Genetic loci involved in resistance to six races of stem rust in a wild emmer wheat panel assessed at seedling stage

Mastrangelo AM[1], Roncallo P[2], Matni O[3], Steffenson B[3], Ozkan H[4], Cakir E[4], Echenique V[2], Barabaschi D[5], Cattivelli L[5], **Elisabetta Mazzucotelli[5]**

[1]CREA-Research Centre for Cereal and Industrial Crops, Foggia, Italy; [2]Depto. de Agronomía, Universidad Nacional del Sur and Centro de Recursos Naturales Renovables de la Zona Semiárida (CERZOS-CONICET), Bahía Blanca, Argentina; [3]Department of Plant Pathology, University of Minnesota, St Paul, Minnesota; [2]Department of Field Crops, University of Çukurova, Adana, Turkey; CREA-Research Centre for Genomics and Bionformatics, Fiorenzuola d'Arda (PC), Italy;
Contact: **elisabetta.mazzucotelli@crea.gov.it**

The wild emmer wheat, *Triticum dicoccoides* (AABB), is the progenitor of the domesticated durum wheat (AABB), and fully fertile in cross with it. Wild emmer harbors a wide spectrum of genes and alleles lost during domestication and breeding, nevertheless a number of genes and alleles positively contributing to biotic and abiotic stress tolerance, yield components and quality have been identified in wild emmer and introgressed in cultivated wheats. Stem rust, caused by the fungus *Puccinia graminis f. sp. tritici (Pgt)*, is one of the most important wheat disease worldwide. Also in Italy, climate weather conditions of the last years favored a persistent presence in Sicily. The majority of the known resistance genes are already ineffective against current races of *Pgt*. A wide collection of wild emmer accessions, equipped with genotypic data, has been established to recover useful genetic diversity for cultivated wheats. This collection is being exploited for genome wide association (GWAS) for a number of traits, including stem rust resistance. This research has been supported by the project CerealMed (PRIMA2019).

THE COLLECTION

The CREA - Genomics and Bioinformatic Research Centre held a collection of wild emmer accessions. The accessions have originated from all environments of the Fertile Crescent countries where wild emmer naturally occurs (Fig.1). Color codes is according to DAPC groups (Fig. 2)

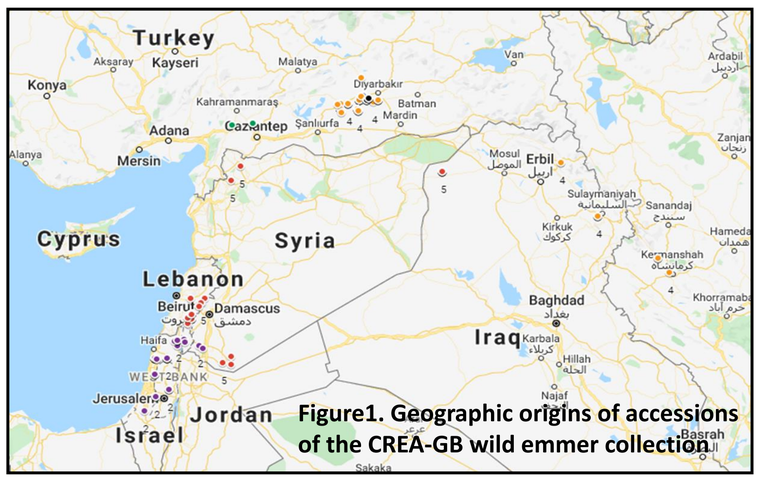

Figure1. Geographic origins of accessions of the CREA-GB wild emmer collection

GENOTYPING AND PHYSICAL MAPPING OF SNPs

After one cycle of single seed descent, 285 wild emmer accessions were genotyped with the **Axiom 35k Wheat Breeders' Array**. Upon step-wise filtering based on call rate (>97%), MAF (>0.05), and SNP classification according to clustering quality, a total of 12100 poly-high resolution SNPs were retained.

These SNPs were physically mapped on the wild emmer genome through Blast against the **Zavitan reference genome** v2 (Zhu et al., 2019). For 6931 SNPs an unambiguous genome position was found, with an average of with an average density of one SNP per 1.26 Mb.

THE POPULATION STRUCTURE OF THE COLLECTION REFLECTS WILD EMMER RACES AND GEOGRAPHIC ORIGINS OF THE ACCESSIONS

The multivariate Discriminant Principal Components Analysis (DPCA) was firstly conducted to analyse the population structure of the collection. The first component (IPC1) separated accessions mostly based on the two known wild emmer races, with the north-eastern race on the left side of the DAPC graph, and the western race on the right side. The second component (IPC2) identified clusters related to specific geographic origins.

