DIVERSITY ANALYSIS OF USDA MINI-CORE PANEL OF RICE GERMPLASM USING BIOCHEMICAL AND *WAXY* GENE LINKED DNA MARKERS

HAFSA IFTIKHAR [1, 2], UMER MAQSOOD [1], RIDA TARIQ [1] AND MUHAMMAD ARIF [1*]

[1] National Institute for Biotechnology and Genetic Engineering (NIBGE), Jhang Road, Faislabad, Pakistan
[2] Department of Biology, SBSSE, Lahore University of Management Sciences (LUMS), Lahore, Pakistan

[*] Corresponding author's Email: marif_nibge@yahoo.com

Abstract

Eating and cooking quality of rice is mainly determined by its amylose (major starch component) content. Waxy gene, located on chromosome 6 is involved in starch biosynthesis in rice endosperm. It is very important to know the amylose content of rice grain accurately because it is one of the major determinants of rice quality. It helps to classify rice varieties and its accurate evaluation is a real challenge in development of rice varieties with desirable cooking quality. USDA mini-core panel of rice germplasm is a rich source of genetically diverse rice accessions, which can serve as a good starting point in development of rice varieties with desirable eating and cooking quality. Current study aimed to characterize amylose content of USDA mini-core collection of rice germplasm by *waxy* gene linked molecular markers and their validation. Biochemical analysis was performed through iodine binding method to determine amylose content of rice endosperm while genotypic analysis was performed by using microsatellite (SSR) and single nucleotide polymorphism (SNP) based molecular markers. Statistical tools (DARwin and PowerMarker) were used to check the consistency of results. Results revealed that mini-core is diverse collection of rice germplasm with respect to amylose content and based on these findings accessions were classified into three groups (low, intermediate and high amylose containing rice) while most of accessions fall in intermediate amylose containing category. According to PIC value (from PowerMarker), RM190 was found to be highly informative marker for screening of mini-core panel. This study will be important in selecting and utilizing intermediate amylose containing genotypes for future breeding programs.

Key Words: Amylose; Biochemical methods; Eating and cooking quality; Molecular markers; SSR; SNP; USDA mini-core; Waxy gene.

Introduction

Agriculture has been playing a vital role in development of human population since very beginning of human civilization. According to an estimate human population will exceed the gigantic figure of 10 billion by 2050 (Tilman *et al*., 2002). Large amount of food will be required to feed this huge population and in near future food shortage is going to be a serious threat to humanity. In the light of this reality plant scientists are striving for the improvement of yield of some major staple food crops including wheat, rice, barley maize etc.(Curtis *et al*., 2002, Mekonnon, 2014, Khan *et al*., 2015, Sarvari and Pepo, 2014). Yield of these crops can be increased by developing the varieties having not only the potential to tolerate biotic and abiotic stress but also have desirable grain quality and nutritious attributes.

Rice is a monocotyledonous plant, a member of family *Gramineae* and one of the widely consumed crops in the world. It is cultivated in many tropical and sub-tropical countries that cover approximately 11% of total cultivated area of Earth, under diverse eco-geographic conditions (Causse *et al*., 1994, Khush, 2005). Out of 25 species of genus *Oryza* only two are cultivated ones known as *O. sativa* and *O. glaberrima* while rest of others are wild. *O. sativa* is mostly grown in Asia, Europe, Africa, South and North America while *O.glaberrima* is grown mostly in West Africa (Linares, 2002). Rapid digestive nature of rice starch as compared to other starchy foods makes it distinctive among other cereals (Kent, 1978, Hu *et al*., 2014, Feng *et al.*, 2017).

Although rice genome is smaller in size as compared to other cereal crops but is highly diverse

in nature and as a matter of fact, in any crop improvement program the main source of variability is genetic diversity (Chang, 1976). Natural genetic diversity, which has developed as a result of survival extinct of plant in different agro-climatic conditions, has been exploited from the very beginning of agriculture within the crop species to get the desirable crop traits. USDA rice mini-core collection can serve as the diverse gene pool for such crop improvement practices. The USDA rice mini-core (URMC) subset is representative of 18,000 conserved accessions in the USDA global gene-bank of rice and was established from core collection that was comprised of 1,794 accessions (Yan *et al*., 2007). This URMC collection consists of 217 accessions which are originated from 76 countries covering about 15 geographical regions. These facts about URMC collection reveal that it can serve as reservoir of many valuable genes including not only biotic and abiotic stress (pests, disease, salinity and drought) resistance genes but also the genes for improving the cooking, processing and nutritional quality of rice. Increasing knowledge of the rice genome has accelerated the process of gene discovery that will facilitate the transfer of genes of breeder's interest in an elite genetic background through molecular assisted selection and other such approaches (Garris *et al*., 2005, Li *et al*., 2017).

Traditionally focus of plant breeders was on breeding to increase the pest resistance and yield of crop but recently trend has shifted towards the incorporation of preferred quality traits to increase the overall economic value of rice. Various genetic and environmental factors are involved in determining the quality of rice (Giri and Laxmi, 2000) while Eating and Cooking Quality (ECQ) of rice is mostly dependent upon the starch content of grain. Starch is composed of two major polysaccharides, amylose and amylopectin. Of these two amylose content (AC) plays more important role in classifying the rice varieties. *Waxy* gene which encodes granule bound starch synthase (GBSS) is located on short arm of chromosome 6 in rice genome is major contributor in controlling the amylose content in rice grains (Umemoto et al., 2002). However, it was reported by Isshiki *et al.* (2000) that some other genes also play roles in controlling amylose biosynthesis.

It is very important to know the amylose content of rice grain accurately because it is one of the major determinants of rice quality (Okpala et al., 2020). Different methods have been reported for the determination of amylose content such as iodine binding, size exclusion chromatography, near infrared spectroscopy and asymmetrical flow field flow fractionation (Juliano, 1971b, Wesley *et al*., 2003, Ward *et al*., 2006, Chiaramonte *et al*., 2012b). There are some limitations associated with these methods which are in need to be overcome (Caffagni *et al*., 2013). Molecular markers can be used to overcome the limitations associated with aforementioned methods. Variety of molecular markers has been reported including Simple Sequence Repeats (SSRs) and Single Nucleotide Polymorphism (SNP) based markers. The information obtained by using these molecular markers is of great value for the scientists to incorporate desirable genes through Marker Assisted Breeding schemes.

Biochemical and molecular marker based characterization of diverse mini-core panel of rice germplasm to identify the intermediate amylose content containing accessions was the main objective of current research. We also estimated the capability of already reported SSR and SNP based markers in screening this diverse collection of rice germplasm on the basis of amylose content. The rice lines with desirable grain quality attributes and agronomic characters will be used in future breeding program for improvement of rice quality. On a broader aspect, it is important to improve the rice quality because it will help in expanding the export market of Pakistan globally.

Materials and Methods

Plant material

217 accessions of USDA rice mini-core panel were requested and obtained from USDA in 2014. The seeds were multiplied in kharief seasons of 2015 and 2016. Some accessions showed poor grain filling and low seed setting due to different climatic conditions in Pakistan from the countries of which these lines were native. Out of 217 lines 207 produced sufficient grains to be used in current study.

Biochemical Analysis of Amylose Content by Colorimetric Method

Amylose content was determined by standard Iodine binding method described earlier by Juliano (1971) and Perez & Juliano (1978).

To describe precisely standard solutions were prepared by using 0.04 g of potato amylose (Sigma-Aldrich CAS No. 9005-82-7). In weighed potato amylose 9 mL of 1 M NaOH and 1 mL of 95% ethanol were added. The mixture was shaken gently and placed into water bath already set at 65°C. The mixture was heated for 10 minutes to gelatinize the potato amylose. After that contents were allowed to cool for one hour at room temperature. Distilled water was added to make the volume up to 100 mL and flask was shaken well. Amylose solution volume 1, 2, 3, 4, 5 mL were taken into different flasks. 2 mL of iodine solution (2g potassium iodide & 0.2 g iodine in 100 mL distilled water) and 0.2, 0.4, 0.6, 0.8 and 1 mL of 1 M acetic acid solution were added to each flask respectively. Distilled water was added to make the volume up to 100 mL and left for 20 minutes. Afterwards absorbance of standards was used to plot standard curve. The absorbance values at 620 nm (using Bio-Rad SmartSpec. 3000 UV/Vis Spectrophotometer) were plotted against the concentration of potato starch by using Microsoft Excel (2010) and the conversion factor was determined. The dilution factor value 20 was included in the conversion factor. After defining conversion factor and by using absorbance values percentage amylose content of samples was calculated by using equation 1 after performing the colorimetric analysis and taking absorbance of samples (de-husked, polished and ground rice seeds).

Amylose Content (%) = (y0.0547x+ 0.0439) × Absorbance at 620 nm… (1)

Genomic DNA Extraction

Seeds from all 207 mini-core accessions were soaked in water (in petri plates) to germinate in growth chamber at 28°C. Young green leaves from 14 days old seedlings were collected for isolation of genomic DNA using a CTAB (Cetyl Trimethyl Ammonium Bromide) DNA extraction method as described by Sika *et al.*, (2015)

Quality of DNA was checked by agarose gel electrophoresis, using 0.8% agarose (Sigma-Aldrich CAS No. 9012-36-6) in 0.5X TAE buffer. Gel was visualized and image was captured on gel documentation system. DNA concentration was estimated by observing the intensity of DNA bands while the exact DNA concentration was determined by using NanoDrop spectrophotometer (ThermoScientific NanoDrop, 2000). Samples were stored at -20°C for further use.

