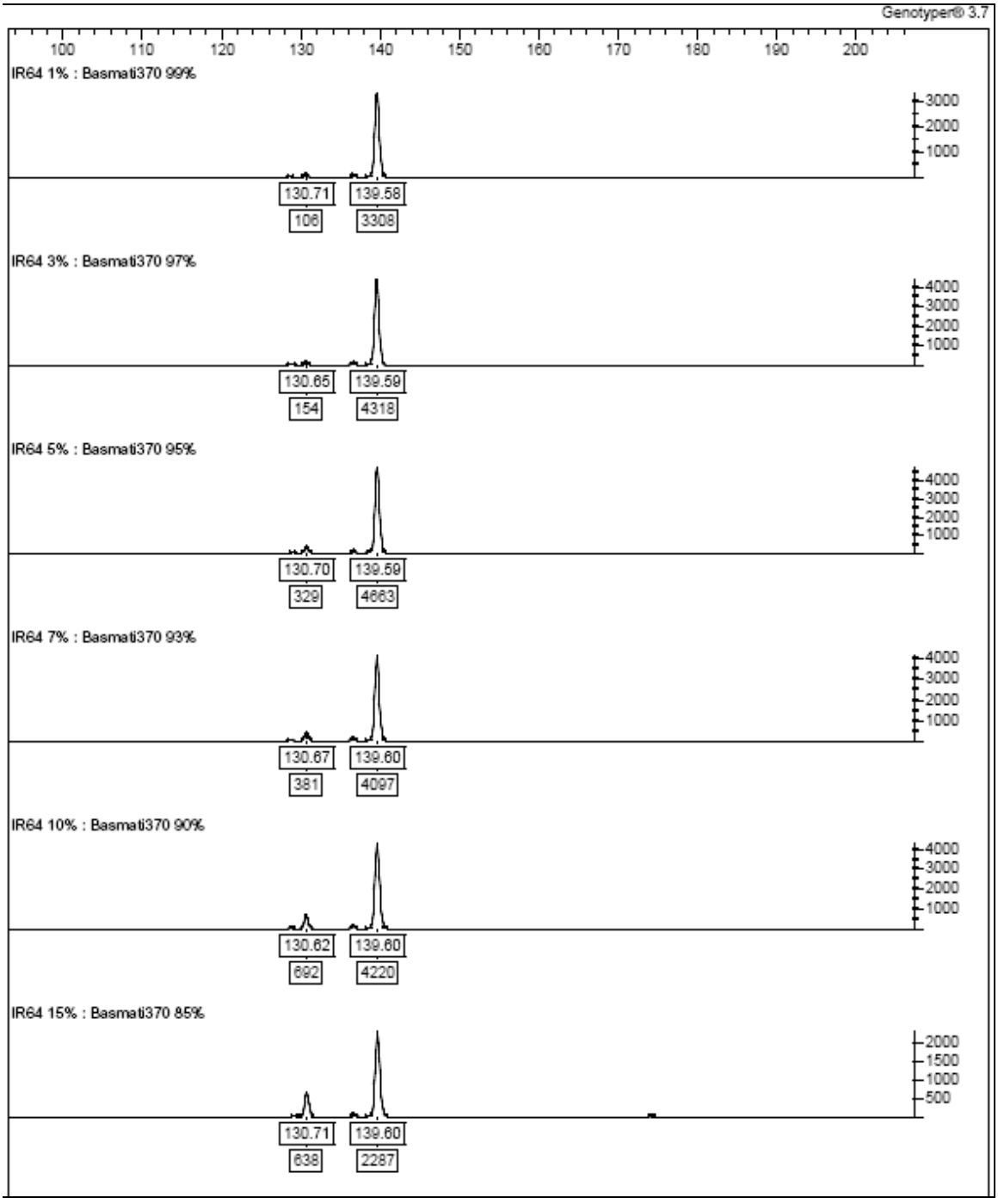

11 Jaya
106.08
157.69
322.01
127.91
162.37
103.22
130.49
236.72
12 IR64
105.07
163.83
346.44
131.98
185.85
103.20
130.53
229.63

