




Article

QTL Analysis for Bread Wheat Seed Size, Shape and Color Characteristics Estimated by Digital Image Processing

Mian Abdur Rehman Arif ^{1,†}, Evgenii G. Komyshev ^{2,3,4,†}, Mikhail A. Genaev ^{2,3,4}, Vasily S. Koval ^{2,4}, Nikolay A. Shmakov ^{2,3,4}, Andreas Börner ^{5,*} and Dmitry A. Afonnikov ^{2,3,4,*}

¹ Nuclear Institute for Agriculture and Biology, Faisalabad 38000, Pakistan

² Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

³ Faculty of Natural Sciences, Novosibirsk State University, 630090 Novosibirsk, Russia

⁴ Kurchatov Genomics Center, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

⁵ Leibniz Institute of Plant Genetics and Crop Plant Research, 06466 Seeland, Germany

* Correspondence: boerner@ipk-gatersleben.de (A.B.); ada@bionet.nsc.ru (D.A.A.); Tel.: +49-394825229 (A.B.); +7-(383)-363-49-63 (D.A.A.)

† These authors contributed equally to this work.

Abstract: The size, shape, and color of wheat seeds are important traits that are associated with yield and flour quality (size, shape), nutritional value, and pre-harvest sprouting (coat color). These traits are under multigenic control, and to dissect their molecular and genetic basis, quantitative trait loci (QTL) analysis is used. We evaluated 114 recombinant inbred lines (RILs) in a bi-parental RIL mapping population (the International Triticeae Mapping Initiative, ITMI/MP) grown in 2014 season. We used digital image analysis for seed phenotyping and obtained data for seven traits describing seed size and shape and 48 traits of seed coat color. We identified 212 additive and 34 pairs of epistatic QTLs on all the chromosomes of wheat genome except chromosomes 1A and 5D. Many QTLs were overlapping. We demonstrated that the overlap between QTL regions was low for seed size/shape traits and high for coat color traits. Using the literature and KEGG data, we identified sets of genes in *Arabidopsis* and rice from the networks controlling seed size and color. Further, we identified 29 and 14 candidate genes for seed size-related loci and for loci associated with seed coat color, respectively.

Keywords: wheat; seed size; seed shape; seed coat color; phenotyping; candidate genes; QTLs



Citation: Arif, M.A.R.; Komyshev, E.G.; Genaev, M.A.; Koval, V.S.; Shmakov, N.A.; Börner, A.; Afonnikov, D.A. QTL Analysis for Bread Wheat Seed Size, Shape and Color Characteristics Estimated by Digital Image Processing. *Plants* **2022**, *11*, 2105. <https://doi.org/10.3390/plants11162105>

Academic Editor: Abdelmajid Kassem

Received: 22 July 2022

Accepted: 10 August 2022

Published: 12 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bread wheat (*Triticum aestivum* L.) is a major staple crop. Millions of people depend on its production (<https://www.fao.org/faostat/en/#data>, accessed on 20 January 2022). This has led to an ongoing search for and study of genes that control wheat yield traits. Some of them are the characteristics of wheat seeds (size and shape) which have been shown to be related to seed weight [1–4]. Seed size and shape have also been shown to be related to flour quality and composition: small kernels can contribute to enhancing the bread-making quality of flour while having a detrimental effect on the milling yield [5]. To find genes that control these traits in wheat, QTL analysis is used. This analysis makes it possible to identify sets of markers that are associated with seed size or shape traits. Studies have shown that seed size and shape in wheat are controlled by a large number of loci located on almost all chromosomes [6–13]. Identification of these loci combined with molecular analysis can identify genes that are involved in controlling seed weight or size [14–19]. Based on genetic and molecular studies in both the model organism *Arabidopsis thaliana* and cereals, it is now established that seed weight is affected by multiple molecular and genetic aspects that lead to dynamic changes in cell division, expansion, and differentiation during seed development. Several important biological pathways contribute to seed weight, such

as ubiquitination, phytohormones, G-proteins, photosynthesis, epigenetic modifications, and microRNAs [20,21]. Knowledge of the pathways controlling seed development in well-studied organisms allows the prioritization of candidate genes controlling these traits in wheat as well for their further study by molecular methods [15].

Another important characteristic of wheat seeds is the color of the shell. It characterizes the pigments and metabolites it contains. Purple and blue coloring of seeds is determined by the presence of anthocyanins. A yellowish color may be due to the presence of carotenoids. A reddish brown or dark brown coloration of the seeds is due to the presence of flavonoids such as proanthocyanidins and phlobaphenes [22]. Genetic control of color formation in both seeds and other plant organs is carried out by genes encoding the enzymes involved in pigment biosynthesis as well as regulatory genes [23]. For a number of pigments, these genes have been well studied; however, for some pigments, the molecular mechanisms of biosynthesis are still poorly understood [24].

The presence of pigments in the seed coat affects various technological properties of the seed [25] and is associated with antioxidant properties [26]. Therefore, varieties and lines with diverse seed coloration are of active interest in the food industry [27,28]. Seed shell color in wheat is also associated with important characteristics such as germination ability and pre-harvest sprouting (PHS). Red seeds are less susceptible to PHS [29]. QTL searches for seed color and PHS resistance are often simultaneously performed [30,31].

Recently, genotyping technologies have made great progress and include diversity array technology, genotyping-by-sequencing [32], and SNPs [33,34]. High-throughput genotyping can achieve high-density marker mapping [35]. This allows more QTLs to be obtained and, as a result, more accurately establish the molecular mechanisms controlling important plant phenotypic traits [8,32,33,36,37].

In the present work, we performed a SNP-based QTL search for seven traits of seed size/shape and 48 traits of shell color evaluated on the basis of digital image analysis on a set of 114 recombinant inbred lines (RILs) of the “International Triticeae Mapping Initiative” mapping population (ITMI/MP) and their parental plants.

2. Results

2.1. Analysis of the Seed Traits in ITMI Population

Figure 1 shows the distribution of six of the fifty-five seed characteristics in the ITMI population. Three of them characterize size (sL, sW, sA), and three characterize color (Lab_mL, Lab_ma, Lab_mb). The distributions were bell-shaped, and the hypothesis of normality was not rejected for the characteristics of seed length and area (Shapiro–Wilks test, $p < 0.05$), but for the width and color characteristics (Figure 1). Overall, the hypothesis of normality was not rejected based on this test for 22 of the 55 traits.

In order to visualize the distribution of genotypes in the space of the considered traits, we performed principal component analysis for the traits of shape/size (all seven traits) (Figure 2), color (12 traits of average values of color components in four-color spaces) of seeds independently (Figure 3) and all these 19 traits simultaneously (Figure S1). From the diagram of the principal components in the size/shape feature space, we see that the first component characterizes the roundness of the seeds and is correlated with circularity. The second component characterizes seed size (most related to width and area). The most characteristic genotypes are ITMI_082 (the most rounded seeds), ITMI_075 (large area), ITMI_048 (small area), and ITMI_111 and Synthetic_W7984 (most elongated). The second parental genotype is located in this diagram on the far-right side of the diagram, close to the X-axis, i.e., it has a rounded seed shape. It is difficult to distinguish any noticeable clusters in this diagram: it is a cloud with some distant genotypes. Notably, this cloud is more sparse in the upper half-plane ($PC2 > 0$) and compact in the lower half-plane ($PC2 < 0$).

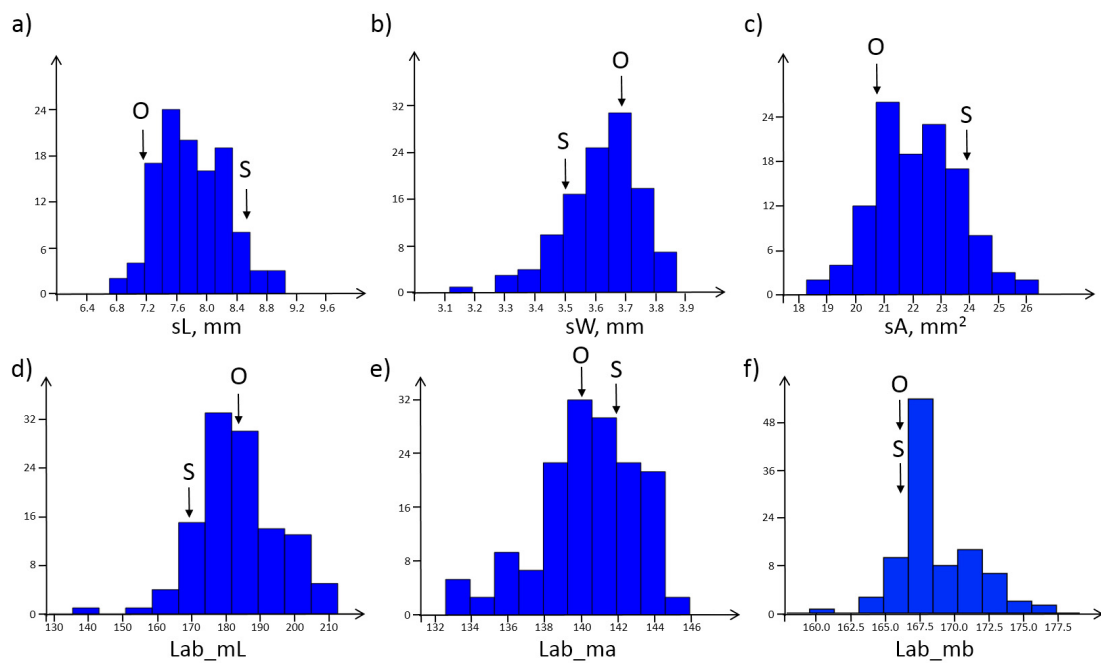


Figure 1. Distribution of six characteristics of seed size and color. The X-axis is the value of the characteristic, and the Y-axis is the frequency in the sample. (a) seed length, sL; (b) seed width, sW; (c) area of seed projection in the image, sA; (d) intensity of component L of Lab color space (lightness); (e) intensity of component a of Lab color space (redness); (f) intensity of component b of Lab color space (yellowness). The arrows show the characteristic values for the parental genotypes Opata (O) and Synthetic (S).

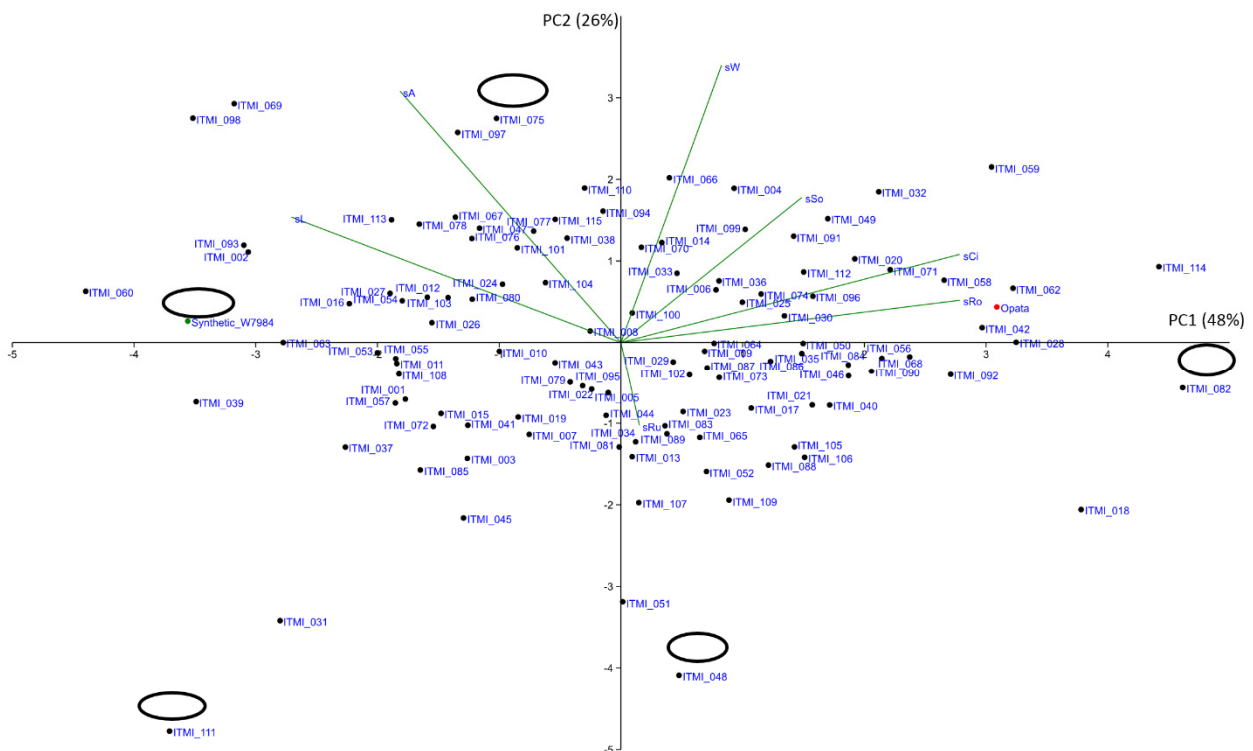


Figure 2. PCA biplot of seed size and shape of ITMI/MP performed using seed 7 coat traits. Ellipses represent seed size and shape for some contrast genotypes in the same scale. Parent genotypes are shown by green (Synthetic_W7984) and red (Opata) dots. PC1, PC2 axes denote principal components 1 and 2, percentage of variance explained shown in parentheses.

