Analysis of Genetic Diversity and Population Structure of Spring Bread Wheat (*Triticum aestivum* L.) from Bangladesh Using SSR Markers

Md. Ashraful Alam[1]*, Md. Abdul Hakim[1], Muhammad Rezaul Kabir[1], MSN Mandal[1], Akbar Hossain[1]

[1]Bangladesh Wheat and Maize Research Institute, Nashipur, Dinajpur 5200, Bangladesh;

*Correspondence: e-mail: ashrafulw@yahoo.com

Abstract

Analysis of genetic diversity and population structure of crop cultivars are the key to design a successful breeding program. In this context, genomic diversity and population structure of 30 bread wheat genotypes were assessed by using 10 simple sequence repeat (SSR) markers. After observation, 50 alleles were recognized from these genotypes, among them, 3 to 9 alleles were observed in each locus with a mean hereditary lushness of 5.0. Generally, markers selection of genomic studies, the informativeness of genotypes can be quantitatively detected by the polymorphism information content (PIC). In the study, the PIC of all genotypes fluctuated from 0.32 to 0.86. Among these utilized markers, the Xgwm186 marker showed the lowest number of alleles (3 alleles) locus^{-1}, but the maximum PIC value (0.86). In the case of the population structure analysis, all wheat genotypes were grouped into two reconstructed panmictic populations (RPPs). The group RPP I consisted of 17 genotypes of which 15 genotypes were pure and two were admixed. The group RPP II consist of a total of 13 genotypes out of which one variety was admixed. Based on the direction of data entrance, the tree by using Neighbor joining (NJ) method generally shows >2 topologies of "tie trees" from a single source of data; however in the current study NJ method showed two groups of genotypes. Among these groups, Group-I consisted of 21 and Group-II consisted of 09 genotypes. Among these studied genotypes, the maximum genetic similarity (94%) was detected in genotypes Barkat and Ananda (similar 47 loci), followed by genotypes Akbar and Inia (92%) and BARI Gom 32 and BAW 1203 (92%). In the case of the principal component analysis (PCA), four clusters were identified among the genotypes by which was similar to the results of STRUCTURE analysis. This study clearly depicted that microsatellite markers are promising in the characterization and differentiation of wheat genotypes, which provides a foundation for the selection of parental material for the genetic improvement of wheat.

Introduction

Genetic diversity is the primary requirement to initiate a successful breeding programme for enhancing wheat productivity. Enriching germplasm collection with 'superior' gene/alleles is an essential part of regional genetic improvement programs including Bangladesh. Due to limited germplasm collection in wheat breeding program in Bangladesh and subsequent selection of cultivars with desirable traits have further narrowed the genetic diversity over time. Microsatellite markers are useful for their multi-allelic and chromosome specificity. SSR marker have already been utilized for identifying specific genes, QTLs and marker-assisted selection in wheat. SSR markers are effective tool for diversity studies. SSR markers have additional benefits because of their high rate of polymorphism i.e., high PIC value, co-dominant nature, selective neutrality, genome-wide distribution, and cost and labour efficiency. At present, SSR markers are established as one of the most effective molecular markers for inter- or intra-specific genetic differentiation. The SSR markers have major applications over extremely mutable and manifold-allelic PCR-constructed heritable markers as they are ubiquitously distributed within the eukaryotic genomes.

Materials and Methods

Plant materials
Thirty wheat genotypes including 29 genotypes (released by Bangladesh Wheat and Maize Research Institute (BWMRI)) were used in this study.

Genomic DNA extraction
Ten seeds of each wheat genotype were grown in pots and DNA isolation was performed from 15 days old wheat seedlings using the CTAB method with some modifications. Ten SSR markers (xcfd18, xwmc667, xgwm46, xgwm44, xgwm533, xgwm642, xgwm111, xgwm369, xgwm108 and xgwm164) were randomly selected representing different locations of chromosomes.

