

Candidate Genes and Stable QTL for Grain Yield and Seed Size in Durum Wheat

Giacomo Mangini (✉ giacomo.mangini@ibbr.cnr.it)

Università degli Studi di Bari Aldo Moro <https://orcid.org/0000-0002-6600-2178>

Antonio Blanco

Università degli Studi di Bari Aldo Moro

Domenica Nigro

Università degli Studi di Bari Aldo Moro

Massimo Antonio Signorile

Università degli Studi di Bari Aldo Moro

Rosanna Simeone

Università degli Studi di Bari Aldo Moro

Research article

Keywords: wheat, yield, grain size, stable QTL, candidate genes

Posted Date: January 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-34145/v3>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: In wheat grain yield is expressed as the product of different components. Among these, thousand kernels weight (TKW) reflects the combination of several grain related traits including grain length (GL), grain width (GW) and area. Grain weight is also affected by phenological traits, such as heading time (HT) and plant height (PH). To detect stable QTL and candidate genes involved in phenotypic control of grain yield, a recombinant inbred line (RIL) population derived from two elite durum wheat cultivars (Liberdur and Anco Marzio) was evaluated for yield components and grain related traits for three growing seasons in southern Italy. The mapping population was genotyped with a 90K SNP array and a high-density genetic linkage map with 5134 markers was obtained.

Results: A total of 30 QTL were detected on the durum RIL population including 9 stable QTL for TKW (2 QTL), GL, GW (2 QTL), AREA, HT and PH (2 QTL) distributed on 1B, 2A, 3A and 6B chromosomes. Interestingly, a QTL cluster mapped on 2A included a major QTL for HT explaining at least 70% of phenotypic variance and co-located with a QTL for YLD, TKW, GL and GW and AREA, respectively. In the physical position of this QTL cluster a *photoperiod sensitivity* gene (*Ppd-A1*) was found. *Serine carboxypeptidase*, *Big Grain 1* and β -*fructofuranosidase* candidate genes were mapped in clusters containing stable QTL. Candidate genes involved in auxin metabolism were also found in QTL clusters in which a QTL for AREA was declared.

Conclusions: This study showed that yield components and phenological traits had higher inheritances than grain yield, allowing an accurate stable QTL cluster detection. This was a powerful requisite to physically map QTL on the reference durum wheat genome and to identify candidate genes strongly affecting the genetic grain yield network.

Background

Durum wheat (*Triticum turgidum* L. ssp. *durum*) is grown on about 17 million hectares of land worldwide [1]. The Mediterranean Basin is the main durum wheat producing area, as well as its most significant import market and its largest consumer, being durum wheat primarily used for pasta, couscous, bulgur and farik production. The projected world's population growth will result in a higher production demand, thus increasing the grain yield per unit area will be of great importance to face this mounting challenge [2]. The rate of wheat production has been increasing by only 0.9% per year, which is much less than the required 1.8%, suggesting that the improvement of wheat yield must be further increased [3]. Grain yield is conventionally expressed as the product of different number of sub-traits named "yield components", covering two main parameters: number of spikes per area and grain yield per spike. Grain yield per spike comprises grain number per spike and grain weight, usually expressed as thousand kernels weight (TKW). Grain weight is mainly and tightly underpinned by grain morphology including area, grain length and grain width. In domestication process and breeding history, grain size was a major selection and breeding target, and has been widely selected and manipulated to increase the related grain yield in wheat [4].

Understanding the genetic and molecular determinants of grain size might provide valuable information on the markers to be used for improving grain yield. As extensively reviewed by Nadolska-Orczyk et al. [5], the major genes determining yield-related traits can be classified in several groups. Among these, there are: transcription factors, which can affect grain number by regulating spike development; genes involved in metabolism or signaling of growth regulators, affecting plant architecture; genes determining cell division and proliferation, related to grain size; floral regulators, which regulate inflorescence architecture and seed number, and genes involved in carbohydrate metabolism, affecting plant architecture and grain yield.

The most advanced knowledge on the genetic factors controlling grain size were detected in rice. Indeed, quantitative genetic studies and map-based cloning have allowed to isolate several genes associated with grain size and weight in rice [6]. Rice *OsGW2* gene was found to affect both grain weight and width [7], as well as its wheat orthologue, named *TaGW2*, which was mapped on chromosome 6A [8] and found to be related to grain weight, grain width and length [9]. Similarly, allelic variation at *TaGS-D1*, the wheat orthologue of rice *GS3* [10], showed main effects on grain weight and kernel size [11]. Genes involved in starch and sucrose metabolism pathways were also shown to affect grain weight, such as *TaSus1* and *TaSus2* [12, 13] and *TaCwi-A1* [14]. In addition, major phenology loci can exhibit pleiotropic effects on spike and kernel traits. These include the locus *Rht1*, which have been intensively used since the Green Revolution in breeding programs worldwide, and the major photoperiod sensitivity locus *Ppd1*, which is known to affect heading time as well as a range of other traits [15]. In particular, *Ppd1* genes play an important role in wheat growth and development regulation, by affecting the accumulation and distribution of dry matter within the plant and modifying source-sink equilibrium [16, 17]. Indeed, it has been estimated that up to 35% of bread wheat's grain yield increment observed in Europe has been due to photoperiod insensitivity [18].

In the last decade, SNP markers have become fundamental for both genetic studies and breeding programs, and the development of wheat high-density SNP array [19] provided the most innovative tool for candidate genes searching by both QTL mapping or GWAS studies. Moreover, the recently released durum wheat genome, could provide powerful information to decipher complex marker-traits associations and simplify the searching and discovery of genes underlying important agronomic traits [20].

In the present study, a high-density genetic linkage map based on iSelect 90K SNP markers was realized using the durum wheat Liberdur x Anco Marzio recombinant inbred line (RIL) population that was evaluated to detect QTL controlling grain yield components and grain size. Furthermore, the availability of a full reference durum wheat genome [20], allowed the physical projection of the previously identified QTL genetic intervals on the reference Svevo genome, as well as the detection of candidate genes involved in phenotypic control of grain yield and grain size.

Results

Phenotypic characterization

The two parental line, Liberdur and Anco Marzio, and the RILs mapping population were evaluated for yield components and grain size traits (YLD, PH, HT, TKW, GL, GW and AREA) for three growing seasons (2016, 2017 and 2018) in southern Italy (Valenzano, Bari, Italy). Analysis of variance revealed highly significant differences among RILs for all traits in each year (Additional file 1: Table S1), while the combined analysis across years revealed significant effects of RILs, years and a strong genotype x year interaction (Table 1). However, although the strong season effect, the genotype variability was higher than genotype x year component for all traits with the exception of YLD.