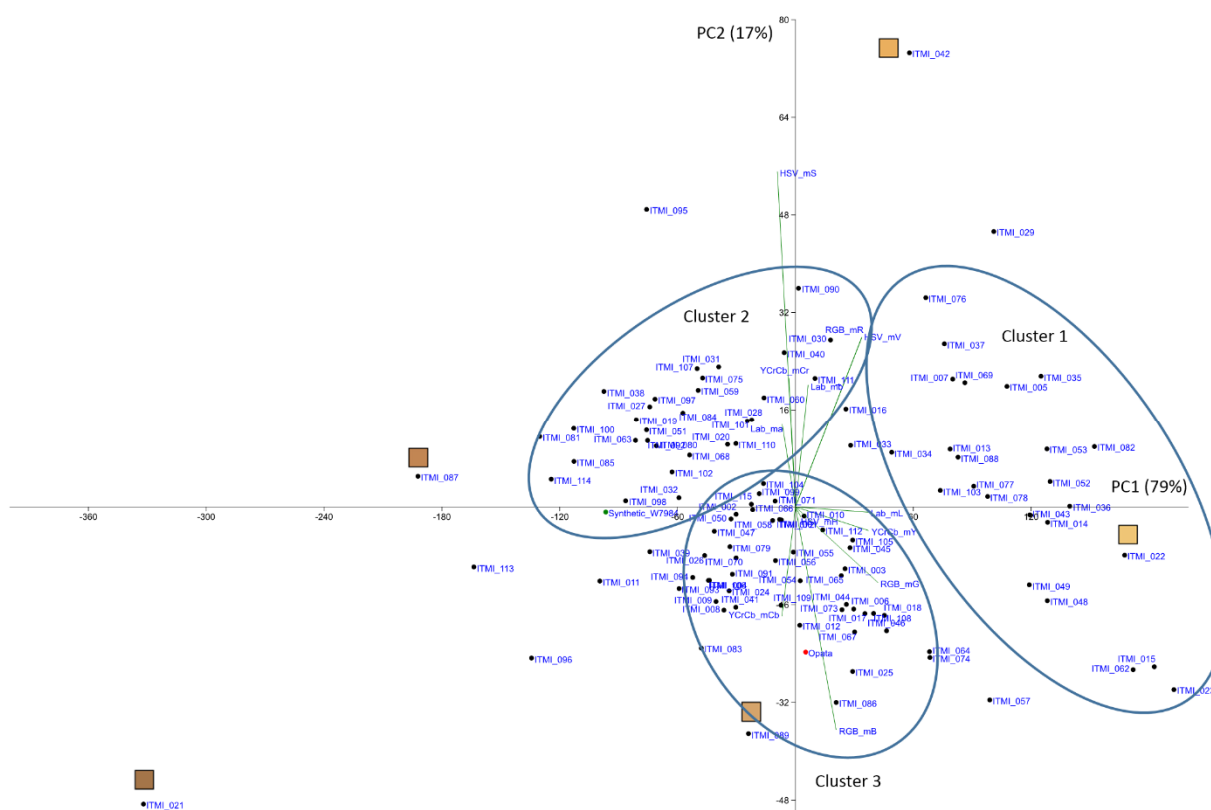


Figure 3. PCA biplot of seed coat color of ITMI/MP performed using seed coat traits (mean values for 12 color components of four-color spaces). Color bars represent coat color for some contrast genotypes. Parent genotypes shown by green (Synthetic_W7984) and red (Opatá) dots. PC1 and PC2 axes denote principal components 1 and 2. Explained percentage of variance shown in parentheses. Three clusters of genotypes shown by ellipses.

The diagram of the principal components in the color feature space shows that the first component primarily characterizes the lightness of the shell (correlates with Lab_mL and YCrCb_mY). The second component characterizes seed color saturation and reddish shade (positively correlates with HSV_ms, RGB_mr, and Lab_ma characteristics). The most characteristic genotypes are ITMI_2 (the lightest shell), ITMI_042 (the most saturated color), ITMI_088 (the palest shell hue), and ITMI_021 and ITMI_087 (the darkest shell hue). Three clusters can be distinguished on the plot (Figure 3). Seeds from plants of the first cluster have a lighter color (large values of PC1), with a large part of them having a more reddish color (positive values of PC2). Seeds from plants of the second cluster have darker (negative values of PC1) and more reddish color (positive values of PC2). Seeds from plants of the second cluster have less reddish color (positive values of PC2 and PC1 values dispersed about 0 value). Interestingly, parent genotypes Synthetic_W7984 and Opatá fall into distinct clusters on the plot (Cluster 2 and 3, respectively).

The diagram of the principal components in the seven size/shape and 12 color feature space (Figure S1) shows that the first component characterizes the color of the seed shell: the negative values are characteristics of reddish color (PC1 positively correlates with Lab_mb and negatively correlates with Lab_ma). The second component characterizes seed size/shape (positively correlated with roundness, sRo, and circularity, sCi and negatively correlates with area, sA, and length, sL). No clear clusters were detected in this plot for genotypes.

2.2. QTL Analysis

Genetic analysis of the three characteristics of seed size (sL, sW, and sA), four characteristics of shape (sCi, sRo, sRu, and sSo), and 48 characteristics of color (12 characteristics each of RGB, HSV, L*a*b, and YCrCb) revealed a total of 20, 22 and 170 QTLs (212 in total) (Figure 4, Table S1), correspondingly, on all the chromosomes of wheat genome except chromosomes 1A and 5D. The number of QTLs varied from one (characteristics: HSV_mS, HSV_dCH_2, HSV_dCS_2, and HSV_dCV_2) to ten (characteristic: sA) for one single trait. The majority of the traits yielded three (13 characters) to four QTLs (19 characters).

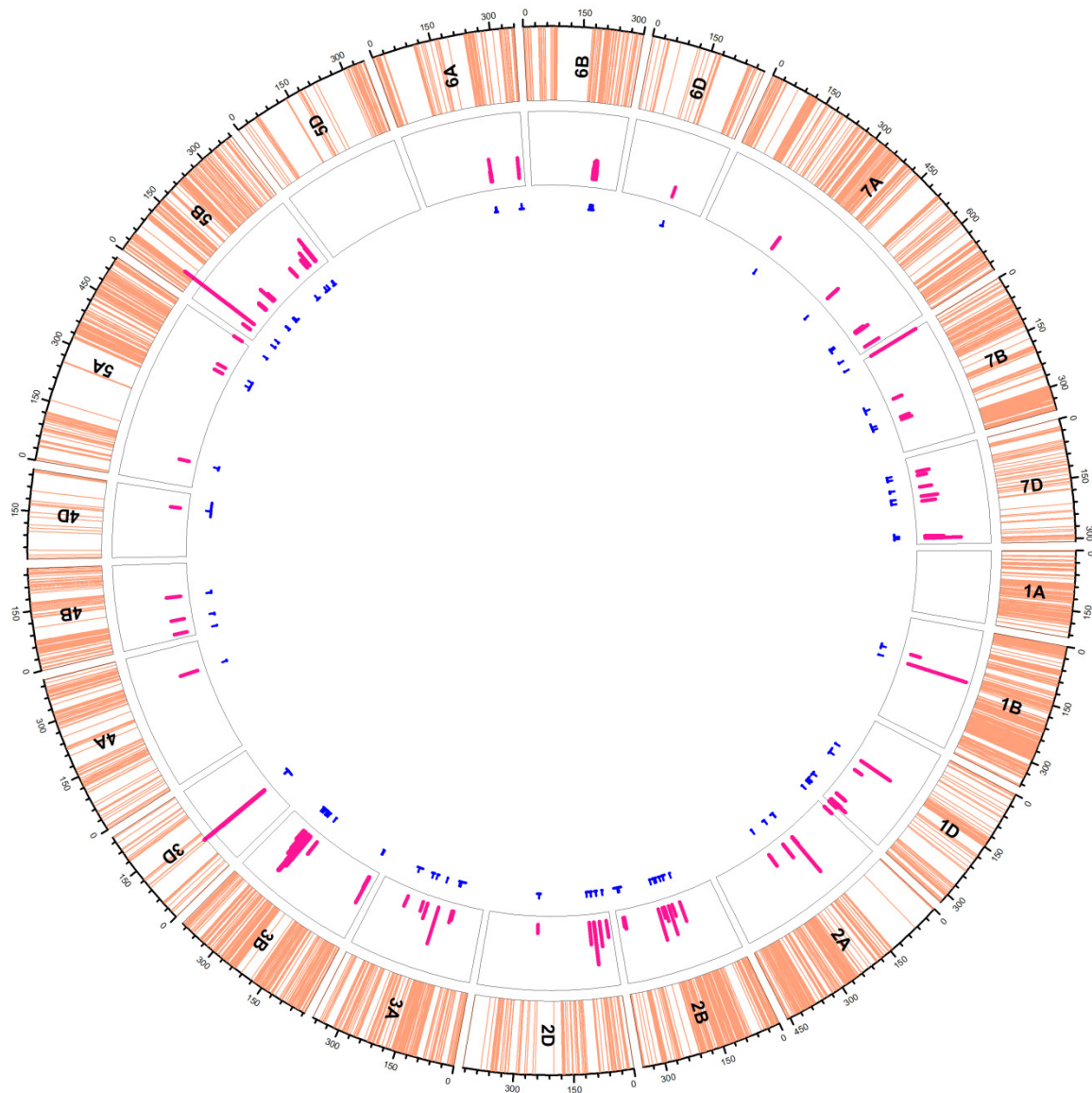


Figure 4. Distribution of additive QTLs (blue lines in the inner circle). Light orange lines in the outer track indicate the SNP positions on each chromosome. Pink bars in the second circle indicate the LOD values of QTLs. The blue lines under the track circle indicate the confidence interval of QTLs with small vertical lines pointing to the peak position of QTL. For details, see Table S1.

Among chromosomes, the highest number of QTLs was observed on chromosome 3B (46 QTLs), followed by chromosomes 3D and 6B with 34 and 27 QTLs, correspondingly. Chromosome 5B carried 15 QTLs, and the chromosome 2B carried 14 QTLs, whereas chromosome 7A carried 12 QTLs. This was followed by chromosome 1D with 11 QTLs. Chromosomes 3A and 7D carried nine QTLs each, and chromosome 6A carried seven QTLs. Chromosomes 6A and 2D carried seven and six QTLs, respectively. Five QTLs resided on

each of chromosomes 2A and 4B, whereas four QTLs resided on each of chromosomes 5A and 7B. On the other hand, two QTLs were detected on each of chromosomes 1B and 6D. Finally, chromosomes 4A and 4D carried one QTL each. In terms of groups, group 3 chromosomes carried the highest number of QTLs (89), whereas group 4 chromosomes carried the least number of QTLs (five). Group 6 chromosomes carried 36 QTLs, whereas each of group 2 and 7 chromosomes carried 25 QTLs. On the other hand, group 5 chromosomes carried 19 QTLs, and group 1 chromosomes carried 13 QTLs.

Additionally, we were able to detect 34 pairs of epistatic QTLs controlling at least 22 characters in our RILs, with five characters under the influences of more than one pair of epistatic QTLs (Figure 5, Table S2). These QTLs involved all the wheat chromosomes except chromosomes 1A, 4A, 4B, and 6A. The most frequently involved chromosome was 3D (12 times), followed by chromosomes 3A (11 times), and 3B (nine times). Chromosome 2D was involved six times, whereas chromosomes 5B and 5D were involved four times each. Chromosomes 1D, 2B, and 7A were involved three times each. Two times involvement was observed for chromosomes 1B, 2A, 4D, 6B, and 6D, whereas the chromosomes 5A, 7B, and 7D were only involved one time.

