Exploiting old wheat landraces from the Carpathian Basin in present breeding

Project ID: NKFI-K-129221

Final report

Wheat (*Triticum aestivum* L., ABD, 2n=6x=42) stands as Hungary's most crucial food crop and a major export commodity of Hungarian agriculture. The challenge of securing Hungarian food supplies and wheat production capabilities while preserving essential natural resources and biodiversity must be addressed. Concurrently, the impact of climate change, characterized by rising temperatures and shifts in rainfall patterns and intensity, is already affecting agriculture, presenting an additional obstacle to efficient wheat production. Expanding the genetic diversity of cultivated wheat will directly impact food security and the sustainability of future agriculture. Genetic variation serves as the foundation for the crop's adaptation to any forthcoming challenges. Studies on collections of bread wheat landraces reveal a high degree of genetic diversity and suggest a rich source of alleles that have not been utilized in modern breeding (Huang et al. 2002; Horvath et al. 2009; Wingen et al. 2014). The Hungarian wheat landrace collection from the Carpathian Basin offers a substantial and largely untapped reservoir of genetic variation that can be harnessed for wheat improvement.

The objective of this study was to deliver comprehensive genotyping and phenotyping data for the entire landrace collection, which includes 200 landraces, alongside 70 modern wheat cultivars. Genotyping was conducted using a high-density 20K Illumina SNP genotyping array with the utilization of 17,905 gene-based SNP probes. We have analysed the geographical distribution of genetic variability and the distribution of polymorphic markers along the chromosomes. During the phenotyping work, the main emphasis has been placed on traits that have significantly contributed to ecological adaptation, biotic and abiotic stress tolerance, and end-use quality.

Objectives

I.) Genotyping and phylogenetic analyses of the Hungarian landrace collection in parallel to modern wheat cultivars

We analysed the population stucture, the genetic variability and the distribution of polymorphic markers along the chromosomes. Neighbor-Joining (NJ) tree analyses and Principal Component Analysis (PCA) were conducted to delineate the population structure of the wheat collection, using a set of 15,784 SNP markers.

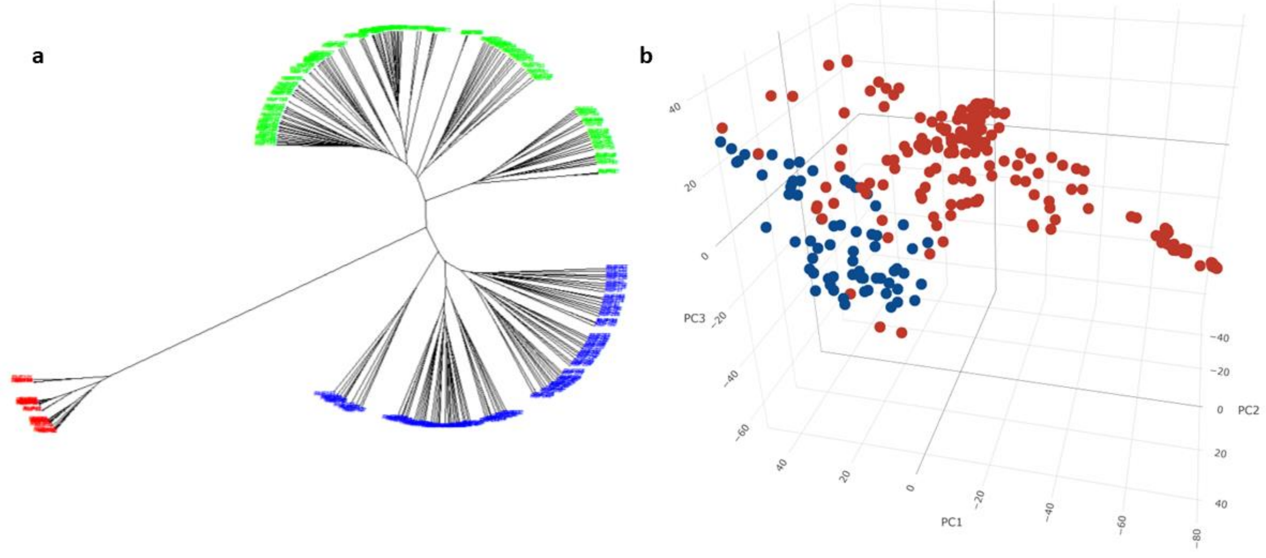


Figure 1. Genetic structuring of wheat accessions. a) An unrooted Neighbor-Joining (NJ) tree of wheat accessions. The tree branches are colored based on the genetic grouping. The blue group consists 67 modern cultivars and 30 landraces; there are 129 landraces within the green group and the red group encompasses a total of 37 landraces. b) The 3D PCA plot showing the relationship between the accessions belonging to the landraces (red dots) and modern varieties (blue dots).

