Universitat Politècnica de València

Council For Agricultural Research And Economics (CREA) - Research Centre For Genomics And Bioinformatics

"Identification and mapping of quantitative trait loci for kernel-related traits in a durum wheat x *T. dicoccum* segregating population"

Erasmus Mundus Master's degree in Plant Breeding

2019-2021

Final Master's Thesis presented by:

Ana Paola Valladares García

Supervised by:

Dr. Elisabetta Mazzucotelli

Dr. Santiago Vilanova

Dr. Luigi Cattivelli

September 13th, 2021

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

COMAV
Instituto de Conservación y Mejora de la Agrodiversidad Valenciana

El/los Doctor/es D./Dª. Elisabetta MAZZUCOTELLI/Luigi CATTIVELLI/Santiago VILANOVA profesor/es del Master Oficial en Mejora Genética Vegetal, en calidad de director/es del Trabajo de Fin de Máster, por la Presente,

RECONOCE/N

Que el Trabajo Fin de Máster realizado por el/la alumno/a D./Dª.: Ana Paola VALLADARES GARCÍA, con el título: "Identification and mapping of quantitative trait loci for kernel-related traits in a durum wheat x T.dicoccum segregating population" y realizado bajo mi/nuestra dirección, reúne las condiciones necesarias para completar la formación del alumno y por tanto,

AUTORIZA/N

La presentación del citado Trabajo Final de Máster para su defensa ante el correspondiente Tribunal.

Y para que así conste a los efectos oportunos así lo firma/n,

SANTIAGO VILANOVA NAVARRO - NIF: 22566711P

Digitally signed by SANTIAGO VILANOVA NAVARRO - NIF: 22566711P
DN: CN=SANTIAGO VILANOVA NAVARRO - NIF:22566711P, SERIALNUMBER=22566711P, G=SANTIAGO, SN=VILANOVA NAVARRO, OU=Ciudadanos, O=ACCV, C=ES
Reason: I am the author of this document
Location: your signing location here
Date: 2021.09.09 09:41:36+02'00'
Foxit Reader Version: 10.1.1

Luigi Cattivelli
07.09.2021
13:08:26
GMT+01:00

Fdo:

Máster Oficial en Mejora Genética Vegetal

Valencia, 07 de Septiembre de 2021

Acknowledgements

Fiorenzuola d'Arda, 2021

I dedicate this thesis to my family, my mother, father and brother, who taught me to persevere and always give the best version of me, anywhere. Their encouragement to follow my dreams and unconditional love accompanied me through this master and there are no words that can explain how thankful I am for that.

I would like to express my deep gratitude to all the CREA - Centro di ricerca Genomica e Bioinformatica - members, especially Dr. Luigi, who opened the doors of this institution to me; Dr. Elisabetta and Dr. Francesca, for their patient guidance and support given in every step of this thesis; and finally to Alessia and Afi for their sincere friendship.

Special thanks to the emPLANT consortium and professors in France and Spain, which were always there supporting us, even during a pandemic, they performed amazingly. Thank you to Dr. Santiago, my tutor from the UPV, for helping me during the administrative process of this thesis.

Last but not least, I would like to thank my emPLANT family and especially my best/boy-friend Sergio, for their love and for all the experiences shared, which enriched my life both professionally and personally.

“Identification and mapping of quantitative trait loci for kernel-related traits in a durum wheat x *T. dicoccum* segregating population”

Abstract

Durum wheat (*Triticum turgidum* ssp. *durum*), which constitutes the raw material of pasta, is the 10th most important cereal worldwide. A key goal in order to meet its upcoming demand while coping with climate change, is to understand the genetic control behind thousand kernel weight (TKW), a major component of grain-yield. A strategy to achieve this is to explore new genetic resources as domesticated emmer (*T. turgidum* ssp. *dicoccum*) to discover favorable alleles that affect kernel morphological factors, which have a determining role on TKW. Therefore, the present study aimed to explore the genetic network responsible for kernel size components (length, width, perimeter and area) and kernel shape (width-length ratio and form coefficient) and their relationships with kernel weight and heading date. QTL mapping was performed on a segregating population of 110 recombinant inbred lines, derived from a cross between *T. dicoccum* accession MG5323 x *T. durum* cv. Latino, evaluated in 4 different environments. A total of 20 QTL were found environmentally stable and further grouped in 6 clusters on chromosomes 2B, 3A, 3B, 4B, 6B and 7A. Among them, a QTL cluster on 4B chromosome was associated with kernel size traits and TKW, where the parental MG5323 contributed the favorable allele, highlighting its potential to improve durum wheat germplasm. Further, the physical positions of the clusters, defined by the projection on the *T. durum* reference genome, overlapped with already known genes, such as *BIG GRAIN PROTEIN 1* on chromosome 4B. These results might provide genome-based guidance for the efficient exploitation of *T. dicoccum* variability in wheat, possibly through yield-related molecular markers.

Keywords: Durum wheat, *Triticum dicoccum*, Quantitative Trait Locus, kernel size and shape, kernel weight

Resumen

El trigo duro (*Triticum turgidum* ssp. *durum*), materia prima de la pasta, es el décimo cereal más importante a nivel mundial. A fin de satisfacer su futura demanda a pesar del cambio climático, es necesario entender la base genética de uno de los principales componentes de su rendimiento, el peso del grano (en inglés TKW). Una estrategia para alcanzar dicho objetivo, es el estudio de nuevos recursos genéticos, como el farro (*T. turgidum* ssp. *dicoccum*), con el fin de descubrir alelos favorables que controlan los caracteres morfológicos del grano, los cuales tienen un rol determinante sobre el peso del grano. Con dicho contexto, el presente estudio tuvo como objetivo explorar la red genética responsable de los factores del tamaño (longitud, ancho, perímetro y área) y forma (relación ancho-longitud y coeficiente de forma) del grano y sus relaciones con el peso del grano y la fecha de espigado. Se realizó un mapeo de locus de rasgo cuantitativo (QTL) en una población segregante de 110 líneas endogámicas recombinantes, derivadas de la cruza entre *T. dicoccum* accessión MG5323 x *T. durum* variedad Latino, y cultivada en 4 ambientes. Un total de 20 QTL fueron detectados estables entre los ambientes y agrupados en 6 clusters en los cromosomas 2B, 3A, 3B, 4B, 6B y 7A. Entre ellos, destaca el cluster localizado en el cromosoma 4B y relacionado con los caracteres del tamaño del grano y el peso del grano, cuyo alelo favorable fue donado por el parental *T. dicoccum.* Las posiciones físicas de los clusters, definidas por la proyección de los marcadores en el genoma de referencia, coincidieron con las posiciones de genes ya descritos, por ejemplo, *BIG GRAIN PROTEIN 1* en el cromosoma 4B. Dichos resultados proporcionan información genómica de *T. dicoccum* para facilitar su uso como donador de alelos favorables en el trigo, aplicados a la mejora de su rendimiento por medio de marcadores moleculares.

Palabras clave: Trigo duro, *Triticum dicoccum*, locus de rasgo cuantitativo, tamaño y forma de grano, peso del grano

Table of Contents

1. Introduction

1.1. Crop description: origin, biology and importance

The different species of wheat, which make up the Triticum genus, includes diploid *T. monococcum* L. (2n = 14), tetraploid *T. turgidum* L. (2n = 28) and hexaploid *T. aestivum* L. (2n = 42). Currently, the cultivated tetraploid wheat comprises only five subspecies: durum wheat, ssp. *durum* (Desf.); rivet wheat, ssp. *turgidum* L.; hulled emmer, ssp. *dicoccum* (Schrank) Thell.; Ethiopian Wheat, ssp. *aethiopicum* (Jakubz.) and Khorasan wheat, ssp. *turanicum* (Jakubz.) (Bozzini et al., 2012; De Vita & Taranto, 2019).

Among them, durum wheat (**Tab. 1**) is the most cultivated subspecies as it is an integral component of the Mediterranean diet. It is the primary wheat for the elaboration of pasta, hence it ranks the second most-cultivated wheat, after common wheat (Maccaferri et al., 2015). The center of origin of this cereal is in the 'Fertile Crescent', constituted by Iraq, Syria, Lebanon, Palestine, Israel, Jordan, Turkey and Iran. From here, *T. durum* expanded to Europe, Africa, Asia, and America, and it is today cultivated across the globe (De Vita & Taranto, 2019; Martínez-Moreno et al., 2020).

Cultivated emmer (*T. dicoccum*) is the direct ancestor of durum wheat, which harbors mutations in the *Q* locus, located on chromosome 5A, and in the *Br* locus, on chromosome 3A and 3B, giving way to easy threshing and non-brittle rachis, respectively. Therefore, these mutations allowed the selection and domestication of durum wheat about 10,000 years ago (De Vita & Taranto, 2019; Martínez-Moreno et al., 2020). Important to mention that the origin of the mutation at the *Tg* (tenacious glumes) locus, on chromosome 2A and 2B, which also give rise to free-threshing wheat, is yet unclear. There is no certainty if it occurred during the transition of wild to cultivated emmer, or from cultivated emmer to durum wheat (Faris et al., 2014).

Table 1. Taxonomy of Durum wheat (*ITIS Standard Report Page: Triticum Durum*, 2021)**.**

Durum wheat	
Kingdom	*Plantae*
Division	*Tracheophyta*
Class	*Magnoliopsida*
Superorder	*Lilianae*
Order	*Poales*
Family	*Poaceae*
Genus	*Triticum L.*
Species	*Triticum durum Desf.*

Moreover, tetraploid wheats are disomic allopolyploids, derived from a natural intergeneric hybridisation and polyploidisation event involving the cross between diploid *T. urartu* (genome AA) with an unknown diploid specie related to *Aegilops speltoides* (genome BB). This event led to the actual durum wheat genomes AABB (2n=4x=28) with seven groups of homoeologous chromosomes. Further, the assembly of the genome of durum wheat cultivar Svevo has been described to have a total genome size of 10.45 gigabase (Gb) (Maccaferri et al., 2019).

In detail, durum wheat is a monocotyledonous, mid-tall annual grass with flat leaf blades and a terminal floral spike consisting of perfect flowers; it can produce a mean of three tillers in addition to the main shoot. The inflorescence is a spike with a rachis bearing spikelets, which are formed by two glumes (bracts) enclosing two to five fertile florets, that potentially will produce a one-seeded fruit called caryopsis (Kadkol & Sissons, 2016). These flowers are mostly closed (cleistogamous) and self-fertilized (autogamous), with exceptions (open florets and cross-pollination) as a survival mechanism as a consequence of extreme environmental conditions (Okada et al., 2018).

This crop is widely adapted to semiarid regions and most of its varieties are spring types, hence in southern Europe, North Africa and Australia, these are planted in late autumns or early winter and are harvested in summer. Durum kernels are the largest and hardest from all the other wheats, on this basis the name *durum* comes from the Latin word for hard. Additionally, the kernels are amber-colored with yellow endosperm and contain high protein content and gluten strength. Due to these characteristics, durum is the preferred choice as raw material for pasta (Government of Canada, 2012; Kadkol & Sissons, 2016).

Although durum wheat constitutes only 5–8% of the world wheat production, it is considered as an economically important crop and staple food due to its unique characteristics and diverse use on food products. The main products derived from durum are pasta products (*e.g.* spaghetti, noodles, and macaroni), bulgur (cracked durum wheat), couscous and other semolina-based products, such as the Indian traditional products called *rava idli, upma*, and *halwa*. Whole kernels are also used to make *freekeh* and for the preparation of leavened or unleavened bread (Arriagada et al., 2020; Dhanavath & Rao., 2017; De Vita & Taranto., 2019).

According to the International Grain Council, durum is the 10th most important and commonly cultivated cereal worldwide, planted annually in over almost 17 million ha, with a global production of 38.1 million tonnes in 2019 (Beres et al., 2020; De Vita & Taranto, 2019; Xynias et al., 2020). The European Union (EU) is the largest producer of this crop with 9 million tonnes (in 2018), and followed by Canada, together they account for 60% of the world production. Other producers of durum are Turkey, United States, Algeria, Mexico, Kazakhstan, Syria and India. Furthermore, the cultivation of this cereal is concentrated in the Mediterranean basin, which includes some of the countries that are the largest importers and consumers of durum products. In the EU (**Fig. 1**), Italy is considered the leader of production, which reached an average of 4.26 million tonnes produced in the last decade (in a 1.28 million ha growing area), followed by France with 1.89 million tonnes (0.37 million ha) (De Vita & Taranto., 2019; Xynias et al., 2020).

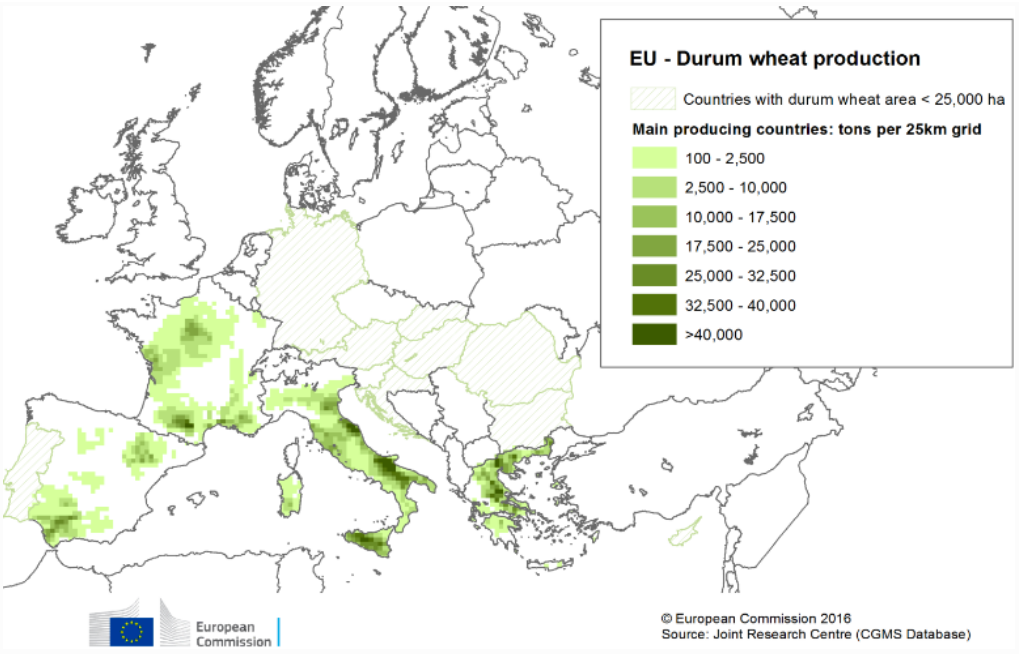


Figure 1. Durum wheat production in the European Union (Willems, 2017).

1.2. Durum wheat breeding and challenges

Since its beginning, in the early decades of the 20th century, durum breeding focused on combining high grain yields and high quality to meet the market demands. Italy was a pioneer in this context, releasing the first variety in 1915 named *Senatore Cappelli* (or *Cappelli*). This variety was a success thanks to its high productivity and quality, covering up to 60% of the Italian growing area from the 1920s to the 1950s. Up to 1955, the breeding programs and scientific research on durum wheat were scarce compared to bread wheat (Martínez-Moreno et al., 2020). Later in 1960, as a result of the Green Revolution driven by the International Maize and Wheat Improvement Center (CIMMYT), new semi-dwarf high yielding varieties were introduced in developing countries (Martínez-Moreno et al., 2020). Due to these efforts, modern durum varieties are shorter, hence more resistant to lodging and better adapted to their actual target environments. However, improving productivity of durum is still essential for the continued viability of its industry, which is constrained by the limited available cropland and the upcoming climate conditions (Arriagada et al, 2020; Kadkol & Sissons, 2016).

According to the European Commision, EU's durum wheat production in 2021 is due to reach 7.3 million tonnes, registering a decline of -3% (year-on-year, **Fig. 2**).

This is a direct consequence of the low yields in France (-20%) in 2020 caused by an excessively dry spring (EC, 2020). Although durum wheat is considered as one of the most drought-tolerant cereal crops, more severe high temperatures and water scarcity, coupled with the emergence of new pests and diseases, jeopardize its cultivation and yield because it is mainly grown under rain-fed conditions. Due to the increased frequency of heat waves in France and Italy, breeding programs are now focusing on increasing biomass and thousand kernel weight, to develop varieties which could indeed outperform under this severe scenario (De Vita & Taranto., 2019; Rehman Arif et al., 2020).

Additionally, efficiency in these breeding programs is required to strengthen food security. The growing demand for cereals is projected to reach 3 billion tonnes by 2050 (Patil et al., 2013), and specifically in the case of wheat products that will account for 20% of protein and calories consumption per capita for a global population of 9.7 billion in 2050 (Beres et al., 2020).

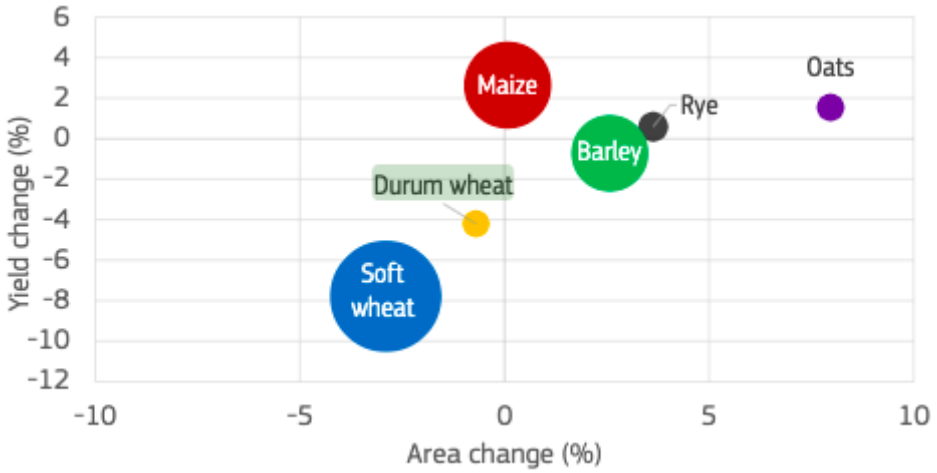


Figure 2. Cereals area and yield change in the European Union (2020/21 vs 2019/20). *The size of the bubbles represents the production levels (EC, 2020).

Presence of genetic diversity is essential for breeding wheat varieties that can overcome these challenges. Nevertheless, the evolution of this crop included multiple bottleneck processes, such as hybridization, polyploidization and domestication, as well as natural and artificial selection, which resulted in a

significant decrease of its genetic variability. These events led firstly to rapid alterations and modifications of the ancestor tetraploid genome, then upon the domestication and human selection, a high level of gene loss occurred and thus, a reduction in the allelic plasticity of durum wheat. Uniformity in the elite cultivars and the displaced cultivation of relatives and landraces have also contributed to the genetic erosion in durum germplasm (Rahman et al., 2020). Hence, new durum varieties are less prone to adapt to fluctuating climates or to tolerate new or re-emerging pests, diseases and weeds (Beres et al., 2020; De Vita & Taranto, 2019; Taranto et al., 2020). Current breeding strategies must necessarily evolve to use genome-based information in order to identify, address and exploit genetic resources. For instance, using wild progenitors might enrich the existing durum genepool, by introducing important wild genes that were changed, modified or lost during domestication. In this way, genetic studies for the identification of QTL/genes for different biotic and abiotic resistance and yield related traits, are important to increase genetic variability of the durum germplasm and consequently improve its productivity (De Vita & Taranto, 2019; Mazzucotelli et al., 2020; Rahman et al., 2020; Xynias et al., 2020).

1.3. Emmer wheat as a promising genetic resource

As mentioned above, emmer wheat (*T. turgidum* ssp. *dicoccum),* known as “farro” in Italy, is the domesticated ancestor of modern bread and durum wheats. In the last few years, a significant number of research studies has been carried out to unleash its potential as material for breeding. Due to these studies it has been demonstrated that emmer could be used to restore genetic variability for both wheats, as it harbors a rich allelic repertoire on many desirable traits (Mohammadi et al., 2021; Rahman et al., 2020). However, there is still a considerable disparity in the number of studies about *T. dicoccum* compared to bread and durum wheats, as depicted in **Fig. 3.**

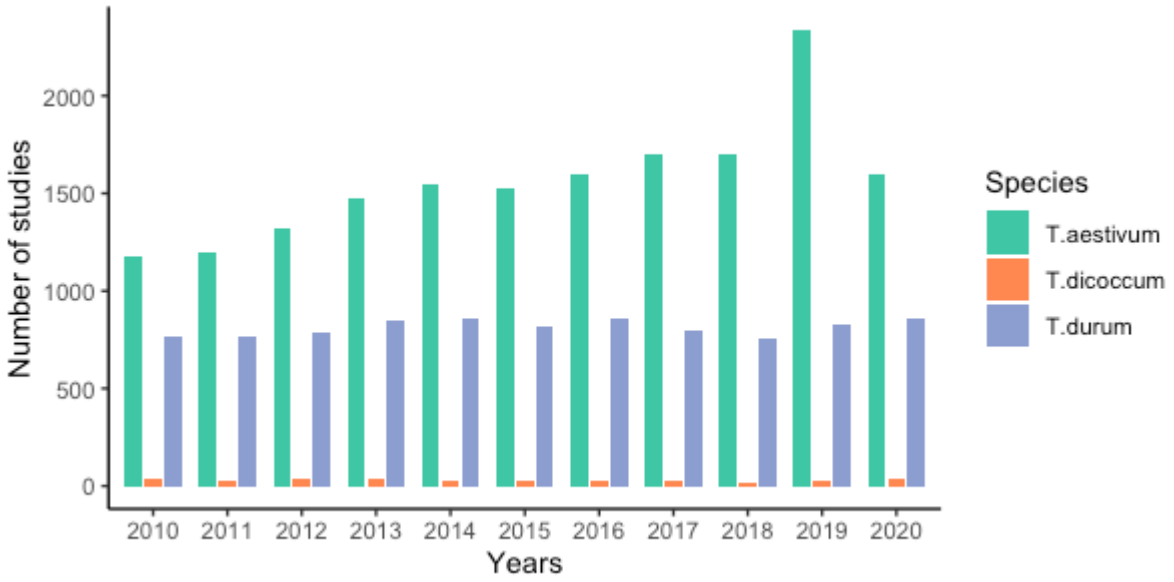


Figure 3. Number of studies on different species of *Triticum* by year of publication. *Retrieved from Scholar.google publication titles in May 2021.