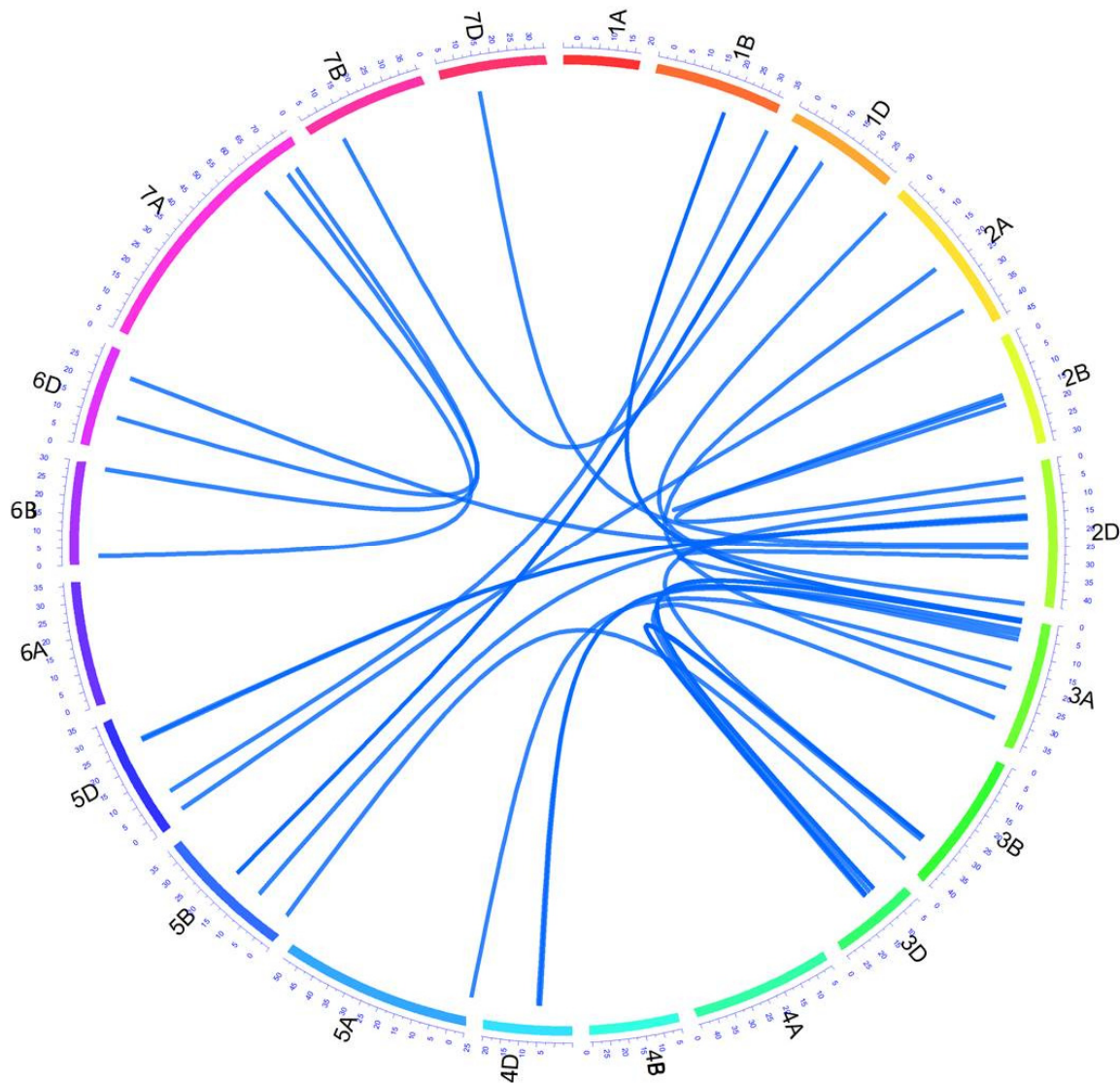


Figure 5. Epistasis QTL network in the ITMI/MP. Outer circular plot represents the hexaploid genome arranged in chromosomes (chrs) 1–21 (1A–7D) in clockwise direction. Numbers on colored outer circle represents cM on respective chrs. Blue-colored connections represent epistasis QTLs controlling different traits. For details, see Table S2.

2.3. Analysis of the Similarity of Traits by QTL Location

We observed remarkable overlap between QTL locations for different traits. For example, two QTLs related to seed shape, *Q.sCi-2B^c* (circularity) and *Q.sSo-2B^c* (solidity), were located in the same position 129 of chromosome 2B. The 3B chromosome has loci with multiple traits associations: position 39.179 (two traits of size), position 298.179 (seven color traits), position 299.179 (twelve color traits), position 300.179 (two color traits), position 306.179 (ten color traits), position 308.179 (two color traits), position 311.179 (two color traits and one shape trait, rugosity), and position 324.179 (three color traits). This is not surprising because our parameters estimated from images represent various quantifications of the same biological seed property (i.e., seed weight, pigment concentrations in the coat, etc.). This suggests that the set of our characteristics is degenerating and that many of them, in fact, are controlled by the same genes.

To evaluate the similarity of various traits under analysis, we hierarchically clustered them by the degree of the overlap of QTL locations (Figure 6). The tree diagram demonstrates several interesting features. Firstly, size/shape characteristics (right part of the tree) are clearly separated from the color traits (with the exception of rugosity, sRu). Secondly, some traits with a small QTL number (one to two) are also separated from other traits. Thirdly, a remarkable number of traits related to yellowness form a large cluster. Finally, traits related to the seed lightness (Lab_mL, HSV_mV, and YCrCb_mY) fall in the same cluster, and their QTLs are highly overlapped. Other color traits are irregularly dispersed on the tree within the large cluster of color traits.

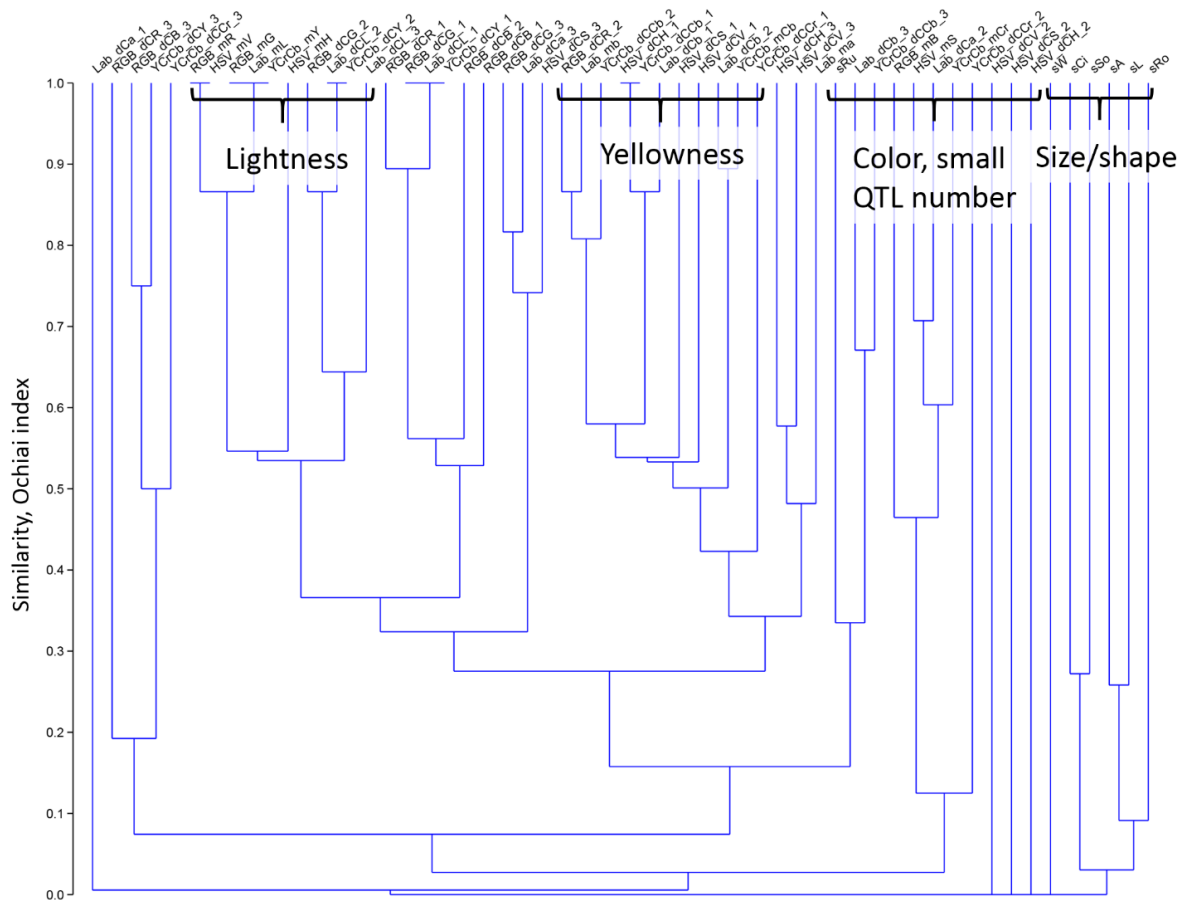


Figure 6. The similarity tree for seed traits obtained by the degree of the overlap between their QTL locations. The vertical axis represents the similarity measure based on the Ochiai index (Y-axis). Leaves correspond to seed traits described in [1] (for trait abbreviation, see Section 4.2. *Quantitative Characteristics of Seed Shape, Size and Color*). Groups of traits with strong overlapping of the QTL locations are shown by curly brackets.

2.4. Gene Prioritization

A search for orthologous groups for the eight pathways of pigment biosynthesis and their precursors identified 307 KEGG orthologs involved in these processes (Table S3). A review of the literature [18,19,38] identified 155 *Arabidopsis* and 42 rice genes involved in the molecular processes of seed development (Table S3). Of these genes, 193 were found to have sequence identifiers in the KEGG database, and 109 of them were associated with KEGG orthologous groups (Table S4).

For prioritization of genes, we used 48 highly significant QTLs with LOD > 3 for which marker positions were identified in the wheat genomic sequence (Table S5). On this basis, we identified 2787 unique genes localized to marker-limited sites. Of these, 1422 genes associated with seed size/shape, and 1365 genes associated with seed color. After filtering by expression level, 823 genes associated with seed color remained (Table S6). For these sets of genes, we performed KEGG orthogroup assignment using BlastKOALA and KofamKOALA services. For 464 genes associated with size trait loci and 321 genes associated with color traits, such orthogroups were found.

For 29 genes from the seed size-related loci, we found a match within the orthogroup list obtained based on the analysis of the literature data (Table 1). Eleven genes identified in this way belong to regulatory proteins (transcription factors EREBP, HD-ZIP, and MYBP; loci on chromosomes 3A, 2B, 2D, and 7D). Six genes belong to translation initiation factors (ELF2C; loci on chromosomes 2B and 7D). Five genes relate to enzymes associated with ubiquitination processes (loci on chromosomes 2D). Four genes have chitinase activity (locus on chromosome 7D), two genes with cytokin dehydrogenase activity (chromosomes 3A and 7D), and one aarF domain-containing kinase gene (chromosome 7D).

Table 1. List of candidate genes from QTLs associated with seed size/shape. Columns of the table contain QTL name (QTL), chromosome and position in cm (Chr/Pos), gene ID, KEGG orthogroup ID, KEGG orthogroup description, and EC number, if provided.

QTL	Chr/Pos	Gene ID	KO ID	Description	EC
Q.sA-3A	3A/155	<i>TraesCS3A03G0787100</i>	K09286	EREBP; EREBP-like factor	-
		<i>TraesCS3A03G0782400</i>	K09338	HD-ZIP; homeobox-leucine zipper protein	-
		<i>TraesCS3A03G0763900</i>	K00279	CKX; cytokinin dehydrogenase	EC:1.5.99.12
Q.sSo-4A	4A/305	<i>TraesCS4A03G1100100</i>	K19045	BB; E3 ubiquitin-protein ligase BIG BROTHER and related proteins	EC:2.3.2.27
Q.sCi-2B ^c ; sSo-2B ^c	2B/129	<i>TraesCS2B03G0313000</i>	K09286	EREBP; EREBP-like factor	-
		<i>TraesCS2B03G1115600</i>	K11593	ELF2C, AGO; eukaryotic translation initiation factor 2C	-
		<i>TraesCS2B03G1114800</i>	K11593	ELF2C, AGO; eukaryotic translation initiation factor 2C	-
		<i>TraesCS2B03G1105600</i>	K09422	MYBP; transcription factor MYB, plant	-
		<i>TraesCS2B03G1109700</i>	K09286	EREBP; EREBP-like factor	-
		<i>TraesCS2B03G1116700</i>	K11593	ELF2C, AGO; eukaryotic translation initiation factor 2C	-
		<i>TraesCS2B03G1104200</i>	K11593	ELF2C, AGO; eukaryotic translation initiation factor 2C	-
		<i>TraesCS2B03G1109900</i>	K09286	EREBP; EREBP-like factor	-
		<i>TraesCS2B03G1106500</i>	K09338	HD-ZIP; homeobox-leucine zipper protein	-
Q.sW-2D	2D/74	<i>TraesCS2D03G0133000</i>	K09422	MYBP; transcription factor MYB, plant	-
		<i>TraesCS2D03G0143000</i>	K09602	OTUB1; ubiquitin thioesterase protein OTUB1	EC:3.4.19.12
		<i>TraesCS2D03G0143400</i>	K09602	OTUB1; ubiquitin thioesterase protein OTUB1	EC:3.4.19.12

Table 1. Cont.

