Available online freely at www.isisn.org

Bioscience Research

Print ISSN: 1811-9506 Online ISSN: 2218-3973

Journal by Innovative Scientific Information & Services Network

RESEARCH ARTICLE BIOSCIENCE RESEARCH, 2021 18(2):1681-1693. **OPEN ACCESS**

Phylogenetics of rice (*Oryza sativa* L.) Genotypes from Pakistan based on diversity of biochemical markers

Haseena Gulzar[1], Sidra Pervez[2], Asad Jan[3] and Muhammad Asif Nawaz[1]*

[1]Department of Biotechnology, Shaheed Benazir Bhutto University, Sheringal Dir Upper, P.O Box 18000 Khyber Pakhtunkhwa, **Pakistan**
[2]Department of Biochemistry, Shaheed Benazir Bhutto Women University, Peshawar, Khyber Pakhtunkhwa, **Pakistan**
[3]Institute of Biotechnology and Genetic Engineering, Agriculture University Peshawar, P.O. Box 706 Khyber Pakhtunkhwa, **Pakistan**

***Correspondence:** asif_biotech33@sbbu.edu.pk, **Received 05-04-2021, Revised: 25-05-2021, Accepted: 05-06-2021 e-Published: 20-06-2021**

In this study phylogenetics and diversity of 150 rice (*Oryza sativa* L.) genotypes originated from different regions of an agricultural country Pakistan, self-sufficient in its production were analyzed. Total seed protein profile of these genotypes was assessed using a technique of sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). The seeds of accessions from different geographical regions of the country were provided by the Institute of Agri-Biotechnology and Genetic Resources, National Agricultural Research Centre (IABGRI-NARC) Islamabad, Pakistan. Total endosperm proteins from each genotype were extracted and used as biochemical markers. Three different protein extraction protocols were followed in comparison to optimize the procedure. An immense content of genetic variability was confirmed based on polymorphism data of polypeptide bands. The Jaccard's dissimilarity coefficient ranged from 0.09-0.94 for all 150 accessions. Based on the neighbour joining method of clustering, the whole germplasm was divided into two clusters, each further subdivided sorting centre of origin and depicting diversity. In the same manner, 126 accessions of province Khyber Pakhtunkhwa (KPK) were also assessed separately. This wealth of knowledge based on proteins profiling will be used for marker assisted breeding/selection, germplasm conservation strategies and easy identification of these genotypes from Pakistan.

Keywords: Rice, Phylogenetics, Genetic diversity, SDS-PAGE, Genotypes

INTRODUCTION

Rice, a member of family *Poaceae* is the staple food for more than half of the world's population and South Asia being the biggest rice producer and consumer. Nevertheless, it is cultivated all over the world on about 154 million hectares every year occupying 11% of the world's cultivated area, and 10% only in Pakistan. It adds 5.9% of value in agriculture and 1.3% in GDP in Pakistan, a value larger than other South Asian countries therefore, also called 'Golden Grain of Pakistan (Simmonds 1986; Shah et al. 2011). Among the cereal crops rice has the smallest genome, well characterized and sequenced thus a model plant for genomic research having 24 chromosomes (Eckardt, 2000; Swain et al. 2017). Pakistan is among one of the few South Asian countries, self-sufficient in rice production with increasing yield since 2001 and is the second largest exporter with a wealth of diverse traditional

landraces still cultivating in villages (Shah et al. 2011). More than 2,500 accessions collected from different areas of Pakistan are preserved in the national gene bank, IABGRI-NARC, Islamabad. Moreover, 100,000 traditional rice varieties are stored in the gene banks throughout the world (Pervaiz et al. 2011; Bishwajit et al. 2013). In the last two decades many researchers have put their efforts to assess the genetic variability present in rice germplasm of Pakistan (Khan et al. 2013; Tahir, 2014; Bibi et al. 2017).

Landraces are valuable genetic resources containing huge genetic variability also hold several unwanted agronomic traits like tall plant stature, long crop duration etc however, also exhibit strong characteristics such as resistance to biotic and abiotic stresses. In the last decade of 20th century, Pakistan has released a number of superior cultivars by crossing advanced breeding lines and traditional varieties, achieving high quality of grains and resistance to biotic and abiotic stresses (Pervaiz et al. 2011; Haque, et al. 2015). SDS-PAGE is useful, easy, low cost and a quick way for studying genetic diversity in numerous genotypes based on polymorphic endosperm storage proteins which are unlike morphological traits not under the influence of environment (Ali et al. 2007). Therefore, the information provided by protein profiling can truly be beneficial in the screening of desirable traits in these germplasm resources for efficient utilization in advanced crop improvement, genetic engineering and breeding strategies (Conrad et al. 2013; Gulzar et al. 2015; Gul et al. 2015).

This work is one of the efforts to attain these goals as to ensure food security, systematic evaluation of the valuable genetic resources is pre-requisite for using in crop improvement programs (Bibi et al. 2017).

MATERIALS AND METHODS

This study was conducted at the Institute of Biotechnology and Genetic Engineering (IBGE), Agriculture University Peshawar, KPK, Pakistan. Rice germplasm was provided by IABGRI-NARC, Islamabad, Pakistan.

The work consisted of two experiments; optimization of protein extraction protocol and biochemical evaluation of 150 rice accessions originated from Pakistan. The accessions belonging to KPK province (126) were analyzed separately including check variety Fakhr-e-Malakand based on SDS-PAGE of endosperm storage proteins.

Optimization of protein extraction protocol

For optimization three different compositions of protein extraction buffer (PEB) were used to extract proteins from rice endosperms in order to obtain clear bands.

PEB 1: 0.5M Tris-HCl pH 8.0, 0.2% SDS, 5M urea, 1% 2-mercaptoethanol (2-ME) (Ranjan et al. 2012).

PEB 2: Phosphate buffered saline (PBS):137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4 and 1.8 mM KH_2PO_4 with pH 7.4 (Walczyk et al. 2017).

