bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Unravelling consensus genomic regions associated with quality traits in wheat (*Triticum aestivum* L.) using meta-analysis of quantitative trait loci

Santosh Gudi[1*], Dinesh K Saini[1], Gurjeet Singh[1], Priyanka Halladakeri[2], Mohammad Shamshad[1], Mohammad Jafar Tanin[1], Pradeep Kumar[1] and Achla Sharma[1]

[1]Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, Punjab, India

[2]Department of Genetics and Plant Breeding, Anand Agricultural University, Gujarat, India

* Corresponding author

E-mail: santosh-pbg@pau.edu (SG)

Author details,

1. **Santosh Gudi**
   Ph.D. student, Punjab Agricultural University, Ludhiana, India
   Email ID.: Santosh-pbg@pau.edu
   Phone no.: +91-7986168478
   ORCID: 0000-0003-1885-8139
   (**Corresponding author**)

2. Dinesh Kumar Saini
   Ph.D. student, Punjab Agricultural University, Ludhiana, India
   Email ID.: dineshsaini-pbg@pau.edu
   Phone no.: +91-6280758417

3. Gurjeet Singh
   Ph.D. student, Punjab Agricultural University, Ludhiana, India
   Email ID.: gurjeet-pbg@pau.edu
   Phone no.: +91-9462966845
   ORCID: 0000-0002-9995-612X

4. Priyanka Halladakeri
   Ph.D. student, Anand Agricultural University, Gujarat, India
   Email ID.: priyankahalladakeri6798@gmail.com
   Phone no.: +91-7204170357

5. Pradeep Kumar
   M.Sc. Student, Punjab Agricultural University, Ludhiana, India
   Email ID.: pra12jangid@gmail.com
   Phone no.: +91-8949474953

6. Mohammad Shamshad
   Research Scholar, Punjab Agricultural University, Ludhiana, India
   Email ID.: shamshad.rattan@gmail.com
   Phone no.: +91-9653188786
   ORCID: 0000-0002-5002-4575

7. Mohammad Jafar Tanin
   Ph.D. student, Punjab Agricultural University, Ludhiana, India
   Email ID.: jafartanin@gmail.com
   Phone no.: +91-8557919975

8. Dr Achla Sharma
   Senior Wheat Breeder, Punjab Agricultural University, Ludhiana, India
   Email ID.: achla@pau.edu
   Phone no.: +91-8146111166

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Main Conclusion

Meta-QTL analysis in wheat for major quality traits identified 110 MQTLs with reduced confidence interval. Candidate gene mining and expression analysis discovered differentially expressed genes involve in quality traits.

Abstract

A meta-analysis of quantitative trait loci (QTLs) associated with following six major quality traits (i) arabinoxylan, (ii) dough rheology properties, (iii) nutritional traits, (iv) polyphenol content, (v) processing quality traits, and (vi) sedimentation volume was conducted in wheat. For this purpose, as many as 2458 QTLs were collected from the 50 mapping studies published during 2013-20. Of the total QTLs, 1126 QTLs were projected on to the consensus map saturated with 2,50,077 markers resulting into the identification of 110 meta-QTLs (MQTLs) with average confidence interval (CI) of 5.6 cM. These MQTLs had 18.84 times reduced CI compared to CI of initial QTLs. Fifty-one (51) MQTLs were also verified with the marker-trait associations (MTAs) detected in earlier genome-wide association studies (GWAS). Physical region occupied by a single MQTL ranged from 0.12 to 749.71 Mb with an average of 130.25 Mb. Candidate gene mining allowed the identification of 2533 unique gene models from the MQTL regions. *In-silico* expression analysis discovered 439 differentially expressed gene models with >2 transcripts per million (TPM) expression in grains and related tissues which also included 44 high-confidence candidate genes known to be involved in the various cellular and biochemical processes related to quality traits. Further, nine functionally characterized wheat genes associated with grain protein content, high molecular weight glutenin and starch synthase enzymes were also found to be co-localized with some of the MQTLs. In addition, synteny analysis between wheat and rice MQTL regions identified 23 wheat MQTLs syntenic to 16 rice MQTLs. Furthermore, 64 wheat orthologues of 30 known rice genes were detected in 44 MQTL regions. These genes encoded proteins mainly belonging to the following families: starch synthase, glycosyl transferase, aldehyde dehydrogenase, SWEET sugar transporter, alpha amylase, glycoside hydrolase, glycogen debranching enzyme, protein kinase, peptidase, legumain and seed storage protein enzyme.