QTL	Chr/Pos	Gene ID	KO ID	Description	EC
Q.sCi-2D ^g ; Q.sRo-2D ^g	2D/58	<i>TraesCS2D03G0107400</i>	K09602	ubiquitin thioesterase protein OTUB1	EC:3.4.19.12
		<i>TraesCS2D03G0107900</i>	K09602	ubiquitin thioesterase protein OTUB1	EC:3.4.19.12
sSo-7D.2	7D/141	<i>TraesCS7D03G1008800</i>	K09286	EREBP; EREBP-like factor	-
		<i>TraesCS7D03G0987700</i>	K09338	HD-ZIP; homeobox-leucine zipper protein	-
		<i>TraesCS7D03G0983800</i>	K09286	EREBP; EREBP-like factor	-
		<i>TraesCS7D03G0972400</i>	K08869	ADCK, ABC1; aarF domain-containing kinase	-
		<i>TraesCS3A03G0763900</i>	K00279	CKX; cytokinin dehydrogenase	EC:1.5.99.12
sL-7D ^{bb} ; sRo-7D ^{bb}	7D/287	<i>TraesCS7D03G1260500</i>	K20547	CHIB; basic endochitinase B	EC:3.2.1.14
		<i>TraesCS7D03G1260300</i>	K20547	CHIB; basic endochitinase B	EC:3.2.1.14
		<i>TraesCS7D03G1260400</i>	K20547	CHIB; basic endochitinase B	EC:3.2.1.14
		<i>TraesCS7D03G1286900</i>	K11593	ELF2C, AGO; eukaryotic translation initiation factor 2C	-
		<i>TraesCS7D03G1260600</i>	K20547	CHIB; basic endochitinase B	EC:3.2.1.14
		<i>TraesCS7D03G1287400</i>	K11593	ELF2C, AGO; eukaryotic translation initiation factor 2C	-

For genes from the loci associated with seed coat color, 14 found a match with the orthogroups of the metabolic pathways of the KEGG database related to pigment biosynthesis (Table 2). Eight genes were involved in the phenylpropanoid biosynthesis (loci on chromosomes 3A, 3B, 6A, and 6B). Two genes were involved in the carotenoid biosynthesis pathway (loci on chromosomes 2A and 6A), and one gene each was involved in flavone and flavonol biosynthesis, flavonoid biosynthesis, tryptophan metabolism, and terpenoid backbone biosynthesis.

Table 2. List of candidate genes from QTLs associated with seed shell color. Columns of the table contain QTL name (QTL), chromosome and position in cm (Chr/Pos), gene ID, KEGG orthogroup ID, KEGG orthogroup description, EC number, KEGG pathway ID and description.

Trait	Chr/Pos	Gene ID	KO ID	KO Description	EC	KEGG Pathway ID	KEGG Pathway Description
Q.YCrCb_dCCr_1-2A.3	2A/196	<i>TraesCS6A03G0725700</i>	K09840	NCED; 9-cis-epoxycarotenoid dioxygenase	EC:1.13.11.51	map00906	Carotenoid biosynthesis
		<i>TraesCS2A03G0158600</i>	K22772	FG2; flavonol-3-O-glucoside L-rhamnosyltransferase	EC:2.4.1.159	map00944	Flavone and flavonol biosynthesis
		<i>TraesCS6A03G0953500</i>	K13065	HCT; shikimate O-hydroxycinnamoyltransferase	EC:2.3.1.133	map00941	Flavonoid biosynthesis
		<i>TraesCS2A03G0099800</i>	K13066	COMT; caffeic acid 3-O-methyltransferase / acetylserotonin O-methyltransferase	EC:2.1.1.68; 2.1.1.4	map00380	Tryptophan metabolism
Q.Lab_dCb_3-3A ^{i*}	3A/195	<i>TraesCS3A03G0925800</i>	K01904	4CL; 4-coumarate-CoA ligase	EC:6.2.1.12	map00940	Phenylpropanoid biosynthesis
		<i>TraesCS3A03G0925900</i>	K01904	4CL; 4-coumarate-CoA ligase	EC:6.2.1.12	map00940	Phenylpropanoid biosynthesis
Q.RGB_dCB_1-3B.1	3B/269.179	<i>TraesCS3B03G1115600</i>	K12355	REF1; coniferyl-aldehyde dehydrogenase	EC:1.2.1.68	map00940	Phenylpropanoid biosynthesis
HSV_dCH_1-3B ^{m**}	306.179	<i>TraesCS3B03G1278200</i>	K01904	4CL; 4-coumarate-CoA ligase	EC:6.2.1.12	map00940	Phenylpropanoid biosynthesis

Table 2. Cont.

Trait	Chr/Pos	Gene ID	KO ID	KO Description	EC	KEGG Pathway ID	KEGG Pathway Description
HSV_dCH_3-6A ^{t***}	6A/246	<i>TraesCS6A03G0725000</i>	K09843	CYP707A; (+)-abscisic acid 8'-hydroxylase	EC:1.14.14.137	map00906	Carotenoid biosynthesis
		<i>TraesCS6A03G0741000</i>	K00021	HMGCR; hydroxymethylglutaryl-CoA reductase (NADPH)	EC:1.1.1.34	map00900	Terpenoid backbone biosynthesis
Q.YCrCb_dCCr_1-6A	6A/340	<i>TraesCS6A03G0953500</i>	K13065	HCT; shikimate O-hydroxycinnamoyltransferase	EC:2.3.1.133	map00940	Phenylpropanoid biosynthesis
		<i>TraesCS2A03G0163500</i>	K00430	peroxidase	EC:1.11.1.7	map00940	Phenylpropanoid biosynthesis
		<i>TraesCS2A03G0164200</i>	K00430	peroxidase	EC:1.11.1.7	map00940	Phenylpropanoid biosynthesis
Lab_dCL_2-6B ^{x,****}	6B/220	<i>TraesCS6B03G0367700</i>	K00430	peroxidase	EC:1.11.1.7	map00940	Phenylpropanoid biosynthesis

* Co-located QTL: Q.YCrCb_dCCb_3-3A ⁱ. ** Co-located QTLs: HSV_dCH_3-3B ^m; HSV_dCS_1-3B ^m; HSV_dCV_1-3B ^m; HSV_dCV_3-3B ^m; Lab_dCb_1-3B ^m; Lab_dCb_2-3B ^m; Lab_mb-3B ^m; RGB_dCR_2-3B ^m; YCrCb_dCCb_1-3B ^m; YCrCb_dCCb_2-3B ^m; YCrCb_dCCr_1-3B ^m; YCrCb_mCcb-3B ^m. *** Co-located QTL: HSV_dCV_3-6A ^t; Lab_ma-6A ^t. **** Co-located QTLs: Lab_dCL_3-6B ^x; Lab_mb-6B ^x; RGB_dCG_2-6B ^x; RGB_dCR_2-6B ^x; YCrCb_dCCb_2-6B ^x; YCrCb_dCY_2-6B ^x; YCrCb_dCY_3-6B ^x.

3. Discussion

3.1. Using Digital Image Analysis for QTL Identification

Based on the analysis of digital images, we identified QTLs associated with quantitative seed characteristics in wheat. With the development of modern phenotyping technologies [39,40], such approaches are increasingly being used [7–11,41]. Modern digital cameras and image processing algorithms have made great progress; they allow us to estimate even small differences in the color characteristics of seeds, their shape and size with high accuracy. In addition, these approaches have one interesting feature: the use of a large number of quantitative characteristics that are essentially derived from the same biological trait of the plant. For example, seed shape and size could be described as the sets of elliptic Fourier components [41] or virtual curves [42,43]. Components of various digital spaces [44,45] represent seed coat color. In the case when quantitative traits are derived from the same biological trait, we can assume that they will be associated with the same loci. Williams and Sorrels [11] used QTL for a set of seed size and shape characteristics derived from the developed seed image in two projections (elliptic Fourier components) as well as length, width, and thousand-kernel weight (TKW). They studied two populations, one of which, SynOpDH, was derived from crosses of the same parents used to obtain the ITMI population, and the other Cayuga × Caledonia was a doubled-haploid mapping population (C × C). Thirty-one loci were identified for the SynOpDH population which controlled from one to four traits per locus. Thirty loci were identified for the C × C population which also controlled from one to four traits per locus.

From our results (Figure 6), it is apparent that many color trait loci overlap between each other but not with the QTLs of size/shape (Table S1; chromosomes 3B, 3D, and 6B). The exception is rugosity which reflects the roughness of the shell and is probably associated with color distortions at the seed boundary on the image background. On the one hand, these features reflect the degeneracy of the evaluated traits which indicates their redundancy. On the other hand, the location of several loci related to color traits in the same region may indicate a more reliable identification of the association of the locus with a particular trait. Using many digital representations of the same trait looks redundant and confusing. The reasonable step would be to select a single or a few numerical characteristics that are most efficient in the identification of QTLs. We believe that various numerical representations of the same biological trait are useful and allow the evaluation of its subtle details. Many traits with QTL in the same locus can support its significance.

3.2. Identification of QTLs Associated with Seed Features

Our analysis allowed us to identify a number of QTLs associated with wheat seed characteristics, shape/size, and shell color. Such analyses have long been intensively conducted based on both QTL and GWAS investigations [8,11,41,46,47]. Reference [11] investigated three-dimensional characteristics of seed size and shape based on the analysis of images of seed obtained in two projections and the use of Fourier analysis-based descriptors using two populations, one of which was SynOpDH. They found a QTL that affects a number of shape characteristics. For seed length, eight QTLs were found for chromosomes 2A, 2D, 4B, 5A, 5B, 6A, 7A, and 7D. In our work, we detected a smaller number of QTLs for this trait located on the chromosomes 2A, 2D, 3B, 5B (2 QTLs) and 7D. For seed width, three QTLs on chromosomes 2A, 5A and 6A were detected by [11], whereas we detected four QTLs for this trait located on the chromosomes 1D, 2D, 3B, and 4D. This anomaly could be due to the use of data from several different environments in many years by reference [11].

Reference [47] previously analyzed 92 accessions from the ITMI population for a large number of traits, including such traits as TKW and kernel color (KC), in different locations and years. We did not find any overlap of QTL for TKW with the traits characterizing seed size and shape in our work. For seed color, 15 QTLs were reported by [47]. A comparison of our results with those from this work showed that of the 15 QTLs, one exactly matched the one found in our work. This is Q.KC_Pu07-3B [47] bounded by markers AX-94979462 and IAAV6088 and located in our work on chromosome 3B at position 306.179 (Table S1). In our work, several QTLs associated with seed shell color characteristics correspond to this locus (see marker HSV_dCH_1-3Bm (Table S1) and also listed in Table 2.) It is also interesting to note the QTLs Q.KC_Mo07-3D and Q.KC_Mo08-3D [47] bounded by markers D_GDS7LZN02IJRXZ_309 on the left and CAP12_c2615_128 on the right located on chromosome 3D at 76 cM. In our work, we found a series of color-related QTL localized on chromosome 3D at position 100–102, bounded by markers CAP12_c2615_128 on the left and BS00067163_51 on the right. Thus, the QTLs from our work and that of reference [47] are in the vicinity on the chromosome. For the other QTLs associated with color, we found no coincidence. For example, reference [47] identified five QTLs associated with shell color on chromosome 5A. However, in our work, only two loci at other positions on this chromosome were associated with color.

