

Genome-Wide Association Analysis of Grain Size Traits in a Global Barley (*Hordeum vulgare* L.) Collection

Outmane Bouhlal^{1,2*}, Miguel Sanchez-Garcia¹, Adil Mulakat^{1,3}, Fatima Ezzahra Rachdad^{1,4}, Houda Hiddar^{1,5}, Mostafa Kandil² & Andrea Visioni¹

¹Biodiversity and Crop Improvement Program (BICP), International Centre for Agricultural Research in the Dry Areas (ICARDA), Rabat, Morocco.

²Anthropogenetics, Biotechnologies and Health Laboratory, Faculty of Sciences, Chouaib Doukkali University, El-Jadida, Morocco.

³Plant and animal production and agro-industry laboratory, Faculty of Science, Ibn Tofail University, Kenitra, Morocco

⁴Environment and Ecology Laboratory, Biology department, Faculty of Sciences Ben M'sik, Hassan II University, Casablanca, Morocco.

⁵Laboratory of Microbiology and Molecular Biology, Faculty of Sciences, BioBio Research Center, University Mohammed V in Rabat, Rabat, Morocco

Abstract | Barley (*Hordeum vulgare* L.) grain size is a critical factor in determining storage capacity during grain filling. Grain size is a complex quantitative trait easily influenced by environmental factors and directly affects yield and malt quality. In this study, we aimed to identify marker-trait associations (MTAs) and putative candidate genes (CGs) related to grain size traits using a global spring barley collection genotyped with a 50K SNP chip. Data on grain size traits was collected from a set of 275 barley accessions at two different sites over two consecutive years in Morocco. The barley lines exhibited significant phenotypic variation for all studied traits. A broad-sense heritability over 60% was observed across all tested environments. Furthermore, highly significant correlations were observed among grain size traits. Population structure analyses showed four distinct sub-populations. Genome-wide association analysis (GWAS) using best linear unbiased estimators (BLUES) unveiled 233 MTAs linked to thousand kernel weight, number of kernels per spike, weight of kernels per spike, and grain shape traits including length, width, perimeter, area, and circularity. By clustering significant markers with high Linkage Disequilibrium (LD) and within a 0.61 Mb distance, 13 QTLs were identified across all environments. Three stable genomic regions, located on chromosomes 1H (488.88 - 489.15 Mb), 2H (581 - 582.67 Mb), and 7H (496.43 - 532.13 Mb), were associated with multiple grain size traits, highlighting their potential significance for Moroccan dry environments. Notably, these genomic regions harbor QTLs relevant to several trait and were enriched with functional proteins involved in diverse cellular processes including plant reproductive development, grain production, and drought tolerance. Overall, the study provides essential information about significant markers and new QTL associated to grain size traits that should be further validated for their potential use in marker-assisted selection.

Keywords: Barley, Grain size, Genome-wide association analysis, population structure Linkage Disequilibrium, QTLs.