For many centuries, emmer has been appreciated mainly for human food, limited to biscuits and traditional cakes due to its poor qualitative gluten composition. It is possible to find emmer bread in Switzerland and Italy, and it has been also used for animal feed to chickens, horses and pigs (De Vita et al., 2006). Covering 1% of the total world wheat area, it is mainly cultivated in Ethiopia, Iran, Morocco, Spain, Albania, eastern Turkey, Switzerland and Italy (Dhanavath & Rao, 2017).

The potential of emmer as a genetic resource relies in its wider genetic variation (compared to bread and durum wheat) in multiple important traits as: drought and heat tolerance (Konvalina et al., 2010; Fu et al., 2015; Wang et al., 2019), disease resistance (Desiderio et al., 2014; Fatima et al., 2020; Liu et al., 2017; Piarulli et al., 2012; Olivera et al., 2014;), insect resistance (Bassi et al., 2019), thousand kernel weight (Russo et al., 2014; Mangini et al., 2018) and protein content (Dhanavath & Rao., 2017; Nigro et al., 2019), which is likely due to its long cultivation in a large range of eco-geographical conditions (Zaharieva et al., 2010).

Nevertheless, compared to its large diversity, it is still required to extend the collection and conservation of *T. dicoccum*, while developing studies to understand the genome features in this crop. At the end, these efforts might facilitate its exploitation in modern breeding programs (Rahman et al., 2020; Zaharieva et al., 2010).

1.4. Yield decomposition overview

Grain yield, the metric and the key trait for productivity in crops, is determined by a number of interrelated plant and grain characters. Yield decomposition refers to the division of this main, quantitative and complex character into several components or subtraits that can allow a high-throughput and accurate phenotyping. The most important grain yield subtraits are: seed weight and number of seeds/m^2, the latter being dissected in spike number per unit area and kernel number per spike. These yield components of quantitative nature are also highly influenced by the environment, which limits the information about the genetic mechanisms behind them. Thus, the improvement of grain yield remains an ongoing challenge for researchers and breeders (Patil et al., 2013; Sun et al., 2020).

Grain weight, measured in Thousand Kernel Weight (TKW), has a direct impact on grain yield and a positive influence on the price sale of the harvest (Patil et al., 2013; Sun et al., 2020). Seed morphology descriptors, including kernel size and shape (*e.g.* length, perimeter, area), have been demonstrated in determining grain weight and therefore grain yield (Cui et al., 2011; Patil et al., 2013). Indeed, these kernel-related traits underwent major changes during domestication, resulting in larger and shorter seeds (with higher width and lower length) with a positive influence on yield (Gegas et al. 2010). Consequently, a prior breeding strategy is to exploit the positive correlation between kernel size and weight to improve yield (Ma et al., 2021; Mohammadi et al., 2021).

Moreover, kernel shape has been studied for other important quality factors of the semolina industry such as test weight for flour yield and milling quality (Tyagi et al., 2015) and ash distribution (Ficco et al., 2020). It has been demonstrated that the optimum grain morphology is large and spherical (thick) shape as it has the highest endosperm-to-bran ratio, while small-sized kernels had the lowest test weight and semolina yield (Ficco et al., 2020). Additionally, larger kernels have a positive influence on the seedling vigor and early growth in different crops, as in rice and wheat (Avni et al., 2018; Sun et al. 2020).

Using the genetic variability between the progenitors of wheat (as emmer wheat) and durum elite material might help in the identification of major regions in

the genome involved in the control of seed morphology traits. This understanding of the genetic mechanisms that regulate grain size and shape may facilitate the selection of the ideal kernel architecture through molecular markers (Russo et al., 2014). However, the path to achieve this is still uncertain, as there is a serious gap in the knowledge regarding kernel-traits control on durum wheat. The majority of studies have been performed in common wheat (*Triticum aestivum L.*) (*e.g.* Cao, et al., 2019; Gupta et al., 2020; Li et al., 2007; Ma et al., 2021; Tyagi et al., 2015; Xin et al., 2020; Zhang et al., 2010), placing durum wheat in a more unexplored terrain (**Fig.3**).

1.5. Genetic basis of kernel-related traits

A Quantitative Trait Locus/Loci (QTL) is defined as a genomic region whose molecular variation (polymorphism) is associated with phenotypic variation. For quantitative traits, a number of QTL is expected to determine the corresponding phenotype. Each QTL underlines many genes, and advanced genetic analyses are employed to determine which one is responsible for the phenotypic variation. In this way, QTL mapping is the first step towards more in-depth and precise studies (Browman et al., 2013; Giri et al., 2018). QTL mapping is a statistical tool which is performed to identify the regions of interest by correlating three interrelated data structures: the phenotypes, the genotypes and a genetic map (marker map) (Browman et al., 2013). For this purpose, there are two main approaches: linkage mapping and association mapping (also known as genome-wide association study -GWAS-). Briefly, linkage mapping uses information from recombination events between markers within a progeny of known pedigree, for instance, a Recombinant Inbred Line (RIL) population. Meanwhile, association mapping employs historical recombination events, measuring deviation from the random occurrence of alleles in a haplotype in unrelated individuals of unknown pedigree (Bartholomé et al., 2016).

Up to date, a total of approximately 300 QTL located on all 14 tetraploid wheat chromosomes have been described for the different kernel-related traits (size and shape) and kernel weight (**Fig. 4**). This information was retrieved by a literature survey from the publicly available linkage and association mapping studies till July 2021 (Summary on **Tab.7**, based on 18 studies reviewed by Maccaferri et al., 2019 -QTLome-; 7 recent studies reviewed by Arriagada et al., 2020; and the studies from

Desiderio et al., 2019 and Mangini et al., 2021). More in detail, the highest number of QTL detected corresponds to TKW, with almost 200 loci, while less information is found about the genetic basis of kernel size and shape (Arriagada et al., 2020). A total of 94 QTL for kernel size factors (length, width, perimeter and area) and 27 QTL for kernel shape (width length ratio and form coefficient) have been described in these studies. Nevertheless, as some of these major QTL are environment-specific, they should be prudently considered in breeding programs (Arriagada et al., 2020). Note to mention that Heading Date (HD) is another yield-related trait considered in these studies, as it delimits grain weight by marking the transition from spike formation to grain-filling period (Mangini et al., 2021).

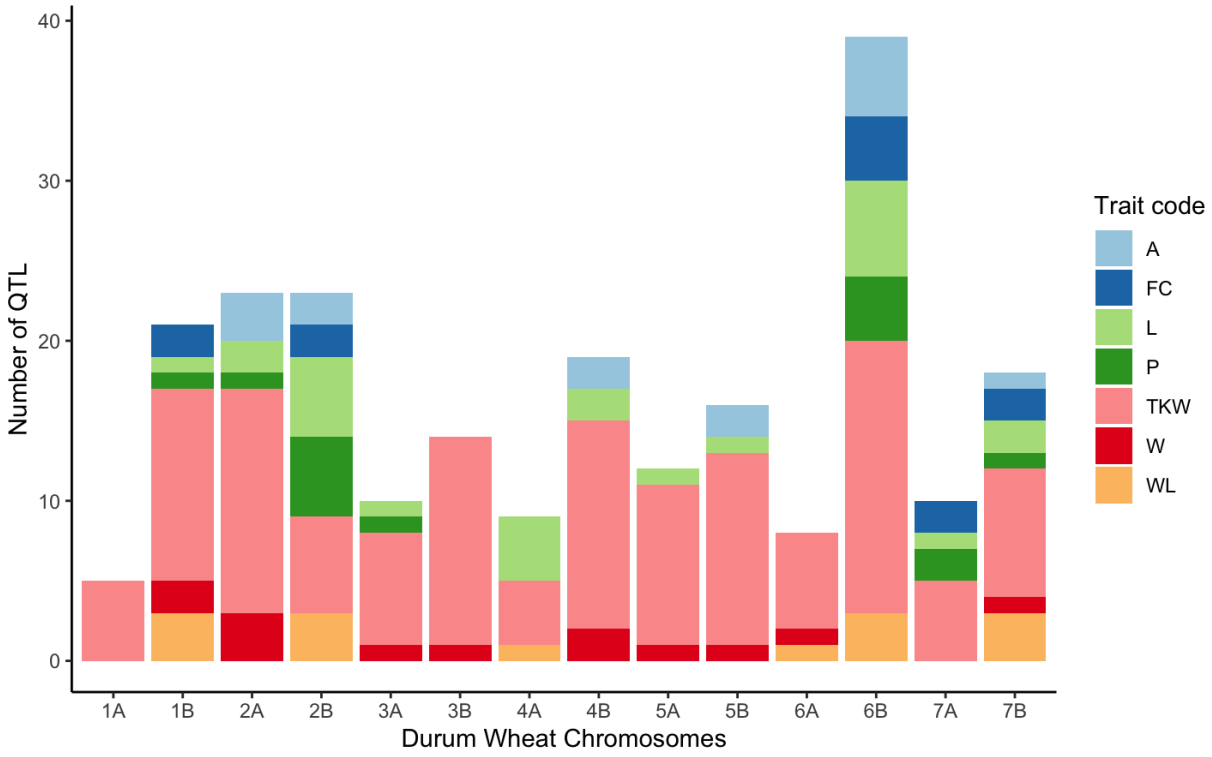


Figure 4. Identified QTL and chromosome positions for kernel-related traits retrieved by a literature survey till July 2021 (Maccaferri et al, 2019; Desiderio et al., 2019; Arriagada et al., 2020 and Mangini et al., 2021). *Acronyms correspond to: A, Area; FC, Form Coefficient; L, Length; P, Perimeter; TKW, Thousand Kernel Weight; W, Width; WL, Width to Length Ratio.

The knowledge for the genes controlling these kernel traits is mostly extended in rice (*Oryza sativa*), with approximately 20 genes already described (**Appendix C**; Avni et al., 2018). The close relationship between wheat and rice has allowed the cloning of the orthologous bread wheat genes, while in durum wheat is yet to be done. For example, *TaGW2*, the orthologue of *OsGW2*, encoding an E3 Ubiquitin Ligase involved in the pathway for cell wall expansion, has been demonstrated to control grain weight and kernel architecture in bread wheat (Zhai et al., 2018). Similarly, *BIG GRAIN 1* has been mapped at chromosome 4B and it is known to be related with auxin transport and regulation of seed growth (Liu et al., 2015). The functions of the genes that play an important role in kernel morphology are highly diverse, including metabolism of growth regulators such as auxins (for example *TaTGW6*); genes determining cell division and proliferation (such as *D1*, *GS2* and *TaGS5*); carbohydrate metabolism as starch and sucrose metabolism pathways (as *TaSus1*, *Tasus2* and *TaCWI-A1*); and genes coding for proteins involved in ubiquitination processes, transcription factors and floral regulators (Arriagada et al., 2020; Desiderio et al., 2019; Mangini et al., 2021).

Specifically, for durum wheat it is worth mentioning that two major advances had opened new avenues for dissecting the functional genomics for different traits. First, the development of a high density marker array (wheat 90K SNP iSelect assay containing 81,587 SNP markers) assisting QTL mapping and GWAS (Wang et al., 2014). Furthermore, an international consortium released a high quality reference sequence of the modern durum wheat cultivar Svevo in 2019 allowing to compare published QTL positions and facilitating the searching of candidate genes (Maccaferri et al., 2019). These tools combined are simplifying the research work on deciphering complex marker-traits associations, which will be translated into more genes being described for different durum wheat characteristics, as kernel-related traits, in the future (Avni et al., 2018; De Vita & Taranto., 2019; Mangini et al., 2021).

2. Objectives

The objective of the current study is to dissect the genetic basis of kernel-related traits in a recombinant inbred line (RIL) population derived from a *T. dicoccum* accession crossed with a durum wheat cultivar, facilitating genomic-information to improve durum wheat germplasm. For this purpose, the following specific objectives, corresponding to consecutive steps of the workflow, are involved:

a. To perform a high throughput phenotyping of kernel size and shape based on digital image analysis from kernel samples obtained in field trials across four different environments. Worth to mention that phenotypic data from three of the environments were already available, completing one environment during the performance of this thesis.
b. To identify quantitative trait loci (QTL) for the kernel size/shape traits, in addition to kernel weight and heading date, and group them in QTL clusters to highlight possible functional relationships among traits.
c. To update the physically anchored QTLome of Maccaferri et al., 2019, related to kernel-related traits, with recently published QTL. Analogously, to update the previous list of genes affecting kernel size, shape and weight identified in bread wheat and/or rice (from Desiderio et al., 2019).
d. To project QTL identified (at step b) on the durum wheat consensus map and the *T. durum* reference genome, allowing the analysis of physical regions: inspection in gene content to hypothesize candidate genes and their comparison with the position of cloned genes (list from step c).
e. To evaluate QTL congruence of identified QTL (at step b) with those already reported by literature (from step c) and highlight stable and novel QTL regions for the mentioned traits.

3. Materials and Methods

3.1. Plant material

A segregant population of 110 RILs developed via single-seed descent of F2 plants obtained from a cross between the accession MG5323 of *Triticum turgidum* ssp. *dicoccum* and the modern durum wheat “Latino” was used for the present study. The *dicoccum* accession MG5323 (USDA accession number PI 94683) shows leaf rust resistance (Desiderio et al., 2014), high resistance to powdery mildew (Piarulli et al., 2012), increased protein yield (De Vita et al., 2006) and high gluten content (Piergiovanni et al., 2009), and has longer and thinner kernels than cv. Latino, as shown in **Fig. 5**.

MG5323 was collected in Armenia and maintained by the National Small Grains Collection (USDA-ARS, Aberdeen, ID, USA). The cultivar Latino (pedigree CAPPELLI/ANHINGA/4/YAKTANA-54//(SEL.14)-NORIN-10/BREVOR/3/ST-64/2*THATCHER) was firstly released by the Federconsorzi (Italy) in 1982 (Desiderio et al., 2014).

Figure 5. Kernel morphology of parental lines used in this study. At the left, *T. turgidum* ssp. *dicoccum* MG5323 and at the right, *T. turgidum* ssp. *durum* cv. Latino.

3.2. Molecular analysis and Genetic map

The Single Nucleotide Polymorphism (SNP) genotyping data was provided by the host institute (CREA, Research Center for Genomics and Bioinformatics). The details of SNP genotyping and dataset filtering were described in Desiderio et al., 2014. Briefly, for each line the SNP genotyping was performed using the wheat 90k iSelect Infinium™ SNP platform developed by Wang et al., 2014. By these data, a high-density genetic map was previously constructed with a total of 10,840 high-quality SNP markers divided in 14 linkage groups corresponding to the 14 durum wheat chromosomes. The overall map length is 2,363.4 cM with the average marker density of 0.23 cM/marker, ranging from 0.16 to 0.35 cM/marker (Desiderio et al., 2014).

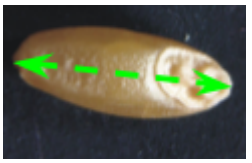
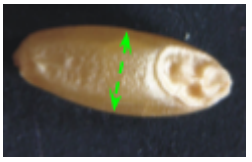
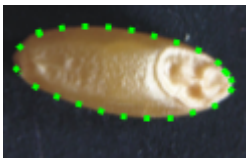
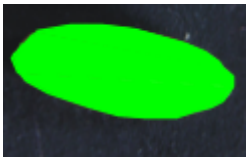
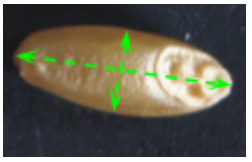
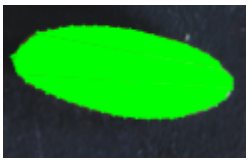
3.3. Experimental design

The parental lines and the RIL population had been previously evaluated in four different environments (location, year): Valenzano (BA, Italy) in 2012-2013 (V13), Bologna (BO, Italy) in 2013-2014 (B14), and in Fiorenzuola d'Arda (PC, Italy) in 2014-2015 (F15) and 2019-2020 (F20) (**Fig. 6**). The experimental design had been a randomized complete block with 2 replications for V13 and B14 trials and 3 replications for F15 and F20. Each experimental unit consisted of a single 1 m row with 20-25 plants each. Trials had been fertilized following the standard agronomic practices for each location, weeds were chemically controlled. For further analysis, 110 RIL had been studied for each environment, except for F20, where only 103 lines had been harvested. Harvest was done at complete maturation, and post-harvest phenotyping was conducted after a long maintenance of all seed samples in a cold room at fixed humidity.

Figure 6. Map of environments considered in this study. *Acronyms correspond to: V13, Valenzano 2013; B14, Bologna 2014; F15 and F20, Fiorenzuola d'Arda 2015 and 2020, respectively.

The phenotypic characterization was performed on a random sample of 100 kernels for each experimental unit. Each sample was scanned by Epson Expression 10000XL. Following, the kernel morphology descriptors, presented in **Tab. 2**., were analyzed by the software WinSEEDLE™ Pro Version 2011a (Regent Instruments Canada Inc.). Additionally, for B14, F15 and F20, Thousand Kernel Weight (TKW) and Heading Date (HD) had been scored for each experimental unit. TKW was recorded as 10-folds the mean of the weight from 3 random samples of 100 kernels, weighted with electronic balance. HD was recorded as the number of days from April 1st to the time when 50% of tillers within a plot have the spike emerged from the flag leaf.

Table 2. Kernel morphological traits considered in this study, ordered from main attributes to derivative attributes (Modified from Desiderio et al., 2019).

Descriptor	Definition	Illustration	Trait category
Length (L)	The straight distance between the two farthest points on the projected image perimeter.		Kernel size (mm)
Width (W)	The maximum width measured perpendicular to length.		
Perimeter (P)	The length of the seed's outline.		
Area (A)	The two-dimensional area occupied by the seed projection.		Kernel size (mm^2)
Width to Length Ratio (WL)	The comparison of the width and length.		Kernel shape
Form Coefficient (FC)	Indicates the seed shape through the formula $4*\pi*A/P^2$, where A is area and P is perimeter; with a value of 0 for a filiform object and 1 for a perfect circle.		

3.4. Statistical analysis

The statistical analyses were all performed in R software (R Core Team, 2020) using the phenotypic data from each environment and for each trait. Student's t-test ($p < 0.05$) was performed to evaluate the parents, in conjunction with descriptive statistical analysis and analysis of variance (One-Way ANOVA, $p < 0.05$), which were used to determine the effect of RILs. Previously, as prerequisites, Shapiro-Wilk test and Levene's test were performed. Broad-sense heritability (H^2)

was estimated according to the R/*metan* package (Olivoto & Lúcio, 2020), using the formula proposed by Wricke and Weber (1986) : H^2=[MSG – MSE/r]/[MSG/r]; where MSG, MSE and r allude to the mean squares of genotypes, mean squares of error and number of replications, respectively for each environment. Next, data across the four environments were analyzed by Two-Way ANOVA (p < 0.05) to assess significance of Genotype (GEN), Environment (ENV), and Genotype × Environment Interaction (GEI). Adjusted overall means were calculated by fitting a model through the Best Linear Unbiased Prediction (BLUP) method using the R/*metan* package, where GEN and GEI were assumed to be random effects and ENV as fixed effect. Additionally, the broad sense heritability was estimated from the BLUP model through the formula H^2=[$\hat{\sigma}^2 g$]/[$\hat{\sigma}^2 g$+$\hat{\sigma}^2 i$+$\hat{\sigma}^2 r$]; where $\hat{\sigma}^2 g$ is the genotypic variance; $\hat{\sigma}^2 i$ is the GEI variance; and $\hat{\sigma}^2 r$ is the residual variance, respectively. Pearson's correlation coefficient analysis among the different examined traits were calculated, using R/*metan,* for all trait combinations based on the data recorded for each environment and across environments (BLUP dataset).