Among the loci associated with color, the site on chromosome 3D (at 100 cM) bounded by the markers CAP12_c2615_128 and BS00067163_51 is perhaps the most interesting. As indicated above, it is located next to the color QTL identified in [47]. In our work, 34 different traits characterizing seed color are associated with it. All of them are color parameters in various color spaces. Our analysis allowed us to localize the physical coordinates of this site on chromosome 3D: 573.6–580.8 Mbp according to IWGS v2.0 genome annotation (Table S5). Interestingly, reference [48] recently performed the analysis of the *PHS-3D* QTL associated with seed resistance to pre-harvest sprouting for a population of synthetic hexaploid wheat [49]. It turned out that on the physical map this region is located on chromosome 3D at positions 571.9–574.3 Mbp which overlaps with the physical localization of the QTL we identified. Reference [48] also showed that plant genotypes susceptible to pre-harvest sprouting are characterized by a ~2.4 Mbp deletion involving 20 genes in this region of the genome. It turned out that the gene encoding the transcription factor Myb10-D, which confers resistance to pre-harvest sprouting by activation of flavonoid and abscisic acid biosynthesis pathways, was located in this region. Note that plants that do not contain deletions in this region and are resistant to pre-harvest sprouting have reddish/brown coloring of the seed shell.

3.3. Epistatic QTLs

In our work, we identified several QTLs whose contribution to the trait are non-additive. Currently, there are only a few examples of epistatic QTLs analysis in wheat [46,50–52]. We found 34 QTL pairs exhibiting epistatic interactions. Our results show a predominance of epistatic QTLs for color features (30 pairs of QTLs). One pair of QTL each was identified

for the area, width, rugosity, and solidity of seed. These results demonstrate a possible interaction of genes located at different loci in the formation of color traits.

Epistatic QTLs for yield, flour color, and seed weight traits were investigated for the RIL population of durum wheat [51]. QTL epistatic interactions on chromosomes 1A and 1B and chromosomes 5B and 7B were determined for thousand-seed weight. Reference [46] analyzed yield traits including 1000-seed weight, seed length, and seed width of bread wheat in the RIL Chuannong18 × T1208 population. Epistatic QTLs were found for 114 QTL pairs, including 10 for seed length, 17 for seed width, and seven for 1000-seed weight. The authors noted the complex nature of the effect of epistatic interactions on the seed properties. Thus, more epistatic pairs for geometric seed traits were identified in these works compared to ours. However, in our work, the most intense epistatic interactions were shown for seed shell color, a trait not reported in [46]. Interestingly, QTL pairs that are localized on chromosomes 3B (position 306) and 3D (position 100), which are also characterized by a large number of additive QTLs, are often represented. This again indicates the importance of these regions for the formation of seed shell color in wheat.

3.4. Gene Prioritization

Our analysis allowed us to identify a number of candidate genes associated with seed size/shape and their color, based on bioinformatics analysis and annotation of genes according to data in the literature and the KEGG database. We identified eight loci associated in the genome with seed size/shape traits for which we found 28 orthologous genes involved in gene networks controlling these traits. Some of genes are transcription factors (EREBP, HD-ZIP, and MYB) that may be involved in the regulation of seed growth and development. In particular, transcription factors associated with the response to ethylene (EREBP) are known to be involved in the determination of seed size, seed weight, and accumulation of seed oil and protein in *A. thaliana* [53]. Reference [54] identified two transcription factors from the AP2/EREBP family, TaPARG, located on 2A and 2D chromosomes of wheat which regulate several yield-related traits, including seed weight.

Several genes represent families of enzymes related to ubiquitin modification (E3 ubiquitin–protein ligase, ubiquitin thioesterase protein OTUB1). Ubiquitins and related enzymes are known to play an important role in seed development by controlling cell proliferation [55]. For example, genes of the E3 ligase family are involved in amylose biosynthesis in wheat [56]. The *TaGW2-6A* gene from this family is shown to control seed size [57].

Another type of enzyme that was frequently found among the candidates we identified are endochitinases. These are enzymes involved in defense against pathogens such as bacteria or fungi in seeds [58,59]. However, this is not their only role in seed formation and function. It has been shown from proteomic data in rice that *chitinase 14* interacts with the *GW2* (*RING-type E3 ubiquitin ligase*) gene. It was also shown that *GW2* controls seed size through the regulation of chitinase 14 and phosphoglycerate kinase levels or activities [60]. Other genes that we detected (cytokinin dehydrogenase, aarF domain-containing kinase, and eukaryotic translation initiation factor 2C) may also be associated with seed development in wheat [20].

For QTLs associated with seed color, we also found a number of possible candidates among genes encoding enzymes of plant pigment biosynthesis pathways. On chromosome 2A, we found several genes that are involved in plant pigment biosynthesis. Among them, one gene, which we annotated as *NCED*, is involved in the carotenoid biosynthesis pathway. In rice, mutants of this gene lead to changes in pericarp seed coloration [61]. The expression of this enzyme is controlled by abscisic acid [62], and *NCED* is also involved in ABA biosynthesis [48]. The functions of this gene in seed development are well known [63]: it is an important regulator in seed development, in the zygotic embryogenesis, and dormancy. Another gene related to carotenoid biosynthesis that we found among the primate genes is *CYP707A* ((+)-abscisic acid 8'-hydroxylase) (Table 2). Its functions are closely related to the *NCED* gene, and its participation in the same processes related to seed development has

been shown [63]. Interestingly, two of these genes are located near loci associated with seed coat redness (*Q.YCrCb_dCCr_1-2A.3* and *HSV_dCH_3-6At/Lab_ma-6At*).

We found two genes that are involved in the flavonoid biosynthesis pathway (Table 2) that provide different coloration of seeds in cereals [64,65]. These include the homologue FG2 (*flavonol-3-O-glucoside L-rhamnosyltransferase*), for which mutations result in a phenotype with seed color change in soybean [66]. *Shikimate O-hydroxycinnamoyltransferase* has been shown to be elevated in expression in wheat plants with high seed antioxidant activity [67].

In the QTL region of *Q.Lab_dCb_3-3Ai*, we found two genes involved in phenylpropanoid biosynthesis. They both encode *4-coumarate-CoA ligase*. This enzyme catalyzes the conversion of p-coumaric acid to p-coumaroyl CoA, which further serves as a source of biosynthesis of both lignin (a structural component of the seed shell) and flavonoids. In transcriptome-wide association studies in *Brassica napus*, 4CL expression during seed development was shown to positively correlate with seed coat content, i.e., the fraction of seed mass attributable to the coat [68]. Interestingly, in gene expression analysis in *B. napus* plants with brown seed coloration, the expression level of genes encoding this enzyme was higher than in plants with yellow coloration [69].

Modern genomics advances in wheat genome sequencing and genetic marker technologies allow QTLs to be linked to the physical coordinates of the wheat genome. Such analysis is now an important complement to QTL identification [70–72]. The genes we have identified as possible candidates associated with seed size/shape and color formation in the ITMI/MP can be further investigated in more detail using genetic and molecular methods to establish the mechanisms controlling these important traits.

4. Materials and Methods

4.1. Materials

We studied seeds from 114 accessions of the well-known ITMI/MP of bread wheat (*T. aestivum* L.). The ITMI mapping population was obtained by pollination of the *T. aestivum* (var. Opata 85) flower with the pollen of the synthetic hexaploid spring wheat W7984 [47]. Plants of each genotype were grown in season in 2014 on the experimental fields of IPK in Gatersleben, Germany.

4.2. Seed Imaging Protocol and Image Processing

Seeds were imaged in March 2020. We supposed that the storage time affected seed traits for different genotypes in the same manner. The imaging of seeds was performed according to the protocol described earlier [1]: seeds were scattered in an amount of up to 20 pieces on the table on a white sheet of A4 paper. A ColorChecker color calibration card (x-rite ColorChecker® Classic Mini, X-Rite, Grand Rapids, MI, USA, <https://xritephoto.com/camera>; accessed on 20 January 2022) was placed in the image area and used for color correction and obtaining image scale. The lighting was adjusted to avoid shadows. Images were taken with a digital camera Canon EOS 600D equipped with a Canon EF 100 mm f/2.8 Macro USM lens and saved in files in JPG or PNG format. Examples of images are shown in Figure S2. Digital image processing was performed using the SeedCounter application for desktop PC [73] with color analysis capabilities [1]. We used two images per genotype for our analysis: 15 and 5 seeds. Splitting was initially used to check for reproducibility of the evaluated trait values. No significant differences between mean values of the seed traits were observed between these replicates according to an *F*-test (results not shown). Therefore, we used the average values of the images of 20 seeds from two replicates as input for QTL analysis.

4.3. Quantitative Characteristics of Seed Shape, Size, and Color

The analysis of digital images for each seed yielded a set of 55 quantitative characteristics described earlier [1]. Size was defined by seed length (sL), width (sW), and projected

area (sA). Seed shape characteristics included circularity (sCi), roundness (sRo), rugosity (sRg), and solidity (sSo).

The circularity and roundness indices reflect how close the shape of a contour is to a circle but are calculated differently. Circularity is a measure of the similarity of a 2D figure to a circle [43]. For objects with rugged contours, the closeness of the shape to a circle is more correctly described by the roundness parameter since this value does not depend on the roughness of the contour line. This index is calculated as the ratio of the area of the shape (area) to the square of the length of the major axis [40]. For a shape other than a circle, the index takes values less than unity. The rugosity index (sRg) is defined as the ratio of the contour perimeter to the convex perimeter [43]. The index of solidity (sSo) is the ratio of the contour area to the area of its convex hull [74].

To describe the color characteristics of the seeds, we used a color representation in the form of four-color spaces: RGB, HSV, Lab, and YCrCb [44,75,76]. Each of them represents color as three components. The component values of one space can be obtained by transforming the component values of the other. The color features included two types of descriptors, which were independently calculated for each color space.

The first type of descriptors: mean values of component intensities for seed pixels. To calculate them, the mean and standard deviations of intensities for each of the color component channels were first estimated, then the pixels whose intensities differ from the mean by more than three standard deviations were excluded from the analysis. The mean value was calculated for the remaining pixels and used further. The descriptors of the average component values are indicated by a small letter m. For example, for the RGB color space, these are the three parameters: RGB_mR, RGB_mG, and RGB_mB. For other spaces, the indications are similar.

The second type of descriptors are dominant seed colors. These descriptors provide an illustration of representative colors in an image or its region [77]. To determine dominant colors, all seed pixels were grouped by color similarity into three clusters. The clusters were ranked by the number of pixels they contained. In each of the three clusters, the values of the three color components for the centroid were determined. This procedure was performed for each color space and resulted, respectively, in nine color descriptors. For example, for the RGB space, these are RGB_dCj_i parameters, where $j = 1,2,3$ is the color component designation, and $i = 1,2,3$ is the number of the dominant cluster. For example, the RGB_dCR_1 parameter is the R component for the first dominant color in the RGB space. The use of three dominant colors allows for a more accurate estimation of the shades of seed coloration.

As a result, three size characteristics, four shape characteristics, and 48 color characteristics were determined for each seed. Characteristics were calculated for each seed of the 114 wheat genotypes. The mean values for the genotype were estimated and used in the QTL analysis.

4.4. Statistical Analysis

To get an idea of the similarity of genotypes in the space of seed traits, we used the principal component method implemented in the PAST program [78].

4.5. Genotyping and QTL Analysis

Fresh flag leaves were used for the DNA extraction for the purpose of genotyping which was performed using Illumina (San Diego, CA, USA) Infinium technology. An optimized array (wheat 20K Infinium SNP array) was used. This array is the refined version of the 15K chip [79] of 90K iSELECT SNP-chip as previously reported [80]. To make it more informative, 5385 markers from the 35K Wheat Breeders Array [81] were also added. All sequences of the markers, a complete genetic map and the list of 92 RILs with genotypic data are available in reference [47].

To capture the variance explained by the molecular markers such as SNPs mapped to any genome, different methods were proposed (such as single marker analyses, inter-

val mapping, and composite interval mapping) and implemented in different computer programs (*Qgene*, *QTL Cartographer* and *PLABQTL*) [82] which have successfully been used to detect QTLs for various traits in wheat including seed-related traits such as seed longevity [37,83,84]. In light of the limitations of the above-mentioned methods, a more refined method known as “inclusive composite interval mapping” was proposed and implemented in the *QTL Ici Mapping 4.2.53* (<http://www.isbreeding.net/>) (latest released in September 2019, accessed on 2 February 2022) which is considered as the most modern method of QTL detection [47]. We have recently detected several QTLs for *Fusarium* head blight [85] and seed longevity [50] in wheat and germination-related traits in tobacco [86] by applying the *QTL Ici Mapping* tool. Therefore, we convened the *Ici Mapping 4.2.53* to detect the putative additive QTLs of the traits under consideration by applying the inclusive composite interval mapping (ICIM) command where 1.0 cM was the walking speed. An LOD score of $>2.0 \leq 3$ was applied to detect QTLs as significant and >3.0 as highly significant [87].