The Neighbor-Joining (NJ) tree analysis divided the population into three main groups (blue, green, red), and these groups effectively illustrate the historical source and origins of the accessions (Figure 1a). The blue group comprises modern varieties and those local landraces that likely participated in breeding programs. The green group is the largest and most diverse group, containing landraces from all countries investigated; however, they likely did not participate in crossing programs for the production of modern cultivars. Genetically, the red group is the most distant, and this group primarily consists of Hungarian landraces, as only four of the 37 accessions originate from other countries. The PCA analysis captures the primary sources of variation within the genomic data, revealing distinctions among the sampled regions and illustrating variation within groups of accessions. . Interestingly, mostly the Austrian and former Yugoslavian landraces grouped together with the modern varieties in the left corner of the PCA plot (Figure 1b), these landraces might have had a greater impact on the genetic structure of the modern wheat cultivars.

The 20 K array was constructed from SNPs that were previously mapped in the wheat genome, enabling us to assess their distribution and allele frequencies among the 21 wheat chromosomes and seven homoeologous groups (Fig. 2a–c). The detected polymorphisms were unevenly distributed along the wheat genome, considering the size of the chromosomes (Figure 2).

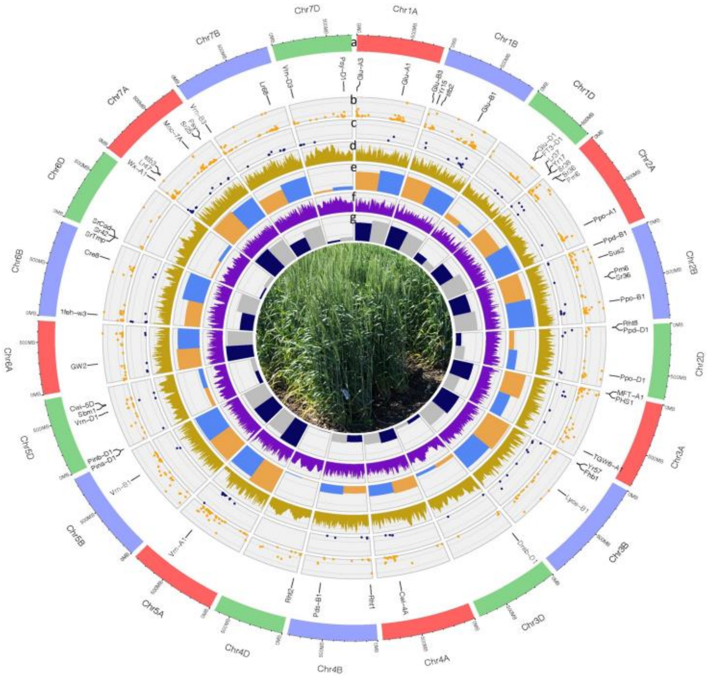


Figure 2. Circos diagram showing the physical map of twenty-one wheat chromosomes. (a) The physical scale (Mbp) of the A (red), B (blue) and D (green) genome chromosomes of wheat. Chromosomes are displayed by size proportioned bars in the outer circle. The position of genes encoding important agronomic traits are marked under the bars (see text for further details). (b) The chromosome position and frequency of polymorphic SNPs (yellow dots) found in Central European landraces compared to modern Western varieties. (c) The chromosome position and frequency of polymorphic SNPs (dark blue dots) found in landraces compared to modern varieties. (d) Nucleotide diversity (π) of Central European landraces compared to modern varieties in 3 Mb window intervals (light brown). (e) Bar chart showing the total number of SNPs compared to the 'MV Ménrót' reference in modern varieties (orange) and landraces (light blue). (f) Nucleotide diversity (π) of Central European landraces in 3 Mb window intervals (purple). (g) Marker density of polymorphic SNPs based on the 'MV Ménrót' reference in modern varieties (grey) and landraces (dark blue). Tracks are marked alphabetically (a-g) from top to bottom.