Keywords

Meta-QTLs, Wheat, Quality traits, MTAs, Candidate genes, Synteny

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Wheat (*Triticum aestivum* L.) is a globally grown cereal crop and is a major contributor of calories and protein to the human diet. Currently, wheat is widely consumed and processed into bread, noodles, cakes, pasta, beer, and other products. Wheat research has greatly contributed to its yield enhancement and disease resistance, but focus on quality of the produce took the back stage while enhancing yield. Hence, developing the high yielding varieties with enhanced quality characters is the foremost concern of the breeders (Nuttall et al. 2017). Improving end-use qualities is a tough endeavour because firstly, it is very difficult to measure the seed quality and rheological properties such as grain protein content (GPC), sedimentation rate (SDS), hectolitre weight (HW), 1000-grain weight (TGW), wet gluten content (WGC), dry gluten content (DGC), flour water absorption (FWA), dough development time (DDT), dough stability time (DST), mixing tolerance index (MTI), break down time (BDT) and kernel hardness (KH) as they are labour intensive and also require much seeds for analysis. Secondly, quality characters are complex traits governed by a plethora of gene networks that are largely influenced by several environmental conditions (Quraishi et al. 2017).

Quantitative trait locus (QTL) analysis has emerged as an effective approach for dissecting complex traits into component loci and studying the relative effects of the loci on the target trait (Doerge 2002). Since the first report on QTL analyses for wheat quality traits published in 1990s (Powell 1990), a plentiful of QTLs have been identified using different mapping populations to underpin the genetic architecture underlying end-use quality, including GPC (Huang et al. 2006; Mann et al. 2009), dough rheological properties (Huang et al. 2006; Mann et al. 2009), SDS (McCartney et al. 2006), falling number (FN) (Fofana et al. 2009) and starch pasting properties (Mohler et al. 2014). However, the rationality of these QTL mapping results is strongly influenced by the experimental conditions, type and size of mapping population, density of genetic markers, statistical methods used among others (Swamy et al. 2011). Thus, the practical implication of these QTLs for quality improvement via molecular QTL cloning and marker-assisted selection has been rather limited (Quraishi et al. 2017). Considering this challenge, it is desirable to identify QTLs that show major effect on target phenotype and are consistently detected across the multiple genetic backgrounds and environments.

Meta-analysis of available QTLs enable the identification of consensus and robust QTLs or MQTL regions that are most frequently associated with trait variation in diverse studies and reduce their confidence intervals (CIs) (Goffinet and Gerber 2000; Veyrieras et al. 2007). Software packages, such as Meta-QTL and BioMercator, facilitate meta-analysis of QTLs derived from independent studies by formulating and embedding specific sets of algorithms for exact evaluation and recalculation of the genetic position for the given QTLs

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(Wang et al. 2014). Among them, BioMercator is the most advanced and commonly used software for projection of initial QTLs on consensus map and identification of MQTLs. In wheat, meta-QTL analysis has already been conducted for different traits, including the ear emergence (S et al. 2009), fusarium head blight (Liu et al. 2009; Venske et al. 2019; Zheng et al. 2021), tan spot resistance (Liu et al. 2020), pre-harvest sprouting tolerance (Tyagi and Gupta 2012), abiotic (drought and heat) stress tolerance (Acuña-Galindo et al. 2015; Darzi-Ramandi et al. 2017; Kumar et al. 2020), yield (Yang et al.; Tyagi et al. 2015; Quraishi et al. 2017; Kumar Saini et al. 2021) and baking-quality and GPC (Quraishi et al. 2017). Quraishi et al. (2017) identified the six and eight MQTLs for GPC and baking quality traits, respectively, using the 155 initial QTLs collected from eight interval mapping studies. S**Error! Bookmark not defined.**ince, large number of QTLs have been reported after the meta-QTL analysis for quality traits by Quraishi et al. (2017) in wheat, the present study involving meta-QTL analysis was performed (based on interval mapping studies published during 2013-20) to supplement the list of MQTLs and candidate genes reported in the earlier MQTL study for quality traits (Quraishi et al. 2017). Further, the results of the meta-analysis were integrated with GWAS and transcriptomics studies to identify the promising genomic regions and important CGs, which affect quality traits in wheat. Further, by utilizing the synteny and collinearity of wheat with other cereals (Sorrells et al. 2003), ortho-MQTL analysis was conducted to see if the generated information can be transferred to other cereals such as rice. The findings of the present study may help in identification of diagnostic markers and aid marker-assisted breeding to improve quality traits in wheat.

