

**MOLECULAR CHARACTERIZATION OF WHEAT
(*Triticum aestivum* L.) GENOTYPES THROUGH SSR MARKERS**

S. ISLAM¹, M. S. HAQUE¹, R. M. EMON²
M. M. ISLAM² AND S. N. BEGUM²

Abstract

A study was undertaken to examine the genetic diversity of 12 wheat (*Triticum aestivum* L.) genotypes, using 4 simple sequence repeats (SSRs). A total of 10 alleles were found. Allele number per locus ranged from 2 to 4 with an average of 2.5. The polymorphic information content (PIC) values ranged from 0.2755 to 0.5411 with an average of 0.3839. The average gene diversity over all SSR loci for the 12 wheat genotypes was 0.4688, ranging from 0.3299 to 0.6042. Cluster analysis based on microsatellite allelic diversity discriminated the varieties into different clusters. Genetic diversity was the highest between variety Gourab and Akbar as well as Gourab and BAW-1064, showing a genetic distance value of 0.4697. The genetic distance was lowest between Balaka and Aghrani as well as Triticale and BAW-1036. Positive correlations were found between gene diversity, number of alleles, the allele size range and the types of repeat motif of microsatellite markers. It was found from this study that microsatellite markers could characterize and discriminate all of the genotypes. More primers should be used for saturation of different regions in further studies.

Keywords: *Triticum aestivum* L., simple sequence repeats, genetic diversity, cluster analysis.

Introduction

Common wheat (*Triticum aestivum*) ($2n = 6x = 42$) belongs to family Poaceae, the most diverse and important family of the plant kingdom. It produces large edible grains and provides about one-half of humans' food calories and a large part of their nutrient requirements. Wheat (*Triticum* spp) is a worldwide cultivated and domesticated grass. The gradual increase in population demands a substantial increase in its productivity. Wheat has always been subjected to extensive and ceaseless research so as to maximize grain production but also to improve grain yield per unit area. However, there is still considerable room for improvement, especially to amplify efforts for continued genetic improvement of wheat to meet the growing requirements of an ever increasing population. Genetic manipulation is the best way to boost up wheat production. Therefore, it is necessary to estimate and study the genetic variation and mode of inheritance in different plant parameters to initiate productive wheat breeding programs. The

¹Dept. of Biotechnology, Bangladesh Agricultural University (BAU), Mymensingh,

²Plant Breeding Division, Bangladesh Institute of Nuclear Agriculture (BINA), Mymensingh, Bangladesh.

stagnant yield of wheat in different countries is due to limited diversity in the germplasm used in breeding programmes.

Through breeding and selection, great numbers of alleles have been lost, so that more difficulties have emerged for wheat improvement in modern agriculture systems (Allard, 1996; Hoisington *et al.*, 1999). Molecular markers can provide detailed characterization of genetic resources. Molecular markers provide a direct measure of genetic diversity and go beyond the indirect diversity measures based on agronomic traits or geographic origin. Microsatellites are simple sequence repeats (SSR) of 1-6 nucleotides. They are abundant, dispersed throughout the genome and show higher levels of polymorphism than other genetic markers. These features, coupled with their ease of detection, have made them useful markers. Their potential for automation and their inheritance in a co-dominant manner are additional advantages when compared with other types of molecular markers, and they cover all 21 wheat chromosomes. SSR markers have been used to characterize genetic diversity in wild relatives (Hammer *et al.*, 2000) and in a seed bank collection of improved wheat germplasm (Börner *et al.*, 2000; Huang *et al.*, 2002). This study was conducted to estimate the genetic divergence among twelve wheat genotypes as well as cultivated varieties of Bangladesh with the help of SSR markers. The present study addressed the utilization of microsatellite markers, to determine genetic diversity and relationships at the molecular level among 12 genotypes of wheat to help in the selection of parents to develop high-yielding varieties in breeding programmes.

Materials and Method

Plant materials collection

Seeds of 12 genotypes (Aghrani, Akbar, Ananda, Balaka, Gourab, BAW-1036, BAW-1064, Pavon-76, BAW-677, BAW-1035, Durum, and Triticale) were collected from the Wheat Research Centre of Bangladesh Agricultural Research Institute (BARI), Joydebpur, Gazipur.