A higher number of polymorphisms was concentrated on the homoeologous group 2 of landraces when the genetic origin of the accession was considered. In the case of group 4, 6 and 7, only chromosomes of the B and D genomes were more diverse in landraces compared to the modern elite varieties. On the contrary, the A genome from all homoeologous groups except 1 and 2 showed more polymorphisms in modern varieties. Mapping the positions of polymorphic markers along several important genes of great agronomic importance indicated that novel polymorphisms co-localize with these genes (Fig. 2). These included reduced height genes (*Rht1*, *Rht2*), photoperiod response genes (*Ppd-B1*, *Ppd-D1*), as well as several genes associated with disease resistance (*Lr, Sr, Yr, Pm*), suggesting that they might have provided the basis of selection during the breeding programs.

This suggests that these landraces possessed many novel polymorphisms compared to modern varieties. Thus, the Central European collection could be a good source to improve the genetic diversity of the elite varieties.

II.) Phenotypic characterizations, cataloguing agronomically useful traits and GWAS

Phenological and morphological traits and end-use quality

We characterised the ecological adaptation of the landraces and those of the wheat cultivars through measuring phenological and morphological traits under field conditions. During the growing seasons, detailed phenological and morphological characterisations of the genotypes were carried out and significant differences have been found between landraces and modern wheat cultivars. At harvest, 6 ears per genotypes were sampled for evaluating yield components from all the genotypes (Figure 3), and then the micro plots were harvested for evaluating the yield and basic quality parameters (Thousand-kernel weight on the main spike, grain number of the main spike, grain weight of the main spike, seed number/spikelet, length of the main spike, grain shape and Hardness, ZELE). We identified more than one hundred marker trait associations with great interest for wheat breading. We present two examples as the results of association mapping. We measured great phenotyping variation while analysing the length of the main spike (Figure 3).

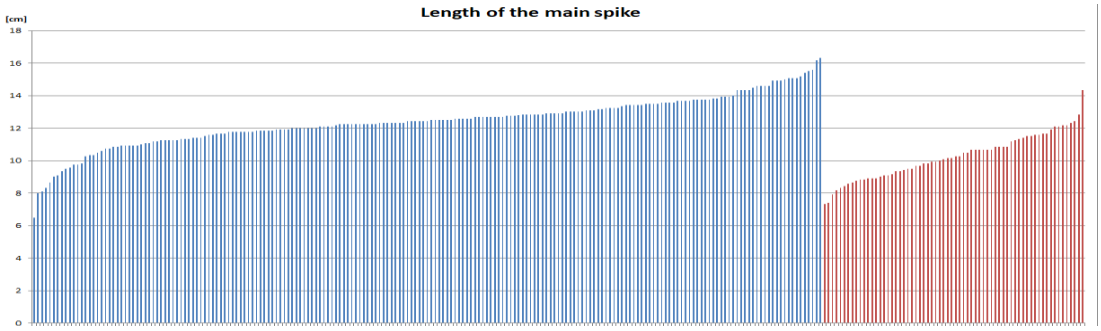


Figure 3. Diagram of the length of the main spike collected in three growing seasons in Martonvásár. Blue lines represent the 200 landraces and red lines show the 70 modern wheat cultivars.

We performed GWAS to identify the genetic background of this trait. We identified one significant marker (AX-95126447) trait association on the 2AS chromosome arm (Figure 4).