3.5. QTL mapping

For each trait, the R/*qtl* package (Broman et al., 2003) was used for QTL analysis with the mean values for each genotype in each single environment and the BLUP values as adjusted mean values for the combined data. The procedure described by Desiderio et al., (2019) was performed as follows: (i) a permutation test to define the LOD (Logarithm of Odds) significance level with a genome-wide significance level of 5% after 1,000 permutations; (ii) initial scan of the genome was carried out using the simple interval mapping with a 1-cM step and the position of the highest LOD was recorded; (iii) the position and effect of the QTL was then evaluated with the multiple imputation method (composite interval mapping) by running the "sim.geno" command followed by the "fitqtl" command; (iv) the "addqtl" command was used to search for additional QTL. If more QTL were identified for the trait under consideration, the "fitqtl" command was used to test a model containing the QTL and their possible interactions were tested by the "addint" command. If these putative loci remained significant, the "refineqtl" command re-evaluated the QTL positions based on the full model. Through the functions listed above, information is obtained relating to the chromosome that contains the QTL and the

relative position on it, the LOD value, the Phenotypic Variability Explained (PVE or R^2) and the additive effect. The confidence interval (CI) of each QTL was determined as proposed by Darvasi and Soller (1997). Next, for each trait, the QTL found in the different environments and across them were grouped per position (overlapping CIs), as considered to correspond to the same QTL provided that the additive effect is conferred by the same parent. Further, these QTL were named according to the rule "Q + trait code + chromosome.locus number", where Q stands for QTL, trait code to its acronym presented in **Tab. 2**, and last the wheat chromosome on which the corresponding QTL is located. If two QTL are on the same chromosome, a consecutive number (".1,.2,.3,") was added.

3.6. Analysis of physical regions and candidate genes

The most significant results of the QTL mapping, the environmentally stable QTL, were compared with the current state of the art (including a QTLome from previous studies and a compilation of cloned genes for the same traits) through a co-location on the durum wheat reference genome. This comparison procedure included: (i) updating the tetraploid QTLome provided by Maccaferri et al, 2019, with the most recent QTL for the traits of interest (Desiderio et al., 2019; Arriagada et al., 2020 and Mangini et al., 2021); (ii) clustering the QTL identified in the present study for different traits which were co-located in the same/partial overlapping region, and defining the best QTL for LOD and PVE for each cluster; (iii) initial anchoring of peak and flanking SNP markers of each QTL (from this study and from the recently published QTL) on the durum wheat consensus map (Maccaferri et al., 2015), allowing to use more and reliable markers as bridge and thus to increase the consistency and accuracy on the next step of projection; (iv) defining the position interval, in terms of genetic position on the consensus map, for each cluster, and selecting the coinciding/nearest QTL from previous studies; (v) projecting the best QTL of each cluster and coinciding QTL on the *T. durum* reference genome sequence (cv. Svevo) (Maccaferri et al., 2019) by using data provided by the CREA institute, which consisted of Blast matches results corresponding to the SNP's sequences and used to determine the physical positions (best alignments were selected based on the percent of identity, e-value and agreement with the genetic linkage maps); (vi) hypothesizing candidate genes within the physical interval of the

QTL clusters by screening High Confidence Svevo genes based on their functional annotation (previously obtained via blast2GO PRO, available at https://figshare.com/s/2629b4b8166217890971); (vii) screening a compilation of common wheat and/or rice cloned genes with known functions affecting kernel size, shape and kernel weight. To this aim, their sequences were blasted against the durum wheat reference genome to define their genomic positions (**Appendix C.,** updated from Desiderio et al., 2019), which were compared with the physical location of the QTL clusters.

4. Results

4.1. Phenotypic characterization

Mean values of the two parents and of the corresponding RIL mapping population in each environment (Valenzano 2013, Bologna 2014, Fiorenzuola 2015 and 2020) and across environments (BLUP) for the kernel size, kernel shape, grain weight and for heading date are presented in **Tab. 3.** The two parental lines showed significant differences ($p < 0.01$) in most of the traits in each of the environments and across them, except for area, which was less consistent by only being significant in one environment (B14). As expected, MG5323 obtained greater values for kernel length and perimeter, and heading date, and lower values for kernel width, WL ratio, form coefficient and kernel weight compared to parent Latino. This means that MG5323 grains are significantly longer and narrower, while Latino grains are more round affecting its grain weight.

About the RILs mean values, in all environments, the length ranged from 7.9 to 8.2 mm, the width from 3.0 to 3.2 mm, the perimeter from 18.8 to 19.6 mm, the area from 18.6 to 20.3 mm^2, WL ratio from 0.38-0.41, form coefficient from 0.65-0.67, TKW from 44.3 to 53.5 gr, and heading date ranging from 30 to 45 days (**Tab.3**).

As depicted in **Fig. 7,** the frequency distribution analysis was performed for the RILs, indicating that each of the traits followed a normal distribution in each environment and across environments, which was also confirmed by the Shapiro-Wilk tests (not shown here). This distribution suggests the contribution of several loci controlling the phenotypic variation for each trait (quantitative nature), including HD. Additionally, high transgressive segregation was observed for all traits, including TKW, which implies the presence of superior alleles for the kernel-related traits in both parents.

Table 3. Summary of the phenotype data for the 8 traits analyzed in the parents and in the MG5323 x Latino recombinant inbred line (RIL) mapping population.

Trait	Environment	Parents			RIL						
		Latino	MG5323	*p*-value	Min	Max	Range	Mean	H^2	SD	CV%
L	V13	7.47	8.64	***	6.693	9.430	2.737	7.908	0.90	0.449	5.7
	B14	7.66	8.61	****	6.782	9.422	2.640	8.167	0.97	0.486	6.0
	F15	7.75	9.12	**	6.863	9.939	3.076	8.275	0.97	0.502	6.1
	F20	7.60	9.19	***	6.892	9.524	2.632	8.061	0.99	0.459	5.7
	BLUP	7.64	8.90	-	6.950	9.385	2.434	8.117	0.96	0.490	6.1
W	V13	3.32	2.92	*	2.708	3.725	1.018	3.158	0.83	0.176	5.6
	B14	3.24	2.72	****	2.218	3.660	1.443	3.060	0.92	0.220	7.2
	F15	3.42	2.96	**	2.698	3.807	1.109	3.245	0.92	0.189	5.8
	F20	3.28	2.82	***	2.364	3.468	1.105	3.016	0.98	0.176	5.8
	BLUP	3.32	2.87	-	2.818	3.451	0.634	3.124	0.83	0.210	6.8
P	V13	18.19	20.10	**	16.465	21.773	5.308	18.792	0.87	0.919	4.9
	B14	18.28	19.69	****	15.651	22.057	6.406	19.059	0.96	1.005	5.3
	F15	18.75	21.07	**	16.576	23.253	6.677	19.569	0.96	1.020	5.2
	F20	18.34	20.91	***	16.503	21.994	5.491	18.941	0.98	0.932	4.9
	BLUP	18.43	20.48	-	16.748	21.584	4.836	19.127	0.95	1.020	5.3
A	V13	19.53	19.04	ns	14.876	24.336	9.460	19.139	0.83	1.443	7.6
	B14	18.73	17.36	****	11.581	22.778	11.196	18.973	0.92	1.791	9.5
	F15	20.34	20.35	ns	14.710	25.621	10.911	20.332	0.94	1.742	8.6
	F20	19.19	19.48	ns	14.023	22.692	8.669	18.564	0.97	1.440	7.8
	BLUP	19.37	19.35	-	15.321	22.132	6.811	19.305	0.87	1.760	9.1

Table 3. Summary of the phenotype data for the 8 traits analyzed in the parents and in the MG5323 x Latino recombinant inbred line (RIL) mapping population - Continued.

Trait	Environment	Parents			RIL						
		Latino	MG5323	*p*-value	Min	Max	Range	Mean	H^2	SD	CV%
WL	V13	0.45	0.34	***	0.322	0.518	0.196	0.407	0.94	0.030	7.4
	B14	0.42	0.32	****	0.290	0.527	0.237	0.376	0.96	0.034	9.0
	F15	0.44	0.33	***	0.321	0.519	0.198	0.394	0.95	0.032	8.1
	F20	0.43	0.31	****	0.304	0.501	0.197	0.376	0.99	0.033	8.7
	BLUP	0.44	0.32	-	0.329	0.505	0.176	0.388	0.94	0.030	8.9
FC	V13	0.73	0.59	***	0.566	0.792	0.225	0.671	0.94	0.036	5.4
	B14	0.70	0.56	****	0.514	0.803	0.289	0.656	0.94	0.039	6.0
	F15	0.73	0.58	***	0.573	0.804	0.231	0.668	0.95	0.037	5.6
	F20	0.72	0.56	***	0.543	0.785	0.242	0.651	0.99	0.040	6.1
	BLUP	0.72	0.57	-	0.583	0.781	0.198	0.661	0.93	0.040	5.9
TKW	B14	51.02	37.38	****	20.167	62.900	42.733	48.009	0.90	6.977	14.6
	F15	59.68	50.43	ns	29.367	68.667	39.300	53.516	0.90	6.640	12.4
	F20	53.53	44.80	**	24.300	56.000	31.700	44.310	0.96	5.435	12.3
	BLUP	54.99	45.22	-	35.520	58.524	23.005	48.794	0.80	7.480	15.3
HD	B14	20.00	40.00	*	13.000	44.000	31.000	30.491	0.94	6.391	21.0
	F15	34.00	44.33	****	19.000	51.000	32.000	38.306	0.87	3.777	9.9
	F20	38.67	52.67	**	34.000	61.000	27.000	45.503	0.95	4.829	10.6
	BLUP	33.52	47.69	-	28.974	48.213	19.239	39.051	0.87	7.640	19.6

Significance is denoted as **** p< 0.0001, *** p< 0.001, ** p<0.01, * p<0.05 between parental lines based on Student's t-test; ns: not significant; (-) : not available. RIL: recombinant inbred lines; BLUP: best linear unbiased prediction; H^2: the broad-sense heritability; SD: population standard deviation; CV: variation coefficient.

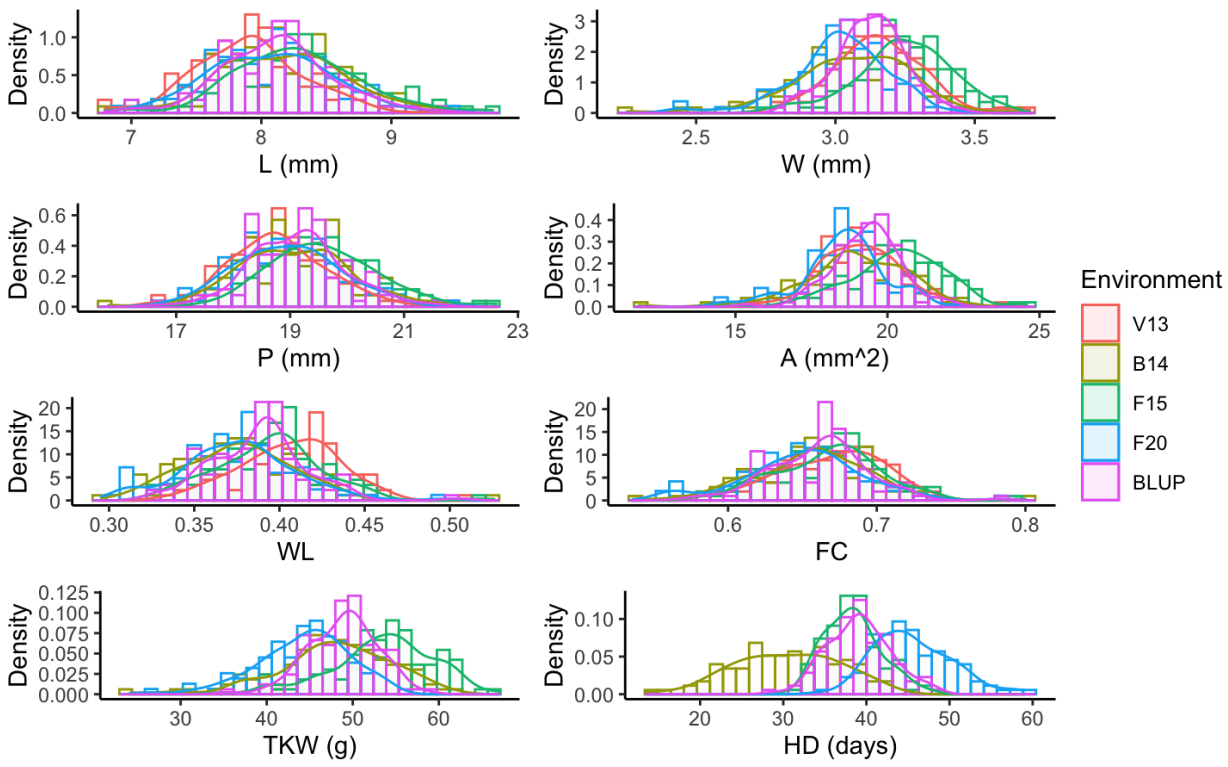


Figure 7. Frequency distribution for the 8 traits analyzed for each environment (V13, B14, F15 and F20) and across environments (BLUP) for this study.

The analysis of variance (One-Way ANOVA) detected highly significant differences among RILs for all traits in each environment ($p < 0.0001$, **Appendix A**), which indicates that genetic factors are contributing to the large phenotypic variability detected. However, for environment F15 the replication factor was also significant and higher than the genotype factor, which could imply experimental error as a non-homogeneous experimental field, for that matter the data on this environment was taken into account carefully.

Furthermore, the analysis of variance for the overall dataset across environments (Two-Way ANOVA) revealed significant effects of RILs, environments and genotype × environment interaction (GEI), as shown in **Tab. 4**. Being the GEI highly significant could imply that the phenotypic expression of one genotype might be superior to another genotype in one environment but inferior in a different environment, therefore the use of BLUP was performed to get adjusted means.

As observed, there is a large environmental effect (p < 0.001), ENV, accounting for most of the variability, which confirms that the environments considered were enough to differentiate the possible GEI effect on the target traits. Despite this, the genotype variability (GEN) is higher than the GEI component for all traits. Indeed, high values of broad sense heritability were obtained for all traits, ranging from 0.80 to 0.99, with the highest values obtained by the kernel shape traits (**Tab.3**). Additionally, from the Two-Way ANOVA it can be seen that the effect of the replications within each environment, REP(ENV), is significant for all traits (p < 0.001), however it is always lower than the genotype variability.

Table 4. Mean squares from the overall analysis of variance for the 8 traits analyzed in the MG5323 x Latino recombinant inbred line (RIL) mapping population.

Source of variation	*DF*	L	W	P	A	WL	FC	*DF*	TKW	HD
GEN	*109*	1.89***	0.21***	7.54***	16.73***	0.0083***	0.011***	*109*	199.14***	137.46***
ENV	*3*	6.23**	3.14**	32.29**	173.76**	0.0559***	0.025**	*2*	6505.31**	15020.57***
GEI	*320*	0.08***	0.04***	0.4***	2.16***	0.0005***	0.001***	*218*	41.39***	18.67***
REP(ENV)	*6*	0.31***	0.12***	1.79***	11.18***	0.0003**	0.001***	*5*	138.17***	31.14***
Residuals	*640*	0.02	0.01	0.10	0.39	0.0001	0.0001	*545*	7.44	4.00

Statistical significance is denoted as *** p < 0.001 and ** p < 0.01.

Correlation analysis was performed for the phenotypic data on each environment (**Appendix B**) and across environments (**Fig. 8**) among the eight evaluated traits. These kernel-related traits can be distinguished in main or primary (length and width), and secondary (perimeter, area, WL ratio and FC) being derived by combinations of the main traits. As expected, these biological and geometrical relationships between traits were inherently correlated in all environments.

Regarding the across environments analysis (**Fig. 8**), from the 28 possible correlation pairs, 18 were found highly significant (p < 0.001). In detail, for kernel size traits, kernel length is the main feature related to its secondary features area (r ≈ 0.7) and perimeter (r ≈ 1). Meanwhile, kernel width is the main trait for WL ratio and form coefficient attributes (r ≈ 0.7). Interestingly, TKW showed a highly statistical positive

correlation to area ($r \approx 0.9$) and kernel width ($r \approx 0.8$) in all environments and across; also a moderate significant and positive correlation to length ($r \approx 0.4$) and perimeter ($r \approx 0.5$). The correlations between TKW and kernel shape traits were significant ($p < 0.05$) with r values lower than the traits mentioned before. Heading date was correlated positively with length and perimeter ($r \approx 0.3$), negatively correlated with width ($r \approx -0.3$), WL ratio ($r \approx -0.4$) and form coefficient ($r \approx -0.5$); and non-correlated to TKW.

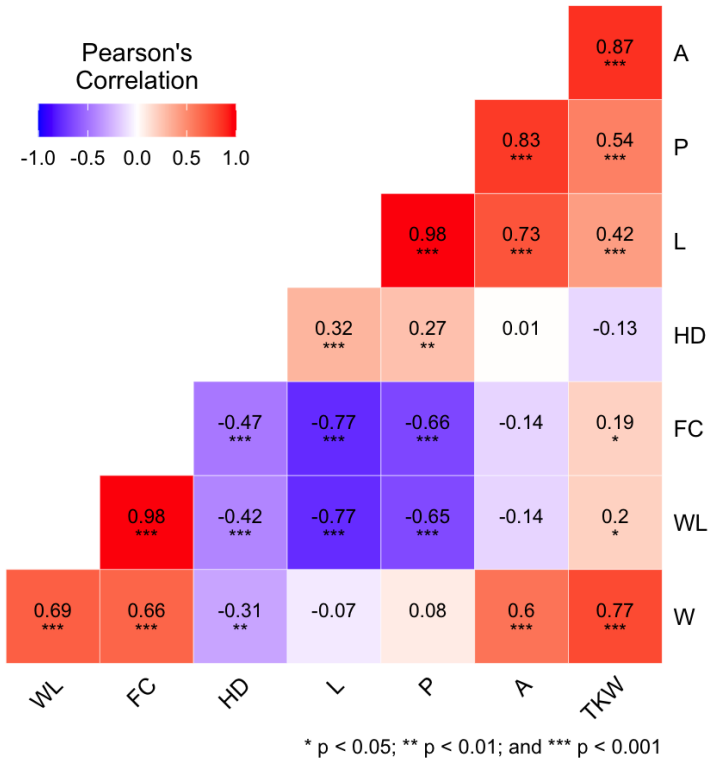


Figure 8. Pearson correlations coefficients (r) among the phenotypic traits analyzed using overall data across environments (from BLUP).

4.2. Quantitative Trait Loci (QTL) Detection

The analysis of QTL was performed for the 8 traits, including kernel size and kernel shape parameters, kernel weight and heading date using phenotypic data from 4 single environments (V13, B14, F15 and F20) and across environments (by BLUP), by using R/*qtl* software. For each trait, loci whose peaks were less than 10 cM faraway and/or have overlapping CIs were considered to correspond to the same

QTL, provided that the additive effect is conferred by the same parent, further these QTL were named as shown in **Tab.5**.

In this way, a total of 41 different QTL were found significant at a LOD ≥ 3.0, distributed on 11 of the 14 chromosomes of the MG5323 x Latino linkage map. The chromosomes reported with the highest number of associated regions were 2B and 2A with 7 and 6 loci, respectively. No QTL were located on chromosomes 1A, 1B and 6A.

Among the 41 QTL, seven were retrieved in the 4 environments and across them (BLUP dataset). Additionally, twenty-two loci out of the 41, were found with a positive additive effect, meaning the allele involved is conferred by MG5323 parent (*T. dicoccum*). In detail, 21 QTL were associated with kernel length, perimeter, area, width and heading date, one to TKW and none to kernel shape traits (WL ratio and form coefficient). Major/moderate QTL (with Percentage of Variance Explained - PVE or R^2 - above 15%) were found, including 3 for kernel length, 1 for width, 5 for perimeter, 2 for area, 3 for WL ratio, 1 for form coefficient, 1 for TKW and 3 for heading date. No significant epistatic interactions were identified in this study.

In the next sections, the associations from these 41 QTL were fully described.

Table 5. Quantitative Trait Loci (QTL) for the 8 traits analyzed detected in the MG5323 x Latino recombinant inbred line (RIL) mapping population per environment and across environments (BLUP).

				Environments												Across environments (BLUP)		
				V13			B14			F15			F20					
Trait	QTL name	Chr.	Peak Pos. Interval (cM)	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add
L	QL-2A	2A	78,2-83,1	3.77	12.75	0.16	*2.82*	*8.14*	*0.15*	-	-	-	-	-	-	-	-	-
	QL-2B	2B	121.1	-	-	-	-	-	-	5.54	15.37	0.20	-	-	-	-	-	-
	QL-4B	4B	78,8-79,5*	-	-	-	3.51	10.27	0.16	4.86	13.29	0.18	5.79	18.63	0.20	4.08	13.54	0.15
	QL-6B	6B	49.6	-	-	-	-	-	-	3.96	10.62	-0.16	-	-	-	-	-	-
	QL-7A	7A	77,4-90,6	4.17	14.21	0.16	3.69	10.84	0.16	3.64	9.69	0.15	6.68	21.95	0.21	3.69	12.15	0.15
		Model		7.02	25.46		10.34	**35.12**		12.63	**41.08**		10.30	**36.92**		7.64	**27.38**	
W	QW-2B	2B	31.2	-	-	-	-	-	-	-	-	-	3.07	10.81	-0.06	-	-	-
	QW-3A	3A	77.4	-	-	-	-	-	-	-	-	-	3.48	12.37	-0.06	-	-	-
	QW-4B	4B	21	4.02	13.06	0.06	-	-	-	-	-	-	-	-	-	-	-	-
	QW-5A	5A	164.8	-	-	-	-	-	-	3.69	11.90	0.06	-	-	-	-	-	-
	QW-6B	6B	66,5-67,1	-	-	-	3.77	14.59	-0.08	*2.96*	*9.42*	*-0.05*	-	-	-	-	-	-
	QW-7A	7A	77.4	4.97	16.50	-0.07	-	-	-	3.81	12.32	-0.06	-	-	-	3.75	14.53	-0.05
		Model		8.09	**28.73**		3.77	14.59		8.07	**28.68**		6.86	**26.42**		3.75	14.53	
P	QP-2A	2A	111.2	3.08	9.72	0.27	-	-	-	-	-	-	-	-	-	-	-	-
	QP-2B	2B	43.2	3.07	9.7	0.27	-	-	-	-	-	-	-	-	-	-	-	-
	QP-4A	4A	105.2	-	-	-	3.90	12.48	0.35	-	-	-	-	-	-	-	-	-
	QP-4B	4B	79,5*	-	-	-	5.04	16.56	0.40	4.02	15.50	0.38	6.61	22.06	0.44	4.59	15.56	0.32
	QP-7A.1	7A	78,1-85,6	-	-	-	-	-	-	-	-	-	5.71	18.68	0.40	*2.89*	*9.45*	*0.25*
	QP-7A.2	7A	103,4	3.71	11.90	0.30	-	-	-	-	-	-	-	-	-	-	-	-
		Model		8.27	**29.28**		8.36	**29.52**		4.02	15.5		9.91	**35.81**		7.39	**26.61**	

Table 5. QTL for the 8 traits analyzed detected in the RIL population - Continued.

