Genetic variability and population structure analysis of Bulgarian bread wheat germplasm as revealed by SNP markers

Vladimir Aleksandrov[1], Tanya Kartseva[1], Krasimira Tasheva[1], Konstantina Kocheva[1], Andreas Börner[2], Svetlana Misheva[1]

[1]Institute of Plant Physiology and Genetics, Bulgarian Academy of Sciences Sofia, Bulgaria

[2]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Germany

Introduction

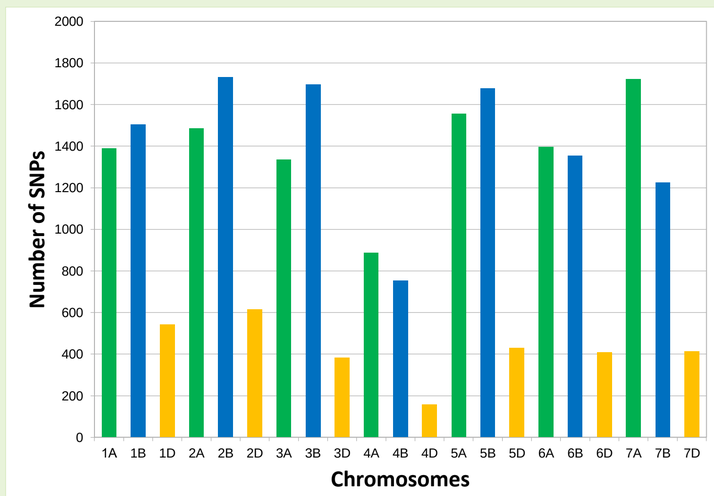
Information on the amount of genetic variation in a crop is fundamental to its improvement, for accurate registration, identification, control, and protection of cultivars. It is also an important consideration for efficient conservation and utilization of genetic resources in breeding programmes. In this study, we analysed the genetic diversity and population structure of a collection of 184 Bulgarian bread wheat accessions using single nucleotide polymorphism (SNP) markers.

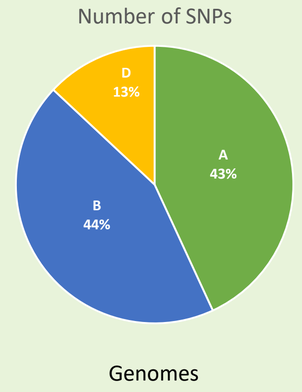
Results

Single nucleotide polymorphism distributions (Fig. 1)

The distribution of identified SNPs on the 21 wheat chromosomes ranged from 159 on chromosome 4D to 1733 on chromosome 2B; the A and B genome had almost equal amount of SNPs.

Fig. 1





Population heterozygosity

The statistically expected level of heterozygosity in the population, expressed by the Wright's "fixation index" F was 0.17. The percentage of heterozygosity was 62.4, and the percentage of homozygosity 37.6, $\chi 2 = 0.09$.

Population structure (Fig. 2, 3)

The number of subgroupings was highest at K=5, denoting 5 recognizable groups (Fig. 2a). The majority of landraces formed a distinct cluster, whereas the old and modern cultivars were structured according to the releasing breeding center/geographic region and formed sub-populations (Fig. 2b). When applying principal component analysis (PCA) to the data, the 184 genotypes formed 184-dimensional space and the identified subgroupings were similar to those elucidated by the K means clustering algorithm (Fig. 3). In PCA, the two principal components explained 12.2 % of the total variation (PC1 8.6 % and PC2 3.6 %).