Selection of Primers

In this study *Waxy* (*Wx*) gene associated simple sequence repeats (SSR) and single nucleotide polymorphism (SNP) based markers were selected from literature on the basis of their polymorphism and linkage with *waxy* and *GBSS1* gene as these have major role in determining amylose content in grain. Those primer pairs and related information is listed in the Table 1.

Table 1. SSR and SNP markers used for molecular marker analysis of amylose content

Sr. #	Marker type	Primer	Ch. No.	Primer Sequence 5'-3'	Source
1	SSR	RM25	8	F-GTAATCGATGCTGTGGGAAG- GAGTCATGTGATAGCCGATATG	(Jain *et al.*, 2006)
2		RM190	6	F-TTTGTCTATCTCAAGACAC R-TTGCAGATGTTCTTCCTGATG	(Kottearachchi *et al.*, 2014)

3		RM339	8	F-GTACGACTACGAGTGTCACCAA R-GTCTTCGCGATCACTCGC	(Temnykh *et al*., 2000)
4		RM440	5	F-CATGCAACAACGTCACCTTC R-ATGGTTGGTAGGCACCAAAG	(Cheng *et al*., 2014)
5		GBSS1	6	F-CAAATAGCCACCCACACCAC R-CTTGCAGATGTTCTTCCTGATG	(Fernando *et al*., 2015)
7		GF TR GR TF	6	F-TACAAATAGCCACCCACA R-GATCAGCCTAACCAAACA R-GGGAAACAAAGAATTATAAACAT ATATGTACAC F-CATCAGGAAGAACATCTGCAAGT	(Cai *et al*., 2015)
8	**SNP**	WxIn1 (ln1-F-G) Common primer GBSS-W2R In1-F-T	6	F-CAGGAAGAACATCTGCACGG R-TTTCCAGCCCAACACCTTAC F-ATCAGGAAGAACATCTGCACGT	(Chen *et al*., 2010)
9		WxEx6 (e6-F-C) Common primer e6-R e6-F-A	6	F-CAACCCATACTTCAAAGGAACATC R-AGTCGTTGCAGACGAACACAAC F-AACAACCCATACTTCAAAGGAACTTA	(Chen et al., 2010)
10		WxEx10 (e10-R-C) Common primer e10-F e10-R-T	6	R-GCGGCCATGACGTCTGG F-TCAGGCAATCGAGGCGAAG R-GGCGGCCATGACGTCAGA	(Chen *et al*., 2010)

PCR Parameters Optimization for Molecular Marker Analysis

PCR reaction mixture and profile was optimized. Conditions on which desired results were obtained are mentioned below:

Single nucleotide polymorphism (SNP) markers and microsatellite (SSR) markers used in this study were optimized by using different concentrations of reagents (green master mix (2X), forward and reverse primers (10 µM), template DNA (15 ng/µL) and deionized water to make reaction mixture up to 18 µL) to get the desired results. Three controls IR6, Super Basmati (SB), IRBB57 (having high, intermediate and low amylose content respectively) were used to validate the results obtained by using different markers.

In present study during optimization of PCR profile for SNP and SSR markers three different annealing temperatures were used to get the desired amplified product by using (BIORAD, iCycler PCR Thermal Cycler). After optimization, the PCR profiles that was used for SNP and SSR markers is shown (Fig. 1 a & b**)**

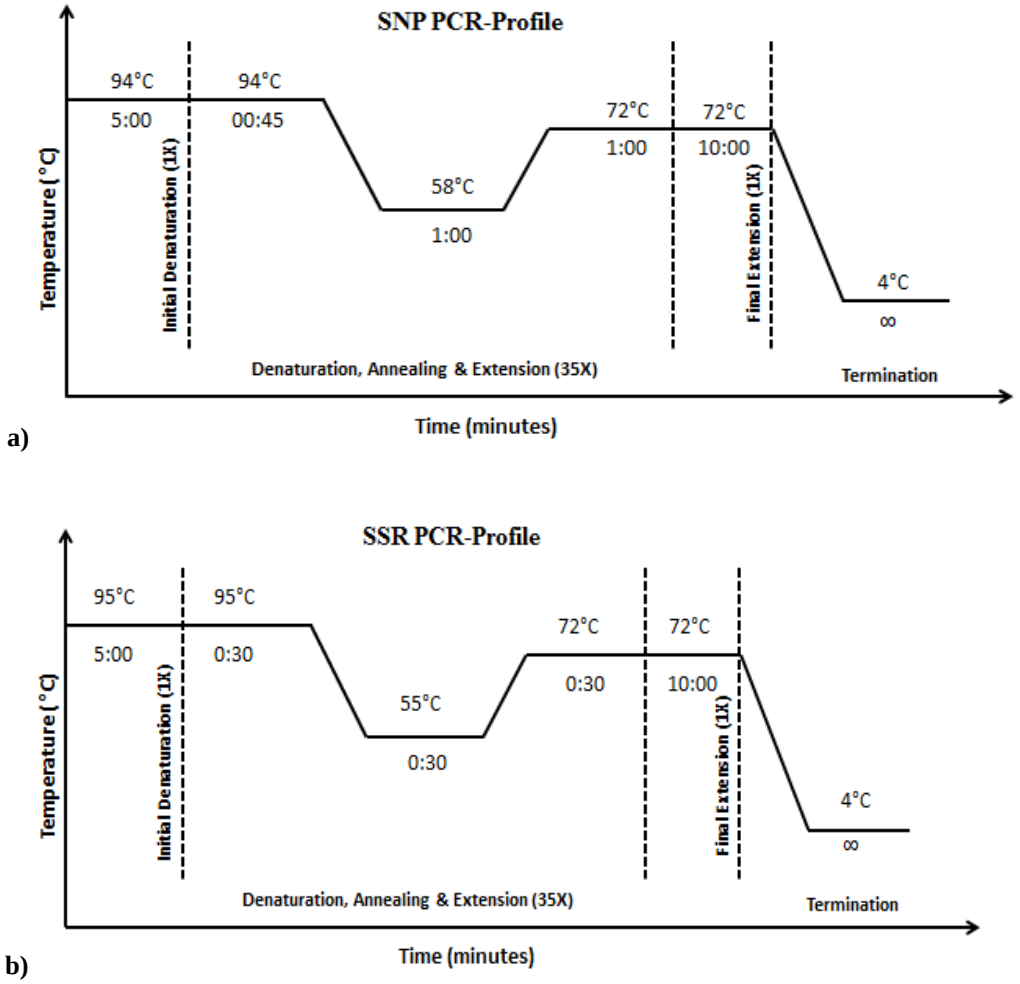


Figure 1. a) PCR profile for SNP markers; b) PCR profile for SSR markers

Above mentioned profile was used for other SNP markers with a few variations in annealing temperature. Annealing temperature used for WxIn1 and WxEx6 was 60°C while that for WxEx10 SNP markers was found to be 62°C. While for all SSR markers same annealing temperature was used.

Molecular Marker Based Analysis of USDA Mini-core

Molecular marker based analysis by PCR amplification of USDA mini-core collection of rice germplasm was carried out by using SNP and SSR markers, revealing the amylose content diversity of this rice collection.

Amplified PCR product (5 µL) was resolved on 2.5% agarose gel in the presence of 0.5X TAE buffer. Gel was visualized and image was captured using Gel Documentation System (High-Performance UV Transilluminator UVP). Fragment size of each sample was compared with a 50 bp DNA ladder (Promega, Cat. No. G4521). Bands were identified on the basis of size and scored visually for further evaluation.

Allele Scoring and Statistical Analysis

Scoring for PowerMarker and DARwin software

PCR amplified fragments were scored to perform statistical analysis through Power Marker (V3.25) and DARwin (V6.0.014) software. For all SSR markers (RM25, RM190, RM339, RM440 and GBSS1) and one SNP based marker against intron one of *waxy* gene, amplified bands were scored on the basis of size. Other SNP markers (WxIn1, WxEx6 and WxEx10) were scored on the basis of presence of certain allele.

Diversity Analysis by PowerMarker

Genetic diversity parameters (such as number of alleles, major allele frequency, polymorphic information content (PIC) value was estimated for a set of 207 mini-core rice accessions by using PowerMarker (V3.25). Allele frequency shows the frequency of particular allele for each of the markers while polymorphic information content (PIC) value refers to a comparative measure of extent of distinguishing ability of a marker among different populations.

Phylogenetic Analysis by DARwin

Genetic diversity analysis of USDA mini-core rice germplasm was conducted on the basis of amylose content by constructing a phylogenetic tree by using DARwin (Dissimilarity Analysis and Representation for Windows) software (V6.0.014). For this purpose, input file comprised of allelic data that was generated after interpretation of gel banding pattern developed by using 9 molecular markers (5SSR and 4SNP). DARwin uses dissimilarity matrix for clustering of 207 mini-core lines on the basis of un-weighted neighbour-joining method. Dissimilarity index was calculated by using simple matching coefficient.

