Theor Appl Genet
DOI 10.1007/s00122-016-2820-1

ORIGINAL ARTICLE

Genome-wide association study reveals genetic architecture of coleoptile length in wheat

Genqiao Li[1,3] · Guihua Bai[2] · Brett F. Carver[3] · Norman C. Elliott[1] · Rebecca S. Bennett[1] · Yanqi Wu[3] · Robert Hunger[4] · J. Michael Bonman[5] · Xiangyang Xu[1]

Received: 3 June 2016 / Accepted: 27 October 2016

Abstract

Key message **Eight QTL for coleoptile length were identified in a genome-wide association study on a set of 893 wheat accessions, four of which are novel loci.**

Abstract Wheat cultivars with long coleoptiles are preferred in wheat-growing regions where deep planting is practiced. However, the wide use of gibberellic acid (GA)-insensitive dwarfing genes, *Rht-B1b* and *Rht-D1b*, makes it challenging to breed dwarf wheat cultivars with long coleoptiles. To understand the genetic basis of coleoptile length, we performed a genome-wide association study on a set of 893 landraces and historical cultivars using 5011 single nucleotide polymorphism (SNP) markers. Structure analysis revealed four subgroups in the association panel. Association analysis results suggested that *Rht-B1b* and *Rht-D1b* genes significantly reduced coleoptile length, and eight additional quantitative trait loci (QTL) for coleoptile length were also identified. These QTL explained 1.45–3.18 and 1.36–3.11% of the phenotypic variation in 2015 and 2016, respectively, and their allelic substitution effects ranged from 0.31 to 1.75 cm in 2015, and 0.63–1.55 cm in 2016. Of the eight QTL, *QCL.stars-1BS1*, *QCL.stars-2DS1*, *QCL.stars-4BS2,* and *QCL.stars-5BL1* are likely novel loci for coleoptile length. The favorable alleles in each accession ranged from two to eight with an average of 5.8 at eight loci in the panel, and more favorable alleles were significantly associated with longer coleoptile, suggesting that QTL pyramiding is an effective approach to increase wheat coleoptile length.

Abbreviations

FDR	False discovery rate
GWAS	Genome-wide association study
LD	Linkage disequilibrium
LnP(D)	Log probability of data
MAGIC	Multiparent advanced generation intercross
MLM	Mixed linear model
NSGC	National Small Grains Collection
PCR	Polymerase chain reaction
QTL	Quantitative trait loci
SNP	Single-nucleotide polymorphism

Introduction

The wheat coleoptile is a sheath-like structure that helps deliver the stem and first leaf from embryo to the soil surface. The coleoptile is critical for crop establishment because its length only determines the maximum depth at which seeds can be sown (Rebetzke et al. 2007, 2014). In the low-precipitation dryland regions of the Great Plains and Pacific

Communicated by S. Dreisigacker.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-016-2820-1) contains supplementary material, which is available to authorized users.

✉ Xiangyang Xu
xiangyang.xu@ars.usda.gov

[1] USDA-ARS Wheat, Peanut, and Other Field Crops Research Unit, Stillwater, Oklahoma, USA

[2] USDA-ARS Hard Winter Wheat Genetics Research Unit, Manhattan, Kansas, USA

[3] Plant and Soil Science Department, Oklahoma State University, Stillwater, Oklahoma, USA

[4] Entomology and Plant Pathology Department, Oklahoma State University, Stillwater, Oklahoma, USA

[5] USDA-ARS Small Grains and Potato Germplasm Research Unit, Aberdeen, Idaho, USA

Published online: 14 November 2016

Northwest, deep sowing is required to utilize moisture in the soil for wheat germination (Budak et al. 1995; Schillinger et al. 1998). Deep sowing also avoids animal predation of seeds (Brown et al. 2003) and phytotoxicity associated with some pre-emergent herbicides (O'Sullivan et al. 1985). However, sowing depth exceeding the coleoptile length results in poor stand establishment, late emergence, and slow early leaf development (Whan 1976a; Schillinger et al. 1998). Thus, increased coleoptile length is essential for deep-sown wheat, and its importance is most pronounced for dual-purpose wheat production systems where winter wheat is planted early for both forage and grain production. Short coleoptile and deep seed placement in hot soils together may have devastating effects on stand establishment of dual-purpose wheat (Stockton et al. 1996; Bai et al. 2004).

Genetic variation for coleoptile length has been reported (Whan 1976a, b; Schillinger et al. 1998). Previous studies showed that coleoptile length is under polygenic control with high heritability and strong additive effects (Rebetzke et al. 2004, 2007; Murphy et al. 2008), indicating the potential for genetic increase of coleoptile length. Reduced plant height (*Rht*) genes such as *Rht-B1b* (*Rht1*) and *Rht-D1b* (*Rht2*) have large pleiotropic effects on coleoptile length in wheat (Ellis et al. 2004; Rebetzke et al. 2007; Yu and Bai 2010; Li et al. 2011). Unlike standard cultivars that are responsive to gibberellic acid (GA), wheat varieties with *Rht-B1b* or *Rht-D1b* are insensitive to endogenous GA, thus producing not only short plants, but also seedlings with short coleoptiles (Keyes et al. 1989; Addisu et al. 2009). In contrast, *Rht* genes *4*, *5*, *7*, *8*, *9*, *12*, *13,* and *14* reduce plant height in wheat but do not affect coleoptile length (Ellis et al. 2004; Botwright et al. 2005). These genes provide an opportunity to breed dwarf cultivars with long coleoptiles.

Quantitative trait loci (QTL) for coleoptile length have been mapped to chromosomes 1A, 1B, 2B, 2D, 3D, 4A, 4B, 4D, 5A, 5B, 5D, 6A, and 6B, with two major QTL mapped on 4BS (*Rht-B1b*) and 4DS (*Rht-D1b*) (Rebetzke et al. 2001, 2007, 2014; Spielmeyer et al. 2007; Yu and Bai, 2010; Li et al. 2011). The *Rht-B1b* and *Rht-D1b* genes from Norin 10 contributed greatly to the Green Revolution and are widely distributed in modern wheat cultivars. Thus, it is a challenge to breed semi-dwarf long-coleoptile wheat cultivars using these cultivars. More recently, a QTL on 1AS with a large effect on coleoptile length and more than 10 QTL with minor effects were identified in a multi-parent advanced generation intercross (MAGIC) population, and when combined, the minor QTL showed the potential increase of coleoptile length up to 70 mm (Rebetzke et al. 2014), suggesting that pyramiding multiple QTL for long coleoptile can be an effective strategy to increase coleoptile length.

