

RESEARCH ARTICLE

# Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats

Giacomo Mangini<sup>1</sup>, Agata Gadaleta<sup>2\*</sup>, Pasqualina Colasuonno<sup>2</sup>, Iaria Marcotuli<sup>2</sup>, Antonio M. Signorile<sup>1</sup>, Rosanna Simeone<sup>1</sup>, Pasquale De Vita<sup>3</sup>, Anna M. Mastrangelo<sup>3</sup>, Giovanni Laidò<sup>3</sup>, Nicola Pecchioni<sup>3</sup>, Antonio Blanco<sup>1</sup>

**1** Department of Soil, Plant & Food Sciences, Genetics and Plant Breeding Section, University Aldo Moro, Bari, Italy, **2** Department of Agricultural & Environmental Science, Research Unit of “Genetics and Plant Biotechnology”, University Aldo Moro, Bari, Italy, **3** Council for Agricultural Research and Economics—Cereal Research Centre, Foggia, Italy

\* [agata.gadaleta@uniba.it](mailto:agata.gadaleta@uniba.it)



**OPEN ACCESS**

**Citation:** Mangini G, Gadaleta A, Colasuonno P, Marcotuli I, Signorile AM, Simeone R, et al. (2018) Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats. *PLoS ONE* 13(1): e0190162. <https://doi.org/10.1371/journal.pone.0190162>

**Editor:** Aimin Zhang, Institute of Genetics and Developmental Biology Chinese Academy of Sciences, CHINA

**Received:** October 24, 2017

**Accepted:** December 9, 2017

**Published:** January 11, 2018

**Copyright:** © 2018 Mangini et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This research was supported by grants from MIUR, Italy, projects ‘PON-01\_01145 – ISCOCEM’ to AG.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Increasing grain yield potential in wheat has been a major target of most breeding programs. Genetic advance has been frequently hindered by negative correlations among yield components that have been often observed in segregant populations and germplasm collections. A tetraploid wheat collection was evaluated in seven environments and genotyped with a 90K SNP assay to identify major and stable quantitative trait loci (QTL) for grain yield per spike (GYS), kernel number per spike (KNS) and thousand-kernel weight (TKW), and to analyse the genetic relationships between the yield components at QTL level. The genome-wide association analysis detected eight, eleven and ten QTL for KNS, TKW and GYS, respectively, significant in at least three environments or two environments and the mean across environments. Most of the QTL for TKW and KNS were found located in different marker intervals, indicating that they are genetically controlled independently by each other. Out of eight KNS QTL, three were associated to significant increases of GYS, while the increased grain number of five additional QTL was completely or partially compensated by decreases in grain weight, thus producing no or reduced effects on GYS. Similarly, four consistent and five suggestive TKW QTL resulted in visible increase of GYS, while seven additional QTL were associated to reduced effects in grain number and no effects on GYS. Our results showed that QTL analysis for detecting TKW or KNS alleles useful for improving grain yield potential should consider the pleiotropic effects of the QTL or the association to other QTLs.

## Introduction

Increasing grain yield in cereal crops has been a major goal of most breeding programs, and the effects of genetic improvement on yield potential has been reported in several studies [1–

6]. Improving grain yield has always been a difficult task as it is a typical quantitative trait controlled by several genes, strongly influenced by environmental factors and crop management. In the last five decades, genetic gains in common and durum wheat have been mainly obtained by increasing harvest index associated to the gradual reduction in plant height. Further improvement are expected by increasing the biomass production and the radiation use efficiency without any reduction of the harvest index of modern cultivars [7–10]. Anyway, grain yield potential is the final product of plant growth and development, and several others complex factors, such as abiotic stress tolerance, adaptation to different soils and climate changes, disease resistances, contribute to plant productivity.

The primary numerical components of grain yield are the number of spikes per unit area, the average number of kernel per spike (KNS) and the average kernel weight, usually determined as one-thousand kernel weight (TKW). The product of kernel number and kernel weight is grain yield per spike (GYS). The importance of TKW also derives from being a marketing standard directly related to milling quality. KNS and TKW are quantitatively inherited, while the number of spikes per unit area, in the intensive cropping systems, mainly depends from planting density. Several authors (e.g. [11–12]) reported that an increase in seeding densities is almost invariably associated with a linear increase of the number of spikes per square meter across a wide range of seed rates.

Genetic advance in wheat breeding has been frequently hindered by negative correlations among yield components that have been often observed in segregating populations and germplasm collections (see reviews [6,13]). Particularly, the phenotypic correlation between KNS and TKW has been generally found negative even though not always consistent [14–16], while the correlations between GYS and the sub-components KNS and TKW have been always found consistent and positives. Phenotypic correlations may be attributed to genetic linkage, pleiotropy, environmental factors, and yield component compensations due to competition for a common limited nutrient supply. One of the most likely and widely accepted hypothesis for explaining the negative relationship between grain number and grain weight is that the increase of grain number per unit area or per spike produced by a genotype results in a lower availability of photo-assimilates synthesized during grain filling for each grain, which leads to decreases in individual grain weight due to competition effects [13,17–18]. Numerous results in the literature (e.g. [7,19–21]) on the effects of the introduction of the *rht* genes in breeding lines and in the released semi-dwarf cultivars during the 1960s and 1970s demonstrated that grain yield progress was associated with the increased number of grains per square meter and the increased number of kernels per spike. The same studies found a tendency of a concomitant decrease of grain weight in the widely cultivated semi-dwarf varieties. However, some investigations on common wheat [22–23] underlined how in most conditions the availability of assimilates was not the limiting factor for grain growth, and the negative relationship between KNS and TKW could not be attributed to competition among grains for limited assimilates. Moreover, recently studies [6,24–27] reported that in different countries grain yield progress of advanced breeding lines and cultivars was mainly associated with increased grain weight.

The recent development and increased availability of high throughput genotyping technologies have significantly contributed to the genetic dissection of complex traits into discrete quantitative trait loci (QTL). The classical QTL mapping approach is conducted in segregating populations resulting from crossing two parental lines different for the trait of interest, and exploiting the genetic association between molecular markers and QTL [28]. However, the development of large biparental populations is time-consuming, QTL analysis can allow detecting loci in genomic regions containing polymorphisms restricted to the parental lines, and the resolution power is rather low due to the limited number of crossing-over and recombination. The recent, alternative approach of the genome-wide association study (GWAS),

based on the linkage disequilibrium present in natural germplasm (landraces, breeding lines and elite cultivars), provides higher resolution in QTL analysis and increased power in loci with modest-size effects detection [29–30]. The limitation of GWAS in detecting a high frequency of false-positive and false-negative marker-trait associations is usually overcome by appropriate statistical methodologies, which take into account the population structure and relative kinship among individuals, and the multiple testing of thousands markers [31].

Mapping studies for the several components of grain yield potential have identified QTL on all 21 chromosomes of wheat genome by classical linkage mapping using biparental populations and by GWAS using wheat collections (see reviews [32–34]). Major QTL associated to thousand grain weight were detected on chromosomes 1B, 2A, 2D, 3A, 3B, 4B, 4D, 5A, 5B, 6A, 7B and 7D [16,35–52], and QTL for grain number per spike on chromosomes 2B, 3B, 4A, 4B, 4D, 5A, 5D, 7A and 7B [33,36–38, 43,48,53–59]. Current development of advanced biotechnologies have recently allowed the identification of some key genes for significant increases in grain weight or grain number per spike in wheat by comparative genomics with rice [60].

A better understanding of the genetic architecture of yield components is particularly necessary to break the negative relationship between number of KNS and TKW, and then to achieve further genetic progress in wheat breeding programs. The objectives of this study were: a) to identify major and stable QTL for KNS and for TKW by GWAS in a tetraploid wheat collection evaluated in several environments; b) to analyse the genetic relationships between GYS, KNS and TKW at QTL level; c) to identify molecular markers tightly linked to grain yield components. The identification of stable QTL that increase one component without decreasing others across environments is essential for improving wheat grain yield by traditional and genomic selection programs.