Factorial Analysis by DARwin

DARwin uses a dissimilarity matrix calculated with a simple-matching index to perform factorial analysis. This analysis was performed for 207 mini-core lines by using the scoring data obtained from 9 molecular markers with the help of DARwin (V6.0.014).

Results

Bio-chemical Analysis Based diversity of Mini-core

Findings from biochemical analysis of amylose content that was performed by colorimetric method are discussed in the following sections:

Standard Curve for Estimation of Amylose Content

As discussed in section earlier standards were generated from potato starch (Fig. 2) to prepare standard curve (Fig. 3) by using their absorbance values (Table 2) and concentration. Solutions from S1-S5 contain 5-1 mg/mL potato amylose. Colours from dark blue to lemon yellow are associated with gradual decrease in amylose content while blank solution is without any amylose.

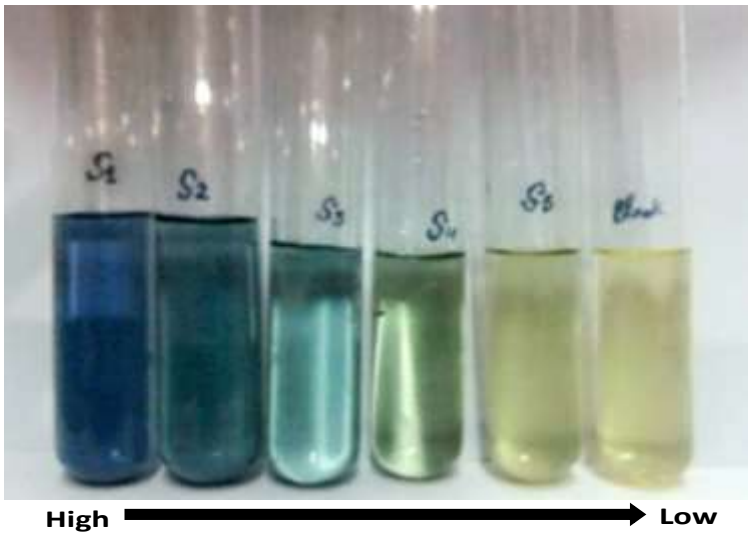


Figure 2. Standard solutions showing gradual decrease in amylose concentration

Table 2. Absorbance Values of Standard Solutions in Triplicate Manner along with their Respective Concentration

Conc. mg/mL	Abs. 1	Abs.2	Abs.3	Average
1	0.1	0.098	0.095	0.098

2	0.161	0.156	0.158	0.158
3	0.205	0.21	0.208	0.208
4	0.255	0.251	0.253	0.253
5	0.326	0.322	0.324	0.324

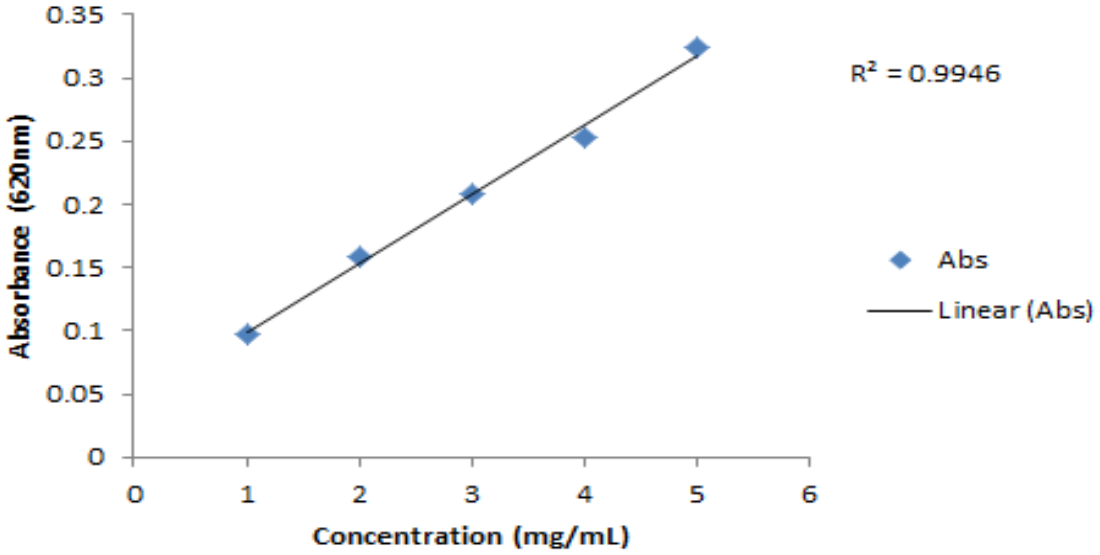


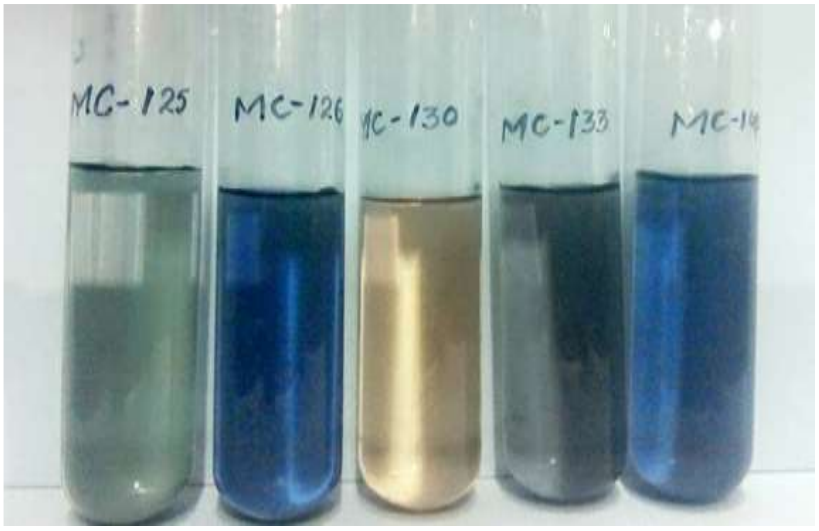
Figure 3. Standard curve representing amylose concentration. Concentration (mg/mL) along horizontal axis (x-axis) while absorbance (620 nm) along vertical axis (y-axis)

Value of linear regression was 0.99 that was obtained from graph. Percentage amylose content of each sample was determined.

Variation in Amylose Content of USDA Mini-core Rice Germplasm

Amylose content of 207 accessions was determined along with varieties with low, intermediate and high amylose content (IRBB57, SB and IR6 respectively) as check (Fig. 4a) as described by Juliano (Juliano, 1971a). Brownish yellow colour was associated with very low amylose, bluish green with intermediate and dark blue was found to be associated with high amylose content (Fig. 4b)

(a)



(b)

Figure 4. (a)Varieties (IR6, IRBB57 & SB) after colour development by iodine binding method (b) Mini-core accessions after colour development. MC-130 representing low amylose content, MC-125 and MC-133 representing intermediate amylose content, MC-130 and MC-140 representing high amylose content

Iodine binding method results revealed that this collection of rice is highly diverse with respect to amylose content and members of all amylose classes are part of it. It was analysed that out of 207, 13.65% accessions were found associated with high, 55.6% with intermediate and 29.75% were found associated with low amylose content.

SNP based Molecular Marker Analysis of Mini-core

For the purpose of genotyping of mini-core samples four SNP based markers were used which were selected after reviewing literature. All of the SNP markers used in this study were located in waxy gene (Intron 1, exon 6 and exon 10) on chromosome number 6.

Two primer sets that were used to detect intron 1 SNPs had two SNP alleles (SNP allele G and T). G SNP was found to be associated with intermediate/high amylose content while T SNP was associated with glutinous/low amylose containing rice varieties. Location of this SNP is shown in Fig. 5. In case of WxIn1 marker, we used only G SNP allele carrying primer that's why band was observed only in those varieties having SNP allele G while T SNP carrying varieties no band was observed (Fig. 6 a). while in case of other intron 1 based marker use of two primer pairs (mentioned in table 2, associated

with one specific SNP G/T) resulted in three different sized products (199 bp, 207 bp and 235 bp) that were found corresponding with the high, intermediate and low amylose content respectively (Fig. 6 b).

WxEx6 SNP marker is associated with A/C substitution that results in codon change and ultimately change in amino acid from serine to tyrosine. SNP allele A was associated with high and low while C was associated with glutinous and intermediate amylose containing varieties (Fig. 8). Location of this SNP is shown in Fig. 7. Primer set that was associated with SNP allele 'C' (Table 1) was used and amplification was observed only in those samples carrying that certain SNP.

WxEx10 SNP marker was found associated with C to T substitution. SNP allele C in case of this marker was associated with high while T was associated with glutinous, low and intermediate amylose containing rice varieties (Fig. 10). Location of this SNP is represented in Fig. 9. SNP allele C carrying primer set (Table 1) was used and amplification was observed in only those samples having that SNP.

