





## Article

# Genetic Mapping of Flavonoid Grain Pigments in Durum Wheat

Natalia Sgaramella <sup>1</sup>, Domenica Nigro <sup>1</sup> , Antonella Pasqualone <sup>1</sup> , Massimo Antonio Signorile <sup>1</sup>, Barbara Laddomada <sup>2</sup> , Gabriella Sonnante <sup>3</sup> , Emanuela Blanco <sup>3</sup> , Rosanna Simeone <sup>1,\*</sup> and Antonio Blanco <sup>1,\*</sup> 

<sup>1</sup> Department of Soil, Plant and Food Sciences (DiSSPA), University of Bari Aldo Moro, Via Amendola 165/A, 70126 Bari, Italy; natalia.sgaramella@gmail.com (N.S.); domenica.nigro@uniba.it (D.N.); antonella.pasqualone@uniba.it (A.P.); massimoantonio.signorile@uniba.it (M.A.S.)

<sup>2</sup> Institute of Sciences of Food Production (ISPA), National Research Council (CNR), Via Prov.le Monteroni, 73100 Lecce, Italy; barbara.laddomada@ispa.cnr.it

<sup>3</sup> Institute of Biosciences and Bioresources, National Research Council (CNR), Via Amendola 165/A, 70126 Bari, Italy; gabriella.sonnante@ibbr.cnr.it (G.S.); emanuela.blanco@ibbr.cnr.it (E.B.)

\* Correspondence: rosanna.simeone@uniba.it (R.S.); antonio.blanco@uniba.it (A.B.); Tel.: +39-080-5442993 (R.S.)

**Abstract:** Pigmented cereal grains with high levels of flavonoid compounds have attracted the attention of nutritional science backing the development of functional foods with claimed health benefits. In this study, we report results on the genetic factors controlling grain pigmentation in durum wheat using a segregant population of recombinant inbred lines (RILs) derived from a cross between an Ethiopian purple grain accession and an Italian amber grain cultivar. The RIL population was genotyped by the wheat 25K SNP array and phenotyped for total anthocyanin content (TAC), grain color, and the L\*, a\*, and b\* color index of wholemeal flour, based on four field trials. The mapping population showed a wide variation for the five traits in the different environments, a significant genotype x environment interaction, and high heritability. A total of 5942 SNP markers were used for constructing the genetic linkage map, with an SNP density ranging from 1.4 to 2.9 markers/cM. Two quantitative trait loci (QTL) were identified for TAC mapping on chromosome arms 2AL and 7BS in the same genomic regions of two detected QTL for purple grain. The interaction between the two QTL was indicative of an inheritance pattern of two loci having complementary effects. Moreover, two QTL for red grain color were detected on chromosome arms 3AL and 3BL. The projection of the four QTL genomic regions on the durum wheat Svevo reference genome disclosed the occurrence of the candidate genes *Pp-A3*, *Pp-B1*, *R-A1*, and *R-B1* involved in flavonoid biosynthetic pathways and encoding of transcription factors bHLH (*Myc-1*) and MYB (*Mpc1*, *Myb10*), previously reported in common wheat. The present study provides a set of molecular markers associated with grain pigments useful for the selection of essential alleles for flavonoid synthesis in durum wheat breeding programs and enhancement of the health-promoting quality of derived foods.

**Keywords:** purple grains; wheat; grain color; nutritional quality; anthocyanins; transcription factors; regulatory genes; candidate genes



**Citation:** Sgaramella, N.; Nigro, D.; Pasqualone, A.; Signorile, M.A.; Laddomada, B.; Sonnante, G.; Blanco, E.; Simeone, R.; Blanco, A. Genetic Mapping of Flavonoid Grain Pigments in Durum Wheat. *Plants* **2023**, *12*, 1674. <https://doi.org/10.3390/plants12081674>

Academic Editors: Yong-Gu Cho, Yu-jin Jung and Me-Sun Kim

Received: 14 March 2023

Revised: 10 April 2023

Accepted: 12 April 2023

Published: 17 April 2023



**Copyright:** © 2023 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

Wheat is one of the major cereal crops, grown on about 210 million hectares all over the world, and represents a primary source of carbohydrates and proteins for the human population. The major cultivated wheat species is common wheat (*Triticum aestivum* L. ssp. *aestivum*), whereas durum wheat (*T. turgidum* L. ssp. *durum*) accounts for about 6% of the total cultivated wheat area and 5% of total wheat grain production [1]. Durum wheat, more adapted to semi-arid climates than common wheat, is mainly grown in countries of the Mediterranean basin, in the Middle East, USA, Canada, Mexico, and some areas in Kazakhstan, Australia, India, and Argentina. Quality traits of durum grain such as protein

content and gluten strength make durum wheat more suitable for preparing typical end products consumed worldwide, such as pasta, couscous, bulghur, and several types of leavened and flat breads [2].

Most of the modern durum cultivars are characterized by grains with yellow to amber color as a result of the intense breeding activity for this trait over the last decades to improve the yellow color of semolina and end products generally preferred by consumers. Several quantitative trait loci (QTL) and candidate genes for the carotenoid content were detected and successfully deployed in durum wheat breeding programs to improve the yellow color and the nutritional value of end products [3].

More recently, other wheat grain pigments, such as flavonoids (i.e., anthocyanins, proanthocyanidins and phlobafenes), have attracted the attention of nutritional science backing the development of functional foods with claimed health benefits [4]. Several lines of evidence suggest that a regular intake of flavonoid-rich food contributes to the reduction of overweight and obesity as well as several associated chronic diseases [5]. The functional properties of flavonoids are ascribed to their antioxidant activity, which is associated with some health benefits both for human and animal populations [6,7].

Among the natural sources of flavonoids, some wild wheats and tetraploid wheat landraces and varieties are characterized by red, brown, purple, black, and blue grains due to the accumulation of different flavonoid compounds in the bran and/or aleurone layer [8]. Several studies have been concerned with the development and production of purple wheat products, including bread, pancakes, biscuits, porridge, crackers, chapati, fresh and dried pasta, and beer, and their characterization and evaluation as functional foods [9–14].

Studies on wheat-colored grains started in the 19th century by using the red and purple characters as morphological markers to estimate the wheat cross-pollination rate, gamete transmission, production of hybrid seeds, as well as the identification of double haploids from anther culture, histological observation, and pigment localization [15]. The genetic control of purple grain color, determined by the presence of anthocyanins in the pericarp, was studied in common wheat, with reported results from one or two independent dominant genes to two complementary dominant loci for purple grain [8,16,17]. The use of molecular markers supported the genetic control of two dominant genes with complementary effects, *Pp-A3* and *Pp-B1*, located on chromosomes 2A and 7B, respectively [18]. In common wheat, subsequent studies confirmed *Pp-A3* mapping on the centromeric region of chromosome 2A and the presence of three homoeoloci, namely *Pp-A3*, *Pp-B3*, and *Pp-D1*, on the short arm of homoeologous chromosomes 7A, 7B, and 7D, respectively [19,20]. Red grain color, determined by the presence of proanthocyanidins and catechins in the aleurone layer of the grain, was found to be controlled in common wheat by homoeoloci *R-A1*, *R-B1*, and *R-D1*, located on the long arm of chromosomes 3A, 3B, and 3D, respectively [21]. It was found that the dominant alleles have additive effects, while a single locus is sufficient to confer the red color on the grains [22]. The dark blue color of the grains is determined by a high content of delphinidins in the aleurone layer [23]. The blue color is controlled by the dominant gene *Ba1*, transferred from chromosome 4e1 of *Thinopyrum ponticum* (Podp.) into the long arm of chromosome 4B [24]. A second gene for the blue grain (*Ba2*) was transferred to common wheat from *T. boeoticum* as a disomic substitution of 4A (4AmL) [25].

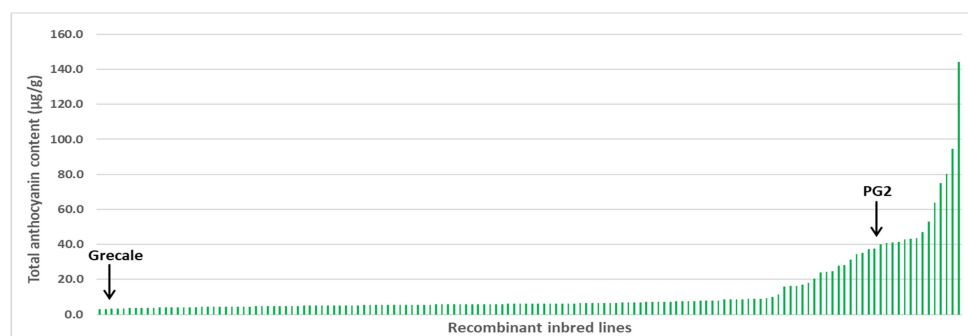
The purple grain color was primarily found in tetraploid wheat populations grown on limited areas in Ethiopia and then transferred to common wheat in early breeding programs [26]. Studies on the inheritance of grain pigments concerned F<sub>3</sub> segregating progenies of common wheat, with phenotypic assessment of grain color visually scored; only a few studies focused on durum wheat [19,27]. Therefore, the aim of this work is to study the genetic control of grain pigmentation in durum wheat. For this purpose, a segregant population of recombinant inbred lines (RIL), derived from crossing an Ethiopian purple accession of durum wheat and a modern durum variety, was used. The mapping population was genotyped with SNP markers and evaluated in four environments for total anthocyanin content (TAC), grain color, and the L\*, a\*, b\* color index of whole meal flour.

Moreover, a candidate gene analysis was conducted to uncover putative genes involved in the regulation of flavonoid biosynthesis. The reported results and the identified markers linked to TAC and grain color could support marker-assisted selection (MAS) in durum wheat breeding programs aimed to improve the quality of durum wheat grains and derived foods and to improve sustainable food and nutritional security.

## 2. Results

### 2.1. Phenotypic Variations of Total Anthocyanin Content (TAC) and Grain Color

The durum wheat RIL population obtained by crossing PG2 × Grecale and the two parental lines were evaluated for TAC and grain color in four replicated field trials carried out at Valenzano (Italy) in 2020, 2021, and 2022 and at Bari (Italy) in 2021. Highly significant differences ( $p < 0.01$ ) among RIL genotypes in each of the two replicated trials were detected by analysis of variance for TAC (Table S1); the combined analysis across environments revealed significant differences for years, genotypes, and environment × genotypes interaction (Table S2). Table 1 reports the mean of the parental lines and the mean, standard error, range, genetic variance, and broad-sense heritability of the RILs in each environment. The parental lines were always significantly different for TAC in each environment; the Ethiopian line PG2 always had higher values (from 25.83 to 52.05  $\mu\text{g/g}$ ) than those of Grecale (from 1.41 to 4.29  $\mu\text{g/g}$ ). High broad-sense heritability values were found in the two replicated trials (0.81 in VAL\_2022 and 0.91 in VAL\_2021) and across environments (0.94). The RIL population showed mean values (8.06 to 17.91  $\mu\text{g/g}$ ) that were nearer to the means of Grecale than to those of PG2, while transgressive segregation was observed at high values falling outside the range of the parental lines in all environments. The frequency distributions (Figures 1 and S1) were shifted toward low TAC values, and significant deviations from normal distribution were detected for the four trials and across environments. The normality of the original TAC values was not improved by the square root, arcsine, and log transformation (data not shown). Significant correlations for TAC were observed among the four field trials (Table S3), with correlation coefficients ranging from  $r = 0.90^{**}$  to  $r = 0.95^{**}$ .



**Figure 1.** Total anthocyanin content in the PG2 × Grecale RIL population and parental lines. The reported values are the mean of four field trials carried out in southern Italy across 2020 and 2022.

Grain color was visually determined by using a 0–3 scale based on the intensity of the color (0 = amber, 1 = red, 2 = red-brownish, 3 = dark purple) (Figure 2), with the parental lines PG2 and Grecale always showing 0 and 3 scores, respectively. The grain color ranged from 0 to 3 in the whole segregant population, with values shifted toward low scores and with a high frequency of score 1 (from  $n = 64$  to  $n = 88$  in the four trials) (Figure 3). The color scores were highly consistent among the four experiments (Figure S3) with correlation coefficients, ranging from  $r = 0.82^{**}$  to  $r = 0.88^{**}$ , indicating a relatively low genotype × environment interaction (Table S3).