## Materials and methods

### Plant materials and phenotypic trait evaluation

The tetraploid wheat (*Triticum turgidum* L.,  $2n = 4x = 28$ ; AABB genome) collection used in this study was comprised of 233 accessions (S1 Table) chosen to represent the phenotypic variability for the grain yield component traits that were evaluated in this study. The panel, including wild and cultivated accessions of seven subspecies (*durum*, *turanicum*, *polonicum*, *turgidum*, *carthlicum*, *dicoccum* and *dicoccoides*), has been characterized in terms of genetic diversity and population structure [61] and used for genome-wide association mapping of loci controlling agronomic traits [62] and some qualitative traits, such as  $\beta$ -glucan content [63] and carotenoid content [64].

The whole wheat collection was grown in southern Italy in the experimental fields of the University of Bari at Valenzano (Bari) for five years (2009, 2010, 2012, 2013 and 2014, hereafter reported as V09, V10, V12, V13, V14), at Foggia in 2012 (hereafter reported as F12) and at Gaudio (Potenza) in 2013 (hereafter reported as G13). A randomized complete block design with four replications (V09, V10) or three replications (V12, F12, V13, G13, V14) and plots consisting of 1-m rows, 30 cm apart, with 50 germinating seeds per plot, was used in all field experiments. During the growing season, 100 kg/ha of N was applied and standard cultivation practices were adopted. Plots were hand-harvested at maturity and GYS was determined dividing grain yield per row by the number of spikes per row (about 70–80 spikes). A 15-g seed sample per plot was used to determine the TKW. KNS was calculated as the ratio between the average GYS and the average TKW.

### DNA extraction and SNP genotyping

Genomic DNA was isolated from freeze-dried young leaf tissues using the protocol as described by Sharp et al. [65]. DNA quality was checked by electrophoresis on 1.0% agarose

gels, and the concentration determined with a NanoDrop spectrophotometer. DNA of each sample was diluted to 50 ng  $\mu\text{l}^{-1}$  and genotyped for single-nucleotide polymorphism (SNP) using the wheat 90K Infinium iSelect array containing 81,587 gene-associated SNP markers [66]. Genotyping procedure was performed at TraitGenetics Laboratory, Gatersleben, Germany (<http://www.traitgenetics.de>) following the manufacturer's recommendations as described in Akhunov et al. [67]. The genotyping assays were carried out to the Illumina iScan reader and performed using Genome Studio software version 2011.1.

## Statistical and association mapping analyses

In the current study, each year-location combination was considered as an environment. Standard procedures for analysis of variance for each yield component was carried out with MSTAT-C software. Genetic variance ( $\sigma^2_G$ ), environmental variance ( $\sigma^2_E$ ), and variance due to genotypic x environment interaction ( $\sigma^2_{G \times E}$ ) were obtained by using the combined analysis of variance. Broad-sense heritability ( $h^2_B$ ) was estimated by the ratio  $\sigma^2_G/\sigma^2_P$ , where  $\sigma^2_P$  is phenotypic variance ( $\sigma^2_P = \sigma^2_G + \sigma^2_{G \times E} + \sigma^2_E$ ). Pearson phenotypic correlation coefficients were calculated between GYS, TKW and KNS.

SNP markers with >10% missing data points and markers with a minimum allele frequency (MAF) of less than 5% were removed from the data matrix prior to GWAS. Association between SNP markers and individual grain yield component traits was tested by using: a) the general linear model (GLM) and the GLM including the Q-matrix derived from the principal component analysis (PCA) as implemented in TASSEL (GLM+Q), and b) the mixed linear model (MLM) based on the kinship-matrix (MLM+K) and the MLM based on both the K-matrix and the Q-matrix (MLM+K+Q). Marker-trait associations (MTA) were considered significant at threshold  $-\log_{10}(P) \geq 3.0$  determined by the modified Bonferroni correction as implemented in Genstat software [68]. In agreement to the linkage disequilibrium (LD) estimates determined by Laidò et al. [62], the value of 8 cM was used as the support interval to declare significant SNPs associated with the examined yield component traits. QTL for GYS, KNS and TKW were considered stables when detected at  $-\log_{10}(P) \geq 3.0$  in at least three environments or two environments and the mean across environments, while were considered suggestive QTL at the sub-threshold  $2.0 < -\log_{10}(P) < 3.0$ .

The consensus high-density linkage map of durum wheat described by Maccaferri et al. [69] was used as reference map for chromosome localization and map position of SNP markers associated to QTL for each trait. The proportion of phenotypic variance explained by a single QTL was determined by the square of the partial correlation coefficient ( $R^2\%$ ). Additive effect for the bi-allelic SNP markers was estimated by TASSEL as the difference of the phenotypic effect between the most frequent allele toward the less frequent allele; positive or negative sign indicates the increasing or decreasing effect of the SNP allele with higher frequency. Graphical representation of linkage groups and QTL was carried out using MapChart 2.2 software [70].

## Results

### Phenotypic variation of grain yield components

Analysis of variance revealed highly significant differences between genotypes for GYS, KNS and TKW in each environment, while the combined analysis across environments revealed significant effects of genotypes, environments, and a strong genotype x environment interaction (G by E) (S2 Table). Mean GYS ranged from 1.76 g to 2.52 g in each environment with the average GYS across the seven environments being 2.08 g (Table 1). Mean TKW ranged from 46.7 to 59.5 g and mean KNS from 35.1 to 46.2 in each environment. The wide range of each trait in each environment can be attributed to the composition of the collection, which include

**Table 1. Means, ranges, coefficients of variation (CV), genetic variance ( $\sigma^2_G$ ) and heritability ( $h^2_B$ ) of grain yield per spike, kernel number per spike and thousand-kernel weight in a tetraploid wheat collection evaluated in seven environments.**

Traits	Environments							Across environments
	V09	V10	F12	V12	V13	G13	V14	
<i>Grain yield per spike (g)</i>								
Mean	1.95	2.19	2.52	1.76	2.30	2.12	2.01	2.08
Range	(0.44–3.24)	(0.70–3.59)	(0.70–4.28)	(0.52–3.08)	(0.74–3.62)	(0.74–3.48)	(0.81–3.38)	(0.80–3.25)
CV	15.9	11.3	13.2	14.7	8.8	9.0	11.2	11.3
$\sigma^2_G$	0.207	0.305	0.391	0.235	0.287	0.269	0.238	0.164
$h^2_B$	0.68	0.83	0.78	0.78	0.88	0.88	0.82	0.52
<i>Kernel per spike (n)</i>								
Mean	41.7	41.3	46.2	36.0	41.9	35.1	42.6	40.3
Range	(13.5–63.7)	(21.8–61.3)	(20.5–88.9)	(16.8–58.1)	(23.8–82.0)	(18.1–63.3)	(20.8–66.5)	(20.7–62.5)
CV	17.9	8.7	12.2	12.9	8.2	8.8	9.5	9.6
$\sigma^2_G$	55.862	55.795	102.625	63.228	64.360	48.215	66.188	39.554
$h^2_B$	0.70	0.81	0.91	0.75	0.85	0.84	0.80	0.49
<i>Thousand-kernel weight (g)</i>								
Mean	46.7	52.8	54.9	48.8	54.3	59.5	46.9	51.3
Range	(24.3–73.5)	(26.5–88.8)	(26.4–86.2)	(26.3–74.9)	(25.3–88.3)	(29.5–93.5)	(21.4–78.9)	(41.2–61.1)
CV	16.9	5.5	5.4	5.4	4.5	4.5	5.9	4.9
$\sigma^2_G$	80.789	105.128	104.848	73.451	123.981	151.705	105.630	81.698
$h^2_B$	0.92	0.93	0.92	0.91	0.95	0.95	0.93	0.78