Genotyping by these SNP markers helped in identification that which SNP allele is present in respective sample and basically facilitated in finding the class of amylose of mini-core samples. Detailed results of mini-core classification revealed by these markers are listed in Table 3 and Table 4.

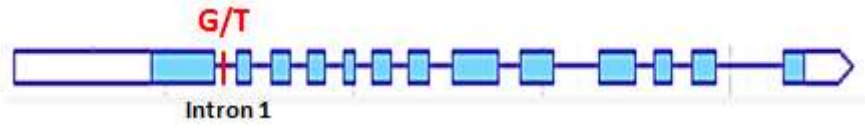


Figure 5. Representation of location of SNP in intron 1 of *waxy* gene

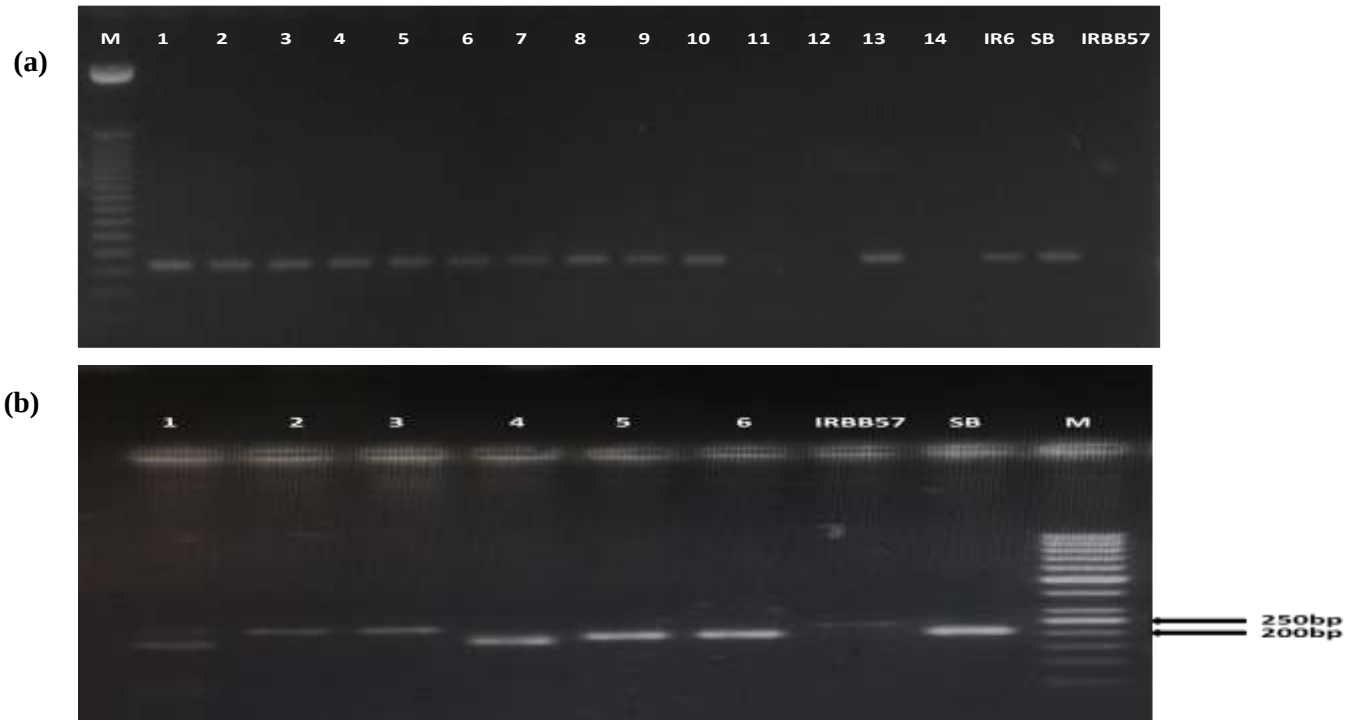


Figure 6. Agarose gel (2.5%) showing amplification of genomic DNA of samples having SNP in intron 1 of *waxy* gene.
a) Amplification with primers only with SNP allele G; b) M: 50 bp DNA ladder; (a) (Lane 1-14 representing mini-core samples from120 to133; last three lanes representing controls; (b) Lane 1-6 representing mini-core samples from1 to 6; IRBB57 and SB representing controls; (IR6: high, SB: intermediate, IRBB57: low amylose)

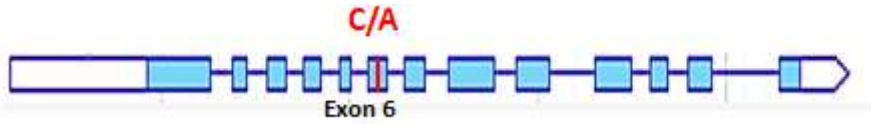


Figure7. Representation of location of SNP in exon 6 of *Waxy* gene

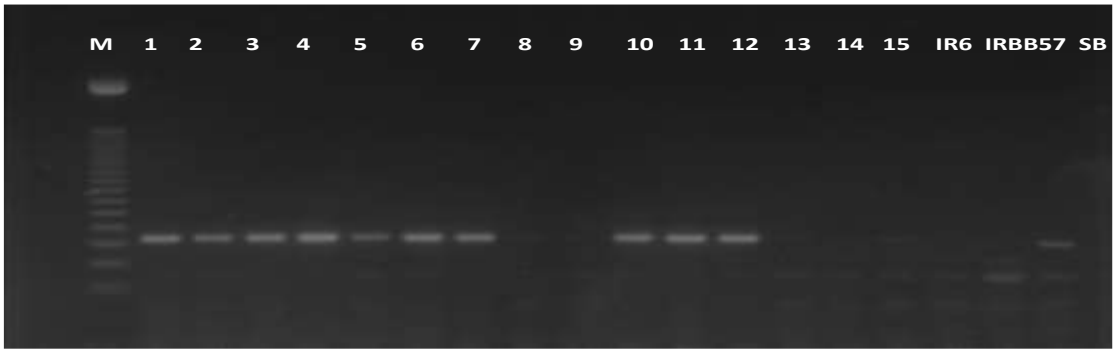


Figure 8. Agarose gel (2.5%) showing amplification of genomic DNA of samples having SNP allele C
M: 50 bp DNA ladder; Lane 1-15 representing mini-core samples from 33 to 48; last three lanes representing controls (IR6: high, SB: intermediate, IRBB57: low amylose)

Figure 9. Representation of location of SNP in exon 10 of *Waxy* gene

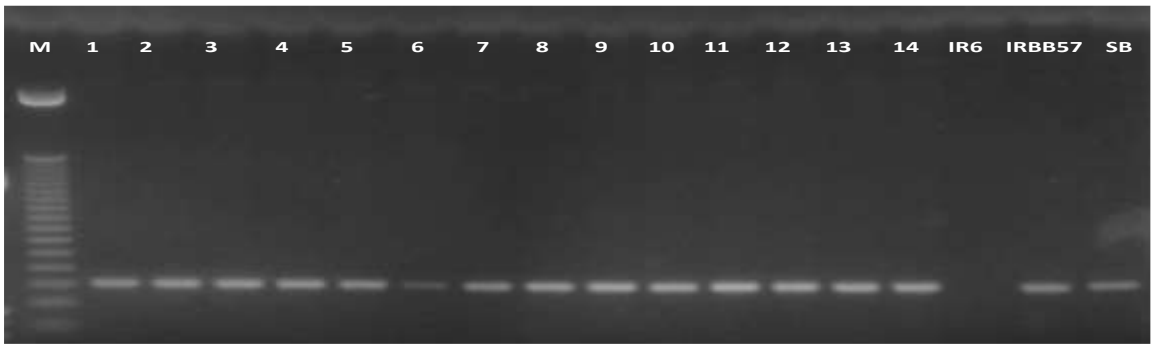


Figure 10. Agarose gel (2.5%) showing amplification of genomic DNA of samples having SNP allele C
M: 50 bp DNA ladder; Lane 1-14 representing mini-core samples from 122 to 135; last three lanes representing controls (IR6: high, SB: intermediate, IRBB57: low amylose)