PEB 3: Cold spring harbour protocol for 2X sample buffer; 1M Tris, pH 6.8, 50% glycerol, 10% SDS, 0.5% bromophenol blue (BPB) and 0.4 ml, 2-ME (Cold Spring Harb Protoc, 2009)

Using PEB 1 and 3, the protocol followed include dehusking of rice seeds, ground into fine powder with morter and pestle and centrifugation (0.01 g seed meal/400 µl PEB) at 10,000 rpm for 10 minutes to obtain supernatant and stored at -20°C (Pervaiz et al. 2011).

While using PEB 3 or PBS, the protocol followed include the steps of dehusking rice seeds, ground into fine meal, defatting rice meal with n-hexane twice, protein extraction by shaking 0.01 g defatted seed meal/500 µl PBS for 2 hours at 4°C and centrifugation at 12,000 rpm for 10 minutes to obtain supernatant and stored at -20°C until use. Before loading into gel, the PEB 1 was used as loading dye with proteins extracted in PBS in the same ratio (1:1) (Walczyk et al. 2017).

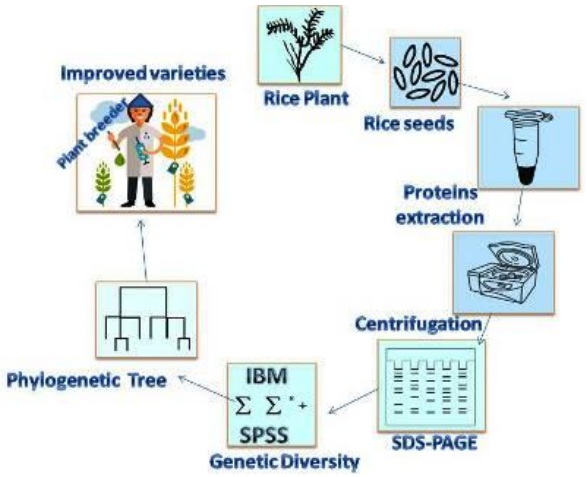
Biochemical characterization using SDS-PAGE

Total rice seed proteins were evaluated by one dimensional discontinuous system of SDS-PAGE with some modifications in Leammli's method (1970). Extracted proteins (12µl each) were loaded in 5% stacking and 12% separating gel, run at 25mA and100V for 60 to 90 min. Bio-Rad dual precision protein ladder was used as molecular weight marker (10 to 250 kDa). The gels were stained overnight in 0.02% Comassie Brilliant Blue (CBB) by continuous shaking and destained in distilled water until the background was clear (Khan et al. 2013).

Data analysis for phylogenetics and diversity

An electrophoregram was analyzed for the presence/absence (1/0) of polypeptide bands in each accession. The binary data made was used to generate similarity and dissimilarity matrix which was further processed to construct a phylogenetic tree by the unweighed pair group method with arithmetic means (UPGMA)

introduced by Sneath and Sokal in (1973) using a Statistical Programme for Social Scientists (SPSS) and Ward's method (Ali et al., 2007; Rahman et al. 2010).



RESULTS

In this paper 150 different local rice (*Oryzasativa* L.) accessions originated from different regions of Pakistan—including north western regions (Province of KPK), including Parachinar and Kurrum agency, northern areas (Gilgit-Baltistan) and Punjab province—were analyzed for phylogenetic relationship and genetic diversity using SDS-PAGE of total seed proteins (Table 1). Most of these were from KPK Province (126) including a check variety Fakhr-e-Malakand and only 24 from the other regions of the country.

Optimization of protein extraction method

Three different protein extraction protocols were used and the procedure optimized. The electrophoregrams generated by first protocol and third (Ranjan et al. 2012), resulted in clear bands however, we recommend the use of second protocol with PBS pH 7.4 (Walczyk et al (2017) for genetic variability studies in rice which gave us more clear and sharp bands (Fig. 1 & 2). Moreover, up to 12 samples can be run in a single gel when working with protocol 2, thus saving time and resources.While with protocol 1, only six samples can be run at a time because it produce more spread out and denser bands (Fig. 1 & 2).

We found that defatting seed meal with n-hexane improved protein extraction process with no floating fats on surface and gave clear bands.

SDS-PAGE based Polymorphism

In this study the polymorphism present in 150 local rice (*Oryza sativa* L.) accessions of Pakistan were biochemically analyzed for seed storage proteins using SDS-PAGE. A huge content of genetic variability was seen based on polypeptides banding pattern, size (10 to 250kDa) and number (14 to 27). Both major and minor bands were kept under consideration, thus results showed comparatively significant variation in minor bands although some variation in major bands was there (Fig. 1 and 2) (Tahir et al. 2014). For ease of scoring the entire banding pattern was divided into four zones, zone A (60-to-250kDa), B (30-to-60-kDa), C (18-to-30kDa) and D (10-to-18kDa) as shown in Fig. 1 and 2. Maximum number of bands were found in zone A and B (each 11 bands) followed by C (08) and D (04) in different accessions. Therefore, during analysis 34 bands were observed with minimum number of 14 bands and maximum of 27 bands in different accessions (Table 2). Out of 34 bands 10 (29.4%) were monomorphic and 24 (70.5%) were polymorphic.

In zone A, nine bands showed polymorphism while eight, six and one band was polymorphic in zone B, C and D respectively. In present investigation, maximum number of bands were seen in the accessions belonging to province KPK, Chitral including accession no.7284 with 27 bands followed by 25 bands in accessions no. 7626, 7805, 7794. While 24 bands were observed in accession no. 7598 (KPK, Malakand), 17940 (KPK, Swat), 7886 (Pakistan Agriculture Research Council (ARC) and 7801, 7791, 7792, 7808, 7811 accessions from Chitral District (Tables 1, 2, 3 & 5). Minimum number of bands i.e. 14 was observed in accessions number 7281 (KPK, Mansehra), 7291 & 7292 (KPK, Swat) and 7829 (KPK). It shows that the genetic resources of rice originated from district Chitral possess more diversity than that of other districts of the country (Bibi et al. 2017). In a check variety Fakhr-e-Malakand 17 (2, 6, 6, 3) bands were observed corresponding to zones A –D (Table 1-3, Fig. 1 & 2).