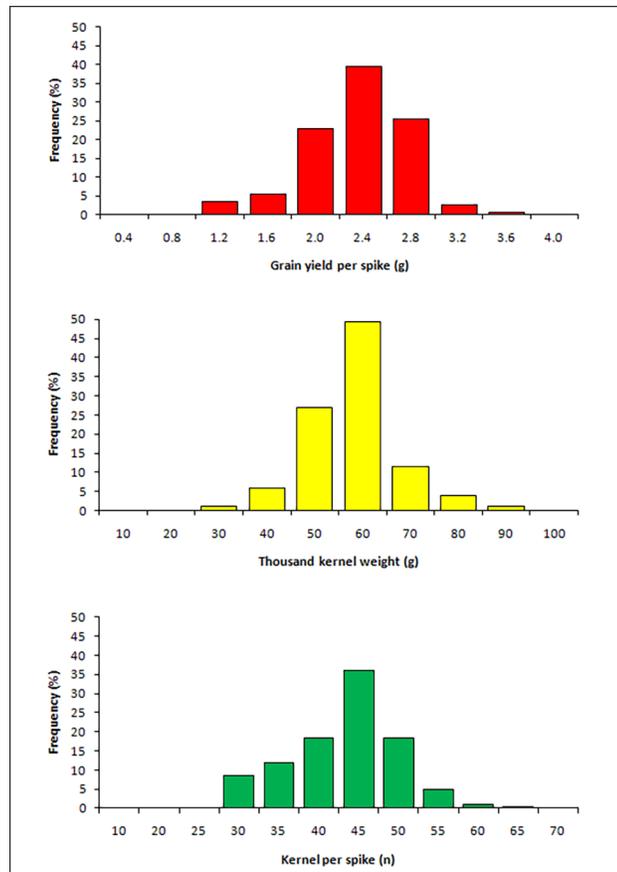
<https://doi.org/10.1371/journal.pone.0190162.t001>

wild and semi-domesticated accessions, breeding lines and modern cultivars. The normal distribution pattern and the large variation indicated the polygenic control of the examined yield components in the tetraploid wheat collection (Fig 1). Estimates of broad sense heritability showed high values in individual environments, but relatively low values for GYS (0.52) and KNS (0.49) and high for TKW (0.78) across environments.

Phenotypic correlations between the three yield components in each environment are reported in Table 2. GYS was significantly positively correlated with KNS ( $r$  values ranging from 0.61 to 0.83) and TKW ( $r$  from 0.53 to 0.71) in all environments and across environments. Correlations between KNS and TKW achieved at each environment ranged from statistically not significant values ( $r = 0.06$  to  $-0.11$ ) to negative and statistically significant coefficients ( $r = -0.23$  at V14). KNS was negatively related to TKW across environments at 0.01P ( $r = -0.22$ ).

### Detection of QTL for grain yield components

The tetraploid wheat collection was analysed with the 90K iSelect genotyping array including 81,587 gene-associated SNPs recently developed by Wang et al. [66]. A total of 15,211 markers mapped in the durum consensus map [69] were retained for the GWAS analysis after removing failed and monomorphic markers, SNPs with more than 10% of missing data and SNPs with a minor allele frequency lower than 0.05. Marker-trait associations were determined by four statistical models: the GLM, the GLM+Q model taking into account the population structure (determined by the principal component analysis), and the MLM incorporating the K matrix (MLM+K) and the K and Q matrices (MLM+K+Q) in order to consider the confounding effects of both population structure and relative kinship, and minimize type 1 errors (false-positive associations). Inspection of Q-Q plots and Manhattan plots for each yield component trait in each environment and across environments (Fig 2 and S1 Fig) indicated strong deviations of the observed  $-\log_{10}(P)$  values from the expected  $-\log_{10}(P)$  distributions for the GLM



**Fig 1. Frequency distributions of mean values of grain yield per spike (g), thousand-kernel weight (g) and kernel number per spike (n) in a tetraploid wheat collection evaluated in seven field experiments.**

<https://doi.org/10.1371/journal.pone.0190162.g001>

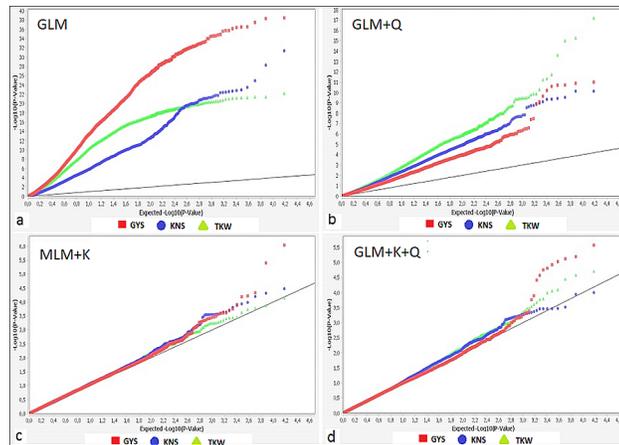
and GLM+Q models, while the closer observed and expected distributions of the  $-\log_{10}(P)$  values in the MLM+K and MLM+K+Q models suggested the reduction of potential spurious marker-trait associations. The last two models produced similar results and the MLM+K model was definitively used in the GWAS analysis.

**Table 2. Correlations between grain yield per spike (GYS), thousand-kernel weight (TKW) and kernel number per spike (KNS) in a tetraploid wheat collection evaluated in seven environments.**

Environment	GYS-TKW	GYS-KNS	TKW-KNS
V09	0.63***	0.70***	-0.02
V10	0.713***	0.72***	0.06
F12	0.53***	0.83***	0.00
V12	0.59***	0.70***	-0.12
V13	0.66***	0.66***	-0.11
G13	0.68***	0.69***	-0.05
V14	0.57***	0.65***	-0.23**
Across environments	0.63***	0.61**	-0.22**

\*\* , \*\*\*: Significant at  $P < 0.01$  and  $P < 0.001$ , respectively

<https://doi.org/10.1371/journal.pone.0190162.t002>



**Fig 2.** Quantile-quantile plots by four statistical GWAS models for yield component traits (mean values across seven environments): a) general linear model (GLM); b) GLM including the Q-matrix derived from the principal component analysis (GLM+Q); c) mixed linear model (MLM) based on the kinship-matrix (MLM+K); d) MLM based on both the K-matrix and the Q-matrix (MLM+K+Q). GYS = Grain yield per spike, KNS = Kernel number per spike, TKW = Thousand kernel weight.

<https://doi.org/10.1371/journal.pone.0190162.g002>

Several QTL significant at  $-\log_{10}(P) \geq 3.0$  in one or two environments were detected for each grain yield component on all 14 chromosomes; these QTL were considered environment-specifics and not reported in the present work because we were interested in stable QTL potentially useful for wheat breeding programs. QTL for GYS, KNS and TKW detected at  $-\log_{10}(P) \geq 3.0$  in at least three environments or two environments and the mean across environments are reported in S3 Table and illustrated in Fig 3. As one of the objective of the present work was to analyze the variation of one component in relation with the other traits change,  $P$  value, additive effect and  $R^2$  value for the three yield components were reported for each significant stable QTL across all environments. Ten stable QTL were detected for GYS, individually accounting for 5.0 to 9.4% of the phenotypic variation and consistent at  $-\log_{10}(P)$  ranging from 3.0 to 5.4. The QTL *QGys.mgb-5B* on the long arm of chromosome 5B was consistent in five environments and the mean across environments. Eight QTL were detected for KNS and 11 QTL for TKW. The  $-\log_{10}(P)$  scores for the KNS QTL ranged between 3.1 to 5.9, with each QTL explaining 5.0–11.6% of the phenotypic variation. The QTL for TKW were detected at  $-\log_{10}(P)$  from 3.0 to 5.6, each accounting 4.8–10.7% of phenotypic variation. All detected QTL were generally consistent at  $-\log_{10}(P) \geq 3.0$  in 3–4 environments and in additional 1–3 environments at sub-threshold  $2.0 < -\log_{10}(P) < 3.0$ , with additive effects in the same direction.