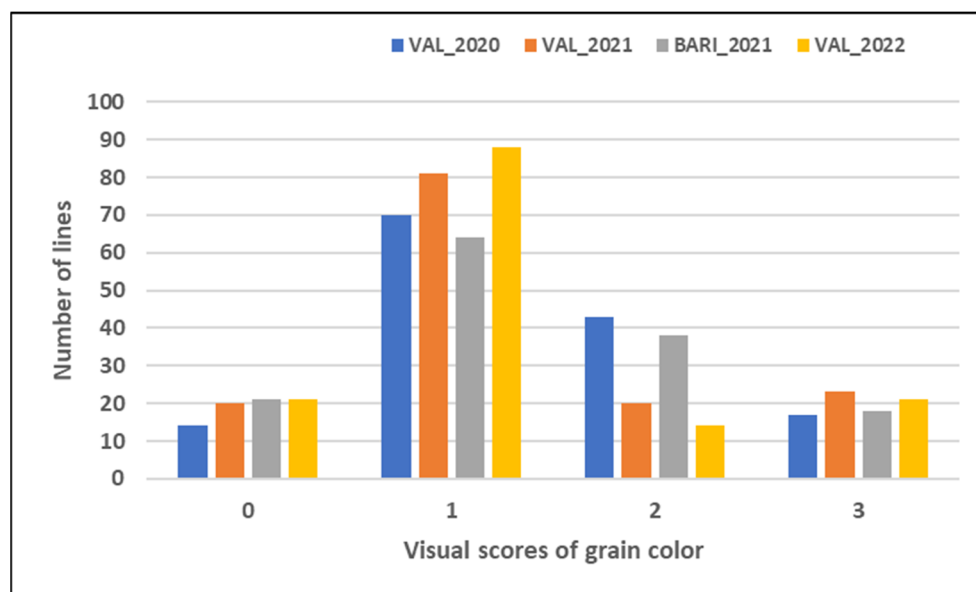
**Table 1.** Descriptive statistics and heritability of total anthocyanin content ( $\mu\text{g/g}$ ), grain color (0–3 scores), and  $L^*$ ,  $a^*$ , and  $b^*$  color values of the recombinant inbred line population derived from the PG2  $\times$  Grecale cross grown across four environments.

Trait	Environment	Durum Wheat Parental Lines				RIL Population			
		PG2	Grecale	Mean	Standard Error	Minimum	Maximum	Variance	$h^2$
Anthocyanin content ( $\mu\text{g/g}$ )	VAL_2020	25.83 **	3.92	8.86	0.69	1.50	58.39	68.71	$^{\circ}$
	VAL_2021	52.05 **	4.29	17.91	2.37	2.08	198.56	808.78	0.91
	BA_2021	42.72 **	1.41	8.06	1.04	0.70	103.61	154.02	$^{\circ}$
	VAL_2022	39.32 **	3.05	12.10	1.48	2.66	150.33	316.74	0.81
	Across environments	39.98 **	3.17	13.33	1.60	2.96	144.30	370.60	0.94 $^{\circ\circ}$
Grain color (0–3)	VAL_2020	3 **	0	1.46	0.82	0.00	3.00	0.67	0.93
	VAL_2021	3 **	0	1.34	0.90	0.00	3.00	0.80	0.97
	BA_2021	3 **	0	1.36	0.87	0.00	3.00	0.75	0.95
	VAL_2022	3 **	0	1.26	0.84	0.00	3.00	0.71	0.95
	Across environments	3 **	0	1.37	0.82	0.00	3.00	0.68	0.97
$L^*$ color index	VAL_2020	73.83 **	81.9	78.6	1.8	73.7	83.1	3.2	0.88
	VAL_2021	75.4 **	84.0	80.7	2.2	74.3	84.7	4.7	0.97
	BA_2021	73.6 **	82.5	79.4	2.1	73.6	83.9	4.4	0.79
	VAL_2022	76.1 **	83.7	80.6	1.9	75.1	84.8	3.7	0.96
	Across environments	74.7 **	83.0	79.8	1.9	74.3	84.1	3.6	0.96
$a^*$ color index	VAL_2020	2.36 **	0.36	1.57	0.44	0.09	2.44	0.20	0.81
	VAL_2021	2.60 **	-0.11	1.38	0.58	-0.08	3.02	0.34	0.92
	BA_2021	2.64 **	0.62	1.65	0.50	0.43	2.91	0.25	0.86
	VAL_2022	2.38 **	0.02	1.36	0.50	0.04	2.71	0.25	0.91
	Across environments	2.49 **	0.22	1.49	0.47	0.19	2.55	0.22	0.94
$b^*$ color index	VAL_2020	11.7 **	16.8	14.6	1.3	10.6	18.2	1.7	0.92
	VAL_2021	12.1 **	18.2	15.3	1.8	8.9	18.9	3.1	0.95
	BA_2021	11.9 **	18.5	15.6	1.6	10.2	19.5	2.6	0.95
	VAL_2022	12.5 **	18.3	15.5	1.6	9.6	18.3	2.4	0.96
	Across environments	12.0 **	17.9	15.2	1.5	9.9	18.7	2.3	0.97

\*\* Significant at  $p < 0.01$  with a Student's  $t$  test.  $^{\circ}$   $h^2$  estimate is missing because TAC analysis was performed on mixed samples of three replications.  $^{\circ\circ}$   $h^2$ , broad-sense heritability determined across two environments (VAL\_2021 and VAL\_2022).



**Figure 2.** Seed color of test samples: (a) amber (score 0); (b) red (score 1); (c) red-brownish (score 2); (d) purple (score 3).



**Figure 3.** Frequency distribution of grain color scores evaluated using a 0 to 3 scale based on the deepness of the color (amber, light red, red-brownish and dark purple) in the biparental mapping population PG2 × Grecale grown in four environments (VAL\_2020, VAL\_2021, BA\_2021, VAL\_2022).

Color values  $L^*$ ,  $a^*$ , and  $b^*$  were also measured in each trial to investigate their possible relationships with TAC (Table 1). The parental lines were always different for each color value, while the RIL population showed mid-parental means. High broad-heritability values were observed for each color index. Correlation coefficients among all five traits (TAC, grain color, and  $L^*$ ,  $a^*$ , and  $b^*$ ) in each environment are reported in Table S4; correlation coefficients across environments are shown in Table 2. Across environments,  $L^*$  and  $b^*$  were found significantly negatively correlated with TAC and grain color, while  $a^*$  showed positive correlation with TAC and grain color. TAC and grain color demonstrated a significant positive correlation ( $r = 0.63^{**}$ ).

**Table 2.** Simple correlation coefficients among  $L^*$ ,  $a^*$ , and  $b^*$  color values, total anthocyanin content (TAC) (Pearson correlation coefficients), and grain color (Sperman rank correlation coefficients) from combined data of the recombinant inbred line population derived from the PG2 × Grecale cross grown in four environments.

Trait	$L^*$	$a^*$	$b^*$	Grain Color	TAC
$L^*$	1.00 **				
$a^*$	−0.85 **	1.00 **			
$b^*$	0.77 **	−0.76 **	1.00 **		
Grain color	−0.87 **	0.72 **	−0.72 **	1.00 **	
TAC	−0.73 **	0.47 **	−0.72 **	0.63 **	1.00 **

\*\* : significant at 0.01  $p$ .

## 2.2. Genetic Linkage Map

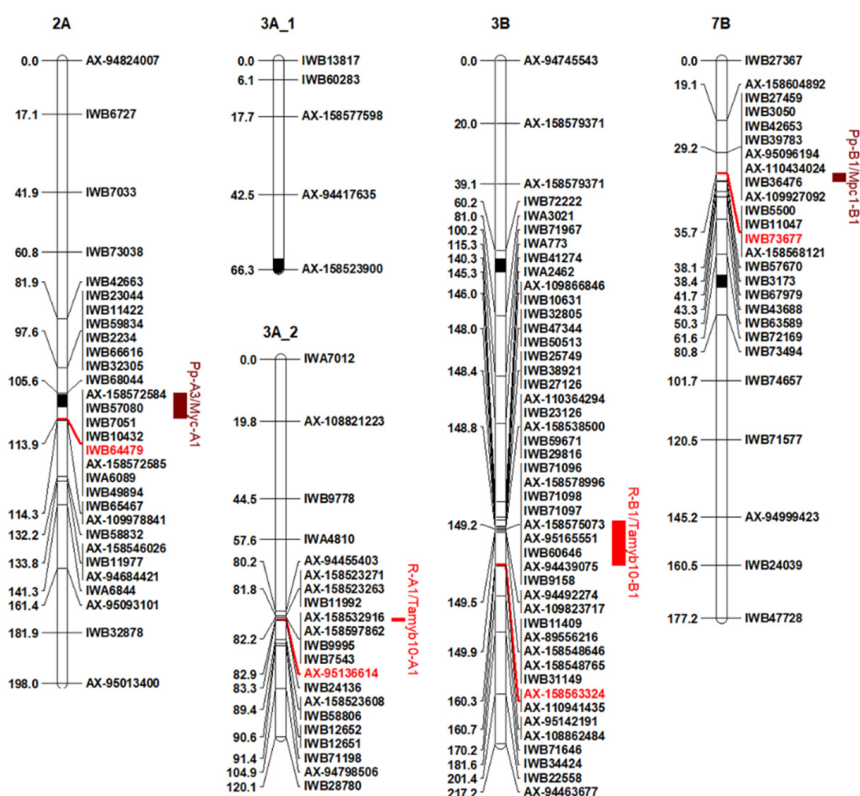
The RIL population and the parental lines were genotyped by the wheat 25K SNP Illumina Infinium Array [28], including 24,145 single nucleotide polymorphisms (SNPs). After SNP data quality evaluation and control, a total of 5942 (24.6%) polymorphic markers were used to obtain a genetic linkage map via the software IciMapping v. 4.2 [29]. The ultimate map comprised 11 LGs assigned to individual chromosomes and 6 LGs assigned to 3 chromosomes (2B\_1, 2B\_2, 3A\_1, 3A\_2, 5B\_1, 5B\_2) (Table S5). The total length of the A genome chromosomes was 1380.9 cM and included 2600 markers, while the total length of the B genome chromosomes was 1309.5 cM and included 3234 markers. Individual chromosomes had a length ranging from 166.1 cM (4B) to 240.2 cM (7A), and a



number of markers from 265 (4B) to 604 (2B). The SNP density varied from a minimum of 1.4 markers/cM for chromosome 4A and a maximum of 2.9 markers/cM for chromosomes 5B and 6B.

### 2.3. Detection of QTL for Total Anthocyanin Content (TAC)

QTL for TAC were determined via the Inclusive Composite Interval Mapping (ICIM) method [29], and QTL detected in at least two environments with LOD values > 3.0 were considered. Putative QTL for TAC and their main characteristics (LOD score, genetic and physical map position, closest markers, phenotypic variance explained, additive effects) in individual environments and across environments are reported in Table 3 and Figure 4. Two QTL (*QTAC.mgb-2A* and *QTAC.mgb-7B*) localized on chromosome arms 2AL and 7BS were identified via the ICIM-ADD method. These QTL were consistent in all four environments, and each explained 8.8–13.4% (*QTAC.mgb-2A*) and 11.3–16.9% (*QTAC.mgb-7B*) of TAC variation. Both QTL were also significant across environments with a phenotypic variance explained (PVE) of 10.5% and 16.7%, respectively. The positive alleles (elevated TAC) were both contributed by Ethiopian line PG2, with allelic effects ranging from 2.9  $\mu\text{g/g}$  (*QTAC.mgb-2A* at VAL\_2020) to 11.2 TAC  $\mu\text{g/g}$  (*QTAC.mgb-7B* at VAL\_2021) in the different environments.



**Figure 4.** Schematic representation of four chromosomes of the durum PG2 x Grecale linkage map with positions of QTL for total anthocyanin content (TAC) and grain color. Each chromosome map is represented by the first and the last SNP marker, and by an SNP marker roughly every 20 cM. Markers are indicated on the right side and cM distances on the left side of the bar. QTL are represented by bars on the right of each chromosome bar. The closest marker to each QTL is indicated in red. The black segment on each chromosome bar represents the centromere.