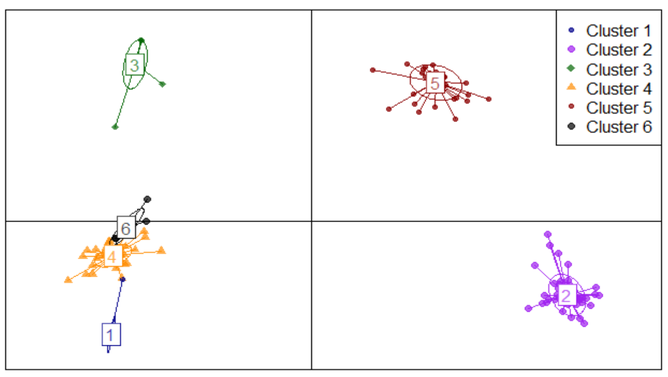

Figure 2. Graphical representation of DAPC principal component 1 (IPC1) and 2 (IPC2) distances for 6 sub-populations. Clusters include accession origins as follows; Cluster 1: Turkey-Diyarbakir; Cluster 3: Turkey-Gaziantep; Cluster 6: Turkey-Pirinclik; Cluster 4: Turkey-different areas, Iraq and Iran; Cluster 2: Israel; Cluster 5: Lebanon and Syria.

A subset of 1521 SNPs, selected based on physical position in order to evenly span all chromosomes, were used for **STRUCTURE analysis**. STRUCTURE revealed two groups of genetically similar individuals (K1 and K2), mostly corresponding to the two known wew races (North-Eastern race-NE, and Western race-W). Further analysis **identified subgroups mostly corresponding to origin countries** (Fig 3). A **phylogenetic analysis** was conducted through MEGA4 software (Fig. 3). The phylogenetic tree confirmed the strong differentiation between the two wild emmer races (NE and W) and highlighted a general higher genetic diversity within the NE race than in the W race can be observed.

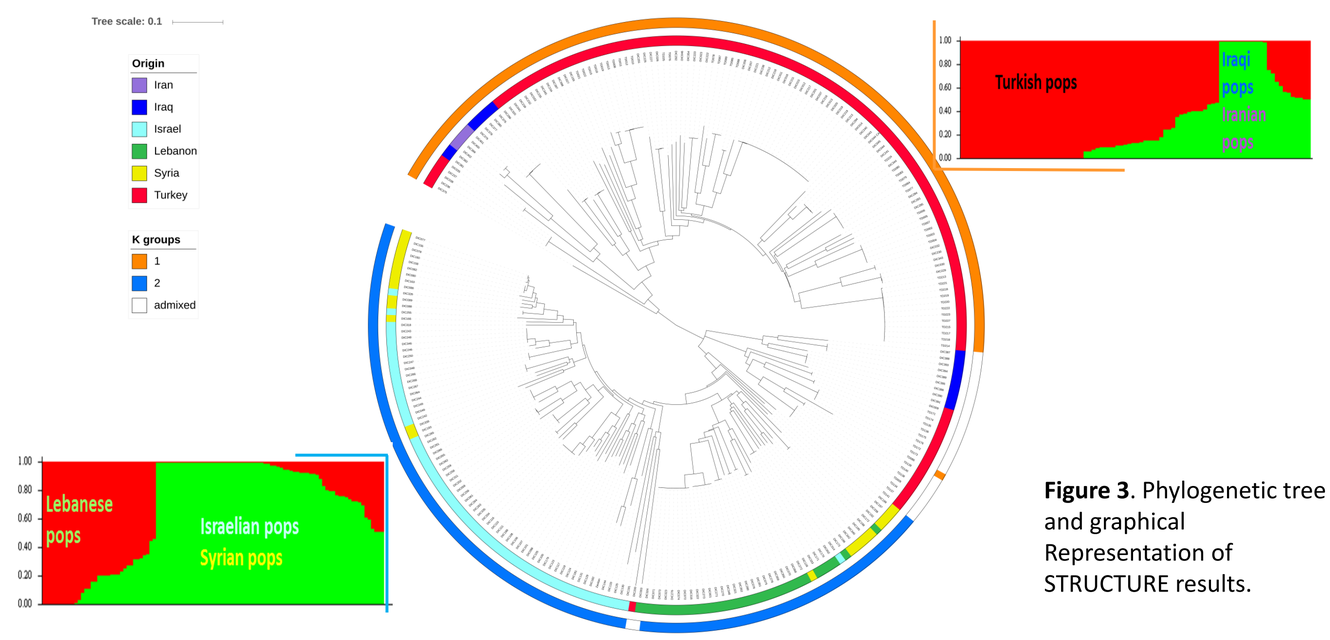

Figure 3. Phylogenetic tree and graphical Representation of STRUCTURE results.

FAST LD DECAY SUPPORTING HIGH RESOLUTION MAPPING

Significant LD r^2 values ($p<0.01$) were plotted against physical distance to calculate the LD decay rate.