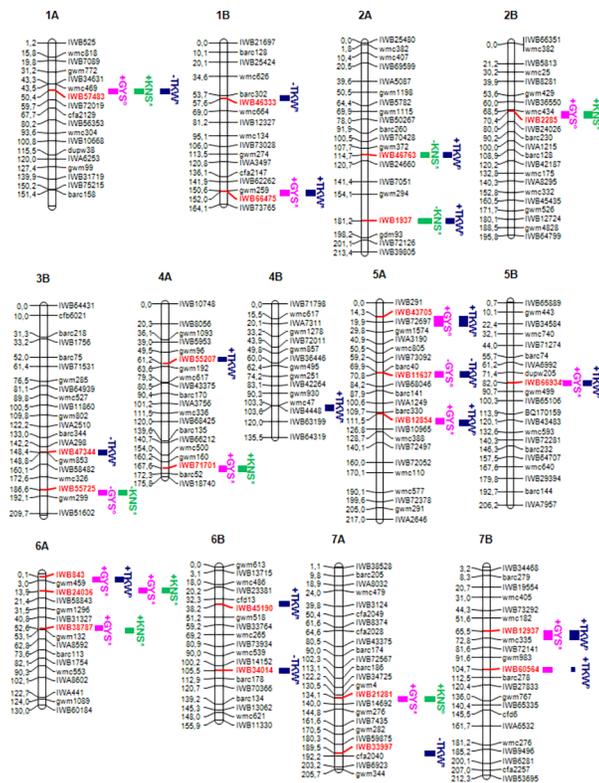
More interesting was the fact that among the eight QTL for KNS, six QTL (*QKns.mgb-1A*, *QKns.mgb-2B*, *QKns.mgb-3B*, *QKns.mgb-4A*, *QKns.mgb-6A.1*, *QKns.mgb-6A.2*) co-located with QTL for GYS at  $-\log_{10}(P) \geq 3.0$  on chromosomes 4A and 6A (2 QTL), and at  $2.0 < -\log_{10}(P) < 3.0$  on chromosomes 1A, 2B and 3B. One additional suggestive KNS QTL was co-located with a significant GYS QTL on chromosome 7A. Among the 11 QTL for TKW, four QTL (*QTKw.mgb-1A*, *QTKw.mgb-1B.2*, *QTKw.mgb-5A.1* and *QTKw.mgb-5A.2*) co-located with QTL for GYS at  $-\log_{10}(P) \geq 3.0$  on chromosomes 1A, and at  $2.0 < -\log_{10}(P) < 3.0$  on chromosomes 1B and 5A (two QTL). Five additional suggestive TKW QTL (*QTKw.mgb-5A.3*, *QTKw.mgb-5B*, *QTKw.mgb-6A*, *QTKw.mgb-7B.1* and *QTKw.mgb-7B.2*) were detected in the same chromosome location with significant QTL for GYS on chromosomes 5A, 5B, 6A and 7B (two QTL). Only the GYS QTL *QGys.mgb-1A* mapped with one QTL for both NKS and TKW (*QKns.mgb-1A* and *QTKw.mgb-1A*, respectively) on chromosome 1A (S3 Table).

### Discussion

During the last decades, grain yield improvement of wheat has been the major focus of most breeding programs in several countries of the world. Significant yield increases were achieved during the 1960s and 1970s transferring the semi-dwarfing *Rht* genes associated to reduced plant height, resistance to lodging, higher number of grain per spike e per unit area, and higher harvest index [71]. Later, the genetic progress obtained by selecting for yield *per se* has been weak, inducing hard effort for breeding work due to the quantitative nature of the trait controlled by a complex genetic system and the strong environmental factors and agronomic management influence. Grain yield increases using indirect selection of yield components has been difficult to achieve due to the complex interactions between plant developmental traits and yield components. Molecular markers are useful tools both for the genetic dissecting of complex traits and for analyzing the relationships of the different yield components at QTL level. Understand the genetic and physiological bases of grain yield may contribute to overcome the negative relationships among some components, and to develop appropriate and efficient breeding strategies for further yield improvement.

### Detection of stable QTL for yield components

The present study was designed to detect stable QTL for KNS and TKW in order to analyze the genetic basis of the interdependence between the above grain yield per spike sub-



**Fig 3. Schematic representation of A and B genome chromosomes of the durum consensus linkage map [69] with map positions of QTL for grain yield component traits.** Each chromosome map is represented by the first and the last SNP marker, and by a SNP marker every about 20 cM. SSR markers have been also inserted every about 20 cM to compare the consensus SNP map with published SSR-based maps. Markers are indicated on the right side and cM distances on the left side of the bar. QTLs are represented by bars on the right of each chromosome bar. QTL names indicate the trait (GYS for Grain Yield per Spike, TKW for Thousand Kernel Weight, KNS for Kernel Number per Spike). The closest SNP marker is indicated in red.

<https://doi.org/10.1371/journal.pone.0190162.g003>

components in a tetraploid wheat collection evaluated in seven replicated field trials. The observed wide phenotypic variation for each yield component (Table 1) can be attributed to the composition of the collection including wild and semi-domesticated accessions, landraces and modern durum cultivars. A total of 15,211 SNP-derived genes were used for the genome wide-association study by using the GLM and the MLM models taking into account the confounding effect of population structure and the relative kinship. Q-Q plots (Fig 2) for the three examined yield-related traits indicated that the models MLM+K e MLM+K+PCs including the kinship matrix detected less potential spurious marker-trait associations than the models GLM and GLM+Q. Moreover, the MLM+K model performed slightly better than MLM+K+Q and was used as the most suitable model for the association mapping analysis of KNS, TKW and GYS, thus confirming other results of GWAS on grain quality traits carried out on the same tetraploid collection [64], and supporting previous findings on the efficiency of the MLM+K model for correcting cryptic relatedness for most of the traits [34,72].

Eighteen stable QTL for the two GYS sub-component traits, distributed on 13 of the 14 chromosomes of the A and B genomes, were detected at  $-\log_{10}(P) \geq 3.0$  in at least three environments or two environments and the mean across environments (Fig 3 and S2 Table). Ten consistent QTL for GYS co-located with four KNS QTL and six TKW QTL. Each detected QTL was often consistent at sub-threshold  $2.0 > -\log_{10}(P) > 3.0$  in the remaining environments, and all had additive effects in the same direction. The consensus high-density linkage map for durum wheat described by Maccaferri et al. [69] was used in the current study as reference map for chromosome localization and map position of SNP markers associated to QTL. This consensus map includes both SNP and SSR markers and enabled us to compare the genomic regions involved in the quantitative expression of grain yield components found in the tetraploid wheat collection with map position of QTL found in previous analysis. Several experiments mapped QTL for yield components on all wheat chromosomes (see reviews by Araus et al. [32], Cui et al. [33] and Gupta et al. [34]). Major QTL associated to thousand grain weight were detected on chromosomes 1B, 2A, 2D, 3A, 3B, 4B, 4D, 5A, 5B, 6A, 7B and 7D [16,35–52], and major QTL for grain number per spike on chromosomes 2B, 3B, 4A, 4B, 4D, 5A, 5D, 7A and 7B [33,36–38,43,48,53–59]. Most QTL were detected in individual environments and/or single mapping population and will hardly be employed in wheat breeding programs with success. Differences in number and map position of QTL detected in the above studies may be attributed to the high number of effective genes controlling grain yield, the different genotypes of the mapping population parental lines, the G by E, the marker coverage of linkage maps used in QTL analyses, the statistical methodologies and the threshold values employed for the statistical significance of marker-trait associations.

While many of the QTL identified in the current study have been described previously, two stable QTL for KNS detected on chromosome arms 2AL and 3BL (*QKns.mgb-2A*, *QKns.mgb-3B*), and four stable QTL for TKW detected on 3BL, 6AS, 6BL and 7BL (*QTKw.mgb-3B*, *QTKw.mgb-6A.1*, *QTKw.mgb-6B.1*, *QTKw.mgb-7B.2*) were totally new. Moreover six QTL for KNS detected on chromosome arms 1AS, 2AL, 2BL, 4AL, 6AS, 6AL and 7AL, and eleven QTL for TKW detected on 1BS, 1BL, 4AL, 4BL, 5AS (three), 5BL, 6BS, 7AL and 7BS validated previously detected QTL in different genetic backgrounds, and they can be considered stable and useful for MAS in breeding programs. Interestingly, the QTL *QKns.mgb-4A* was repeatedly detected in several experiments in different environments and according to Cui et al. [58] should be subjected to map-based cloning.