Recently, the genome-wide association study (GWAS) has been widely used to identify allelic variants associated with traits of agronomic importance, because it is relatively easy to implement, has the potential to improve mapping resolution, and can detect genes present in multiple germplasm lines. A gene-level mapping resolution of GWAS has been achieved in rice and maize (Huang et al. 2012; Chen et al. 2014; Chia et al. 2012). In wheat, GWAS has also been used to identify resistance loci for soil-borne mosaic virus (Zhang et al. 2011), leaf rust (Maccaferri et al. 2010; Kertho et al. 2015; Li et al. 2016), stripe rust (Maccaferri et al. 2015), and stem rust (Yu et al. 2012), as well as loci for yield (Sukumaran et al. 2015), quality traits (Breseghello and Sorrells 2006), and plant height (Zanke et al. 2014). In addition, a large set of germplasm collected at National Small Grain Collection (NSGC) has been genotyped with wheat 9 K SNP arrays, providing valuable resources for wheat genetic studies (Bonman et al. 2015).

To date, coleoptile length QTL mapping studies have usually used bi-parental or multiple parental populations. The objectives of this study were to explore allelic diversity in a large germplasm collection for QTL associated with long coleoptile using GWAS, and to gain insight into the genetic basis of coleoptile length in wheat.

Materials and methods

Plant materials

The worldwide collection of 893 winter or facultative wheat accessions used in this study consists of 12 accessions from 7 countries in Africa, 216 from 19 countries in Asia, 534 from 38 countries in Europe, 47 from 5 countries in South America, 73 from the USA and Canada, and 11 from Australia, New Zealand and Cuba. The collection also includes 159 historical breeding lines, 285 registered cultivars, 110 accessions of uncertain improvement status, 331 landraces, and 8 genetic stocks (Table 1). All accessions were selected from the USDA-ARS NSGC core subset of wheat germplasm, and seeds were provided by the NSGC.

Evaluation of coleoptile length

Coleoptile lengths of 893 accessions were evaluated in 2015 using a modified blotter-paper germination protocol (Yu and Bai 2010), and the experiment was repeated in 2016. There were three replications for each experiment. In brief, a paper germination towel was soaked in distilled water and then placed on top of a piece of wax paper. Fifteen seeds from each accession were lined up at 1 cm apart and 7 cm from the bottom. The paper towel and wax paper were rolled loosely, secured with a rubber band, and arranged vertically in a plastic box. The plastic boxes were covered with a black bag and put in a refrigerator at 4 °C for 24 h to break dormancy. The boxes were then moved to

Table 1 Distribution of 893 wheat accessions used in this study. The number of countries in each continent is given in the parentheses

Improvement status	Africa (7)	Asia (19)	Europe (38)	Northern America (2)	Southern America (5)	Others (3)
Breeding lines	4		71	50	30	4
Cultivars	1	26	235	13	5	5
Uncertain	1	16	86	2	3	2
Landraces	6	174	142	0	9	
Genetic stocks				8		
Total	12	216	534	73	47	11

a growth chamber for 7 days at 15 °C, followed by 7 days at 20 °C. Coleoptile lengths were measured with a ruler as the distance from the scutellum to the tip of coleoptile.

Genotyping the association panel with perfect markers for *Rht-B1b* and *Rht-D1b*

A previously described protocol (Dubcovsky et al. 1994) was used to extract genomic DNA from 2-week-old leaves of each accession. The association panel was genotyped with molecular markers specific for the *Rht-B1b* and *Rht-D1b* genes, as well as for the corresponding tall alleles *Rht-B1a* and *Rht-D1a*, using the protocol described by Ellis et al. (2002). PCR products were separated on a 1% agarose gel and visualized after the gels were stained with ethidium bromide.

Data analysis

The SAS PROC ANOVA and PROC CORR procedures (http://www.sas.com; SAS, Cary, NC, USA) were used to analyze coleoptile length variance and the correlation between two experiments, respectively. The broad-sense heritability of coleoptile length was estimated based on ANOVA results.

The NSGC subcore set of wheat germplasm was genotyped previously using the wheat 9 k SNP arrays and DArT Arrays (Bonman et al. 2015). Among the 5011 SNP markers scored (http://www.triticeaecap.org), 4716 were previously mapped (Wang et al. 2014) and used in linkage disequilibrium (LD) analysis in this study. The R package (https://www.r-project.org/) was used to calculate intra-chromosome pair-wise LD values and estimate genome-wide LD decay by plotting LD r^2 from all 21 chromosomes against the corresponding genetic distances. The LD values in the A, B, and D genomes were also estimated.

A set of 4091 non-redundant markers were identified using the tagger function $r^2 = 1.0$ implemented in HAPLOVIEW 4.2 (Barrett et al. 2005), and were used to analyze the population structure. The population structure was analyzed using the admixture model in STRUCTURE 2.3.4 (Pritchard et al. 2000) with a burn-in length of 100,000 and a total of 50,000 Markov chain Monte Carlo iterations to test for a K value (the number of clusters) in the range of 2–15. A method based on the rate of change in the log probability of data (LnP[D]) between successive K values (ΔK) was employed to determine the number of clusters (Evanno et al. 2005).

The mixed linear model (MLM) approach implemented in TASSEL 5.0 was used to identify molecular markers associated with coleoptile length (Yu et al. 2006; Bradbury et al. 2007). Molecular markers, together with the Q matrix inferred from population analysis, were used as fixed effects. A kinship matrix was further calculated with TASSEL 5.0, and used as a random-effects component in MLM analysis to correct for family relatedness. The procedure developed by Benjamini and Hochberg (1995) was used to determine the significant levels of marker–trait association.

Results

Linkage disequilibrium and population structure

A set of 4716 mapped SNPs, including 2220 (47.1%) on the A genome, 2200 (46.6%) on the B genome, and 296 (6.3%) on the D genome, were used to estimate LD values. Supplemental Figure S1 shows the scatter plots of LD values, represented as squared allele-frequency coefficients, between intra-chromosomal SNPs against the genetic distance. The fitted model suggested that LD decayed to <0.1 at 9.7, 9.9, and 10.7 cM in the A, B, and D sub-genome, respectively (Fig. S1 a–c). The genome-wide LD was about 10 cM (Fig. S1 d). Thus, SNPs significantly associated with coleoptile length on the same chromosome were considered to represent the same locus if either the genetic distance between them was less than 10 cM or the LD between them was greater than 0.1.