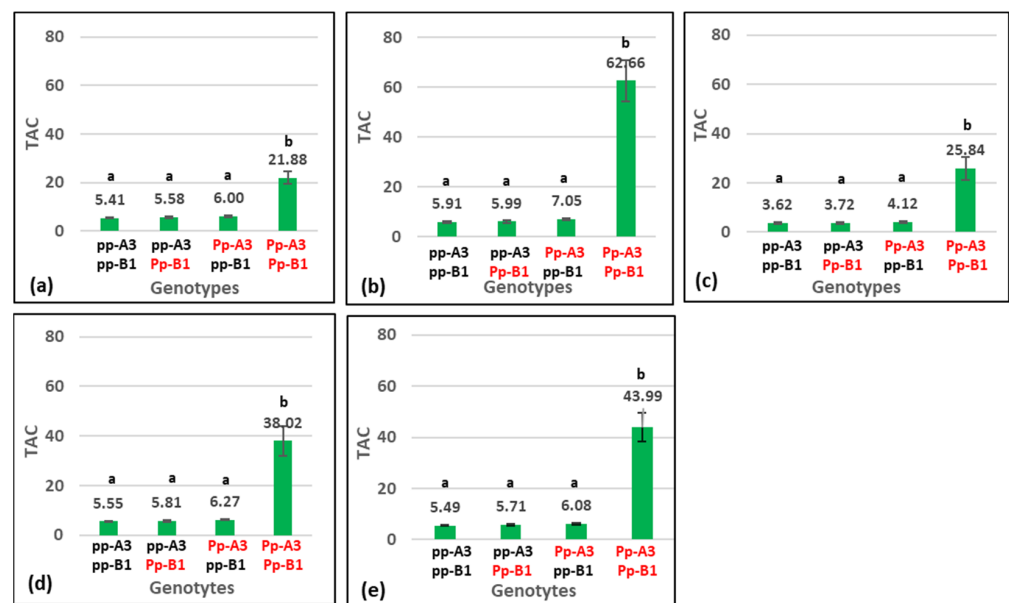
**Table 3.** Quantitative trait loci (QTL) for total anthocyanin content and grain color detected in the RIL population of durum wheat PG2 × Grecale via single-environment and cross-environment QTL analyses (ICIM-ADD).

QTL	Environment	Chrom. Arm	Genetic Position (cM)	Closest Marker	Physical Position (bp) <sup>a</sup>	LOD	PVE (%)	Additive Effect
Total anthocyanin compounds								
<i>QTAC.mgb-2A</i>	VAL_2020	2AL	113	IWB64479	670,992,300	4.3	12.0	2.9
	VAL_2021	2AL	115	AX-109978841	672,230,615	4.7	13.4	9.9
	BA_2021	2AL	113	IWB64479	670,992,300	2.9	8.8	3.7
	VAL_2022	2AL	113	IWB64479	670,992,300	3.6	10.5	5.7
	Across environments	2AL	113	IWB64479	670,992,300	4.6	12.8	6.9
<i>QTAC.mgb-7B</i>	VAL_2020	7BS	39	IWB3173	89,972,586	5.0	14.4	3.2
	VAL_2021	7BS	38	IWB73677	83,850,422	5.9	16.9	11.2
	BA_2021	7BS	38	IWB73677	83,850,422	3.8	11.3	4.2
	VAL_2022	7BS	38	IWB73677	83,850,422	5.1	15.1	6.9
	Across environments	7BS	38	IWB73677	83,850,422	6.0	16.7	7.9
Grain color								
<i>QGc.mgb-2A</i>	BA_2021	2AL	113	IWB64479	670,992,300	2.9	6.3	0.2
	VAL_2021	2AL	118	AX-109978841	672,230,615	5.7	14.8	0.4
	VAL_2022	2AL	118	AX-109978841	672,230,615	5.5	13.6	0.3
	Across environments	2AL	118	AX-109978841	672,230,615	4.9	12.4	0.3
	<i>QGc.mgb-7B</i>	VAL_2020	7BS	38	IWB73677	83,850,422	7.0	18.4
VAL_2021		7BS	38	IWB73677	83,850,422	13.0	27.2	0.5
BA_2021		7BS	38	IWB73677	83,850,422	10.4	23.4	0.4
VAL_2022		7BS	38	IWB73677	83,850,422	13.1	28.7	0.5
Across environments		7BS	38	IWB73677	83,850,422	12.4	27.3	0.4
<i>QGc.mgb-3B</i>	VAL_2020	3BL	161	AX-108862484	775,341,071	4.0	10.1	0.3
	VAL_2021	3BL	148	IWB47344	756,364,388	3.9	6.8	0.2
	BA_2021	3BL	161	AX-108862484	775,341,071	6.8	14.7	0.3
	VAL_2022	3BL	148	IWB47344	756,364,388	3.0	5.5	0.2
	Across environments	3BL	148	IWB47344	756,364,388	4.5	8.5	0.2
Grain color *								
<i>QGc.mgb-3A</i>	VAL_2020	3AL_2	83	AX-95136614	695,008,707	8.4	18.4	0.3
	VAL_2021	3AL_2	83	AX-95136614	695,008,707	4.6	14.2	0.2
	BA_2021	3AL_2	83	AX-95136614	695,008,707	8.6	23.1	0.3
	VAL_2022	3AL_2	83	AX-95136614	695,008,707	5.5	17.6	0.2
	Across environments	3AL_2	83	AX-95136614	695,008,707	7.8	20.5	0.2
<i>QGc.mgb-3B</i>	VAL_2020	3BL	158	AX-158563324	774,808,623	6.4	15.2	0.3
	VAL_2021	3BL	148	IWB47344	756,364,388	6.3	20.0	0.2
	BA_2021	3BL	148	IWB47344	756,364,388	8.2	21.8	0.3
	VAL_2022	3BL	148	IWB47344	756,364,388	5.1	16.2	0.2
	Across environments	3BL	148	IWB47344	756,364,388	8.1	21.3	0.2

<sup>a</sup> The physical location is based on the durum wheat Svevo reference genome v1 [30]. \* QTL analysis on a subset of 110 RILs excluding the lines with anthocyanin compounds. LOD, logarithm of odds; PVE, phenotype variance explained.

The ICIM-ADD method used for biparental populations can reliably detect QTL with additive effects [31], while the detection of epistasis effects is hampered by the complex pattern of epistasis, the large number of model effects, and the lack of efficient statistical methods [32,33]. Considering these issues and knowledge of the inheritance of purple grain color in common wheat and other cereal species [20,34], the interaction between the two detected QTL was analyzed by comparing the TAC means of the four groups of RILs with a different genotype at the two TAC loci. The genotype assignment to each RIL was carried out based on the closest SNP marker to each QTL. Mean values of TAC of the four groups of RILs with different genotypes at *QTAC.mgb-2A* and *QTAC.mgb-7B* in each of the four environments and across environments are reported in Figure 5.

By comparing the four RIL groups using Fisher's LSD test, highly significant differences at  $p < 0.01$  were found between the group with the dominant loci *QTAC.mgb-2A* and *QTAC.mgb-7B* and the other three groups having one or two recessive loci in individual environments and across environments, whereas non-significant differences were detected among the last three groups. The QTL analysis and the interaction between the detected loci are indicative of an inheritance pattern of two loci with complementary effect on TAC. The two QTL most likely correspond to the *Pp3* and *Pp1* genes for purple pericarp identified in common wheat [20], and, according to the catalogue of gene symbols for wheat [21], hereafter will be designated *Pp-A3* and *Pp-B1*.



**Figure 5.** Mean value of total anthocyanin content ( $\mu\text{g/g}$ ) (TAC) of groups of recombinant inbred lines with different genotypes at *Pp3* and *Pp1* loci grown in four environments: (a) Valenzano 2020; (b) Valenzano 2021; (c) Bari 2021; (d) Valenzano 2022; (e) across environments. *Pp-A3* and *Pp-B1* highlighted in red are dominant to *pp-A3* and *pp-B1*, respectively. Bars indicate mean value  $\pm$  standard error. Different letters show significant differences ( $p < 0.05$ ).



#### 2.4. Detection of QTL for Grain Color

Three QTL for grain color were detected on chromosome arms 2AL, 7BS, and 3BL. The first two QTL, *QGc.mgb-2A* and *QGc.mgb-7B*, significant in environments two and four, respectively, and with PVE ranging from 13.6% to 28.7%, were in the same genetic regions of the two detected loci for TAC (*QTAC.mgb-2A* and *QTAC.mgb-7B*) and should correspond to genes *Pp-A3* and *Pp-B1* identified in common wheat (Table 3). This validated the indication that the purple grain color can be attributed to TAC, as also indicated by the significant positive correlation between color intensity and TAC ( $r = 0.70^{***}$ ) (Table 2). The third QTL for grain color (*QGc.mgb-3B*), located on chromosome arm 3BL, was consistent in the four environments and across environments, and explained a phenotypic variance of 6.8–14.7%. This QTL is responsible for the red color categorized 1 and 2 in the current manuscript and should correspond to the locus *R-B1* identified in common wheat as responsible for the red grain color determined by proanthocyanidin and catechin compounds in the testa of seeds [22].

Considering the relatively low content of proanthocyanidin and catechin compounds determining the red color and the high content of TAC determining the dark purple color [35], and that this pigment could affect and partially mask the red color, a second QTL analysis was performed on a RIL subset excluding the homozygous lines for *Pp-A3/Pp-B1*-containing anthocyanin compounds. This further analysis on a subset of 110 RILs confirmed the presence of the red color locus *R-B1* on 3BL and detected a second QTL for the red grain color (*R-A1*) on chromosome arm 3AL. The two QTL were significant in all four environments and across environments, with LOD ranging from 4.6 to 8.6 and PVE from 14.8% to 23.1%. The statistical analysis of the four RIL groups differing in their *R-1* genotype showed significant differences in the double recessive genotype (*r-A1/r-B1*) compared to the other three genotypes (*R-A1/r-B1*, *r-A1/R-B1*, *R-A1/R-B1*) for grain color and for  $L^*$ ,  $a^*$ , and  $b^*$  color values, while no significant difference was found between the three genotypes having one or both dominant *R-1* alleles (Table 4).

**Table 4.** Mean values across four environments of grain color and  $L^*$ ,  $a^*$ , and  $b^*$  color values of groups of recombinant inbred lines with different genotypes at the *R-1* locus. Values followed by the same letter are not significantly different ( $p < 0.001$ ).

Genotype	Grain Color	$L^*$	$a^*$	$b^*$
<i>R-A1/R-B1</i>	1.4 b	79.8 b	1.6 b	15.41 b
<i>R-A1/r-B1</i>	1.2 b	80.0 b	1.5 b	15.63 b
<i>r-A1/R-B1</i>	1.2 b	80.1 b	1.5 b	15.55 b
<i>r-A1/r-B1</i>	0.3 a	82.4 a	0.8 a	16.71 a

#### 2.5. Candidate Genes

The 100 bp sequences of the closest SNP marker to each detected QTL on chromosomes 2A, 3A, 3B, and 7B (Table 3) were used to identify the corresponding genomic physical regions via a BLAST analysis against the durum wheat Svevo reference genome [30] (Table 3). The four identified QTL regions were further investigated within a linkage disequilibrium (LD) window of <5 cM to uncover candidate genes involved in the flavonoid synthetic pathways. Several dozen high-confidence protein-coding genes were found to be annotated, including the occurrence of previously reported candidate genes for purple and red grain color in common wheat. Table 5 reports the detected QTL for TAC and grain color and the corresponding durum wheat candidate genes along with the orthologous genes retrieved from the wild emmer wheat Zavitan reference genome [36] and from the common wheat Chinese Spring reference genome [37] for comparison. Phylogenetic studies were also performed to facilitate durum wheat candidate gene identification and classification.

**Table 5.** Candidate genes for the detected QTL for total anthocyanin content and red grain color in durum wheat (*Triticum turgidum* ssp. *durum*) and related homoeologous genes in wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*) and common wheat (*Triticum aestivum*).