Mean values of Liberdur, Anco Marzio, and RILs in single season and across the three field trials are reported in Table 2. The two parents showed significant differences for HT, PH, AREA in all three seasons, and for TKW, GL and GW in two trials. In particular, Anco Marzio has generally larger grains compared to those of Liberdur, which were narrower. Analysis of the frequency distribution of traits in the RIL population was performed to have a preliminary idea of the genetic basis for each trait. A bimodal distribution was observed for HT indicating a single locus segregating in the RIL mapping population (Fig. 1). By contrast, a normal distribution obtained for YLD, PH, TKW and grain size traits would be indicative of several to many loci each contributing to a small proportion of the total variation observed. High transgressive segregation was recorded for all traits suggesting the presence of superior alleles for grain yield components in both parental lines. Low values of broad sense heritability were estimated for YLD, confirming that the trait was strongly affected by environmental conditions. High heritability values, exceeding 0.55, were found for HT, GL and AREA while TKW heritability values ranged from 0.45 to 0.72 (Table 2).

Phenotypic correlation among the traits across the three seasons are reported in Table 3. HT was negatively correlated to TKW ($r = -0.24$), GW ($r = -0.36$) and AREA ($r = -0.19$), while a positive correlation was shown with YLD ($r = 0.55$). As expected, the grain size traits were inherently correlated, such as AREA and GL ($r = 0.86$). TKW showed a high positive correlation with AREA ($r = 0.95$), GL ($r = 0.70$) and GW ($r = 0.83$).

Genetic Linkage Map

Out of 81,587 SNPs assayed, 5543 (6.8%) resulted as failed and 67,999 (83.3%) were monomorphic across the mapping population. The remaining 8,045 (9.86%) were polymorphic; however, 2,686 had more than 10% missing data and 225 with distorted segregation (at $p \geq 0.05$ value) were excluded from further analysis. Hence, 5,134 (6.29%) markers were used for the genetic map construction. The 5,134 loci were grouped in 21 linkage groups when using a LOD score 5. Linkage groups were assigned to the A and B genome chromosomes using the durum consensus map [21]. Eight chromosomes (1B, 2A, 3B, 4B, 6A, 6B, 7A and 7B) were assembled in a single linkage group (Table 4). Twenty-three loci assembled in a linkage group of 4A chromosome resulted to be coincident in the same position. Therefore, this linkage group was discarded from the QTL analysis. A total of 2085 markers were localized on the A genome with a total length of 1,145.25 cM, whereas 3,049 were mapped on the B genome (total length 1,062.69 cM). The

entire map covered 2,207.94 cM with an average chromosome length of 157.71 cM. The lengths of individual chromosomes varied from 73.25 cM (chromosome 5B) to 197.22 cM (chromosome 7A). The overall SNP density was 2.3 markers/cM, with a maximum of 3.8 for chromosome 1B and a minimum of 1.2 for chromosome 4A. The SNP markers were generally well distributed throughout the genome, although some chromosomes exhibited higher densities.

QTL detection

A total of 30 putative QTL were detected on ten chromosomes in the durum wheat Liberdur x Anco Marzio RIL mapping population (Table 5). Two YLD QTL were located on chromosomes 2A and 7B respectively, explaining individually 11.9-67.9% of the phenotypic variance. *QYLD.mgb-2A*, declared in 2016 and across years, showed the additive effect being provided by the parental line Liberdur (Fig. 2). *QYLD.mgb-7B* was detected in 2018 and displayed the additive effect being conferred by Anco Marzio.

Two QTL conferring HT were detected on chromosomes 2A and 7A, accounting for 2.4 to 85.7% of the explained phenotypic variance. The $-\log_{10}(P)$ scores of *QHT.mgb-2A* ranged from 44.6 to 69.7 and explained over 70% of the phenotypic variance in each year. This locus was stably detected in three years and across years, resulting as a major QTL, which additive effect was contributed by Liberdur.

Four QTL were found to be significantly associated with PH on chromosomes 2B, 3A, 6B and 7A, accounting for 7.8 to 17.5% of the explained phenotypic variance. *QPH.mgb-3A* and *QPH.mgb-6B* were stably detected in two field experiments and across years. The additive effects at *QPH.mgb-3A* and *QPH.mgb-6B* were provided by Liberdur and Anco Marzio, respectively.

Five TKW QTL were detected on chromosomes 1A, 2A, 3A and 6B, which accounted for 2.9-35.8% of the explained phenotypic variance. Of them, two stable QTL were mapped on chromosome 2A, delimited by genetic intervals 35.6-36.4 cM (*QTKW.mgb-2A.1*) and 87.5-89.6 cM (*QTKW.mgb-2A.2*), respectively. *QTKW.mgb-2A.1* and *QTKW.mgb-2A.2* additive effects were conferred by Anco Marzio and Liberdur, respectively.

Six QTL were found to be significantly associated with GL on 1B, 2A (2 QTL), 5A and 6B (2 QTL). Among them, *QGL.mgb-2A.2* was declared in all years accounting for 39.1-53.8% of the explained phenotypic variance, thus resulting as a stable QTL. Additive effect at this locus derived from Liberdur. The *QGL.mgb-1B* and *QGL.mgb-5A* were declared in two years (both 2016 and 2018), while the remaining three QTL (*QGL.mgb-2A.1*, *QGL.mgb-6B.1* and *QGL.mgb-6B.2*) were detected in a single year. For these five QTL, the alleles for longer grain were carried by the parental line Anco Marzio.

Six QTL were found to be significantly associated with GW on 1B, 2A (2 QTL), 4B, 6B and 7B chromosomes. The *QGW.mgb-1B* and *QGW.mgb-2A.2* were stably detected and accounted for 8.4-16.3% of explained phenotypic variance. Both QTL showed the additive effect of Liberdur's allele. The

QGW.mgb-2A.1 explained a large phenotypic variation (42%), although it was detected in a single year and across years.

Five QTL were identified for AREA on 2A (2 QTL), 3A, 4B and 7B. The *QAREA.mgb-2A.2* was found in all years and across years accounting for 25.8-47.8% of the phenotypic variance, resulting to be a stable QTL. *QAREA.mgb-4B* was found in two years (2017 and 2018) while the other three QTL (*QAREA.mgb-2A.1*, *QAREA.mgb-3A* and *QAREA.mgb-7B*) were identified in a single year. Liberdur carried the allele for larger area for all QTL, except for *QAREA.mgb-2A.1*.