Material and methods

Bibliographic search and QTL data cumulation

A comprehensive bibliographic search and data retrieval on the published papers from 2013 to 2020 enabled us to compile information from the available studies. From each independent study, the following information was obtained: (i) QTL name (wherever available), (ii) flanking or closely linked markers, (iii) peak position and CI, (iv) LOD score, (v) phenotypic variation explained (PVE) or R^2 value for each QTL and (vi) type and size of the mapping population (S1 and S2 Tables). Wherever, information about the peak position of the QTL was missing, it was calculated as the mid-point of the two markers flanking the QTL, and when LOD scores of individual QTLs were unavailable from a particular study, LOD score of 3 was considered for all the QTLs identified in the concerned study.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

All the quality traits were grouped into the following six major trait categories: (i) *Arabinoxylan* (*Ax*) recorded as total Ax, sucrose solvent retention capacity (SRC), sucrose softness equivalent SRC, water-extractable and water-unextractable Ax; (ii) *Dough rheology properties (DRP)* recorded as mixolab, mixograph, mixogram, alveograph, farinograph, water SRC, alkaline SRC, dough mixing, and pasting properties; (iii) *Nutritional traits (NT)* recorded as lipoxygenase activity, yellow pigment, Fe, Zn, β-glucan, protein, and starch content; (iv) *Polyphenol content* (*PPO*) recorded as polyphenol oxidase activity, flour colour, and loaf colour; (v) *Processing quality traits (PQT)* recorded as milling and baking qualities; and (vi) *Sedimentation volume* (*SV*) recorded as SV, sedimentation rate and zeleny SV (S3 Table).

Construction of consensus map

An R package LPmerge (Endelman and Plomion 2014) was used to construct a consensus map (named as Wheat_Reference_GeneticMap-2021) by merging the following linkage maps: "Wheat_Composite_2004" (http://wheat.pw.usda.gov), "Wheat_Consensus_SSR_2004" (Somers et al. 2004); "Durum wheat integrated map" (Marone et al. 2013); and four SNP array based maps developed using "Illumina 9K iSelect Beadchip Array" (Cavanagh et al. 2013), "Illumina iSelect 90K SNP Array" (Wang et al. 2014), "Wheat 55K SNP array" (Winfield et al. 2016) and "AxiomR, Wheat 660K SNP array" (Cui et al. 2017). Markers flanking the initial QTLs identified from different mapping studies were also integrated on to the consensus map (S4 Table).

QTL Projection, meta-QTL and ortho-MQTL analysis

For the QTLs with no CI available, CI (95%) was calculated by using the following population-specific equations (Darvasi and Soller 1997; Venske et al. 2019).

For RIL,

$$CI\ (95\%) = \frac{163}{Population\ size} \times R^2$$

For DH,

$$CI\ (95\%) = \frac{287}{Population\ size} \times R^2$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Further, careful examination and evaluation of individual QTLs for genetic map or QTL-related information led us to exclude 472 QTLs from the further analysis. Information on LOD score, PVE, genetic position (i.e., CI and peak position) from the remaining 1986 QTLs were compiled and used for projection on consensus map using the QTLProj tool of BioMercator V4.2. Meta-analysis was then performed for the individual chromosomes using Veyrieras two-step approach available in the software. In the first step, the best QTL model (used to ascertain the number of MQTL per chromosome) were computed by using the following five criteria: (i) Akaike Information Criterion (AIC), (ii) corrected Akaike Information Criterion (AICc), (iii) Akaike Information Criterion 3 (AIC3), (iv) Bayesian Information Criterion (BIC) and (v) Approximate Weight of Evidence Criterion (AWE). QTL model which had the least criterion value and zero delta value was considered as the best model for the analysis. In the second step, selected model was considered to determine the number of MQTLs on each chromosome (based on number of input QTLs on common genetic map), their consensus positions (based on variance of input QTL positions) and 95% CI (based on variance of input QTL intervals) (Sosnowski et al. 2012).

To investigate the ortho-MQTLs for quality characters, a previous meta-QTL study (Youlin et al. 2021) conducted in rice for starch pasting properties was used. To mine the ortho-MQTLs between wheat and rice, orthologous regions between the two species were investigated by using EnsemblPlants database. Candidate genes available in selected MQTL regions were used to search the corresponding rice genes. MQTL regions with at least 4 conserved genes present in wheat and rice genome were considered as the ortho-MQTL (Youlin et al. 2021).

Delineating physical location and candidate gene mining in MQTL regions

Nucleotide sequences of the markers flanking the MQTLs were retrieved from either of the following databases/websites: (i) the GrainGenes (https://wheat.pw.usda.gov/GG3), (ii) Diversity array technology (https://www.diversityarrays.com) and (iii) CerealsDB (https://www.cerealsdb.uk.net/cerealgenomics/CerealsDB/indexNEW.php). These sequences were used to obtain the physical positions of markers by BLASTN searches against Wheat Chinese Spring IWGSC RefSeq v1.0 genome assembly available in EnsemblPlants database (http://plants.ensembl.org/index.html). Physical positions of some of the SNP markers were directly obtained from the JBrowse-WHEAT URGI database (https://urgi.versailles.inra.fr/jbrowseiwgsc/).