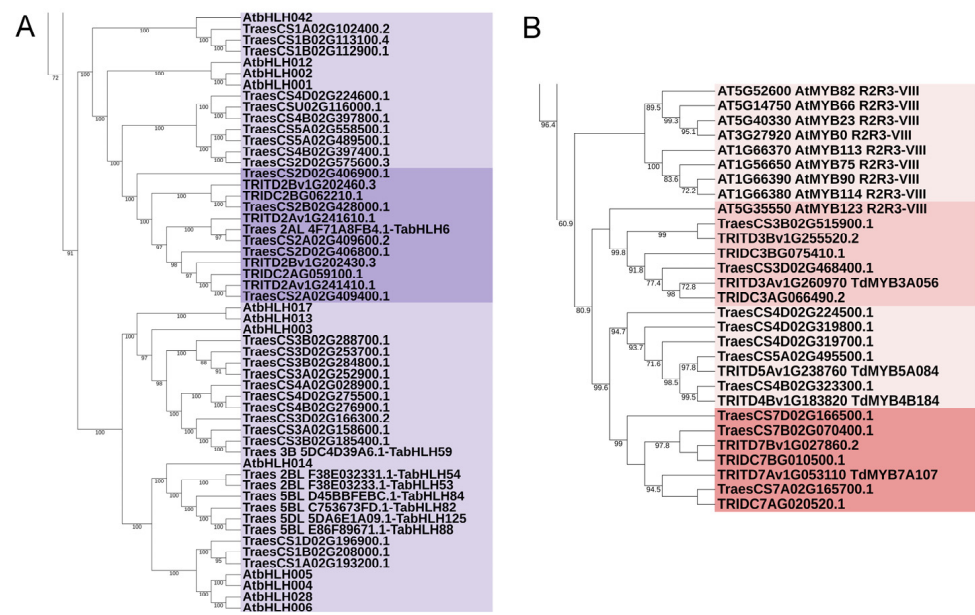
QTL	Species *	Ch.	Gene	<i>Triticum</i> ortholog TF **	ID	Position	Annotation
QTAC.mgb-2A/QGC.mgb-2A	<i>T. t. durum</i>	2A	<i>Pp-A3</i>	<i>Myc-A1.1</i>	<i>TRITD2Av1G241410</i>	661,787,635-661,794,834	bHLH DNA-binding superfamily protein.
	<i>T. t. durum</i>	2A	<i>Pp-A3</i>	<i>Myc-A1.2</i>	<i>TRITD2Av1G241610</i>	662,425,745-662,427,498	bHLH DNA-binding superfamily protein.
	<i>T. t. durum</i>	2B	<i>Pp-B3</i>	<i>Myc-B1.1</i>	<i>TRITD2Bv1G202430</i>	604,115,921-604,121,056	bHLH transcription factor.
	<i>T. t. durum</i>	2B	<i>Pp-B3</i>	<i>Myc-B1.2</i>	<i>TRITD2Bv1G202460</i>	604,213,485-604,217,695	bHLH transcription factor.
	<i>T. t. dicoccoides</i>	2A	<i>Pp-A3</i>	<i>Myc-A1</i>	<i>TRIDC2AG059100</i>	660,742,692-660,747,726	
	<i>T. t. dicoccoides</i>	2B	<i>Pp-B3</i>	<i>Myc-B1</i>	<i>TRIDC2BG062210</i>	611,414,598-611,419,239	
	<i>T. aestivum</i>	2A	<i>Pp-A3</i>	<i>Myc-A1.1</i>	<i>TraesCS2A02G409400</i>	667,010,726-667,015,774	MYC-like regulatory protein. <i>TaMYC1</i> [38]; <i>TaMYC3-A1</i> and <i>TaMYC3-A2</i> [39]; <i>TaPpb1</i> [40].
	<i>T. aestivum</i>	2A	<i>Pp-A3</i>	<i>Myc-A1.2</i>	<i>TraesCS2A02G409600</i>	667,647,089-667,652,609	MYC-like regulatory protein. <i>TaMYC1</i> [38]; <i>TaMYC3-A1</i> and <i>TaMYC3-A2</i> [39]; <i>TaPpm1</i> [40].
	<i>T. aestivum</i>	2B	<i>Pp-B3</i>	<i>Myc-B1</i>	<i>TraesCS2B02G428000</i>	615,352,721-615,356,928	<i>TaMYC3-B</i> [41]
	<i>T. aestivum</i>	2D	<i>Pp-D3</i>	<i>Myc-D1</i>	<i>TraesCS2D02G406900</i>	522,521,462-522,526,938	
<i>T. t. durum</i>	7A	<i>Pp-A1</i>	<i>Mpc1-A1</i>	<i>TRITD7Av1G053110</i>	117,987,002-117,987,868	MYB transcription factor. <i>TdMYB7A107</i> [42].	

Table 5. Cont.

QTL	Species *	Ch.	Gene	<i>Triticum</i> ortholog TF **	ID	Position	Annotation
QTAC.mgb-7B/QTAC.mgb-7B	<i>T. t. durum</i>	7B	<i>Pp-B1</i>	<i>Mpc1-B1</i>	TRITD7Bv1G027860	77,346,121-77,348,220	MYB transcription factor.
	<i>T. t. dicoccoides</i>	7A	<i>Pp-A1</i>	<i>Mpc1-A1</i>	TRIDC7AG020520	118,943,357-118,944,094	
	<i>T. t. dicoccoides</i>	7B	<i>Pp-B1</i>	<i>Mpc1-B1</i>	TRIDC7BG010500	81,313,184-81,314,083	
	<i>T. aestivum</i>	7A	<i>Pp-A1</i>	<i>Mpc1-A1</i>	TraesCS7A02G165700	121,397,856-121,399,015	
	<i>T. aestivum</i>	7B	<i>Pp-B1</i>	<i>Mpc1-B1</i>	TraesCS7B02G070400	77,430,167-77,432,792	<i>TaPpb1</i> [40].
	<i>T. aestivum</i>	7D	<i>Pp-D1</i>	<i>Mpc1-D1</i>	TraesCS7D02G166500	117242813-117245836	
QGc.mgb-3A	<i>T. t. durum</i>	3A	<i>R-A1</i>	<i>Myb10-A1</i>	TRITD3Av1G260970	694,983,038-694,991,683	MYB transcription factor. <i>TdMYB3A056</i> [42].
QGc.mgb-3B	<i>T. t. durum</i>	3B	<i>R-B1</i>	<i>Myb10-B1</i>	TRITD3Bv1G255520	769,811,153-769,814,329	MYB transcription factor.
	<i>T. t. dicoccoides</i>	3A	<i>R-A1</i>	<i>Myb10-A1</i>	TRIDC3AG066490	701,352,578-701,365,988	
	<i>T. t. dicoccoides</i>	3B	<i>R-B1</i>	<i>Myb10-B1</i>	TRIDC3BG075410	771,960,350-771,962,476	
	<i>T. aestivum</i>	3A			-		
	<i>T. aestivum</i>	3B	<i>R-B1</i>	<i>Myb10-B1</i>	TraesCS3B02G515900	757,918,264-757,920,082	<i>Tamyb10-B1</i> [22,43].
	<i>T. aestivum</i>	3D	<i>R-D1</i>	<i>Myb10-B1</i>	TraesCS3D02G468400	570,801,163-570,803,376	<i>Tamyb10-D1</i> [22,43].

\* *T. t. durum* = *Triticum turgidum* ssp. *durum*; *T. t. dicoccoides* = *Triticum turgidum* ssp. *dicoccoides*. TF \*\*, Transcription Factor *Myc-1* designation according to [44]; *Mpc-1* designation according to [45,46]; *Myb10* designation according to [22,43].

Within the chromosomal location of QTAC.mgb-2A/QGc.mgb-2A, a *Pp-A3* gene was identified, TRITD2Av1G241410, annotated as a member of the basic helix-loop-helix (bHLH) DNA-binding protein superfamily involved in flavonoid biosynthetic pathways [44]. A paralogous putative bHLH transcription factor (TF) gene, TRITD2Av1G241610, was also found within the same QTL interval. TRITD2Av1G241410 was found to be homoeologous to TRITD2Bv1G202430 located on the 2B chromosome, which is also a paralog of TRITD2Bv1G202460, another bHLH TF found in the same 2B region at about 98 kb. TRITD2Av1G241610 had no homoeolog on the 2B chromosome. A phylogenetic tree (Figures 6 and S2) was generated by processing these putative durum wheat bHLH TFs together with their corresponding orthologs from wild emmer and bread wheat bHLH TFs [47], along with the recently identified *TaMYC* TFs [38,41], and all the annotated Arabidopsis [48]. The phylogenetic analysis placed the four putative durum wheat bHLH TFs identified here into the cluster containing the majority of MYC TFs; in particular, they clustered with TraesCS2A02G409400 (*TaMYC1*), [44] and with the Arabidopsis *AtbHLH42* [49].



**Figure 6.** Phylogenetic relationships of durum wheat candidate transcription factors (TFs) with their homologs. **(A)** A section of the bHLH phylogenetic tree (complete tree in Figure S2), including the putative durum wheat TFs (TRITD) identified in this study, their wild emmer (TRIDC) and bread wheat (TraesCS) orthologs, and the *Arabidopsis thaliana* (AtbHLH) TFs. Purple: putative MYC TFs; dark purple: group with the Myc1 candidate genes identified here. **(B)** A portion of the R2R3-MYB phylogenetic tree (complete tree in Figure S3), including durum wheat (TRITD), their wild emmer (TRIDC) and bread wheat (TraesCS) ortholog TFs putatively involved in the regulation of phenylpropanoid biosynthesis, and *A. thaliana* (AtMYBs). Brown: putative MYB TFs; medium brown: group with the Myb10 candidate genes; dark brown: group with the Mpc1 candidate genes. Numbers on the branch nodes indicate bootstrap values (only bootstraps > 60 are shown).

A low-confidence gene, *TRITD7Bv1G027860*, was identified in the physical interval underlying the QTL *QTAC.mgb-7B/QGc.mgb-7B*. Interestingly, this gene showed a high sequence similarity to the durum wheat *TRITD7Av1G053110*, annotated as MYB transcription factor *TdMYB7A107* [42].

Candidate gene investigation was also performed for the physical regions of QTL for grain color detected on chromosomes 3A (*QGc.mgb-3A*) and 3B (*QGc.mgb-3B*). The *TRITD3Av1G260970* gene, also reported as *TdMYB3A056* and annotated as an MYB transcription factor [42], was found physically located within the chromosomal interval of *QGc.mgb-3A*. No orthologous gene was reported in common wheat, and no homoeologous gene on the B genome in durum wheat. A BLAST analysis allowed the identification of a high-similarity sequence, annotated as a low-confident gene on chromosome 3B, *TRITD3Bv1G255520*, also a MYB TF, which mapped within the *QGc.mgb-3B* interval. Taking advantage of a recent study on durum wheat MYB TFs [42], the putative durum wheat MYB TFs identified here were subjected to a phylogenetic analysis, including the annotated *Arabidopsis AtMYBs*, the emmer orthologs, and a subset of bread wheat MYB genes putatively involved in the regulation of phenylpropanoid biosynthesis. The durum wheat candidate genes and their wild emmer and bread wheat orthologs clustered with *AtMYB* TFs participating in the regulation of phenylpropanoid production [50] (Figures 6 and S3).

### 3. Discussion

In recent years, durum wheat breeding programs have been focusing on several aspects of grain quality, such as protein content, starch content and composition, and grain color, a trait related to the presence of natural pigments, particularly carotenoids and anthocyanins. Pigmented wheats with red, purple, blue, and black grains have attracted the attention of the food industry for their large content of flavonoid compounds and associated antioxidant activity [51]. In particular, colored whole grains are regarded as promising new functional foods due to their anthocyanin content located in outer layers (pericarp, aleurone, testa) of grains [52]. Numerous *in vivo* and *in vitro* studies have demonstrated the high antioxidative activity of anthocyanins [23,53] and their health effects, including cardio-protection, anti-diabetes, anti-obesity, anti-cancer, and anti-aging-affiliated pathology [51,52,54]. As a result, the utilization of pigmented wheat grains to produce more nutritious and healthy food products has been considered [51]. To achieve this aim, a number of studies developed methods to carefully select the most appropriate milling fractions to produce flours rich in anthocyanins and other bioactive compounds present in the outer layers of the kernel without lowering the quality of end products [55–59].

Genetic, physiological, and biochemical studies have also shown that flavonoid compounds are involved in several biological activities, including plant defense responses to abiotic and biotic stress conditions [60,61]. In particular, anthocyanin-rich genotypes maintain significantly higher dry matter production under salt stress conditions [62]. Moreover, it was found that colored wheat genotypes are able to activate biotic and abiotic stress-responsive genes in response to drought [63], and to increase resistance to seed dormancy and pre-harvest sprouting [39,64]. In fact, anthocyanin biosynthesis and accumulation are enhanced by a number of hormones and environmental factors, including light, UV irradiation, high temperature, and heavy metals, besides drought and salinity [65].