PCR amplification and electrophoresis separation
PCR amplifications were performed in 10 μL volume using a Verity Thermal Cycler (Apliedbiosystems, USA). The total volume of reaction mixture was 10μL which contained 20 to 40 ng DNA, 10 mM/L Tris-HCl, 2 mM/L MgCl2, 50 mM/L KCl, 200 μmol/L of each dNTPs, 250 nM/L of each primer and 0.25 U Taq DNA polymerase. The PCR was run following a program of initial cycle of 95° C for 3 min; 35 cycles of 94° C for 1 min, 50 to 60° C for 1.5 min and 72° C for 1 min; and a final extension at 72° C for 10 min. The PCR products were mixed with 2μL loading buffer. Finally, gels were silver stained and photographed.

Statistical analysis
Every band was considered as a single locus. All the scorable loci were considered for the generation of bivariate 1-0 data matrix and genetic distances (GD) among the genotypes were estimated using Unweighted Pair Group Method with Arithmetic Mean (UPGMA). MetaboAnalyst software was used to construct the dendrogram for the evaluation of genetic diversity. To assess the degree of genotypes concentration changes (heat maps) and PCA clustering were performed using the MetaboAnalyst 2.0 software. Heatmaps were created based on the Pearson distance. Principal component analysis (PCA) serves as a platform to provide a spatial illustration of the comparative genetic distances between the populations. It also assesses the robustness of the differentiation among the groups classified by the dendrogram. A model-based Bayesian procedure implemented in the program STRUCTURE v. 2.3.3 was used to estimates structure in the group of wheat genotypes

Results and Discussion

SSR marker-based DNA fingerprinting analysis
A total of 10 microsatellite primers were used in this study. The genomic multiplicity of 30 bread wheat genotypes was studied by using 10 simple sequence repeat (SSR) markers. Fifty alleles were detected from all cultivars, where 3 to 9 alleles were observed in each locus with a mean hereditary lushness of 5.0. The polymorphic microsatellite marker Xwmc667 showed the maximum number of alleles (9 alleles), but the PIC value was estimated to be 0.60 (Figure 1), followed by Xgwm 642 (7 alleles). However, the marker Xgwm186 showed the lowest number of alleles (3 alleles) locus^{-1}, but the maximum PIC value (0.86).

Population structure of wheat genotypes
For population genetic structure analysis, Bayesian clustering modelling was executed in the STRUCTURE software using data generated by SSR marker-based profiling and diversity analysis of 30 wheat accessions . As the clustering model presumes the underlying existence of K clusters, Ln (PD) derived ΔK was schemed in contradiction of the K to regulate the numeral of reconstructed panmictic populaces (RPPs). The highest log-likelihood was performed which provided the maximum peak at K=2, proposing the existence of 2 key RPPs, indicating that all genotypes represent two distinct clusters. The structure analysis was performed by locating the range of the promising number of groups (K) since 2 to 14. In structure analysis, RPP I consisted of 56.7% of genotypes (17 genotypes: 'Kylansona', 'BARI Gom 31', 'BARI Gom29', 'Sonora', 'BAW 1203', 'Protiva', Sourav, 'BARI Gom 30', 'BARI Gom 32', 'Aghrani', 'Pavon', 'Sonalika', 'Kanchan', 'BARI Gom 25', 'BARI Gom 21', 'BARI Gom 23', 'Balaka'), where 15 were pure with qi > 0.80 (80%) and 2 were admixed with qi < 0.80. RPP II comprised a total of 13 genotypes (Barkat, Ananda, Akbar, BARI Gom 24, BAW 1194, BARI Gom 26, Inia, Khery, BARI Gom 22, BARI Gom 27, BARI Gom 28, Gorob, BARI Gom 33), among which only one variety was found admixed with qi < 0.80 (Figure 2).