Trait	QTL name	Chr.	Peak Pos. Interval (cM)	Environments												Across environments (BLUP)		
				V13			B14			F15			F20					
				LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add
A	QA-2A	2A	111.2	*2.38*	*9.49*	*0.41*	-	-	-	-	-	-	-	-	-	-	-	-
	QA-3B	3B	114.3	-	-	-	-	-	-	3.46	10.26	-0.52	-	-	-	3.35	9.86	-0.36
	QA-4A	4A	22.7	-	-	-	3.37	9.89	0.63	-	-	-	-	-	-	-	-	-
	QA-4B	4B	79,5-81,5*	-	-	-	3.14	9.17	0.53	3.90	11.60	0.54	4.19	17.10	0.58	3.94	11.77	0.39
	QA-6B	6B	66.5	-	-	-	6.57	20.68	-0.83	4.11	12.28	-0.57	-	-	-	4.14	12.42	-0.40
		Model		*2.38*	*9.49*		10.16	**34.64**		10.11	**34.50**		4.19	17.1		10.09	**34.44**	
WL	QWL-2A	2A	43.4	-	-	-	-	-	-	-	-	-	3.51	8.52	-0.01	-	-	-
	QWL-2B	2B	114,9-122,4	3.80	9.70	-0.01	3.39	10.98	-0.01	3.60	11.38	-0.01	3.13	7.55	-0.01	4.25	10.81	-0.01
	QWL-3A	3A	77.4	3.59	9.14	-0.01	-	-	-	-	-	-	3.59	8.74	-0.01	4.31	10.99	-0.01
	QWL-7A	7A	77,4-86,9	7.36	20.32	-0.01	4.25	13.99	-0.01	4.85	15.78	-0.01	4.87	12.22	-0.01	5.99	15.85	-0.01
		Model		13.75	**43.75**		7.88	**28.1**		8.51	**29.97**		15.4	**49.78**		14.04	**44.44**	
FC	QFC-2A	2A	43.4	-	-	-	-	-	-	-	-	-	4.04	10.05	-0.01	-	-	-
	QFC-2B	2B	43.4	3.53	7.64	-0.01	-	-	-	-	-	-	-	-	-	3.49	7.60	-0.01
	QFC-2B	2B	114,9-122,4	3.22	6.92	-0.01	3.84	12.24	-0.01	3.30	10.27	-0.01	*2.88*	*6.97*	*-0.01*	4.49	9.98	-0.01
	QFC-3A	3A	77.4	5.84	13.29	-0.01	-	-	-	-	-	-	3.42	8.38	-0.01	5.39	12.19	-0.01
	QFC-7A	7A	77,4-86,2	5.67	12.85	-0.01	4.41	14.22	-0.01	5.40	17.58	-0.01	4.42	11.08	-0.01	4.45	9.88	-0.01
		Model		17.57	**52.07**		8.45	**29.80**		8.75	**30.68**		15.18	**49.27**		17.42	**51.78**	

Table 5. QTL for the 8 traits analyzed detected in the RIL population - Continued.

				Environments												Across environments (BLUP)		
				V13			B14			F15			F20					
Trait	QTL name	Chr.	Peak Pos. Interval (cM)	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add	LOD	R^2 (%)	Add
	QTKW-3B.1	3B	114,3				-	-	-	4.00	12.76	-2.18	-	-	-	-	-	-
	QTKW-3B.2	3B	154,9				-	-	-	-	-	-	-	-	-	4.19	13.34	-1.51
TKW	QTKW-4B	4B	82,1*				-	-	-	3.29	10.36	1.96	3.02	12.65	1.87	3.08	9.56	1.27
	QTKW-6B	6B	66,5-67,1				3.99	15.39	-2.64	3.49	11.02	-2.02	-	-	-	4.26	13.60	-1.48
		Model					3.99	15.39		8.50	**29.94**		3.02	12.65		8.67	**30.44**	
	QHD-2A	2A	34,9-35,6				17.30	42.76	4.05	5.20	9.58	1.10	-	-	-	13.16	26.12	1.88
	QHD-2B	2B	46.4				*2.79*	*4.99*	*1.37*	5.70	10.61	1.16	9.00	27.79	2.51	6.58	11.27	1.29
	QHD-3A	3A	5.6				-	-	-	3.88	6.95	0.95	-	-	-	-	-	-
HD	QHD-5A	5A	124.4				4.97	9.31	1.76	-	-	-	-	-	-	4.88	8.06	1.06
	QHD-5B	5B	85.2				-	-	-	5.21	9.61	1.17	-	-	-	4.17	6.79	1.11
	QHD-7B	7B	39,9-48,1				3.85	7.03	1.53	6.90	13.20	1.24	4.10	11.37	1.57	6.98	12.06	1.29
		Model					21.76	**59.78**		22.25	**60.61**		11.90	**39.23**		24.70	**64.45**	

Not available

(-) : Not significant; LOD: Logarithm of Odds; R^2: Percentage of the phenotypic variance explained (= PVE); Add: additive effect of a QTL, where the absence of sign indicates alleles from parent MG5323 which are increasing the trait scores, while the negative sign (-) indicates alleles from parent Latino. Suggestive/Putative QTL, below the threshold (LOD < 3.0), are reported in italics. The best models with phenotypic variation explained over 25% are reported in bold. Note: (*) For the QTL on 4B, the peak position in the F20 environment was located at 63 cM, while in the rest of environments and across data was located at about 79-82 cM.

4.2.1. QTL for Kernel Size

In this study, for the kernel size traits (length, width, perimeter and area) a total of 22 QTL were found and mapped on chromosomes 2A, 2B, 3A, 3B, 4A, 4B, 5A, 6B and 7A. Specifically, 5 loci were detected for kernel length, from which only 2 were environmentally stable (identified by more than one environment and across environments on BLUP dataset) at chromosomes 4B and 7A, being “QL-7A” the major QTL with the highest percentage of variance explained (R^2 ranging from 9.69 to 21.95%) and a LOD ranging from 3.6 to 6.8. For kernel width, 6 regions were associated. Among them, only “QW-7A" was found stable, with a R^2 ranging from 12.32 to 16.50% and a LOD from 3.8 to 5.0, and which was also coinciding with the above mentioned “QL-7A”. Related to the perimeter, 6 QTL were found, only 2 were environmentally stable in chromosomes 4B and 7A, coinciding with QTL of length. ”QP-4B” was the major QTL, explaining 15.50 to 22.06% of the phenotypic variance and with a LOD of 4.0 to 6.6. In the case of kernel area, 3 QTL, from a total of 5, were environmentally stable and mapped in chromosomes 3B, 4B and 6B, this last one was the major QTL for this trait (“QA-6B”, R^2 = 12.28 to 20.68%, LOD = 4.1 to 6.5). As seen, the stable QTL related to kernel size traits were mainly mapped on chromosome 4B at the same position (about 79-82 cM). Additionally, for all these traits, a total of 15 QTL showed a positive additive effect, meaning the alleles for increasing the target trait were contributed by the parental line MG5323 (*T. dicoccum).*

4.2.2. QTL for Kernel Shape

Nine loci were detected in total for kernel shape traits (WL ratio and form coefficient), located in chromosomes 2A, 2B, 3A and 7A. For WL ratio, 4 QTL were detected, among them, 3 were environmentally stable and located on chromosomes 2B, 3A and 7A. The major QTL was “QWL-7A”, with the highest R^2 (12.22 to 20.32%) and LOD value (4.2 to 7.3). From the total of 5 loci found for form coefficient, 3 were environmentally stable and distributed in chromosomes 2B, 3A and 7A. The highest R^2 was also shown by the coincident QTL on 7A, “QFC-7A”, ranging from 11.08 to 17. 58%, and with a LOD value of 4.4 to 5.6. These results

show the inherent correlation between the traits, as expected. The parental Latino carried all the alleles for increasing kernel shape traits and thus for conferring more roundness to the kernels.

4.2.3. QTL for Grain Weight (TKW) and Heading Date

Ten chromosome positions were found for TKW and heading date in this study. For TKW, 4 QTL were detected, 2 located on chromosome 3B ("QTKW-3B.1" and "QTKW-3B.2"), and 2 were environmentally stable in chromosomes 4B and 6B. The major QTL was "QTKW-6B" with the highest phenotypic variance explained (R^2 = 11.02 - 15.39%) and LOD values (3.5 - 4.3). Noteworthy, three of these loci identified ("QTKW-3B.1", "QTKW-4B" and "QTKW-6B") were co-located with QTL for area, confirming the highly significant correlation between these traits. Important to mention that the locus in 4B was stably detected in 2 environments (F15 and F20) and across them, however the peak position in F20 was slightly moved (from 63 cM to 81 cM). Interestingly, the parent line Latino contributed the positive alleles at most of the loci, except for the one in 4B ("QTKW-4B"), contributed by parental MG5323.

From the 6 QTL found for heading date, 5 loci were environmentally stable on chromosomes 2A, 2B, 5A, 5B and 7B, being the major QTL "QHD-2A", with a R^2 varying from 9.6 to 42.8% and a LOD value from 5.2 to 17.3. The positive alleles from these QTL were all derived from parent MG5323, which is indeed the late parent.

4.3. Cluster of QTL

Due to the geometrical or biological nature of the relationships between the traits under examination, it was expected to find coincident loci between different traits, as also suggested by the correlation analysis. This implies the pleiotropic effect of a single gene or a set of linked genes for the different traits. Therefore, six QTL clusters were defined as regions with two or more overlapping QTL for different traits (on closer/same positions of the Latino x MG5323 map). Further, the genetic position of each cluster on the tetraploid wheat consensus map (Maccaferri et al., 2015) was obtained by projecting the flanking markers of each locus (left and rigth

markers defined by the previous calculation of their CIs), corroborating the co-location of these loci, besides following analysis described below. The six clusters identified were associated with the 7 kernel morphological traits and TKW, located on chromosomes 2B, 3A, 3B, 4B, 6B and 7A (**Tab. 6** and **Fig. 9**). There were no co-locating QTL found with the heading date.

The clusters 1, 2 and 6 (on chromosomes 2B, 3A and 7A, respectively) highlighted the expected relationship between the main traits (length and width) and their derivative ones (perimeter, area, WL ratio, form coefficient). Meanwhile, clusters 3, 4 and 5 (on chromosomes 3B, 4B and 6B, respectively) are associated with kernel size/shape and TKW, which can confirm the causal relationship between these traits. Worth mentioning that cluster 3 is considered as a putative one, due to the slight change of positions for the loci related to TKW in the across environments model.

Notably, an interesting relationship was found on chromosome 4B (cluster 4) between kernel size traits (length, perimeter and area) and TKW, where the positive alleles of the QTL were donated by the emmer parent (MG5323).

Table 6. Consensus positions, physical positions and candidate genes of QTL clusters detected in the MG5323 x Latino recombinant inbred line (RIL) mapping population for kernel-related traits.

QTL Cluster	Traits involved	Chr.	Donor	Position on Latino x MG5323 map (cM)			Position on Consensus map (cM)		Physical Position on Reference genome (Mbp)		Candidate Genes		
				Peak	Left CI	Right CI	Left CI	Right CI	Start	End	Quantity or Gene ID	Start position (Mbp)	Function (Known genes)
1	L, WL, **FC**	2B	MG Latino	114.9	107.5	122.3	112.3	128.0	537.6	629.3	564	-	-
2	W, WL, **FC**	3A	Latino	77.4	71.3	83.5	65.9	77.8	439.3	534.9	576	462.0	Regulation of Cell Division and Elongation (*D61*)
3*	A, **TKW**	3B	Latino	114.3	108.5	120.1	129.6	138.9	691.3	741.6	TRITD3Bv1G229090,TRITD3Bv1G229910, TRITD3Bv1G235190	695.9 698.3 717.9	Response to Auxin
											TRITD3Bv1G231370	702.8	Regulation of Cell Division
											TRITD3Bv1G239650	729.7	Ubiquitination and Auxin Regulation
4	L, **P**, A, TKW	4B	MG	79.5	75.0	83.9	77	84.9	594.7	619.2	TRITD4Bv1G175480, TRITD4Bv1G179270	595.1 605.8	Auxin Regulation
											TRITD4Bv1G177190	600.5	Regulation of Cell Proliferation
											TRITD4Bv1G171270	582.0	Auxin Transport and Seed Growth Regulation (*BIG GRAIN PROTEIN 1*)
5	W, **A**, TKW	6B	Latino	66.5	62.9	70.1	73.4	80.9	263.2	467.3	582	300.8 373.4	Regulation of Cell Growth (*GW2)* Regulation of Cell Division (*FUWA*)
6	L, P, W, **WL**, FC	7A	MG Latino	79.3	75.7	82.9	92.9	107.9	113.9	167	TRITD7Av1G052720, TRITD7Av1G055870	117.2 125.6	Regulation of Cell Growth and Differentiation
											TRITD7Av1G050690	111.5	Sucrose Metabolism (*TaSus1*)
											TRITD7Av1G071860	168.5	Heat Acclimatization (*TaGASR7-A1*)

Best QTL selected by the highest LOD and PVE within the cluster are shown in bold, for these QTL the positions were retrieved. (-) : Not available. For physical intervals higher than 60 Mbp, only the quantity of genes within and known genes are shown. (*) : Putative cluster as the position of the QTL for TKW changes position on the BLUP model. 1 Mbp (Megabase pair) = 1 000 000 bp (base pairs). The start position of candidate genes refers to its position on the reference genome.

4.4. Comparative analysis of QTL clusters with previously published QTL

The availability of the durum wheat reference genome (*cv.* Svevo) allowed to define the physical interval of the clusters identified. To this aim, for the most consistent QTL (best QTL selected by the highest LOD and PVE) within each cluster, the molecular markers (retrieved from the consensus map) closest to the extremes of the QTL confidence interval were projected on the genome, by identifying their best BLAST hit. In this way, the largest clusters were detected on chromosome 2B (cluster 1), 3A (cluster 2) and 6B (cluster 5), which spanned for more than 90 Mbp (119 Mbp, 95 Mbp and 216 Mbp, respectively). Cluster 3 and 6 spanned for approximately 53 Mbp. Regarding cluster 4 on chromosome 4B, it spanned for approximately 25 Mbp.

Additionally, the comparison of physical positions of the clusters detected in this study with QTL from previous studies (cited on section 6, Materials and Methods) was performed to assess the novelty of our results (**Tab. 7**). The regions found in this study were already described for most of the traits considered, as described below.

In detail, the physical interval of cluster 1 overlapped with a QTL from Desiderio et al., 2019, related to the same traits (length and WL ratio), confirming the association between these traits and the loci position 537 - 629 Mbp on chromosome 2B.

Regarding cluster 2, on chromosome 3A, its physical interval (439 - 534 Mbp) coincided with 2 loci for TKW detected by Avni et al., 2018 and Sun et al., 2020, meanwhile, the association with this trait was missing in this study. However, 2 loci described in Wang et al., 2019 in association with kernel width coincided with our result.

In the cluster 3, detected on chromosome 3B, there were moderate coincidences (with a light shift) at its physical interval (691 - 742 Mbp) with the loci detected by Faris et al., 2014 and Mangini et al., 2018 both for TKW, while no loci

related to area were previously found in this chromosome region. However, it is important to mention that the position of the QTL in this study requires further refinement, as extended phenotyping or more experimental locations, due to the fact that the position of the locus for TKW was slightly moved for the across-environments dataset.

The cluster 4, spanning from 595 to 619 Mbp, overlapped with 3 QTL from Blanco et al., 2012 (related to TKW), Elouafi et al., 2004 (TKW) and Mangini et al., 2021 (related to area and width), which is consistent with the traits associated with this cluster in this study (area and TKW). This result confirms the relationship between these traits and the considered loci. However, in the specific case of length and perimeter, there were no coincidences with previously described QTL.

The physical interval of cluster 5 (263 - 467 Mbp) overlapped with a total of three known regions, one from Desiderio et al., 2019 related to area, one from Tzarfati et al., 2014 related to TKW, and one from Sun et al., 2020 associated with WL ratio (derivative trait from width), therefore, they coincided with our related traits for this region (area, width and TKW).

Lastly, only two QTL were found coincident with the physical interval of cluster 6 (114 - 167 Mbp), being one related to TKW (Patil et al., 2013) and one found in association with kernel width (Sun et al., 2020). This last is consistent with some of the traits related for this cluster in our work (kernel width, WL ratio, and form coefficient), while no coincidences were found for length and perimeter (**Tab. 7**).

Table 7. Comparison of positions from previously reported quantitative trait loci (QTL) for kernel-related traits. Overlapping or closer literature QTL are shown.

Reference	Trait	Chr.	Left marker on reference genome		Right marker on reference genome	
			Marker ID	Physical position (Mbp)	Marker ID	Physical position (Mbp)
This study : Cluster 1	**L, WL, FC**	2B	**IWB29112**	**537.6**	**IWA2130**	**629.3**
Desiderio et al., 2019	L, WL		IWB39200	448.4	IWB69139	546.4
This study : Cluster 2	**W, WL, FC**		**IWA2095**	**439.3**	**IWA5316**	**534.9**
Avni, et al., 2018	TKW		IWB16112	487.2	IWB20961	521.7
Blanco et al., 2012	*TKW*	3A	*IWB66938*	*543.5*	*IWB44737*	*568.7*
Sun et al., 2020	TKW		N/A[a]	419.1	N/A	521.1
Wang et al., 2019	W		N/A[b]	447.5	N/A	466.7
Wang et al., 2019	W		N/A[c]	448.0	N/A	467.2
This study : Cluster 3	**A, TKW**		**IWB11298**	**691.3**	**IWB24723**	**741.6**
Desiderio et al., 2019	*TKW*	3B	*IWB9399*	*781.1*	*IWB71782*	*817.5*
Faris et al., 2014	TKW		IWA5510	741.5	IWA1094	778.4
Mangini et al., 2018	TKW		wPt-7145	742.6	IWA1745	774.1
This study : Cluster 4	**L, P, A, TKW**		**IWA1382**	**594.7**	**IWA8591**	**619.2**
Blanco et al., 2012	TKW		IWA2398	555.1	IWB59718	582.3
Elouafi et al., 2004	TKW	4B	IWB34975	501.2	IWB8082	599.3
Graziani et al., 2014	*TKW*		*IWB71667*	*629.0*	*IWB32544*	*654.0*
Mangini et al., 2021	A, W		IWB38381	567.5	IWB17082	598.5

Table 7. Comparison of positions from previously reported quantitative trait loci (QTL) for kernel-related traits. Overlapping or closer literature QTL are shown - Continued.

Reference	Trait	Chr.	Left marker on reference genome		Right marker on reference genome	
			Marker ID	Physical position (Mbp)	Marker ID	Physical position (Mbp)
This study : Cluster 5	**A, W, TKW**	6B	**IWA3632**	**263.2**	**IWB28348**	**467.3**
Desiderio et al., 2019	A		IWB5586	150.1	IWB73148	449.7
Tzarfati et al., 2014	TKW		IWB58306	443.0	IWB73374	562.8
Sun et al., 2020	WL		N/A[d]	301.4	N/A	403.4
This study : Cluster 6	**L, W, WL, P, FC**	7A	**IWB65337**	**113.9**	**IWB46718**	**167.0**
Desiderio et al., 2019	*FC*		*IWB53096*	*673.0*	*IWB39743*	*673.0*
Patil et al., 2013	TKW		IWB14901	106.1	IWB47160	123.3
Sun et al., 2020	W		N/A[e]	106.6	N/A	208.6

Clusters found in this study are shown in bold. Near QTL but not overlapped to the cluster's positions are shown in italics. Trait's acronyms are explained in Table 2. Chr. refers to Chromosome. 1 Mbp (Megabase pair) = 1 000 000 bp (base pairs). Note : For the studies Sun et al., 2020 and Wang et al., 2019, extension and position of the confidence interval around each associated marker was calculated based on LD extension, that was 51 Mbp and 9.6 Mbp, respectively. The physical position of each associated marker is as follows :

a) BE425919_3_A_592 at 470.072 Mbp; b) IWA2069 at 457.07 Mbp; c) IWA5616 at 457.64 Mbp; d) BE404912_6_B_Y_488 at 352.3 Mbp; e) BE499652_7_A_Y_391 at 157.5 Mbp.