The seedlings of 12 wheat genotypes were raised in pots at the glasshouse of Bangladesh Institute of Nuclear Agriculture on May in 2010. Young, vigorously growing fresh leaf samples from these seedlings were collected from 21 days old seedlings to extract genomic DNA. Initially, healthy portion of the youngest leaf of the tiller were cut apart with sterilized scissors and washed in distilled water and ethanol and dried on fresh tissue paper to remove spore of microorganisms and any other sources of foreign DNA. The collected leaf samples were then kept in polythene bags and for avoiding any damage of the leaf tissues the bags were placed in an ice box to carry it in Lab. and finally, the samples were stored in -80°C freezer.

DNA extraction and SSR analysis

Genomic DNA of 12 wheat genotypes was extracted from fresh leaves, using the mini preparation CTAB method as described by Khan *et al.* (2004). Following steps were followed for DNA extraction. The leaf sample were cut into 2-3 cm pieces and the sample was ground .670µl extraction buffer and 50 µl 20% SDS were added. Then the mixture was Vortexed for 20 second and incubated for 10 minutes at 65⁰C in hot water bath. 100 µl 5M NaCl was added and inverted gently to suspend the samples evenly. Then 100 µl CTAB was added and mixed well. Again, the mixture was vortexed for 20 seconds and incubated for 10 minutes at 65⁰C in hot water bath. 900 µl chloroform (chloroform: Isoamyl alcohol = 24:1, v/v) was added and mixed well. The samples were spinned down at 12000 rpm for 15 minutes. Then the supernatant was transferred into a new eppendorf tube and 600 µl Ice cold Isopropanol was added to the supernatant and was shaken well. The mixture was again spinned down at 12000 rpm for 15 minutes by centrifuge. The supernatant was discarded and the pellet was washed with 200 µl 70% ethanol. At last the sample was spinned down again at 12000 rpm for 5 minutes, the ethanol was removed and the pellets were allowed for air-drying for 1hour. The pellet was then suspended in 30 µl 1X TE buffer. Finally, the DNA samples were stored at -20⁰C.

Isolated genomic DNA contains a large amount of RNA and pigments, which cause over estimation of DNA concentration during spectrophotometer reading. Therefore, the DNA samples were evaluated both qualitatively and quantitatively using spectrophotometer and agarose gel electrophoresis, respectively.

For SSR analysis, a total of 4 screened primer pairs were used for polymerase chain reaction (PCR) amplification. PCR conditions were maintained as described by Roder *et al.* (1998). The amplification protocol as follows: 5 min at 94 °C, followed by 34 cycles with 1 min at 94 °C, 1 min 60 °C. Polymeraization at 72 °C for 2 min and a final extension step of 7 min at 72 °C.

Each PCR was carried out in a 15 µL reaction volume, containing 8.25 µL double-distilled deionized H₂O, 1.5 µL 10X buffer, 0.75 µL dNTPs (conc. 10 mM), 0.5 µL Taq polymerase (conc. 5 U/µl), 1 µL of each primer of a primer pair, and 2 µL DNA (conc. 25 ng/ µl).

Following steps were followed for Preparation of cocktail of SSR markers for each sample. At first, 8.25 µl of sterilized ddH₂O, 1.5 µl of 10 x buffer, 0.75 µl of dNTPs were taken in a 1.5 ml PCR tube. Then 1 µl of primer forward and 1 µl of primer reverse were added. The mixture was then vortexed. At last 0.5 µl of Taq DNA polymerase was mixed with it. Thus the PCR cocktail was ready to use. 2.0 µl of each template DNA samples were pipetted into the wells of the PCR tubes for PCR amplification.

The PCR products were electrophoresed on 3% agarose gels containing at 100 V for 2 h, and observed under a UV transilluminator.

Bands were counted and the presence and absence of bands were scored as 1 and 0, respectively. The data were collected and aligned for the construction of cluster analysis. The summary statistics including the number of alleles per locus, major allele frequency, gene diversity and polymorphism information content (PIC) values were determined using POWER MARKER version 3.23 (Liu and Muse 2005), a genetic analysis software.

Bootstrapping of the UPGMA tree was performed using POWER MARKER with 1,000 iterations followed by the PHYLIP CONSENSE module with the majority rule setting (Felsenstein, 1985). A model based cluster analysis was then performed using the programme STRUCTURE version 2.1 (Pritchard *et al.*, 2000).

The 12 wheat genotypes were clustered based on the matrix of genetic similarities using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA). The cluster analysis and dendrogram construction were performed with NTSYS-PC (version 2.1).