Genetic diversity among the genotypes
Based on allelic information among the wheat genotypes, a dendrogram was constructed using Neighbor joining (NJ) method (Figure 3). The cluster analysis naked two major clusters with a similarity coefficient ranged from 0.0 to 0.61 indicating significant genetic variation among the wheat accessions studied. Cluster one consisted of 21 genotypes, while cluster 2 contained 9 wheat genotypes. The tow clusters separated into four sub-clusters which subcluster-1 comprised of 5 genotypes and subcluster-2 comprised of 16 genotypes, whereas subcluster-3 and subcluster-4 comprised of 7 and 2 genotypes, respectively (Figure 3). The highest genetic similarity was observed between Barkat and Ananda (similar 47 loci), while BARI Gom 26 and BARI Gom 28 showing a minimum genetic similarity (Figures 3 and 4).

Two-dimensional scaling obtained using principal component analysis (PCA) analysis also showed the same grouping pattern as dendrogram and STRUCTURE analysis. In PCA scatterplots, the first three principal components had Eigenvalues of 19.3, 13.3 and 9.7%, and a total cumulative dissimilarity of 42.3% were recorded with the initial three components of PCA (Figure 5). Similar to the result obtained in UPGMA-tree, genotypes were clearly separated by PCA and found 4 distinct clusters with considerable genetic differences among the genotypes. The Sub-clustering presented in different colours in Figure 3 corresponds with Figure 5 (red, green, blue and light blue).
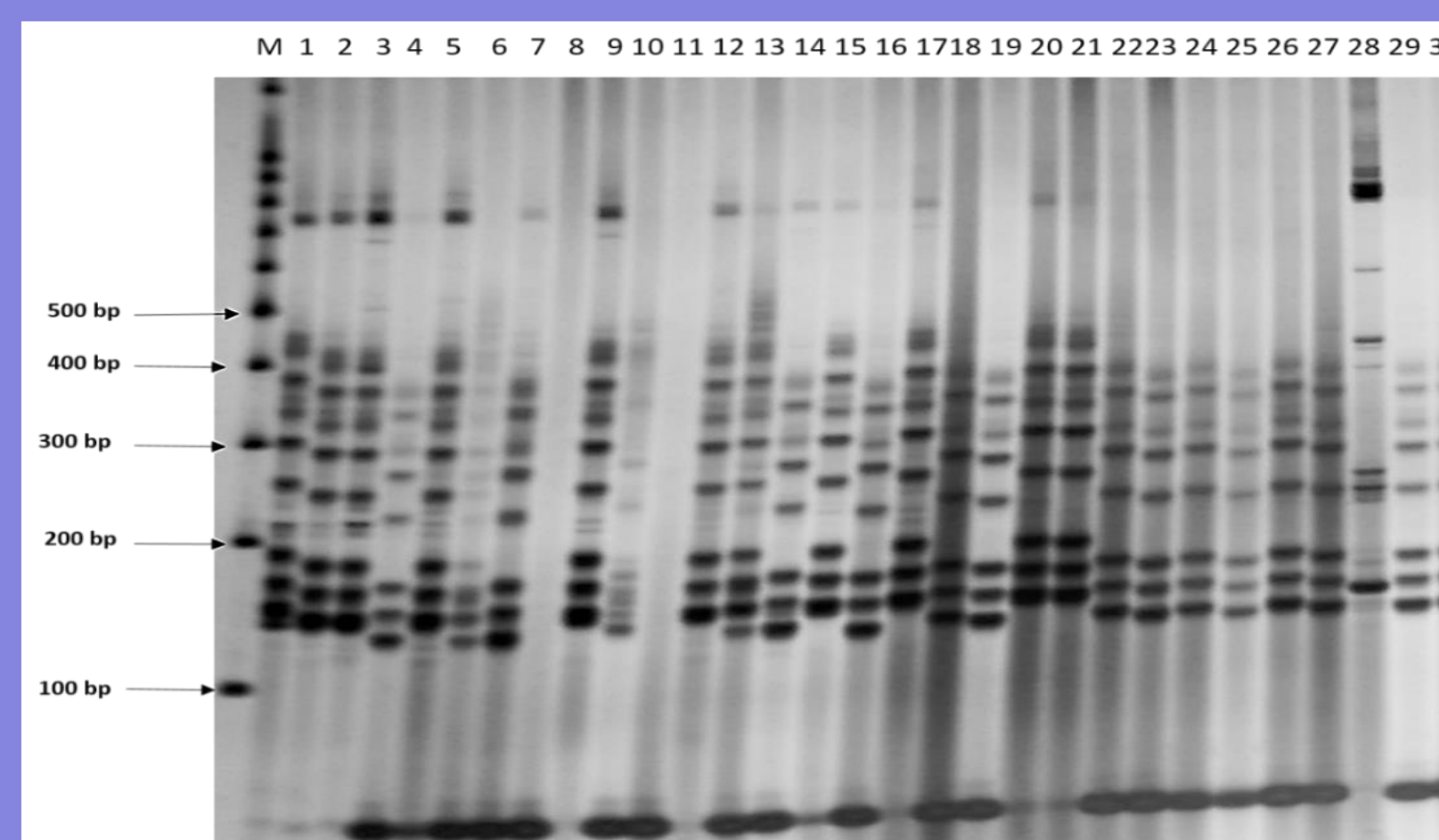