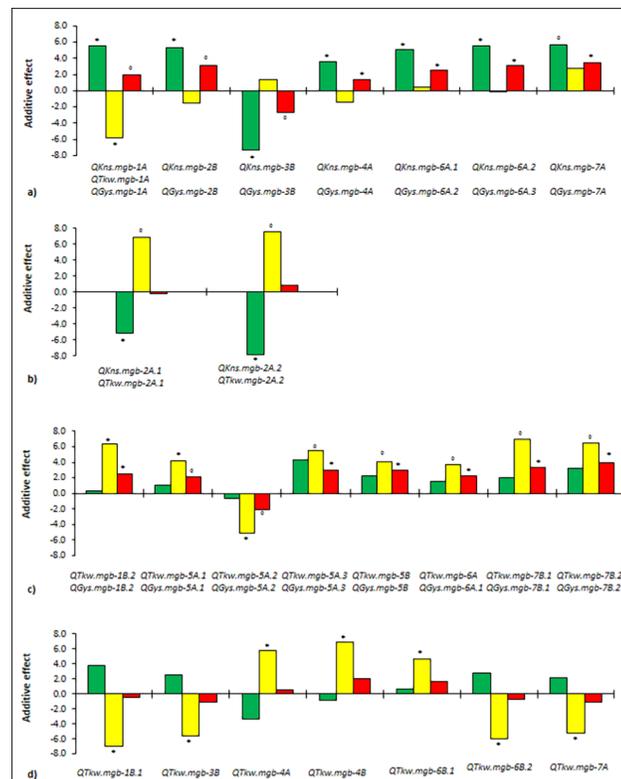
Recently, some candidate genes for significant increases in grain weight were identified in wheat by comparative genomics with rice, such as *TaCwi-A1* [73], *TaGS-D1* [74], *TaSus1* and *TaSus2* [75], *TaGS5-3A* [76], *TaCYP78-A3* [77], *TaCWI-5D* [78], *TaTGW6-A1* [79–80]. Using the same comparative genomic approach, Zheng et al. [74] identified the transcript elongation

factor *TaTEF-7A* having a strong effect on grain number per spike. Some of the above candidate genes, such as *TaCwi-A1* and *TaGW6-A1*, were located in or near the QTL that were found in the current study and in previous experiments; however, further studies on the characterization of these genes on the tetraploid wheat collection and/or suitable biparental populations are needed to ascertain their implication on the detected QTL.

### Relationships between yield components

The primary components of grain yield potential are GYS and number of spikes per unit area. The latter depends on sowing density and it is strongly affected by environmental factors and agro-technique practices. The sub-components of GYS are the KNS and the TKW. In the current study GYS was significantly positively correlated with KNS and with TKW in all environments and across environments, while the correlations between KNS and TKW achieved at each environment ranged from statistically not significant values to negative and statistically significant coefficients (Table 2). Similar results were previously found in segregating materials and germplasm collections in common wheat [14–16,43,48,81–82]. The predominantly negative correlation between KNS and TKW has been the often observed cause of the increases in KNS partially counterbalanced by the reduction of average grain weight, and vice versa.

The QTL detected in the current study could be divided into 4 different groups based on their effects on grain yield per spike (Fig 4):



**Fig 4. Additive effects of QTL for grain yield per spike (expressed in dg, red bars), number of kernels per spike (green bars) and kernel weight (expressed in mg, yellow bars) identified by GWAS in a tetraploid wheat collection (mean values across seven environments).** a) QTL for kernel number per spike associated to QTL for grain yield per spike; b) QTL for kernel number per spike associated to QTL for kernel weight in the opposite direction and without effects on grain yield per spike; c) QTL for kernel weight associated to QTL for grain yield per spike; d) QTL for kernel weight without significant effects on grain yield per spike. The positive or negative additive value refers to the SNP allele with higher frequency. \* Significant at  $-\log_{10}(P) \geq 3.0$ , ° significant at  $2.0 < -\log_{10}(P) < 3.0$ .

<https://doi.org/10.1371/journal.pone.0190162.g004>

1. QTL for KNS (*QKns.mgb-1A*, *QKns.mgb-2B*, *QKns.mgb-3B*, *QKns.mgb-4A*, *QKns.mgb-6A.1*, *QKns.mgb-6A.2*, *QKns.mgb-7A*) with a significant effect on GYS. The effects of these QTL on KNS were independent from TKW variation, or the variation in KNS tended to be only partially compensated by not significant variation in TKW in the opposite direction. The QTL *QKns.mgb-1A* on chromosome 1A increasing KNS was associated with *QTkw.mgb-1A* for a decreasing effect on TKW; the compensating effect was partial and the QTL *QGys.mgb-1A* for GYS was found significant at the sub-threshold  $2.0 > -\log_{10}(P) > 3.0$ . The genomic regions involved in the quantitative expression of KNS found in the tetraploid wheat collection were compared with the map position of QTL found in previous QTL analyses on the same yield components. The QTL *QKns.mgb-1A*, *QKns.mgb-2B*, *QKns.mgb-6A.1*, *QKns.mgb-6A.2* for increased NKS and GYS were located in or near the QTL regions that were found previously by linkage analysis on biparental populations [15,33,45,57–59]. So far, in our knowledge, *QKns.mgb-3B* was not reported in linkage map or association studies before. These seven KNS QTL with direct genetic effects on GYS because of increasing the number of seeds, *i.e.* likely directly linked to spikelet fertility, are particular interesting in wheat breeding since they would allow increasing KNS without a related decrease in grain weight.
2. QTL for KNS (*QKns.mgb-2A.1* and *QKns.mgb-2A.2*) associated to significant QTL with opposite effects on TKW (*QTkw.mgb-2A.1*, *QTkw.mgb-2A.2*). In such cases, the increase of KNS implied a decrease of TKW and the QTL did not contribute to grain yield per spike. *QKns.mgb-2A.1* was previously detected with a similar compensating effect in the same genomic region of chromosome 2A by Yao et al. [83], Blanco et al. [45] and Jia et al. [60]. These QTL have no actual interest for wheat breeding, but it might be interesting to identify their molecular bases to identify the reasons of the physiological trade-off.
3. QTL for TKW (*QTkw.mgb-1B.2*, *QTkw.mgb-5A.1*, *QTkw.mgb-5A.2*, *QTkw.mgb-5A.3*, *QTkw.mgb-5B*, *QTkw.mgb-6A*, *QTkw.mgb-7B.1*, *QTkw.mgb-7B.2*) with significant effects on GYS. The effects of these TKW QTL were independent from KNS, or partially compensated by not significant variation in KNS (with opposite effects), resulting in significant increases of GYS. These QTL could represent genes involved in carbohydrate and/or storage protein synthesis with direct consequences on grain weight. Previously described QTL for grain yield associated with grain weight without pleiotropic effects on grain number [15,33,39,43,45,57–59,84–85] were detected in our study near the QTL *QTkw.mgb-1B.2*, *QTkw.mgb-5A.1*, *QTkw.mgb-5A.2*, *QTkw.mgb-5A.3*, *QTkw.mgb-5B*, *QTkw.mgb-7B.1* and *QTkw.mgb-7B.2* on chromosomes 1B, 5A, 5B and 7B. *QTkw.mgb-6A* was not reported in linkage analysis or in association mapping studies before. These eight TKW QTL are interesting for wheat breeding since they would allow increasing GYS without an associated decrease in grain number.
4. QTL for TKW (*QTkw.mgb-1B.1*, *QTkw.mgb-3B*, *QTkw.mgb-4A*, *QTkw.mgb-4B*, *QTkw.mgb-6B.1*, *QTkw.mgb-6B.2*, *QTkw.mgb-7A*) with no effects on GYS. These TKW QTL likely represent genes with pleiotropic effects in the opposite direction on KNS, or genes that are tightly linked to minor QTL for KNS with reduced effects below the detection threshold, thus resulting in complete compensating effects and no influence on GYS. The likelihood to detect a QTL in the QTL mapping is dependent on the ratio between the variance caused by the QTL effect and the total variance of the trait [86]. Minor QTL with limited effects on a trait could have a low variance, therefore such QTL would remain below the statistical detection threshold. The QTL *QTkw.mgb-1B.1*, *QTkw.mgb-4A*, *QTkw.mgb-4B* and *QTkw.mgb-6B.2* were found to be located in the same genomic regions on chromosomes 1B, 4A, 4B and 6B with QTL for TKW detected previously [15,45,57–59,85]. These seven QTL

increasing TKW without direct consistent effects on GYS could be important in wheat breeding to improve the marketing value of wheat grain as TKW is directly related to milling quality.