Out of the 30 detected QTL, ten were co-located in two different regions of chromosome 2A (Fig. 2). The first region included six QTL (*QYLD.mgb-2A*, *QHT.mgb-2A*, *QTKW.mgb-2A.1*, *QGL.mgb-2A.1*, *QGW.mgb-2A.1* and *QAREA.mgb-2A.1*) in the genetic interval between 34.4 and 36.4 cM, while the second comprised four stable QTL (*QTKW.mgb-2A.2*, *QGL.mgb-2A.2*, *QGW.mgb-2A.2* and *QAREA.mgb-2A.2*) in the interval between 87.5 and 89.6 cM. Interestingly, among the five QTL for TKW, two co-located together with QTL for GL, GW and AREA on chromosome 2A (both regions), and two co-located with QTL for GL or AREA on chromosomes 3A and 6B, respectively. Four out of 6 QTL for GL resulted co-located with QTL for GW on chromosomes 1B, 2A (two QTL) and 6B. QTL for AREA always co-located with grain traits QTL (TKW, GL, and GW) on chromosomes 2A (2 QTL), 3A, 4B and 7B.

Candidate genes involved in grain yield related traits

To identify candidate genes in physical regions underlying the QTL detected in the RILs mapping population, each region of interest was projected on the recently released reference durum wheat genome of cv. Svevo [20]. Regions with two or more co-locating QTL were considered as a QTL cluster (Table 6), and the sequences of flanking markers of each of them were anchored to their physical position on durum wheat genome. A total of eight clusters were searched for gene content, and the retrieved IDs genes were screened for their annotated functional role and involvement in cellular metabolisms, which were confirmed by searching for paralogues and orthologous genes in closely related species. The QTL cluster size ranged from 0.4 Mbp (on chromosome 3A, where two QTL for TKW and AREA co-located) to 57.1 Mbp (on chromosome 6B, with a QTL for TKW and a GL co-locating).

Among the several candidate genes localized within the QTL cluster regions, three were particularly noteworthy, as previously reported to be directly involved in grain yield: *TRITD2Av1G019250*, encoding for pseudo-response regulator (*Ppd-A1*) in the 2A cluster that included *QHT.mgb-2A* (Table 5); *TRITD4Bv1G171270*, encoding for a Big Grain 1 protein in the 4B cluster, and two candidate genes encoding for an *acid β -fructofuranosidase* (*TRITD6Bv1G005370* and *TRITD6Bv1G005450*), paralogues of the *cell wall invertase* gene, both in the 6B cluster containing QTL for GL, GW and PH.

Specific protein classes were frequently observed, such as proteins involved in ubiquitination processes, including E3 ubiquitin-protein ligase and RING U-box superfamily proteins (identified in six QTL clusters),

cytochrome P450 (identified in four QTL clusters) and thioredoxins (identified in two QTL clusters), as well as serine carboxypeptidase proteins (identified in two QTL clusters).

Interestingly, in all cluster in which a QTL for AREA was detected, candidate genes involved in auxin metabolism were found: *TRITD1Bv1G118820*, *TRITD2Av1G189400*, *TRITD7Bv1G173200*, encoding for protein involved in auxin response; *TRITD4Bv1G175480*, involved in auxin signalling, and *TRITD3Av1G012070*, a paralog gene of the YUC family encoding for a *Flavin-containing monooxygenase*, which is directly involved in auxin biosynthesis. Similarly, except for one cluster on 2A chromosome, in all QTL clusters in which the QTL for GW were co-located, candidate genes encoding for cytochrome P450 were found.

Discussion

Grain yield is considered a typical quantitative trait, controlled by a complex genetic system and strongly influenced by both environmental factors and agronomic management. Grain yield reflects the combination of thousand kernels weight and grain number per area. Besides being one of the key components of grain yield, thousand kernels weight is also often used as standard parameter both for flour-milling yield and marketing standard. Thousand kernels weight is mainly and tightly underpinned by grain morphology including grain length, grain width and area. Grain yield can be maximized by growing varieties which heading time allows the crop to avoid stresses during grain-filling phase [17]. Indeed, heading time is a critical stage that delimits the duration of spike formation and marks the transition into the grain-filling period, during which grain per spike and grain weight are both defined. In addition, in the last century, grain yield improvement was obtained through increased harvest index and straw strength with the introduction of major genes conferring reduced plant height.

In this study, YLD, TKW, morphology grain traits (GL, GW and AREA), HT and PH were investigated by using a RILs mapping population consisting of 133 lines derived by crossing the elite durum wheat cultivars Liberdur and Anco Marzio. Plant materials were grown for three years in field trials.

A total of 30 QTL were identified and localized on 10 out of 14 durum wheat chromosomes. Among them, there were 9 stable QTL for TKW (2 QTL), GL, GW (2 QTL), AREA, HT and PH (2 QTL), distributed on 1B, 2A, 3A and 6B chromosomes, which were detected at $-\log_{10}(P) > 3.0$ in at least two years and across years (Table 5). Interestingly, 6 out of 9 stable QTL were co-located in QTL clusters on chromosome 2A, while the other three were detected on 1B, 3A and 6B chromosomes (Fig. 2). Furthermore, 4 out of 8 QTL clusters were mapped on 2A and 6B chromosomes, and the remaining ones located on 1B, 3A, 4B and 7B.

The first QTL cluster on chromosome 2A, comprised within 35.6-36.4 cM, included two stable QTL (*QHT.mgb-2A* and *QTKW.mgb-2A.1*) along with a QTL for YLD, GL, GW and AREA. In this region, the *QHT.mgb-2A* can be considered a major QTL involved in the phenotypic control of HT as it was declared in each field trial with a $-\log_{10}(P)$ ranging from 44.0 to 69.7, and a R^2 from 74.3 to 85.7, respectively. The candidate genes analysis performed on durum wheat genome detected a pseudo-response regulator

protein (*TRITD2Av1G019250*), corresponding to a photoperiod sensitivity gene (*Ppd-A1*), within the physical position of this QTL, thus confirming the effect of *Ppd-A1* gene on YLD. *QHT.mgb-2A.1* was previously detected in the same genomic region co-locating with a QTL for YLD in a durum multi-parental cross population [22]. Wilhelm et al. [23], found two large deletions within the *Ppd-A1* gene in durum wheat (1027 and 1117 bp deletion designated as alleles 'GS-100' and 'GS-105, respectively), which remove a common region from the wild-type sequence. Therefore, our results suggested that the parental lines Liberdur and Anco Marzio could be different for the alleles at *Ppd-A1* locus. Similar conclusions were reported by Maccaferri et al. [24], who identified a QTL reducing heading date associated with *Ppd-A1* gene in a RILs population derived from the cross 'Kofa' ('GS-100' allele) × 'Svevo' ('GS-105' allele), thus suggesting that these alleles decrease photoperiod sensitivity at different degrees. Heading time resulted strongly related with grain yield ($r=0.55$), underlining the importance of phenology fitting in maximizing grain yield as observed by Royo et al. [25]. The QTL for HT (*QHT.mgb-2A*) and YLD (*QYLD.mgb-2A*) were associated to significant QTL with opposite effects on TKW (*QTKw.mgb-2A.1*, *QTKW.mgb-2A.1*) and grain size (*QGL.mbg-2A.1*, *QGW.mgb-2A.1*, *QAREA.mgb-2A.1*). Moreover, HT showed significant negative correlations with TKW, GW and AREA, while YLD did not correlate to grain related traits. These results suggested that the *Ppd-A1* gene showed negative pleiotropic effects on TKW and grain size, but positively affected others yield components such as grain number per unit area [26].