Rf values of the last bands of each zone (A, B, C, D) were recorded as 0.32, 0.58, 0.75 and 0.96 respectively, shown in Fig. 1 & 2 (Chen et al. 1987). The sum total of maximum bands observed in each zone is equal to 34 (A; 11, B; 11, C: 08, D; 04).

Heirarchical Clustering:

Cluster analysis for the whole germplasm

Cluster analysis (Ward's method) Ward,

(1963) for the whole germplasm consisting of 150 rice landraces of Pakistan was carried out using binary data processed by SPSS software. A dendrogram was developed which divided the whole Germplasm into two major clusters; cluster A consisted of 58 accessions while the rest 92 genotypes grouped in cluster B, each further classified into two (A1, A2 & B1, B2) sub-clusters having 42, 16, 48 and 44 accessions respectively. A check variety Fakhr-e-Malakand (FM) was grouped in cluster B2 as shown in table 4. The Jaccard's dissimilarity coefficient ranged from 0.09 (lowest) to 0.94 (highest), while that of similarity coefficient from 0.009 to 0.635, revealing highest genetic diversity among the rice landraces of Pakistan.

Table 1: Passport information of 150 Rice (*Oryza sativa* L.) landraces of Pakistan characterized for total seed proteins

Province/District	Accession numbers									
Khyber Pakhtunkhwa	7237	7238	7239	7241	7242	7271	7280	7281	7284	7285
	7286	7287	7288	7289	7290	7291	7292	7293	7294	7295
	7296	7297	7397	7593	7594	7595	7597	7598	7599	7600
	7601	7603	7604	7605	7606	7608	7609	7610	7611	7612
	7613	7615	7626	7628	7630	7633	7634	7635	7637	7638
	7639	7640	7641	7642	7648	7649	7650	7651	7652	7653
	7654	7655	7657	7737	7738	7753	7754	7755	7756	7760
	7762	7783	7784	7785	7786	7787	7788	7789	7790	7791
	7792	7793	7794	7795	7796	7797	7798	7799	7800	7801
	7803	7804	7805	7806	7807	7808	7809	7811	7812	7813
	7814	7815	7817	7818	7819	7820	7821	7822	7824	7825
	7829	7831	7832	7833	7836	7839	7937	8083	8085	17940
	17941	17942	17943	17944	17946	Fakhre	Mkd			
RGP-ARC	7847	7850	7859	7863	7865	7869	7886			
Unknown	7411	7413	7414	7416	7418	7419				
Punjab	7225	7673	7763	7765						
Parachinar	7258	7261	7262							
AJ Kashmir	7735	7736								
Gilgit, Baltistan	7420									
Kurrum Agency	7260									

Table2 :A table complementary to table 1 showing number of polypeptide bands obtained in each of the 150 rice (*Oryza sativa* L.) accessions of Pakistan

	No. of polypeptide bands obtained in each rice accession of table 1									
Province/District	Accession numbers									
Khyber Pakhtunkhwa	18	15	17	16	18	19	15	14	27	19
	21	18	21	22	23	14	14	15	15	18
	18	20	17	21	21	21	20	24	17	21
	21	21	21	22	19	20	20	20	19	19
	15	19	25	19	22	20	20	23	15	20
	23	20	20	19	20	19	16	16	17	23
	18	18	19	17	20	22	19	21	15	17
	18	19	18	17	19	22	21	20	16	24
	24	20	25	22	23	17	21	23	22	24
	22	21	25	23	22	24	19	24	19	20
	22	19	23	19	23	18	16	17	17	21
	14	19	19	19	21	15	22	19	17	24
	21	17	19	23	20	17				
RGP-ARC	19	17	19	19	16	15	24			
Unknown	18	17	19	18	18	19				
Punjab	21	20	19	18						
Parachinar	20	15	18							
AJ Kashmir	18	18								
Gilgit, Baltistan	17									
Kurrum Agency	18									

Table 3 : A table complementary to table 1 & 2 showing number of polypeptide bands obtained in each of the 150 rice (*Oryza sativa* L.) accessions of Pakistan where, A, B, C, D = No. of bands in each zone (as shown in fig . 1 & 2) and T = Total no of bands in each accession.