Figure 1. SSR profile of 30 wheat genotypes with Xwmc 667 in PAGE gel M, 100 bp marker. (As seen, the marker showed different polymorphism patterns within wheat genotypes. Legend: 1-Khery, 2-Sonora 64, 3-Kalyansona, 4-Inia 66, 5-Sonalika, 6-Balaka, 7-Pavon 76, 8-Ananda, 9-Kanchan, 10- Akbar, 11-Barkat, 12-Aghrani, 13-Protiva, 14-BARI Gom 19 (Sourav), 15-BARI Gom 20 (Gourab), 16-BARI Gom 21 (Shatabdi), 17-BARI Gom 22 (Sufi), 18-BARI Gom 23 (Bijoy), 19-BARI Gom 24 (Prodip), 20-BARI Gom 25, 21-BARI Gom 26, 22-BARI Gom 27, 23-BARI Gom 28, 24-BARI Gom 29, 25-BARI Gom 30, 26-BARI Gom 31, 27-BARI Gom 32, 28BARI Gom 33, 29-BAW 1194, 30-BAW 1203)
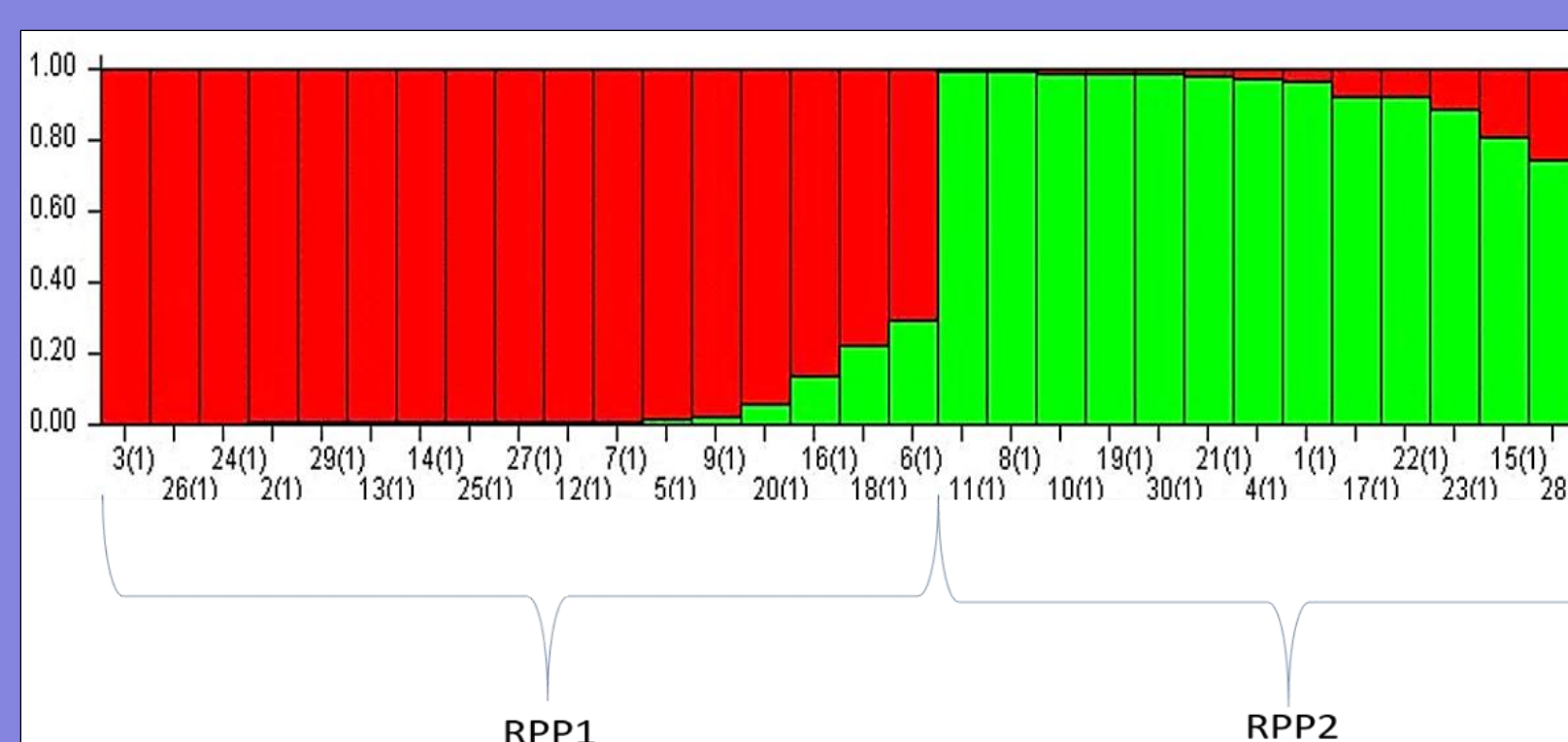