The genetic stratification was analyzed using the STRUCTURE program to assign fractional membership of accessions to a number of subpopulations (K). Analysis of the rate of change in LnP[D] suggested that the optimum number of subpopulations is 4 (S1–S4), with 51 accessions in S1, 297 accessions in S2, 237 accessions in S3, and 308

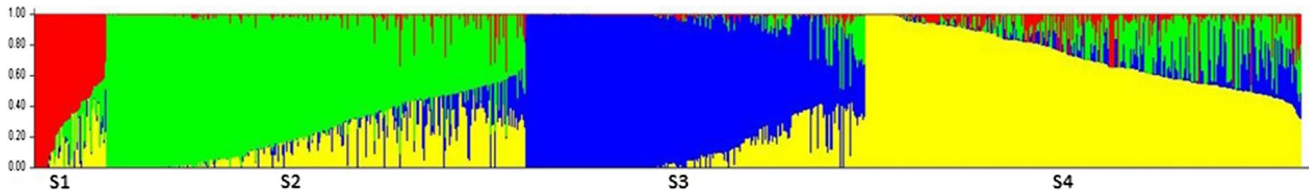

Fig. 1 Structure analysis revealed four subpopulations (*S1–S4*). *S1* and *S3* represent eastern European cultivar group and Asian landrace group, respectively. European cultivars are predominant in *S2*. Subpopulation S4 mainly consists of European cultivars, European landraces and US cultivars. *Vertical lines* show membership coefficients of each accession corresponding to the four subpopulations

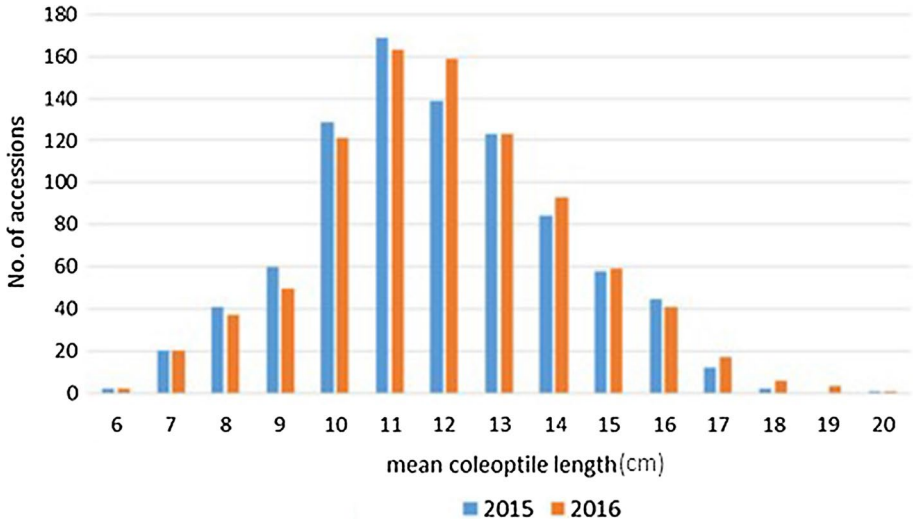

Fig. 2 Distribution of mean coleoptile lengths of the association panel in the 2015 and 2016 experiments

accessions in S4 (Fig. 1). Wheat accessions included in S1 were mainly developed in Romania, Bulgaria, Croatia, Poland, Hungary, Georgia, and the Czech Republic (eastern European cultivar group). In subpopulation S2, cultivars or breeding lines developed in European countries predominated, accounting for 77.1%. Another 13.8% cultivars were developed in Chile, and the remaining originated primarily from the USA and New Zealand. Most of wheat accessions included in S3 were landraces (77.4%), which mainly originated from Asia, including Iran, China, Afghanistan, Pakistan, Bhutan, and Korea. Others in subpopulation S3 were cultivars released from Azerbaijan, Japan, South Korea and China (Asian landrace group). Subpopulation S4 mainly consisted of European cultivars or breeding lines (39.6%), European landraces (27.9%), and U.S. breeding lines (18.2%). A small number of cultivars or landraces that were either released or collected from Argentina, Peru, China, Kenya, and South Korea were also included in this group.

Phenotypic analysis

ANOVA analysis revealed significant variation for coleoptile length in the GWAS panel ($p < 0.0001$), ranging from 5.7 to 20.2 cm with a mean of 11.4 cm in 2015, and 5.7 to 20.1 cm with a mean of 11.5 cm in 2016 (Fig. 2). Significant correlation was observed for coleoptile length in two experiments ($r = 0.84195$, $p < 0.0001$), and the estimated broad-sense heritability was 0.85.

The mean coleoptile lengths of S1 to S4 over two years were 10.3, 10.3, 13.3 and 11.2 cm, respectively. The subpopulation S3 had a significantly longer mean coleoptile length than all other three subpopulations ($p < 0.001$), and the mean coleoptile length of subpopulation S4 was significantly greater than those of S1 and S2 ($p < 0.001$) (Fig. 3). Further analysis revealed that the longer coleoptiles of subpopulations S3 and S4 were due to the higher percentage of landraces in these groups. The landraces had significantly ($p < 0.001$) longer coleoptiles (mean length = 13.33 cm) than the breeding lines or cultivars (mean length = 10.28 cm). Landraces accounted for 77.4 and 32.1% of the total accessions in S3 and S4, respectively, compared to 15.7 and 13.1% in S1 and S2, respectively.

Rht1 and *Rht2* genes are closely associated with coleoptile length

GWAS analysis showed that *Rht2* was most significantly associated with coleoptile length in two experiments

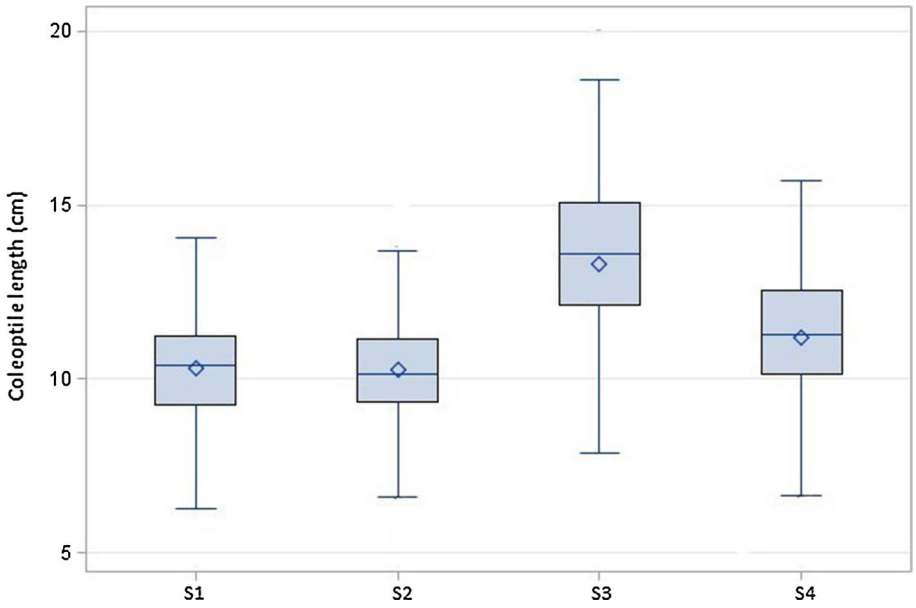

Fig. 3 Box plot showing the distribution of coleoptile length of accessions over 2 years in four subpopulations. The diamond represents mean coleoptile length of each subpopulation

(Table 2). At the *Rht2* locus, 5.3% of accessions carried the dwarf allele, *Rht-D1b*, and the remaining 94.7% of accessions carried the tall allele *Rht-D1a*. The frequencies of *Rht-D1b* allele were 0, 4.5, 5.6, and 7% in subpopulations S1 to S4, respectively. Allele variation at the *Rht2* locus accounted for 6.62 and 7.88% of total phenotypic variation in the two experiments, respectively. The allele substitution effects were 2.19 cm in the 2015 experiment, and 2.32 cm in the 2016 experiment.