Province/ District	A,B,C,D =T	A,B,C,D =T	A,B,C,D =T	A,B,C,D =T	A,B,C,D =T	A,B,C,D =T	A,B,C,D =T	A,B,C,D =T	A,B,C,D =T	A,B,C,D =T
	7,6,2,3 =18	6,4,3,2 =15	8,4,2,3 =17	7,5,2,2 =16	8,5,2,3 =18	6,5,5,3 =19	3,4,5,3 =15	5,3,3,3 =14	10,11,3, 3=27	7,4,5,3 =19
Khyber Pakhtunkhwa	7,8,3,3 =21	6,6,3,3 =18	8,5,5,3 =21	9,7,3,3 =22	11,5,4, 3=23	3,4,4,3 =14	4,2,5,3 =14	3,4,5,3 =15	3,4,5,3= 15	7,4,4,3 =18
	6,5,4,3 =18	8,4,5,3 =20	6,5,3,3 =17	8,7,3,3 =21	7,7,4,3 =21	9,5,4,3 =21	7,5,4,4 =20	9,9,4,3 =24	7,5,2,3= 17	9,6,3,3 =21
	9,6,3,3 =21	9,6,3,3 =21	8,7,3,3 =21	8,6,4,4 =22	6,5,5,3 =19	7,5,5,3 =20	7,5,5,3 =20	6,6,5,3 =20	6,4,6,3= 19	4,4,4,3 =15
	7,5,4,3 =19	6,6,4,3 =19	9,6,6,4 =25	8,5,3,3 =19	6,9,4,3 =22	7,5,5,3 =20	7,5,5,3 =20	8,5,7,3 =23	5,4,4,2= 15	7,5,5,3 =20
	7,6,7,3 =23	7,4,6,3 =20	7,5,5,3 =20	7,5,4,3 =19	8,5,4,3 =20	6,5,4,4 =19	5,4,4,3 =16	6,5,2,3 =16	7,5,2,3= 17	8,7,5,3 =23
	5,5,5,3 =18	4,5,6,3 =18	7,5,4,3 =19	5,4,5,3 =17	7,5,5,3 =20	9,4,6,3 =22	7,5,4,3 =19	9,4,5,3 =21	5,4,3,3= 15	6,5,3,3 =17
	7,6,2,3 =18	7,5,4,3 =19	7,5,3,3 =18	7,5,6,3 =17	7,5,4,3 =19	7,6,6,3 =22	9,5,4,3 =21	7,5,5,3 =20	5,4,4,3= 16	8,6,7,3 =24
	8,5,5,3 =21	6,6,5,3 =20	8,7,7,3 =25	7,6,6,3 =22	8,6,6,3 =23	4,5,5,3 =17	7,6,5,3 =21	8,7,5,3 =23	8,6,5,3= 22	7,6,8,3 =24
	7,7,5,3 =22	6,6,6,3 =21	7,8,7,3 =25	7,7,6,3 =23	6,6,7,3 =22	10,6,5, 3=24	8,5,3,3 =19	9,6,6,3 =24	6,6,4,3= 19	7,5,5,3 =20
	8,6,5,3 =22	7,5,4,3 =19	9,5,6,3 =23	6,6,4,3 =19	5,9,6,3 =23	8,7,6,3 =18	5,5,5,1 =16	4,5,5,3 =17	5,5,4,3= 17	7,6,5,3 =21
	5,3,3,3 =14	6,5,5,3 =19	7,6,3,3 =19	6,5,5,3 =19	6,6,6,3 =21	5,4,3,3 =15	6,7,6,3 =22	6,5,5,3 =19	6,4,4,3= 17	11,5,5, 4=24
	5,6,7,3 =21	2,5,7,3 =17	5,4,7,3 =19	11,4,5, 3=23	7,6,4,3 =20	2,6,6,3 =17				
RGP-ARC	7,5,4,3 =19	4,4,6,3 =17	7,5,4,3 =19	5,5,5,3 =19	7,5,5,3 =16	4,4,5,2 =15	8,6,7,3 =24			
Unknown	6,4,5,3 =18	6,4,4,3 =17	7,4,5,3 =19	5,5,5,3 =18	7,4,4,3 =18	5,6,5,3 =19				
Punjab	7,6,5,3 =21	7,5,5,3 =20	6,6,4,3 =19	6,4,5,3 =18						
Parachinar	7,5,5,3 =20	5,4,3,3 =15	6,4,5,3 =18							
AJ Kashmir	6,5,4,3 =18	9,5,1,3 =18								
Gilgit-Baltistan	7,4,3,3 =17									
Kurrum-Agency	7,5,3,3 =18									

Table4 :Based on phylogenetic relationship assemblage of rice (*Oryza sativa* L.) landraces/accessions of Pakistan (as in Fig. 3) into two groups (A & B) and sub-group (A1, A2, B1, B2)

Groups	Accession numbers									
Group A1	7863	7831	7859	7836	7755	7765	7420	7239	7238	7800
	7833	7813	7937	8083	7286	7287	7798	7642	7296	7811
	7799	7785	7801	7258	7633	7594	7641	7595	7260	7806
	7808	7814	7601	7754	7626	7787	7788	7635	7805	7794
	7753	7414								
Group A2	7599	7237	7289	7598	7653	7284	7600	7603	7604	7241
	7736	7613	7784	7242	7821	7605				
Group B1	7418	7611	7610	7639	7608	7807	7815	7597	7792	7738
	7886	7290	7673	7640	7397	7630	7762	7786	7797	7737
	7735	7652	7593	7803	7634	17946	7609	7789	7812	7818
	7791	7796	7817	17940	17944	7295	7606	7638	7262	7804
	7793	7225	7825	7649	7847	7648	7820	7795		
Group B2	7294	17942	7941	7655	7869	7416	7280	7865	7612	7419
	FM	7657	7756	7824	7829	7292	7628	7832	7288	7763
	7783	7819	7297	7615	7809	8085	7411	7291	7293	7850
	7285	7650	7822	7839	7271	7790	7261	7413	17943	7654
	7760	7651	7281	7637						

Table5 :Passport information of 126 Rice (*Oryza sativa* L.) landraces of province Khyber Pakhtunkhwa (KPK) Pakistan characterized for total seed proteins

KPK Districts	Accession numbers												
KPK, Swat	7290	7291	7292	7293	7294	7295	7397	7633	7634	7635	7637	7638	7639
	7640	17940											
KPK,Chitral	7284	7285	7286	7287	7288	7289	7611	7612	7613	7615	7626	7628	7630
	7783	7784	7785	7786	7787	7788	7789	7790	7791	7792	7793	7794	7795
	7796	7797	7798	7799	7800	7801	7803	7804	7805	7806	7807	7808	7809
	7811	7812	7813	7814	7815	7817	7818	7819	7820	7821	7822	7824	7825
	7937	8083	8085	17943	17944								
KPK(Unknown)	7237	7238	7239	7241	7242	7271	7280	7641	7642	7737	7738	7756	7760
	7762	7829	7831	7832	7833	7836	7839						
KPK, Mansehra	7281	7296	7297	7648	7649	7650	7651	7652	7653	7654	7655	7657	17946
KPK, Malakand	7593	7594	7595	7597	7598	7599	7600	Fakhre	Mkd				
KPK, Dir	7601	7603	7604	7605	7606	7608	7609	7610					
KPK, Besham	7753	7754	7755										
KPK, Nowshehra	17941	17942											

Table 6: Based on phylogenetic relationship, a dendogram (fig. 5) classify 126 rice (*Oryza sativa* L.) landraces of province Khyber Pakhtunkhwa Pakistan into 2 groups (A & B) and sub-groups, A1a, A1b, A2a, A2b, B1, B2