Peak physical positions of MQTLs were calculated by using the following formula:

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$$Peakphysicalposition(bp) = Physicalstart(bp) + \frac{(Physicalstart - Physicalend)}{(Geneticstart - Geneticend)} \times \frac{CI(95\%)}{2}$$

Gene models available in 2 Mb genomic region i.e., 1 Mb region on either side of the MQTL peak position were retrieved using the 'BioMart' tool of EnsemblPlant database (https://plants.ensembl.org/biomart/). Functional characterization and gene ontology (GO) analysis for available gene models was also carried out using the same tool BioMart.

Expression analysis for gene models

An adaptable platform to create a gene expression interface **'**Wheat expression browser' powered by expVIP was used to analyse the expression of gene models available in the MQTL regions (Borrill et al. 2016). Following relevant gene expression datasets: (i) grain tissue-specific developmental time course (CS grain, grain dissection, early grain) (Gillies et al. 2012; Li et al. 2013; Pfeifer et al. 2014); (ii) grain tissue-specific expression at 12 days post anthesis (Pearce et al. 2015) and (iii) grain developmental time course with 4A dormancy QTL (Barrero et al. 2015) were used for this purpose. These studies provided the expression data on "morphological stages of developing wheat grain", "inner pericarp, outer pericarp and endosperm layers from developing grain of bread wheat at 12 days post-anthesis", "aleurone and starchy endosperm tissues of the wheat seed at aleurone layer development time of 6, 9 and 14-days post anthesis" and "candidate genes underlying the grain dormancy in wheat", respectively.

Only the gene models with ≥ 2 TPM (transcripts per million) expression were considered in the present study (GP et al. 2013). Further, protein sequences of the known genes associated with quality traits (collected from the literature) were retrieved and BLASTP searches were made against the wheat genome database available in EnsemblPlants to find their physical positions. Then, the physical positions of these genes were compared with the physical coordinates of the MQTLs to determine their co-localization, if any.

Validating MQTLs with MTAs identified in earlier GWA studies

Information from the 11 independent GWA studies on most stable and significant MTAs were collected to validate the MQTLs identified in the present study. These GWA studies involved one durum wheat population (with population size 194), four spring wheat populations (with population size ranging from 189 to 2038), three winter wheat populations (with population size ranging from 267 to 1325) and three mixed population of spring and winter wheat (with population size ranging from 163 to 4095). P**Error! Bookmark not defined.**hysical positions of each significant and stable SNPs associated with the trait were obtained from

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the respective studies and/or JBrowse-WHEAT URGI database (https://urgi.versailles.inra.fr/jbrowseiwgsc/). Finally, physical positions of these MTAs were compared with the physical coordinates of the MQTLs, to find any MQTL co-localizing with at least one MTA and such MQTLs were considered as a GWAS-validated/verified MQTL.

MQTL characterization using cloned genes from the rice

Functionally validated rice genes associated with quality traits (involving starch and protein metabolism, embryo and endosperm development, sugar transportation, grain development, etc.) were collected from the available literature and their protein sequences were retrieved from the Rice Annotation Project Database (rap-db) (https://rapdb.dna.affrc.go.jp/index.html). The protein sequences of rice genes were then BLAST against wheat genome to identify their corresponding wheat homologues available in the MQTL regions.

Results

Characterization of QTL studies involving quality traits

Characteristics of 50 independent mapping studies involving 63 bi-parental populations were rigorously reviewed to compile information on available QTLs. In total, 2458 QTLs associated with quality traits were collected and they were found to be distributed unequally on all the wheat chromosomes (Fig 1a). Maximum number of QTLs were present on the sub-genome B (995 QTLs), while the minimum number of QTLs were present on the sub-genome D (607 QTLs). There were 45 sets of RIL populations (population size ranging from 83 to 171) and 18 sets of DH populations (population size ranging from 94 to 192) (Fig 1b, c; S1 and S2 Tables). The number of studies involving different DH and RIL populations, molecular markers and mapping methods are presented in Fig 1c, d and respectively. Most of the mapping studies used SSR as the markers and composite interval mapping (CIM) as method for QTL mapping, respectively (Fig 1d, e). Number of QTLs associated with the different quality traits also varied (from 36 QTLs for Ax content to 962 QTLs for DRP). PVE values and LOD score for individual QTLs ranged from 0.01 to 95.24 per cent with a mean of 9.31 per cent and from 0.99 to 55.32 with an average of 5.74, respectively (Fig 1f, g; S1 Table). Peak positions of the initial QTLs also varied (ranging from 0 to 678.2 cM with a mean of 108.18 cM) (Fig 1h).

Construction of consensus genetic map

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The consensus map, "Wheat_Reference_GeneticMap-2021" showed significant variation for the genetic lengths of the individual linkage groups/chromosomes (ranging from 361.8 cM for 1A to 808.6 cM for 2A, with a mean of 649.4 cM) and for the number of markers positioned on a single chromosome (ranging from 2,225 on 4D to 21,253 on 3B with an average of 11908.4 markers per chromosome) (Fig 2). The total length of the consensus map measured 13,637.4 cM which included a total of 2,50,077 markers of different types such as AFLP, RFLP, SSR, SNP, etc. Since, different genetic maps with varying number and type of markers were used to construct the consensus map, the distribution of markers on the chromosomes was uneven and density of the markers was comparatively higher at the fore-end of the chromosome (Fig 2; S4 Table). Marker density on individual chromosomes ranged from 5.7 markers/cM on 4D to 46.2 markers/cM on 1A with a mean of 18.3 markers/cM on whole genome.