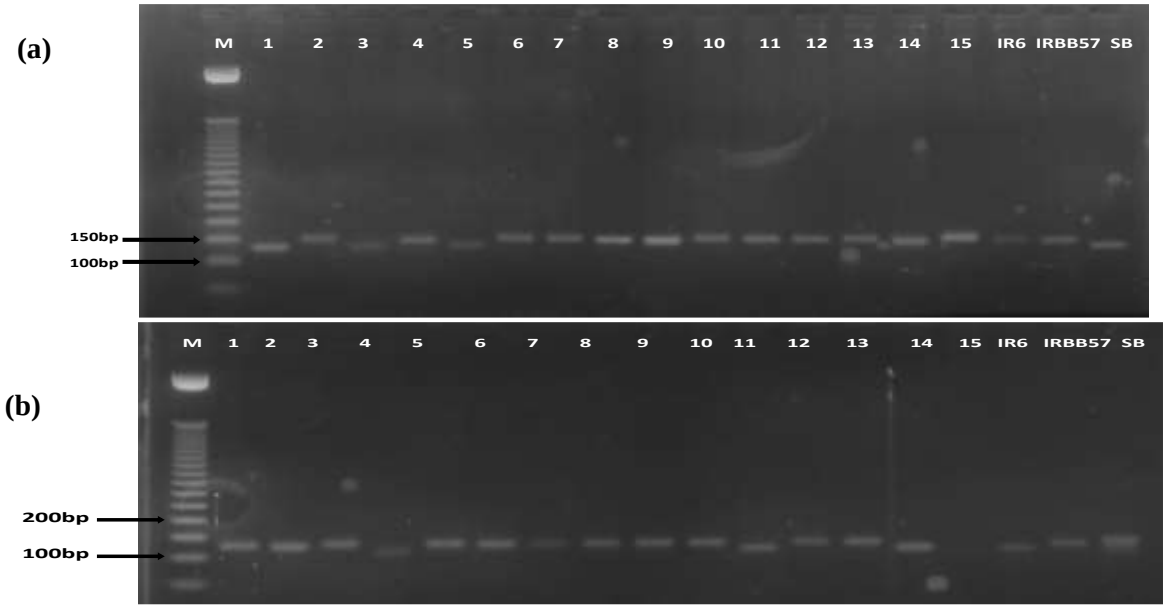
Table 3. Classification of Mini-core Samples on the Basis of Amylose Content by using WxIn1, WxEx6 and WxEx10 SNP Based Markers

Marker name	Allele type	Interpretation	No. of Mini-core samples
WxIn1	G	High/intermediate	176

WxEx6	T	Low/glutinous	31
	A	High/low	56
	C	Intermediate/glutinous	151
WxEx10	T	High	6
	C	Intermediate/low/glutinous	201

SSR based Molecular Marker Analysis of Mini-core

Genotyping of mini-core samples was also performed by 5 SSR markers (RM25, RM190, RM339, RM440 and GBSS1). Some of these markers were located on chromosome number 6 while some on chromosome 5 and 8 in rice genome. Amplification with RM25 result in two different sized bands. Smaller sized band was associated with intermediate amylose content while larger sized band was with high and low amylose content (Fig. 11a). RM190 that is a well-known SSR marker used in various studies to evaluate amylose content showed three different band sizes including 100 bp, 150 bp and 200 bp that were found associated with high, low and intermediate amylose content respectively (Fig. 11b). RM339 is also a polymorphic marker for amylose content and helped in separating the intermediate amylose containing genotypes from the rest of others (Fig. 11c). RM440 played important role in this diversity analysis and showed three product sizes. Amplification product of 170 bp was corresponding with intermediate while 220 bp and 250 bp were with low and high amylose content respectively (Fig. 11d). GBSS1 is one of the important candidates for determination of amylose content because granule bound starch synthase enzyme is considered to be an important enzyme in starch biosynthesis pathway to control the amylose content in rice grain. Three different sized amplification product were obtained. Out of those, smallest and largest sized bands were associated with high and intermediate amylose content respectively while the band that is of size in between these two bands was found associated with low amylose content (Fig. 11e). Detailed results of mini-core classification by using these markers are listed in Table 4.



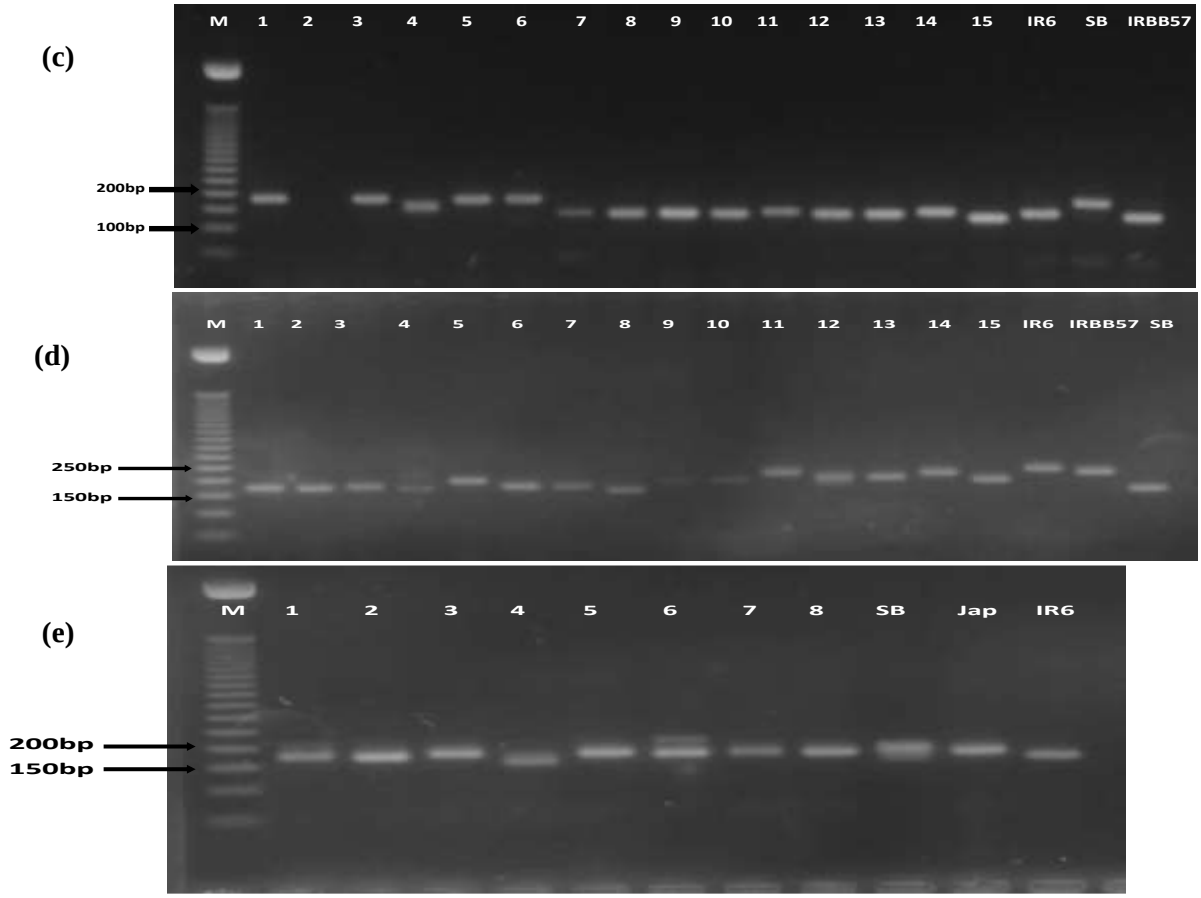


Figure 11. Agarose gel (2.5%) showing amplification of genomic DNA of samples by using SSR markers RM25, RM190, RM339, RM440 and GBSS1. M: 50 bp DNA ladder; (a) Lane 1-15 representing amplification in mini-core samples from 66 to 80 by using RM25 (b) Lane 1-15 representing amplification in mini-core samples from 1 to 16 by using RM190; (c) Lane 1-15 representing amplification in mini-core samples from 1 to 16 by using RM339; (d) Lane 1-15 representing amplification in mini-core samples from 1 to 16 by using RM440; (e) Lane 1-8 representing amplification in mini-core samples from 1 to 9 by using GBSS1 marker; last three lanes representing controls (IR6: high, SB: intermediate, Japonica/ IRBB57: low amylose)

Table 4. Classification of Mini-core Samples on the Basis of Amylose Content by using Wx (G/T) In1, RM25, RM190, RM339, RM440 and GBSS1 Markers

Marker Type	Marker Name	Product Size Range	Amylose Classes	No. of Mini-core samples
SNP	Wx(G/T)In1 based	190-235 bp	High	51
			Intermediate	103
			Low	47
SSR	RM25	125-146 bp	High & low	138
			Intermediate	67
	RM190	100-200 bp	High	70
			Intermediate	69

		Low	63
		High & low	155
RM339	126-189 bp	Intermediate	43
		High	71
RM440	180-250 bp	Intermediate	62
		Low	68
		High	61
GBSS1	150-200 bp	Intermediate	81
		Low	62

3.1 Diversity Analysis

Diversity analysis of 207 mini-core accessions by PowerMarker V3.25 provided the information regarding PIC values (Polymorphism Information Content), Number of Alleles and Major Allele Frequency (Table 5).

This analysis was performed by using all the markers used in this study (4 SNP and 5 SSR markers) which produced a total of 22 alleles. The number of alleles per loci ranged from 2 to 3 while on average 2.44 alleles per locus was observed. Some of the markers, as shown in table 5 were showing two alleles while all other markers including Wx G-T (intron1), RM190, RM440 and GBSS1 were showing 3 alleles. The major allele frequency variation was observed from lowest 0.3465 to the highest value of 0.9662 with an average of 0.6235.

Polymorphic information content (PIC) value is the representative of relative information provided by each marker and in present study, the average PIC value was found to be 0.3941. The highest genetic diversity was explained by RM190 with a PIC value of 0.5919, which shows that this marker was found to be the most informative one. PIC value ranged between 0.0632 for WxEx10 to 0.5919 for RM190.