Groups	Accession numbers												
A1a	7609	7789	7812	7818	7634	17946	7593	7803	7787	7788	7738	7635	7290
	7640	7806	7808	7814	7601	7754	7755	7626	7791	7796	7817	17940	17944
A1b	7608	7611	7639	7292	7606	7783	7825	7638	7793	7805	7794	7753	7800
	7833	7813	7937	8083	7397	7630	7819	7628	7832	7762	7652	7648	7820
	7795	7649	7597	7792	7807	7815	7804						
A2a	7756	7829	7824	17941	8085	7809	7657	7786	7595	7633	7613	7799	7285
	7801	7594	7641										
A2b	7294	17942	7822	7839	7280	7655	FM	7797	7737	7292	7612	7650	7297
	7615	7271	7790	7291	7293	7836	7831	7285	7288	7238	7760	7651	7281
	7637	17943	7654	7610									
B1	7286	7287	7798	8084	7296	7811							
B2	7821	7605	7784	7242	7239	7241	7598	7653	7284	7600	7603	7604	7599
	7237	7289											

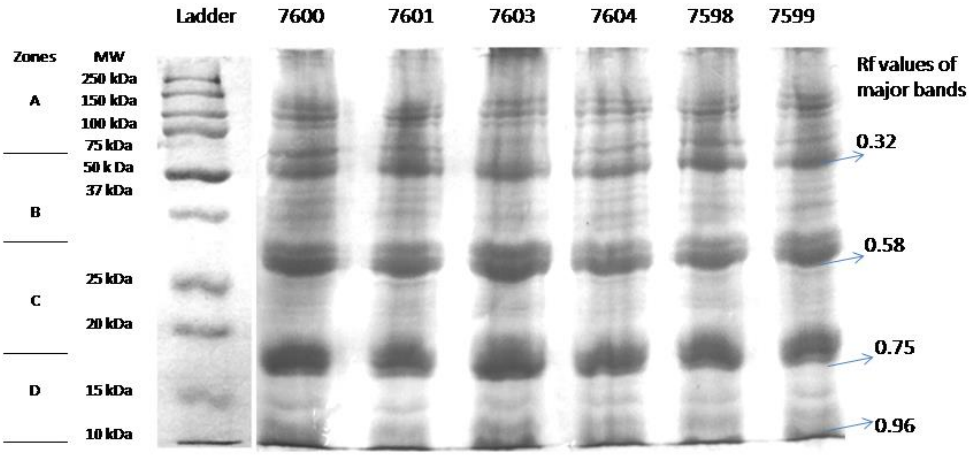


Figure 1: Rice (*Oryza Sativa* L.) electrophoregram of endosperm proteins from Pakistan s accessions (7600-7599) extracted by protocol 1, divided into four zones (A, B, C, D) and Rf values of major bands given

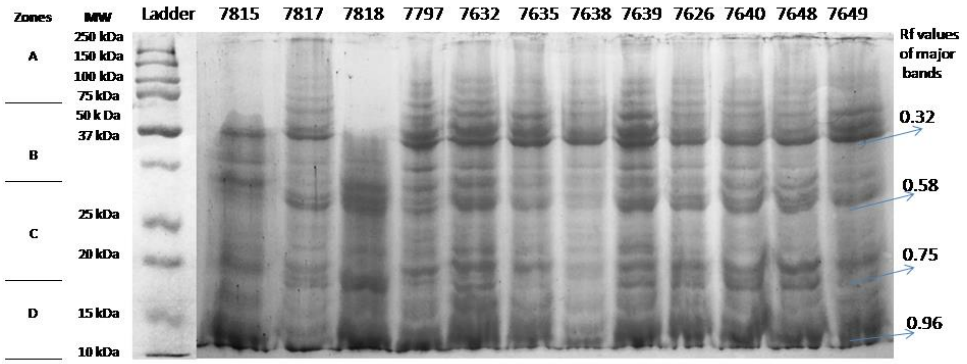


Figure 2: Rice (*Oryza Sativa* L.) electrophoregram of endosperm proteins from 12 Pakistans accessions (7815-7849) extracted by protocol 2, showing variation divided into four zones ($A_2B_2C_2D_2$) and Rf values of major bands given

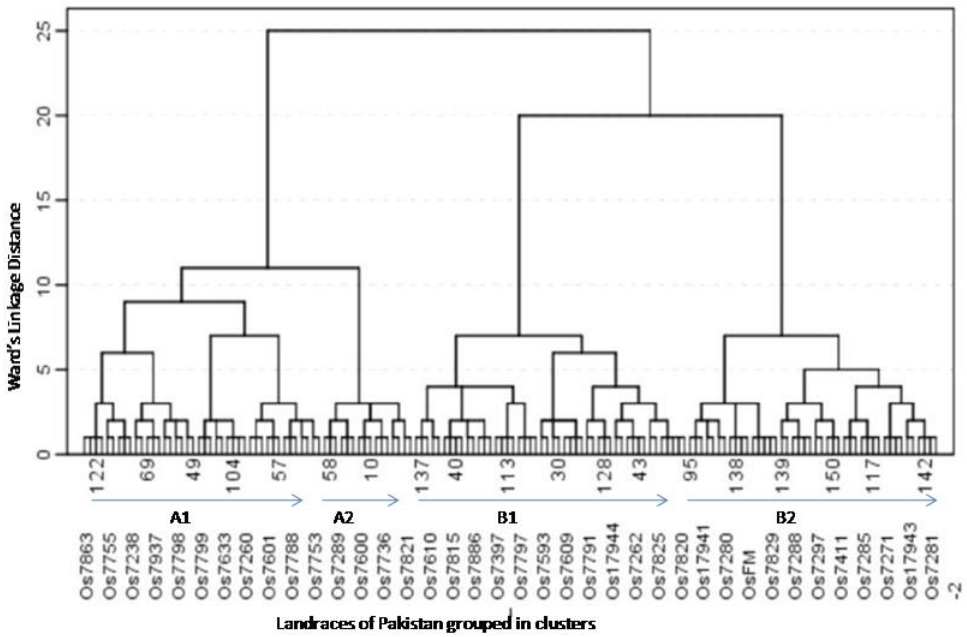


Figure 3: A dendogram showing phylogenetic relationship and genetic diversity among 150 rice (*Oryza Sativa* L.) accessions of Pakistan based on SDS-PAGE