In order to discover digenic epistasis QTLs to find clues for latent variation, the ICIM-EPI command was used where LOD was kept at 5.0 cM. Here, the epistasis QTLs with LOD ≥ 5 and explaining $\geq 5\%$ phenotypic variance were reported. All QTLs were assigned names according to the rules set out in the Catalog of Gene Symbols [88]. All additive and epistasis QTLs were visualized using the “circlize” package in R [89].

4.6. Seed Traits Similarity by QTL Location

Preliminary analysis of QTLs demonstrated that loci for several traits often overlap. In particular, the same locus may associate with several characteristics of the seed shell color. In this regard, we decided to analyze the similarity of traits by their localization in the genome. To do this, we compiled a list of loci with which they were associated for each of the 55 features. Based on the overlap of the list of these loci, we calculated the Ochiai index [90] for each pair of features. This index was suggested for ecological studies to estimate associations between species and groups of sites representing habitat types. In our work, the index reflects the degree of overlap of the lists of loci between two traits. The greater the similarity between sets of loci for two traits, the greater the index value. It equals 1 when loci are identical and 0 when there is no loci overlap. Based on this measure, we performed clustering of features in the PAST [78] package and built a tree of the trait similarity.

4.7. QTL Gene Prioritization

In order to identify possible candidate genes associated with seed traits, we prioritized them based on several conditions and using the gene annotation provided in the KEGG database [91]. The analysis included loci for which the LOD value exceeded three and consisted of several steps.

In the first step of the analysis, we determined the physical localization of the markers by aligning their sequences to the IWGS 2.1 wheat genome assembly sequence [92]. Genome sequence and annotation data were obtained from URGI (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Assemblies/v2.1; accessed on 10 January 2022). We considered “high confidence” gene annotations only. Marker sequences were obtained from reference [80] and the Gramene marker database (<https://archive.gramene.org/markers/>; accessed on 10 January 2022) [93]. Marker sequences were aligned using BLASTn of the BLAST+ package [94] using $e\text{-value} = 1 \times 10^{-17}$ (other parameters were set by default). This allowed us to search similar sequences with length above 50 nt and avoid noise. Thus, for each of our selected QTLs, we obtained a list of IWGS 2.1 wheat genome annotation genes. Note that it was not possible to determine the physical boundaries of the QTLs for several loci because, for one of the sequences, the alignment did not occur on the chromosome corresponding to the marker.

Since plant pigments can be synthesized in various tissues and organs, when prioritizing genes for seed color QTL, we additionally filtered genes by expression level in the

seed. For this purpose, we used the expression data presented for wheat in the expVIP database [95]. Data in text format were downloaded from URGI (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.1/iwgsc_refseqv1.1_rnaseq_mapping_2017July20.zip; accessed on 10 January 2022). We used data from RNA-seq experiments in which the column “High level tissue” contains “seed”. We selected genes as expressed if their TPM ≥ 1 in these experiments. To perform filtering, we developed scripts in Python, taking into consideration the Gene ID conversion between annotation ver. 2.1 (genome) and 1.2 (transcriptome).

In the second stage of analysis, we generated a list of orthologous protein groups from the KEGG database [91], which are associated with the formation of the traits of seed size and color. It is well known that the color of the seed shell is determined by the presence of specific plant pigments in it [22]. Therefore, we selected orthogroups involved in KEGG pathways of the biosynthesis of these pigments and a number of their precursors. The list of such pathways includes tryptophan metabolism (map00380), terpenoid backbone biosynthesis (map00900), carotenoid biosynthesis (map00906), phenylpropanoid biosynthesis (map00940), flavonoid biosynthesis (map00941), anthocyanin biosynthesis (map00942), isoflavonoid biosynthesis (map00943), and flavone and flavonol biosynthesis (map00944). We obtained 307 KEGG orthogroups for these pathways.

Seed size depends on a multitude of biological processes occurring at the molecular level, including protein ubiquitination, response to hormonal signals, protein biosynthesis and transport, etc. Therefore, it was not possible to isolate the pathways corresponding to these processes based only on their description in the KEGG database. However, the genes involved in seed development have been fairly well experimentally studied in *A. thaliana* and rice (*Oryza sativa*). Therefore, we used three recent literature reviews describing the molecular processes of seed development in *Arabidopsis* and rice [20,21,38]. We combined set of genes from three papers and removed duplicated IDs. During compilation, we converted gene IDs from reference [20] from RAP to MSU format using the “ID converter” tool at the website of the OryzaExpress database ([96]; http://bioinf.mind.meiji.ac.jp/OryzaExpress/ID_converter.php; accessed on 20 January 2022). We identified KEGG orthogroups for the selected genes and used them for our analysis.

The assignment of KEGG orthologous groups to wheat genes by their sequence was performed using BlastKOALA [97] and KofamKOALA [98] tools. The orthogroups were assigned to genes by at least one of the methods. We prioritized genes whose orthogroups were in the lists associated with traits of seed coat color and size.

5. Conclusions

A QTL search for seven traits of seed size/shape and 48 traits of coat color evaluated on the basis of digital image analysis of the ITMI/MP identified 212 additive and 34 pairs of epistatic QTLs on all the chromosomes of wheat genome except chromosomes 1A and 5D. The number of QTLs varied from one to ten for one single trait. The majority of the traits yielded three to four QTLs. We demonstrated that one locus could control dozens of seed characteristics. Analysis of the loci overlap showed that this is typical for color traits and rarely occurred for seed size/shape traits. For a number of highly significant QTLs, we identified the physical location of their markers on the wheat chromosomes. Additionally, we demonstrated that the overlap between QTL regions was low for seed size/shape traits and high for coat color traits. Using the literature and KEGG data, we identified sets of genes in *Arabidopsis* and rice from the networks controlling seed size and color. This information along with the coordinates of the markers in the wheat genome was used for the prioritization of wheat genes within QTL regions. We identified 29 candidate genes from the seed size-related loci and 14 for genes from the loci associated with seed coat color. The genes we have identified as possible candidates associated with seed size/shape and color formation in the ITMI/MP can be further investigated in more detail using genetic and molecular methods to establish the mechanisms controlling these important traits. Our results demonstrate the complex nature of the genetic control of the wheat seed

traits and the efficiency of the image analysis methods for obtaining novel QTLs for seed characteristics.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/plants11162105/s1>: Figure S1: PCA biplot of seed size/shape and color traits of ITMI/MP performed using seed coat traits (mean seven values for seed size/shape and 12 values for color components of four-color spaces); Figure S2: Examples of seed images used for digital phenotyping; Table S1: Complete list of QTLs identified through composite interval mapping. Left and right flanking markers linked to QTL are also provided along with LOD, PVE%, and additive effect (+ = provided by Opata and - = provided by W7984 parent). Similar asterisks (also highlighted in similar color) indicate likely identical loci. Markers highlighted in red are involved in multiple QTLs. Table S2: Pairs of epistatic quantitative trait loci detected in the ITMI/MP. Table S3: List of 197 (42 from rice and 155 from Arabidopsis) genes involved in molecular processes of the grain development. Input Gene ID = Gene ID from review papers; KEGG Gene ID = GID for gene in KEGG database; NA: not found in KEGG; KEGG Gene Name = Gene name in KEGG database and KEGG Orthology ID = Orthogroup ID in KEGG database. Table S4: Pathway ID and KEGG orthologous groups. Pathway KEGG = pathway ID and KEGG orthology group = Orthogroup ID in KEGG database. Table S5: List of QTLs used in gene prioritization. Table S6: List of 823 genes associated with grain color.

Author Contributions: Conceptualization, M.A.R.A., A.B. and D.A.A.; methodology, M.A.R.A., V.S.K. and E.G.K.; software, E.G.K.; validation, M.A.R.A., E.G.K. and D.A.A.; formal analysis, M.A.R.A., E.G.K., M.A.G. and N.A.S.; investigation, M.A.R.A., E.G.K., N.A.S. and D.A.A.; resources, A.B.; data curation, M.A.R.A., E.G.K. and M.A.G.; writing—original draft preparation, M.A.R.A., E.G.K. and D.A.A.; writing—review and editing, M.A.R.A., A.B. and D.A.A.; visualization, M.A.R.A. and E.G.K.; supervision, A.B. and D.A.A.; project administration, A.B. and D.A.A.; funding acquisition, A.B. and D.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding, whereas the publication of this article was funded by the Open Access Fund of the Leibniz Association, Russian Science Foundation (21-76-30003) and Russian Ministry of Science and Higher Education (FWNR-2022-0020).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Seed phenotyping was supported by the RSF project 21-76-30003. The data analysis was performed using computational resources of the ICG SB RAS “Bioinformatics” Joint Computational Center, supported by the budget project № FWNR-2022-0020.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Afonnikov, D.A.; Komyshev, E.G.; Efimov, V.M.; Genaev, M.A.; Koval, V.S.; Gierke, P.U.; Börner, A. Relationship between the Characteristics of Bread Wheat Grains, Storage Time and Germination. *Plants* **2021**, *11*, 35. [[CrossRef](#)] [[PubMed](#)]
2. Brinton, J.; Uauy, C. A reductionist approach to dissecting grain weight and yield in wheat. *J. Integr. Plant Biol.* **2019**, *61*, 337–358. [[CrossRef](#)] [[PubMed](#)]
3. Huang, Y.; Kong, Z.; Wu, X.; Cheng, R.; Yu, D.; Ma, Z. Characterization of three wheat grain weight QTLs that differentially affect kernel dimensions. *Theor. Appl. Genet.* **2015**, *128*, 2437–2445. [[CrossRef](#)] [[PubMed](#)]
4. Zhang, X.; Deng, Z.; Wang, Y.; Li, J.; Tian, J. Unconditional and conditional QTL analysis of kernel weight related traits in wheat (*Triticum aestivum* L.) in multiple genetic backgrounds. *Genetica* **2014**, *142*, 371–379. [[CrossRef](#)] [[PubMed](#)]
5. Baasandorj, T.; Ohm, J.B.; Manthey, F.; Simsek, S. Effect of kernel size and mill type on protein, milling yield, and baking quality of hard red spring wheat. *Cereal Chem.* **2015**, *92*, 81–87. [[CrossRef](#)]
6. Li, S.; Wang, L.; Meng, Y.; Hao, Y.; Xu, H.; Hao, M.; Lan, S.; Zhang, Y.; Lv, L.; Zhang, K.; et al. Dissection of genetic basis underpinning kernel weight-related traits in common wheat. *Plants* **2021**, *10*, 713. [[CrossRef](#)]
7. Ali, A.; Ullah, Z.; Alam, N.; Naqvi, S.M.; Jamil, M.; Bux, H.; Sher, H. Genetic analysis of wheat grains using digital imaging and their relationship to enhance grain weight. *Sci. Agric.* **2020**, *77*, e20190069. [[CrossRef](#)]
8. Alemu, A.; Feyissa, T.; Tuberosa, R.; Maccaferri, M.; Sciara, G.; Letta, T.; Abeyo, B. Genome-wide association mapping for grain shape and color traits in Ethiopian durum wheat (*Triticum turgidum* ssp. durum). *Crop J.* **2020**, *8*, 757–768. [[CrossRef](#)]