Figure 3. Molecular phylogenetic tree of the 30 wheat genotypes based on the banding pattern of the 10 SSR markers using Neighbor joining method.

Figure 5. Principal coordinate analysis of 30 wheat genotypes based on SSR markers

Figure 2. Model-based population structure plot for each genotype with K=2, using Structure software program with SSR marker data. Color codes are as follows: Reconstructed panmictic population (RPP) I (red), RPP II (green). The Code of each variety corresponds to the description in Figure 1.
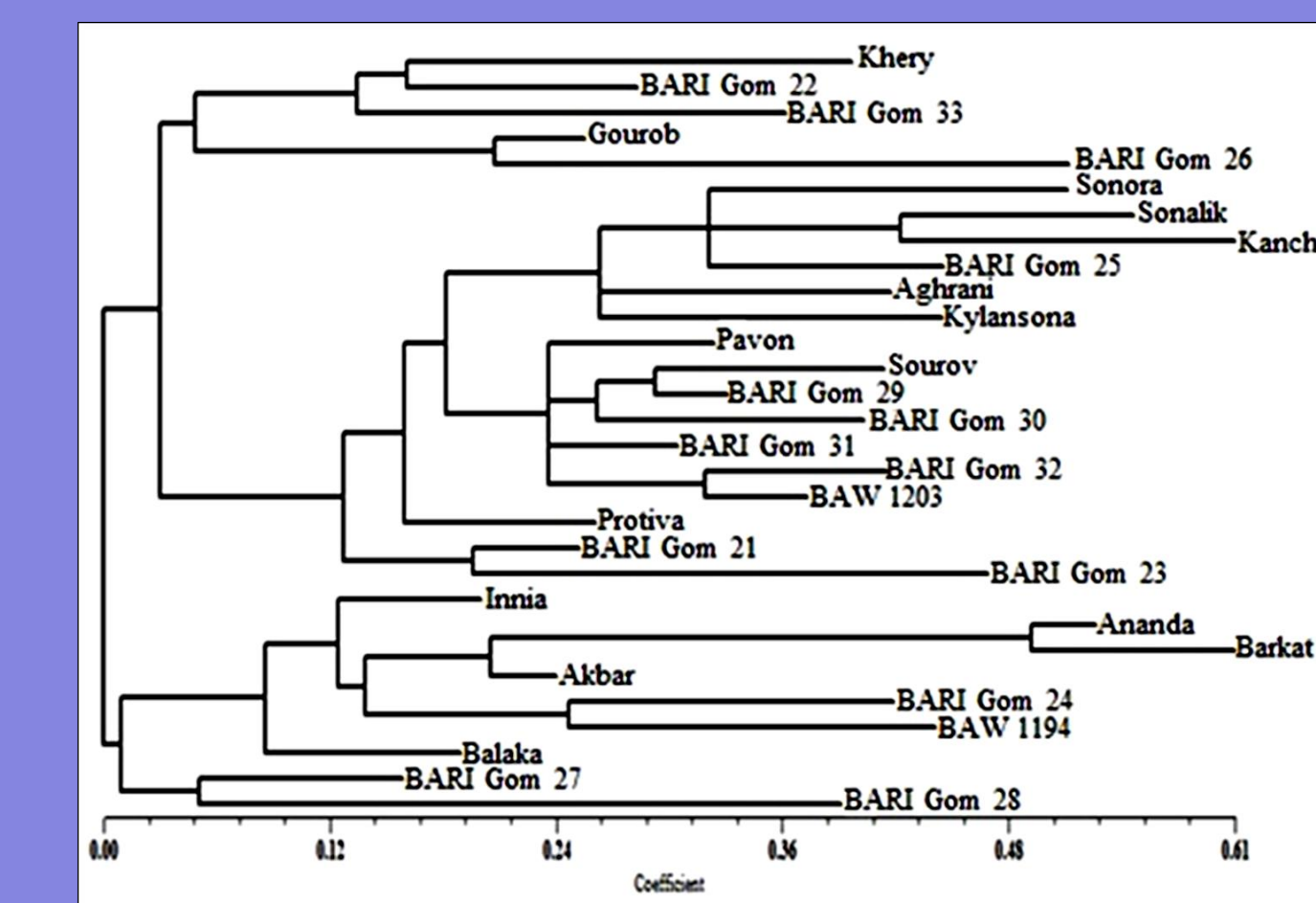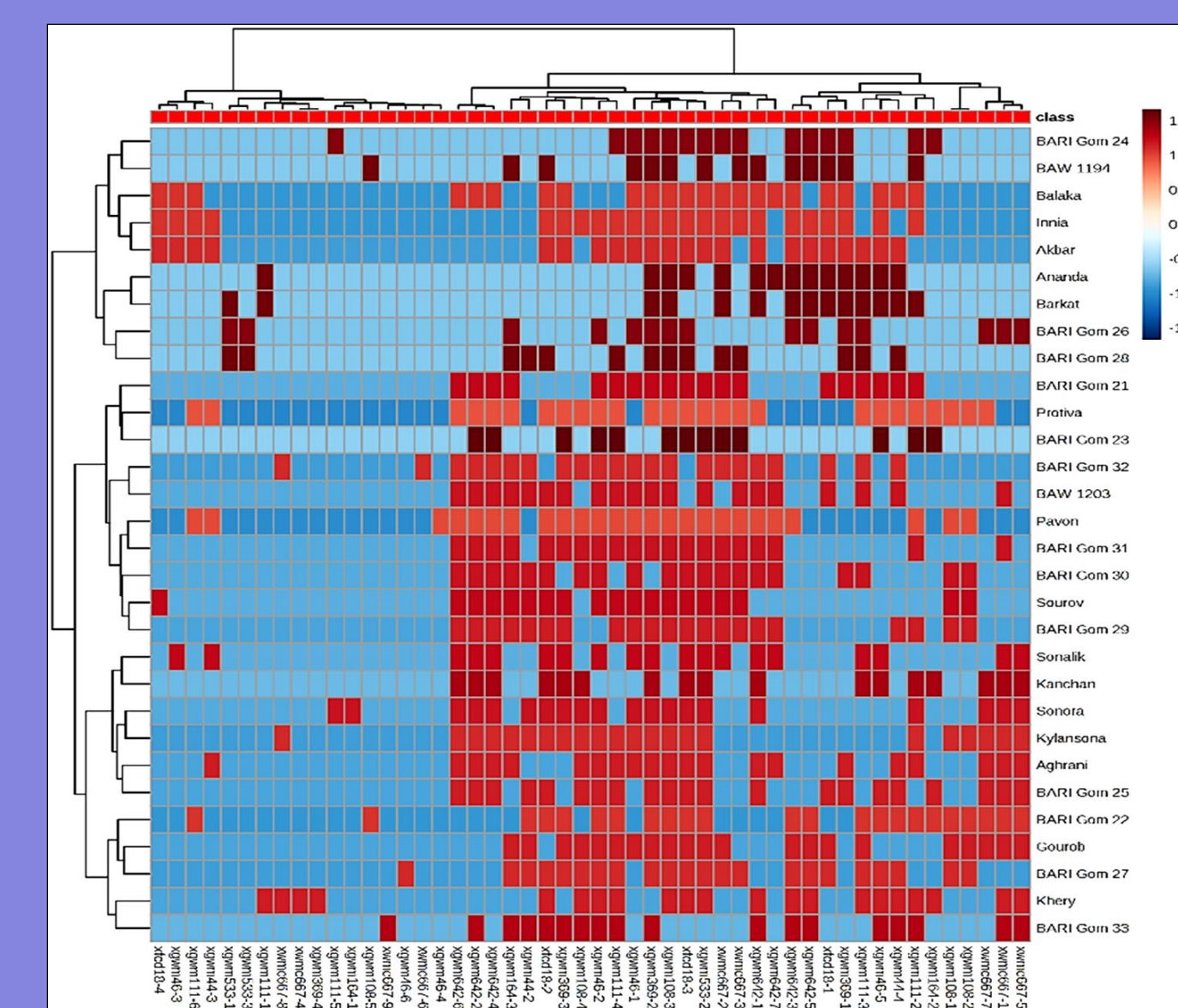

