

UPGRADING THE DURUM WHEAT GENOMIC RESOURCES: FROM THE PLATINUM-QUALITY SVEVO GENOME ASSEMBLY AND ANNOTATION TO THE TETRAPLOID WHEAT PANGENOME

FORESTAN C.*, BOZZOLI M.*, MACCAFERRI M.*, TUBEROSA R.*, MAZZUCOTELLI E.**,
DESIDERIO F.***, FACCIOLI P.***, CATTIVELLI L.***, ENS J.***, SHARPE A.***,
POZNIAK C.***, CHAWLA H.****, WALKOWIAK S.*****, PIRONA R.*****,
CERIOTTI A.*****, MORGANTE M.*****, MASCI S.*****, SESTILI F.*****
, GIULIANO G.***** , GADALETA A.***** , PÈ M. E.***** ,
PECCHIONI N.***** , BASSI F. M.***** ,
DISTELFELD A.***** , RUSHOLME-PILCHER R.***** ,
HALL A.***** , SWARBRECK D.***** ,
SPANNAGL M.***** , NAVRATILOVA P.***** ,
ŠIMKOVÁ H.***** , SILVESTRI M.***** ,
ZASTROW-HAYES G.***** , LLACA V.***** ,
FENGLER K.***** ,
THE SVEVO PLATINUM CONSORTIUM ***** ,
THE TETRAPLOID WHEAT PANGENOME CONSORTIUM *****

*) Department of Agricultural and Food Science (DISTAL) - Università di Bologna - Italy

**) Research centre for Genomics & Bioinformatics - Council for Agricultural Research and Economics (CREA) - Italy

***) Wheat Genetics and Breeding - Crop Development Centre - University of Saskatchewan, - Canada

****) University of Manitoba – Canada

*****) Canadian Grain Commission – Canada

*****) National Research Council–Institute of Agricultural Biology and Biotechnology (CNR-IBBA), Milano -Italy

*****) Department of Agricultural, Food, Environmental and Animal Sciences (Di4A), University of Udine - Italy

*****) Department of Agriculture and Forest Sciences, University of Tuscia, Viterbo - Italy

*****) Biotechnology and Agro-Industry Division, ENEA, Casaccia Research Center, Rome, Italy

*****) Department of Soil, Plant and Food Science (Di.S.S.P.A.),

University of Bari Aldo Moro - Italy

*****)) Center of Plant Sciences, Scuola Superiore Sant'Anna, Pisa, Italy

*****)) International Centre for the Agricultural Research in the Dry Areas, BIGM, Rabat, Morocco

*****)) Department of Evolutionary and Environmental Biology, University of Haifa – Israel

*****)) Earlham Institute, Norwich Research Park, Norwich, United Kingdom

*****)) Helmholtz Zentrum München, Plant Genome and Systems Biology, Neuherberg, Germany

*****)) Institute of Experimental Botany of the Czech Academy of Sciences, Olomouc - Czech Republic

*****)) Barilla G. e R. F.lli S.p.A, Parma, Italy

*****)) Biotechnology-GT-Genomics – Corteva Agriscience - 8325 NW 62nd Ave, Genomics Lab, Johnston, IA, USA

*****)) The Svevo Platinum Consortium

*****)) The Tetraploid Wheat Pangenome Consortium

*****)) Research Centre for Cereal and Industrial Crops - Council for Agricultural Research and Economics (CREA), Italy

durum wheat, genome assembly, pangenome, transcriptome, short and long read sequencing

Durum wheat (*Triticum turgidum* L. ssp. *durum*) is a major cereal and staple in the semi-arid regions of the Mediterranean Basin. It originates from BBAA wild tetraploid domesticated in Neolithic era, later evolving to domesticated emmer and then to up to 11 *T. turgidum* subspecies, including durum wheat landraces and modern cultivars. Tetraploid wheat is the donor of the A and B genomes of hexaploid bread wheat (DDAABB), representing therefore a valuable source of genetic variability and beneficial alleles for both durum and bread wheat breeding.

The reference durum wheat genome (cv Svevo) was previously sequenced and assembled using a short-read sequencing approach (Maccaferri et al., 2009), representing an important milestone for wheat genomics. In an effort to improve this release, an international consortium was established to produce a platinum-quality reference genome, that accomplish with the Contiguity, Completeness, Correctness ultimate RefSeq2.0 requirements. PACBIO HiFi long read 35X sequencing was coupled to BIONANO Optical Mapping, producing 259 Hybrid scaffolds (N50 = 112.3 Mb) ordered by Hi-C data in a 10.4 Gb assembly. A complete and accurate gene prediction was then obtained by coupling Illumina RNASeq and Nanopore Isoseq sequencing of

30 plant samples representative of a range of tissues and developmental stages grown in normal growth condition and of 28 samples obtained under biotic, abiotic and nutrient stress conditions. The expression of the 68,154 high confidence genes will be integrated in transcriptome atlas, together with more than 100,000 low confidence, TE-related or long non-coding genes.

Gene regulatory networks associated to spike and seed development are under investigation and will be associated with the chromatin accessibility obtained by ATAC-Seq assays from the same samples, providing an exceptional starting point to study how enhancers, promoters, transcription factors binding cooperatively regulate gene expression during development.

Finally, the preliminary results coming from the sequencing of 24 tetraploid wheat (representative of wild emmer, domesticated emmer, turgidum and turanicum subspecies, durum wheat landraces and modern varieties) pangenome will be presented. Genomic and transcriptomic data coming from this project will allow to fully explore the huge tetraploid genome diversity available for the identification of beneficial alleles enhancing both durum and bread wheat resilience to abiotic stresses and disease resistance.

This research has been supported by: PRIN project “PanWheatGrain - Grain Pangenomics for Durum Wheat Sustainable Production” funded by the Italian Ministry of University Education and Research; Agritech National Research Center funded by the European Union Next-Generation-EU (PNRR - MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022) and by the Canadian project “Canadian Tetraploid Pan Genomics Project”.