#### 3.1. Trait Variation and Genotype $\times$ Environment Interaction

Total anthocyanin content and composition of pigmented wheat grains and end products vary greatly depending on cultivar genotypes, pedological and climatic factors, agrotechnical management, extraction method, and processing conditions [66–68]. In the current work, an Ethiopian purple durum wheat line was crossed with an Italian amber durum wheat to obtain the RIL mapping population PG2  $\times$  Grecale, evaluated for five traits in four environments. The RIL population showed a wide variation for TAC in the different environments (ranging from 1.50–58.39  $\mu\text{g/g}$  at VAL\_2020 to 2.08–198.56  $\mu\text{g/g}$  at VAL\_2021), while the Ethiopian purple durum line PG2 had an average TAC value ranging from 25.83  $\mu\text{g/g}$  at VAL\_2020 to 52.05  $\mu\text{g/g}$  at VAL\_2021, thus indicating the influence of environmental factors on phenotypic expression of TAC. Despite the significance of genotype  $\times$  environment interaction, TAC heritability showed high values (0.81–0.91 in individual environments and 0.94 across environments measured on a mean basis). TAC was always positively correlated with grain color and  $a^*$ , and negatively correlated with  $L^*$  and  $b^*$  in all environments. The grain color, detected via a 0–3 scale based on color intensity, showed high heritability values (0.70–0.96) and significant interactions with environmental factors, especially for the variation of classes 1 and 2 in the four environments (Figure 3).

#### 3.2. QTL for TAC and Grain Color

Early studies on the genetics of purple grain color in common wheat reported various results by using different purple wheat accessions in crosses with white grains: one dominant gene for purple grain (ratio 3:1 in  $F_2$  populations) [69–71], two duplicate dominant genes (ratio 11:5) [72], two independent genes (15:1) [66,73], and two complementary dominant genes (ratio 9:7) [72,74]. The use of molecular markers supported the genetic control of two dominant genes with complementary effects, *Pp-A3* and *Pp-B1*, located on chromosomes 2A and 7B, respectively [18]. Subsequent studies confirmed the *Pp-A3* mapping on the centromeric region of chromosome 2A and the presence of three homoeoloci,



*Pp-A3*, *Pp-B3*, *Pp-D1*, on the short arms of the homoeologous chromosomes 7A, 7B, and 7D, respectively [19,20].

Most of these investigations were carried out in common wheat by visually scoring the grain color on F<sub>3</sub> segregating populations. In contrast, very few genetic studies have considered the pigmented durum wheat [19,27]. The present research was conducted on an RIL population of durum wheat genotyped by the SNP 25K chip array, and simultaneously assessing the TAC, grain color, and the L\*, a\*, and b\* color values for their possible correlation. QTL analysis identified, and validated, in durum wheat the two *Pp-A3* and *Pp-A1* loci, located on the long arm of chromosome 2A close to the centromere and on the short arm of chromosome 7B, respectively. The subsequent genetic analysis of the interaction of the two loci showed the absence of anthocyanin compounds in the three groups of RILs with a homozygous recessive genotype or with one dominant allele (*pp-A3/pp-B1*, *Pp-A3/pp-B1*, *pp-A3/Pp-B1*) and a high TAC content in the group of RILs with both dominant alleles (*Pp-A3/Pp-B1*), thus demonstrating the complementary action of the two loci on the phenotypic expression of TAC. The wide variation of TAC in the *Pp-A3/Pp-B1* RILs can be likely attributed to the segregation of the structural genes coding for the flavonoid compounds.

The genetics of the red grain color has been investigated since the beginning of the last century, with the well-known study of Nilsson-Ehle [75] on common wheat reporting the genetic control by three loci with additive effects; the experiment is considered a classic case of genetic dissection of a quantitative trait. Subsequent cytogenetic analyses located the 3 loci, designated *R-A1*, *R-B1*, *R-D1*, on the homoeologous chromosomes 3A, 3B, and 3D, respectively, and showed that one functional dominant locus is sufficient for the expression of the red color [76–78]. A series of two-way crosses and the color categorization into six classes, based on color intensity after NaOH treatment of the seeds, showed that grain color is under maternal inheritance and that the red color intensity was related to the number of R-1 loci; the grain color reaction to NaOH of each R-1 gene was not determined [79]. The quantification of proanthocyanidins of single, double, and triple homozygous R-1 genotypes has recently showed that the proanthocyanidin level and the red color intensity are closely linked to the number of dominant *R-1* alleles [80]. Moreover, the significantly higher proanthocyanidin content detected in genotypes *R-B1* and *R-D1* compared to that of *R-A1* suggested that each locus could differently contribute to the flavonoid biosynthesis [80]. Our results validate earlier observations supporting that one functional dominant allele is sufficient for determining the red grain color [79] and that genotypes with two or more functional alleles could only slightly increase the color intensity [80]. However, the observed results could also be due to genotype–environment interactions and/or to errors of the visual categorization of the grain color (particularly the 1 and 2 scores).

Overall, the use of a high-density SNP map determined the precise genetic and physical localization of the *Pp3*, *Pp-B1*, and *R-1* loci, and identified several tightly associated SNP markers to be used in MAS in durum wheat breeding programs.

### 3.3. Candidate Genes and Transcription Factors

The focus of this study was to determine the genetic control of grain pigments in durum wheat. Using an RIL population evaluated in four environments, we identified two QTL for total anthocyanin compounds (*QTAC.mgb-2A* and *QTAC.mgb-7B*) and two QTL for purple grain color (*QGc.mgb-2A* and *QGc.mgb-7B*), mapping in the same genetic and physical regions of chromosomes 2A and 7B. Two other QTL for red grain color (*QGc.mgb-3A* and *QGc.mgb-3B*) were detected on chromosome arms 3AL and 3BL. The projection of these four QTL regions on the durum wheat Svevo reference genome [30] disclosed the occurrence of candidate genes for purple and red grain color underneath their physical interval. Within the chromosomal location of *QTAC.mgb-2A/QGc.mgb-2A* and *QTAC.mgb-7B/QGc.mgb-7B*, the genes *Pp-A3* and *Pp-B1*, respectively, were identified on the corresponding physical interval on the Svevo genome. These dominant genes

with complementary effects were firstly reported as responsible for grain color in bread wheat [18]. Later studies [27] reported that the Pp genes acted as transcriptional factors of an anthocyanin synthesis network in the pericarp. More recent studies on the anthocyanin biosynthesis in common wheat by multi-omic approaches have established that the key loci Pp-A3 and Pp-A1, Pp-B1, and Pp-D1 correspond to the regulatory genes encoding the transcription factors MYC (*Myc-A1*) and MYB (*Mpc1-A1*, *Mpc1-B1*, *Mpc1-D1*) localized on chromosomes 2A and 7A, 7B, and 7D, respectively [44–46,49,81].

Two paralogous genes were identified in durum wheat within the *QTAC.mgb-2A/QGc.mgb-2A* interval, *TRITD2Av1G241410* and *TRITD2Av1G241610*, annotated as a bHLH DNA-binding superfamily protein and bHLH transcription factor, respectively. The two sequences are orthologs to the bread wheat *TraesCS2A02G409400* and *TraesCS2A02G409600*, two tandem duplicated genes reported as *TaMYC1* (Chen et al., 2019) or *TaMYC3-A1* and *TaMYC3-A2* [41]. Two sequences were also identified on durum wheat chromosome 2B: *TRITD2Bv1G202430* (ortholog of *TraesCS2B02G428000*) and *TRITD2Bv1G202460*, two bHLH TFs which showed very high sequence similarity and gene structure conservation. No duplication was observed for the homoeologous genes in emmer and bread wheat. The existence of these close genes on chromosomes 2A and 2B could represent tandem duplications, as observed on 2A in bread wheat, and are considered the main cause of the increase of MYC members in wheat [41]. As reported by Bai et al. [41], the duplication of genes as well as of chromosomal segments represents the primary factors of genome evolution in plants. Phylogenetic analyses placed the durum wheat candidate TFs into a cluster, including four closely related Arabidopsis members (*AtbHLH42*, *AtbHLH12*, *AtbHLH001*, *AtbBHLH002*), which were found to physically interact with R2R3–MYB proteins, and all participating in the regulation of flavonoid pathways, namely proanthocyanidin and anthocyanin biosynthesis [82].

A durum wheat low-confidence gene for MYB TF, *TRITD7Bv1G027860*, was found in the physical interval of the projected *QTAC.mgb-7B/QGc.mgb-7B* QTL, while on the A genome, a high-similarity sequence was annotated as *TRITD7Av1G053110*, an MYB transcription factor reported as *TdMYB7A107* [42]. The missing annotation of the homoeologous gene on the durum wheat 7B chromosome might be depending on mutations and/or polymorphism at the *Pp-B1* locus. In fact, we found a truncated R2R3 domain at the N terminus of the Svevo sequence. A recent paper [40] reported several sequence variations of *Pp3* (*TaPpm1*) and *Pp1* (*TaPpb1*) loci in different colored wheat lines, identifying four allelic variants within the coding region of *Pp3* and two allelic variants in the promoter region for *Pp1*. Alignments of the orthologous genes *TraesCS7B02G070400*, *TRIDC7BG010500*, and *TRITD7Bv1G027860* with the four *Pp3* allelic variants showed that common wheat, wild emmer, and durum wheat *Pp3* genes were different from the four variants identified in Jiang et al. [40].

Candidate gene investigation performed for the physical regions of QTL for grain color on chromosomes 3A (*QGc.mgb-3A*) and 3B (*QGc.mgb-3B*) uncovered MYB-like transcription factors (TFs) at first designated *Myb10* [22,43]. The gene *TRITD3Av1G260970* was found to be physically located within the chromosomal interval of *QGc.mgb-3A* and reported as an MYB transcription factor (*TdMYB3A056*) [42]. No orthologous gene was reported in bread wheat, and no homoeologous gene on the durum wheat B genome. However, a BLAST analysis identified a high-similarity sequence annotated as a low confident-gene, *TRITD3Bv1G255520*, also reported as an MYB TF, which mapped within the *QGc.mgb-3B* interval. Previously, in bread wheat, the R genes for the red grain color, located on the long arms of homoeologous chromosomes 3A, 3B, and 3D, were found to correspond to the three MYB-type transcription factors *Tamyb10-A1*, *Tamyb10-B1*, and *Tamyb10-D1*, respectively, and involved in the activation of the flavonoid biosynthetic pathway in developing grain [22,43]. *Tamyb10* genes encode R2R3-type MYB domain proteins, and sequence variations were found within *Tamyb10-A1* and *Tamyb10-B1* in recessive *R-A1* and *R-B1* lines [22]. A deletion of the first half of the R2-repeat of the MYB region in *Tamyb10-A1* in Chinese Spring caused a loss of function, which explained why the gene

is not annotated in the bread wheat genome. A deletion of 19 bp was also found in the third exon of *Tamyb10-B1* genes of recessive *R-B1* bread wheat lines in comparison with the sequence of the red grain cv. Norin 61 [37]. In Chinese Spring, the deletion was located at 757,919,833 bp in the gene *TraesCS3B02G515900*. We found that the same 19 bp deletion occurred in Svevo (*TRITD3Bv1G255520*, position 3B: 769,813,185), and a shorter one in the wild emmer Zavitan (9 bp deletion in *TRIDC3BG075410*, position 3B: 771,961,919) (Figure S4). Phylogenetic analyses highlighted that the above durum candidate TFs and their wild emmer and bread wheat orthologs fall into the Arabidopsis functional group of flavonoid regulators *AtMYB75*, *AtMYB90*, *AtMYB113*, and *AtMYB114*, which are known to control anthocyanin biosynthesis [50]. In particular, they cluster with the Arabidopsis *AtMYB123*, ortholog of maize *ZmC1*, which controls the biosynthesis of proanthocyanidins in the seed coat of *A. thaliana* [49], thus providing clues as to their putative function.