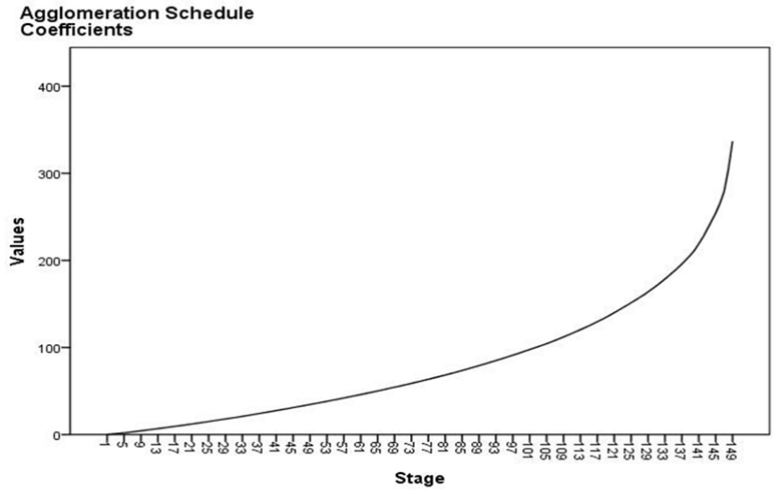


Figure 4: Graph showing relationship between Agglomeration schedule coefficients (values) and stages of clustering revealing diversity in 150 Pakistani rice (*Oryza Sativa* L.) accessions

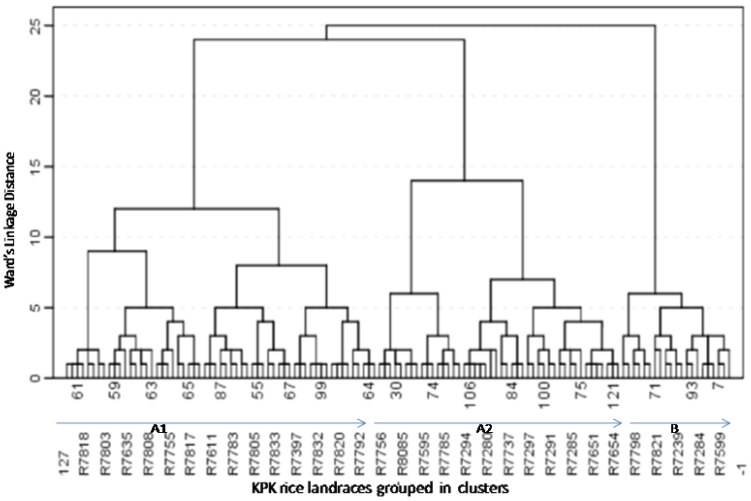


Figure 5: Dendrogram showing phylogenetic relationship and genetic diversity among 126 rice (*Oryza Sativa* L.) accessions of province Khyber Pakhtunkhwa (KPK) Pakistan based on SDS-PAGE

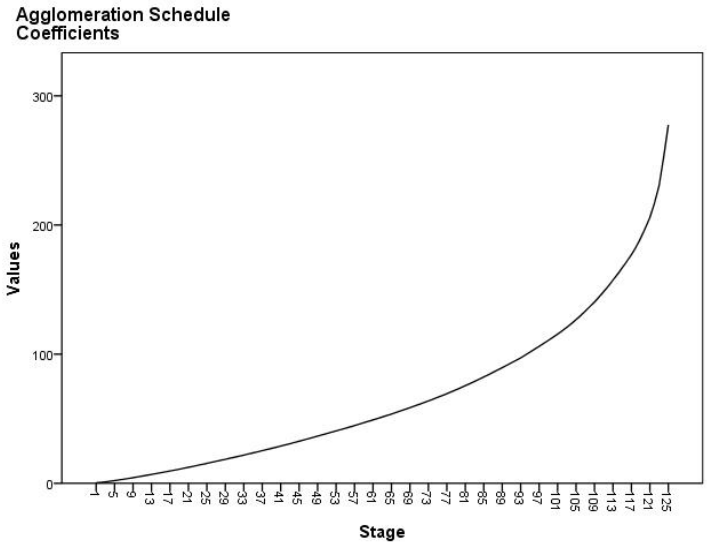


Figure 6 :A graph showing relationship between agglomeration schedule coefficients (values) and stages of clustering revealing diversity in 126 rice (*Oryza Sativa* L.) accessions of KPK province from Pakistan

Cluster analysis for the accessions of Province KPK

In addition to whole germplasm molecular polymorphism studies, phylogenetics studies for 126 accessions of KPK were also carried out and separately analyzed by cluster analysis.

A dendrogram was developed which divided the KPK Germplasm into two major clusters; the major cluster A consisted of 105 accessions and cluster B consisted of 21 accessions. Cluster A was further divided into two (A1, A2) sub-clusters with 59 (A1a; 26, A1b; 33) and 46 (A2a;16, A2b; 30) accessions while cluster B into two (B1, B2) sub-clusters with 06 and 15 accessions respectively. A check variety Fakhr-e-Malakand (FM) was grouped in A2b (Fig. 5 and Table 5 & 6). Pair-wise genetic distance calculated to comprehend the genetic diversity ranged from 0.185 (lowest) to 0.965 (highest) and similarity coefficient from 0.018 to 0.632, thus revealing intraspecific genetic diversity within the accessions of KPK.

DISCUSSION

In this study 150 diverse local rice (*Oryza sativa* L.) accessions originated from different regions of Pakistan were analyzed for genetic diversity using SDS-PAGE. Maximum number (126) of accessions was from province KPK, including a check variety Fakhr-e-Malakand. Fakhr-e-Malakand, a new cold tolerant and early maturing japonica rice variety was introduced in 2003 by the government of KPK, Pakistan, cultivated at more than 1000 meters altitude, and irrigated by snow fed rivers. It gives higher head rice (56.8%) on milling, higher quality index (1.33) and the best results in comparison to other rice varieties under the system of rice intensification (SRI) practices (Fayaz and Hamayoon, 2016). Fakhr-e-Malakand produced heavier grains and high yield as compared to other varieties when studied their sowing dates in relation to bacterial leaf blight (Rafi et al. 2013). The outstanding features of Fakhr-e-Malakand like higher yield, lodging and pests resistance, cold tolerance, along with more Benefit Cost Ratio of 3.24, made it most popular and economical commercial variety of the rice farming community (Ahmad et al. 2015).