Fig. 2a

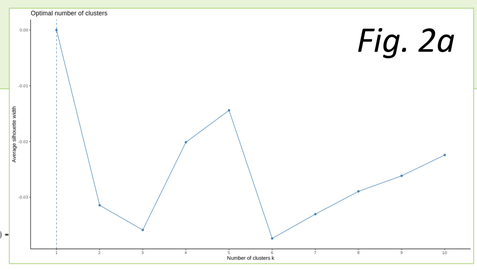
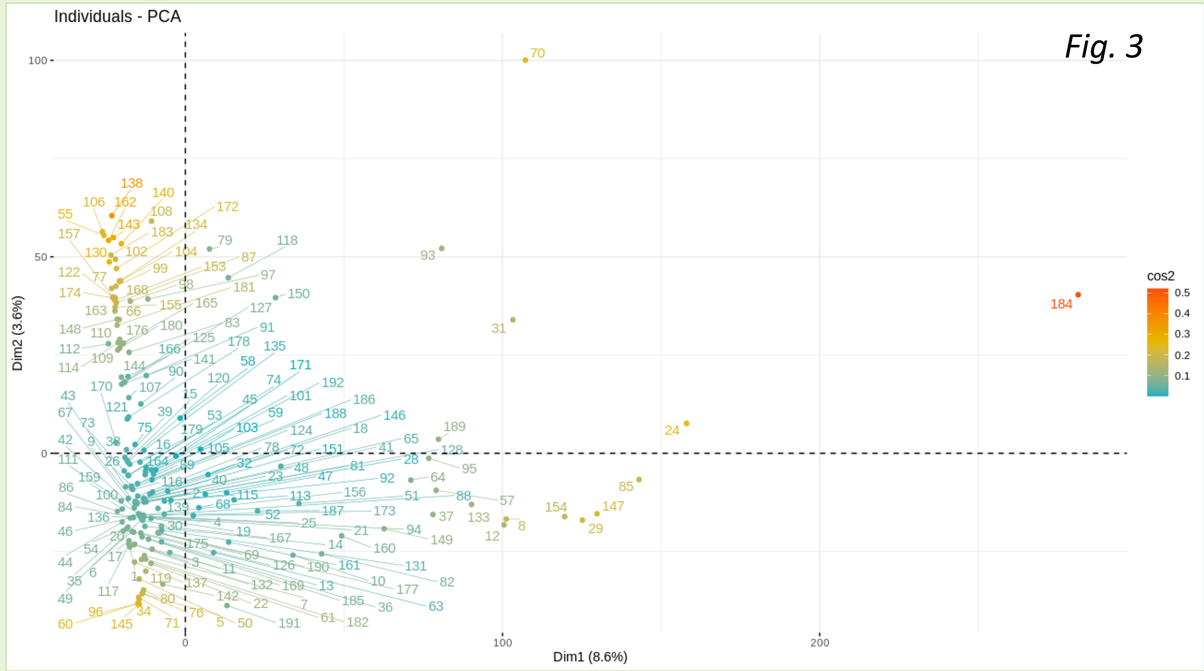


Fig. 2b

Fig. 3



Experimental material and methods

- 184 Bulgarian wheat (*Triticum aestivum* L.) accessions, including landraces, old traditional cultivars and modern releases.
- For whole-genome genotyping, an optimized 25K wheat Illumina Infinium array (TraitGenetics GmbH, Germany) was used, resulting in 24 145 scorable SNP markers.
- Population structure analysis was performed by the K means clustering algorithm in package R.
- PCA analysis in package R was performed to determine subpopulations.

Conclusions

- The study identified **24 145 SNPs**, distributed as follows:

  44% in B genome, 43% in A genome, and 13 in D genome;
- The pool of Bulgarian wheats contains **17 % fewer heterozygotes** than would be expected in a population that is in Hardy–Weinberg equilibrium;
- **Five distinct clusters** were identified. The landraces formed the most divergent group.

FACULTY OF BIOLOGY SOFIA UNIVERSITY

SCIENTIFIC CONFERENCE KLIMENT'S DAYS (KD 2020)
FACULTY OF BIOLOGY
SOFIA UNIVERSITY St. KLIMENT OHRIDSKI
5 November, 2020

BULGARIAN SCIENCE FUND

Acknowledgements. The study was supported by the Bulgarian NSF (contract KP-06-N31/17).