Results and Discussion

Using 4 SSR markers, a total of 10 alleles were detected among 12 wheat genotypes. The average number of allele per locus was 2.5 with a range of 2 (Xgwm192, Xgwm642 and Wmc296) to as many as 4 (Xgwm533) (Table 1). Similar observation was reported by Salem *et al.* (2008), where they found that the number of alleles per locus ranged from 2 alleles to 7 alleles with an average of 3.2 alleles per locus. Jain *et al.* (2004) also reported similar observation of number allele per locus ranged from 3 to as high as 22 with an average of 7.8 alleles per locus.

Table 1. No. of alleles, no. of rare alleles, gene diversity (GD) and PIC value found in 12 wheat genotypes for 4 SSR markers.

Locus	Repeat type	Alleles number	Rare alleles	Allele size and ranges (bp)	Difference (bp)	Gene diversity	PIC
Xgwm192	(CT)46	2		43-149	106	0.4965	0.3733
Xgwm533	(CT)18(CA)20	4	1	137-513	376	0.6042	0.5411
Xgwm642	(GT)14	2		157-195	38	0.4444	0.3457
Wmc296	(GT)28	2		49-172	123	0.3299	0.2755
Mean		2.5	0.25			0.4688	0.3839

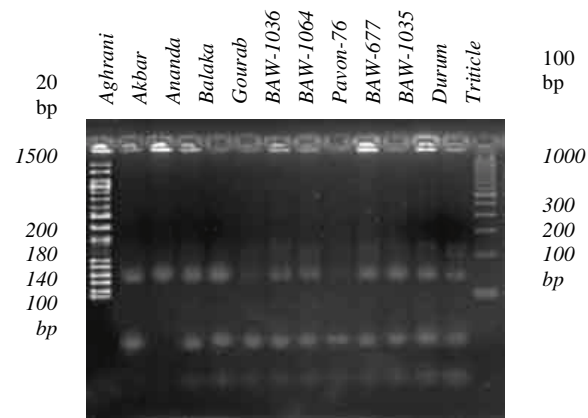


Fig. 1. Microsatellite profiles of 12 wheat genotypes at loci Xgwm192.

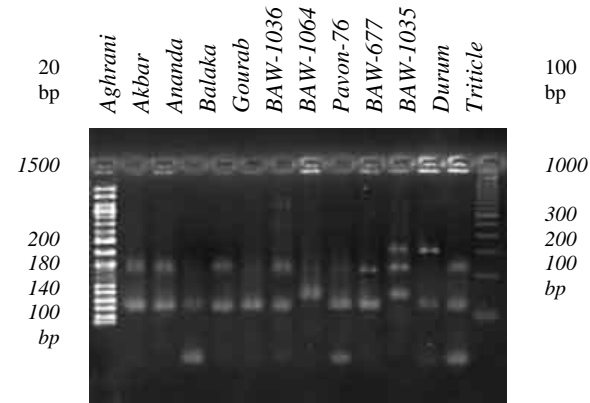


Fig. 2. Microsatellite profiles of 12 wheat genotypes at loci Xgwm-533.

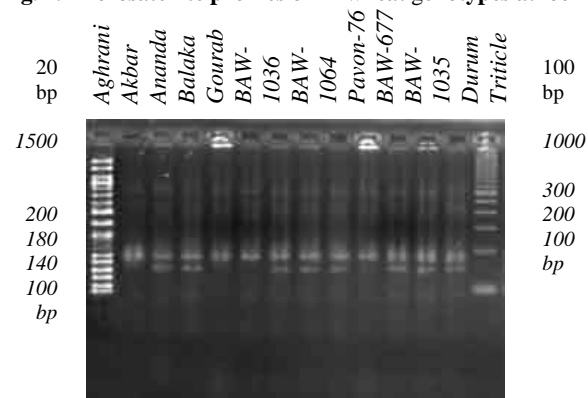


Fig. 3. Microsatellite profiles of 12 wheat genotypes at loci Xgwm-642.

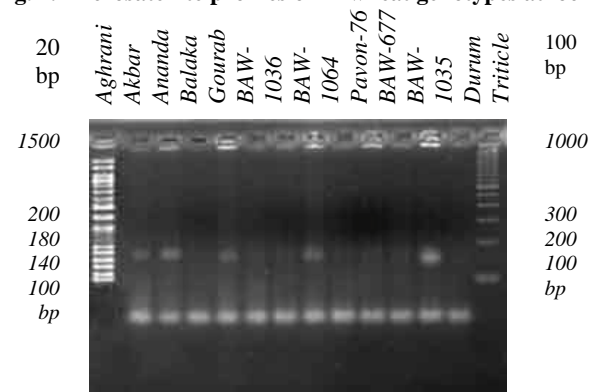


Fig. 4. Microsatellite profiles of 12 wheat genotypes at loci WMC 296.