As expected, the *Rht1* gene was also very significantly associated with coleoptile length (Table 2). The dwarf allele *Rht-B1b* was detected in 7.6% of accessions in the association panel. The frequencies of *Rht-B1b* allele were 9.8, 7.2, 8.4, and 6.7% in S1 to S4 subpopulations, respectively. The *Rht1* gene explained 3.12–4.16% of total phenotypic variation with substitute effects of 1.23–1.37 cm in the two experiments.

QTL for coleoptile length

Nine SNPs on chromosomes 1A, 1B, 2D, 4B, 4D, and 5B, as well as seven SNPs that were not previously mapped to the 90 K SNP consensus map, were significantly associated with coleoptile length. In the 2015 experiment, 7, 8, and 15 SNPs were significantly associated with coleoptile length at a false discovery rate (FDR) of 0.05, 0.1 and 0.2, respectively, and the corresponding SNP numbers in the 2016 experiment were 6, 9 and 13. Five SNPs, including *IWA5381*, *IWA8564*, *IWA2963*, *IWA1698,* and *IWA2355*, consistently showed significant association with coleoptile length at $FDR_{0.05}$ in both years. *IWA3798* was significant at $FDR_{0.1}$, while *IWA4903*, *IWA3580*, and *IWA2205* were significant at $FDR_{0.2}$ in both experiments. *IWA8304*, *IWA1799,* and *IWA2931* were significant at varying significant levels in the two experiments (Table 2). Also, *IWA2577*, *IWA2578*, *IWA4678*, and *IWA2855* were significant in only one experiment at $FDR_{0.2}$, and these markers were excluded from further consideration.

For SNPs that were not previously mapped to the 90 K SNP consensus map, alignment of SNP sequences with Chinese Spring survey sequence identified *IWA2355*, *IWA2931*, and *IWA3798* in the wheat contigs 1BS_3483406, 5BL_10804384, and 4BS_4952369, respectively (https://triticeaetoolbox.org/wheat/genotyping/marker_report_ref.php?uid2=35). Thus, we tentatively locate *IWA2355*, *IWA2931*, and *IWA3798* to chromosome 1BS, 5BL, and 4BS, respectively. However, the *IWA3580* sequence was found in contigs 5BL_10804384, 5BL_10903610, and 5DL_4543633. Thus, *IWA3580* may be locatfed on 5BL or 5DL. *IWA1698* and *IWA2205* were identified in several contigs and their genomic locations are yet to be determined.

Table 2 lists the 12 SNPs associated with coleoptile length in both experiments. Of these, *IWA8564*, *IWA2963,* and *IWA3798* were mapped on chromosome 4B, and the genetic distance between *IWA8564* and *IWA2963* was 5.9 cM on the 90 K SNP consensus map. LD analysis suggested that *IWA8564* and *IWA2963* were closely associated with an r^2 of 0.296 (Supplemental Fig. S2). Thus, *IWA8564* and *IWA2963* represent the same locus, designated as *QCL.stars-4BS1*. The precise location of *IWA3798* on the 90 K consensus map has not been determined. However, the LD values (r^2) between *IWA3798* and the two SNPs underlying *QCL.stars-4BS1*, *IWA8564* and *IWA2963,* were 0.073 and 0.041, respectively (Supplemental Fig. S2), suggesting that IWA3798 may represent a locus independent of *QCL.*

Table 2 Designated name, linked SNP(s), chromosome arm, map location on the 90 K SNP consensus map, alleles, $-\log_{10}(p)$ value, R^2 value, and allele effect of each gene or QTL identified in 2015 and 2016

G Gene/QTL	Marker	Chrom. arm[a]	Map location (cM)	Allele[b]	2015			2016		
					$-\text{Log}_{10}$(p) [c]	R^2 (%)	Allelic effects (cm)	$-\text{Log}_{10}$ (p)[c]	R^2 (%)	Allelic effects (cm)
Rht2		4DS	31.2	**Rht-D1a** Rht-D1b	*11.92****	6.62	2.19	14.10***	7.88	2.32
Rht1		4BS	43.9	**Rht-B1a** Rht-B1b	*5.71****	3.12	1.23	7.58***	4.16	1.37
QCL.stars-4DC1	IWA5381	4DC	80.7	**A**/G	6.05***	3.18	1.00	5.64***	2.96	0.94
QCL.stars-4BS1	IWA8564	4BS	64.5	A/**G**	5.68***	3.01	1.23	5.86***	3.11	1.24
	IWA2963	4BC	70.3	**C**/T	4.84***	2.53	0.73	4.35***	2.27	0.71
QCL.stars-2DC1	IWA8304	2DC	114.7	A/**G**	4.50***	2.34	0.58	3.86**	2.01	0.63
QCL.stars-2DS1	IWA1799	2DS	6.7	**A**/G	4.14***	1.78	0.73	3.72**	1.58	0.67
QCL.stars-5BL1	IWA4903	5BL	222.9	C/T	3.38*	1.43	0.52	3.28*	1.38	0.49
	IWA2931	5BL		A/**G**	3.56*	1.85	0.68	4.07***	2.12	0.71
QCL.stars-1BS1	IWA2355	1BS		A/**G**	4.53***	1.97	1.15	4.42***	1.92	1.11
	IWA1698	UNK		**C**/T	4.53***	1.97	1.15	4.42***	1.92	1.11
	IWA2205	UNK		**C**/T	3.28*	1.36	1.04	3.17*	1.30	0.98
QCL.stars-4BS2	IWA3798	4BS		**C**/T	3.86**	1.66	1.35	3.84**	1.65	1.27
QCL.stars-5B/5D	IWA3580	5BL/5DL		**A**/G	3.44*	1.45	1.65	3.26*	1.36	1.55

[a] Chromosome arm. L, S, and C refer to long and short chromosome arms, and centromere, respectively

[b] The favorable alleles are in bold type

[c] *, **, *** Significant level of $FDR_{0.2}$, $FDR_{0.1}$ and $FDR_{0.05}$, respectively

stars-4BS1. Thus, we designated the QTL associated with *IWA3798* as *QCL.stars-4BS2*.