Other studies have demonstrated that mutations in the promoters or coding sequences of bHLH or MYB proteins affected the ability of TFs to activate the downstream structural genes, thus being responsible for diverse color distributions or intensities in fruits and rice [82–85]. In fact, all these regulatory genes (bHLH and MYB) work together to activate the expression of the structural genes (*PAL*, *CHS*, *CHI*, *F3H*, *F3'H*, *DFR*, *F3'5'H*, and *FLS*) involved in the flavonoid biosynthesis and accumulation in vegetative tissues and seeds [46,65]. Further sequencing analysis would be useful to define eventual different genome variations in both colored and non-colored grains of different species.

### 3.4. Concluding Remarks

Cultivated cereal species include genotypes with pigmented grain (red, purple, blue, black) having high levels of flavonoid compounds, which possess beneficial effects on human and animal health. Colored grain can be consumed as whole grain, utilized to produce a variety of fortified functional foods, or for preparation of antioxidant dietary supplements. Using a biparental segregating population of durum wheat genotyped with the 25K wheat SNP arrays and evaluated for TAC and grain color, four QTL corresponding to genes *Pp-A3*, *Pp-B1*, *R-A1*, and *R-B1*, responsible for the purple and red grain and involved in flavonoid biosynthesis pathways, were identified. The four experimental trials revealed the effect of environmental factors on total grain anthocyanin content. The availability of the genome sequence in durum wheat allowed for the identification and precise localization of the candidate genes corresponding to transcription factors and a comparison of them with the orthologous genes in emmer and common wheat. The present study provides a set of molecular markers to select the essential alleles for flavonoid pigment synthesis in durum wheat breeding programs.

## 4. Materials and Methods

### 4.1. Plant Materials

A segregant population of 144 recombinant inbred lines (RILs) was developed by crossing the Ethiopian durum wheat line PG2 with purple grain with Grecale, an Italian durum cultivar with amber grain, and by advancing single F<sub>2</sub> plants to the F<sub>7</sub> generation via the single seed descent method. The PG2 line was obtained by genealogical selection carried out in the durum landrace accession CI 14629 collected at Shewa, Ethiopia, and kindly provided by the United States Department of Agriculture (USDA-ARS), USA. Grecale is an elite durum cultivar grown in Italy for its quality traits, including a relatively high grain protein content. The RIL population and the parental lines were grown in the experimental fields of the University of Bari, DISSPA Dept. (Bari, Italy), across four field trials carried out for three growing seasons (2020–2022) at Valenzano, Italy, (designated VAL\_2020, VAL\_2021, VAL\_2022) and for one year (2021) at Bari (Italy) (designated BA\_2021) under rainfed conditions. The experimental design for each trial was a randomized complete block with three replications. The parental lines PG2 and Grecale were repeated three times in each replication. The experimental unit consisted of a 1 m row, 30 cm apart, with

30 germinating seeds per plot. At maturity, the plots were hand-harvested, and the spikes shelled with a micro-thresher.

#### 4.2. Phenotypic Analysis

Grain color intensity of each RIL and parental line was visually determined by using a 0–3 scale based on the deepness of the color: 0 = amber, 1 = red, 2 = red-brownish, 3 = dark purple (Figure 1). Grains were ground to wholemeal flour by means of a laboratory mill (Cyclotec Sample Mill, Tecator Foss, Hillerød, Denmark) equipped with a 1 mm sieve. Wholemeal samples were stored at 4 °C to minimize pigment degradation by oxidative enzymes for a maximum of 24 h before analysis. Colorimetric indexes were determined using the reflectance colorimeter Chroma Meter CR-300 (Minolta, Osaka, Japan) equipped with a pulsed xenon lamp. Color values  $L^*$ , measuring brightness,  $a^*$ , measuring red to green, and  $b^*$ , measuring yellow to blue [86], were used in subsequent analysis.

Total anthocyanin content (TAC) was determined according to Abdel-Aal and Hucl [87]. Briefly, 10 mL of 85:15 (*v/v*) methanol/1 M HCl was added to 1 g of each sample, and TACs were extracted by an orbital shaker at 500 rpm for 30 min; the supernatant was collected after centrifugation at  $12,000\times g$  for 5 min. The pellet was re-extracted under the same conditions. The two supernatants were combined, and the absorbance determined at 535 nm by a Cary 60 UV-Vis spectrophotometer (Agilent Technologies, Santa Clara, CA, USA). All determinations were done at least in triplicate, and analytical results expressed on a dry matter basis as  $\mu\text{g/g}$  of cyanidin 3-O-glucoside (Phytoplan, Heidelberg, Germany) used as standard in the previously set calibration curve. Each replication of all the RIL and parental lines were evaluated for TAC in the trials performed at Valenzano 2021 and Valenzano 2022, while a mixture of wholemeal of the three replications was assessed for the trials conducted at Valenzano 2020 and Bari 2021.

#### 4.3. DNA Extraction and Molecular Marker Analysis

DNA of each RIL and parental lines PG2 and Grecale was extracted from fresh leaves by using the GeneElute Plant Genomic Miniprep Kit (Sigma, Waltham, MA, USA). After checking DNA concentration and quality by agarose gel-electrophoresis and NanoDrop2000 (Thermo Scientific™, Waltham, MA, USA), each DNA sample was diluted to 50 ng/ $\mu\text{L}$  and sent to TraitGenetics GmbH (Gatersleben, Germany) [88] for the genotyping assays with the wheat SNP 25K chip array developed by Illumina CPro<sup>R</sup> (San Diego, CA, USA), as described by [28].

#### 4.4. Genetic Linkage Map Construction

The RIL population and the parental lines were genotyped by the wheat 25K SNP Illumina Infinium Array [28], including 24,145 single nucleotide polymorphisms (SNPs). A total of 2142 (9.1%) markers failed, 602 (2.4%) markers had more than 10% missing data, 13,141 (54.4%) markers were monomorphic, and 2268 (9.4%) markers showed distorted segregation (at  $p \geq 0.001$  value). The remaining 5942 (24.6%) polymorphic markers were used to obtain a genetic linkage map. Seven RILs with more than 20% missing data were also excluded. Linkage analysis between markers and linear order determination of loci were performed by the software QTL IciMapping v. 4.2 [29] using, initially, an LOD value of 6.0 for grouping the markers. SNP data from the reference consensus durum wheat map [89] were used as anchor loci as well as for assigning linkage groups to specific chromosomes. Linkage groups belonging to the same chromosome were then re-analyzed at LOD 3.0 for possible assignment to a single linkage group. A total of 108 markers remained unlinked or assembled in small LGs with any chromosomal anchor locus or assembled in an LG with the same map position, and therefore not further considered. Map distances were calculated by the Kosambi mapping function. Linkage groups were designated according to the wheat chromosome nomenclature; chromosome name is followed by a consecutive number when two or more linkage groups are present for a chromosome.



#### 4.5. Candidate Genes and Transcription Factors

The QTL genomic regions identified on chromosomes 2A, 3A, 3B, and 7B were projected on the durum wheat Svevo reference genome [30,90]. The physical intervals were then investigated to validate the occurrence of purple and red grain color candidate genes encoding for bHLH and MYB TFs within these intervals. Furthermore, orthologous genes in durum wheat were detected by searching the wild emmer wheat Zavitan reference genome [36,90] and the bread wheat Chinese Spring reference genome v1.0 [37,91].

The candidate durum and wild emmer wheat genes were translated, and the first longest variant was considered for phylogenetic analyses. Multiple sequence alignments were performed by MAFFT v. 7 online service v. 7 [92] using MAFFT alignment with L-INS-i preset. Full-length amino acid sequences of putative bHLH genes were aligned with *Arabidopsis AtbHLH* proteins [48], bread wheat *TabHLH* [47], and *TaMYC* TFs [38,41]. The longest isoform of putative durum and wild emmer wheat MYB proteins were aligned with the annotated *Arabidopsis AtMYBs* retrieved from Plant TFDB and annotated according to Jiang and Rao [93], the durum wheat *TdMYB* TFs [42], and a subset of bread wheat MYB genes, putatively involved in the regulation of phenylpropanoid biosynthesis, selected from the phylogenetic analyses carried by Blanco and coworkers [42]. Phylogenetic studies based on maximum likelihood (ML) analyses were performed using W-IQ-TREE [94], with 1000 ultrafast bootstrap replicates. Phylogenetic trees were imported and annotated with iTOL v. 6.7 [95]. The trees were rooted with CDC5 sequences.

#### 4.6. Statistical and QTL Analyses

Analysis of variance (ANOVA) for each trait for each field trial and for the combined data across environments was performed by the software CoStat v. 6.4 [96]. Data are reported as means  $\pm$  standard error (es). Genetic variance, environmental variance, genotypic  $\times$  environmental variance, and broad-sense heritability were obtained by using variance component estimates. Means of parental lines were compared by an independent two-sample Student's *t*-test. Comparison of the four RIL groups with different genotypes was carried out by Fisher's LSD test. Simple phenotypic correlation coefficients among  $L^*$ ,  $a^*$ , and  $b^*$  color values, TAC (Pearson correlation coefficients), and grain color (Sperman rank correlation coefficients) were calculated for each environment and across environments. QTL mapping was carried out via the Inclusive Composite Interval Mapping (ICIM) method using the software IciMapping v. 4.2 [29].

QTL analysis was conducted for the mean values of each phenotypic trait for each field trial and for the overall mean across the four trials. A QTL was considered significant when one or more markers were associated with the trait at a threshold *p* value of 0.001 ( $-\log_{10}(P) \geq 3.0$ ). QTL are reported when detected in at least two environments to reduce the detection of false-positive QTL. The proportion of phenotypic variance explained (PVE%) and the additive effect were estimated for each detected QTL. The International Rules of Genetic Nomenclature for wheat were followed for the QTL designation [21]. The software MapChart v. 2.2 [97] was used for the graphical representation of linkage groups and QTL.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/plants12081674/s1>: Figure S1: Frequency distributions of total anthocyanin compounds (TAC) in four field trials carried out at Valenzano 2020 (a), Valenzano 2021 (b), Bari 2021 (c), Valenzano 2022 (d). Figure S2: Phylogenetic analysis of bHLH proteins from wheat and *Arabidopsis thaliana*. Maximum likelihood tree, including the putative durum TFs (TRITD) identified in this study, their wild emmer orthologs (TRIDC), bread wheat TabHLH and subclass TaMYC proteins (TraesCS), and the *Arabidopsis thaliana* AtbHLH TFs. Purple: putative MYC proteins; dark purple: group with the MYC candidate genes identified here. The tree is rooted with the AtCDC5 protein. Numbers on the branch nodes indicate bootstrap values (only bootstraps > 60 are shown). Figure S3: Phylogenetic analysis of R2R3-MYB proteins from wheat and *Arabidopsis thaliana*. Maximum likelihood tree, including durum wheat (TRITD), their wild emmer orthologs (TRIDC), bread wheat (TraesCS) proteins putatively involved in the regulation of phenylpropanoid



biosynthesis, and *A. thaliana* AtMYBs. Brown: group with the putative MYB genes identified in this study. The tree is rooted with the AtCDC5 protein. Numbers on the branch nodes indicate bootstrap values (only bootstraps >60 are shown). Figure S4: Alignment of *Tamyb10-B1* gene sequences from the red grain bread wheat cv. Norin 61 [43], bread wheat Chinese Spring, wild emmer wheat Zavitan, and durum wheat Svevo. Table S1: Analysis of variance of five traits evaluated in the RIL population PG2 × Grecale in four environments. Table S2: Analysis of variance of five traits evaluated in the RIL population PG2 × Grecale across four environments. Table S3. Pairwise correlation coefficients among environments for L\*, a\*, and b\* color values, total anthocyanin compounds (TAC) (Pearson correlation coefficients), and grain color (Sperman rank correlation coefficients) in the recombinant inbred line population PG2 × Grecale. Table S4: Simple correlation coefficients among L\*, a\*, and b\* color values, total anthocyanin compounds (TAC) (Pearson correlation coefficients), and grain color (Sperman rank correlation coefficients) in individual environments and across environments in the recombinant inbred line population PG2 × Grecale grown in four environments. Table S5: Number and distribution of SNP markers in the PG2 × Grecale map.

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

**Funding:** This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions; neither the European Union nor the European Commission can be considered responsible for them.