In order to obtain of clear bands on the polyacrylamide gel, protein extraction method was first optimized. Three different protocols gave different results. One protocol gave denser bands and another sharp bands, on the basis of which, we proposed the use of second protocol with PBS pH 7.4 (Ranjan et al. 2012; Walczyk et al. 2017). Defatting seed meal with n-hexane improved protein extraction process and gave clear bands. Hirano et al. 2000; Walkzyk et al. 2017) reported the same results and found that defatting peanut meal with n-hexane resulted in significantly higher yields of crude proteins and allergens than other methods. Isotonic solutions like PBS is extensively used in the extraction of lectin proteins and food allergens which resulted in greater sensitivity of their molecular analysis than with Tris-Borate EDTA (TBS) or Tris. The yields of the compounds of interest were also affected by the composition of different buffers, their pH and other conditions.

One hundred and fifty local rice (*Oryza sativa* L.) accessions of Pakistan, analyzed for seed storage proteins using SDS-PAGE showed significant polymorphism. A total of 14-27 bands were found in each accession with size range from 10 to 250 kDa. Pervaiz et al. (2010-11) made one of the efforts to explore the genetic polymorphism of some Pakistan's landraces based on biochemical and molecular markers including Wx gene products (60kDa), glutelin, albumin, globulin and prolamine polypeptide subunits as well as RAPD fragments.

Shah et al. (2011) also reported 34 bands in different species of rice including *Oryza sativa* L. However, we have not seen 34 polypeptide bands in any accession as we have observed a maximum of 27 bands. Rice accessions originated from district Chitral showed more genetic variation as compared to other districts of the country (Bibi et al., 2017). Pakistan's rice germplasm has proved to be such a diverse that there can possibly be more than what is explored to date. After Shah et al., (2011), maximum number of polypeptide bands in various Pakistan's landraces has been reported by Habib et al. (2000), equal to 32, Pervaiz et al. in (2011), 25, followed by Asghar et al. (2004) who observed 22 bands and Bibi et al. (2017) reported overall 18 bands. Sharief et al. (2005) reported a minimum of 12 bands. The primary reason for variation in number of polypeptide subunits reported by other authors is the diversity of the rice germplasm of Pakistan from different geographical locations. Moreover, it can also be due to difference in protein extraction buffer as we have used three different compositions, and other lab protocols and gel composition for the evaluation rice seed proteins (Asghar et al. 2004).

No major relationship of geographical location and variation in seed protein profile was examined. The rice genotypes of China have been found relatively low in genetic diversity as compared to the landraces of the South Asian countries (Swain et al. 2017; Warusawithana et al. 2017).

A phylogenetic tree was developed for 150 rice accessions which divided the whole Germplasm into two major groups with 58 and 92 genotypes respectively, including check variety Fakhr-e-Malakand in second group (Ward and Hook, 1963).The Jaccard's dissimilarity coefficient (0.09 to 0.94) and similarity coefficient (0.009 to 0.635), revealed highest genetic diversity among the rice genotypes of Pakistan. The beneficial characteristics of FM are expected to be present in the same group of accessions on the basis of hierarchical clustering, in which it has been grouped due to less genetic distance (Ahmad et al. 2016). Similar work has been done in these references on *Oryza sativa* L. (Bibi et al. 2017). The accessions from different geographical zones were clustered together in same groups. This

demonstrates that expression profile is least affected by environment, a statement given by a range of studies on other cereal crops and rice. Some work also linked protein markers to the agronomic traits clustering in the same group which need to be explored further (Pervaiz et al. 2011). Cluster analysis for 126 accessions of KPK were also carried out and a cladogram was developed which divided the KPK Germplasm into two major groups consisting of 105 accessions and 21 accessions respectively including check variety Fakhr-e-Malakand in the first group (A2b). Pair-wise genetic distance (0.185-0.965) and similarity coefficient (0.018 to 0.632), was also found to express intraspecific genetic diversity within.

As FM is an improved variety with desirable characteristics, the same traits are also expected to be present in group A2b, due to lowest genetic distance with them (Ahmad et al. 2015; Ahmad et al. 2016). Perwaiz et al. (2011) reported a similarity coefficient of 0.81; this low genetic distance may be credited to the same genetic background, long-established farming practices and consumer preferences.

CONCLUSION

The findings in this work unveiled a considerable amount of genetic diversity in 150 landraces of the *O. sativa* L. from Pakistan and that of Province Khyber Pakhtunkhwa in particular. These facts offer an opportunity to be used for marker aided selection or breeding strategies and improvement of modern rice cultivars very low in genetic variability. The information obtained from clustering behaviour of accessions can be helpful to design plans for their management in the gene-bank. The desirable and rich traits in these landraces can also be identified and exploited in order to ensure food security and to cope with changing environmental conditions thus providing an efficient utilization and emphasizing on the conservation of germplasm resources of Pakistan for the betterment of humanity.

CONFLICT OF INTEREST

The authors declared that present study was performed in absence of any conflict of interest.

ACKNOWLEGEMENT

The authors gratefully acknowledge the cooperation of IABGRI-NARC, Islamabad, Pakistan for making available the seed material of all rice genotypes. We are also gratified to the Institute of Biotechnology and Genetic Engineering (IBGE), Khyber Pakhtunkhwa Agriculture University Peshawar for providing all the laboratory facilities.

AUTHOR CONTRIBUTIONS

HG conceived and designed research, HG and AJ conducted experiments while MAS, AJ and SP analysed the data. HG and MAN wrote the manuscript. All authors read and approved the manuscript.

Copyrights: © 2021@ author (s).

This is an open access article distributed under the terms of the **Creative Commons Attribution License (CC BY 4.0)**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

REFERENCES

Ahmad N, Ahmad Zada, AA, Junaid M, 2015. Effect of earthing up procedure on enhancement in yield of different groundnut varieties planted under agro-climatic conditions of malakand division. Journal of Agriculture and Environmental Sciences 4: 181-184.