MQTLs identified for quality traits

Of the total 1986 QTLs with complete information available for analysis, only 1128 (56.79% QTLs were projected onto the consensus map. Among the 1128 QTLs, 1121 QTLs were grouped into 110 MQTLs, 7 QTLs remained as singletons (Table 1; Fig 3 and 4). Eight hundred and fifty-eight (850) QTLs were not projected onto the consensus map owing to either of the following reasons: (i) there were no common markers between the consensus and original maps, and (ii) the QTLs had low R^2 values, resulting in a large CI. The number of MQTLs on individual chromosomes ranged from 2 (on chromosomes 5B, 6B and 7A) to 9 (on chromosome 1A) (Fig 3a; S5 Table). The number of clustered QTLs per MQTL ranged from 2 (in several MQTLs) to 85 (in MQTL1B.1) with 11 MQTLs involving at least 20 initial QTLs (Fig 3d; S5 Table). Of these 11 MQTLs, four were located on chromosome 1B and remaining each of the seven MQTLs were located on different chromosomes. Since, these 11 MQTLs integrated the QTLs from different populations, they were considered to be more stable and reliable MQTLs for wheat quality breeding. The chromosome-wise average CI (95%) for the initial QTLs and MQTLs (presented in the Fig 3c) ranged from 14.87 cM (1D) to 95.55 cM (3D) for QTLs and from 0.55 cM (1D) to 34.30 cM (2D) for MQTLs. The average CI of MQTLs (5.56 cM) was 18.84-fold less than that of initial QTLs (40.35 cM), and there were significant differences among different chromosomes (Fig 3c). The average CI of MQTLs reduced maximum for chromosome 5B (72.55 times) and minimum for chromosome 2D (1.99 times) compared to the CI of initials QTLs located on corresponding chromosomes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Individual QTLs for processing quality traits were found to be associated with as many as 98 MQTLs and for that of the other traits are given in Fig 3b and S5 Table. All MQTLs were associated with at least one quality trait (Fig 3e). As many as 38 MQTLs were exclusively associated with only a single quality trait, whereas, four MQTLs (viz., 2D.2, 5A.2, 5D.4, and 7B.3) were found to be associated with 5 quality related traits (S5 Table). None of the MQTL included the QTLs associated with all of the six studied traits. MQTLs associated with multiple traits are considered to be more important for the breeding programmes aiming towards the simultaneous improvement of multiple quality traits. In this regard, MQTLs (viz., 2D.2, 5A.2, 5D.4, and 7B.3) were considered as more promising and breeders can use them to enhance multiple quality traits simultaneously.

Table 1. Summary of wheat MQTLs associated with different quality traits

S. No.	MQTLs	Genetic position (cM)		Flanking genetic markers		Physical position (bp)		Number of QTLs projected
		Start	End	Start	End	Start	End	
1	MQTL 1A.1	14.7	15.1	IWB72800	D-1377838	9.5	36.6	24
2	MQTL 1A.2	22.3	22.7	3980487_1al_11151	IWB35066	39.6	41.7	20
3	MQTL 1A.3	36.1	36.4	BobWhite_c5793_372	3955448_1al_1513	3.5	22.7	10
4	MQTL 1A.4	42.6	43.1	3881592_1al_184	wPt2311	22.7	23.7	14
5	MQTL 1A.5	68.6	69.1	AX-109031595	wPt8172	300.4	489.5	6
6	MQTL 1A.6	85.5	87.8	Tdurum_contig47183_205	Xwpt3698	30.5	42.3	2
7	MQTL 1A.7	92.9	94.8	Xbarc174	wPt1906	41.9	42.3	15
8	MQTL 1A.8	118.4	123.1	snp2584	wPt7030	550.3	551.3	10
9	MQTL 1A.9	156.6	156.9	D_3022884	Xcfe0242b	516.4	530.3	12
10	MQTL 1B.1	41.2	41.4	AX-95133874	Xwmc44	635.0	635.1	85
11	MQTL 1B.2	66.3	66.8	IWB65744	AX-94462160	442.2	646.1	21
12	MQTL 1B.3	77.0	77.6	IWB13393	IWA7298	542.4	543.0	27
13	MQTL 1B.4	110.0	111.7	Ku_c241_460	IWB6906	328.0	548.6	12
14	MQTL 1B.5	195.5	197.3	wPt1973	wPt_2315	7.2	8.2	21
15	MQTL 1B.6	428.3	439.3	Kukri_rep_c111991_498	XPaggMcgg6	408.6	553.0	9
16	MQTL 1D.1	43.6	43.9	Xbarc229.2	cfd92	28.7	31.0	21
17	MQTL 1D.2	47.2	47.6	2257978_1dl_15913	2281626_1dl_570	420.7	449.0	11
18	MQTL 1D.3	60.1	60.7	D_1073588	D_1234123	433.5	435.8	14