**Data Availability Statement:** All available data are contained within the article.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Abbreviations

QTL	Quantitative trait loci
LOD	Logarithm of odds
PVE	Phenotype variance explained
LG	Linkage group
TAC	Total anthocyanin content
ICIM	Inclusive Composite Interval Mapping
TF	Transcription factor
SNP	Single nucleotide polymorphism

## References

- International Grains Council. World Grain Statistics. Available online: <https://www.igc.int/en/default.aspx> (accessed on 1 December 2022).
- FAOSTAT. Production Domain. Crops. Rome: FAO. Available online: <https://www.fao.org/faostat/> (accessed on 17 January 2023).
- Colasuonno, P.; Marcotuli, I.; Blanco, A.; Maccaferri, M.; Condorelli, G.E.; Tuberosa, R.; Gadaleta, A. Carotenoid pigment content in durum wheat (*Triticum turgidum* L. ssp. *durum*): An overview of quantitative trait loci and candidate genes. *Front. Plant Sci.* **2019**, *10*, 1347. [[CrossRef](#)]
- Weststrate, J.A.; Poppel, G.; Verschuren, P.M. Functional foods, trends and future. *Br. J. Nutr.* **2002**, *88*, 233–235. [[CrossRef](#)]
- Dwivedi, S.L.; Mattoo, A.K.; Garg, M.; Dutt, S.; Singh, B.; Ortiz Rios, R.O. Developing germplasm and promoting consumption of anthocyanin-rich grains for health benefits. *Front. Sustain. Food Syst.* **2022**, *6*, 137. [[CrossRef](#)]
- Petroni, K.; Pilu, R.; Tonelli, C. Anthocyanins in corn: A wealth of genes for human health. *Planta* **2014**, *240*, 901–911. [[CrossRef](#)] [[PubMed](#)]
- Blesso, C.N. Dietary Anthocyanins and human health. *Nutrients* **2019**, *11*, 2107. [[CrossRef](#)]

8. Ficco, D.B.M.; Mastrangelo, A.M.; Trono, D.; Borrelli, G.M.; DeVita, P.; Fares, C.; Beleggia, R.; Platani, C.; Papa, R. The colors of durum wheat: A review. *Crop Pasture Sci.* **2014**, *65*, 1–15. [[CrossRef](#)]
9. Pasqualone, A.; Bianco, A.M.; Paradiso, V.M.; Summo, C.; Gambacorta, G.; Caponio, F.; Blanco, A. Production and characterization of functional biscuits obtained from purple wheat. *Food Chem.* **2015**, *180*, 64–70. [[CrossRef](#)] [[PubMed](#)]
10. Ficco, D.B.M.; Simone, V.; Leonardis, A.M.; Giovanniello, V.; Nobile, M.A.D.; Padalino, L.; Lecce, L.; Borrelli, G.M.; Vita, P. Use of purple durum wheat to produce naturally functional fresh and dry pasta. *Food Chem.* **2016**, *205*, 187–195. [[CrossRef](#)]
11. Gamel, T.H.; Wright, A.J.; Pickard, M.; Abdel-Aal, E.S.M. Characterization of anthocyanin-containing purple wheat prototype products as functional foods with potential health benefits. *Cereal Chem.* **2020**, *97*, 34–38. [[CrossRef](#)]
12. Kumari, A.; Sharma, S.; Sharma, N.; Chunduri, V.; Kapoor, P.; Kaur, S.; Goyal, A.; Garg, M. Influence of biofortified colored wheats (purple, blue, black) on physicochemical, antioxidant and sensory characteristics of chapatti (Indian flatbread). *Molecules* **2020**, *25*, 5071. [[CrossRef](#)] [[PubMed](#)]
13. Saini, P.; Kumar, N.; Kumar, S.; Mwaurah, P.W.; Panghal, A.; Attkan, A.K.; Singh, V. Bioactive compounds, nutritional benefits and food applications of colored wheat: A comprehensive review. *Crit. Rev. Food Sci. Nutr.* **2021**, *61*, 3197–3210. [[CrossRef](#)]
14. Romano, G.; Tufariello, M.; Calabriso, N.; Del Coco, L.; Fanizzi, F.P.; Blanco, A.; Carluccio, M.A.; Grieco, F.; Laddomada, B. Pigmented cereals and legume grains as healthier alternatives for brewing beers. *Food Biosci.* **2023**, *52*, 102463. [[CrossRef](#)]
15. Zeven, A.C. Wheats with purple and blue grains: A review. *Euphytica* **1991**, *56*, 243–258. [[CrossRef](#)]
16. Shoeva, O.Y.; Khlestkina, E.K. The specific features of anthocyanin biosynthesis regulation in wheat. In *Advances in Wheat Genetics: From Genome to Field*; Springer: Tokyo, Japan, 2015; pp. 147–157.
17. 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](#)]
18. Dobrovolskaya, O.; Arbuzova, V.S.; Lohwasser, U.; Röder, M.S.; Börner, A. Microsatellite mapping of complementary genes for purple grain color in bread wheat (*Triticum aestivum* L.). *Euphytica* **2006**, *150*, 355–364. [[CrossRef](#)]
19. Khlestkina, E.K.; Röder, M.S.; Börner, A. Mapping genes controlling anthocyanin pigmentation on the glume and pericarp in tetraploid wheat (*Triticum durum* L.). *Euphytica* **2010**, *171*, 65–69. [[CrossRef](#)]
20. Gordeeva, E.I.; Shoeva, O.Y.; Khlestkina, E.K. Marker-assisted development of bread wheat near-isogenic lines carrying various combinations of purple pericarp (*Pp*) alleles. *Euphytica* **2015**, *203*, 469–476. [[CrossRef](#)]
21. McIntosh, R.A.; Hart, G.E.; Devos, K.M.; Gale, M.D.; Rogers, W.J. Catalogue of gene symbols for wheat. In *Proceedings of the Ninth International Wheat Genetics Symposium*; Saskatoon, SK, Canada, 2–7 August 1998, University of Saskatchewan Extension Press: Saskatoon, SK, Canada, 1998; Volume 5. Available online: <http://wheat.pw.usda.gov/ggpages/pubs.html> (accessed on 20 November 2022).
22. Himi, E.; Maekawa, M.; Miura, H.; Noda, K. Development of PCR markers for *Tamyb10* related to *R-1*, red grain color gene in wheat. *Theor. Appl. Genet.* **2011**, *122*, 1561–1576. [[CrossRef](#)]
23. Abdel-Aal, E.S.M.; Abou-Arab, A.A.; Gamel, T.H.; Hucl, P.; Young, J.C.; Rabalski, I. Fractionation of blue wheat anthocyanin compounds and their contribution to antioxidant properties. *J. Agric. Food Chem.* **2008**, *56*, 11171–11177. [[CrossRef](#)]
24. Qualset, C.O.; Soliman, K.M.; Jan, C.C.; Dvorak, J.; McGuire, P.E.; Vogt, H.E. Registration of UC66049 *Triticum aestivum* blue aleurone genetic stock. *Crop Sci.* **2005**, *45*, 432. [[CrossRef](#)]
25. Singh, K.; Ghai, M.; Garg, M.; Chhuneja, P.; Kaur, P.; Schnurbusch, T.; Keller, B.; Dhaliwal, H.S. An integrated molecular linkage map of diploid wheat based on a *Triticum boeoticum*–*T. monococcum* RIL population. *Theor. Appl. Genet.* **2007**, *115*, 301–312. [[CrossRef](#)] [[PubMed](#)]
26. Eticha, F.; Grausgruber, H.; Siebenhandl-Ehn, S.; Berghofer, E. Some agronomic and chemical traits of blue aleurone and purple pericarp wheat (*Triticum* L.). *J. Agric. Sci. Technol.* **2011**, *1*, 48–58.
27. Tereshchenko, O.Y.; Arbuzova, V.S.; Khlestkina, E.K. Allelic state of the genes conferring purple pigmentation in different wheat organs predetermines transcriptional activity of the anthocyanin biosynthesis structural genes. *J. Cereal Sci.* **2013**, *57*, 10–13. [[CrossRef](#)]
28. 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](#)]
29. 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](#)]
30. 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.; et al. Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* **2019**, *51*, 885–895. [[CrossRef](#)]
31. Li, H.; Ye, G.; Wang, J. A Modified algorithm for the improvement of composite interval mapping. *Genetics* **2007**, *175*, 361–374. [[CrossRef](#)]
32. Frankel, W.N.; Schork, N.J. Who's afraid of epistasis. *Nat. Genet.* **1996**, *14*, 371–373. [[CrossRef](#)]
33. Li, H.; Ribaut, J.M.; Li, Z.; Wang, J. Inclusive Composite Interval Mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. *Theor. Appl. Genet.* **2008**, *116*, 243–260. [[CrossRef](#)]
34. Lap, B.; Rai, M.; Tyagi, W. Playing with colors: Genetics and regulatory mechanisms for anthocyanin pathway in cereals. *Biotechnol. Genet. Eng. Rev.* **2021**, *37*, 1–29. [[CrossRef](#)]