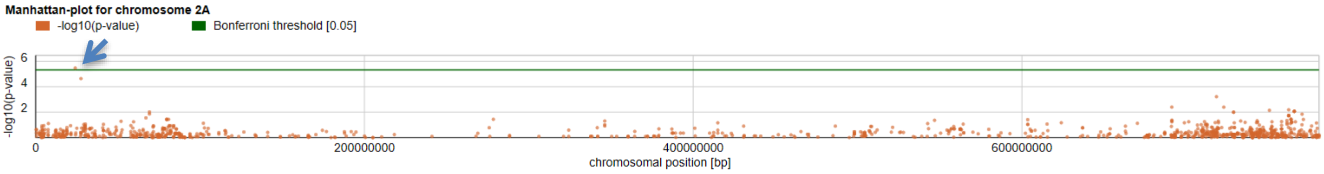


Figure 4. Manhattan-plot of significant SNPs associated with length of the main spike. Marker-trait association analyses were conducted using the easyGWAS software, where EMMAX test was fitted based on SNP data. We applied a GWAS threshold of the Bonferroni correction to declare significant associations.

The allele and phenotype information were helped to identify the favourable allele for breeders (Figure 5).

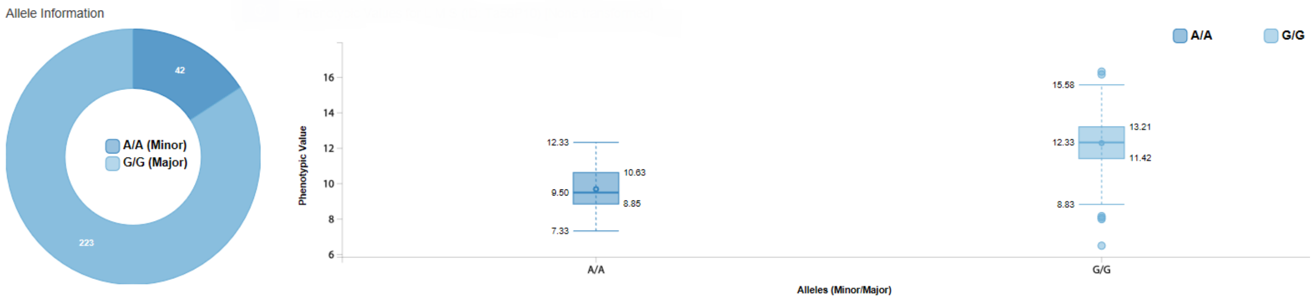
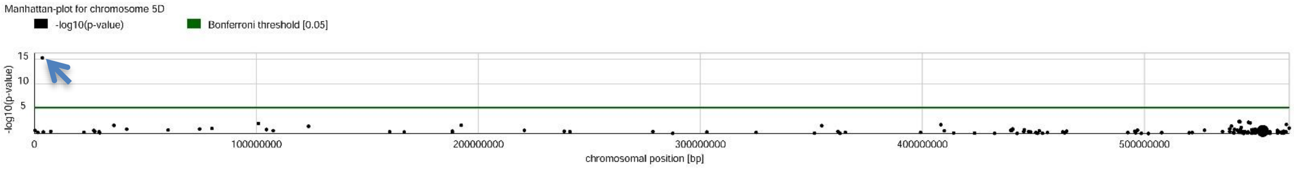


Figure 5. Allele and phenotype information of AX-95126447 SNP marker.

We used the online Ensembl Plants database (https://plants.ensembl.org) to identify the linked wheat gene. This marker is represented within the TraesCS2A02G057100.1 gene encoding an uncharacterized protein with ATPase, ATPase-coupled transmembrane transporter and ATP binding activity. During the marker-assisted selection (MAS) breeding process this marker could be useful to select the lines with longer spike phenotype.

In the first half of the 20th century Hungarian wheats were known worldwide for their special quality parameters. The seed hardness strongly correlates with the bread making quality parameters. Harder seed usually means better end use quality. We selected landraces with height seed hardness and very good bread making quality parameters to improve the modern wheat cultivars by traditional crossing and MAS. For MAS we identified two SNP markers with the help of GWAS (Figure 6) and we found that they represent two alleles of the *Pinb* gene (*Pinb-A1*, *Pinb-D1*) placed on 5DS chromosome arm.