4.5. Analysis of QTL physical intervals to hypothesize candidate genes

To obtain gene content and hypothesize candidate genes, we inspected the functional annotations of High Confidence Svevo genes (based on Gene Ontology -GO- terms) within the physical intervals of the most consistent QTL (best QTL) for each cluster. More attention was addressed to those genes whose GO terms could be likely associated with the kernel development and grain yield based on previous knowledge. Therefore, functional categories considered were related to hormone pathways and sugar metabolism, since these protein classes have been already associated with grain size determination, grain weight and seed morphology.

Additionally, a previous list of durum wheat genes orthologous to genes cloned in rice and/or wheat for phenotypic effects on kernel size/shape and weight (Desiderio et al., 2019) was updated (**Appendix C**). Then, the position of these orthologous genes on the durum wheat reference genome was also used to retrieve candidate genes on each cluster.

As positive control, the physical position of the environmentally stable QTL localized in 2A and 2B for heading date was compared to the known positions of phenology major genes *Ppd*-A1 and *Ppd*-B1, which are key components in the photoperiod/flowering regulatory pathway, and are located at 36.6 Mbp and 56.3 Mbp on the Svevo genome, respectively (Willhelm et al., 2009; Takenaka et al., 2012). The physical interval detected in this study for heading date QTL, “QHD-2A”, was from 23.4 to 34.1 Mbp on 2A, and for “QHD-2B”, 41.0 to 52.6 Mbp on 2B. As seen, a light position shift (+2/3 Mbp approximately) was detected, which can be a consequence of the process of anchoring of QTL on the consensus map/reference genome. Although markers could look to be co-segregant in a genetic map, their physical position on the genome will be slightly different, also based on the recombination rate of the target region.

On clusters 3, 4 and 6, the most important candidate genes retrieved and known genes are listed (**Tab. 6** and **Fig. 9**). For clusters 1, 2 and 5, a high number of annotated genes (around 500) were present due to the big physical interval

detected. The gene content of these regions was not inspected, only the comparison with known genes from **Appendix C**. were reported.

More in detail, cluster 1 did not correspond to any orthologous gene (**Fig. 9**). For the physical interval of cluster 2, a co-location was found with *D61* rice gene (Os01g0718300, orthologous located at 462 Mbp on Svevo genome), which is known to be involved in the regulation of cell division (Nakamura et al., 2006).

Among the several candidate genes identified in the interval of cluster 3 on chromosome 3B, four genes were found to be related to auxin metabolism (TRITD3Bv1G229090, TRITD3Bv1G229910, TRITD3Bv1G235190 and TRITD3Bv1G239650) and one related to cell division (TRITD3Bv1G231370), while no coincidences were found with orthologous of rice/wheat cloned genes.

In cluster 4, two genes related to auxin regulation (TRITD4Bv1G175480 and TRITD4Bv1G179270) and one for cell proliferation (TRITD4Bv1G177190) overlapped with the target interval. Additionally, with a light position shift, we found TRITD4Bv1G171270 close to this region on chromosome 4B. This position refers to the known rice gene *BIG GRAIN PROTEIN 1* (*BG1*). The orthologous gene has been identified in common wheat, where it was involved with the production of larger seed size (Milner et al., 2021). This gene overlapped with the position of the loci related to TKW in this cluster (**Fig. 9**), therefore it was considered within the cluster's interval.

Cluster 5 was found overlapping within the paralogue genes *GW2* (located at 300 Mbp) and *FUWA* (located at 373 Mbp) coincided within its physical interval. *GW2* is known to be a negative regulator of grain weight and its mutation increments TKW (Zhai et al., 2018). *FUWA* is related to controlling grain size, and its mutations increase grain width and thickness, but decrease grain length (Chen et al., 2015). Lastly, the region of cluster 6 encompasses two genes related to regulation of cell growth (TRITD7Av1G052720 and TRITD7Av1G055870) and two known genes were detected: *TaSus*1 (TRITD7Av1G050690) and *TaGASR7* (TRITD7Av1G071860). The first one encodes a sucrose synthase associated with TKW (Mohler et al., 2016), while *TaGASR7* is considered as a genetic determinant of grain length in wheat (Dong et al., 2014, **Fig. 9**).

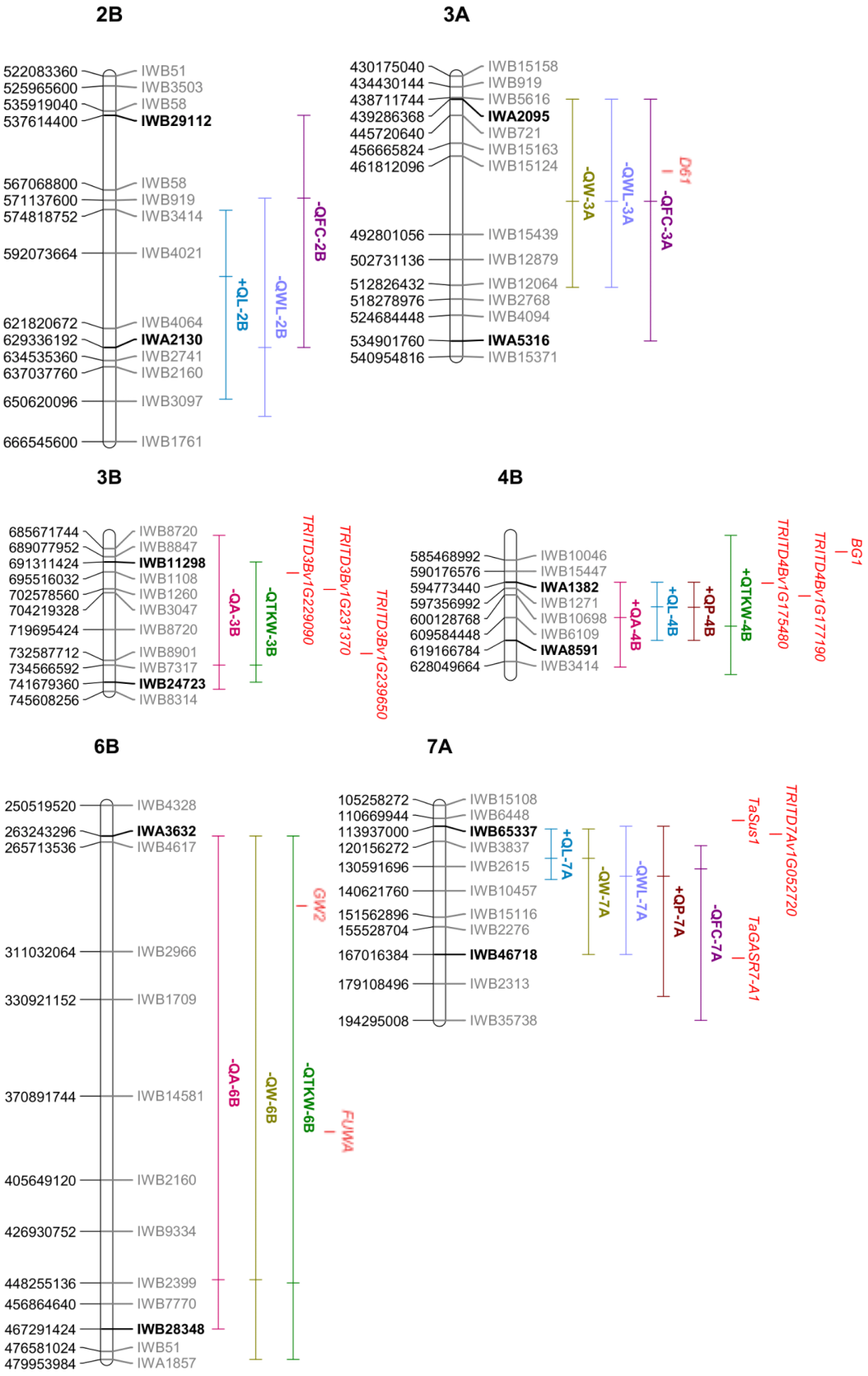


Figure 9. Schematic representation of Clusters of QTL on Durum Wheat reference genome. Part of the chromosomes are represented by including some markers surrounding the QTL clusters; marker IDs are on the right and in bold are the flanking markers of the cluster's interval. The + or − signs preceding the QTL name indicate the positive or negative additive effect of the parental line MG5323. The known genes and candidate genes hypothesized are shown in red. Complete information is on Tab. 6. (Figure performed using MapChart version 2.3).

5. Discussion

The improvement of wheat yield has stagnated since the mid-1990s, after the great impact caused by the integration of dwarfing genes, which led to the so-called Green Revolution (Milner et al., 2021). Nowadays, minor advances have been made and upcoming challenges, such as climate change, are demanding new strategies in breeding programs. In order to drive grain yield improvement, unraveling the genetic basis determining yield components, such as thousand-kernel weight (TKW) or spike fertility, is an ongoing and essential task for researchers.

In this way, attention should be put on kernel size and shape factors, which are important parameters for grain weight and have been manipulated as a consequence of domestication and selection for grain yield. The molecular mechanisms behind these traits have been mainly studied in bread wheat, while in durum wheat there is still a huge terrain to cover (Gegas et al., 2010; Cui et al., 2011; Patil et al., 2013; Desiderio et al., 2019; Mangini et al., 2021). Moreover, wheat ancestors, as cultivated emmer (*T. dicoccum*), should be considered as promising genetic resources to be employed for restoring durum wheat diversity (Mohammadi et al., 2021; Rahman et al., 2020).

Under this context, the present study was conceived to dissect the genetic network behind kernel size and shape traits, kernel weight (TKW) and heading date (HD), by performing QTL mapping on a RIL population derived from a *T. dicoccum* accession.

5.1. Detection of environmentally stable QTL and trait-relationships

In order to assess the environmental stability of the QTL, data from 4 different environments (location-year) across Italy was considered. In this study, the analysis of variance across the environments showed that all effects (GEN, ENV and GEI) were statistically significant, however the genotypic effect was higher than the genotype x environment interaction effect for all traits, coinciding with the heritability values obtained. Thus, we were able to detect environmentally stable QTL. The more stability of these loci across the environments implies the more importance of themselves in their determination of the considered traits.

Overall, twelve loci were detected in more than two/three environments and across them, in association for most of the kernel morphological factors and located on chromosomes 2B, 3A, 4B, 6B and 7A. The trait with the poorest stability was kernel width, where from 6 QTL found, 5 had poor stability (only detected in one or two environments). This could imply that kernel width might be controlled by minor effect genes under a relatively higher environmental effect, as exposed in two previous studies in durum wheat, where low heritability was also detected for this trait (Sun et al., 2020; Desiderio et al., 2019). Regarding TKW and HD, most of the loci found were detected in more than two environments and across them, indicating their stability, and were located on chromosomes 4B and 6B for TKW, and on 2A, 2B and 7B for HD.

Further, the objective of this study was to understand the genetic interdependence of the kernel morphological traits between them and with kernel weight and heading date. Therefore, 20 of the 41 detected loci (49%) were found relevant to more than one trait, meaning they were co-locating loci. Hence, these QTL were grouped in 6 clusters, located on chromosomes 2B at 112.3 cM, 3A at 65.9 cM, 3B at 129.6 cM, 4B at 77 cM, 6B at 73.4 cM and 7A at 92.9 cM (referring to consensus map positions).

As expected, the clusters 1, 2 and 6 show the inherent relationships between main kernel traits (length and width) and their mathematically derivative ones (area, perimeter, form coefficient and WL ratio), also detected from Pearson's correlation coefficients. Further, the relationship between kernel length and width is more intriguing. Indeed, the identification of loci that independently control these two kernel traits might allow the use of this genome-based information to obtain the kernel ideotype: longer and rounded. On the other hand, one locus determining both traits may allow one to focus on only one genomic region to efficiently increase kernel area. In this study, as previously shown, width and length were independent for most of the clusters, and Pearson's correlation analysis showed no significant correlation between these two traits, so the independence of both characters could be implied as in previous studies (Desiderio et al., 2019, Mangini et al., 2021). The

only exception is the cluster 6 on chromosome 7A, which suggests pleiotropic effects of the responsible gene(s).

The most interesting clusters were 3, 4 and 5, which were correlated to both kernel size/shape traits and TKW and located on chromosomes 3B, 4B and 6B, respectively. The highest significant positive relationship between a size trait and TKW, was detected from Pearson's correlation analysis for kernel area ($r \approx 0.9$), and further confirmed by the coincident loci detected for both traits in the above mentioned chromosomes. There is compelling evidence for this relationship, suggesting that TKW improvement could be due to the kernel area increase (Russo et al., 2014; Desiderio et al., 2019; Mangini et al., 2021; Sun et al., 2021). The other significant positive relationship found by Pearson's correlation value was between kernel width and TKW ($r \approx 0.8$), which was confirmed only by one cluster on chromosome 6B in this work. This result can be due to the above explained low stability of kernel width. Noteworthy, from the two-location study elaborated by Russo et al., 2014, with a population derived from a *T.dicoccum* line (named Molise Colli), the same correlations were found, being TKW highly correlated to kernel surface area and width ($r = 0.63$ and 0.49, respectively), validating our results. Notwithstanding, cluster on chromosome 4B may also suggest an effect of kernel length on TKW, although a low correlation ($r \approx 0.4$) was found, likely through an effect of increasing kernel area.

Further, these findings suggest that genes responsible for variation of kernel size/shape and for kernel weight express pleiotropy and/or are closely linked. Some recent examples about this assumption have been documented in both bread and durum wheat (Avni et al., 2018; Wang et al., 2019; Xin et al., 2020; Li et al., 2021; Desiderio et al., 2019, Mangini et al., 2021; Ma et al., 2021) and in some cases it has been also confirmed by the cloning of the candidate genes in rice and wheat (Yamamuro et al., 2000; Chen et al., 2015; Zhang et al., 2017). Therefore, the presented clusters (3, 4 and 5) could be of great value for marker-assisted breeding, allowing the selection of varieties carrying the pleiotropic/closed linked gene(s), which could impact on their grain weight potential (Ma et al., 2021).

Intriguingly, no cluster was detected combining QTL for the mentioned traits and heading date, in other words, the QTL for kernel-related traits and heading date were located in different marker intervals, which could indicate that they are genetically controlled independently by each other. However, in the study by Mangini et al., 2021, using a RIL population derived from a cross between 2 durum wheat lines, a cluster on chromosome 2A was associated with heading date, kernel area and kernel length, showing a relationship not found in this work, which might be due to the difference of genetic backgrounds. Considering that *T. dicoccum* is the most late parent, the independence of the control behind the kernel-related traits and heading date is a good sign to exploit this genetic resource to select for grain size and shape factors, without any linkage drag for late maturity.

5.2. Favorable alleles from *T. dicoccum*

The phenotypic description of the different kernel morphological factors for the parent lines considered was consistent with previous studies, being *T. dicoccum* (accession MG5323) kernel larger but narrower (higher kernel length, smaller width) compared to the durum cultivar (Latino), and with a significantly lower TKW and late in maturation (Gegas et al., 2010; Russo et al., 2014). Consistently, the contribution of superior alleles for kernel size traits (all except width) and heading date was from MG5323, while the superior alleles for kernel shape and most of the alleles for kernel width and TKW were donated by the parent Latino.

Here, despite the lower TKW value of MG5323, a high transgressive segregation was observed in the RIL population. This was confirmed by identifying a major and consistent QTL on chromosome 4B, explaining from 9.6 to 12.7% of the kernel weight variance, with the favorable allele derived from MG5323. As mentioned, this loci also explained a significant phenotypic variation (up to 22%) of the kernel size traits (length, perimeter and area), being the favorable alleles involved also deployed by cultivated emmer. This result is consistent with several studies supporting the involvement of *T. dicoccum* as donor of valuable alleles, highlighting its potential to increase seed size and weight for modern durum wheat breeding programs (Faris et al., 2014; Russo et al., 2014; Thanh et al., 2013; Wang et al., 2019). In summary, this result implies that the gene(s) underlying the loci on

chromosome 4B has pleiotropic effects on kernel length, perimeter and area and on the weight of the seed. Moreover, it suggests that the increase of kernel size from a *T. dicoccum* line might have a positive impact on the kernel weight of durum wheat.

5.3. Co-location of already published QTL

A comparison between previously known regions (from both linkage and association mapping) and the QTL clusters found in this work was performed by a projection on the durum wheat reference genome, through their corresponding SNP's markers. Most of the QTL here identified validated QTL previously detected for kernel-related traits, despite the different genetic backgrounds, the diversity of mapping populations, experimental conditions and/or genetic coverage of the maps used in each of the studies. For example, the loci related to kernel length and WL ratio found on chromosome 2B (cluster 1) and the QTL related to area located on chromosome 6B (cluster 5) coincided with the regions found by Desiderio et al., (2019) related to the same traits. Parallelly, QTL related to kernel width located on chromosome 3A (cluster 2), coincided with the loci found by Wang et al., (2019), also related to this trait. Lastly, on chromosome 4B (cluster 4), the loci related to area coincided with the QTL for the same trait detected by Mangini et al., (2021), and further described below.

Noteworthy, most of our QTL clusters for kernel size and shape factors overlapped with published QTL for TKW. In detail, the cluster 2 overlapped with QTL related to TKW from 2 different studies (Avni et al., 2018 and Sun et al., 2020), cluster 3 on 3B with QTL from Faris et al., (2014), cluster 4 with QTL from 2 studies (Blanco et al., 2012; Elouafi et al., 2004), cluster 5 on chromosome 6B with one QTL from Tzarfati et al., (2014), and cluster 6 on 7A with one QTL detected by Patil et al., (2013). This result confirms the common assumption, presented above, that the genetic control on kernel factors is also responsible for variability in kernel weight, also applied to the case of durum wheat.

Further, the cluster 3 (on chromosome 3B), which was taken as a putative one in this study, was validated through this approach, as the same loci associated with TKW was detected by Faris et al., (2014), using also a *T. dicoccum* derived

population. However, in this study the mentioned QTL was also found related to area, while no other overlapping published QTL were found related to this trait, which could be an indication of a likely new relationship involving the two traits found by this study.

Regarding the region on cluster 4, we validated our result in the case of area through the QTL found by Mangini et al., 2021, however the positive allele donated in that study was from a durum cultivar (named Liberdur), while in our case the donor is *T. dicoccum* (MG5323). Then, taking into consideration the nature of the population, Russo et al., (2014), identified a QTL on chromosome 4B related to TKW, kernel area and width, where the favorable allele was donated from the *T. dicoccum* line. However, this locus is located at about 27 Mbp on the reference genome and thus is unlikely to overlap with our cluster (594 Mbp - 619 Mbp on chromosome 4B). Overall, such comparisons suggest that the cluster 4 detected in this work is likely to be new for the relationships found (between kernel length, perimeter, area and TKW), but it was validated by QTL already found independently for TKW (Blanco et al., 2012; Elouafi et al., 2004) and kernel area (Mangini et al., 2021).

5.4. Candidate genes for the six QTL clusters

By exploiting available genomic tools and knowledge, we were able to hypothesize candidate genes from the physical intervals of each cluster and corroborate them with a list of known genes controlling kernel related traits and kernel weight already described in rice and/or wheat.

Noteworthy, in all clusters associated with kernel width (cluster 2, 5 and 6), genes involved in cell development were retrieved, strengthening the chances of being potential candidates for this trait. These genes were: the known gene *D61*, encoding a BR insensitive (BRI)-like leucine-rich repeat (LRR) receptor kinase (Avni et al., 2018) associated with cell elongation and located at cluster 2; the known genes *GW2* and *FUWA,* at cluster 5, controlling grain size by regulating cell division (Zhai et al., 2018; Chen et al., 2015); and lastly, TRITD7Av1G052720 and

TRITD7Av1G055870 at cluster 6, whose annotations mention the regulation of cell growth and cell differentiation. Parallely, in most of the clusters related to kernel area (cluster 3 and 4), genes associated to auxin metabolism were retrieved (TRITD3Bv1G229090, TRITD3Bv1G229910, TRITD3Bv1G235190, TRITD3Bv1G239650, TRITD4Bv1G175480, TRITD4Bv1G179270 and TRITD4Bv1G171270), which could imply the importance to consider these as candidates for this trait. Further, several lines of evidence have determined that auxins play an important role in organ size determination by affecting cell expansion, cell division and differentiation thus affecting stem elongation, lateral branching, vascular development, growth responses and various aspects of seed development, including development of the embryo, endosperm, and seed coat (Cao et al., 2020; Teale et al., 2006; Zhao, 2010).

Interestingly, near to cluster 4, we found the position of the gene called *BIG GRAIN 1* (*BG1*), which encodes a plasma membrane-associated protein (Liu et al., 2015), and could be involved in the control of the relationships found for this cluster. *BG1* is a known gene in rice (GenBank accession Q10R09.1), whose function has been described as a positive regulator of the auxin signaling pathway involved in gravitropism, plant growth and grain development. The over-expressing dominant mutants of this gene in rice showed increased grain size with bigger length, width and area, associated with longer epidermis cells and higher number of parenchyma cells in both the palea and lemma in the spikelet hull (Liu et al., 2015). Nevertheless, a recent study about this orthologous gene in common wheat showed that even if the overexpression of *BG1* led to larger seed size, it also triggered the reduction in seed number per plant (fewer grains), thus causing no significant overall increase in yield, and also was related to a lower concentration of essential elements (zinc and phosphorus) and protein content (Milner et al., 2021).