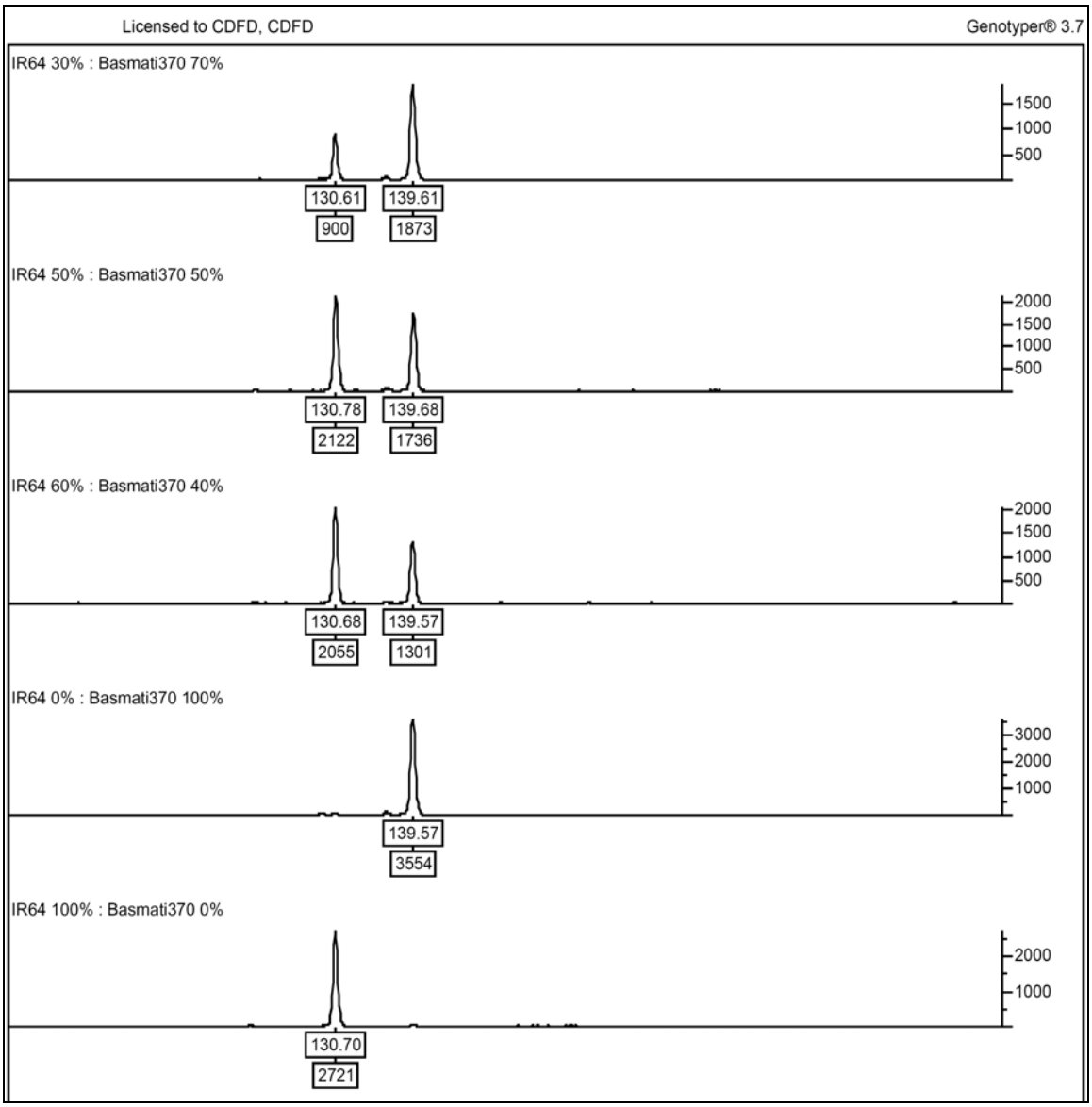
SI Figure 2.2 Aligned sequences of variety specific alleles of 8 microsatellite loci used in the multiplex assay




RM202
IR64 : 61
BASMATI370 : 59
DEHRADUN : 59
TARAORI : 59
BASMATI386 : 59
RANBIR : 59
BASMATI217 : 59
PUSA : 59
SHARBATI : 59
SUPER : 41
HARYANA : 37
JAYA : 37
TT AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAgaga GGA
RM171
BASMATI370 : 30
DEHRADUN : 30
TARAORI : 30
BASMATI386 : 30
RANBIR : 30
BASMATI217 : 38
HARYANA : 38
PUSA : 38
SUPER : 38
SHARBATI : 26
JAYA : 26
IR64 : 42
AC GATGGATGGATGGATGGATG**** G TT
RM241
RANBIR : 75
BASMATI370 : 71
DEHRADUN : 71
IR64 : 63
TARAORI : 59
BASMATI386 : 59
BASMATI217 : 59
HARYANA : 59
PUSA : 59
SUPER : 59
JAYA : 59
SHARBATI : 57
AA GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAga T
RM348
BASMATI370 : 54
DEHRADUN : 54
TARAORI : 54
BASMATI386 : 54
RANBIR : 54
BASMATI217 : 45
HARYANA : 45
PUSA : 45
SUPER : 45
SHARBATI : 45
JAYA : 45
IR64 : 45
AAGC AGAACAGCAAACACACACAGATTA CAGCAGCAGCAG CGAA