Figure 4. The Heatmap imagining and hierarchical clustering-analysis using Metabo-Analyst's data annotation tool constructed based on the different genotypes for 50 alleles. Rows: Genotypes; Columns: SSR markers; Color key indicates variety value, blue: lowest, red: highest.
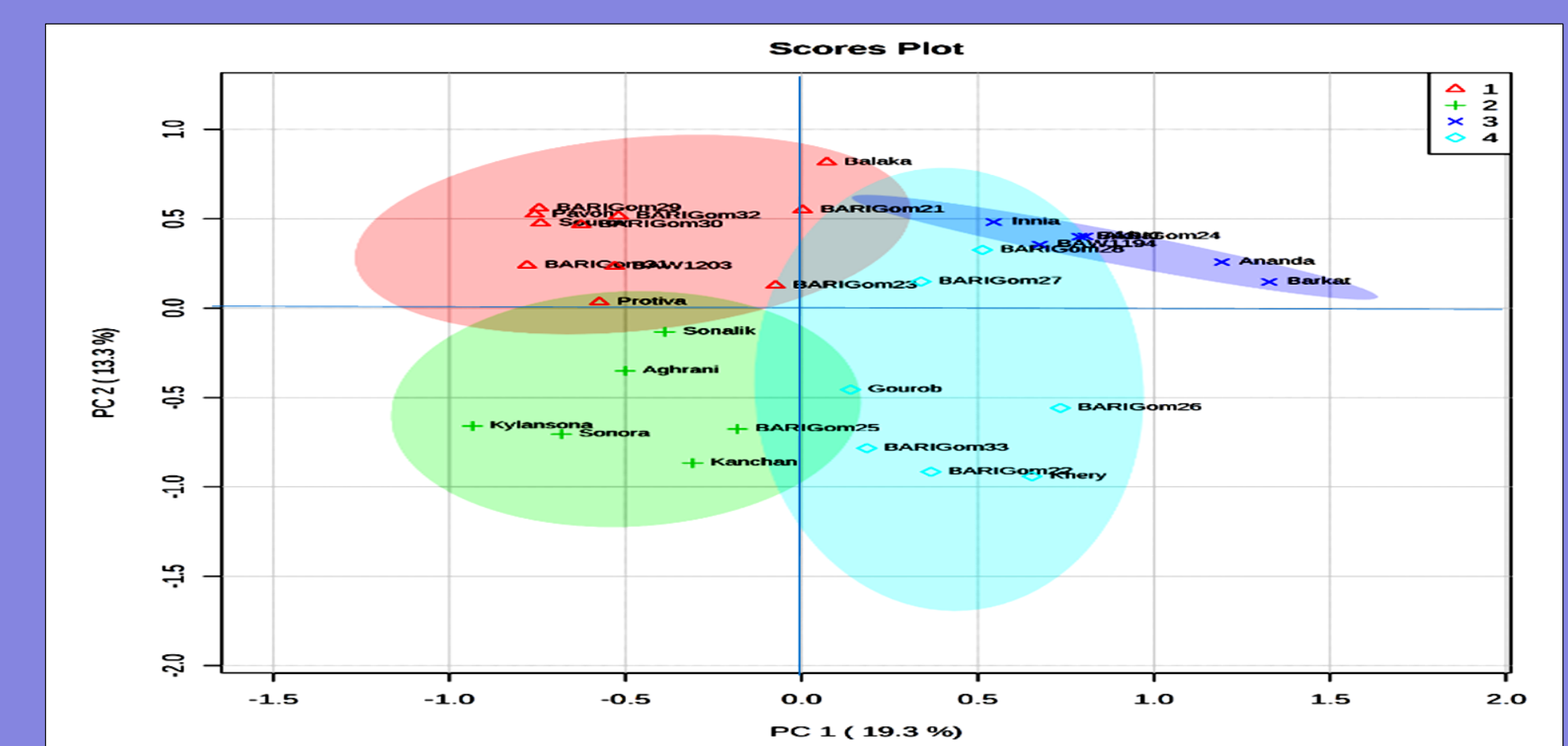
Conclusion

The present study evaluated the variability of 30 genotypes including popular and historic wheat genotypes and advanced lines using in the breeding program of BWMRI. Wheat genotypes were characterized using chromosome-specific SSR markers. A total of 10 SSR markers were used to discriminate the genotypes. Our data showed significant variation in microsatellite DNA polymorphisms among different wheat genotypes tested. Our results reconfirmed that microsatellites used for the genetic diversity studies in 30 wheat genotypes were found to be highly polymorphic, genome-specific, abundant and co-dominant thus, serve as promising genetic markers in wheat breeding programs. Genetic similarity and dissimilarity among the genotypes would be a source of information that could be useful in utilizing these genotypes in the breeding program. However, for designing an efficient breeding program, more genotypes could be included in the diversity study.