All together, it is clear that to increase yield potential it is needed a combination of different loci/alleles and traits that can show synergic relationships instead of known trade-offs. Therefore, it is key to comprehend the genetic bases of these complex quantitative traits, together with new alleles from less cultivated germplasm which can contribute to model the interactions between the yield components and the nutritional status, the activity of particular phytohormones or the

synchronization of floret development (Desiderio et al., 2019). An example of this, is proposed by Nigro et al., 2019 consisting in the identification of genetic sources, from a durum wheat collection (which included *T. dicoccum*) of elevated protein content without negative pleiotropic effects. The authors found four QTL associated with higher levels of protein and without affecting final grain yield per spike, showing it is possible to increase both traits.

In addition, we found genes impacting grain size which participate in other biological processes not involved in plant hormone regulation. For example, the above mentioned *GW2* gene, which encodes an E3 RING ligase and mediates ubiquitination in the ubiquitin-26S proteasome system. This gene has been shown to negatively regulate grain size in rice and has been already found present in bread wheat (Nadolska-Orczyk et al., 2017; Hong et al., 2014; Simmonds et al., 2016). We also detected *TaSus1* located within the physical interval of cluster 6. This gene encodes a sucrose synthase, catalyzing the first step in the conversion of sucrose to starch and has been correlated with thousand kernel weight, as starch is the main component of grain endosperm (70%) (Nadolska-Orczyk et al., 2017). As seen, the major genes determining yield-related traits can be classified in several groups, from which the ones detected in this study were related to metabolism or signaling of growth regulators, cell division, cell proliferation and carbohydrate metabolism.

Regarding the results for heading date, several QTL were detected but only the ones detected in chromosome 2A and 2B were co-located with known genes *Ppd*-A1 and *Ppd*-B1, consistent with observations made by Maccaferri et al., 2010. There was no coincidence detected with any of the known *Vrn* loci, which could be because none of the entries have any vernalization requirement (Wang et al., 2019).

6. Conclusions

Uncovering the genetic mechanisms of the relationships between kernel weight and seed size/shape, taking into the basis a *T. dicoccum* derived population, is of great significance for improving wheat yield. In this way, the current study contributes to lay the foundations on how grain weight can be improved through its components, by identifying 3 clusters of co-locating loci on chromosomes 3B, 4B and 6B. Especially, the stable loci on 4B were repeatedly detected in two or more environments and across them for the kernel size traits (kernel length, area and perimeter) and kernel weight, being the superior allele donated by the *T. dicoccum* line. Therefore, this study further supports the underlying possibility of this ancestral species as a source of favorable alleles for durum wheat germplasm.

Thanks to the available consensus map and reference genome, validation with previous QTL and identification of candidate genes was facilitated. A good candidate for kernel size is proposed here, being the rice orthologue *BG1*, detected within the loci found on chromosome 4B and whose role is to regulate auxin transport. The availability of SNP markers within candidate gene sequences might represent a breeding strategy based on functional markers, however it is important to also contemplate the trade-offs with other yield components and find synergic collaborations between them.

As for future perspectives, it is important to further validate the detected QTL in this study, considering techniques such as fine mapping and cloning, to confirm the correspondent identity of the genes that were hypothesized here and study their interactions with other genes and traits. Additionally, it is needed to deepen the understanding of the dependency of these QTL to the environment. For this aim, more analysis to understand the Genotype x Environment interaction (GEI) is required, for example, applying different models to explain the environmental differences and models with epistatic interactions, in order to assure the efficiency and stability of the QTL as targets in future breeding programs.

7. Bibliography

Alemu, A., Feyissa, T., Tuberosa, R., Maccaferri, M., Sciara, G., Letta, T., & Abeyo, B. (2020). Genome-wide association mapping for grain shape and color traits in Ethiopian durum wheat (*Triticum turgidum* ssp. *durum*). *The Crop Journal*, *8*(5), 757–768. https://doi.org/10.1016/j.cj.2020.01.001

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research.*, 19. https://doi.org/10.1101/gr.094052.109

Arriagada, O., Marcotuli, I., Gadaleta, A., & Schwember, A. R. (2020). Molecular Mapping and Genomics of Grain Yield in Durum Wheat: A Review. *International Journal of Molecular Sciences*, *21*(19). https://doi.org/10.3390/ijms21197021

Arzani, A. (2019). Chapter 7—Emmer (*Triticum turgidum* ssp. *dicoccum*) Flour and Bread. In V. R. Preedy & R. R. Watson (Eds.), *Flour and Breads and their Fortification in Health and Disease Prevention (Second Edition)*. Academic Press. 89–98. https://doi.org/10.1016/B978-0-12-814639-2.00007-1

Avni, R. (2014). Ultra-dense genetic map of durum wheat×wild emmer wheat developed using the 90K iSelect SNP genotyping assay. *Molecular Breeding*, 34. https://doi.org/10.1007/s11032-014-0176-2

Avni, R. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, *357*. https://doi.org/10.1126/science.aan0032

Avni, R., Oren, L., Assili, S., Pozniak, C., Hale, I., Ben-David, R., Peleg, Z., & Distelfeld, A. (2018). Genome Based Meta-QTL Analysis of Grain Weight in Tetraploid Wheat Identifies Rare Alleles of GRF4 Associated with Larger Grains. *Genes*, *9*(12), 636. https://doi.org/10.3390/genes9120636

Bartholomé, J., Bink, M. C., Heerwaarden, J. van, Chancerel, E., Boury, C., Lesur, I., Isik, F., Bouffier, L., & Plomion, C. (2016). Linkage and Association Mapping for Two Major Traits Used in the Maritime Pine Breeding Program: Height Growth and Stem Straightness. PLOS ONE, 11(11), e0165323. https://doi.org/10.1371/journal.pone.0165323

Bassi, F. M., Brahmi, H., Sabraoui, A., Amri, A., Nsarellah, N., Nachit, M. M., Al-Abdallat, A., Chen, M. S., Lazraq, A., & El Bouhssini, M. (2019). Genetic identification of loci for Hessian fly resistance in durum wheat. *Molecular Breeding*, 39(2), 24. https://doi.org/10.1007/s11032-019-0927-1

Beier, S. (2017). Construction of a map-based reference genome sequence for barley, Hordeum vulgare L. *Science Data*, 4. https://doi.org/10.1038/sdata.2017.44

Beres, B. L., Rahmani, E., Clarke, J. M., Grassini, P., Pozniak, C. J., Geddes, C. M., Porker, K. D., May, W. E., & Ransom, J. K. (2020). A Systematic Review of Durum Wheat: Enhancing Production Systems by Exploring Genotype, Environment, and Management (G × E × M) Synergies. *Frontiers in Plant Science*, 11. https://doi.org/10.3389/fpls.2020.568657

Bozzini, A., David, J., & Natoli, V. (2012). CHAPTER 1—Origin and Distribution of Durum Wheat Genetic Diversity in the World. In M. Sissons, J. Abecassis, B. Marchylo, & M. Carcea (Eds.), *Durum Wheat (Second Edition)*. AACC International Press. 1–14 https://doi.org/10.1016/B978-1-891127-65-6.50006-4

Broman, K. W., Wu, H., Sen, S., & Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics (Oxford, England)*, *19*(7), 889–890. https://doi.org/10.1093/bioinformatics/btg112

Brozynska, M., Furtado, A., & Henry, R. J. (2016). Genomics of crop wild relatives: Expanding the gene pool for crop improvement. *Plant Biotechnology Journal*, 14. https://doi.org/10.1111/pbi.12454

Cakmak, I. (1996). Zinc deficiency as a critical problem in wheat production in Central Anatolia. *Plant Soil*, 180. https://doi.org/10.1007/BF00015299

Cao, J., Li, G., Qu, D., Li, X., & Wang, Y. (2020). Into the Seed: Auxin Controls Seed Development and Grain Yield. *International Journal of Molecular Sciences*, 21(5), 1662. https://doi.org/10.3390/ijms21051662

Cao, J., Shang, Y., Xu, D., Xu, K., Cheng, X., Pan, X., Liu, X., Liu, M., Gao, C., Yan, S., Yao, H., Gao, W., Lu, J., Zhang, H., Chang, C., Xia, X., Xiao, S., & Ma, C. (2020). Identification and Validation of New Stable QTLs for Grain Weight and Size by Multiple Mapping Models in Common Wheat. *Frontiers in Genetics*, 11, 584859. https://doi.org/10.3389/fgene.2020.584859

Chapman, J. A. (2015). A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biology*, 16. https://doi.org/10.1186/s13059-015-0582-8

Chen, H., Patterson, N., & Reich, D. (2010). Population differentiation as a test for selective sweeps. *Genome Research*, 20. https://doi.org/10.1101/gr.100545.109

Chen, J., Gao, H., Zheng, X.-M., Jin, M., Weng, J.-F., Ma, J., Ren, Y., Zhou, K., Wang, Q., Wang, J., Wang, J.-L., Zhang, X., Cheng, Z., Wu, C., Wang, H., & Wan, J.-M. (2015). An evolutionarily conserved gene, FUWA, plays a role in determining panicle architecture, grain shape and grain weight in rice. *The Plant Journal*, 83(3), 427–438. https://doi.org/10.1111/tpj.12895

Cheng, X., Xin, M., Xu, R., Chen, Z., Cai, W., Chai, L., Xu, H., Jia, L., Feng, Z., Wang, Z., Peng, H., Yao, Y., Hu, Z., Guo, W., Ni, Z., & Sun, Q. (2020). A Single Amino Acid Substitution in STKc_GSK3 Kinase Conferring Semispherical Grains and Its Implications for the Origin of Triticum sphaerococcum[OPEN]. *The Plant Cell*, *32*(4), 923–934. https://doi.org/10.1105/tpc.19.00580

Colasuonno, P., Marcotuli, I., Gadaleta, A., & Soriano, J. M. (2021). From Genetic Maps to QTL Cloning: An Overview for Durum Wheat. *Plants*, 10(2), 315. https://doi.org/10.3390/plants10020315

Cui, F., Ding, A., Li, J., Zhao, C., Li, X., Feng, D., Wang, X., Wang, L., Gao, J., & Wang, H. (2011). Wheat kernel dimensions: How do they contribute to kernel weight at an individual QTL level? *Journal of Genetics*, 90(3), 409–425. https://doi.org/10.1007/s12041-011-0103-9

De Vita, P., Riefolo, C., Codianni, P., Cattivelli, L., & Fares, C. (2006). Agronomic and qualitative traits of *T. turgidum* ssp. *dicoccum* genotypes cultivated in Italy. *Euphytica*, 150(1–2), 195–205. https://doi.org/10.1007/s10681-006-9107-6

De Vita, P., & Taranto, F. (2019). Durum Wheat (*Triticum turgidum* ssp. *durum*) Breeding to Meet the Challenge of Climate Change. In J. M. Al-Khayri, S. M. Jain, & D. V. Johnson (Eds.), *Advances in Plant Breeding Strategies: Cereals: Volume 5*, 471–524. Springer International Publishing. https://doi.org/10.1007/978-3-030-23108-8_13

Desiderio, F., Guerra, D., Rubiales, D., Piarulli, L., Pasquini, M., Mastrangelo, A. M., Simeone, R., Blanco, A., Cattivelli, L., & Vale', G. (2014). Identification and mapping of quantitative trait loci for leaf rust resistance derived from a tetraploid wheat Triticum dicoccum accession. *Molecular Breeding*, 17.

Desiderio, F., & Mazzucotelli, E. (2019). Genomic Regions From an Iranian Landrace Increase Kernel Size in Durum Wheat. *Frontiers in Plant Science*, *10*, 21.https://doi.org/10.3389/fpls.2019.00448

Dhanavath, S., & Rao, U. J. S. P. (2017). Nutritional and Nutraceutical Properties of Triticum dicoccum Wheat and Its Health Benefits: An Overview. *Journal of Food Science*, 82(10), 2243–2250. https://doi.org/10.1111/1750-3841.13844

Dong, L., Wang, F., Liu, T., Dong, Z., Li, A., Jing, R., Mao, L., Li, Y., Liu, X., Zhang, K., & Wang, D. (2014). Natural variation of TaGASR7-A1 affects grain length in common wheat under multiple cultivation conditions. *Molecular Breeding*, 34(3), 937–947. https://doi.org/10.1007/s11032-014-0087-2

Dubcovsky, J., & Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, 316. https://doi.org/10.1126/science.1143986

Dvorak, J., McGuire, P. E., & Cassidy, B. (1988). Apparent sources of the A genomes of wheats inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. *Genome*, 30. https://doi.org/10.1139/g88-115

EC. (2020). *Short-Term Outlook for EU Agricultural Markets in 2020*. European Commission, DG Agriculture and Rural Development, Brussels. https://ec.europa.eu/info/sites/default/files/food-farming-fisheries/farming/documents/short-term-outlook-summer-2020_en.pdf

Ellis, J. G., Lagudah, E. S., Spielmeyer, W., & Dodds, P. N. (2014). The past, present and future of breeding rust resistant wheat. *Frontiers Plant Science*, 5. https://doi.org/10.3389/fpls.2014.00641

Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M., & Servin, B. (2013). Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics*, 193. https://doi.org/10.1534/genetics.112.147231

Faris, J. D., Zhang, Z., & Chao, S. (2014). Map-based analysis of the tenacious glume gene Tg-B1 of wild emmer and its role in wheat domestication. *Gene*, 542. https://doi.org/10.1016/j.gene.2014.03.034

Fatima, F., McCallum, B. D., Pozniak, C. J., Hiebert, C. W., McCartney, C. A., Fedak, G., You, F. M., & Cloutier, S. (2020). Identification of New Leaf Rust Resistance Loci in Wheat and Wild Relatives by Array-Based SNP Genotyping and Association Genetics. *Frontiers in Plant Science*, 11, 1728. https://doi.org/10.3389/fpls.2020.583738

Fatiukha, A., Filler, N., Lupo, I., Lidzbarsky, G., Klymiuk, V., Korol, A. B., Pozniak, C., Fahima, T., & Krugman, T. (2020). Grain protein content and thousand kernel weight QTLs identified in a durum × wild emmer wheat mapping population tested in five environments. *Theoretical and Applied Genetics*, 133(1), 119–131. https://doi.org/10.1007/s00122-019-03444-8

Ficco, D. B. M., Beleggia, R., Pecorella, I., Giovanniello, V., Frenda, A. S., & Vita, P. D. (2020). Relationship between Seed Morphological Traits and Ash and Mineral Distribution along the Kernel Using Debranning in Durum Wheats from Different Geographic Sites. *Foods*, 9(11). https://doi.org/10.3390/foods9111523

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*. https://doi.org/10.1534/genetics.113.160572

Fu, J., Bowden, R. L., Prasad, P. V. V., & Ibrahim, A. M. H. (2015). Genetic Variation for Heat Tolerance in Primitive Cultivated Subspecies of Triticum turgidum L. *Journal of Crop Improvement*, 29(5), 565–580. https://doi.org/10.1080/15427528.2015.1060915

Gaut, B. S. (2015). Evolution is an experiment: Assessing parallelism in crop domestication and experimental evolution. *Molecular Biology and Evolution*, 32. https://doi.org/10.1093/molbev/msv105

Gegas, V. C., Nazari, A., Griffiths, S., Simmonds, J., Fish, L., Orford, S., Sayers, L., Doonan, J. H., & Snape, J. W. (2010). A Genetic Framework for Grain Size and Shape Variation in Wheat. *The Plant Cell*, 22(4), 1046–1056. https://doi.org/10.1105/tpc.110.074153

Giri, P., Yadav, M. L., & Mohapatra, B. (2018). QTL Linkage Analysis. In J. Vonk & T. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior*. Springer International Publishing. 1–6. https://doi.org/10.1007/978-3-319-47829-6_161-1

Government of Canada, C. F. I. A. (2012, March 5). *The Biology of Triticum turgidum* ssp. *durum (Durum Wheat)* [Reference material]. https://inspection.canada.ca/plant-varieties/plants-with-novel-traits/applicants/directive-94-08/biology-documents/triticum-turgidum-ssp-durum/eng/1330983955477/1330984025320

Guan, P., Lu, L., Jia, L., Kabir, M. R., Zhang, J., Lan, T., Zhao, Y., Xin, M., Hu, Z., Yao, Y., Ni, Z., Sun, Q., & Peng, H. (2018). Global QTL Analysis Identifies Genomic Regions on Chromosomes 4A and 4B Harboring Stable Loci for Yield-Related Traits Across Different Environments in Wheat (Triticum aestivum L.). *Frontiers in Plant Science*, *0*. https://doi.org/10.3389/fpls.2018.00529

Gupta, P. K., Balyan, H. S., Sharma, S., & Kumar, R. (2020). Genetics of yield, abiotic stress tolerance and biofortification in wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *133*(5), 1569–1602. https://doi.org/10.1007/s00122-020-03583-3

Harris, N. S., & Taylor, G. J. (2013). Cadmium uptake and partitioning in durum wheat during grain filling. *BMC Plant Biology*, 13. https://doi.org/10.1186/1471-2229-13-103

Hart, J. J., Welch, R. M., Norvell, W. A., Clarke, J. M., & Kochian, L. V. (2005). Zinc effects on cadmium accumulation and partitioning in near-isogenic lines of durum wheat that differ in grain cadmium concentration. *New Phytologist*, 167. https://doi.org/10.1111/j.1469-8137.2005.01416.x

He, X. Y., He, Z. H., Ma, W., Appels, R., & Xia, X. C. (2009). Allelic variants of phytoene synthase 1 (Psy1) genes in Chinese and CIMMYT wheat cultivars and development of functional markers for flour colour. *Molecular Breeding*, 23. https://doi.org/10.1007/s11032-009-9255-1

Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: Defining, estimating and interpreting FST. *Natural Reviews Genetics*, 10. https://doi.org/10.1038/nrg2611

Hong, Y., Chen, L., Du, L., Su, Z., Wang, J., Ye, X., Qi, L., & Zhang, Z. (2014). Transcript suppression of TaGW2 increased grain width and weight in bread wheat. *Functional & Integrative Genomics*, 14(2), 341–349. https://doi.org/10.1007/s10142-014-0380-5

Hu, M.-J., Zhang, H.-P., Cao, J.-J., Zhu, X.-F., Wang, S.-X., Jiang, H., Wu, Z. Y., Lu, J., Chang, C., Sun, G.-L., & Ma, C.-X. (2016). Characterization of an IAA-glucose hydrolase gene TaTGW6 associated with grain weight in common wheat (Triticum aestivum L.). *Molecular Breeding*, 36(3), 25. https://doi.org/10.1007/s11032-016-0449-z

Hu, M.-J., Zhang, H.-P., Liu, K., Cao, J.-J., Wang, S.-X., Jiang, H., Wu, Z.-Y., Lu, J., Zhu, X. F., Xia, X.-C., Sun, G.-L., Ma, C.-X., & Chang, C. (2016). Cloning and Characterization of TaTGW-7A Gene Associated with Grain Weight in Wheat via SLAF-seq-BSA. *Frontiers in Plant Science*, *0*. https://doi.org/10.3389/fpls.2016.01902

ITIS Standard Report Page: Triticum durum. (n.d.). Retrieved May 9, 2021, from https://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=42240#null

Jiang, Y., Jiang, Q., Hao, C., Hou, J., Wang, L., Zhang, H., Zhang, S., Chen, X., & Zhang, X. (2015). A yield-associated gene TaCWI, in wheat: Its function, selection and evolution in global breeding revealed by haplotype analysis. *Theoretical and Applied Genetics*, 128(1), 131–143. https://doi.org/10.1007/s00122-014-2417-5

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, 11. https://doi.org/10.1186/1471-2156-11-94

Kadkol, G., & Sissons, M. (2016). Durum Wheat: Overview. In *Reference Module in Food Science*, 117–124. https://doi.org/10.1016/B978-0-08-100596-5.00024-X

Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., & Chen, L. (2011). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nature Biotechnology.*, 30. https://doi.org/10.1038/nbt.2057

Knox, R. E. (2009). Chromosomal location of the cadmium uptake gene (Cdu1) in durum wheat. *Genome*, 52. https://doi.org/10.1139/G09-042

Konvalina, P., Moudrý, J., Dotlačil, L., & Stehno, Z. (2010). Drought Tolerance of Land Races of Emmer Wheat in Comparison to Soft Wheat. *Cereal Research Communications*, 38(3), 429–439.

Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, 8. https://doi.org/10.1371/journal.pgen.1002453

Li, N., Xu, R., Duan, P., & Li, Y. (2018). Control of grain size in rice. *Plant Reproduction*, 31(3), 237–251. https://doi.org/10.1007/s00497-018-0333-6

Li, S., Jia, J., Wei, X., Zhang, X., Li, L., Chen, H., Fan, Y., Sun, H., Zhao, X., Lei, T., Xu, Y., Jiang, F., Wang, H., & Li, L. (2007). A intervarietal genetic map and QTL analysis for yield traits in wheat. *Molecular Breeding*, *20*(2), 167–178. https://doi.org/10.1007/s11032-007-9080-3

Li, S., Wang, L., Meng, Y., Hao, Y., Xu, H., Hao, M., Lan, S., Zhang, Y., Lv, L., Zhang, K., Peng, X., Lan, C., Li, X., & Zhang, Y. (2021). Dissection of Genetic Basis Underpinning Kernel Weight-Related Traits in Common Wheat. *Plants*, 10(4), 713. https://doi.org/10.3390/plants10040713

Liu, L., Tong, H., Xiao, Y., Che, R., Xu, F., Hu, B., Liang, C., Chu, J., Li, J., & Chu, C. (2015). Activation of Big Grain1 significantly improves grain size by regulating auxin transport in rice. *Proceedings of the National Academy of Sciences*, 112(35), 11102–11107. https://doi.org/10.1073/pnas.1512748112

Liu, R., Hou, J., Li, H., Xu, P., Zhang, Z., & Zhang, X. (2021). Association of TaD14-4D, a Gene Involved in Strigolactone Signaling, with Yield Contributing Traits in Wheat. *International Journal of Molecular Sciences*, 22(7), 3748. https://doi.org/10.3390/ijms22073748

Liu, S., Ke, S., Tang, G., Huang, D., Wei, M., Zhang, Y., Qin, G., & Zhang, X.-Q. (2021). OsPEX1, a leucine-rich repeat extensin protein, plays an important role in the regulation of caryopsis development in rice. *Research Square.* https://doi.org/10.21203/rs.3.rs-51139/v1

Liu, W., Maccaferri, M., Chen, X., Laghetti, G., Pignone, D., Pumphrey, M., & Tuberosa, R. (2017). Genome-wide association mapping reveals a rich genetic architecture of stripe rust resistance loci in

emmer wheat (*Triticum turgidum* ssp. *dicoccum*). *Theoretical and Applied Genetics*, 130(11), 2249–2270. https://doi.org/10.1007/s00122-017-2957-6

Luo, M. C. (2007). The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. *Theoretical and Applied Genetics*, 114. https://doi.org/10.1007/s00122-006-0474-0

Ma, J., Ding, P., Qin, P., Liu, Y.-X., Xie, Q., Chen, G., Li, W., Jiang, Q., Chen, G., Lan, X.-J., Wei, Y.-M., Liu, C., & Zheng, Y.-L. (2017). Structure and expression of the TaGW7 in bread wheat (*Triticum aestivum L.*). *Plant Growth Regulation*, 82(2), 281–291. https://doi.org/10.1007/s10725-017-0258-3

Ma, J., Qu, X., Liu, J., Xie, X., Xu, Q., Tang, H., Mu, Y., Pu, Z., Li, Y., Ma, J., Gao, Y., Jiang, Q.-T., Liu, Y., Chen, G., & Wang, J. (2021). Genetic mapping and validation of loci for kernel-related traits in wheat *(Triticum aestivum L.)*. *Frontiers in Plant Science,* 49. https://doi.org/10.3389/fpls.2021.667493

Maccaferri, M. (2008). A major QTL for durable leaf rust resistance widely exploited in durum wheat breeding programs maps on the distal region of chromosome arm 7BL. *Theoretical and Applied Genetics*, *117*. https://doi.org/10.1007/s00122-008-0857-5

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., Ormanbekova, D., Lux, T., Prade, V. M., Milner, S. G., Himmelbach, A., Mascher, M., Bagnaresi, P., Faccioli, P., Cozzi, P., Lauria, M., Lazzari, B., Stella, A., Manconi, A., … Cattivelli, L. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics*, 51(5), 885–895. https://doi.org/10.1038/s41588-019-0381-3

Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., Casadio, R., Akhunov, E., Scalabrin, S., Vendramin, V., Ammar, K., Blanco, A., Desiderio, F., Distelfeld, A., Dubcovsky, J., Fahima, T., Faris, J., Korol, A., Massi, A., … Tuberosa, R. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnology Journal*, 13(5), 648–663. https://doi.org/10.1111/pbi.12288

Malosetti, M., Ribaut, J.-M., & van Eeuwijk, F. A. (2013). The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology*, 4. https://doi.org/10.3389/fphys.2013.00044

Mangini, G., Blanco, A., Nigro, D., Signorile, M. A., & Simeone, R. (2021). Candidate Genes and Quantitative Trait Loci for Grain Yield and Seed Size in Durum Wheat. *Plants*, 10(2), 312. https://doi.org/10.3390/plants10020312

Mangini, G., Gadaleta, A., Colasuonno, P., Marcotuli, I., Signorile, A. M., Simeone, R., De Vita, P., Mastrangelo, A. M., Laidò, G., Pecchioni, N., & Blanco, A. (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

Martínez-Moreno, F., Solís, I., Noguero, D., Blanco, A., Özberk, İ., Nsarellah, N., Elias, E., Mylonas, I., & Soriano, J. M. (2020). Durum wheat in the Mediterranean Rim: Historical evolution and genetic resources. *Genetic Resources and Crop Evolution*, 67(6), 1415–1436. https://doi.org/10.1007/s10722-020-00913-8

Mascher, M. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature*, *544*. https://doi.org/10.1038/nature22043

Mazzucotelli, E., Sciara, G., Mastrangelo, A. M., Desiderio, F., Xu, S. S., Faris, J., Hayden, M. J., Tricker, P. J., Ozkan, H., Echenique, V., Steffenson, B. J., Knox, R., Niane, A. A., Udupa, S. M., Longin, F. C. H., Marone, D., Petruzzino, G., Corneti, S., Ormanbekova, D., Bassi, F. M. (2020). The Global Durum Wheat Panel (GDP): An International Platform to Identify and Exchange Beneficial Alleles. *Frontiers in Plant Science*, 11. https://doi.org/10.3389/fpls.2020.569905

Mérida-García, R., Bentley, A. R., Gálvez, S., Dorado, G., Solís, I., Ammar, K., & Hernandez, P. (2020). Mapping Agronomic and Quality Traits in Elite Durum Wheat Lines under Differing Water Regimes. *Agronomy*, 10(1), 144. https://doi.org/10.3390/agronomy10010144

Miao, L., Li, Y., Zhang, H., Zhang, H., Liu, X., Wang, J., Chang, X., Mao, X., & Jing, R. (2021). TaSnRK2.4 is a vital regulator in control of thousand-kernel weight and response to abiotic stress in wheat. *Journal of Integrative Agriculture*, 20(1), 46–54. https://doi.org/10.1016/S2095-3119(19)62830-3

Miao, L., Mao, X., Wang, J., Liu, Z., Zhang, B., Li, W., Chang, X., Reynolds, M., Wang, Z., & Jing, R. (2017). Elite Haplotypes of a Protein Kinase Gene TaSnRK2.3 Associated with Important Agronomic Traits in Common Wheat. *Frontiers in Plant Science*, *0*. https://doi.org/10.3389/fpls.2017.00368

Miles, C. & Wayne, M. (2008). *Quantitative Trait Locus (QTL) Analysis | Learn Science at Scitable.*. Retrieved June 3, 2021, from https://www.nature.com/scitable/topicpage/quantitative-trait-locus-qtl-analysis-53904/

Mills, R. F. (2005). The plant P1B-type ATPase AtHMA4 transports Zn and Cd and plays a role in detoxification of transition metals supplied at elevated levels. *FEBS Letters*, 579. https://doi.org/10.1016/j.febslet.2004.12.040

Milner, M. J., Bowden, S., Craze, M., & Wallington, E. J. (2021). TaBG1 Increases Seed Size and Alters Nutritional Characteristics of the Grain in Wheat But Does Lead to Increased Yields. *BMC Plant Biology*. https://doi.org/10.21203/rs.3.rs-523405/v1

Miyadate, H. (2011). OsHMA3, a P1B-type of ATPase affects root-to-shoot cadmium translocation in rice by mediating efflux into vacuoles. *New Phytologist*, 189. https://doi.org/10.1111/j.1469-8137.2010.03459.x

Mohammadi, M., Mirlohi, A., Majidi, M. M., & Soleimani Kartalaei, E. (2021). Emmer wheat as a source for trait improvement in durum wheat: A study of general and specific combining ability. *Euphytica*, 217(4), 64. https://doi.org/10.1007/s10681-021-02796-x

Mohler, V., Albrecht, T., Castell, A., Diethelm, M., Schweizer, G., & Hartl, L. (2016). Considering causal genes in the genetic dissection of kernel traits in common wheat. *Journal of Applied Genetics*, 57(4), 467–476. https://doi.org/10.1007/s13353-016-0349-2

Montenegro, J. D. (2017). The pangenome of hexaploid bread wheat. *Plant Journal*, 90. https://doi.org/10.1111/tpj.13515

Montesinos-López, O. A., Montesinos-López, A., Tuberosa, R., Maccaferri, M., Sciara, G., Ammar, K., & Crossa, J. (2019). Multi-Trait, Multi-Environment Genomic Prediction of Durum Wheat With Genomic Best Linear Unbiased Predictor and Deep Learning Methods. *Frontiers in Plant Science*, 10. https://doi.org/10.3389/fpls.2019.01311

Nadolska-Orczyk, A., Rajchel, I. K., Orczyk, W., & Gasparis, S. (2017). Major genes determining yield-related traits in wheat and barley. *Theoretical and Applied Genetics*, 130(6), 1081–1098. https://doi.org/10.1007/s00122-017-2880-x

Nakamura, A., Fujioka, S., Sunohara, H., Kamiya, N., Hong, Z., Inukai, Y., Miura, K., Takatsuto, S., Yoshida, S., Ueguchi-Tanaka, M., Hasegawa, Y., Kitano, H., & Matsuoka, M. (2006). The role of OsBRI1 and its homologous genes, OsBRL1 and OsBRL3, in rice. *Plant Physiology*, 140(2), 580–590. https://doi.org/10.1104/pp.105.072330

Nave, M., Avni, R., Ben-Zvi, B., Hale, I., & Distelfeld, A. (2016). QTLs for uniform grain dimensions and germination selected during wheat domestication are co-located on chromosome 4B. *Theoretical and Applied Genetics*, 129(7), 1303–1315. https://doi.org/10.1007/s00122-016-2704-4

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences- USA*, 12. https://doi.org/10.1073/pnas.70.12.3321

Nigro, D., Gadaleta, A., Mangini, G., Colasuonno, P., Marcotuli, I., Giancaspro, A., Giove, S. L., Simeone, R., & Blanco, A. (2019). Candidate genes and genome-wide association study of grain protein content and protein deviation in durum wheat. *Planta*, 249(4), 1157–1175. https://doi.org/10.1007/s00425-018-03075-1

Okada, T., Jayasinghe, J. E. A. R. M., Nansamba, M., Baes, M., Warner, P., Kouidri, A., Correia, D., Nguyen, V., Whitford, R., & Baumann, U. (2018). Unfertilized ovary pushes wheat flower open for cross-pollination. *Journal of Experimental Botany*, *69*(3), 399–412. https://doi.org/10.1093/jxb/erx410

Olivera, Pablo, J., Yue. (2014). *Genetic resources for stem rust resistance in cultivated and wild tetraploid wheats.* Options Méditerranéennes, A No. 110, 2014 - Proceedings of the International Symposium on Genetics and breeding of durum wheat.

Olivoto, T., & Lúcio, A. (2020). metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution*, 11, 783–789. https://doi.org/10.1111/2041-210x.13384

Özkan, H., Brandolini, A., Schäfer-Pregl, R., & Salamini, F. (2002). AFLP analysis of a collection of tetraploid wheats indicates the origin of emmer and hard wheat domestication in southeast Turkey. *Molecular Biology and Evolution*, 19. https://doi.org/10.1093/oxfordjournals.molbev.a004002

Pankin, A., Altmüller, J., Becker, C., & Korff, M. (2018). Targeted resequencing reveals genomic signatures of barley domestication. *New Phytologist*, 218. https://doi.org/10.1111/nph.15077

Patil, R. M., Tamhankar, S. A., Oak, M. D., Raut, A. L., Honrao, B. K., Rao, V. S., & Misra, S. C. (2013). Mapping of QTL for agronomic traits and kernel characters in durum wheat (*Triticum durum* Desf.). *Euphytica*, 190(1), 117–129. https://doi.org/10.1007/s10681-012-0785-y

Penner, G. A., Clarke, J., Bezte, L. J., & Leisle, D. (1995). Identification of RAPD markers linked to a gene governing cadmium uptake in durum wheat. *Genome*, 38. https://doi.org/10.1139/g95-070

Piarulli, L., Gadaleta, A., Mangini, G., Signorile, M. A., Pasquini, M., Blanco, A., & Simeone, R. (2012). Molecular identification of a new powdery mildew resistance gene on chromosome 2BS from *Triticum turgidum* ssp. *dicoccum*. *Plant Science*, 196, 101–106. https://doi.org/10.1016/j.plantsci.2012.07.015

Piepho, H. P., Möhring, J., Melchinger, A. E., & Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*, 161(1), 209–228. https://doi.org/10.1007/s10681-007-9449-8

Piergiovanni A.R., Simeone, R., Pasqualone, A. (2009). Composition of whole and refined meals of Kamut under southern Italian conditions. *Chemistry Engineering Transactions 17, 891–896*.

R Core Team. (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing,* Vienna, Austria.*https://www.R-project.org/*.

Rahman, S., Islam, S., Yu, Z., She, M., Nevo, E., & Ma, W. (2020). Current Progress in Understanding and Recovering the Wheat Genes Lost in Evolution and Domestication. *International Journal of Molecular Sciences*, 21(16). https://doi.org/10.3390/ijms21165836

Rehman Arif, M. A., Attaria, F., Shokat, S., Akram, S., Waheed, M. Q., Arif, A., & Börner, A. (2020). Mapping of QTLs Associated with Yield and Yield Related Traits in Durum Wheat (Triticum durum Desf.) Under Irrigated and Drought Conditions. *International Journal of Molecular Sciences*, 21(7). https://doi.org/10.3390/ijms21072372

Russo, M. A., Ficco, D. B. M., Laidò, G., Marone, D., Papa, R., Blanco, A., Gadaleta, A., De Vita, P., & Mastrangelo, A. M. (2014). A dense durum wheat × T. dicoccum linkage map based on SNP markers for the study of seed morphology. *Molecular Breeding*, 34(4), 1579–1597. https://doi.org/10.1007/s11032-014-0181-5

Sabeti, P. C. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449. https://doi.org/10.1038/nature06250
Saintenac, C., Jiang, D., & Akhunov, E. D. (2011). Targeted analysis of nucleotide and copy number variation by exon capture in allotetraploid wheat genome. *Genome Biology*, 12. https://doi.org/10.1186/gb-2011-12-9-r88
Segami, S., Takehara, K., Yamamoto, T., Kido, S., Kondo, S., Iwasaki, Y., & Miura, K. (2017). Overexpression of SRS5 improves grain size of brassinosteroid-related dwarf mutants in rice (*Oryza sativa* L.). *Breeding Science*, 67(4), 393–397. https://doi.org/10.1270/jsbbs.16198
Shi, W., Hao, C., Zhang, Y., Cheng, J., Zhang, Z., Liu, J., Yi, X., Cheng, X., Sun, D., Xu, Y., Zhang, X., Cheng, S., Guo, P., & Guo, J. (2017). A Combined Association Mapping and Linkage Analysis of Kernel Number Per Spike in Common Wheat (*Triticum aestivum* L.). *Frontiers in Plant Science*, 8, 1412. https://doi.org/10.3389/fpls.2017.01412
Simmonds, J., Scott, P., Brinton, J., Mestre, T. C., Bush, M., del Blanco, A., Dubcovsky, J., & Uauy, C. (2016). A splice acceptor site mutation in TaGW2-A1 increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theoretical and Applied Genetics*, 129(6), 1099–1112. https://doi.org/10.1007/s00122-016-2686-2
Steffenson, B. J. (2007). A walk on the wild side: Mining wild wheat and barley collections for rust resistance genes. *Crop and Pasture Science.*, 58. https://doi.org/10.1071/AR07123
Sun, L., Huang, S., Sun, G., Zhang, Y., Hu, X., Nevo, E., Peng, J., & Sun, D. (2020). SNP-based association study of kernel architecture in a worldwide collection of durum wheat germplasm. *PLOS ONE*, 15(2), e0229159. https://doi.org/10.1371/journal.pone.0229159
Takenaka, S., & Kawahara, T. (2012). Evolution and dispersal of emmer wheat (*Triticum* sp.) from novel haplotypes of Ppd-1 (photoperiod response) genes and their surrounding DNA sequences. *Theoretical and Applied Genetics*, 125(5), 999–1014. https://doi.org/10.1007/s00122-012-1890-y
Taranto, F., D'Agostino, N., Rodriguez, M., Pavan, S., Minervini, A. P., Pecchioni, N., Papa, R., & De Vita, P. (2020). Whole Genome Scan Reveals Molecular Signatures of Divergence and Selection Related to Important Traits in Durum Wheat Germplasm. *Frontiers in Genetics*, 11. https://doi.org/10.3389/fgene.2020.00217
Teale, W. D., Paponov, I. A., & Palme, K. (2006). Auxin in action: Signalling, transport and the control of plant growth and development. *Nature Reviews. Molecular Cell Biology*, *7*(11), 847–859. https://doi.org/10.1038/nrm2020
THE INTERNATIONAL WHEAT GENOME SEQUENCING CONSORTIUM (IWGSC). (2018). Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. *Science*, 361. https://doi.org/10.1126/science.aar7191
Tyagi, S., Mir, R. R., Balyan, H. S., & Gupta, P. K. (2015). Interval mapping and meta-QTL analysis of grain traits in common wheat (*Triticum aestivum* L.). *Euphytica*, 201(3), 367–380. https://doi.org/10.1007/s10681-014-1217-y
Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., & Dubcovsky, J. A. (2006). NAC gene regulating senescence improves grain protein, zinc and iron content in wheat. *Science*, 314. https://doi.org/10.1126/science.1133649
Ueno, D. (2010). Gene limiting cadmium accumulation in rice. *Proceedings of the National Academy of Sciences - USA*, 107. https://doi.org/10.1073/pnas.1005396107
Ur Rehman, S., Wang, J., Chang, X., Zhang, X., Mao, X., & Jing, R. (2019). A wheat protein kinase gene TaSnRK2.9-5A associated with yield contributing traits. *Theoretical and Applied Genetics*, 132(4), 907–919. https://doi.org/10.1007/s00122-018-3247-7
Varshney, R. K. (2017). Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. *Nature Genetics,* 49. https://doi.org/10.1038/ng.3872
Vendramin, V., Ormanbekova, D., Scalabrin, S., Scaglione, D., Maccaferri, M., Martelli, P., Salvi, S., Jurman, I., Casadio, R., Cattonaro, F., Tuberosa, R., Massi, A., & Morgante, M. (2019). Genomic tools for durum wheat breeding: De novo assembly of Svevo transcriptome and SNP discovery in elite germplasm. *BMC Genomics*, 20(1), 278. https://doi.org/10.1186/s12864-019-5645-x
Wang, L., Yang, X., Cui, S., Mu, G., Sun, X., Liu, L., & Li, Z. (2019). QTL mapping and QTL × environment interaction analysis of multi-seed pod in cultivated peanut (Arachis hypogaea L.). *The Crop Journal*, *7*(2), 249–260. https://doi.org/10.1016/j.cj.2018.11.007
Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., Maccaferri, M., Salvi, S., Milner, S. G., Cattivelli, L., Mastrangelo, A. M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., International Wheat Genome Sequencing Consortium, Lillemo, M., Mather, D., … Akhunov, E. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnology Journal*, 12(6), 787–796. https://doi.org/10.1111/pbi.12183
Wang, S., Xu, S., Chao, S., Sun, Q., Liu, S., & Xia, G. (2019). A Genome-Wide Association Study of Highly Heritable Agronomic Traits in Durum Wheat. *Frontiers in Plant Science*, 10, 919. https://doi.org/10.3389/fpls.2019.00919
Wang, W., Pan, Q., Tian, B., He, F., Chen, Y., Bai, G., Akhunova, A., Trick, H. N., & Akhunov, E. (2019). Gene editing of the wheat homologs of TONNEAU1-recruiting motif encoding gene affects grain shape and weight in wheat. *The Plant Journal*, 100(2), 251–264. https://doi.org/10.1111/tpj.14440

Wang, W., Simmonds, J., Pan, Q., Davidson, D., He, F., Battal, A., Akhunova, A., Trick, H. N., Uauy, C., & Akhunov, E. (2018). Gene editing and mutagenesis reveal inter-cultivar differences and additivity in the contribution of TaGW2 homoeologues to grain size and weight in wheat. *Theoretical and Applied Genetics*, 131(11), 2463–2475. https://doi.org/10.1007/s00122-018-3166-7

Wicker, T. (2018). Impact of transposable elements on genome structure and evolution in bread wheat. *Genome Biology*, 19. https://doi.org/10.1186/s13059-018-1479-0

Wiebe, K. (2010). Targeted mapping of Cdu1, a major locus regulating grain cadmium concentration in durum wheat (*Triticum turgidum* L. var *durum*). *Theoretical and Applied Genetics*, 121. https://doi.org/10.1007/s00122-010-1370-1

Wilhelm, E. P., Turner, A. S., & Laurie, D. A. (2009). Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (*Triticum durum* Desf.). *Theoretical and Applied Genetics*, 118(2), 285–294. https://doi.org/10.1007/s00122-008-0898-9

Willems, E. (2017). *Durum wheat market: An EU perspective*. www.italmopa.com/wp-content/uploads/2017/05/144_all_2.pdf

Xin, F., Zhu, T., Wei, S., Han, Y., Zhao, Y., Zhang, D., Ma, L., & Ding, Q. (2020). QTL Mapping of Kernel Traits and Validation of a Major QTL for Kernel Length-Width Ratio Using SNP and Bulked Segregant Analysis in Wheat. *Scientific Reports*, 10(1), 25. https://doi.org/10.1038/s41598-019-56979-7

Xynias, I. N., Mylonas, I., Korpetis, E. G., Ninou, E., Tsaballa, A., Avdikos, I. D., & Mavromatis, A. G. (2020). Durum Wheat Breeding in the Mediterranean Region: Current Status and Future Prospects. *Agronomy*, 10(3), 432. https://doi.org/10.3390/agronomy10030432

Yamamuro, C., Ihara, Y., Wu, X., Noguchi, T., Fujioka, S., Takatsuto, S., Ashikari, M., Kitano, H., & Matsuoka, M. (2000). Loss of Function of a Rice brassinosteroid insensitive1 Homolog Prevents Internode Elongation and Bending of the Lamina Joint. *The Plant Cell*, 12(9), 1591–1606.

