

Genome wide association study of free asparagine content in durum wheat grain

Andrea Tafuri^{1,2,3}, Raul Pirona¹, Aldo Ceriotti¹, Melania Zuccaro¹, Stefano Ravaglia², Mariaros Giordano², Laura Ruth Cagliani⁴, Roberto Consonni⁴, Agostino Fricano⁵, Elisabetta Mazzucottelli⁵, Elena Baldoni¹.

¹ Institute of Agricultural Biology and Biotechnology (IBBA), National Research Council (CNR), via A. Corti, 12, 20133, Milano (MI), Italy - ² SIS Società Italiana Sementi, Via Mirandola, 5, 40068, San Lazzaro di Savena (BO), Italy - ³ Department of Agriculture and Forest Sciences, University of Tuscia, Via San Camillo de Lellis, 01100, Viterbo (VT), Italy - ⁴ Institute of Chemical Sciences and Technologies "G. Natta" (SCITEC), National Research Council (CNR), via A. Corti, 12, 20133, Milano (MI), Italy - ⁵ Council for Agricultural Research and Economics (CREA), Research centre for Genomics & Bioinformatics, Via S. Protaso, 302, 29017, Fiorenzuola d'Arda (PC), Italy.

BACKGROUND

Acrylamide (AA), known to be neurotoxic and suspected to be carcinogenic, has been detected in many foods after high temperature processing. In wheat derivatives, the free asparagine (fAsn) level has been identified as a key factor in AA formation. Exploring natural variation for fAsn content in wheat genotypes could help breeders to identify novel beneficial traits and genes for breeding programs.

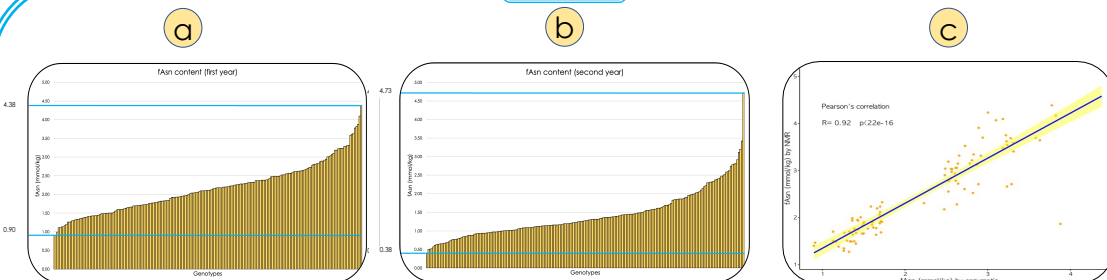
RESEARCH AIM

To investigate the variability in fAsn accumulation in wheat grain in order to identify associated markers and candidate genes controlling fAsn content by a Multi-Locus Genome Wide Association Study (ML-GWAS),

EXPERIMENTAL PLAN

215 durum wheat (*Triticum turgidum* L. ssp. *durum* (Desf.)) genotypes, including 193 landraces, were selected from the international Global Durum Panel (GDP) and grown in Fiorenzuola d'Arda (PC, Italy) for three years. Our experimental plan is to search for candidate genes or chromosome regions associated with fAsn accumulation in the grain through a GWAS.

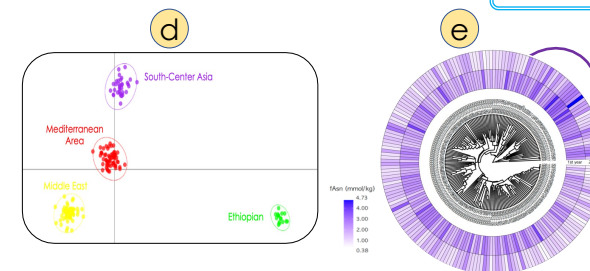
RESULTS



a fAsn quantification in the whole-grain samples was performed using K-ASNA enzymatic assay (Megazyme) for the first two years of field trials. The analysis indicated the presence of a good variability for fAsn content in the population. The analysis of the third year is still in progress.

c 100 genotypes were selected from first year field trial to be investigated by High Resolution ¹H-NMR (Proton Nuclear Magnetic Resonance) spectroscopy. Pearson's correlation between fAsn quantified with the enzymatic assay and ¹H-NMR indicated a good correlation between those two different methodologies applied to fAsn quantification.

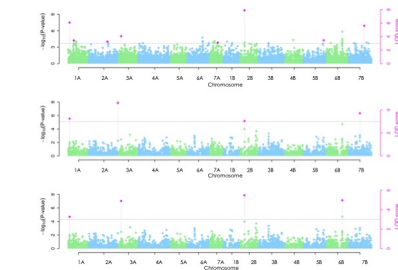
RESULTS



d Structure population analysis confirms conservation of about 95.5% of the original genetic variability in the sub-population and identified four groups, matching the geographic origin of the genotypes.

e A Phylogenetic tree was constructed using poppr (R package). Neighbour-joining phylogenetic tree was constructed based on SNPs data with 1000 bootstrap replications. The associated heatmap shows two years fAsn values. A high fAsn content clade was found (indicated by the purple line)

RESULTS



Linkage disequilibrium (LD) was calculated using TASSEL5 and measured using r^2 . This analysis showed that in our collection LD extends till 4.4 Mbp. A ML-GWAS analysis was performed using mrMLM (R package) which contains six ML-GWAS algorithms (FASTmrEMMA, FASTmrMLM, ISIS EM-BLASSO, mrMLM, pKwEB, pLARM).

Manhattan plots show an example of our results from GWAS for fAsn content. SNPs in purple are considered good markers associated with the trait. This analyses allowed us to identify 6 QTNs (Quantitative Trait Nucleotides).

SUMMARY AND CONCLUSION

Population structure analysis indicated that the selected panel maintained a high level of genetic variability, comparable to that observed for the original GDP population. fAsn quantification indicates the presence of a good variability for this trait in the chosen population. Despite to the strong effects of the environment on fAsn accumulation, well described by several authors, the identification of a high fAsn content clade in the phylogenetic tree suggests the presence of genetic determinants related to the fAsn content in the grain. The preliminary ML-GWAS on the data related to the first two years identified 6 QTNs associated to 144 genes for a total of 398 stable transcripts.

ACKNOWLEDGEMENTS



This work was partially funded by the European Regional Development Fund under the ROP of the Lombardy Region ERDF 2014-2020—Axis I "Strengthen technological research, development and innovation"—Action 1.b.1.3 "Support for co-operative R&D activities to develop new sustainable technologies, products and services"—Call Hub.