Considering the map position of all detected QTL, except for *QKns.mgb-1A*, *QKns.mgb-2A.1* and *QKns.mgb-2A.2*, most of the reported QTL for TKW and for KNS was found to be located in different marker intervals, indicating that KNS and TKW are genetically controlled independently each other. The predominantly negative correlation, not always statistically consistent, between KNS and TKW found in the current study and in previous experiments both in segregating materials and wild and cultivated germplasm [14–16,43,48,81–82], could be the result of genetic and/or environment interactions with the availability of photo-assimilates during grain filling [13,18]. However, this widely accepted explanation on the relationships between grain number and grain weight do not excludes a priori the hypothesis of clusters of tightly linked genes with opposite effects on grain yield and/or single genes with pleiotropic effects. Detailed molecular, physiological and biochemical analyses on specific genetic materials, such as near-isogenic lines, could ascertain if the negative correlation is ascribable to pleiotropic effects, tightly linked QTL or compensating effects on the traits.

### Implications of detected QTL for wheat improvement

Grain yield potential is a complex trait with low to intermediate heritability and strong G by E. Phenotypic selection of single plants and/or in single environments in the early segregating generations has not been effective. Molecular markers associated to agronomical valuable QTL can be successfully used in early generation selection in wheat breeding programs for developing improved cultivars [87]. However, QTL analysis aiming to detect useful alleles to be transferred in commercial cultivars by MAS and/or genomic selection schemes should contemplate whether the QTL have pleiotropic effects or are tightly linked to other QTL that will affect negatively the total performance of the genotypes. Moreover, the QTL could play a relevant role in breeding programs if they are stable in different environments and consistently expressed in different genetic backgrounds. In the current study, we detected QTL for two GYS sub-components consistent in at least three environments or in two environments and in the mean across environments. All the detected QTL were also consistent at sub-threshold  $2.0 > -\log_{10}(P) > 3.0$  in other environments, and all of them had additive effects in the same direction. Most of the QTL for TKW and KNS were found located in different marker intervals, indicating that they are genetically controlled independently by each other. Seven stable QTL for grain number per spike and eight QTL for grain weight were found to be significantly associated to increases of grain yield per spike, indicating that selecting for KNS or TKW or for both yield components could contribute to increase final plant productivity. The identified SNP markers could be usefully deployed in wheat breeding programs for the genetic improvement of grain yield potential. Moreover, QTL accounting for increased TKW without consistent decreasing GYS can be potentially useful to improve the marketing value of wheat grain as TKW is directly related to milling quality.

### Supporting information

**S1 Table. List of accessions of *T. turgidum* subspecies included in the wheat collection.** (PDF)

**S2 Table. Mean squares from the combined analysis of variance for grain yield per spike, thousand kernel weight and kernel per spike across environments in a tetraploid wheat collection evaluated in seven field experiments (accessions evaluated in the seven trials are**

included in the ANOVA).  
(PDF)

**S3 Table. Marker-trait associations for grain yield per spike (GYS), number of kernels per spike (KNS) and thousand kernel weight (TKW) identified by GWAS (model MLM+K) for each of the seven environments (V09, V10, F12, V12, V13, V13\_45, V14) and for the mean across environments in a tetraploid wheat collection.**  
(PDF)

**S1 Fig.** Manhattan plot from MLM+K model for the mean values across environments of grain yield per spike (a), thousand kernel weight (b) and number of kernels per spike (c).  
(TIF)

## Acknowledgments

This research was supported by grants from MIUR, Italy, projects 'PON-01\_01145 – ISCOCEM'.

## Author Contributions

**Conceptualization:** Giacomo Mangini, Agata Gadaleta, Antonio Blanco.

**Data curation:** Giacomo Mangini, Agata Gadaleta, Pasqualina Colasuonno, Iliaria Marcotuli, Antonio M. Signorile, Antonio Blanco.

**Formal analysis:** Giacomo Mangini, Agata Gadaleta.

**Funding acquisition:** Agata Gadaleta.

**Investigation:** Giacomo Mangini, Agata Gadaleta, Pasqualina Colasuonno, Iliaria Marcotuli, Antonio M. Signorile, Rosanna Simeone, Pasquale De Vita, Anna M. Mastrangelo, Giovanni Laidò, Nicola Pecchioni, Antonio Blanco.

**Methodology:** Giacomo Mangini, Iliaria Marcotuli.

**Project administration:** Agata Gadaleta.

**Resources:** Agata Gadaleta, Rosanna Simeone.

**Software:** Giacomo Mangini.

**Supervision:** Agata Gadaleta, Antonio Blanco.

**Validation:** Giacomo Mangini, Agata Gadaleta.

**Writing – original draft:** Giacomo Mangini, Agata Gadaleta, Pasquale De Vita, Anna M. Mastrangelo, Nicola Pecchioni, Antonio Blanco.

**Writing – review & editing:** Giacomo Mangini.

## References

1. De Vita P, Li Destri Nicosia O, Nigro F, Platani C, Riefolo C, Di Fonzo N, et al. Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th. century. *Eur J Agron.* 2007; 26: 39–53.
2. Calderini DF, Reynolds MP, Slafer GA. Genetic gains in wheat yield and associated physiological changes during the twentieth century. In: Satorre E, Slafer G, editors. *Wheat: ecology and physiology of yield determination.* New York: Haworth Press Inc; 1999. pp 503.
3. Peltonen-Sainio P, Jauhiainen L, Laurila IP. Cereal yield trends in northern European conditions: Changes in yield potential and its realisation. *Field Crops Res.* 2009; 110: 85–90.

4. Sadras VO, Lawson C. Genetic gain in yield and associated changes in phenotype, trait plasticity and competitive ability of South Australian wheat varieties released between 1958 and 2007. *Crop Pasture Sci.* 2011; 62: 533–549.
5. Xiao YG, Qian ZG, Wu K, Liu JJ, Xia XC, Ji WQ, et al. Genetic gains in grain yield and physiological traits of winter wheat in Shandong Province, China, from 1969 to 2006. *Crop Sci.* 2012; 52: 44–56.
6. Aisawi KAB, Reynolds MP, Singh RP, Foulkes MJ. The physiological basis of the genetic progress in yield potential of CIMMYT spring wheat cultivars from 1966 to 2009. *Crop Sci.* 2015; 55: 1749–1764
7. Shearman VJ, Sylvester-Bradley R, Scott RK, Foulkes MJ. Physiological processes associated with wheat yield progress in the UK. *Crop Sci.* 2005; 45: 175–185.
8. Foulkes MJ, Reynolds MP, Sylvester-Bradley R. Genetic improvement of grain crops: yield potential. In: Sadras V, Calderini D, editors. *Crop Physiology: Applications for genetic improvement and agronomy.* Amsterdam: Academic Press; 2009. pp. 355–385.
9. Parry MAJ, Reynolds M, Salvucci ME, Raines C, Andralojc PJ, Zhu XG, et al. Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *J Exp Bot.* 2011; 62: 453–467.
10. Reynolds M, Foulkes J, Furbank R, Griffiths S, King J, Murchie E, et al. Achieving yield gains in wheat. *Plant Cell Environ.* 2012; 35: 1799–1823. <https://doi.org/10.1111/j.1365-3040.2012.02588.x> PMID: [22860982](https://pubmed.ncbi.nlm.nih.gov/22860982/)
11. Joseph KDSM Alley MM, Brann DE Gravelle WD. Row spacing and seeding rate effects on yield and yield components of soft red winter wheat. *Agron J.* 1985; 77: 211–214.
12. Lloveras J, Manent J, Viudas J, Lopez A, Santiveri F. Seeding rate influence on yield and yield components of irrigated winter wheat in a Mediterranean climate. *Agron J.* 2004; 96: 1258–1265.
13. Foulkes MJ, Slafer GA, Davies WJ, Berry PM, Sylvester-Bradley R, Martre P, et al. Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. *J Exp Bot.* 2011; 62: 469–86.
14. Borrás L, Slafer GA, Otegui ME. Seed dry weight response to source-sink manipulations in wheat, maize and soybean: a quantitative reappraisal. *Field Crops Res.* 2004; 86: 131–146.
15. Zhang H, Chen J, Li R, Deng Z, Zhang K, Liu B, et al. Conditional QTL mapping of three yield components in common wheat (*Triticum aestivum* L.). *Crop J.* 2016; 4: 220–228.
16. Zhang H, Zhang F, Li G, Zhang S, Zhang Z, Ma L. Genetic diversity and association mapping of agronomic yield traits in eighty-six synthetic hexaploid wheat. *Euphytica* 2017; 213: 111.
17. Adams MV. Basis of yield component compensation in crop plants with special reference to the field bean, *Phaseolus vulgaris*. *Crop Sci.* 1967; 7: 505–510.
18. Slafer GA, Calderini DF Miralles DJ. Increasing Yield Potential in Wheat: Breaking the barriers. In: Reynolds M, editor. *Proc CIMMYT International Symposium: Generation of yield components and compensation in wheat: opportunities for further increasing yield potential.* Mexico: CIMMYT 1996. pp. 101–133.
19. Waddington SR, Ransom JK, Osmanzai M, Saunders DA. Improvement in the yield potential of bread wheat adapted to northwest Mexico. *Crop Sci.* 1986; 26: 698–703.
20. Perry MW, D'Antuono MF. Yield improvement and associated characteristics of some Australian spring wheat cultivars introduced between 1860 and 1982. *Aust J Agr Res.* 1989; 40: 457–472.
21. Sayre KD, Rajaram S, Fischer RA. Yield potential progress in short bread wheats in northwest Mexico. *Crop Sci.* 1997; 37: 36–42.
22. Miralles DJ, Slafer GA. Yield, biomass and yield components in dwarf, semi-dwarf and tall isogenic lines of spring wheat under recommended and late sowing dates. *Plant Breed.* 1995; 114: 392–396.
23. Acreche MM, Slafer GA. Grain weight, radiation interception and use efficiency as affected by sink-strength in Mediterranean wheats released from 1940 to 2005. *Field Crop Res.* 2009; 110: 98–105.
24. Morgounov A, Zykin V, Belan I, Roseeva L, Zelenskiy Y, Gomez-Becerra HF, et al. Genetic gains for grain yield in high latitude spring wheat grown in Western Siberia in 1900–2008. *Field Crop Res.* 2010; 117: 101–112.
25. Tian Z, Jing Q, Dai T, Jiang D, Cao W. Effects of genetic improvements on grain yield and agronomic traits of winter wheat in the Yangtze River Basin of China. *Field Crop Res.* 2011; 124: 417–425.
26. Zheng TC, Zhang XK, Yin GH, Wang LN, Han YL, Chen L, et al. Genetic gains in grain yield, net photosynthesis and stomatal conductance achieved in Henan Province of China between 1981 and 2008. *Field Crop Res.* 2011; 122: 225–233.
27. Lopes MS, Reynolds MP, Manes Y, Singh RP, Crossa J, Braun HJ. Genetic yield gains and changes in associated traits of CIMMYT spring bread wheat in a “Historic” set representing 30 years of breeding. *Crop Sci.* 2012; 52: 1123–1131.