Yan, J. (2016). A loss-of-function allele of OsHMA3 associated with high cadmium accumulation in shoots and grain of Japonica rice cultivars. *Plant Cell Environ.*, *39*. https://doi.org/10.1111/pce.12747

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S., & Dubcovsky, J. (2006). The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proceedings of the National Academy of Sciences- USA*, 103(51), 19581–19586. https://doi.org/10.1073/pnas.0607142103

Yan, X., Zhao, L., Ren, Y., Dong, Z., Cui, D., & Chen, F. (2019). Genome-wide association study revealed that the TaGW8 gene was associated with kernel size in Chinese bread wheat. *Scientific Reports*, *9*(1), 2702. https://doi.org/10.1038/s41598-019-38570-2

Zaharieva, M., Ayana, N. G., Hakimi, A. A., Misra, S. C., & Monneveux, P. (2010). Cultivated emmer wheat (*Triticum dicoccon* Schrank), an old crop with promising future: A review. *Genetic Resources and Crop Evolution*, 57(6), 937–962. https://doi.org/10.1007/s10722-010-9572-6

Zhai, H., Feng, Z., Du, X., Song, Y., Liu, X., Qi, Z., Song, L., Li, J., Li, L., Peng, H., Hu, Z., Yao, Y., Xin, M., Xiao, S., Sun, Q., & Ni, Z. (2018). A novel allele of TaGW2-A1 is located in a finely mapped QTL that increases grain weight but decreases grain number in wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 131(3), 539–553. https://doi.org/10.1007/s00122-017-3017-y

Zhang, L., Zhao, Y.-L., Gao, L.-F., Zhao, G.-Y., Zhou, R.-H., Zhang, B.-S., & Jia, J.-Z. (2012). TaCKX6-D1, the ortholog of rice OsCKX2, is associated with grain weight in hexaploid wheat. *The New Phytologist*, 195(3), 574–584. https://doi.org/10.1111/j.1469-8137.2012.04194.x

Zhang, L.-Y., Liu, D.-C., Guo, X.-L., Yang, W.-L., Sun, J.-Z., Wang, D.-W., & Zhang, A. (2010). Genomic distribution of quantitative trait loci for yield and yield-related traits in common wheat. *Journal of Integrative Plant Biology*, 52(11), 996–1007. https://doi.org/10.1111/j.1744-7909.2010.00967.x

Zhang, P., He, Z., Tian, X., Gao, F., Xu, D., Liu, J., Wen, W., Fu, L., Li, G., Sui, X., Xia, X., Wang, C., & Cao, S. (2017). Cloning of TaTPP-6AL1 associated with grain weight in bread wheat and development of functional marker. *Molecular Breeding*, 37(6), 78. https://doi.org/10.1007/s11032-017-0676-y

Zhang, W. (2017). Identification and characterization of Sr13, a tetraploid wheat gene that confers resistance to the Ug99 stem rust race group. *Proceedings of the National Academy of Sciences- USA*, 114. https://doi.org/10.1073/pnas.1706277114

Zhang, Y., Li, T., Geng, Y., Wang, Y., Liu, Y., Li, H., Hao, C., Wang, H., Shang, X., & Zhang, X. (2021). Identification and development of a KASP functional marker of TaTAP46-5A associated with kernel weight in wheat (*Triticum aestivum*). *Plant Breeding*. https://doi.org/10.1111/pbr.12922

Zhao, D.-S., Li, Q.-F., Zhang, C.-Q., Zhang, C., Yang, Q.-Q., Pan, L.-X., Ren, X.-Y., Lu, J., Gu, M.-H., & Liu, Q.-Q. (2018). GS9 acts as a transcriptional activator to regulate rice grain shape and appearance quality. *Nature Communications*, 9(1), 1240. https://doi.org/10.1038/s41467-018-03616-y

Zhao, Y. (2010). Auxin biosynthesis and its role in plant development. *Annual Review of Plant Biology*, 61, 49–64. https://doi.org/10.1146/annurev-arplant-042809-112308

Zhu, X.-F., Zhang, H.-P., Hu, M.-J., Wu, Z.-Y., Jiang, H., Cao, J.-J., Xia, X.-C., Ma, C.-X., & Chang, C. (2016). Cloning and characterization of Tabas1-B1 gene associated with flag leaf chlorophyll content and thousand-grain weight and development of a gene-specific marker in wheat. *Molecular Breeding*, 36(10), 142. https://doi.org/10.1007/s11032-016-0563-y

8. Appendices

Appendix A. Analysis of variance per environment for the 8 traits considered in the MG5323 x Latino recombinant inbred line (RIL) mapping population.

Trait	Environment	Source	DF	MS	F ratio	Pr>F
A	V13	line	109	3.58	5.8	<.0001
		rep	1	1.172	1.91	0.17
	B14	line	109	6.01	13.1	<.0001
		rep	1	0.214	0.47	0.496
	F15	line	109	7.69	17.2	<.0001
		rep	2	32.770	73.41	<.0001
	F20	line	102	5.93	34.0	<.0001
		rep	2	0.082	0.47	0.63
L	V13	line	109	0.37	9.8	<.0001
		rep	1	0.030	0.78	0.38
	B14	line	109	0.46	29.8	<.0001
		rep	1	0.010	0.65	0.42
	F15	line	109	0.70	31.0	<.0001
		rep	2	0.900	39.85	<.0001
	F20	line	102	0.63	99.8	<.0001
		rep	2	0.004	0.63	0.53
W	V13	line	109	0.05	5.8	<.0001
		rep	1	0.008	0.86	0.35
	B14	line	109	0.09	12.1	<.0001
		rep	1	0.017	2.23	0.14
	F15	line	109	0.09	12.6	<.0001
		rep	2	0.357	50.96	<.0001
	F20	line	102	0.09	43.6	<.0001
		rep	2	0.001	0.61	0.54

Appendix A. Analysis of variance per environment -Continued.

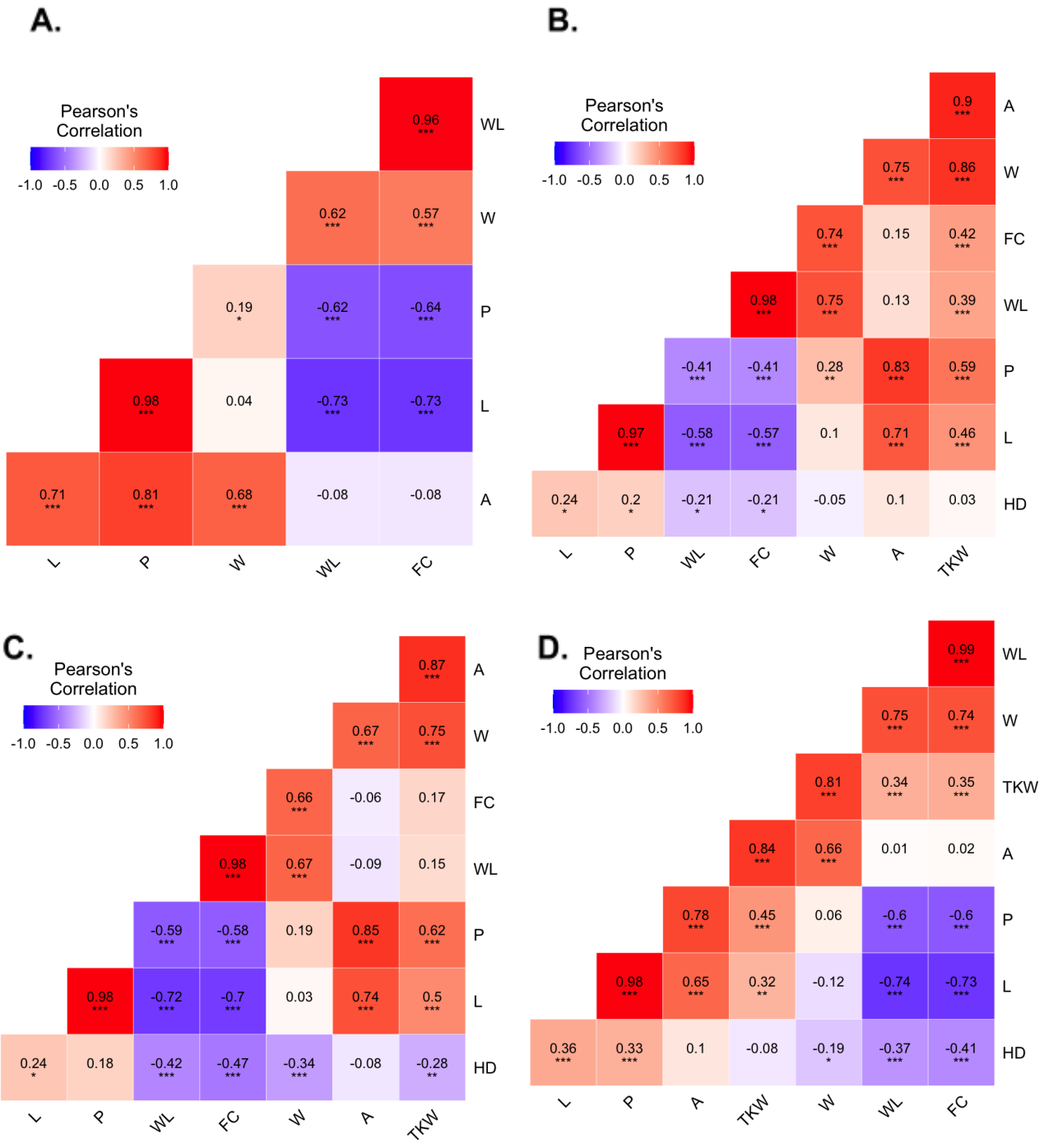
Trait	Environment	Source	DF	MS	F ratio	Pr>F
WL	V13	line	109	1.73E-03	17.9	<.0001
		rep	1	2.20E-06	0.02	0.88
	B14	line	109	2.21E-03	23.5	<.0001
		rep	1	4.25E-04	4.51	0.04
	F15	line	109	2.83E-03	21.8	<.0001
		rep	2	7.03E-04	5.41	0.005
	F20	line	102	3.18E-03	183.2	<.0001
		rep	2	2.75E-06	0.16	0.85
P	V13	line	109	1.50	7.4	<.0001
		rep	1	0.371	1.84	0.17
	B14	line	109	1.96	24.4	<.0001
		rep	1	0.054	0.67	0.41
	F15	line	109	2.84	26.3	<.0001
		rep	2	5.127	47.49	<.0001
	F20	line	102	2.55	64.6	<.0001
		rep	2	0.049	1.23	0.29
FC	V13	line	109	2.44E-03	16.0	<.0001
		rep	1	6.34E-05	0.41	0.52
	B14	line	109	2.89E-03	17.3	<.0001
		rep	1	9.01E-04	5.40	0.022
	F15	line	109	3.78E-03	20.7	<.0001
		rep	2	1.50E-03	8.24	0.0003
	F20	line	102	4.71E-03	92.8	<.0001
		rep	2	7.04E-05	1.39	0.25

Appendix A. Analysis of variance per environment -Continued.

Trait	Environment	Source	DF	MS	F ratio	Pr>F
TKW	B14	line	109	89.00	9.6	<.0001
		rep	1	2.870	0.31	0.58
	F15	line	109	107.00	10.3	<.0001
		rep	2	340.00	32.80	<.0001
	F20	line	102	82.70	24.6	<.0001
		rep	2	4.340	1.29	0.278
HD	B14	line	109	77.10	15.5	<.0001
		rep	1	36.800	7.40	0.0076
	F15	line	109	33.70	7.6	<.0001
		rep	2	30.300	6.80	0.0014
	F20	line	102	64.00	21.0	<.0001
		rep	2	29.100	9.57	0.0001

Environments are indicated as following : V13 for Valenzano 2012-2013, B14 for Bologna 2013-2014, F15 and F20 for Fiorenzuola d'Arda 2014-2015 and 2019-2020, respectively. DF refers to Degrees of Freedom and MS to Mean Squares. Note : In the case of B14, as two repetitions per genotype were provided, the analysis was performed on the complete dataset (100 seeds/repetition).

Appendix B. Pearson correlations coefficients (*r*) among the phenotypic traits analyzed using single environment data : A) Valenzano 2012-2013, B) Bologna 2013-2014, C) Fiorenzuola d'Arda 2014-2015 and D) Fiorenzuola d'Arda 2019-2020.



Statistical significance is denoted as *** $p < 0.001$, ** $p < 0.01$ and * $p < 0.05$.

Appendix C. Genes associated with kernel-related traits identified in bread wheat (*T. aestivum*) and/or rice (*O. sativa*).

Bread Wheat / Rice* gene name (GenBank Acc.)	Wheat Chr.	Start-End on Svevo genome (bp)	ID of the Svevo gene	Identity (%)	References
SRS3 (Os05g0154700)*	1A	88748570-88747063	TRITD1Av1G038700	80.67	Kitagawa et al., 2010
	1B	140384039-140385548	TRITD1Bv1G051340	80.67	
GW5 (Os05g0187500)*	1A	138091027-138092490	TRITD1Av1G057710	76	Shomura et al., 2008; Weng et al., 2008
	1B	181527596-181529089	TRITD1Bv1G065410	76	
TaSnRK2.3-1A (KJ018721.1)	1A	375178508-375178586	TRITD1Av1G139010	87.34	Miao et al. 2017
GIF2/AGPL2 (Os01g0633100)*	1A	568600653-568600462	TRITD1Av1G222050	80.73	Wei et al., 2017
TaDA1 (KM005099.1)	2A	6821815-6828178	TRITD2Av1G003900	99.53	Liu et al., 2019
TaSus2 (AJ000153)	2A	120336304-120335983	TRITD2Av1G053920	97.83	Jiang et al., 2011; Hou et al., 2014
	2B	169017304-169016983	TRITD2Bv1G065410	99.07	
SRS1 (Os07g0616000)*	2A	121243934-121241806	TRITD2Av1G054260	77.21	Abe et al., 2010
	2B	169982961-169980629	TRITD2Bv1G065910	76	
TaGW7	2A	134684064-134685727	TRITD2Av1G059560	80.46	Wang et al., 2019
	2B	180008653-180006995	TRITD2Bv1G069960	80.65	
D11 (Os04g0469800)*	2A	560602960-560604961	TRITD2Av1G202210	77	Tanabe et al., 2005
	2B	489842099-489844658	TRITD2Bv1G165560	77	
GS2/GLW2/GRF4 (Os02g0701300)*	2A	681757822-681758011	TRITD2Av1G251570	90.53	Sun et al., 2017; Li et al., 2017
TaFlo2	2A	737202308-737200621	TRITD2Av1G276740	78.19	Sajjad et al., 2017; She et al., 2010
	2B	728958324-728956637	TRITD2Bv1G240770	78.37	
GLW7/OsSPL13 (Os07g0505200)*	2B	260530937-260530704	TRITD2Bv1G096640	81.86	Si et al., 2016
Tabas1-B1 (AB000405.1)	2B	441278486-441278613	TRITD2Bv1G148880	97.66	Zhu et al. 2016
D2 (Os01g0197100)*	3A	62643468-62642711	TRITD3Av1G031480	74	Hong et al., 2003
TaCKX6-D1 (JQ797673.1)	3A	107426071-107424567	TRITD3Av1G048340	86.76	Zhang et al., 2012
	3B	164239367-164240706	TRITD3Bv1G059800	91.52	
TaGS5 (KX219726.1)	3A	182251443-182253720	TRITD3Av1G076080	99.82	Wang et al., 2016
	3B	231034675-231032624	TRITD3Bv1G081720	93.9	
D61 (Os01g0718300)*	3A	462031527-462027681	TRITD3Av1G163790	81	Yamamuro et al., 2000
	3B	447492899-447496733	TRITD3Bv1G143630	82	
TaSnRK2.4 (GQ384359.1)	3A	622924591-622925119	TRITD3Av1G227410	99.62	Miao et al., 2021
OsLG3/ERF62 (Os03g0183000)*	4A	768032-767789	TRITD4Av1G000320	85.25	Yu et al., 2017
TaSDIR1-4A (MK419003.1)	4A	102455621-102455878	TRITD4Av1G043310	100	Wang et al., 2020
TaCWI-4A (AF030420.1)	4A	603888870-603889886	TRITD4Av1G209690	96.7	Jiang et al., 2015
TaTGW6 (KT582298.1)	4A	666395284-666396939	TRITD4Av1G238170	99.21	Hu et al., 2016

Appendix C. Genes associated with kernel-related traits identified in bread wheat / rice - Continued.

Bread Wheat / Rice* gene name (GenBank Acc.)	Wheat Chr.	Start-End on Svevo genome (bp)	ID of the Svevo gene	Identity (%)	References
6-SFT-A2 (FJ228688.1)	4A	730723774-730726913	TRITD4Av1G261410	89.97	Yue et al., 2015
SRS5 (Os11g0247300)*	4A	563051888-563050402	TRITD4Av1G188370	85.91	Segami et al., 2017
	4B	45209595-45211077	TRITD4Bv1G018150	85.88	
TaD14-4D (Os03g0203200*)	4B	528353376-528351851	TRITD4Bv1G153980	9221	Liu et al., 2021
Big Grain 1 (BG1) (Os - Q10R09.1*)	4B	582040779-582041714	TRITD4Bv1G171270	81.3	Milner et al., 2021
TaSnRK2.9-5A (MH844552.1)	5A	80346088-80346379	TRITD5Av1G035270	100	Rehman et al., 2019
TaTAP46-5A (EF101900.1)	5A	444617166-444621885	TRITD5Av1G161680	81.44	Zhang et al. 2021
GL3A (KY865328.1)	5A	533654548-533659331	TRITD5Av1G198570	99.87	Yang et al., 2019
qGL3/OsPPKL1 (Os03t0646900)*	5B	550457327-550458383	TRITD5Bv1G192150	81.41	Zhang et al., 2012
TaCWI-5B (AF030420.1)	5B	693942186-693939429	TRITD5Bv1G250300	99.5	Jiang et al., 2015
TaGW2-A1, TaGW2-6A (KF176551.1)	6A	235270703-235295248	TRITD6Av1G091060	99	Zhai et al., 2018; Simmonds et al., 2016; Hong et al., 2014
	6B	300808374-300791561	TRITD6Bv1G096950	99	
FUWA (Os02g0234200)*	6A	270173147-270190995	TRITD6Av1G100490	82	Chen et al., 2015
	6B	373411421-373410761	TRITD6Bv1G115800	82.6	
TaTPP-6AL1 (Os - AB120515.1*)	6A	458111300-458111522	TRITD6Av1G158120	85.2	Zhang et al. 2017
TaGS1b (DQ124210)	6A	525603850-525603539	TRITD6Av1G185050	93.84	Bernard et al., 2008
	6B	553792479-553792174	TRITD6Bv1G174320	99.79	
D1 (Os05g0333200)*	7A	1598058-1597571	TRITD7Av1G000840	73	Ashikari et al., 1999
6-SFT-A2 (FJ228688.1)	7A	2804757-2807893	TRITD7Av1G001560	91.97	Yue et al., 2015
TaGS D1 (KF687956.1)	7A	6467996-6467368	TRITD7Av1G003700	91.73	Zhang et al., 2014
TaSus1 (AJ001117.1)	7A	111462985-111462663	TRITD7Av1G050690	98.15	Jiang et al., 2011; Mohler et al., 2016
	7B	68819710-68820032	TRITD7Bv1G024970	98.77	
TaGASR7-A1 (KJ000052.1)	7A	168476784-168476262	TRITD7Av1G071860	100	Dong et al. 2014
TaTGW-7A (KT582299.1)	7A	204055853-204061744	TRITD7Av1G085470	97.58	Hu et al., 2016
TaGW8 (MK388407.1)	7A	249768275-249768729	TRITD7Av1G103100	96.27	Yan et al., 2019
	7B	223504720-223509529	TRITD7Bv1G077140	99.46	
TaCYP78A3 (KP768392)	7A	276836292-276835153	TRITD7Av1G113010	94.38	Ma et al., 2015
TEF-7A (CJ655632)	7A	558165268-558165407	TRITD7Av1G207620	88.57	Zheng et al., 2014
TaSAP-A1 (KC193579)	7A	581186556-581185767	TRITD7Av1G217610.1	96.84	Chang et al., 2013
	7B	528343886-528343099	TRITD7Bv1G166760	97.21	

(*) Genes identified in *O. sativa.* Chr. refers to the wheat chromosome. Physical positions (bp = base pairs) on the *T. durum* cv. Svevo genome were retrieved from BLASTn (Intranet of Durum Wheat Genome Data ,https://www.interomics.eu/durum-wheat-genome-intranet). Genes included in the QTL physical intervals found in this study are shown in bold.