Table 5. Diversity Analysis Information Generated by PowerMarker V3.25

Marker	Major Allele Frequency	Allele No.	PIC
WxIn1	0.8502	2.0000	0.2222
WxEx6	0.7295	2.0000	0.3168
WxEx10 modified	0.9662	2.0000	0.0632
SNP F(GT) R(GT)	0.5124	3.0000	0.5488
RM339	0.7828	2.0000	0.2822
RM440	0.3532	3.0000	0.5915
RM25	0.6732	2.0000	0.3432

RM190	0.3465	3.0000	0.5919
GBSS1	0.3971	3.0000	0.5867

Phylogenetic Analysis

UPGMA (Un-weighted Pair Group Method with Arithmetic mean) neighbour-joining phylogenetic tree that was generated by DARwin (Dissimilarity Analysis and Representation for Windows) software (V6.0.014) showed the diversity in 207 mini-core rice lines on the basis of amylose content. These 207 accessions are divided into three major clusters (Fig. 12). Cluster-1 comprised mainly of those lines (labelled as red) which are high in amylose content including CAROLINO-164, Karayal, Santhi-990 and other high amylose containing accession. It is the smallest of clusters containing approximately less than 25% mini-core lines. More than 50% of mini-core rice accessions are grouped in the cluster-2 and major portion of this cluster is comprised of intermediate amylose containing lines (shown in blue colour) including Dular, Simpor, Djimoron and many more. Third cluster is comprised of those lines (shown in green) containing low amylose content including Pan Ju, SHIMIZU MOCHI, Ragasu and other such accessions.

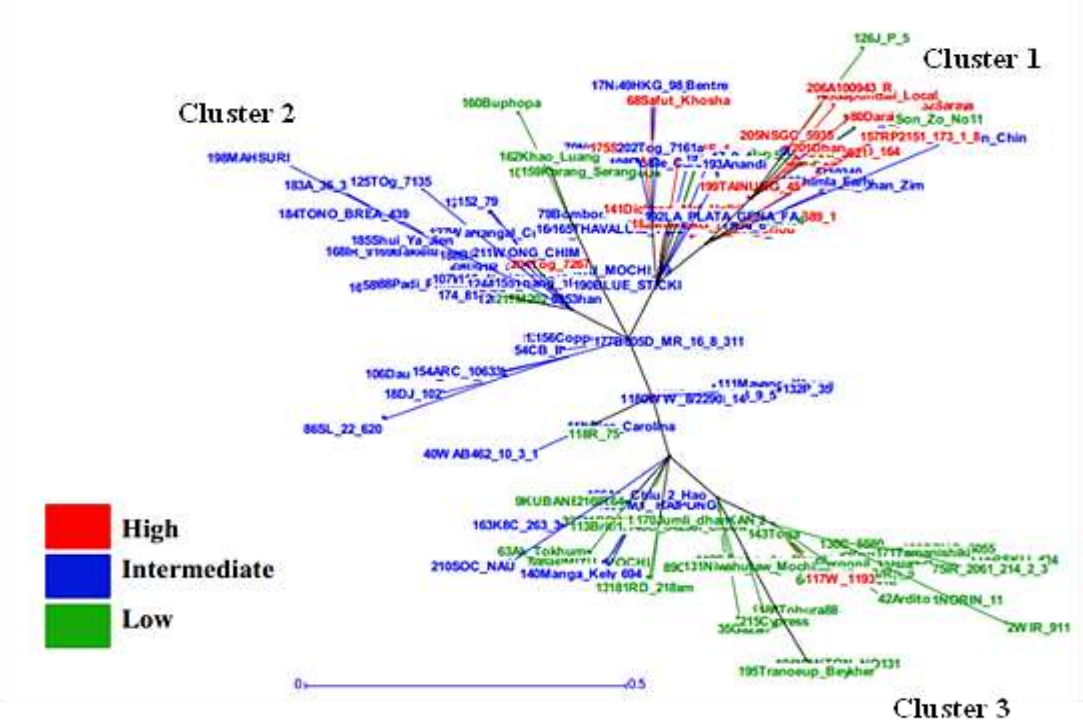


Figure 12. Phylogenetic tree based on the Neighbour-Joining method.
Tree representing genetic dissimilarity between rice accessions, based on SNP and SSR markers

Factorial Analysis

Factorial analysis was performed for 207 mini-core rice lines based on 9 markers (4 SNP and 5 SSR markers) as a result two dimensional scatter plot was obtained that showed separation of these lines into three groups on the basis of amylose content (Fig. 13). This distribution of accessions was found in agreement with that observed by phylogenetic analysis. Sections marked by red, blue and green coloured names representing high, intermediate and low amylose containing accessions respectively.

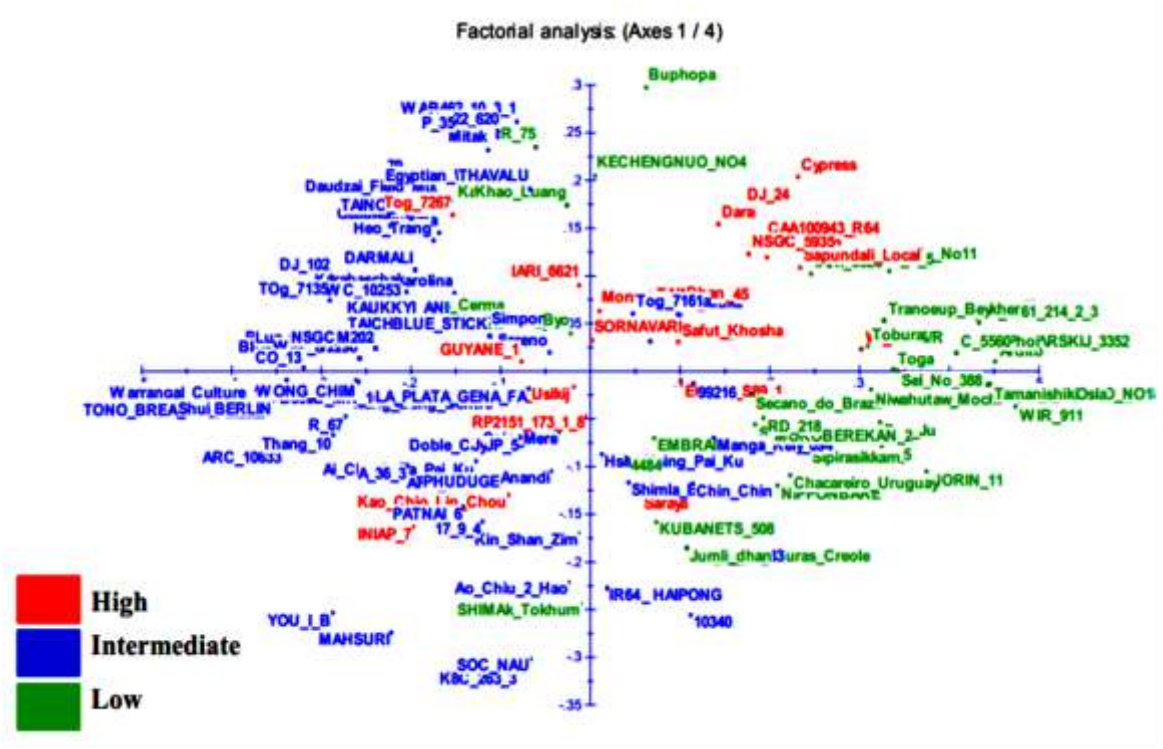


Figure 13. Factorial analysis by DARwin software
Analysis based on two-dimensional plots of mini-core rice accessions. Accessions representatives of high (red), intermediate (blue) and low (green) amylose containing groups

Discussion:

Rice amylose content is one of the important agronomic traits that affect not only the nutritional but also the processing quality of rice grains (Juliano, 1971a). Therefore, it is a major goal of agriculture research to develop an understanding about the mechanism behind variable amylose content and its consequences. This information will be of prime importance in designing breeding strategies aiming to improve both the nutritional quality and yield of this valuable food crop. Various studies (Ong *et al*., 2012, Wanchana *et al*., 2003, Prathepha, 2003, Caffagni *et al*., 2013) have attempted to reveal the importance of waxy gene in controlling the diversity of amylose content in rice endosperm by using large rice populations but all those efforts have not been successful. The 217 USDA mini-core accessions of rice are diverse with respect to geographical origin and ancestry (Agrama *et al*., 2009, Bryant *et al*., 2013, Li *et al*., 2010) that is how most of the genetic diversity has been captured in this small but comprehensive collection that is representative of whole collection.

A wide range of diversity in different phenotypic traits has been observed including the variation in amylose content (Yang *et al*., 1991). This phenotypic diversity is basically representing the diversity that exists at genetic level. The information obtained from the diversity analysis of different phenotypic traits can be of great interest to plant breeders in developing the crops with desirable traits (Maduakor and Lal, 1989).

A number of methods has been reported for estimation of amylose content such as iodine binding method (colorimetric analysis), size exclusion chromatography, near infrared spectroscopy and asymmetric flow field flow fractionation (Juliano, 1971a, Wesley *et al*., 2003, Ward *et al*., 2006, Chiaramonte *et al*., 2012a). All of these methods except colorimetric analysis are not cost effective and require high tech equipment (Fitzgerald *et al*., 2009). Out of these techniques only iodine binding method is the one which is recommended for routine laboratory use (Caffagni *et al*., 2013). Due to this reason, present study was focussed on characterization of diverse USDA mini-core panel of rice germplasm on the basis of amylose content by this colorimetric method.