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

19	MQTL 1D.4	76.0	76.0	TA004476-0719	Glu_D1	411.6	411.7	2
20	MQTL 1D.5	77.0	77.0	TA004476-0719	Glu_D1	411.6	411.7	16
21	MQTL 1D.6	109.4	110.6	AX-95218664	AX-94769043	494.6	573.1	10
22	MQTL 1D.7	156.9	158.3	IWB9092	IWB56444	459.0	460.1	12
23	MQTL 2A.1	8.1	11.9	Xgwm328_2A	GBS_210	8.8	16.8	11
24	MQTL 2A.2	28.4	28.6	Kukri_c36139_292	BS00063632_51	14.8	19.4	15
25	MQTL 2A.3	81.2	82.3	RAC875_c17787_274	TA004602-1630	7.6	14.8	14
26	MQTL 2A.4	99.5	100.3	wPt-8049	IWB23030	57.8	61.3	9
27	MQTL 2A.5	248.2	260.8	BS00082084_51	RAC875_c54668_102	732.6	770.9	2
28	MQTL 2A.6	292.9	292.9	RAC875_c54668_102	wsnp_Ku_c1292_2572110	27.9	45.1	10
29	MQTL 2B.1	0.7	0.9	M35745	D_1162944	32.5	34.4	12
30	MQTL 2B.2	25.3	29.2	Xwmc435	IWA897	58.8	217.3	6
31	MQTL 2B.3	49.9	50.3	Ex_c55735_1012	Xwpt8004	556.2	681.5	11
32	MQTL 2B.4	73.2	75.0	Xgwm55b	Excalibur_c33221_681	532.1	736.4	6
33	MQTL 2B.5	93.9	95.2	CAP7_6910_523	8086989_2bl_2189	31.0	32.5	9
34	MQTL 2B.6	136.7	139.7	8074934_2bl_7628	Kukri_c9507_495	236.3	784.9	12
35	MQTL 2B.7	228.7	230.0	Xgwm1273	U296	41.2	790.9	8
36	MQTL 2D.1	28.8	32.1	2DS_5366150_1491	AX-95093513	48.3	579.9	6
37	MQTL 2D.2	63.7	65.8	AX-110230565	Xctg05205	429.2	584.5	14
38	MQTL 2D.3	112.5	113.6	Kukri_c44769_750	BS00011109_51	409.6	608.8	5
39	MQTL 2D.4	140.0	140.1	IWB64805	IWB34403	428.9	542.3	2
40	MQTL 2D.5	163.8	169.7	IWB29964	IWB32041	574.4	593.5	2
41	MQTL 2D.6	337.3	530.5	IWB44461	GENE_4086_115	604.1	629.2	3
42	MQTL 3A.1	0.0	3.5	Excalibur_c12875_1573	4397491_3al_1806	513.6	571.4	3
43	MQTL 3A.2	25.5	25.6	AX-109316906	P41/M41-4	614.0	733.5	19
44	MQTL 3A.3	65.7	70.5	IWB54878.1	KUKRI_REP_C69970_717	650.4	724.6	4
45	MQTL 3A.4	110.5	111.0	XPaggMcgg1	Excalibur_c5416_846	569.4	737.3	9
46	MQTL 3B.1	34.0	36.7	XPaggMctg16	wPt7502	616.9	760.6	10
47	MQTL 3B.2	72.4	73.7	IWB6207	D-4329487	31.7	32.3	6
48	MQTL	111.5	113.6	AX-109869742	AX-109842601	773.8	787.0	11