(a)



(b)

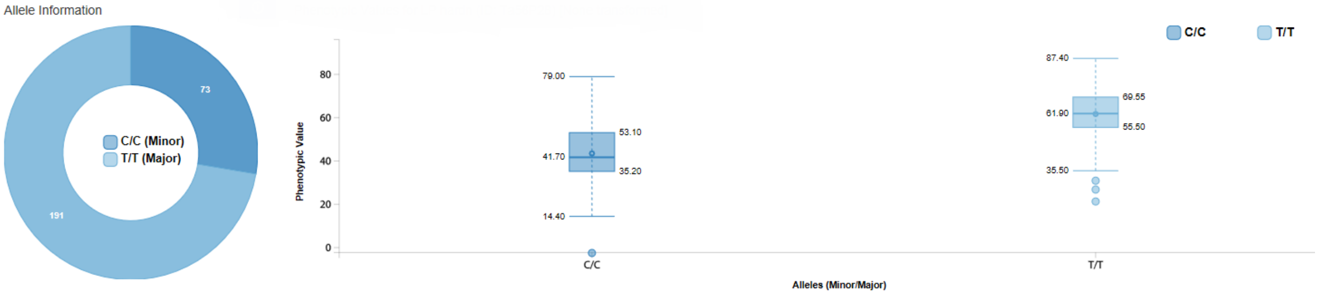


Figure 6. (a) Manhattan-plot of significant SNPs associated with length of the main spike. Marker-trait associations analyses were conducted using the easyGWAS software, where EMMAX test was fitted based on SNP data. We applied a GWAS threshold of the Bonferroni correction to declare significant associations. (b) Allele and phenotype information of BS00000020_51 and TG0028 SNP markers.

Biotic stress tolerance

On the whole set of 270 genotypes evaluation of disease resistance has been carried out for three pathogens – leaf rust, stem rust and Fusarium head blight – under field conditions in artificially inoculated nurseries. In the case of Fusarium head blight, the type II resistance (spreading of the disease in the spike) has been evaluated in greenhouse test. During the GWAS analysis we didn't find significant marker trait associations only strong correlations. We selected 14 Fusarium resistance landraces and four-four leaf rust and stem rust resistance landraces for resistance breeding and developing of MAGIC mapping population for genetic studies.

Abiotic stress tolerance

Drought tolerance

In Europe and worldwide periods of heat and drought are projected to be more frequent, longer, and occurring earlier, which could deleteriously affect the productivity of cool-season crops including bread wheat. The drought tolerance of the lines was established in the rain-out shelter, applying three water regimes: (1) irrigated till ZD45 stage of the plants then not irrigated, (2) fully irrigated, and (3) naturally rain fed. To estimate the effect of drought stress conditions on yield and yield components of each genotype, drought stress tolerance index was calculated. We performed GWAS to identification the genetic background of drought tolerance. We calculated tolerance index from the data of main spike grain weight and grain weight per plant. We identified significant association with two SNPs (BS00049210_51, Kukri_c41603_111) on 2A and 7A chromosomes (Figure 7).

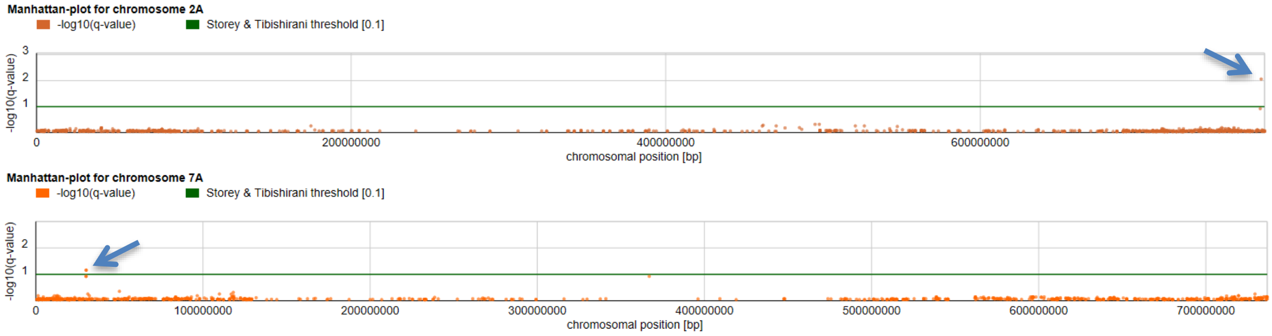


Figure 7. Manhattan-plots of significant SNPs associated with length of the main spike. Marker-trait associations analyses were conducted using the easyGWAS software, where EMMAX test was fitted based on SNP data. We applied a GWAS threshold of the Storey & Tibishirani correction to declare significant associations.