The second QTL cluster located on chromosome 2A included four stable QTL (*QTKW.mgb-2A.2*, *QGL.mgb-2A.2*, *QGW.mgb-2A.2*, *QAREA.mgb-2A.2*) declared in three years and across years. The positive additive effect was provided by the parental line Liberdur for all four QTL. As shown by correlation analysis data, TKW was always positively and significantly correlated to GL, GW and AREA. Considering both these results, it can be suggested that TKW improvement could be due to the grain size increase. A QTL for TKW was previously detected in this region by a genome wide association mapping study on a tetraploid wheat collection [27], thus validating the presence of a metaQTL regardless of the identification approach. Six putative candidate genes were included in this QTL cluster, and among them, a serine carboxypeptidase protein. Previous studies have shown that serine carboxypeptidase enzymes are involved in several biological processes, including development of plant organs [28, 29], cell division [30] and cell elongation [31]. Indeed, by promoting cell division, these enzymes determine larger grain size due to an increased cell number [30]. So far, our result was not unexpected considering that also other studies on dissecting grain size detected a strong relation with serine carboxypeptidase proteins [32, 33].

A QTL cluster on 6B chromosome included the stable QTL *QPH.mgb-6B*, which was detected in the same region where QTL for PH were already reported by Canè et al. [34] and Soriano et al. [35]. Two additional QTL co-localized in this cluster, *QGL.mgb-6B.1* and *QGW.mgb-6B*, where Canè et al. [34] detected a QTL for test weight, a trait which is known to be affected by grain size [36]. Therefore, it could be supposed that high values of test weight might be determined by increasing GL and GW. The co-localization of QTL affecting different traits such as PH, GL and GW, implies closely linked genes involved in different biological processes related to yield [37]. Among the identified putative candidates, two genes, *TRITD6Bv1G005370* and *TRITD6Bv1G005450*, both encoding for acid β -fructofuranosidase enzyme. These two genes have 22 paralogues in the B genome of durum wheat, which probably diverged from a

common ancestral gene and evolved by duplication. More interesting is the fact that these genes are paralogues of *cell wall invertase* genes. β -fructofuranosidase (EC 3.2.1.26) are indeed commonly known as invertase in plants. These genes are involved in the carbohydrate metabolic process and have a common molecular function, as invertase/fructofuranosidase belong to the Glycoside hydrolase family 32 [38]. Ma et al. [14] reported an association between a *cell wall invertase* gene (*TaCwi-A1*) and TKW, useful to improve grain yield. Li et al. [39] showed that expression of the *cell wall invertase* genes significantly improved shoot growth, grain yield and starch content in transgenic maize plants, and specifically increased both grain size and grain number.

QTKW.mgb-6B and *QGL.mgb-6B.2* were co-located on the second QTL cluster on 6B, suggesting that TKW improvement could be due to GL increase. In the same region, Elouafi and Nachit [40] reported a QTL for TKW using a linkage mapping population. Notably, the positive additive effect of the five QTL co-locating in the QTL clusters on 6B chromosome, was provided by the parental line Anco Marzio, indicating that this durum cultivar has potentially useful alleles for grain yield improvement.

Two QTL for grain size (*QGW.mgb-4B*, *QAREA.mgb-4B*) were found to co-localize on 4B chromosome. No QTL for grain size has been previously mapped in this region, indicating that could be considered a new QTL cluster associated to grain size. The positive additive effect was provided by the parental line Liberdur. Among the candidate genes identified within the 4B cluster, one was particularly noticeable, a Big Grain 1 protein (*TRITD4Bv1G171270*), which has been annotated as a positive regulator of auxin response and transport, as well as a regulator of grain size. The rice orthologue gene has been well characterized by Liu et al. [41], who showed how its activation significantly improved grain size. This protein, localized in plasma membrane, is induced by auxin treatment and its expression in vascular tissues could improve plant productivity, the most significant change being observed for increased grain size and grain weight.

The QTL cluster located on 1B was reported for the first time in this study, as no previous QTL for GW and GL have been reported in this region. Nevertheless, the positive additive effect of the stable GW QTL negatively affected the GL, suggesting that linked and/or pleiotropic genes for grain size map in this region.

QTL clusters on 3A and 7B chromosomes included a QTL for AREA, which co-localized with a TKW and a GW QTL, respectively, confirming the positive and highly significant correlation among TKW and grain size.

Candidate genes searching highlighted two more interesting regions on 3A and 5A chromosomes, respectively coincident with a PH QTL and a GL QTL (Fig. 2). Indeed, the candidate genes screening on 3A chromosome identified a *serine carboxypeptidase*, *TRITD3Av1G187020*, and a NADH-dependent *glutamate synthase* (NADH-GOGAT), *TRITD3Av1G177110*, confirming their involvement in grain yield [32, 33, 42]. Additionally, an interesting gene was found in the GL QTL region on 5A chromosome, *TRITD5Av1G037950*, encoding for an expansin protein, a paralogue of *TRITD2Av1G019110* gene,

reported in the first 2A QTL cluster, which belong to a gene family previously found related to both grain size and yield [43, 44].

Conclusions

In this study showed how yield improvement could be pursued considering yield components (TKW, GL, GW and AREA), as well as phenology related traits (PH and HT). These yield sub-components showed higher inheritance than grain yield, allowing a more accurate and powerful stable QTL detection. Physical anchoring of these QTL on the reference durum wheat genome cv. Svevo, enabled the identification of candidate genes strongly affecting the genetic grain yield network. Therefore, the availability of SNPs markers within candidate genes sequences, might represent a new breeding strategy based on functional markers, determining a more efficient grain yield genetic gain.

Methods

Genetic materials and phenotypic analysis

The mapping population used in this study consists of 133 recombinant inbred lines (RILs) developed from a cross between two durum wheat cultivars, Liberdur and Anco Marzio, differing for yield related traits, by advancing random individual F₂ plants to the F₇ generation by the single seed descent procedure. The RILs population was evaluated in open-field conditions over three consecutive growing seasons (2016, 2017, 2018) at Valenzano (Bari, Italy). The experimental design was a complete randomized block design with three replicates with each experimental unit consisting of a 5 m² plot. Sowing density was always 350 seeds m². The field experiments were supplied with 45 kg/ha N and 115 Kg/ha P₂O₅ in pre-sowing and 85 kg/ha N in top dressing each year.