Ali Z, Qureshi AS, Ali W, Gulzar H, Nisar M, Ghafoor A, 2007. Evaluation of genetic diversity present in pea (*Pisum sativum* L.) germplasm based on morphological traits, resistance to powdery mildew and molecular characteristics. Pakistan Journal of Botany 39: 2739-2747.

Asghar R, Siddique R, Afzal M, Akbar S, 2004. Inter and intra-specific variations in SDS-PAGE of total seed protein in rice (*Oryza sativa* L.) germplasm. Pakistan journal of Biological Sciences 7: 139-143.

Ashraf K, 2018. A comprehensive review on *Curcuma longa* Linn.: Phytochemical, pharmacological, and molecular study. International Journal of Green Pharmacy 11: 71-85.

Bibi A, Rabbani MA, Abbasi FM, Niaz IA, Bibi K, 2017. Assessment of Genetic diversity in indigenous rice accessions of northern Pakistan using biochemical markers. Pakistan Journal of Botany 49: 155-

161.

Bishwajit G, Sarker S, Kpoghomou MA, Gao H, Jun L, Yin D, Ghosh S, 2013. Self-sufficiency in rice and food security: a South Asian perspective. Agriculture and Food Security 2: 1-6.

Chen LF, Cheng MC, Chen SC, 1987. Similarity and diversity of seed proteins in rice varieties. Botanical bulletin of Academia sinica 28: 169-183.

Eckardt NA, 2000. Sequencing the rice genome. The Plant Cell 12: 3-9.

Fayaz A, Hamayoon K, 2016. Effect of different fertilizer treatments on the performance of some local rice varieties under SRI (system of rice intensification) and conventional management practices at district Swat. Pure and Applied Biology 5: 37-47.

Gul J, Munir I, Gulzar H, Hassan G, 2015. Agrobacterium-mediated transformation of *Spinacia Oleracea* L. through a synthetic chitinase gene inducing resistance to fungal pathogens. Pakistan Journal of Botany 47: 93-198.

Haque S, Roy P, Anandan A, Samantaray S, Pradhan SK, Singh ON, 2015. Genetic diversity study of seed proteins in rice drought tolerant donor accessions. Rice Genomics and Genetics 6: 1-5.

Hirano K, Teraoka T, Yamanaka H, Harashima A, Kunisaki A, Takahashi H, Hosokawa D, 2000. Novel mannose-binding rice lectin composed of some isolectins and its relation to a stress-inducible salT gene. Plant and Cell Physiology 41: 258-267.

Jamil M, Bashir S, Anwar S, Bibi S, Bangash A, Ullah F, Rha ES 2012. Effect of salinity on physiological and biochemical characteristics of different varieties of rice. Pakistan Journal of Botany 44: 7-13.

Khan SA, Shinwari Z K, Rabbani MA, 2013. Study of total seed protein pattern of rice (*Oryza sativa* L.) breeding lines of Pakistan through SDS-PAGE. Pakistan Journal of Botany 45: 871-876.

Laemmli UK, 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Mehfooza H, Wani SA, Zargar GH, 2000. Seed protein profile and isozyme polymorphism as markers for identification of some important rice cultivars. Applied Biological Research 2: 55-59.

Omonhinmin CA, Ogunbodede OO, 2013. Genetic diversity, taxonomy and legumins implications of seed storage protein profiling in Fabaceae. African Journal of Biotechnology 12: 2157-2163.

Pervaiz ZH, Tehrim S, Rabbani MA, Masood MS, Malik SA, 2011. Diversity in major seed storage proteins of rice landraces of Pakistan. Pakistan Journal of Botany 43: 1607-1612.

Rafi A, Hameed A, Akhtar MA, Akmal M, Bibi A, Ali M, Rahman H, Junaid M, 2013. Effect of planting dates on bacterial leaf blight incidence and yield performance of rice cultivars in different location of Khyber Pakhtunkhwa, Pakistan. Sarhad Journal of Agriculture 29: 407-414.

Rahman MM, Ansari A, Rashid MM, 2010. Diversity analysis in rice using GENSTAT and SPSS programs. The Agriculturists 8: 14-21.

Ranjan S, Matcha R, Madhuri B, Babu P, 2012. Comparative evaluation of protein extraction methods from few leguminous seeds. International Journal of Advanced Biotechnology and Research 3: 558-563.

Shah SMA, Rahman H, Abbasi FM, Rabbani MA, Khan IA, Shinwari ZK, Shah Z, 2011. Inter specific variation of total seed protein in wild rice germplasm using SDS-PAGE. Pakistan Journal of Botany 43: 2147-2152.

Simmonds NW, 1993. Introgression and incorporation. Strategies for the use of crop genetic resources. Biological Reviews 68: 539-562.

Sneath PH, Sokal RR, 1973. Numerical taxonomy. A Book of Systematic Zoology: The principles and practice of numerical classification, Ed 1 Vol 24. Taylor & Francis, Ltd, San Francisco, United states of America, pp 263-268.

Swain R, Mohapatra S, Roy P, Swain D, Singh ON, Meher J, Dash SK, Rao GJN, Subudhi HN, 2017. Assessment of genetic diversity in wild rice of eastern India using SSR markers. Journal of Agricultural Science 9: 239-250.

Tahir NAR 2014. Comparison of RAPD–PCR and SDS–PAGE techniques to evaluate genetic variation among nine barley varieties (*Hordeum* spp). Malaysian Applied Biology 43: 109-119.

Walczyk NE, Smith PM, Tovey ER, Roberts TH, 2017. Peanut protein extraction conditions strongly influence yield of allergens Ara h 1 and 2 and sensitivity of immunoassays. Food Chemistry 221: 335-344.

Ward JH, Hook ME 1963. Application of an hierarchical grouping procedure to a problem of grouping profiles. Educational and Psychological Measurement 23: 69-81.

Ward Jr, JH, 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58: 236-244.

Warusawithana TM, Samarasinghe WLG, Dassanayake PN, Ubeysekara NM, Jayarathna KGCN, 2017. Genetic diversity analysis of traditional rice variety 'Pachchaperumal' using SSR markers. International Journal of Scientific and Technology Research 6: 229-237.