IWA4903 and *IWA2931* on 5BL had an LD value of 0.241 (Supplemental Fig. S2), indicating that they were closely linked and represent the same QTL, designated as *QCL.stars-5BL1*. *IWA3580* was previously mapped on both 5BL and 5DL. The LD values between *IWA3580* and the two other SNPs on 5BL, *IWA4903* and *IWA2931*, were 0.024 and 0.066, respectively (Supplemental Fig. S2). Given that these LD values were less than 0.1, we conclude that *IWA3580* is independent of *QCL.stars-5BL1*, and tentatively designated it as *QCL.stars-5B/5D*. Linkage mapping of this locus will be necessary to determine its chromosomal location.

The chromosome locations of *IWA1698* and *IWA2205* are unknown. However, LD analysis revealed that *IWA1698* co-segregated with *IWA2355* ($r^2 = 1$), and both of them were closely linked to *IWA2205* with an r^2 value of 0.922 (Supplemental Fig. S2), suggesting that these three SNPs likely represent the same locus. Thus, we tentatively assigned *IWA1698* and *IWA2205* to 1BS, and designated this locus as *QCL.stars-1BS1*.

Each of these QTL explained 1.45–3.18% of the phenotypic variation in 2015, and 1.36–3.11% in 2016. Their allele substitution effects ranged from 0.58 to 1.65 cm for coleoptile length in 2015, and 0.63–1.55 cm in 2016 (Table 2). The favorable alleles in each accession ranged from 2 to 8 with an average of 5.8 at the eight loci in this association panel (Fig. 5), and a significant association between the number of favorable alleles and coleoptile length was observed (2015: $r = 0.511$, $p < 10^{-7}$; 2016: $r = 0.524$, $p < 10^{-7}$). The coleoptile length was also significantly correlated with the predicated phenotypes represented as the sum of allele substitution effects (2015: $r = 0.476$, $p < 10^{-7}$; 2016: $r = 0.493$, $P < 10^{-7}$), suggesting that pyramiding multiple QTL with marker-assisted selection may effectively increase coleoptile length. Indeed, the accessions that carry only two favorable alleles had a mean coleoptile length of 7.8 cm, while those accessions carrying all eight favorable alleles had an average of 14.5 cm over two years (Fig. 5).

Discussion

Comparison of QTL identified with known loci for coleoptile length

Rht-B1b and *Rht-D1b* are GA-insensitive dwarfing genes encoding proteins involved in GA signal transduction (Peng et al. 1999). *Rht-B1b* and *Rht-D1b* are probably present in around 90% of the world's semi-dwarf wheat cultivars. These cultivars have lower risk of lodging and greater partitioning of assimilates to grains and are responsible for the large increases in wheat yield (Gale and Youssefian 1985). However, these genes decrease epidermal cell length in leaf and stem tissue, resulting in semi-dwarfing stature, short coleoptile, smaller leaf area, and reduced seedling vigor (Wright 1966; Keyes et al. 1989; Ellis et al. 2004). These undesirable side-effects, especially the short coleoptile, are major issues for wheat production under certain conditions (Rebetzke et al. 1999; Bai et al. 2004). Moreover, a strong association between plant height and coleoptile length was observed in GA-insensitive semi-dwarf wheat populations, indicating that it is difficult to obtain long coleoptile lines without increasing plant height if GA-insensitive semi-dwarf genes are used (Whan 1976a; Keyes et al. 1989; Rebetzke et al. 1999). GA-insensitive semi-dwarf genes also include *Rht11* and *Rht17* (Ellis et al. 2004).

Rht-D1b showed the most significant effects on coleoptile length (Table 2). Of QTL identified in this study, *QCL.stars-4DC1* was mapped to the centromeric region of chromosome 4D, and also showed very significant association with coleoptile length in both 2015 and 2016. The genetic distance between *Rht-D1b* and *QCL.stars-4DC1* is 49.5 cM on the 90 K SNP consensus map (Wang et al. 2014) (Fig. 4). Thus, *QCL.stars-4DC1* is independent of *Rht-D1b*. In addition, *QCL.stars-4DC1* coincides with SNP markers *IWA4580* (*wsnp_Ex_c683_1341113*) and *IWA4180* (*wsnp_Ex_c14026_21924297*) on the 90 K consensus map, and these two SNPs were significantly associated with coleoptile length in a MAGIC population (Rebetzke et al. 2014), lending credence to *QCL.stars-4DC1*.

In agreement with previous studies (Rebetzke et al. 2007; Li et al. 2011), the *Rht-B1b* gene on chromosome 4BS significantly reduced coleoptile length in the present study. We identified two QTL for coleoptile length on 4BS, *QCL.stars-4BS1* and *QCL.stars-4BS2*. The genetic distance between *Rht-B1b* and *QCL.stars-4BS1* is 20.6 cM on the 90 K SNP consensus map. However, the LD between *Rht-B1b* and the SNP underlying *QCL.stars-4BS1*, *IWA8564*, was as high as 0.2, suggesting that further study is needed to determine whether *QCL.stars-4BS1* represents the *Rht-B1b* gene. Also, another marker on 4BS, *IWA1846* (*wsnp_Ex_c14026_21924297*), was significantly associated with coleoptile length at both 12 and 20 °C in a MAGIC population (Rebetzke et al. 2014). Given that *IWA1846* was 0.33 cM distant from *Rht-B1b* on the 90 K SNP consensus map (Wang et al. 2014) (Fig. 5), *IWA1846* likely represented the *Rht-B1b* gene in the MAGIC population (Rebetzke et al. 2014). The precise location of *QCL.stars-4BS2* is not available. However, this QTL is likely a novel locus independent from *Rht-B1b*, evidenced by the fact that the LD between them was less than 0.1 (0.08).