During each field trial heading time (HT) was recorded as the number of days from March 1st, to 50% ear-emergence, corresponding to stage 55 on the Zadoks et al. [45] scale. Plant height (PH) was measured at complete maturity of plants, and grain yield (YLD) was measured after harvesting the plots. The morphometric grain related traits were determined by digital imaging analysis. For each replication of each line, 10 g of kernels were scanned using high resolution scanner-based image analysis. The images were processed using the Image-Pro Plus 7.0 software (Media Cybernetics, USA). Grain length (GL), grain width (GW), grain area (AREA), and kernels number were measured. The kernels number was used to calculate the thousand kernels weight (TKW).

SNP genotyping and linkage map construction

Genomic DNA from each RIL and parental line (Liberdur and Anco Marzio) was diluted to 50 ng/μL and further analyzed with the wheat 90K iSelect array [19]. Genotyping was performed by TraitGenetics GmbH [46] following the manufacturer's recommendations as described in Akhunov et al. [47]. The genotyping

assays were carried out using the Illumina iScan reader and performed using GenomeStudio software version 2011.1 (Illumina, San Diego, CA, USA).

Chi-squared tests were used to determine the goodness-of-fit at $p > 0.001$ of segregation ratios to expected 1:1 ratio for each SNP. All markers with more than 10% missing data or segregating as presence/absence in the mapping population were excluded from further analysis. Linkage analysis between markers and determination of the linear order of loci was performed by QTL IciMapping 4.1 using the BIN and MAP functions [48]. Grouping was performed using the independence LOD parameter, with groups showing a LOD 5. The Kosambi mapping function was used to calculate map distances. SNPs data from the durum consensus map [21] were used as anchor loci and for assigning linkage groups to specific chromosomes. Linkage groups were named accordingly to the wheat chromosome nomenclature followed by a number linkage groups within chromosomes.

Statistical analysis and QTL detection

Analysis of variance (ANOVA) was performed to test the significance of differences among the RILs and replications, using the MSTAT-C software. The effects of replicates and genotypes were considered in the model in each year. All data from each field trial were tested by combined analysis of variance using the statistical model as follows:

$$Y_{ijk} = \mu + g_i + \alpha_k + \delta_{ik} + \beta_{jk} + \epsilon_{ijk} \quad (1)$$

Where μ is the general mean, g_i is effect of the i^{th} line, α_k is effect of the k^{th} year, δ_{ik} is interaction effect between the i^{th} line and the k^{th} year, β_{jk} is effect of the j^{th} replication within the k^{th} year. The g_i , δ_{ik} and ϵ_{ijk} 's effects are assumed independently and randomly distributed, with zero means and variances of σ^2_G , σ^2_{GY} and σ^2_{ϵ} , respectively.

Genetic variance (σ^2_G), environmental variance (σ^2_{ϵ}), and variance due to genotypic x year interaction (σ^2_{GY}) were obtained by using the combined analysis of variance (Additional file 2: Table S2). Broad-sense heritability (h^2_B) was estimated by the ratio σ^2_G / σ^2_P , where σ^2_P is phenotypic variance ($\sigma^2_P = \sigma^2_G + \sigma^2_{GY} + \sigma^2_{\epsilon}$) as reported by Singh & Ceccarelli [49]. Pearson phenotypic correlation coefficients (r) were calculated for all the traits across the years [50]. Mean of parental lines and RIL population, range, coefficient of variation (C.V.), were calculated for each field trials and across years.

The phenotypic traits' average values of each field trial (2016, 2017 and 2018) and the mean across trials were used for QTL analysis. Inclusive Composite Interval Mapping (ICIM) method was employed for QTL mapping using IciMapping 4.1 software [48]. A threshold P value of 0.001 ($-\log_{10}(P) \geq 3.0$) was used for QTL detection [51-53], while suggestive QTL were considered at the sub-threshold $2.5 < -\log_{10}(P) < 3.0$

when declared at least in one year. A QTL was considered stable when detected at $-\log_{10}(P) \geq 3.0$ in at least two years [35, 53-55]. The phenotypic variation explained ($PVE = R^2$) and additive effect were estimated for each detected QTL. Positive or negative additive effect indicates the increasing or decreasing effect of the parental line Liberdur. QTL were named according to the rules of International Rules of Genetic Nomenclature [56]. The QTL name combined the traits evaluated ('YLD', 'PH', 'HT', 'TKW', 'GL', 'GW', and 'AREA') and the Research Institute (Genetics and Plant Breeding Section, University of Bari, 'mgb') that carried out the experiments. Graphical representation of linkage groups and QTL was carried out using MapChart 2.2 software [57].

Candidate genes research

QTL intervals detected on genetic map were physically mapped on the durum wheat reference genome Svevo [58]. Left and right flanking markers of each confidence interval were first searched in the durum consensus map [21], and then projected on Svevo genome. When the marker on RIL mapping population was not mapped on the consensus map, the closest one was chosen.

The projected flanking markers were searched and positioned on the durum reference genome, and the annotated genes within each interval were screened based on their confidence and functional annotation. Candidate genes potentially involved in yield related traits control were further investigated by analyzing gene structure and syntenic relation with other *Triticeae* as well as orthologous search in other grass species [59].

Abbreviations

QTL: quantitative trait loci; RILs recombinant inbred lines; SNPs: single nucleotide polymorphisms; YLD: grain yield; HT: heading time; PH: plant height; TKW: thousand kernels weight; GL: grain length; GW: grain width (GW); AREA: grain area.

Declarations

Authors' contributions

GM and AB contributed to the conception; GM, MAS and DN performed the acquisition data; GM and AB cured the interpretation of data; GM, AB, DN and RS writing the original manuscript; GM Writing, review and editing the manuscript. All the authors read and approved the final manuscript

Funding

This research was supported by grants from ATENEO project 2014: Advanced genetic methods to study biotic and abiotic stresses in durum wheat.

Availability of data and materials

All data generated or analysed during this study are included in this article and its supplementary information files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest.