We performed the GWAS analysis using the plant total seed weight and the strongest correlation was identified with the Kukri_c41603_111 markers on the 7A chromosome. These results indicate that 2A and 7A specific SNP markers would facilitate to transfer drought tolerance from landraces into modern breeding lines.

Heat stress tolerance

Determination of the high temperature tolerance of the landraces collection was carried out in greenhouse and controlled cabinet tests at the developmental phase of DEV 49 (First awns visible). The plants were raised to the given phenophase in greenhouse. The control plants stayed in the greenhouse throughout the development, while part of the plants were exposed to heat stress of 34°C for 14 days in the growth cabinet. The seed number of the main spike indicated the heat stress tolerance of the lines. We identified three significant markers on 2B, 6D and 7A chromosomes which were responsible for the majority of the phenotypic variance (Figure 8). These markers could be useful during the selection of the heat tolerant lines by MAS.

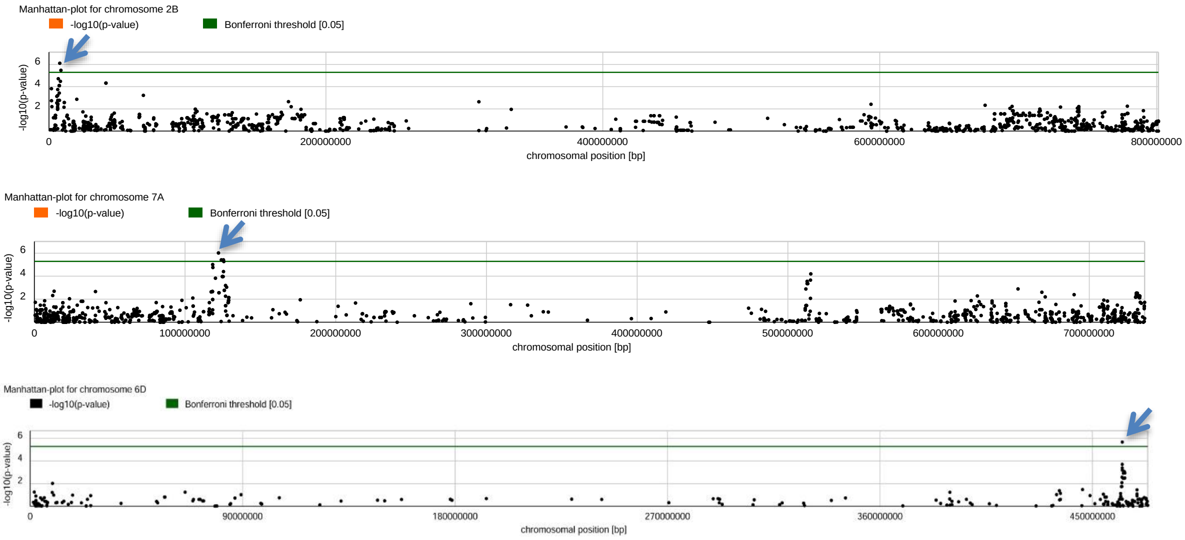


Figure 8. Manhattan-plots of significant SNPs associated with heat stress tolerance. Marker-trait associations analyses were conducted using the easyGWAS software, where EMMAX test was fitted based on SNP data. We applied a GWAS threshold of the Bonferroni correction to declare significant associations.

Conclusions

Our work focusing on the genetic diversity of Central European wheat provided valuable information for understanding the relationships between landraces and modern elite varieties. We facilitated their characterization and determined their population structure and ancestral origins. Our results could enrich breeding strategies for future crop improvement through helping breeders to develop new varieties by reducing pre-breeding activities. Our data can be used, for example, to explore selective sweeps for any specific gene or chromosome region, analyze footprints defining divergence of landraces from distinct ecologies, or identify germplasm groups conserving allelic diversity missing in current breeding programs. We used GWAS to identify SNPs associated with phenotyped traits. We generated a comprehensive database relating SNP polymorphisms with phenotypes and compared the co-localization with previously mapped genes and QTLs from literature. The most important SNP markers have been converted to KASPar markers. The development and validation of a co-dominant SNP marker set are suitable for high-throughput selection in further breeding programs.