9. Kumari, S.; Jaiswal, V.; Mishra, V.K.; Paliwal, R.; Balyan, H.S.; Gupta, P.K. QTL mapping for some grain traits in bread wheat (*Triticum aestivum* L.). *Physiol. Mol. Biol. Plants* **2018**, *24*, 909–920. [[CrossRef](#)] [[PubMed](#)]
10. Kumar, A.; Mantovani, E.E.; Seetan, R.; Soltani, A.; Echeverry-Solarte, M.; Jain, S.; Simsek, S.; Doehlert, D.; Alamri, M.S.; Elias, E.M.; et al. Dissection of genetic factors underlying wheat kernel shape and size in an elite × nonadapted cross using a high density SNP linkage map. *Plant Genome* **2015**, *9*, 0081. [[CrossRef](#)]
11. Williams, K.; Sorrells, M.E. Three-dimensional seed size and shape QTL in hexaploid wheat (*Triticum aestivum* L.) populations. *Crop Sci.* **2014**, *54*, 98–110. [[CrossRef](#)]
12. Gegas, V.C.; Nazari, A.; Griffiths, S.; Simmonds, J.; Fish, L.; Orford, S.; Sayers, L.; Doonan, J.H.; Snape, J.W. A genetic framework for grain size and shape variation in wheat. *Plant Cell* **2010**, *22*, 1046–1056. [[CrossRef](#)] [[PubMed](#)]
13. Brescghello, F. MESorrells QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. *Field Crops Res.* **2007**, *101*, 172–179. [[CrossRef](#)]
14. Huang, X.; Börner, A.; Röder, M.; Ganai, M. Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theor. Appl. Genet.* **2002**, *105*, 699–707. [[CrossRef](#)] [[PubMed](#)]
15. Ma, L.; Li, T.; Hao, C.; Wang, Y.; Chen, X.; Zhang, X. Ta GS 5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. *Plant Biotechnol. J.* **2016**, *14*, 1269–1280. [[CrossRef](#)]
16. Ma, M.; Wang, Q.; Li, Z.; Cheng, H.; Li, Z.; Liu, X.; Song, W.; Appels, R.; Zhao, H. Expression of Ta CYP 78A3, a gene encoding cytochrome P450 CYP 78A3 protein in wheat (*Triticum aestivum* L.), affects seed size. *Plant J.* **2015**, *83*, 312–325. [[CrossRef](#)]
17. Ma, D.; Yan, J.; He, Z.; Wu, L.; Xia, X. Characterization of a cell wall invertase gene TaCwi-A1 on common wheat chromosome 2A and development of functional markers. *Mol. Breed.* **2012**, *29*, 43–52. [[CrossRef](#)]
18. Su, Z.; Hao, C.; Wang, L.; Dong, Y.; Zhang, X. Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* **2011**, *122*, 211–223. [[CrossRef](#)]
19. Tikhenko, N.; Alqudah, A.M.; Borisjuk, L.; Ortleb, S.; Rutten, T.; Wu, D.; Nagel, M.; Himmelbach, A.; Mascher, M.; Röder, M.S.; et al. DEFECTIVE ENDOSPERM-D1 (Dee-D1) is crucial for endosperm development in hexaploid wheat. *Commun. Biol.* **2020**, *3*, 791. [[CrossRef](#)]
20. Chen, K.; Łyskowski, A.; Jaremko, Ł.; Jaremko, M. Genetic and molecular factors determining grain weight in rice. *Front. Plant Sci.* **2021**, *12*, 605799. [[CrossRef](#)]
21. Li, N.; Xu, R.; Li, Y. Molecular networks of seed size control in plants. *Annu. Rev. Plant Biol.* **2019**, *70*, 435–463. [[CrossRef](#)] [[PubMed](#)]
22. Lachman, J.; Martinek, P.; Kotíková, Z.; Orsák, M.; Šulc, M. Genetics and chemistry of pigments in wheat grain: A review. *J. Cereal Sci.* **2017**, *74*, 145–154. [[CrossRef](#)]
23. Khlestkina, E. Genes determining the coloration of different organs in wheat. *Russ. J. Genet. Appl. Res.* **2013**, *3*, 54–65. [[CrossRef](#)]
24. Glagoleva, A.Y.; Shoeva, O.Y.; Khlestkina, E.K. Melanin pigment in plants: Current knowledge and future perspectives. *Front. Plant Sci.* **2020**, *11*, 770. [[CrossRef](#)] [[PubMed](#)]
25. Ma, D.; Wang, C.; Feng, J.; Xu, B. Wheat grain phenolics: A review on composition, bioactivity, and influencing factors. *J. Sci. Food Agric.* **2021**, *101*, 6167–6185. [[CrossRef](#)]
26. Li, Y.; Ma, D.; Sun, D.; Wang, C.; Zhang, J.; Xie, Y.; Guo, T. Total phenolic, flavonoid content, and antioxidant activity of flour, noodles, and steamed bread made from different colored wheat grains by three milling methods. *Crop J.* **2015**, *3*, 328–334. [[CrossRef](#)]
27. Loskutov, I.G.; Khlestkina, E.K. Wheat, barley, and oat breeding for health benefit components in grain. *Plants* **2021**, *10*, 86. [[CrossRef](#)]
28. Khlestkina, E.K.; Pshenichnikova, T.A.; Usenko, N.I.; Otmakhova, Y.S. Prospects of molecular genetic approaches in controlling technological properties of wheat grain in the context of the “grain–flour–bread” chain. *Vavilov J. Genet. Breed.* **2016**, *20*, 511–527. [[CrossRef](#)]
29. Groos, C.; Gay, G.; Perretant, M.R.; Gervais, L.; Bernard, M.; Dedryver, F.; Charmet, G. Study of the relationship between pre-harvest sprouting and grain color by quantitative trait loci analysis in a white × red grain bread-wheat cross. *Theor. Appl. Genet.* **2002**, *104*, 39–47. [[CrossRef](#)]
30. Lin, M.; Zhang, D.; Liu, S.; Zhang, G.; Yu, J.; Fritz, A.K.; Bai, G. Genome-wide association analysis on pre-harvest sprouting resistance and grain color in US winter wheat. *BMC Genom.* **2016**, *17*, 794. [[CrossRef](#)]
31. Kumar, A.; Kumar, J.; Singh, R.; Garg, T.; Chhuneja, P.; Balyan, H.S.; Gupta, P.K. QTL analysis for grain colour and pre-harvest sprouting in bread wheat. *Plant Sci.* **2009**, *177*, 114–122. [[CrossRef](#)]
32. Pang, Y.; Liu, C.; Wang, D.; Amand, P.S.; Bernardo, A.; Li, W.; He, F.; Li, L.; Wang, L.; Yuan, X.; et al. High-resolution genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. *Mol. Plant* **2020**, *13*, 1311–1327. [[CrossRef](#)] [[PubMed](#)]
33. Guo, Y.; Zhang, G.; Guo, B.; Qu, C.; Zhang, M.; Kong, F.; Zhao, Y.; Li, S. QTL mapping for quality traits using a high-density genetic map of wheat. *PLoS ONE* **2020**, *15*, e0230601. [[CrossRef](#)] [[PubMed](#)]
34. Chu, J.; Zhao, Y.; Beier, S.; Schulthess, A.W.; Stein, N.; Philipp, N.; Röder, M.S.; Reif, J.C. Suitability of single-nucleotide polymorphism arrays versus genotyping-by-sequencing for genebank genomics in wheat. *Front. Plant Sci.* **2020**, *11*, 42. [[CrossRef](#)]
35. Colasuonno, P.; Marcotuli, I.; Gadaleta, A.; Soriano, J.M. From genetic maps to QTL cloning: An overview for durum wheat. *Plants* **2021**, *10*, 315. [[CrossRef](#)]

36. Arif, M.A.R.; Börner, A. An SNP based GWAS analysis of seed longevity in wheat. *Cereal Res. Commun.* **2020**, *48*, 149–156. [[CrossRef](#)]
37. Arif, M.A.R.; Börner, A. Mapping of QTL associated with seed longevity in durum wheat (*Triticum durum* Desf.). *J. Appl. Genet.* **2019**, *60*, 33–36. [[CrossRef](#)] [[PubMed](#)]
38. Li, N.; Xu, R.; Duan, P.; Li, Y. Control of grain size in rice. *Plant Reprod.* **2018**, *31*, 237–251. [[CrossRef](#)] [[PubMed](#)]
39. Yang, W.; Feng, H.; Zhang, X.; Zhang, J.; Doonan, J.H.; Batchelor, W.D.; Xiong, L.; Yan, J. Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. *Mol. Plant* **2020**, *13*, 187–214. [[CrossRef](#)] [[PubMed](#)]
40. Afonnikov, D.A.; Genaev, M.A.; Doroshkov, A.V.; Komyshev, E.G.; Pshenichnikova, T.A. Methods of high-throughput plant phenotyping for large-scale breeding and genetic experiments. *Russ. J. Genet.* **2016**, *52*, 688–701. [[CrossRef](#)]
41. Williams, K.; Munkvold, J.; Sorrells, M. Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (*Triticum aestivum* L.). *Euphytica* **2013**, *190*, 99–116. [[CrossRef](#)]
42. Martín-Gómez, J.J.; Rewicz, A.; Goriewa-Duba, K.; Wiwart, M.; Tocino, Á.; Cervantes, E. Morphological description and classification of wheat kernels based on geometric models. *Agronomy* **2019**, *9*, 399. [[CrossRef](#)]
43. Cervantes, E.; Martín, J.J.; Saadaoui, E. Updated methods for seed shape analysis. *Scientifica* **2016**, *2016*, 5691825. [[CrossRef](#)] [[PubMed](#)]
44. Komyshev, E.; Genaev, M.; Afonnikov, D. Analysis of color and texture characteristics of cereals on digital images. *Vavilov J. Genet. Breed.* **2020**, *24*, 340. [[CrossRef](#)]
45. Goriewa-Duba, K.; Duba, A.; Wachowska, U.; Wiwart, M. An evaluation of the variation in the morphometric parameters of grain of six *Triticum* species with the use of digital image analysis. *Agronomy* **2018**, *8*, 296. [[CrossRef](#)]
46. Ren, T.; Fan, T.; Chen, S.; Li, C.; Chen, Y.; Ou, X.; Jiang, Q.; Ren, Z.; Tan, F.; Luo, P.; et al. Utilization of a Wheat55K SNP array-derived high-density genetic map for high-resolution mapping of quantitative trait loci for important kernel-related traits in common wheat. *Theor. Appl. Genet.* **2021**, *134*, 807–821. [[CrossRef](#)] [[PubMed](#)]
47. Arif, M.A.R.; Shokat, S.; Plieske, J.; Ganal, M.; Lohwasser, U.; Chesnokov, Y.V.; Kocherina, N.V.; Kulwal, P.; Kumar, N.; McGuire, P.E.; et al. A SNP-based genetic dissection of versatile traits in bread wheat (*Triticum aestivum* L.). *Plant J.* **2021**, *108*, 960–976. [[CrossRef](#)] [[PubMed](#)]
48. Lang, J.; Fu, Y.; Zhou, Y.; Cheng, M.; Deng, M.; Li, M.; Zhu, T.; Yang, J.; Guo, X.; Gui, L.; et al. Myb10-D confers PHS-3D resistance to pre-harvest sprouting by regulating NCED in ABA biosynthesis pathway of wheat. *New Phytol.* **2021**, *230*, 1940–1952. [[CrossRef](#)]
49. Yang, J.; Tan, C.; Lang, J.; Tang, H.; Hao, M.; Tan, Z.; Yu, H.; Zhou, Y.; Liu, Z.; Li, M.; et al. Identification of qPHS. sicau-1B and qPHS. sicau-3D from synthetic wheat for pre-harvest sprouting resistance wheat improvement. *Mol. Breed.* **2019**, *39*, 132. [[CrossRef](#)]
50. Arif, M.A.R.; Agacka-Mołdoch, M.; Qualset, C.O.; Börner, A. Mapping of additive and epistatic QTLs linked to seed longevity in bread wheat (*Triticum aestivum* L.). *Cereal Res. Commun.* **2022**, 1–7. [[CrossRef](#)]
51. Roncallo, P.F.; Akkiraju, P.C.; Cervigni, G.L.; Echenique, V.C. QTL mapping and analysis of epistatic interactions for grain yield and yield-related traits in *Triticum turgidum* L. var. durum. *Euphytica* **2017**, *213*, 277. [[CrossRef](#)]
52. Roncallo, P.F.; Cervigni, G.L.; Jensen, C.; Miranda, R.; Carrera, A.D.; Helguera, M.; Echenique, V. QTL analysis of main and epistatic effects for flour color traits in durum wheat. *Euphytica* **2012**, *185*, 77–92. [[CrossRef](#)]
53. Jofuku, K.D.; Omidyar, P.K.; Gee, Z.; Okamura, J.K. Control of seed mass and seed yield by the floral homeotic gene APETALA2. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3117–3122. [[CrossRef](#)] [[PubMed](#)]
54. Li, B.; Li, Q.; Mao, X.; Li, A.; Wang, J.; Chang, X.; Hao, C.; Zhang, X.; Jing, R. Two novel AP2/EREBP transcription factor genes TaPARG have pleiotropic functions on plant architecture and yield-related traits in common wheat. *Front. Plant Sci.* **2016**, *7*, 1191. [[CrossRef](#)] [[PubMed](#)]
55. Li, N.; Li, Y. Ubiquitin-mediated control of seed size in plants. *Front. Plant Sci.* **2014**, *5*, 332. [[CrossRef](#)] [[PubMed](#)]
56. Parveen, A.; Rahim, M.S.; Sharma, A.; Mishra, A.; Kumar, P.; Fandade, V.; Kumar, P.; Bhandawat, A.; Verma, S.K.; Roy, J. Genome-wide analysis of RING-type E3 ligase family identifies potential candidates regulating high amylose starch biosynthesis in wheat (*Triticum aestivum* L.). *Sci. Rep.* **2021**, *11*, 11461. [[CrossRef](#)]
57. Lv, Q.; Li, L.; Meng, Y.; Sun, H.; Chen, L.; Wang, B.; Li, X. Wheat E3 ubiquitin ligase TaGW2-6A degrades TaAGPS to affect seed size. *Plant Sci.* **2022**, *320*, 111274. [[CrossRef](#)] [[PubMed](#)]
58. Pollard, A.T. Seeds vs fungi: An enzymatic battle in the soil seedbank. *Seed Sci. Res.* **2018**, *28*, 197–214. [[CrossRef](#)]
59. Gomez, L.; Allona, I.; Casado, R.; Aragoncillo, C. Seed chitinases. *Seed Sci. Res.* **2002**, *12*, 217–230. [[CrossRef](#)]
60. Lee, K.H.; Park, S.W.; Kim, Y.J.; Koo, Y.J.; Song, J.T.; Seo, H.S. Grain width 2 (GW2) and its interacting proteins regulate seed development in rice (*Oryza sativa* L.). *Bot. Stud.* **2018**, *59*, 23. [[CrossRef](#)] [[PubMed](#)]
61. Amer Hamzah, M.; Mohd Kasim, N.A.; Shamsuddin, A.; Mustafa, N.; Mohamad Rusli, N.I.; Teh, C.Y.; Ho, C.L. Nucleotide variations of 9-cis-epoxycarotenoid dioxygenase 2 (NCED2) and pericarp coloration genes (Rc and Rd) from upland rice varieties. *3 Biotech* **2020**, *10*, 105. [[CrossRef](#)] [[PubMed](#)]
62. Sano, N.; Marion-Poll, A. ABA metabolism and homeostasis in seed dormancy and germination. *Int. J. Mol. Sci.* **2021**, *22*, 5069. [[CrossRef](#)] [[PubMed](#)]
63. Matilla, A.J.; Carrillo-Barral, N.; Rodríguez-Gacio, M.d.C. An update on the role of NCED and CYP707A ABA metabolism genes in seed dormancy induction and the response to after-ripening and nitrate. *J. Plant Growth Regul.* **2015**, *34*, 274–293. [[CrossRef](#)]