28. Giancaspro A, Giove SL, Zito D, Blanco A, Gadaleta A. Mapping QTLs for fusarium head blight resistance in an interspecific wheat population. *Front Plant Sci.* 2016; 7: 1381 <https://doi.org/10.3389/fpls.2016.01381> PMID: 27746787
29. Rafalski JA. Association genetics in crop improvement. *Curr Opin Plant Biol.* 2010; 13: 174–180.
30. Marcotuli I, Houston K, Waugh R, Fincher GB, Burton RA, Blanco A, et al. Genome wide association mapping for arabinoxylan content in a collection of tetraploid wheats. *PLoS One:* 2015; 10(7): e0132787 <https://doi.org/10.1371/journal.pone.0132787> PMID: 26176552
31. Gupta PK, Kulwal PL, Jaiswal V. Association mapping in crop plants: opportunities and challenges. *Adv Genet.* 2014; 85: 109–147.
32. Araus JL, Slafer GA, Royo C, Serret MD. Breeding for yield potential and stress adaptation in cereals. *Crit Rev Plant Sci.* 2008; 27: 377–412.
33. Cui F, Zhao C, Ding A, Li J, Wang L, Li X, Bao Y, Li J, Wang H. Construction of an integrative linkage map and QTL mapping of grain yield-related traits using three related wheat RIL populations. *Theor Appl Genet.* 2014; 127: 659–675.
34. Gupta PK, Balyan HS, Gahlaut V. QTL analysis for drought tolerance in wheat: present status and future possibilities. *Agronomy* 2017; 7:5 <https://doi.org/10.3390/agronomy7010005>
35. Groos C, Robert N, Bervas E, Charmet G Genetic analysis of grain protein content, grain yield and thousand kernel weight in bread wheat. *Theor Appl Genet.* 2003; 106: 1032–1040.
36. Guo Z, Chen D, Alqudah AM, Röder MS, Ganai MW, et al. Genome-wide association analyses of 54 traits identified multiple loci for the determination of floret fertility in wheat. *New Phytol.* 2017; 214: 257–270.
37. Quarrie S, Pekic Quarrie S, Radosevic R, Rancic D, Kaminska A, et al. Dissecting a wheat QTL for yield present in a range of environments: from the QTL to candidate genes. *J Exp Bot.* 2006; 57: 2627–2637.
38. Quarrie SA, Steed A, Calestani C, Semikbodskii A, Lebreton C, Chinoy C, et al. A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor Appl Genet.* 2005; 110: 865–880.
39. Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang Y. Dissecting gene × environmental effects on wheat yields via QTL and physiological analysis. *Euphytica* 2007; 154: 401–408.
40. Cuthbert JL, Somers DJ, Brûlé-Babel AL, Brown PD, Crow GH Molecular mapping of quantitative trait loci for yield and yield components in spring wheat (*Triticum aestivum* L.). *Theor Appl Genet.* 2008; 117: 595–608.
41. Maccaferri M, Sanguineti MC, Corneti S, Ortega JL, Salem MB, Bort J, et al. Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics* 2008; 178: 489–511.
42. Röder MS, Huang XQ, Börner A. Fine mapping of the region on wheat chromosome 7D controlling grain weight. *Funct Integr Genomics* 2008; 8: 79–86.
43. McIntyre CL, Mathews KL, Rattey A, Chapman SC, Drenth J, Ghaderi M, et al. Molecular detection of genomic regions associated with grain yield and yield-related components in an elite bread wheat cross evaluated under irrigated and rainfed conditions. *Theor Appl Genet.* 2010; 120: 527–541.
44. Sun XC, Marza F, Ma HX, Carver BF, Bai GH. Mapping quantitative trait loci for quality factors in an interclass cross of US and Chinese wheat. *Theor Appl Genet.* 2010; 120: 1041–1051.
45. Blanco A, Mangini G, Giancaspro A, Giove S, Colasuonno P, Simeone R, et al. Relationships between grain protein content and grain yield components through QTL analyses in a RIL population derived from two elite durum wheat cultivars. *Mol Breed.* 2012; 30: 79–92.
46. Mir R, Kumar N, Jaiswal V, Girdharwal N, Prasad M, Balyan HS, et al. Genetic dissection of grain weight in bread wheat through quantitative trait locus interval and association mapping. *Mol Breed.* 2012; 29: 963–972.
47. Wang L, Ge H, Hao C, Dong Y, Zhang X. Identifying loci influencing 1,000-kernel weight in wheat by microsatellite screening for evidence of selection during breeding. *PLoS One* 2012; 2: e29432 <https://doi.org/10.1371/journal.pone.0029432> PMID: 22328917
48. Wu X, Chang X, Jing R. Genetic insight into yield-associated traits of wheat grown in multiple rainfed environments. *PLoS One* 2012; 7: e31249. <https://doi.org/10.1371/journal.pone.0031249> PMID: 22363596
49. Huang X, Zhao Y, Wie X, Li C, Wang A, Zhao Q, et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet.* 2012; 44: 32–39.
50. Huang Y, Kong Z, Wu X, Cheng R, Yu D, Ma Z. Characterization of three wheat grain weight QTLs that differentially affect kernel dimensions. *Theor Appl Genet.* 2015; 128: 2437–2445. <https://doi.org/10.1007/s00122-015-2598-6> PMID: 26334548