Dissemination

Published Scientific papers:

Cseh, András ; Lenykó-Thegze, Andrea ; Makai, Diána ; Szabados, Fanni ; Hamow, Kamirán Áron ; Gulyás, Zsolt ; Kiss, Tibor ; Karsai, Ildikó ; Moncsek, Blanka ; Mihók, Edit et al.: Meiotic instability and irregular chromosome pairing underpin heat-induced infertility in bread wheat carrying the Rht-B1b or Rht-D1b Green Revolution genes, New Phytologist, 2023 IF: 9.4

Cseh András, Poczai Péter, Kiss Tibor, Balla Krisztina, Berki Zita, Horváth Ádám, Kuti Csaba, Karsai Ildikó: Exploring the legacy of Central European historical winter wheat landraces, SCIENTIFIC REPORTS 11: (1) 23915, 2021 IF: 4.996

Andras Cseh, Maria Megyeri, Caiyun Yang, Stella Hubbart-Edwards, Duncan Scholefield, Stephen Ashling, Ian King, Julie King, Surbhi Grewal: Development of a new Am-genome-specific SNP marker set for the molecular characterization of wheat-Triticum monococcum introgression lines, PLANT GENOME 12: (3) 180098, 2019 IF: 4.042

T. Kiss , K. Balla , **A. Cseh**, Z. Berki, Á. Horváth, Gy. Vida, O. Veisz, I. Karsai Assessment of the genetic diversity, population structure and allele distribution of major plant development genes in bread wheat cultivars using DArT and gene-specific, Cereal Research Communications, 2021 IF: 0.850

Lenykó-Thegze Andrea, Fábián Attila, Mihók Edit, Makai Diána, **Cseh András**, Sepsi Adél: Pericentromeric chromatin reorganisation follows the initiation of recombination and coincides with early events of synapsis in cereals, The PLANT JOURNAL, 2021 IF: 6.417

Sepsi Adél, Mihók Edit, Makai Diána, Lenykó-Thegze Andrea, **Cseh András**: Törpe búza az éghajlati változás tükrében: áldás vagy átok?, Agrofórum, 2020, November, 2020

Our manuscript is ready for submission

András Cseh, Gyöngyvér Gell, Christakis George Florides, Zsófia Birinyi, Dalma Nagy-Réder, Ferenc Békés, Ildikó Karsai (2024) Diversities in seed storage compositions in the Panel of Central European landraces and modern wheat genotypes Current Research in Food Science IF: 6.84

Abstract:

Cseh, András ; Gell, Gyöngyvér ; Christakis, George Florides ; Rakszegi, Marianna ; Birinyi, Zsófia ; Nagy-Réder, Dalma ; Békés, Ferenc ; Karsai, Ildikó: A Kárpát-medencei búza tájfajták és modern fajták tartalékfehérje szintű diverzitásának összehasonlítása, XXIX. Növénynemesítési Tudományos Napok : Összefoglaló kötet Martonvásár, Magyarország : MTA Agrártudományi Kutatóközpont (2023) 158 p. pp. 74-74. , 1 p., 2023

Cseh, András ; Poczai, Péter ; Kiss, Tibor ; Balla, Krisztina ; Berki, Zita ; Horváth, Ádám ; Karsai, Ildikó: Genome Diversity of Central European Wheat Landrace Collection Compared to Elite Bread Wheat Cultivars, Plant and Animal Genome XXVIII Conference (PAG), 2020

András Cseh, Krisztina Balla, Tibor Kiss, Marianna Mayer, Zita Berki, Ádám Horváth, Péter Poczai, Ildikó Karsai: SNP Genotyping and Agro-Morphological Characterization of a Central-European Wheat Landrace Collection and Modern Wheat Cultivars, In: International Plant & Animal Genome XXVII, (2019) PE0966, 2019

Andras Cseh, Tibor Kiss, Krisztina Balla, Zita Berki, Ádám Horváth, Péter Poczai, Ildikó Karsai: Exploiting the legacy of Central European wheat landraces for improving the ecological adaptation ability of wheat, In: MonoGram, (2019) P25, 2019