SI Figure 2.3 Quantification of Basmati370+IR64 standard samples at RM348

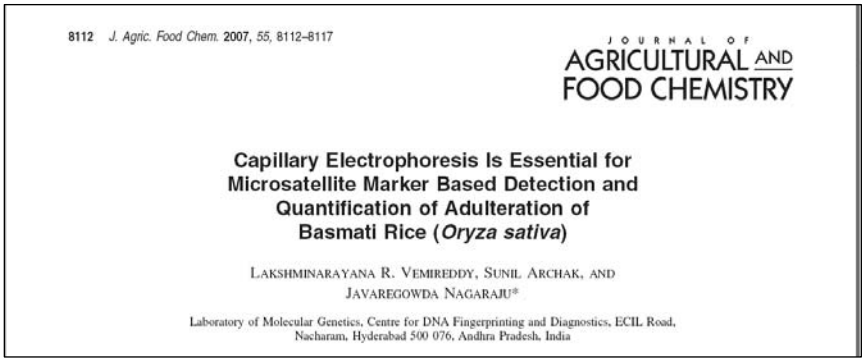



Licensed to CDFD, CDFD
Genotyper® 3.7
IR64 30% : Basmati370 70%
1500
1000
500
130.61
900
139.61
1873
IR64 50% : Basmati370 50%
2000
1500
1000
500
130.78
2122
139.68
1736
IR64 60% : Basmati370 40%
2000
1500
1000
500
130.68
2055
139.57
1301
IR64 0% : Basmati370 100%
3000
2000
1000
139.57
3554
IR64 100% : Basmati370 0%
2000
1000
130.70
2721

Publications

Publications

1. **Lakshminarayana R. Vemireddy,** Archak S., and Nagaraju J (2007) Capillary electrophoresis is essential for microsatellite marker based detection and quantification of adulteration of Basmati rice (*Oryza sativa*) ***J. Agric. Food Chem.***, 55, 8112–8117 (Pubmed)



8112 J. Agric. Food Chem. 2007, 55, 8112–8117

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Capillary Electrophoresis Is Essential for Microsatellite Marker Based Detection and Quantification of Adulteration of Basmati Rice (*Oryza sativa*)

LAKSHMINARAYANA R. VEMIREDDY, SUNIL ARCHAK, AND JAVAREGOWDA NAGARAJU*

Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics, ECIL Road, Nacharam, Hyderabad 500 076, Andhra Pradesh, India

2. Archak S., **Lakshminarayanareddy V.**, and Nagaraju J. (2007) High-throughput multiplex microsatellite marker assay for detection and quantification of adulteration in Basmati rice (*Oryza sativa*) ***Electrophoresis,*** 28, 2396–2405.(Pubmed)

2396

Electrophoresis 2007, *28,* 2396–2405

Sunil Archak
V. Lakshminarayanareddy
Javaregowda Nagaraju

Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics, Nacharam, Hyderabad, India

Research Article

High-throughput multiplex microsatellite marker assay for detection and quantification of adulteration in Basmati rice (*Oryza sativa*)

List of Abbreviations

List of Abbreviations

AFLP	Amplified Fragment Length Polymorphism
APEDA	Agricultural and Processed Food Products Export Development Authority
bad-2	Betaine-aldehyde dehydrogenase-2
bp	Base pairs
cM	Centi morgan
CE	Capillary electrophoresis
CTAB	Hexadecyl-trimethyl-ammonium bromide
DH	Double haploid
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
EB	Evolved Basmati
EDTA	Ethyl diaminetetra acetic acid
EST	Expressed sequence tag
FAM	Carboxyfluorescein
FISSR	Fluorescent labelled inter-simple sequence repeat
g	Gram
GI	Geographical Indication
GM	Genetically modified
HCl	Hydro chloric acid
hr	Hour
ISSR	Inter-simple sequence repeat
JOE	6-carboxy-4'-5'-dichloro-2'-7'-dimethoxyfluorescein
kb	kilo base pairs
LOD	Likelihood of Odds Ratio/ Limit of detection
LOQ	Limit of quantification
M	Molar
mM	Milli molar
MAS	Marker assisted selection
min	Minute

mg	Milli gram
mL	Milli litre
mM	Milli molar
MT	Metric tones
NB	Non-Basmati
PCR	Polymerase chain reaction
PVE	Phenotypic Variance Explained by each QTL
QTL	Quantitative Trait Loci
RAPD	Random Amplification of Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RM	Rice microsatellites
rpm	revolutions per minute
RT PCR	Real Time PCR
SD	Standard deviation
SSR	Simple Sequence Repeat
TAMRA	Carboxytetramethyl rhodamine
Taq	Thermostable DNA polymerase from *Thermus aquaticus*
TB	Traditional Basmati
TE	Tris-EDTA
UV	Ultra violet
w/v (weight/volume)	Percent concentration of a solute in a solution, by weight per unit volume. For example, 20% w/v = 20 g of solute in 100ml of solution.
^{o}C	Degree centigrade
µL	Micro litre
µM	Micro molar
%	Per cent

Filename: Chapter 1 & 2 .doc
Directory: D:\vln D drive\vln documents\Ph.D\Thesis\Thesis final\Thesis aln
Template: C:\Documents and Settings\reddy\Application Data\Microsoft\Templates\Normal.dotm
Title:
Introduction
Subject:
Author: lmg
Keywords:
Comments:
Creation Date: 6/11/2008 9:20:00 AM
Change Number: 543
Last Saved On: 6/13/2008 6:57:00 PM
Last Saved By: reddy
Total Editing Time: 1,355 Minutes
Last Printed On: 6/13/2008 6:59:00 PM
As of Last Complete Printing
Number of Pages: 159
Number of Words: 38,632 (approx.)
Number of Characters: 220,204 (approx.)