Author details

Department of Soil, Plant and Food Sciences, Genetics and Plant Breeding Section, University of Bari, Bari, Italy

References

1. del Pozo A, Matus I, Ruf K, Castillo D, Méndez-Espinoza AM, Serret MD. Genetic advance of durum wheat under high yielding conditions: the case of Chile. *Agronomy*. 2019;9:454.
2. The future of food and agriculture – Trends and challenges. 2017. <http://www.fao.org/3/a-i6583e.pdf>. Accessed 20 May 2020.
3. Ray DK, Mueller ND, West PC, Foley JA. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE*. 2013;8(6):e66428.
4. Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW. A genetic framework for grain size and shape variation in wheat. *Plant Cell*. 2010;22:1046–1059.

5. Nadolska-Orczyk A, Rajchel IK, Orczyk W, Gasparis S. Major genes determining yield-related traits in wheat and barley. *Theor Appl Genet.* 2017;130(6):1081–1098.
6. Zuo JR, Li JY. Molecular genetic dissection of quantitative trait loci regulating rice grain size. *Annu Rev Genet.* 2014;48:99–118.
7. Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. A QTL for rice grain width and weight encodes a previously unknown RINGtype E3 ubiquitin ligase. *Nat Genet.* 2007;39:623–630.
8. Su Z, Hao C, Wang L, Dong Y, Zhang X. Identification and development of a functional marker of *TaGW2* associated with grain weight in bread wheat (*Triticum aestivum*). *Theor Appl Genet.* 2011;122:211–223.
9. Simmonds J, Scott P, Brinton J, Mestre TC, Bush M, Del Blanco A, Dubcovsky J, Uauy C. A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theor Appl Genet.* 2016;129:1099–1112.
10. Wang S, Wu K, Yuan Q, Liu X, Liu Z, Lin X, Zeng R, Zhu H, Dong G, Qian Q, Zhang G., Fu X. Control of grain size, shape and quality by *OsSPL16* in rice. *Nat Genet.* 2012;44:950–954.
11. Zhang Y, Liu J, Xia X, He Z. *TaGS-D1*, an ortholog of rice *OsGS3*, is associated with grain weight and grain length in common wheat. *Mol Breeding.* 2014;34:1097–1107.
12. Hou J, Jiang Q, Hao C, Wang Y, Zhang H, Zhang X. Global selection on sucrose synthase haplotypes during a century of wheat breeding. *Plant Physiol.* 2014;164:1918–1929.
13. Volpicella M, Fanizza I, Leoni C, Gadaleta A, Nigro D, Gattulli B, Mangini G., Blanco A., Ceci LR. Identification and characterization of the sucrose synthase 2 gene (*Sus2*) in durum wheat. *Front Plant Sci.* 2016;7:266.
14. Ma D, Yan J, He Z, Wu L, Xia X. Characterization of a cell wall invertase gene *TaCwi-A1* on common wheat chromosome 2A and development of functional markers. *Mol Breeding.* 2012;29(1); 43–52.
15. Würschum T, Langer SM, Longin CFH, Tucker MR, Leiser WL. A three-component system incorporating *Ppd-D1*, copy number variation at *Ppd-B1*, and numerous small-effect QTL facilitates adaptation of heading time in winter wheat cultivars of worldwide origin. *Plant Cell Environ.* 2018;41(6):1407–1416.
16. Foulkes M, Sylvester-Bradley R, Worland A, Snape J. Effects of a photoperiod response gene *Ppd-D1* on yield potential and drought resistance in UK winter wheat. *Euphytica.* 2004;135:63–73.
17. Kamran A, Iqbal M, Spaner D. Flowering time in wheat (*Triticum aestivum*): a key factor for global adaptability. *Euphytica.* 2014;197:1–26.
18. Worland AJ. The influence of flowering time genes on environmental adaptability in European wheats. *Euphytica.* 1996;89:49–57.
19. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, et al. Characterization of polyploid wheat genomic diversity using a high-density 90000 single nucleotide polymorphism array. *Plant Biotechnol J.* 2014;12: 787–

20. Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, et al. Durum wheat genome reveals past domestication signatures and future improvement targets. *Nat Genet.* 2019; 51, 885–895.
21. Maccaferri M, Ricci A, Salvi S, Milner SG, Noli E, Martelli PL, Casadio R, Akhunov E, Scalabrin S, Vendramin V. A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol J.* 2015;13(5):648–63.
22. Milner SG, Maccaferri M, Huang BE, Mantovani P, Massi A, Frascaroli E, Tuberosa R, Salvi S. A multiparental cross population for mapping QTL for agronomic traits in durum wheat (*Triticum turgidum durum*). *Plant Biotechnol J.* 2016;14(2):735–748.
23. Wilhelm EP, Turner AS, Laurie DA. Photoperiod insensitive *Ppd-A1a* mutations in tetraploid wheat (*Triticum durum*). *Theor Appl Genet.* 2009;118:285–294.
24. Maccaferri M, Sanguineti MC, Corneti S, Ortega JLA, Araus JL, Ben Salem M, et al. Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum*) across a wide range of water availability. *Genetics.* 2008;178:489–511.
25. Royo C, Ammar K, Alfaro C, Dreisigacker S, Garcia del Moral L, Villegas D. Effect of *Ppd-1* photoperiod sensitivity genes on dry matter production and allocation in durum wheat. *Field Crops Res.* 2018;221:358–367.
26. Graziani M, Maccaferri M, Royo C, Salvatorelli F, Tuberosa R. QTL dissection of yield components and morpho-physiological traits in a durum wheat elite population tested in contrasting thermo-pluviometric conditions. *Crop Pasture Sci.* 2014;65(1):80–
27. Mangini G, Gadaleta A, Colasuonno P, Marcotuli I, Signorile AM, Simeone R, De Vita P, Mastangelo AM, Laidò G, Pecchioni N, Blanco A. Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats. *PLoS ONE.* 2018;13(1): e0190162.
28. Dominguez F, Gonzalez MC, Cejudo FJ. A germination-related gene encoding a serine carboxypeptidase is expressed during the differentiation of the vascular tissue in wheat grains and seedlings. *Planta.* 2002;215:727–734.
29. Cercos M, Urbez C, Carbonell J. A serine carboxypeptidase gene (*PsCP*), expressed in early steps of reproductive and vegetative development in *Pisum sativum*, is induced by gibberellins. *Plant Mol Biol.* 2003;51:165–174.
30. Li YB, Fan CC, Xing YZ, Jiang YH, Luo LJ, Sun L, Shao D, Xu C, Li X, Xiao J, He Y, Zhang Q. Natural variation in GS5 plays an important role in regulating grain size and yield in rice. *Nat Genet.* 2011;43:1266–1269.
31. Bienert MD, Delannoy M, Navarre C, Boutry M. NtSCP1 from tobacco is an extracellular serine carboxypeptidase III that has an impact on cell elongation. *Plant Physiol.* 2012;158:1220–1229.
32. Wang S, Zhang X, Chen F, Cui D. A single-nucleotide polymorphism of TaGS5 gene revealed its association with kernel weight in Chinese bread wheat. *Front Plant Sci.* 2015;6:

33. Ma L, Li T, Hao C, Wang Y, Chen X, Zhang X. *TaGS5-3A*, a grain size gene selected during wheat improvement for larger kernel and yield. *Plant Biotechnol J*. 2016;14(5):1269–
34. Canè MA, Maccaferri M, Nazemi G, Salvi S, Francia R, Colalongo C, Tuberosa R. Association mapping for root architectural traits in durum wheat seedlings as related to agronomic performance. *Mol Breeding*. 2014;34:1629–1645.
35. Soriano JM, Malosetti M, Roselló M, Sorrells ME, Royo C. Dissecting the old Mediterranean durum wheat genetic architecture for phenology, biomass and yield formation by association mapping and QTL meta-analysis. *PLoS ONE*. 2017;12(5):
36. Nuttall JG, O'Leary GL, Panozzo JF, Walker CK, Barlow KM, Fitzgerald GJ. Models of grain quality in wheat – A review. *Field Crops Res*. 2017;202:136–
37. Du B, Wang Q, Sun G, Ren X, Cheng Y, Wang Y, Gao S, Li C, Sun D. Mapping dynamic QTL dissects the genetic architecture of grain size and grain filling rate at different grain-filling stages in barley. *Sci Rep*. 2019;9:18823.
38. Roitsch T, González MC. Function and regulation of plant invertases: sweet sensations. *Trends in Plant Sci*. 2004;9(12):606–
39. Li B, Liu H, Zhang Y, Kang T, Zhang L, Tong J, Xiao L, Zhang H. Constitutive expression of cell wall invertase genes increases grain yield and starch content in maize. *Plant Biotechnol J*. 2013;11(9):1080–
40. Elouafi I, Nachit MM. A genetic linkage map of the durum \times *Triticum dicoccoides* backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits. *Theor Appl Genet*. 2004;108:401–413.
41. Liu L, Tong H, Xiao Y, Che R, Xu F, Hu B, Liang C, Chu J, Li J, Chu C. Activation of Big Grain1 significantly improves grain size by regulating auxin transport in rice. *Proc Natl Acad Sci*. 2015;112:11102–11107.
42. Nigro D, Gadaleta A, Mangini G, Colasuonno P, Marcotuli I, Giancaspro A, Giove SL, Simeone R, Blanco A. Candidate genes and genome-wide association study of grain protein content and protein deviation in durum wheat. 2019;249(4):1157–1175.
43. Fu J, Thiemann A, Schrag TA, Melchinger AE, Scholten S, Frisc M. Dissecting grain yield pathways and their interactions with grain dry matter content by a two-step correlation approach with maize seedling transcriptome. *BMC Plant Biol*. 2010;10:63.
44. Lizana XC, Riegel R, Gomez LD, Herrera J, Isla A, McQueen-Mason SJ, Calderini DF. Expansins expression is associated with grain size dynamics in wheat (*Triticum aestivum*). *J Exp Bot*. 2010;4:1147–1157.
45. Zadoks JC, Chang TT, Konzak CF. A decimal code for the growth stages of cereals. *Weed Res*. 1974;14:415–421
46. TraitGenetics GmbH. <http://www.traitgenetics.de>. Accessed 20 May 2020.
47. Akhunov E, Nicolet C, Dvorak J. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theor Appl Genet*. 2009;119(3):507–517.

48. Meng L, Li H, Zhang L, Wang J. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* 2015;3:269–
49. Singh M, Ceccarelli S. Estimation of heritability of crop traits from variety trial data. Technical Manual N. 21. Aleppo, (Syria): International Center for Agricultural Research in the Dry Areas (ICARDA) press: 1996
50. Snedecor GW, Cochran WG. *Statistical Methods.* 6th ed. Ames, Iowa: The Iowa State University Press: 1967
51. Zhang N, Fan X, Cui F, Zhao C, Zhang W, et al. Characterization of the temporal and spatial expression of wheat (*Triticum aestivum* L.) plant height at the QTL level and their influence on yield-related traits. *Theor Appl Genet.* 2017;130, 1235–1252
52. Fan X, Cui F, Ji J, Zhang W, Zhao X, Liu J, Meng D, Tong Y, Wang T and Li J. Dissection of pleiotropic QTL regions controlling wheat spike characteristics under different nitrogen treatments using traditional and conditional QTL mapping. *Front Plant Sci.* 2019;10:17
53. Ma J, Zhang H, Li S, Zou Y, Li T, et al. Identification of quantitative trait loci for kernel traits in a wheat cultivar Chuannong16. *BMC Genet.* 2019;20,77.
54. Guo Y, Zhang G, Guo B, Qu C, Zhang M, et al. QTL mapping for quality traits using a high-density genetic map of wheat. *PLoS ONE* 2020;15(3):e0230601
55. Xin F, Zhu T, Wei S, Han Y, Zhao Y, et al. QTL Mapping of kernel traits and validation of a major QTL for kernel length-width ratio using SNP and bulked segregant analysis in wheat. *Sci Rep.* 2020;10:25
56. GrainGenes a Database for Triticeae and Avena.
<http://wheat.pw.usda.gov/ggpages/wgc/98/Intro.htm>. Accessed 20 May 2020.
57. Voorrips RE. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J Hered.* 2002;93: 77–
58. <https://www.interomics.eu/durum-wheat-genome-intranet>. Accessed 20 May 2020.
59. Ensembl plants. <https://plants.ensembl.org/index.html>. Accessed 20 May 2020.