Table 2. Genetic distance of wheat genotypes based on 4 microsatellite alleles.

OTU	Aghrani	Akbar	Ananda	Balaka	BAW-1035	BAW-1036	BAW-1064	BAW-677	Duram	Gourab	Pavon-76	Tritic-ale
Aghrani	*****											
Akbar	0.1464	*****										
Ananda	0.2197	0.2197	*****									
Balaka	0	0.1464	0.2197	*****								
BAW-1035	0.2714	0.2714	0.25	0.2714	*****							
BAW-1036	0.1464	0.1464	0.0732	0.1464	0.125	*****						
BAW-1064	0.3232	0.3232	0.3232	0.3232	0.1464	0.3232	*****					
BAW-677	0.0732	0.2197	0.1464	0.0732	0.1982	0.0732	0.3964	*****				
Duram	0.1982	0.1982	0.1464	0.1982	0.3232	0.1982	0.25	0.2714	*****			
Gourab	0.2197	0.4697	0.1464	0.2197	0.3964	0.2197	0.4697	0.1464	0.2929	*****		
Pavon-76	0.2929	0.3964	0.0732	0.2929	0.3232	0.1464	0.3964	0.2197	0.2197	0.0732	*****	
Triticale	0.1464	0.1464	0.0732	0.1464	0.125	0	0.3232	0.0732	0.1982	0.1464	0.1464	*****

There was a relationship between the number of alleles detected at a locus and the types of simple repeats within the targeted microsatellite DNA. Thus the larger variation in the repeat motif in the microsatellite DNA, the larger the number of alleles detected. A positive correlation exists between PIC value and the variation in the repeat motif per microsatellite markers. The range of repeat motif $\{(CT)_{18}(CA)_{20}\}$ of Xgwm533 marker is much wider than the other markers and detect the larger number (4) of alleles than the other markers. The PIC value (0.5411) of Xgwm533 is also higher than other markers.

In this study, 12 genotypes of wheat were analyzed using 4 primer pairs (Xgwm192, Xgwm533, Xgwm642, and WMC296). Amplified microsatellite loci were analyzed for polymorphism using agarose gel electrophoresis (Fig.1-4) and the result revealed that all the primer pairs detected polymorphisms among the wheat genotypes analyzed.

According to Nei's (1972), the highest level of genetic diversity value (0.6042) was observed in loci Xgwm533 and the lowest level of genetic diversity value (0.3299) was observed in loci Wmc296 with a mean diversity of 0.4688 (Table 1). It was observed that marker detecting the lower number of alleles showed lower gene diversity than those detected higher number of alleles which revealed higher gene diversity. This result is in consistent with previous work done by Herrera *et al.* (2008), who also observed that the genetic diversity at each SSR locus was significantly correlated with the number of alleles detected, type and range of repeat motif of microsatellite markers and with the allele size range. In this study, we observed that the number of alleles (4) detected by Xgwm533 marker, the range repeat motif $\{(CT)_{18}(CA)_{20}\}$ and the allele size range (137-513) of Xgwm533 marker is higher than other markers. So, the genetic diversity (0.6042) of this marker is the highest which is consistent with the work done by Herrera *et al.* (2008).

As a measure of the informativeness of microsatellites, the PIC values ranged from a low of 0.2755 (Wmc296) to a high of 0.5411 (Xgwm533) and an average of 0.3839 (Table 1), which is close to the result revealed by Zheng *et al.* (2009). PIC also showed a significant, positive correlation with the number of alleles and allele size range for microsatellites evaluated in this study. The allele size range and the number of alleles were themselves also highly correlated.

The values of pair-wise comparison of Nei's (1972) genetic distance (D) between varieties were computed from combined data for the 4 primers, ranged from 0.00 to 0.4697 (Table 2). Comparatively higher genetic distance (0.469) was observed between a number of variety pair. Among them, Gourab vs Akbar and BAW-1064 pair was very important. The higher genetic distance between them indicates that genetically they are diverse compare to lower genetic distance value. Basically this value is an indication of their genetic dissimilarity. Variety

pair with higher value is more dissimilar than a pair with a lower value. The lowest genetic distance (0.000) was found in Balaka vs Aghrani and Triticale vs BAW-1036 variety pairs indicating that they are genetically similar to each other. The average genetic distance among the 12 wheat genotypes was quantified as 0.212.