64. Shoeva, O.Y.; Mock, H.P.; Kukoeva, T.V.; Börner, A.; Khlestkina, E.K. Regulation of the flavonoid biosynthesis pathway genes in purple and black grains of *Hordeum vulgare*. *PLoS ONE* **2016**, *11*, e0163782. [[CrossRef](#)]
65. Khlestkina, E. The adaptive role of flavonoids: Emphasis on cereals. *Cereal Res. Commun.* **2013**, *41*, 185–198. [[CrossRef](#)]
66. Rojas Rodas, F.; Rodriguez, T.O.; Murai, Y.; Iwashina, T.; Sugawara, S.; Suzuki, M.; Nakabayashi, R.; Yonekura-Sakakibara, K.; Saito, K.; Kitajima, J.; et al. Linkage mapping, molecular cloning and functional analysis of soybean gene Fg2 encoding flavonol 3-O-glucoside (1→6) rhamnosyltransferase. *Plant Mol. Biol.* **2014**, *84*, 287–300. [[CrossRef](#)]
67. Ma, D.; Xu, B.; Feng, J.; Hu, H.; Tang, J.; Yin, G.; Xie, Y.; Wang, C. Dynamic Metabolomics and Transcriptomics Analyses for Characterization of Phenolic Compounds and Their Biosynthetic Characteristics in Wheat Grain. *Front. Nutr.* **2022**, *9*, 844337. [[CrossRef](#)]
68. Zhang, Y.; Zhang, H.; Zhao, H.; Xia, Y.; Zheng, X.; Fan, R.; Tan, Z.; Duan, C.; Fu, Y.; Li, L.; et al. Multi-omics analysis dissects the genetic architecture of seed coat content in *Brassica napus*. *Genome Biol.* **2022**, *23*, 86. [[CrossRef](#)]
69. Hong, M.; Hu, K.; Tian, T.; Li, X.; Chen, L.; Zhang, Y.; Yi, B.; Wen, J.; Ma, C.; Shen, J.; et al. Transcriptomic analysis of seed coats in yellow-seeded *Brassica napus* reveals novel genes that influence proanthocyanidin biosynthesis. *Front. Plant Sci.* **2017**, *8*, 1674. [[CrossRef](#)]
70. Sun, L.; Huang, S.; Sun, G.; Zhang, Y.; Hu, X.; Nevo, E.; Peng, J.; Sun, D. SNP-based association study of kernel architecture in a worldwide collection of durum wheat germplasm. *PLoS ONE* **2020**, *15*, e0229159. [[CrossRef](#)]
71. Kiseleva, A.A.; Leonova, I.N.; Pshenichnikova, T.A.; Salina, E.A. Dissection of novel candidate genes for grain texture in Russian wheat varieties. *Plant Mol. Biol.* **2020**, *104*, 219–233. [[CrossRef](#)] [[PubMed](#)]
72. Goel, S.; Singh, K.; Singh, B.; Grewal, S.; Dwivedi, N.; Alqarawi, A.A.; Abd_Allah, E.F.; Ahmad, P.; Singh, N.K. Analysis of genetic control and QTL mapping of essential wheat grain quality traits in a recombinant inbred population. *PLoS ONE* **2019**, *14*, e0200669. [[CrossRef](#)] [[PubMed](#)]
73. Komyshev, E.; Genaev, M.; Afonnikov, D. Evaluation of the SeedCounter, a mobile application for grain phenotyping. *Front. Plant Sci.* **2017**, *7*, 1990. [[CrossRef](#)] [[PubMed](#)]
74. Zdilla, M.J.; Hatfield, S.A.; McLean, K.A.; Cyrus, L.M.; Laslo, J.M.; Lambert, H.W. Circularity, solidity, axes of a best fit ellipse, aspect ratio, and roundness of the foramen ovale: A morphometric analysis with neurosurgical considerations. *J. Craniofacial Surg.* **2016**, *27*, 222. [[CrossRef](#)] [[PubMed](#)]
75. Gowda, S.N.; Yuan, C. ColorNet: Investigating the importance of color spaces for image classification. In *Computer Vision—ACCV 2018, Proceedings of the Asian Conference on Computer Vision, Perth, Australia, 2–6 December 2018*; Springer: Berlin/Heidelberg, Germany, 2018.
76. Busin, L.; Vandenbroucke, N.; Macaire, L. *Volume 151 of Advances in Imaging and Electron Physics, Chapter Chapter 2: Color Spaces and Image Segmentation*; Elsevier Inc.: Amsterdam, The Netherlands, 2008.
77. Cieplinski, L. MPEG-7 color descriptors and their applications. In *Computer Analysis of Images and Patterns, Proceedings of the 9th International Conference, CAIP 2001, Warsaw, Poland, September 2001*; Springer: Berlin/Heidelberg, Germany, 2001.
78. Hammer, Ø.; Harper, D.A.; Ryan, P.D. PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* **2001**, *4*, 9.
79. Soleimani, B.; Lehnert, H.; Keilwagen, J.; Plieske, J.; Ordon, F.; Naseri Rad, S.; Ganai, M.; Beier, S.; Perovic, D. Comparison between core set selection methods using different Illumina marker platforms: A case study of assessment of diversity in wheat. *Front. Plant Sci.* **2020**, *11*, 1040. [[CrossRef](#)]
80. Wang, S.; Wong, D.; Forrest, K.; Allen, A.; Chao, S.; Huang, B.E.; Maccaferri, M.; Salvi, S.; Milner, S.G.; Cattivelli, L.; et al. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* **2014**, *12*, 787–796. [[CrossRef](#)]
81. Allen, A.M.; Winfield, M.O.; BurrIDGE, A.J.; Downie, R.C.; Benbow, H.R.; Barker, G.L.; Wilkinson, P.A.; Coghill, J.; Waterfall, C.; Davassi, A.; et al. Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). *Plant Biotechnol. J.* **2017**, *15*, 390–401. [[CrossRef](#)]
82. Arif, M.A.R.; Afzal, I.; Börner, A. Genetic Aspects and Molecular Causes of Seed Longevity in Plants—A Review. *Plants* **2022**, *11*, 598. [[CrossRef](#)]
83. Agacka-Mołodoch, M.; Arif, M.A.R.; Lohwasser, U.; Doroszewska, T.; Qualset, C.O.; Börner, A. The inheritance of wheat grain longevity: A comparison between induced and natural ageing. *J. Appl. Genet.* **2016**, *57*, 477–481. [[CrossRef](#)]
84. Arif, R.; Nagel, M.; Neumann, K.; Kobiljski, B.; Lohwasser, U.; Börner, A. Genetic studies of seed longevity in hexaploid wheat using segregation and association mapping approaches. *Euphytica* **2012**, *186*, 1–13. [[CrossRef](#)]
85. Sgarbi, C.; Malbrán, I.; Saldúa, L.; Lori, G.A.; Lohwasser, U.; Arif, M.A.R.; Börner, A.; Yannicari, M.; Castro, A.M. Mapping Resistance to Argentinean Fusarium (Graminearum) Head Blight Isolates in Wheat. *Int. J. Mol. Sci.* **2021**, *22*, 13653. [[CrossRef](#)] [[PubMed](#)]
86. Agacka-Mołodoch, M.; Rehman Arif, M.A.; Lohwasser, U.; Doroszewska, T.; Lewis, R.S.; Börner, A. QTL analysis of seed germination traits in tobacco (*Nicotiana tabacum* L.). *J. Appl. Genet.* **2021**, *62*, 441–444. [[CrossRef](#)] [[PubMed](#)]
87. Meng, L.; Li, H.; Zhang, L.; Wang, J. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* **2015**, *3*, 269–283. [[CrossRef](#)]

88. McIntosh, R.A.; Yamazaki, Y.; Dubcovsky, J.; Rogers, W.J.; Morris, C.F.; Sommers, D.J. Catalogue of gene symbols for wheat: 2008. In Proceedings of the 11th International Wheat Genetics, Brisbane, Australia, 24–29 August 2008; Appels, R., Eastwood, R., Lagudah, E., Langridge, Mackay, M., McIntyre, L., Eds.; Sydney University Press: Sydney, Australia, 2008.
89. Gu, Z.; Gu, L.; Eils, R.; Schlesner, M.; Brors, B. Circlize Implements and Enhances Circular Visualization in R. *Bioinformatics* **2014**, *30*, 2811–2812. [[CrossRef](#)]
90. Ochiai, A. Zoogeographic studies on the soleoid fishes found in Japan and its neighbouring regions. *Bull. Jpn. Soc. Sci. Fish.* **1957**, *22*, 526–530. [[CrossRef](#)]
91. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [[CrossRef](#)]
92. Zhu, T.; Wang, L.; Rimbart, H.; Rodriguez, J.C.; Deal, K.R.; De Oliveira, R.; Choulet, F.; Keeble-Gagnère, G.; Tibbits, J.; Rogers, J.; et al. Optical maps refine the bread wheat *Triticum aestivum* cv. *Chin. Spring Genome Assembly. Plant J.* **2021**, *107*, 303–314.
93. Tello-Ruiz, M.K.; Naithani, S.; Gupta, P.; Olson, A.; Wei, S.; Preece, J.; Jiao, Y.; Wang, B.; Chougule, K.; Garg, P.; et al. Gramene 2021: Harnessing the power of comparative genomics and pathways for plant research. *Nucleic Acids Res.* **2021**, *49*, D1452–D1463. [[CrossRef](#)]
94. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [[CrossRef](#)]
95. Borrill, P.; Ramirez-Gonzalez, R.; Uauy, C. expVIP: A customizable RNA-seq data analysis and visualization platform. *Plant Physiol.* **2016**, *170*, 2172–2186. [[CrossRef](#)] [[PubMed](#)]
96. Hamada, K.; Hongo, K.; Suwabe, K.; Shimizu, A.; Nagayama, T.; Abe, R.; Kikuchi, S.; Yamamoto, N.; Fujii, T.; Yokoyama, K.; et al. OryzaExpress: An integrated database of gene expression networks and omics annotations in rice. *Plant Cell Physiol.* **2011**, *52*, 220–229. [[CrossRef](#)] [[PubMed](#)]
97. Kanehisa, M.; Sato, Y.; Morishima, K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* **2016**, *428*, 726–731. [[CrossRef](#)] [[PubMed](#)]
98. Aramaki, T.; Blanc-Mathieu, R.; Endo, H.; Ohkubo, K.; Kanehisa, M.; Goto, S.; Ogata, H. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **2020**, *36*, 2251–2252. [[CrossRef](#)] [[PubMed](#)]