Figures

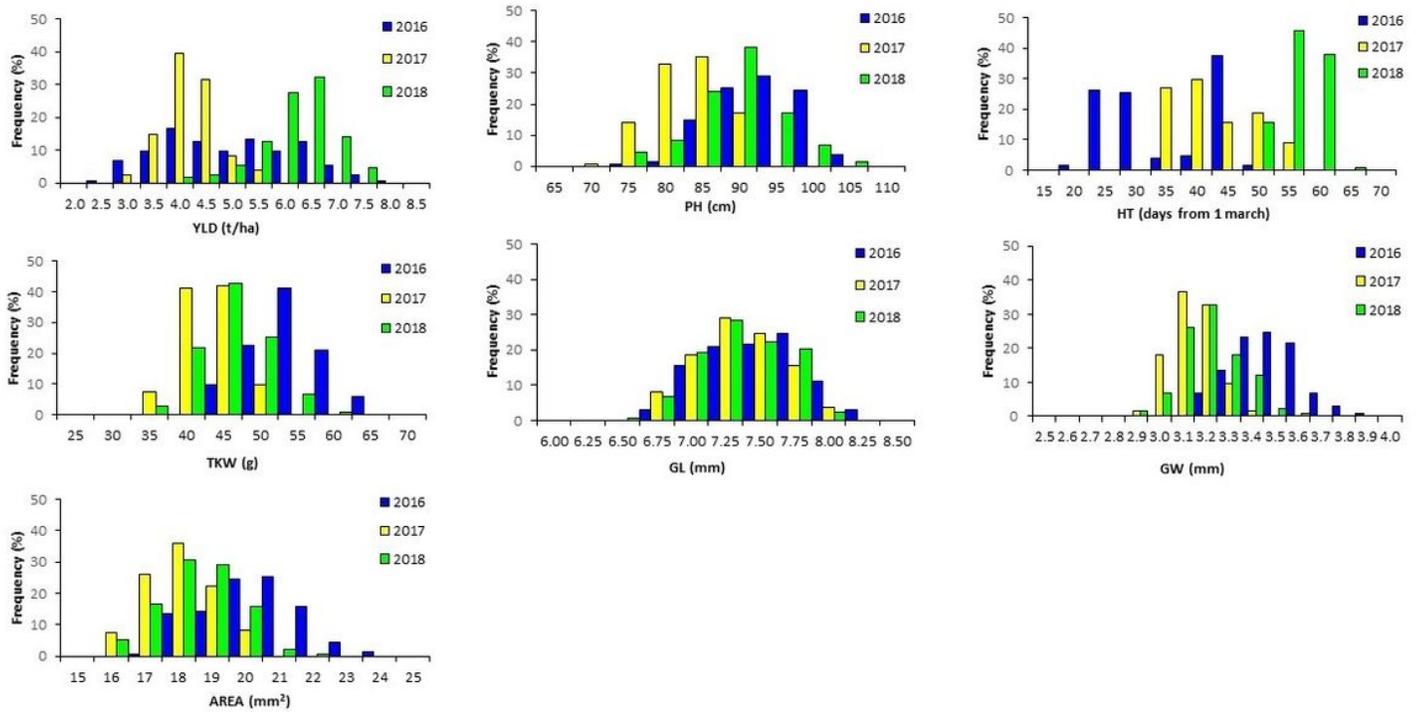


Figure 1

Frequency distribution of grain yield (YLD), plant height (PH), heading time (HT), thousand kernels weight (TKW), grain length (GL), grain width (GW), grain area (AREA) in the Liberdur x Anco Marzio RILs mapping population evaluated at Valenzano (Bari, Italy) for three years.

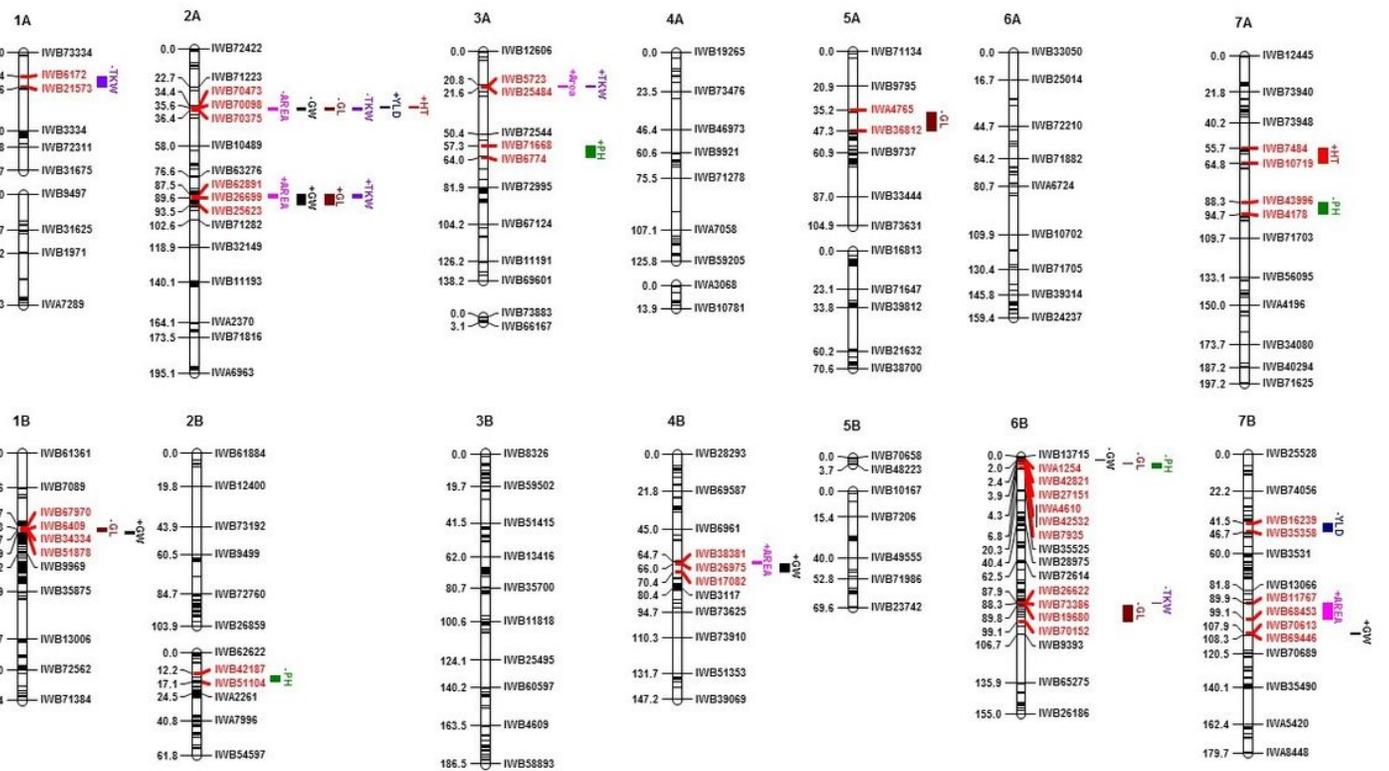


Figure 2

Schematic representation of the durum linkage map Liberdur x Anco Marzio with map positions of QTL for grain yield (YLD), plant height (PH), heading time (HT), thousand kernels weight (TKW), grain length (GL), grain width (GW), grain area (AREA). Each chromosome map is represented by the first and the last SNP marker, and by a SNP marker every about 20 cM. Markers are indicated on the right side and cM distances on the left side of the bar. QTL are represented by bars on the right of each chromosome bar. The + or – preceding the QTL name indicate the positive or negative additive effect of the parental line Liberdur. The left and right SNP markers of the QTL interval are indicated in red.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile1.TableS1.pdf](#)
- [AdditionalFile2.TableS2.pdf](#)