Results of Biochemical characterization of 207 mini-core lines revealed the amylose content based diversity of these rice genotypes by showing a wide range of apparent amylose content (0.3-

28.32%) covering all of the classes generated on the basis of amylose content. Majority of mini-core accessions fall in the category of intermediate amylose content. Previous studies (Juliano, 1985) have also revealed the dramatic variations in apparent amylose content (AAC) among different rice genotypes such as waxy (0 to 2%), very low (5 to 12%), low (12 to 20%), intermediate (20 to 25%) and high (25 to 33%). This variation in amylose content in various rice genotypes might be due to variable expression of waxy gene in rice endosperm.

Comparison between colorimetric analysis results of present study and results obtained from the work of Agrama et al. (Agrama et al., 2009) showed association i.e., mini-core genotypes HB-6-2 and EMBRAPA_1200 were in low amylose containing group. Furthermore, some other genotypes, including Dular and GPNO-25912 were grouped in intermediate and high amylose content groups respectively in both of the studies. Some of variation in amylose content values was also observed although same method was used for the analysis of amylose content. This might be due to the sensitivity of this determinant towards environmental factors during grain filling stages of rice crop (Chen *et al*., 2008).

Previous studies have highlighted five polymorphic sites in the waxy gene having important role in amylose content variations across different rice genotypes (Ayres *et al*., 1997, Larkin and Park, 2003, Wanchana *et al*., 2003, Chen *et al*., 2008, Asante *et al*., 2013, Caffagni *et al*., 2013). In present study out of these five polymorphic sites four (intron 1 G/T SNP, exon 6 G/C SNP, exon 10 C/T SNP and (CT)n repeats) were utilized for amylose content analysis of mini-core accessions

The results that were obtained from biochemical analysis were validated by using already reported waxy gene associated SNP and SSR based molecular markers. Purpose of selecting these markers was that these markers have not been used earlier in any study for genotyping of such a huge and diverse collection of rice accessions. USDA Mini-core collection provided us with an opportunity to have this broad range of genotypic diversity at one place.

Five SSR markers (RM25, RM190, RM339, RRM440 and GBSS1) were successfully employed for genotyping of mini-core collection. Some of these markers were highly informative in explaining the diversity of mini-core with respect to amylose content. From PowerMarker results it was evident that RM190 was highly informative SSR marker with a PIC value of 0.5919, this overall PIC value was similar to the PIC value reported by Palanga *et al.* (Palanga et al., 2016). These markers have confirmed the association between amylose content and *Wx* gene. Different studies have reported the existence of strong relationship between waxy gene associated microsatellite markers and amylose content and these variants have been used in classifying the cultivars that differ in cooking quality (Ayres *et al*., 1997, Bergman *et al.*, 2001). Functional alleles *Wxa* and *Wxb* of this highly informative marker (RM190) came into existence by a change in single base at 5' splice site of intron1 (Hirano and Sano, 1998) and facilitate in distinguishing low amylose content class from high and intermediate ones. These two alleles of microsatellite marker are not sufficient to explain the continuous variation across different classes of amylose and also unable to explain the waxy class of amylose. SNP based markers have ability to overcome this limitation of microsatellite markers by providing the additional information regarding waxy class of amylose due to having superior effect on amylose content (Chen *et al*., 2008).

Identification of three functional SNPs associated with waxy gene, including G to T substitution at 5' splice site of waxy intron 1, A to C substitution in *Waxy* exon 6 and C to T substitution in *Waxy* exon 10, facilitated the classification of rice accessions in four different (glutinous, low, intermediate and high) amylose classes (Wang *et al*., 1995, Larkin and Park, 1999, Bergman *et al.*, 2001).

Five SNP based markers have also been used in the study under discussion to overcome the limitations that were associated with microsatellite markers during further analysis and validation of results obtained from biochemical analysis. Out of these five SNP based markers three were associated with intron1 of *waxy* gene while other two markers were found associated with exon 6 and 10 of *waxy* gene. PowerMarker results revealed that how informative these markers are in explaining diversity of mini-core collection. Out of all these SNP markers, maximum information was provided by *waxy* intron1 based SNP marker with a PIC value of 0.548. Effectiveness of these SNP markers have

also been reported by Biselli *et al.* (Biselli et al., 2014) and Chen *et al.,* (Chen et al., 2010).

Although SNP based markers are more informative according to many previously published studies (Dobo *et al*., 2010, Asante *et al*., 2013) but in present study if we compare the PIC values of all the microsatellites and SNP based markers, information regarding diversity on the basis of amylose content, provided by some of SNP markers was even less than that provided by SSR markers. In the light of these findings it would be a better approach if we use combination of RM190 (SSR marker) and SNP based markers to get better and clear picture of diversity of USDA mini-core panel of rice germplasm.

Findings of phylogenetic and factorial analysis, which was conducted by DARwin software on the basis of 9 molecular marker scoring results, were in agreement with the results of biochemical and genotypic analysis. Both of these showed the distribution of mini-core accessions into three groups on the basis of amylose content. Although this phylogenetic tree was constructed on the basis of results obtained from 9 molecular markers even then clear clustering of different amylose containing accessions was observed. It is expected that in future use of more molecular markers results for this analysis will result in the clearer picture of this phylogenetic tree.

This study has provided us enough useful information that validates the use of these molecular markers for the analysis of amylose content. As these markers are found useful for this diverse collection of rice germplasm then it can be deduced that these markers are capable to analyse any challenging collection of diverse genotypes. Characterization of rice accessions by colorimetric analysis and with functional molecular markers will facilitate the breeding programs aimed at improving grain quality. Improvement in eating and cooking quality of rice will have a broader impact on economy of Pakistan as it will help in expanding the export market of Pakistani rice.

References

Agrama, H., Yan, W., Lee, F., Fjellstrom, R., Chen, M.-H., Jia, M. and Mcclung, A. 2009. Genetic assessment of a mini-core subset developed from the USDA rice genebank. *Crop Science*, 49**,** 1336-1346.

Asante, M. D., Offei, S. K., Gracen, V., Adu-Dapaah, H., Danquah, E. Y., Bryant, R. and Mcclung, A. 2013. Starch physicochemical properties of rice accessions and their association with molecular markers. *Starch-Stärke,* 65**,** 1022-1028.

Ayres, N., Mcclung, A., Larkin, P., Bligh, H., Jones, C. and Park, W. 1997. Microsatellites and a single-nucleotide polymorphism differentiate apparentamylose classes in an extended pedigree of US rice germ plasm. *TAG Theoretical and Applied Genetics*, 94**,** 773-781.

Bergman, C., Delgado, J., Mcclung, A. and Fjellstrom, R. 2001. An improved method for using a microsatellite in the rice waxy gene to determine amylose class. *Cereal Chemistry*, 78**,** 257-260.

Biselli, C., Cavalluzzo, D., Perrini, R., Gianinetti, A., Bagnaresi, P., Urso, S., Orasen, G., Desiderio, F., Lupotto, E. and Cattivelli, L. 2014. Improvement of marker-based predictability of Apparent Amylose Content in japonica rice through GBSSI allele mining. *Rice*, 7**,** 1.

Bryant, R. J., Jackson, A. K., Yeater, K. M., Yan, W. G., Mcclung, A. M. and Fjellstrom, R. G. 2013. Genetic variation and association mapping of protein concentration in brown rice using a diverse rice germplasm collection. *Cereal Chemistry*, 90**,** 445-452.

Caffagni, A., Albertazzi, G., Gavina, G., Ravaglia, S., Gianinetti, A., Pecchioni, N. and Milc, J. 2013. Characterization of an Italian rice germplasm collection with genetic markers useful for breeding to improve eating and cooking quality. *Euphytica*, 194**,** 383-399.

Cai, H., Xu, D., Zhou, L., Cheng, J., Zhang, Z., Wu, J. and You, A. 2015. Development of PCR-based CNP marker of rice Waxy. *Russian Journal of Genetics*, 51**,** 673-676.

Causse, M. A., Fulton, T. M., Cho, Y. G., Ahn, S. N., Chunwongse, J., Wu, K., Xiao, J., Yu, Z., Ronald, P. C. and Harrington, S. E. 1994. Saturated molecular map of the rice genome based on an interspecific backcross population. *Genetics*, 138**,** 1251-1274.

Chang, T. T. 1976. The origin, evolution, cultivation, dissemination, and diversification of Asian and African rices. *Euphytica*, 25**,** 425-441.

Chen, M. H., Bergman, C. J., Pinson, S. R. and Fjellstrom, R. G. 2008. Waxy gene haplotypes: associations with pasting properties in an international rice germplasm collection. *Journal of Cereal Science*, 48**,** 781-788.

Chen, M. H., Fjellstrom, R. G., Christensen, E. F. and Bergman, C. J. 2010. Development of

three allele-specific codominant rice Waxy gene PCR markers suitable for marker-assisted selection of amylose content and paste viscosity. *Molecular Breeding*, 26**,** 513-523.