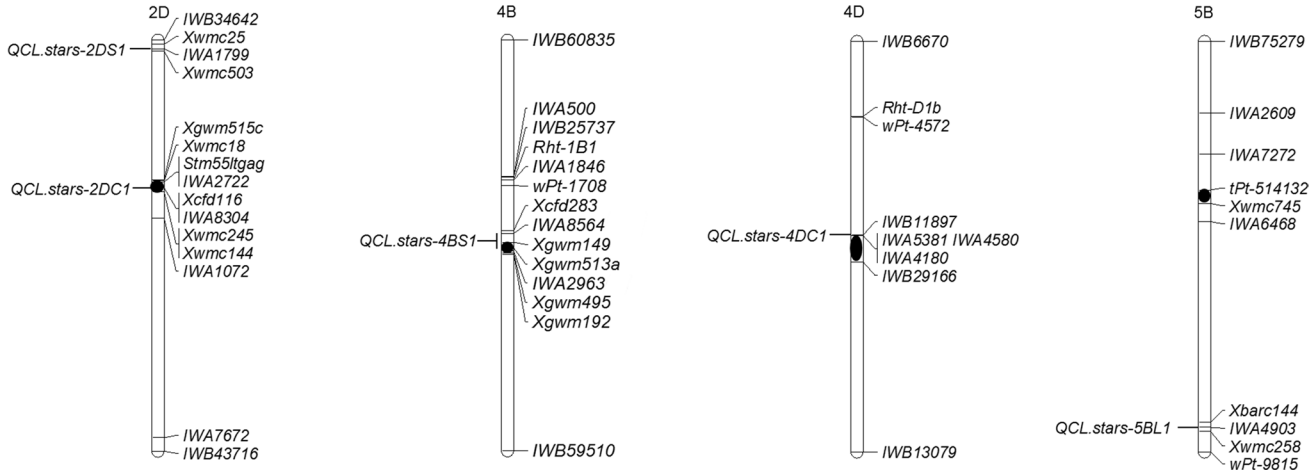

Fig. 4 Chromosome positions of five QTL for coleoptile length identified in this study. Marker locations are based on the SNP 90 K consensus map, and chromosome lengths are standardized to the same relative length. Precise map locations of other three QTL identified in this study, *QCL.stars-1BS1*, *QCL.stars-4BS2,* and *QCL.stars-5B/5D*, are not available

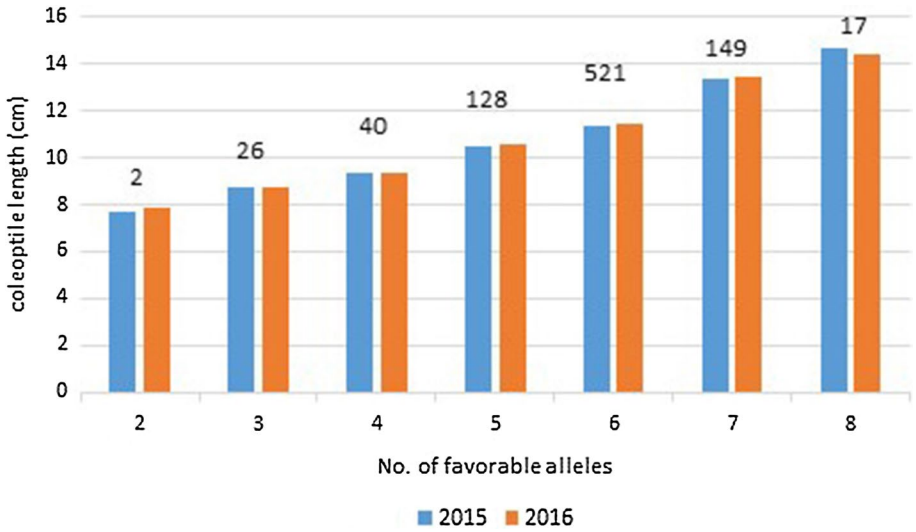

Fig. 5 Mean coleoptile lengths of wheat germplasm with a varying number of favorable alleles at eight loci for coleoptile length in two experiments. The number of accessions for each class is given on the top of bars

QCL.stars-2DS1 and *QCL.stars-2DC1* were present in the terminal region of 2DS and the centromeric region of chromosome 2D, respectively. A QTL for coleoptile length was mapped to chromosome 2D in three DH mapping populations, and the nearest flanking markers included *Xgwm515c*, *Xwmc18*, and *Xstm551tgag* (Rebetzke et al. 2007). These SSR markers were located in the centromeric region of 2D on the 90 K SNP consensus map. In addition, two SNPs on chromosome 2D, *IWA7672* (*wsnp_Ra_c17636_26538543*) and *IWA1072* (*wsnp_CAP7_c2782_1329707*), were associated with coleoptile length, and another SNP, *IWA2722* (*wsnp_Ex_c2251_4218338*), was associated with coleoptile width in a MAGIC population (Rebetzke et al. 2014). Of these, *IWA2722* also resides in the centromeric region of 2D, and is in proximity to *QCL.stars-2DC1* (Fig. 4). *QCL.stars-2DC1* likely represents the locus identified in the MAGIC and DH mapping populations. Since *QCL.stars-2DS1* is far away from any known loci, it is likely a novel QTL for coleoptile length.

Three SNPs on 1BL, *IWA1757* (*wsnp_Ex_c13310_20984763*), *IWA6688* (*wsnp_Ku_c207_407862*), *and IWA725* (*wsnp_Ex_c14760_22866930*), were significantly associated with coleoptile length in a previous study (Rebetzke et al. 2014). Of these, *IWA725* is within the genomic region harboring a QTL for long coleoptile in the wheat cultivar Wheaton, which is flanked by *Xgwm140* and *XpACTG-mCGCT201* (Yu and Bai 2010). The representative SNP of *QCL.stars-1BS1* was identified in a contig mapped to the short arm of chromosome 1B. No locus for coleoptile length has been previously reported on 1BS.

Although the precise locations of the three SNPs underlying *QCL.stars-1BS1* are not yet known, *QCL.stars-1BS1* is most likely a novel locus for coleoptile length.

Previous studies revealed a few SNPs for coleoptile length on chromosome 5B, including *IWA2609* (*wsnp_Ex_c214_421541*), *IWA7272* (*wsnp_Ku_c6464_11320381*), and *IWA6468* (*wsnp_Ku_c12562_20256747*) (Rebetzke et al. 2014). The precise location of another QTL on 5B (Yu and Bai 2010) cannot be assessed in the 90 K SNP consensus map, because it was mainly mapped with AFLP markers. *QCL.stars-5BL1* is likely a new locus, because it resides in the terminal region of 5BL on the 90 K SNP map, and is distant from *IWA2609*, *IWA7272,* and *IWA6468* (Fig. 4).

At present, the locations of *QCL.stars-5B/5D* cannot be determined, and its relationships with previously reported QTL for coleoptile length are also unknown. Further mapping of this QTL may enhance our understanding of the genetic basis of coleoptile length.