The genome wide critical distance for **R^2 decay <0.2 was found at 126 kbp.**

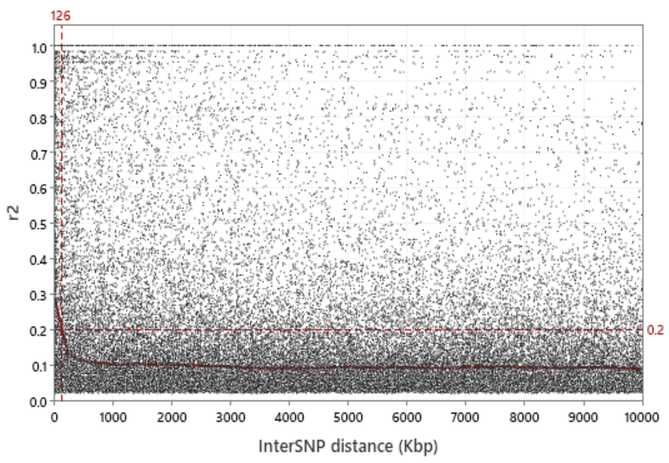

Figure 3. Genome wide LD decay aginst physical distance.

PHENOTYPING FOR STEM RUST RESISTANCE

The collection was evaluated at **seedling stage against six *Pgt* races, namely TPMKC, TTTTF, JRCQC, TRTTF, TTKSK, and TKTTF,** in a greenhouse for controlled conditions at the University of Minnesota, using differential lines as controls. For each race, the frequency distribution of IT scoring values highlighted a prevalence of susceptible genotypes, however 10 accessions were resistant to three or more races. Most of the resistant accessions belonged to the western wild emmer race.

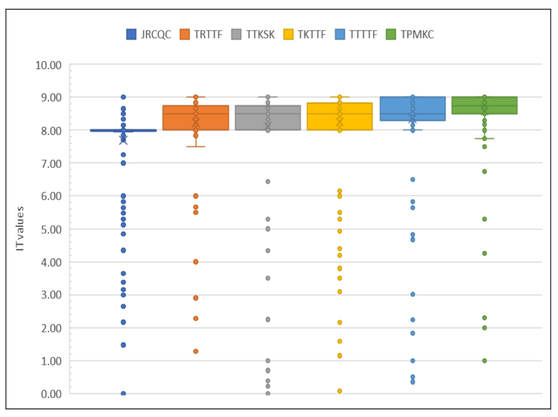

Figure 4. Box plots of wew accessions reaction against the 6 Pgt races

GWAS FOR STEM RUST RESISTANCE: NOVEL LOCI AND CANDIDATE GENES

A GWAS based on a Mixed Linear Model, which considered PCA and Kinship relationships between accessions, was run to identify resistant loci. However, due to the low frequency of resistant lines in the global wew collection and in the *north-eastern* wew race, the GWAS was conducted only on the *western* subset and excluding the TPMKC race, which had only few resistant genotyped lines.

Based on the Bonferroni threshold, 19 MTAs have significant *p*-values, with some MTAs in common between more than one race. The focus was then on 11 MTAs (Tab.1) with a strong effect on the reaction level, that is MTAs for which the mean IT value of one genotypic class was lower than 6. For all MTAs the alternative minor was the allele conferring resistance phenotype.

Trait	Chr	Position (Mb) on Zavitan V2 genome	*p* value	Expl. Variance R^2	Average IT of accessions carrying major allele	Average IT of accessions carrying minor allele	MAF %	n° accessions with the associated allele	Novel/ known loci
TTKSK	1B	54.3	7,3E-10	0,29	8,3	4,5	9	10	Letta 2014
TTKSK	1A	43.1	3,6E-07	0,20	8,2	4,9	8	9	Novel
TTTTF	5B	4.6	3,0E-07	0,21	8,6	5,6	8	10	Novel
JRCQC	1B	54.3	3,2E-18	0,49	7,8	3,7	9	10	Letta 2014
JRCQC	1A	43.1	4,7E-13	0,36	7,8	3,7	8	9	Novel
JRCQC	2A	31.5	1,1E-07	0,22	7,8	4,7	11	13	Letta 2014, Mergessa 2020
JRCQC	7A	45.5	7,3E-07	0,19	7,8	5,5	16	19	Novel
JRCQC	4A§	NM	7,3E-07	0,19	7,8	5,5	16	19	\
JRCQC	6B	35.4	7,3E-07	0,19	7,8	5,5	16	19	Megerssa 2020
TRTTF	5B	4.6	5,8E-06	0,17	8,4	5,6	7	8	Novel
TRTTF	2A	646.4	6,1E-06	0,17	8,4	5	5	6	Novel

Table 1. Summary of MTAs with a strong effect (average IT value <6) on the resistance level.

Results were compared with resistant stem rust loci identified in other wheat panels. Some MTAs located close to loci previously identified (Letta et al 2014, Mergessa et al., 2020), while others looked to be novel. Based on the wew genome annotation, candidate genes related to defense reaction were found for most of the loci within the MTA confidence interval as defined by LD decay.