Cheng, A., Ismail, I., Osman, M. and Hashim, H. 2014. Mapping of quantitative trait loci for aroma, amylose content and cooked grain elongation traits in rice. *Plant Omics*, 7**,** 152.

Chiaramonte, E., Rhazi, L., Aussenac, T. and White, D. R. 2012a. Amylose and amylopectin in starch by asymmetric flow field-flow fractionation with multi-angle light scattering and refractive index detection (AF4–MALS–RI). *Journal of Cereal Science*, 56**,** 457-463.

Chiaramonte, E., Rhazi, L., Aussenac, T. and White Jr, D. R. 2012b. Amylose and amylopectin in starch by asymmetric flow field-flow fractionation with multi-angle light scattering and refractive index detection (AF4–MALS–RI). *Journal of cereal science*, 56**,** 457-463.

Curtis, B. C., Rajaram, S. and Gómez, M. 2002. *Bread wheat: improvement and production*, Food and Agriculture Organization of the United Nations (FAO).

Dobo, M., Ayres, N., Walker, G. and Park, W. D. 2010. Polymorphism in the GBSS gene affects amylose content in US and European rice germplasm. *Journal of cereal science*, 52**,** 450-456.

Feng, F., Li, Y., Qin, X., Liao, Y. and Siddique, K. H. M. 2017. Changes in Rice Grain Quality of Indica and Japonica Type Varieties Released in China from 2000 to 2014. 8.

Fernando, H., Kajenthini, T., Rebeira, S., Bamunuarachchige, T. and Wickramasinghe, H. 2015. Validation of molecular markers for the analysis of genetic diversity of amylase content and gel consistency among representative rice varieties in Sri Lanka. *Tropical Agricultural Research*, 26.

Fitzgerald, M. A., Bergman, C. J., Resurreccion, A. P., Moller, J., Jimenez, R., Reinke, R. F., Martin, M., Blanco, P., Molina, F. and Chen, M.-H. 2009. Addressing the Dilemmas of Measuring Amylose in Rice [Erratum: 2009 Nov-Dec, v. 86, no. 6, p. 714.]. *Cereal Chemistry*, 86**,** 492-498.

Garris, A. J., Tai, T. H., Coburn, J., Kresovich, S. and Mccouch, S. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics*, 169**,** 1631-1638.

Giri, C. and Laxmi, G. V. 2000. Production of transgenic rice with agronomically useful genes: an assessment. *Biotechnology advances*, 18**,** 653-683.

Hirano, H.-Y. and Sano, Y. 1998. Enhancement of Wx gene expression and the accumulation of amylose in response to cool temperatures during seed development in rice. *Plant and cell physiology*, 39**,** 807-812.

Hu, C., Shi, J., Quan, S., Cui, B., Kleessen, S., Nikoloski, Z., Tohge, T., Alexander, D., Guo, L., Lin, H., Wang, J., Cui, X., Rao, J., Luo, Q., Zhao, X., Fernie, A. R. and Zhang, D. 2014. Metabolic variation between japonica and indica rice cultivars as revealed by non-targeted metabolomics. *Scientific Reports*, 4**,** 5067.

Jain, N., Jain, S., Saini, N. and Jain, R. K. 2006. SSR analysis of chromosome 8 regions associated with aroma and cooked kernel elongation in Basmati rice. *Euphytica*, 152**,** 259-273.

Juliano, B. 1971a. A simplified assay for milled-rice amylose. *Cereal Science Today*, 16**,** 334-340,360.

Juliano, B. 1971b. A simplified assay for milled rice amylase. *Cereal Sci. Today*, 16**,** 360.

Juliano, B. 1985. Criteria and tests for rice grain qualities. *Rice: Chemistry and Technology* 443-524.

Kent, N. L. 1978. Technology of cereals; with special reference to wheat.

Khan, M. H., Dar, Z. A. and Dar, S. A. 2015. Breeding strategies for improving rice yield—a review. *Agricultural Sciences*, 6**,** 467.

Khush, G. S. 2005. What it will take to feed 5.0 billion rice consumers in 2030. *Plant molecular biology*, 59**,** 1-6.

Kottearachchi, N., Peiris, R. and Rebeira, S. 2014. Application of DNA markers for the detection of amylose content in Sri lankan rice (*Oryza sativa* L.) varieties. *Asian Journal of Agriculture and Biology*, 2**,** 44-50.

Larkin, P. D. and Park, W. D. 1999. Transcript accumulation and utilization of alternate and non-consensus splice sites in rice granule-bound starch synthase are temperature-sensitive and controlled by a single-nucleotide polymorphism. *Plant molecular biology*, 40**,** 719-727.

Larkin, P. D. and Park, W. D. 2003. Association of waxy gene single nucleotide polymorphisms with starch characteristics in rice (*Oryza sativa* L.). *Molecular Breeding*, 12**,** 335-339.

Li, K., Bao, J., Corke, H. and Sun, M. 2017. Association Analysis of Markers Derived from Starch Biosynthesis Related Genes with Starch Physicochemical Properties in the USDA Rice Mini-Core Collection. 8.

Li, X., Yan, W., Agrama, H., Hu, B., Jia, L., Jia, M., Jackson, A., Moldenhauer, K., Mcclung, A. and Wu, D. 2010. Genotypic and phenotypic characterization of genetic differentiation and diversity in the USDA rice mini-core collection. *Genetica*, 138**,** 1221-1230.

Linares, O. F. 2002. African rice (Oryza glaberrima): history and future potential. *Proceedings of the National Academy of Sciences*, 99**,** 16360-16365.

Maduakor, H. and Lal, R. 1989. Root system topgrowth and yield of cassava (Manihot esculenta Crantz) as affected by plant population in an arid soil. *Journal of Root Crops*, 15**,** 115-122.

Mekonnon, B. 2014. Selection of barley varieties for their yield potential at low Rain Fall Area based on both quantitative and qualitative characters North West Tigray, Shire, Ethiopia. *International Journal of plant breeding and Genetics*, 8**,** 205-213.

Okpala, N. E., Potcho, M. P., An, T., Ahator, S. D., Duan, L. and Tang, X. 2020. Low temperature increased the biosynthesis of 2-AP, cooked rice elongation percentage and amylose content percentage in rice. *Journal of Cereal Science*, 93**,** 102980.

Ong, Y., Ngu, M., Parviz, F., Tilakavati, K. and Wickneswari, R. 2012. Study on Waxy Gene Polymorphism and Amylose Content of Breeding Lines Derived from *Oryza rufipogon* x *Oryza Sativa* CV. MR219. *Pertanika Journal of Tropical Agricultural Science*, 35**,** 833-842.

Palanga, K. K., Traore, K., Bimpong, K., Jamshed, M. and Mkulama, M. A. 2016. Genetic diversity studies on selected rice varieties grown in Africa based on aroma, cooking and eating quality. *African Journal of Biotechnology*, 15**,** 1136-1146.

Prathepha, P. 2003. Characterization of Waxy microsatellite classes that are closely linked to the rice Waxy gene and amylose content in Thai rice germplasm. *Journal of Science and Technology*, 25**,** 1-8.

Sarvari, M. and Pepo, P. 2014. Effect of production factors on maize yield and yield stability. *Cereal Research Communications*, 42**,** 710-720.

Temnykh, S., Park, W. D., Ayres, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y. G., ISHII, T. and Mccouch, S. R. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *TAG Theoretical and Applied Genetics*, 100**,** 697-712.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R. and Polasky, S. 2002. Agricultural sustainability and intensive production practices. *Nature*, 418**,** 671-677.

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. *Theoretical and Applied Genetics*, 104**,** 1-8.

Wanchana, S., Toojinda, T., Tragoonrung, S. and Vanavichit, A. 2003. Duplicated coding sequence in the waxy allele of tropical glutinous rice (*Oryza sativa* L.). *Plant Science*, 165**,** 1193-1199.

Wang, Z. Y., Zheng, F. Q., Shen, G. Z., Gao, J. P., Snustad, D. P., Li, M. G., Zhang, J. L. and Hong, M. M. 1995. The amylose content in rice endosperm is related to the post-transcriptional regulation of the Waxy gene. *The Plant Journal*, 7**,** 613-622.

Ward, R. M., Gao, Q., D. E. Bruyn, H., Gilbert, R. G. and Fitzgerald, M. A. 2006. Improved methods for the structural analysis of the amylose-rich fraction from rice flour. *Biomacromolecules*, 7**,** 866-876.

Wesley, I., Osborne, B., Anderssen, R., Delwiche, S. and Graybosch, R. 2003. Chemometric localization approach to NIR measurement of apparent amylose content of ground wheat. *Cereal chemistry*, 80**,** 462-467.

Yan, W., Rutger, J. N., Bryant, R. J., Bockelman, H. E., Fjellstrom, R. G., Chen, M.-H., Tai, T. H. and Mcclung, A. M. 2007. Development and evaluation of a core subset of the USDA rice germplasm collection. *Crop Science*, 47**,** 869-876.

Yang, R. C., Jana, S. and Clarke, J. 1991. Phenotypic diversity and associations of some potentially drought-responsive characters in durum wheat. *Crop Science*, 31**,** 1484-1491.