Breeding for long-coleoptile wheat

An important strategy for breeding long-coleoptile wheat cultivars is to use GA-responsive dwarfing genes that have no major effect on coleoptile length or leaf growth, such as *Rht4*, *5*, *7*, *8*, *13*, and *14* (Ellis et al. 2004). Rebetzke et al. (1999) demonstrated that plant height and coleoptile length were under independent genetic controls among GA-sensitive wheats, indicating the feasibility of selecting semi-dwarf and long-coleoptile cultivars with improved establishment and seedling vigor using these GA-responsive genes. Indeed, one of these genes, *Rht8*, has been extensively used in southern Europe since the 1920s and was introduced into Australian wheat as a replacement of GA-insensitive dwarfing genes (Spielmeyer et al. 2001). The *Rht8* gene was first identified in the Japanese cultivar Akakomugi, and an allele at this locus has shown moderate reduction in plant height. Additional reductions resulted in plant height when Rht8 was combined with *Ppd-D1*, a closely linked photoperiod-insensitive, height-reducing gene (Worland et al. 1998). Korzun et al. (1998) found that the 192-bp allele of an SSR marker *Xgwm261*, 0.6 cM from the *Rht8* locus, was associated with the more commercially favorable *Rht8* allele, and can be used as an indicator of the *Rht8* gene. This discovery makes it possible to combine *Rht8* with other QTL for long coleoptiles. In this study, we identified a total of eight loci for coleoptile length. Of these, three QTL coincide with known loci and another four are likely novel loci. Although the precise genomic locations of three loci are yet to be discovered, the identified SNPs from SNP arrays can be easily converted to PCR-based assays (such as KASP markers) and used in marker-assisted selection. The combination of these QTL with the *Rht8* gene may result in long-coleoptile wheat cultivars with reduced height.

Notably, not all GA-responsive dwarfing genes are suitable for breeding long-coleoptile wheat. For example, *Rht16* and *Rht18* are GA-responsive, but they also have negative effects on coleoptile length and leaf elongation (Ellis et al. 2004). Another two GA-responsive genes, *Rht9* and *Rht12*, were negatively associated with yield (Worland and Snape 2001). These genes may not be ideal for wheat breeding.

In the present study, a total of 72 accessions, including one breeding line, three cultivars, and 68 landraces, had a mean coleoptile length of 15 cm or greater (Supplemental table S1). Each accession carries six to eight QTL for coleoptile length, and no *Rht-B1b* and *Rht-D1b* genes were identified in these accessions (Supplemental table S1). These accessions can be used in breeding long-coleoptile, semi-dwarf cultivars.

Conclusions

Understanding the genetic basis of coleoptile length is essential for wheat-breeding programs in low-precipitation regions, and GWA is an efficient way to identify QTL for coleoptile length. In this study, *Rht-B1b* and *Rht-D1b* showed very significant effects on coleoptile length, and eight additional QTL for coleoptile length were also identified. Three of these QTL, *QCL.stars-4DC1*, *QCL.stars-4BS1,* and *QCL.stars-2DC1,* were mapped in proximity to known genes or QTL for coleoptile length. Another four QTL, *QCL.stars-1BS1*, *QCL.stars-2DS1*, *QCL.stars-4BS2,* and *QCL.stars-5BL1*, are likely novel loci for coleoptile length. Each QTL explained 1.45–3.18 and 1.36–3.11% of the total phenotypic variation in 2015 and 2016, respectively. The allele substitution effects of these QTL ranged from 0.31 to 1.75 cm in 2015, and 0.63 to 1.55 cm in 2016. The number of favorable alleles in each accession ranged from two to eight, and was significantly correlated with coleoptile length, suggesting that pyramiding multiple QTL is an effective way to enhance the coleoptile length. The combination of the QTL identified in this study with GA-responsive dwarfing genes may result in elite cultivars with long coleoptile and reduced height.

Author contribution statement XX, GL, GB, and BFC designed the research; GL, NE, RB, and RH contributed to the phenotyping of the association panel; GL, YW, and JMB contributed to the genotyping of the association panel; XX wrote the paper. All authors read, revised, and approved the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

Addisu M, Snape JW, Simmonds JR, Gooding MJ (2009) Reduced height (Rht) and photoperiod insensitivity (Ppd) allele associations with establishment and early growth of wheat in contrasting production systems. Euphytica 166:249–267

Bai G, Das MK, Carver BF et al (2004) Covariation for microsatellite marker alleles associated with *Rht8* and coleoptile length in winter wheat. Crop Sci 44:1187–1194

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265. doi:10.1093/bioinformatics/bth457

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57:289–300

Bonman JM, Babiker EM, Cuesta-Marcost A et al (2015) Genetic diversity among wheat accessions from the USDA National Small Grains Collection. Crop Sci 55:1243

Botwright TL, Rebetzke GJ, Condon AG, Richards RA (2005) Influence of the gibberellin-sensitive Rht8 dwarfing gene on leaf epidermal cell dimensions and early vigour in wheat (Triticum aestivum L.). Ann Bot 95:631–639

Bradbury PJ, Zhang Z, Kroon DE et al (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. doi:10.1093/bioinformatics/btm308

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1177. doi:10.1534/genetics.105.044586

Brown PR, Singleton GR, Tann CR, Mock I (2003) Increasing sowing depth to reduce mouse damage to winter crops. Crop Prot 22:653–660. doi:10.1016/S0261-2194(03)00006-1

Budak N, Baenziger PS, Eskridge KM et al (1995) Plant height response of semidwarf and nonsemidwarf wheats to the environment. Crop Sci 35:447–451

Chen W, Gao Y, Xie W et al (2014) Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. Nat Genet 46:714–721. doi:10.1038/ng.3007

Chia JM, Song C, Bradbury PJ et al (2012) Maize HapMap2 identifies extant variation from a genome in flux. Nat Genet 44:803–807. doi:10.1038/ng.2313

Dubcovsky J, Galvez AF, Dvořák J (1994) Comparison of the genetic organization of the early salt- stress-response gene system in salt-tolerant *Lophopyrum elongatum* and salt-sensitive wheat. Theor Appl Genet 87:957–964

Ellis M, Spielmeyer W, Gale K et al (2002) " Perfect" markers for the *Rht-B1b* and *Rht-D1b* dwarfing genes in wheat. Theor Appl Genet 105(6–7):1038–1042

Ellis MH, Rebetzke GJ, Chandler P et al (2004) The effect of different height reducing genes on the early growth of wheat. Funct Plant Biol 31:583–589. doi:10.1071/fp03207

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Gale MD, Youssefian S (1985) Dwarfing genes in wheat. In: Russell GE (Ed). Progress in Plant Breeding, pp 1–35. Butterworths, London

Huang X, Zhao Y, Wei X et al (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat Genet 44:32–39

Kertho A, Mamidi S, Bonman JM et al (2015) Genome-wide association mapping for resistance to leaf and stripe rust in winter-habit hexaploid wheat landraces. PLoS One 10:e0129580. doi:10.1371/journal.pone.0129580