51. Zhang J, Dell B, Biddulph B, Drake-Brockman F, Walker E, Khan N, et al. Wild-type alleles of *Rht-B1* and *Rht-D1* as independent determinants of thousand-grain weight and kernel number per spike in wheat. *Mol Breed*. 2013; 32: 771–783.
52. Mohler V, Albrecht T, Hartl L. Inventory and effects of causal genes for thousand grain weight in winter wheat. In: 65. Tagung der Vereinigung der Pflanzenzüchter und Saatgutkaufleute Österreichs (Raumberg-Gumpenstein), 2014, pp. 47.
53. Su J, Zheng Q, Li H, Li B, Jing R, Tong Y, et al. Detection of QTLs for phosphorus use efficiency in relation to agronomic performance of wheat grown under phosphorus sufficient and limited conditions. *Plant Sci*. 2009; 176: 824–836.
54. Wang RX, Hai L, Zhang XY, You GX, Yan CS, Xiao SH. QTL mapping for grain filling rate and yield related traits in RILs of the Chinese winter wheat population Heshangmai 3 × Yu8679. *Theor Appl Genet*. 2009; 118: 313–325.
55. Wang JS, Liu WH, Wang H, Li LH, Wu J, Yang XM, et al. QTL mapping of yield-related traits in the wheat germplasm 3228. *Euphytica* 2011; 177: 277–292.
56. Deng S, Wu X, Wu Y, Zhou R, Wang H, Jia J, et al. Characterization and precise mapping of a QTL increasing spike number with pleiotropic effects in wheat. *Theor Appl Genet*. 2011; 122: 281–289.
57. Gao F, Wen W, Liu J, Rasheed A, Yin G, Xia X, Wu X, He Z. Genome-wide linkage mapping of QTL for yield components, plant height and yield-related physiological traits in the Chinese wheat cross Zhou 8425B/Chinese Spring. *Front Plant Sci*. 2015; 6, 1099. <https://doi.org/10.3389/fpls.2015.01099> PMID: 26734019
58. Cui F, Zhang N, Fan X, Zhang W, Zhao C, Yang L, Pan R, Chen M, Han J, Zhao X, Ji J, Tong Y, Zhang H, Jia J, Zhao G, Li J. Utilization of a Wheat660K SNP array-derived high-density genetic map for high-resolution mapping of a major QTL for kernel number. *Scientific Reports*. 2017; 7: 3788. <https://doi.org/10.1038/s41598-017-04028-6> PMID: 28630475
59. Jia H, Wan H, Yang S, Zhang Z, Kong Z, Xue S, et al. Genetic dissection of yield-related traits in a recombinant inbred line population created using a key breeding parent in China's wheat breeding. *Theor Appl Genet*. 2013; 126: 2123–2139. <https://doi.org/10.1007/s00122-013-2123-8> PMID: 23689745
60. Nadolska-Orczyk A, Rajchel IK, Orczyk W, Gasparis S. Major genes determining yield-related traits in wheat and barley. *Theor Appl Genet*. 2017; 130: 1081–1098.
61. Laidò G, Mangini G, Taranto F, Gadaleta A, Blanco A, Cattivelli L, et al. Genetic diversity and population structure of tetraploid wheats (*Triticum turgidum* L.) estimated by SSR, DArT and pedigree data. *PLoS One* 2013; 8(6): e67280. <https://doi.org/10.1371/journal.pone.0067280> PMID: 23826256
62. Laidò G, Marone D, Russo MA, Colecchia SA, Mastrangelo AM, De Vita P, et al. Linkage disequilibrium and genome-wide association mapping in tetraploid wheat (*Triticum turgidum* L.). *PLoS One* 2014; 9 (4): e95211. <https://doi.org/10.1371/journal.pone.0095211> PMID: 24759998
63. Marcotuli I, Houston K, Schwerdt JG, Waugh R, Fincher GB, Burton RA, et al. Genetic diversity and genome wide association study of  $\beta$ -glucan content in tetraploid wheat grains. *PLoS One* 2016; 11(4): e0152590. <https://doi.org/10.1371/journal.pone.0152590> PMID: 27045166
64. Colasuonno P, Lozito ML, Marcotuli I, Nigro D, Giancaspro A, Mangini G, et al. The carotenoid biosynthetic and catabolic genes in wheat and their association with yellow pigments. *BMC Genomics* 2017; 18: 122. <https://doi.org/10.1186/s12864-016-3395-6> PMID: 28143400
65. Sharp PJ, Kreis M, Shewry PR, Gale MD. Location of b-amylase sequences in wheat and its relatives. *Theor Appl Genet*. 1988; 75: 286–290.
66. 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: 787–796. <https://doi.org/10.1111/pbi.12183> PMID: 24646323
67. Akhunov V, Nicolet C, Dvorak J. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theor Appl Genet*. 2009; 119: 507–517.
68. GenStat for Windows 14th Edition (2011). VSN International, Hemel Hempstead, UK. Web page: [GenStat.co.uk](http://GenStat.co.uk)
69. Maccaferri M, Ricci A, Salvi S, Milner SG, Noli E, Martelli PL, et al. 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*. 2014; 13: 648–663.
70. Voorrips RE MapChart: Software for the graphical presentation of linkage maps and QTLs. *J Hered*. 2002; 93: 77–78.
71. Fischer RA, Stockman YM. Increase kernel number in Norin 10-derived dwarf wheat: evaluation of the cause. *Aust J of Plant Physiol*. 1986; 13: 767–784.

72. Riedelsheimer C, Lisek J, Czedik-Eysenberg A, Sulpice R, Flis A, Grieder C, et al. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc Natl Acad Sci*. 2012; 109(23): 8872–8877. <https://doi.org/10.1073/pnas.1120813109> PMID: 22615396
73. 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 Breed*. 2012; 29: 43–52.
74. Zheng J, Liu H, Wang Y, Wang L, Chang X, Jing R, et al. *TEF-7A*, a transcript elongation factor gene, influences yield-related traits in bread wheat (*Triticum aestivum* L.). *J Exp Bot*. 2014; 65(18): 5351–5365. <https://doi.org/10.1093/jxb/eru306> PMID: 25056774
75. 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.
76. 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: 1269–1280. <https://doi.org/10.1111/pbi.12492> PMID: 26480952
77. Ma M, Wang Q, Li Z, Cheng H, Li Z, Liu X, et al. Expression of *TaCYP78A3*, a gene encoding cytochrome P450 CYP78A3 protein in wheat (*Triticum aestivum* L.), affects seed size. *Plant J*. 2015; 83: 312–325.
78. Jiang Y, Jiang Q, Hao C, Hou J, Wang L, Zhang H, et al. A yield-associated gene *TaCWI*, in wheat: its function, selection and evolution in global breeding revealed by haplotype analysis. *Theor Appl Genet*. 2015; 128: 131–143.
79. Hanif M, Gao F, Liu J, Wen W, Zhang Y, Rasheed A, et al. *TaTGW6-A1*, an ortholog of rice *TGW6*, is associated with grain weight and yield in bread wheat. *Mol Breed*. 2016; 36: 1. <https://doi.org/10.1007/s11032-015-0425-z>
80. Hu MJ, Zhang HP, Cao JJ, Zhu XF, Wang SX, Jiang H, et al. Characterization of an IAA-glucose hydrolase gene *TaTGW6* associated with grain weight in common wheat (*Triticum aestivum* L.) *Mol Breed*. 2016; 36: 25
81. Kuchel H, Williams KJ, Langridge P, Eagles HA, Jefferies SP Genetic dissection of grain yield in bread wheat: I. QTL analysis. *Theor Appl Genet*. 2007; 115: 1029–1041.
82. Maphosa L, Langridge P, Taylor H, Parent B, Emebiri LC, Kuchel H, et al. Genetic control of grain yield and grain physical characteristics in a bread wheat population grown under a range of environmental conditions. *Theor Appl Genet*. 2014; 127: 1607–1624.
83. Yao J, Wang LX, Liu LH, Zhao CP, Zheng YL. Association mapping of agronomic traits on chromosome 2A of wheat. *Genetica* 2009; 137: 67–75.
84. Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, et al. A genetic framework for grain size and shape variation in wheat. *Plant Cell* 2010; 22: 1046–1056.
85. Zanke CD, Ling J, Plieske J, Kollers S, Ebmeyer E, Korzun V, et al. Analysis of main effect QTL for thousand grain weight in European winter wheat (*Triticum aestivum* L.) by genome-wide association mapping. *Front Plant Sci*. 2015; 6: 644. <https://doi.org/10.3389/fpls.2015.00644> PMID: 26388877
86. Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 1989; 121: 185–199.
87. Collard BC, J Mackill D. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos Trans R Soc Lond B Biol Sci*. 2008; 363: 557–572.