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

	3B.3							
49	MQTL 3B.4	173.9	174.4	ACA.CTGA12	Xctgacg317	553.1	798.3	21
50	MQTL 3D.1	36.2	39.2	AX-95218150	AX-110037813	597.5	597.9	6
51	MQTL 3D.2	67.2	69.2	IWB40541	IWB19457	600.1	600.3	8
52	MQTL 3D.3	122.3	144.9	wPt669255	RFL_Contig5276_1756	21.8	73.9	4
53	MQTL 3D.4	189.1	238.9	RFL_Contig5276_1756	BS00066932_51	73.9	515.7	3
54	MQTL 4A.1	8.9	10.8	Kukri_c17417_797	Xgwm113	24.1	601.3	11
55	MQTL 4A.2	34.7	38.8	wPt7924	wPt0538	671.9	713.0	6
56	MQTL 4A.3	66.7	71.5	wPt0105	snp1066	623.3	699.5	3
57	MQTL 4A.4	113.2	115.3	IAAV5818	SSRabg390.DI	603.5	730.5	30
58	MQTL 4A.5	164.7	165.0	RAC875_c78248_115	BS00067074_51	602.3	609.4	19
59	MQTL 4B.1	12.3	12.5	Xfbb255a	D_3939025	13.9	15.0	15
60	MQTL 4B.2	20.9	21.4	Xmag983	S-3025233	601.9	609.5	6
61	MQTL 4B.3	44.9	45.2	AX-110396575	BS00076259_51	457.9	616.6	11
62	MQTL 4B.4	52.3	52.8	wPt_3908	Xgwm538	32.8	53.6	16
63	MQTL 4B.5	89.3	92.3	IWB73486	IWB15523	660.7	663.1	11
64	MQTL 4B.6	181.8	185.4	Xgwm637b	wsnp_CAP12_c1101_569783	237.0	609.5	10
65	MQTL 4D.1	0.7	1.6	Xwmc617	AX-110527441	14.6	15.0	10
66	MQTL 4D.2	53.6	53.9	IWB16077	Xsrap16	5.6	49.9	5
67	MQTL 4D.3	78.7	80.4	Xgdm129	wPt2379	196.6	484.4	2
68	MQTL 4D.4	83.3	83.9	AX-94881415	AX-94383842	2.7	492.0	8
69	MQTL 4D.5	86.0	86.7	Xsrap11b	IWB62209	490.9	499.1	2
70	MQTL 4D.6	87.9	88.5	AX-110418361	Xgdm125	3.4	12.8	4
71	MQTL 4D.7	93.9	94.5	Xbarc1183	AX-94403999	4.4	5.7	8
72	MQTL 4D.8	205.4	206.0	Xscss30.2.2	Kukri_c40437_66	497.6	508.3	3
73	MQTL 5A.1	8.8	9.5	Xgwm0304b	AX-95186387	6.7	8.1	10
74	MQTL 5A.2	39.0	42.0	D-1206650	XACG.GAC1.2	526.6	528.6	10
75	MQTL 5A.3	103.5	103.8	Xgwm410.3	IWB61598	699.5	700.4	24
76	MQTL 5B.1	8.6	8.8	Xbarc133	GBS_257	577.9	713.0	3
77	MQTL 5B.2	42.7	44.0	wPt5168	IWB36196	14.5	570.6	26

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

78	MQTL 5D.1	9.7	10.9	D-4329273	S-1046500	458.1	463.1	16
79	MQTL 5D.2	28.1	28.3	S-982589	S-1052739	481.4	483.1	5
80	MQTL 5D.3	31.6	31.7	D-4329207	D-1116418	485.5	486.8	4
81	MQTL 5D.4	72.2	72.3	IWB58798	Xswes342a	371.8	531.5	7
82	MQTL 5D.5	168.3	173.0	wPt1197	5DL_4588938_1981	204.9	343.3	12
83	MQTL 5D.6	418.1	432.6	Jagger_c550_91	JD_c4438_839	406.6	546.7	6
84	MQTL 6A.1	0.0	3.3	AX-94788664	D_3952327	16.6	17.6	3
85	MQTL 6A.2	9.4	18.2	IWB44292	D-2278273	14.4	15.8	2
86	MQTL 6A.3	35.3	35.9	Tdurum_contig46828_1430	S_1116215	571.9	614.2	9
87	MQTL 6A.4	40.4	46.3	S_1116215	IWB23460.2	574.5	614.2	3
88	MQTL 6A.5	57.6	70.3	IWA504	wPt4791	253.2	602.5	4
89	MQTL 6A.6	202.0	235.7	104.172	Xcdo1090a	1.9	60.4	2
90	MQTL 6A.7	270.9	277.3	XPaggMctg18	Xbarc146b	1.9	614.7	5
91	MQTL 6B.1	27.0	28.7	S_1054930	GTG.CTT1	3.5	698.6	6
92	MQTL 6B.2	62.7	62.9	Xdupw16	wPt745074	19.5	648.7	24
93	MQTL 6D.1	33.7	38.2	AX-86177993	cfd13c	24.0	89.3	15
94	MQTL 6D.2	75.8	78.7	BobWhite_c13435_700	wPt666008	213.1	471.0	8
95	MQTL 6D.3	116.5	121.7	Xfbb9b	AX-94381525	-	-	5
96	MQTL 6D.4	165.0	165.1	IWB4527	IWB58414	-	-	4
97	MQTL 7A.1	57.1	80.6	M103823	GBS_539	8.8	12.9	2
98	MQTL 7A.2	148.1	154.4	tarc0748	gwm332	42.2	47.4	3
99	MQTL 7B.1	1.1	1.3	Excalibur_c22830_2010	Excalibur_c16580_388	3.1	53.6	6
100	MQTL 7B.2	14.7	15.0	Kukri_c16416_647	Xgwm1250	1.1	6.4	6
101	MQTL 7B.3	46.8	50.4	wPt_0789	snp2273	593.2	718.4	16
102	MQTL 7B.4	78.1	81.8	wPt_2592	D_3938216	587.9	710.7	6
103	MQTL 7B.5	92.3	95.8	wPt_732048	wPt7368	21.7	643.4	3
104	MQTL 7B.6	231.3	234.5	Xgwm0037	RFL_Contig4686_700	0.8	744.3	5
105	MQTL 7D.1	12.7	20.2	Xgwm350	Xgwm974	22.5	25.4	7
106	MQTL 7D.2	40.1	45.8	AX-110940110	wPt664368	58.9	91.7	4
107	MQTL	55.3	64.2	IWB6964	S-1162003	58.9	526.4	2