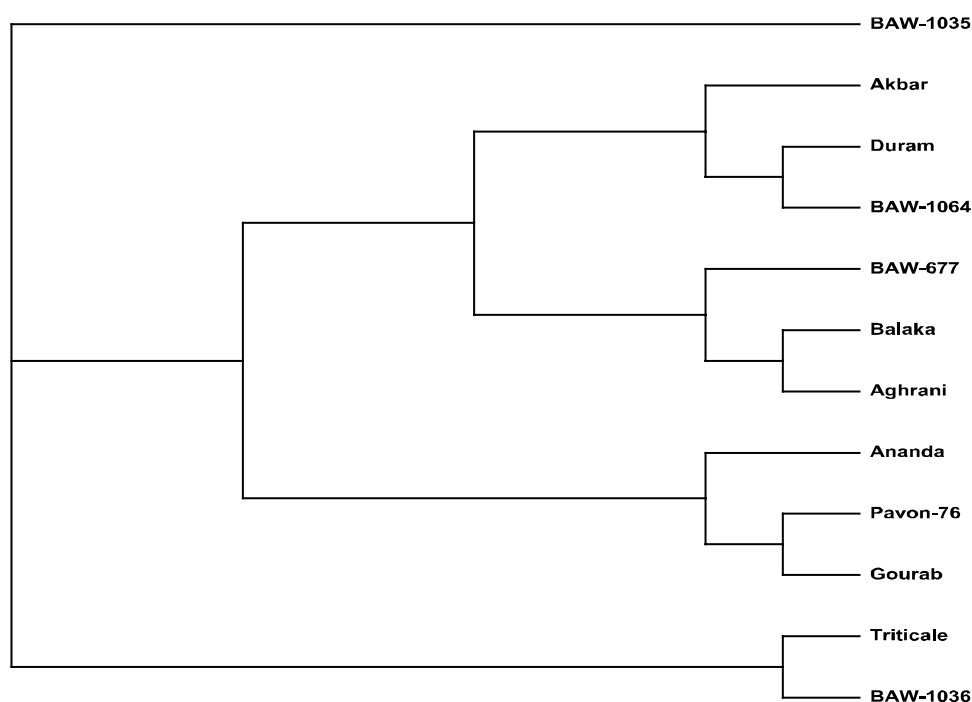


Fig. 5. Dendrogram for 12 wheat germplasms showing the genetic similarity derived from a UPGMA cluster analysis using the Dice similarity coefficient.

A dendrogram was constructed based on the Nei's genetic distance calculated from 10 alleles generated from 12 wheat varieties. All 12 wheat cultivars could be easily distinguished. The UPGMA cluster tree analysis led to the grouping of the 12 wheat varieties in three major clusters (Fig. 5)

The largest cluster comprised of 9 wheat varieties among selected 12 wheat varieties. The fewer advanced line included in this cluster are BAW-1064 and BAW-677 (Fig. 5).

Some popular wheat varieties included in this group are Gourab, Pavon-76, Ananda, Aghrani, Balaka, and Akbar. In this cluster, Balaka and Aghrani were grouped in the same sub-cluster and showed no genetic distance (0.000) (Table 2). The second largest cluster constituted of Triticale and BAW-1036, collected from BARI (Bangladesh Agricultural Research Institute). These two varieties showed their unique origin by showing no genetic distance value (0.000) between them.

In this study, we observed that there is no genetic distance between Balaka and Aghrani, Triticale, and BAW-1036. They also grouped in the same sub-cluster and cluster, respectively. This finding can be explained in such way that, while molecular fingerprinting can be used to distinguish varieties from one another and it can also suggest the possibility that two varieties with different names may be genetically identical. The variety may have been cultivated in different areas of the countries for many years a long time ago. This leads to the possibility that the same variety may be cultivated under a different name in different areas. This may be a case where the same variety is being cultivated under different names. Otherwise, it might be due to the reason that in this study fewer SSR markers were used and by these few markers genomic coverage was not possible.

Several additional sub-clusters were observed within the largest cluster. Durum and BAW-1064, Balaka, and Aghrani, Pavon-76, and Gourab form an individual sub-cluster with each other which remain in the largest cluster. BAW-1035, individually form a single main cluster. The dendrogram revealed that the genotypes that derivatives of genetically similar type clustered together. Ijaz and Khan (2009) classified the 63 wheat genotypes into 03 clusters. Cluster analysis of the 07 wheat varieties showed two major clusters and three sub-clusters (Salem *et al.*, 2008).