Keyes GJ, Paolillo DJ, Sorrells ME (1989) The effects of dwarfing genes *Rht1* and *Rht2* on cellular dimensions and rate of leaf elongation in wheat. Ann Bot 64:683–690

Korzun V, Roder MS, Ganal MW, Worland AJ, Law CN (1998) Genetic analysis of the dwarfing gene (*Rht8*) in wheat. Part I. Molecular mapping of *Rht8* on the short arm of chromosome 2D of bread wheat (*Triticum aestivum* L.). Theor Appl Genet 96:1104–1109

Li P, Chen J, Wu P et al (2011) Quantitative trait loci analysis for the effect of dwarfing gene on coleoptile length and seedling root length and number of bread wheat. Crop Sci 51:2561–2568

Li G, Xu X, Bai G, Carver BF, Hunger R, Bonman M, Kolmer J, and Dong H (2016) Genome-wide association mapping reveals novel quantitative trait loci for leaf rust resistance in a worldwide collection of winter wheat. The Plant Genome (in press)

Maccaferri M, Sanguineti MC, Mantovani P et al (2010) Association mapping of leaf rust response in durum wheat. Mol Breed 26:189–228. doi:10.1007/s11032-009-9353-0

Maccaferri M, Zhang J, Bulli P, et al (2015) A genome-wide association study of resistance to stripe rust (*Puccinia striiformis* f. sp. *tritici*) in a worldwide collection of hexaploid spring wheat (*Triticum aestivum* L.). G3: Genes|Genomes|Genetics 5:449–465. doi: 10.1534/g3.114.014563

Murphy K, Balow K, Lyon SR, Jones SS (2008) Response to selection, combining ability and heritability of coleoptile length in winter wheat. Euphytica 164:709–718

O'Sullivan PA, Weiss GM, Friesen D (1985) Tolerance of spring wheat (*Triticum aestivum* L.) to trifluralin deep-incorporated in the autumn or spring. Weed Res 25:275–280

Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D (1999) 'Green revolution'genes encode mutant gibberellin response modulators. Nature 400(6741):256–261

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959. doi:10.1111/j.1471-8286.2007.01758.x

Rebetzke GJ, Richards RA, Fischer VM, Mickelson BJ (1999) Breeding long coleoptile, reduced height wheats. Euphytica 106:159–168. doi:10.1023/A:1003518920119

Rebetzke GJ, Appels R, Morrison AD et al (2001) Quantitative trait loci on chromosome 4B for coleoptile length and early vigour in wheat (*Triticum aestivum* L.). Aust J Agric Res 52:1221–1234. doi:10.1071/ar01042

Rebetzke GJ, Richards RA, Sirault XRR, Morrison AD (2004) Genetic analysis of coleoptile length and diameter in wheat. Aust J Agric Res 55:733–743

Rebetzke GJ, Richards RA, Fettell NA et al (2007) Genotypic increases in coleoptile length improves stand establishment, vigour and grain yield of deep-sown wheat. F Crop Res 100:10–23. doi:10.1016/j.fcr.2006.05.001

Rebetzke GJ, Verbyla AP, Verbyla KL et al (2014) Use of a large multiparent wheat mapping population in genomic dissection of coleoptile and seedling growth. Plant Biotechnol J 12:219–230. doi:10.1111/pbi.12130

Schillinger WF, Donaldson E, Allan RE, Jones SS (1998) Winter wheat seedling emergence from deep sowing depths. Agron J 90:582–586. doi:10.2134/agronj1998.00021962009000050002x

Spielmeyer W, Bonnett D, Ellis M et al (2001) Implementation of molecular markers to improve selection efficiency in CSIRO

wheat breeding program. 'Proceedings 10th Australian Wheat Breeders Assembly (Wheat Breeding Society of Australia)

Spielmeyer W, Hyles J, Joaquim P et al (2007) A QTL on chromosome 6A in bread wheat (*Triticum aestivum*) is associated with longer coleoptiles, greater seedling vigour and final plant height. Theor Appl Genet 115:59–66

Stockton RD, Krenzer EGJ, Solie J, Payton ME (1996) Stand establishment of winter wheat in Oklahoma: a survey. J Prod Agric 9:571–575

Sukumaran S, Dreisigacker S, Lopes M et al (2015) Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. Theor Appl Genet 128:353–363. doi:10.1007/s00122-014-2435-3

Wang S, Wong D, Forrest K et al (2014) Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. Plant Biotechnol J 12:787–796. doi:10.1111/pbi.12183

Whan BR (1976a) The association between coleoptile length and culm length in semidwarf and standard wheats. Journal of the Australian Institute of Agricultural Science

Whan BR (1976b) The emergence of semidwarf and standard wheats, and its association with coleoptile length. Aust J Exp Agric 16:411–416

Worland T, Snape JW. 2001. Genetic analysis of worldwide wheat varietal improvement. P. – 100. In A.P. Bonjean and W.J. Angus (ed.) The world wheat book: A history of wheat breeding. Lavoisier Publishing, Paris

Worland AJ, Korzun V, Röder MS et al (1998) Genetic analysis of the dwarfing gene *Rht8* in wheat. Part II. The distribution and adaptive significance of allelic variants at the *Rht8* locus of wheat as revealed by microsatellite screening. Theor Appl Genet 96:1110–1120

Wright STC (1966) Growth and cellular differentiation in the wheat coleoptile (*Triticum vulgare*): II. Factors influencing the growth response to gibberellic acid, kinetin and indolyl-3-acetic acid. J Exp Bot 17:165–176

Yu JB, Bai GH (2010) Mapping quantitative trait loci for long coleoptile in Chinese wheat landrace Wangshuibai. Crop Sci 50:43–50. doi:10.2135/cropsci2009.02.0065

Yu J, Pressoir G, Briggs WH et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208. doi:10.1038/ng1702

Yu LX, Morgounov A, Wanyera R et al (2012) Identification of Ug99 stem rust resistance loci in winter wheat germplasm using genome-wide association analysis. Theor Appl Genet 125:749–758. doi:10.1007/s00122-012-1867-x

Zanke CD, Ling J, Plieske J, Kollers S, Ebmeyer E, Korzun V, Argillier O, Stiewe G, Hinze M, Neumann K, Ganal MW (2014) Whole genome association mapping of plant height in winter wheat (Triticum aestivum L.). PLoS One 9(11):e113287

Zhang D, Bai G, Hunger RM et al (2011) Association study of resistance to *Soilborne wheat mosaic virus* in U.S. winter wheat. Phytopathology 101:1322–1329. doi:10.1094/PHYTO-02-11-0041