35. Iannucci, A.; Suriano, S.; Cancellaro, S.; Trono, D. Anthocyanin profile and main antioxidants in pigmented wheat grains and related millstream fractions. *Cereal Chem.* **2022**, *99*, 1282–1295. [[CrossRef](#)]
36. Avni, R.; Nave, M.; Barad, O.; Baruch, K.; Twardziok, S.O.; Gundlach, H.; Hale, I.; Mascher, M.; Spannagl, M.; Wiebe, K.; et al. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* **2017**, *357*, 93–97. [[CrossRef](#)] [[PubMed](#)]
37. International Wheat Genome Sequencing Consortium (IWGSC); Appels, R.; Eversole, K.; Stein, N.; Feuillet, C.; Keller, B.; Rogers, J.; Pozniak, C.J.; Choulet, F.; Distelfeld, A.; et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **2018**, *361*, eaar7191.
38. Chen, S.; Zhao, H.; Luo, T.; Liu, Y.; Nie, X.; Li, H. Characteristics and expression pattern of MYC genes in *Triticum aestivum*, *Oryza sativa*, and *Brachypodium distachyon*. *Plants* **2019**, *8*, 274. [[CrossRef](#)] [[PubMed](#)]
39. Bi, H.H.; Sun, Y.W.; Xiao, Y.G.; Xia, L.Q. Characterization of DFR allelic variations and their associations with pre-harvest sprouting resistance in a set of red-grained chinese wheat germplasm. *Euphytica* **2014**, *195*, 197–207. [[CrossRef](#)]
40. Jiang, W.; Liu, T.; Nan, W.; Jeewani, D.C.; Niu, Y.; Li, C.; Wang, Z. Two transcription factors *Tappm1* and *Tappb1* co-regulate anthocyanin biosynthesis in purple pericarps of wheat. *J. Exp. Bot.* **2018**, *69*, 2555–2567. [[CrossRef](#)]
41. Bai, J.F.; Wang, Y.K.; Guo, L.P.; Guo, X.M.; Guo, H.Y.; Yuan, S.H.; Duan, W.J.; Liu, Z.; Zhao, C.P.; Zhang, F.T.; et al. Genomic identification and characterization of MYC family genes in wheat (*Triticum aestivum* L.). *BMC Genom.* **2019**, *20*, 1032. [[CrossRef](#)]
42. Blanco, E.; Curci, P.L.; Manconi, A.; Sarli, A.; Zuluaga, D.L.; Sonnante, G. R2R3-MYBs in durum wheat: Genome-wide identification, poaceae-specific clusters, expression, and regulatory dynamics under abiotic stresses. *Front. Plant Sci.* **2022**, *13*, 896945. [[CrossRef](#)]
43. Himi, E.; Noda, K. Red grain colour gene (R) of wheat is a Myb-type transcription factor. *Euphytica* **2005**, *143*, 239–242. [[CrossRef](#)]
44. Shoeva, O.; Gordeeva, E.; Khlestkina, E. The regulation of anthocyanin synthesis in the wheat pericarp. *Molecules* **2014**, *19*, 20266–20279. [[CrossRef](#)]
45. Li, W.; Faris, J.D.; Chittoor, J.M.; Leach, J.E.; Hulbert, S.H.; Liu, D.J.; Chen, P.D.; Gill, B.S. Genomic mapping of defense response genes in wheat. *Theor. Appl. Genet.* **1999**, *98*, 226. [[CrossRef](#)]
46. Strygina, K.V.; Khlestkina, E.K. Structural and functional divergence of the *Mpc1* genes in wheat and barley. *BMC Evol. Biol.* **2019**, *19*, 89–99. [[CrossRef](#)] [[PubMed](#)]
47. Wang, L.; Xiang, L.; Hong, J.; Xie, Z.; Li, B. Genome-wide analysis of BHLH transcription factor family reveals their involvement in biotic and abiotic stress responses in wheat (*Triticum aestivum* L.). *3 Biotech* **2019**, *9*, 236. [[CrossRef](#)] [[PubMed](#)]
48. Pires, N.; Dolan, L. Origin and diversification of basic-helix-loop-helix proteins in plants. *Mol. Biol. Evol.* **2010**, *27*, 862–874. [[CrossRef](#)] [[PubMed](#)]
49. Nesi, N.; Jond, C.; Debeaujon, I.; Caboche, M.; Lepiniec, L. The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell* **2001**, *13*, 2099–2114. [[CrossRef](#)]
50. Dubos, C.; Stracke, R.; Grotewold, E.; Weisshaar, B.; Martin, C.; Lepiniec, L. MYB transcription factors in Arabidopsis. *Trends Plant Sci.* **2010**, *15*, 573–581. [[CrossRef](#)] [[PubMed](#)]
51. Gani, A.; Wani, S.M.; Masoodi, F.A.; Hameed, G. Whole-grain cereal bioactive compounds and their health benefits: A review. *J. Food Process. Technol.* **2012**, *3*, 146–156. [[CrossRef](#)]
52. Francavilla, A.; Joye, I.J. Characterization of anthocyanins in purple and blue wheat using fluorescence spectroscopy and reversed phase HPLC. *Cereals Grains* **2020**, *20*.
53. Bueno, J.M.; Ramos-Escudero, F.; Sáez-Plaza, P.; Muñoz, A.M.; Jose Navas, M.; Asuero, A.G. Analysis and antioxidant capacity of anthocyanin pigments. Part I: General considerations concerning polyphenols and flavonoids. *Crit. Rev. Anal. Chem.* **2012**, *42*, 102–125. [[CrossRef](#)]
54. Lu, X.; Zhou, Y.; Wu, T.; Hao, L. Ameliorative effect of black rice anthocyanin on senescent mice induced by D-galactose. *Food Funct.* **2014**, *5*, 2892–2897. [[CrossRef](#)]
55. Giordano, D.; Locatelli, M.; Travaglia, F.; Bordiga, M.; Reyneri, A.; Coisson, J.D.; Blandino, M. Bioactive compound and antioxidant activity distribution in roller-milled and pearled fractions of conventional and pigmented wheat varieties. *Food Chem.* **2017**, *233*, 483–491. [[CrossRef](#)]
56. Cammerata, A.; Laddomada, B.; Milano, F.; Camerlengo, F.; Bonarrigo, M.; Masci, S.; Sestili, F. Qualitative characterization of unrefined durum wheat air-classified fractions. *Foods* **2021**, *10*, 2817. [[CrossRef](#)] [[PubMed](#)]
57. Cammerata, A.; Sestili, F.; Laddomada, B.; Aureli, G. Bran-enriched milled durum wheat fractions obtained using innovative micronization and air-classification pilot plants. *Foods* **2021**, *10*, 1796. [[CrossRef](#)] [[PubMed](#)]
58. Alzuwaid, N.T.; Fellows, C.M.; Laddomada, B.; Sisson, M. Impact of wheat bran particle size on the technological and phytochemical properties of durum wheat pasta. *J. Cereal Sci.* **2020**, *95*, 103033. [[CrossRef](#)]
59. Alzuwaid, N.T.; Fleming, D.; Fellows, C.M.; Laddomada, B.; Sissons, M. Influence of Durum Wheat Bran Particle Size on Phytochemical Content and on Leavened Bread Baking Quality. *Foods* **2021**, *10*, 489. [[CrossRef](#)]
60. Khlestkina, E. The adaptive role of flavonoids: Emphasis on cereals. *Cereal Res. Commun.* **2013**, *41*, 185–198. [[CrossRef](#)]
61. Landi, M.; Tattini, M.; Gould, K.S. Multiple functional roles of anthocyanins in plant-environment interactions. *Environ. Exp. Bot.* **2015**, *119*, 4–17. [[CrossRef](#)]
62. Mbarki, S.; Sytar, O.; Zivcak, M.; Abdelly, C.; Cerda, A.; Brestic, M. Anthocyanins of colored wheat genotypes in specific response to salt stress. *Molecules* **2018**, *23*, 1518. [[CrossRef](#)]

63. Li, X.L.; Xiang, L.Ü.; Wang, X.H.; Qin, P.E.N.G.; Zhang, M.S.; Ren, M.J. Biotic and abiotic stress-responsive genes are stimulated to resist drought stress in purple wheat. *J. Integr. Agric.* **2020**, *19*, 33–50. [[CrossRef](#)]
64. Bassoi, M.C.; Flintham, J. Relationship between grain color and preharvest sprouting-resistance in wheat. *Pesqui. Agropecu. Bras.* **2005**, *40*, 981–988. [[CrossRef](#)]
65. Sharma, A.; Shahzad, B.; Rehman, A.; Bhardwaj, R.; Landi, M.; Zheng, B. Response of phenylpropanoid pathway and the role of polyphenols in plants under abiotic stress. *Molecules* **2019**, *24*, 2452. [[CrossRef](#)] [[PubMed](#)]
66. Knievel, D.C.; Abdel-Aal, E.S.; Rabalski, I.; Nakamura, T.; Hucl, P. Grain color development and the inheritance of high anthocyanin blue aleurone and purple pericarp in spring wheat (*Triticum Aestivum* L.). *J. Cereal Sci.* **2009**, *50*, 113–120. [[CrossRef](#)]
67. Bustos, D.V.; Riegel, R.; Calderini, D.F. Anthocyanin content of grains in purple wheat is affected by grain position, assimilate availability and agronomic management. *J. Cereal Sci.* **2012**, *55*, 257–264. [[CrossRef](#)]
68. Abdel-Aal, E.S.M.; Hucl, P.; Shipp, J.; Rabalski, I. Compositional differences in anthocyanins from blue and purple grained spring wheat grown in four environments in Central Saskatchewan. *Cereal Chem.* **2016**, *93*, 32–38. [[CrossRef](#)]
69. Clark, J.A. Improvement of wheat. In *Yearbook of Agriculture*; U.S. Government Publishing Office: Washington, DC, USA, 1936; pp. 207–302.
70. Sharman, B.C. ‘Purple pericarp’: A monofactorial dominant in tetraploid wheats. *Nature* **1958**, *181*, 929. [[CrossRef](#)]
71. Copp, L.G.L. Purple graininess in hexaploid wheat. *Wheat Inf. Serv.* **1965**, *18*, 19–20.
72. Gilchrist, J.A.; Sorrells, M.E. Inheritance of kernel color in ‘Charcoal’ wheat. *J. Hered.* **1982**, *73*, 457–460. [[CrossRef](#)]
73. Li, X.P.; Lan, S.Q.; Zhang, Y.L.; Liu, Y.P. Identification of molecular markers linked to the genes for purple grain color in wheat (*Triticum aestivum*). *Genet. Resour. Crop Evol.* **2010**, *57*, 1007–1012. [[CrossRef](#)]
74. McIntosh, R.A.; Backer, E.P. Inheritance of purple pericarp in wheat. *Proc. Linn. Soc.* **1967**, *92*, 204–208.
75. Nilsson-Ehle, H. Kreuzungsuntersuchungen an hafer und weizen. Lunds Universitets Arsskrift, N.F. *Afd* **1909**, *5*, 122.
76. Sears, E.R. Cytogenetic Studies with polyploid species of wheat. ii. additional chromosomal aberrations in *Triticum vulgare*. *Genetics* **1944**, *29*, 232–246. [[CrossRef](#)] [[PubMed](#)]
77. Allan, R.E.; Vogel, O.A. Monosomic analysis of red seed color in wheat. *Crop Sci.* **1965**, *5*, 474–475. [[CrossRef](#)]
78. Metzger, R.J.; Silbaugh, B.A. Location of genes for seed coat color in hexaploidy wheat, *Triticum aestivum* L. *Crop Sci.* **1970**, *10*, 495–496. [[CrossRef](#)]
79. Pauw, R.D.; McCaig, T.N. Utilization of sodium hydroxide to assess kernel color and its inheritance in eleven spring wheat varieties. *Can. J. Plant Sci.* **1988**, *68*, 323–329. [[CrossRef](#)]
80. Himi, E.; Matsuura, T.; Miura, H.; Yoshihara, N.; Maekawa, M. A new method for detecting proanthocyanidin content in wheat reveals the relationship between *R-1* gene to grain color deepness. *Cereal Chem.* **2023**, *100*, 512–521. [[CrossRef](#)]
81. Himi, E.; Taketa, S. Isolation of candidate genes for the barley Ant1 and wheat Rc genes controlling anthocyanin pigmentation in different vegetative tissues. *Mol. Genet. Genom.* **2015**, *290*, 1287–1298. [[CrossRef](#)] [[PubMed](#)]
82. Zimmermann, I.M.; Heim, M.A.; Weisshaar, B.; Uhrig, J.F. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J.* **2004**, *40*, 22–34. [[CrossRef](#)]
83. Butelli, E.; Licciardello, C.; Zhang, Y.; Liu, J.; Mackay, S.; Bailey, P.; Reforgiato-Recupero, G.; Martin, C. Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* **2012**, *24*, 1242–1255. [[CrossRef](#)] [[PubMed](#)]
84. Rahman, M.M.; Lee, K.E.; Lee, E.S.; Matin, M.N.; Lee, D.S.; Yun, J.S.; Kim, J.B.; Kang, S.G. The genetic constitutions of complementary genes *Pp* and *Pb* determine the purple color variation in pericarps with cyanidin-3-O-glucoside depositions in black rice. *J. Plant Biol.* **2013**, *56*, 24–31. [[CrossRef](#)]
85. Cheng, J.; Liao, L.; Zhou, H.; Gu, C.; Wang, L.; Han, Y. A small indel mutation in an anthocyanin transporter causes variegated coloration of peach flowers. *J. Exp. Bot.* **2015**, *66*, 7227–7239. [[CrossRef](#)]
86. Commission Internationale de l’Eclairage (CIE). *Colorimetry*, 2nd ed.; No. 15.2; CIE: Vienna, Austria, 1986.
87. Abdel-Aal, E.-S.M.; Hucl, P. A rapid method for quantifying total anthocyanins in blue aleurone and purple pericarp wheats. *Cereal Chem.* **1999**, *76*, 350–354. [[CrossRef](#)]
88. SGS Institut Fresenius. Available online: <https://sgs-institut-fresenius.de/en/gesundheits-und-ernaehrung/traitgenetics> (accessed on 10 December 2022).
89. Maccaferri, M.; Ricci, A.; Salvi, S.; Milner, S.G.; Noli, E.; Martelli, P.L.; Casadio, R.; Akhunov, E.; Scalabrin, S.; Vendramin, V.; et al. A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* **2015**, *13*, 648–663. [[CrossRef](#)] [[PubMed](#)]
90. InterOmics. Intranet of Durum Wheat Genome Data. Available online: <https://www.interomics.eu/durum-wheat-genome-intranet> (accessed on 10 December 2020).
91. Plant Bioinformatics Facility. Available online: <https://urgi.versailles.inrae.fr/jbrowseiwgsc/> (accessed on 10 December 2022).
92. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT Online Service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **2019**, *20*, 1160–1166. [[CrossRef](#)] [[PubMed](#)]
93. Jiang, C.K.; Rao, G.Y. Insights into the Diversification and Evolution of R2R3-MYB Transcription Factors in Plants. *Plant Physiol.* **2020**, *183*, 637. [[CrossRef](#)] [[PubMed](#)]
94. Trifinopoulos, J.; Nguyen, L.T.; von Haeseler, A.; Minh, B.Q. W-IQ-TREE: A Fast Online Phylogenetic Tool for Maximum Likelihood Analysis. *Nucleic Acids Res.* **2016**, *44*, W232–W235. [[CrossRef](#)] [[PubMed](#)]

95. Letunic, I.; Bork, P. Interactive Tree of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)]
96. CoHort Software. Available online: <https://www.cohortsoftware.com/index.html> (accessed on 1 December 2022).
97. Voorrips, R. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* **2002**, *93*, 77–78. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.