Conclusion

The PIC values ranged from a low of 0.2755 (Wmc296) to a high of 0.5411 (Xgwm533) and averaged 0.3839. PIC values also showed a significant positive correlation with the number of alleles and allele size ranged for microsatellites evaluated in this study.

In this study, 4 SSR markers were evaluated in 12 wheat genotypes. All 4 SSRs were polymorphic across the 12 wheat genotypes. A total of 10 alleles were detected with an average number of alleles of 2.5 per locus (range 2-4 per locus). The genetic distance based analysis using Neighbour-Joining clearly separated 12 wheat genotypes into four distinct clusters and eight sub-clusters. The genotypes belonging to the same cluster are genetically more or less similar and the genotypes belonging to the different clusters are genetically different from each other.

Among the various DNA based markers currently available, the set of microsatellite markers, used in the present study, provided a positive assessment of the ability of SSR markers in producing unique DNA profiles and establishing discrete identity of wheat genotypes, which otherwise, were not possible using morphological traits. The present investigations have helped to establish clear-cut identity of all the cultivars under consideration, which will be of great utility for the protection of Plant Breeder's Rights.

References

- Allard, R.W. 1996. Genetic basis of the evolution of adaptedness in plants. *Euphytica*. **92**: 1-11.
- Borner, A., S. Chebotar, and V. Korzun. 2000. Molecular characterization of the genetic integrity of wheat (*Triticum aestivum* L.) germplasm after long-term maintenance. *Theor. Appl. Genet.* **100**: 494-497.
- Felsenstein, J. 1985. Confidence limit on phylogenies: an approach using the bootstrap. *Evol.* **39**: 783-791.
- Hammer, K., A.A. Filatenko and V. Korzun. 2000. Microsatellite markers - a new tool for distinguishing diploid wheat species. *Genet. Resour. Crop Evol.* **47**: 497-505.
- Herrera, T.G., D.P. Duque, I.P. Almeida, G.T. Nunez, A.J. Pieters, C.P. Martinez and J.M. Tohme. 2008. Assessment on genetic diversity in Venezuelan rice cultivars using simple sequence repeats markers. *Electron. J. Biotechnol.* **11**(5): 215-226.
- Hoisington D, M. Khairallah, T. Reeves, J.M. Ribaut, *et al.* 1999. Plant genetic resources: what can they contribute toward increased crop productivity? *Proc. Natl. Acad. Sci. U. S. A.* **96**: 5937-5943.
- Huang, Q., A. Borner, S. Roder and W. Ganal. 2002. Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theor. Appl. Genet.* **105**: 699-707.
- Ijaz, S. and I. A Khan. 2009. Molecular characterization of wheat germplasm using microsatellite markers. *Genetics and Molecular Research* **8** (3): 809-815.
- Jain, S., R.K. Jain and S.R. McCouch. 2004. Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently-labeled microsatellite markers. *Theor. Appl. Genet.* **109**: 965-977.
- Khan I.A, F.S. Awan, A. Ahmad and A.A. Khan. 2004. A modified mini-prep method for economical and rapid extraction of genomic DNA in plants. *Plant Mol. Biol. Rep.* **22**: 89a-89e.
- Liu, K. and S.V. Muse. 2005. Power Marker: Integrated analysis environment for genetic marker data. *Bioinformatics* **21**: 2128-2129.
- Nei, M. 1972. Genetic distance between populations. *American Naturalist* **106**: 283-292.
- Pritchard, J. K., M. Stephens and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.
- Roder, M.S., V. Korzun, K. Wendehake, J. Plaschke, M.H. Tixier, P. Leroy and M.W. Ganal. 1998. A microsatellite map of wheat. *Genetics* **149**: 2007-2023.
- Salem, K.F.M., A.M. El-Zanaty and R.M. Esmail. 2008. Assessing Wheat (*Triticum aestivum* L.) Genetic Diversity Using Morphological Characters and Microsatellite Markers. *World J. Agril. Sci.* **4** (5): 538-544.
- Zheng, M., Ding, H. and Xiao, Z. 2009; Analysis of genetic diversity among Heilongjiang spring wheat cultivars in recent decades based on SSR markers. *J. Triticeae Crops* **29**(3): 409-413.