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

	7D.3							
108	MQTL 7D.4	66.8	75.1	Xgwm0437	D-1135134	2.9	86.1	17
109	MQTL 7D.5	114.9	121.2	AX-110579501	Excalibur_rep_c8301 9_155	15.8	135.3	7
110	MQTL 7D.6	290.6	305.4	CAP8_rep_c9420_ 186	Excalibur_rep_c8589 1_96	211.4	237.5	10

MQTLs validated with GWAS results

The physical coordinates of the MQTLs identified in the present study were also compared with MTAs reported in earlier GWA studies (S6 Table). Among the 108 MQTLs, as many as 43 MQTLs (39.81 %) co-localized with at least one MTA available from the GWAS (Fig 4; S6 Table). There were some MQTLs co-localized with MTAs available from more than one GWA study. For instance, MQTL6A 7 co-located with MTAs reported in seven different GWA studies. Each of the five MQTLs (viz., 1A.1, 1B.2, 3B.4, 5B.2 and 6B.2) involving at least 20 initial QTLs co-localized with multiple MTAs reported from different GWA studies.

Surprisingly, five genes, *TraesCS1A02G040600*, *TraesCS2D02G531100*, *TraesCS3B02G449200*, *TraesCS3D02G095700* and *TraesCS3D02G096000* described by Yang et al. (2021) GWA study were overlapped with four MQTLs identified in the present study. Among them, first three genes were present in the MQTL1A.1, MQTL2D.6, and MQTL3B.1, respectively. While, last two genes located on the 3D chromosomes were present in the MQTL3D.3. Functional annotation and GO study revealed their participation in the various biological processes associated with grain quality in wheat. Genes, namely *TraesCS3D02G095700* and *TraesCS3D02G096000* encode for the wheat allergens and trypsin and alpha-amylase inhibitor present in the seeds.

Wheat homologues of known rice genes in MQTL regions

An extensive search made on known rice genes associated with quality traits led to the collection of information on 34 genes which were further utilized for the identification of their corresponding homologues in wheat MQTL regions. These genes encode proteins/products mainly belonging to the following families: starch synthase, glycosyl transferase, aldehyde dehydrogenase, SWEET sugar transporter, alpha amylase, glycoside hydrolase, glycogen debranching, protein kinase, peptidase, legumain, and seed storage proteins. Wheat homologues of 30 of these rice genes were identified in different MQTL regions (S7 Table); wheat homologues for three rice genes, namely, *RP6*, *RM1*, and *OsAGPL4/OsAPL4* could not be identified due to low sequence identity (<60%) and query coverage (<30%). Further, it has been noticed that, wheat homologues for nine rice

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

genes are identified on the different chromosomes of wheat. For example, wheat homologues for rice gene, *Wx1* are present on the 4th and 7th chromosomes. Sixty-four wheat homologues of these 30 rice genes were detected in the 44 wheat MQTL regions (S7 Table); some MQTL regions harboured more than one wheat homologues of rice genes. For example, wheat homologues of rice genes *wx1*, *OsACS6* and *GBSSII* genes were present within the MQTL4A.2 (S7 Table).

Ortho-MQTL analysis

Extensive investigation of orthologous gene models available from all the 108 wheat MQTLs (with known genomic coordinates) led to the identification of 23 wheat MQTLs (ortho-MQTLs) syntenic to 16 rice MQTLs (Fig 4; S8 Table). In some cases, more than one wheat MQTLs were found to be syntenic to a single rice MQTL. For instance, three wheat MQTLs (viz., 2B.3, 2D.2, and 2D.2) were found to be syntenic to a rice MQTL (i.e., MQTL4.5) (Figs 6 and 8). Conversely, single wheat MQTL was found to be syntenic to multiple rice MQTLs. For instance, wheat MQTL4A.4 was syntenic to two rice MQTLs located on chromosome 8 (viz., MQTL8.6 and MQTL8.8) and MQTL4B.6 was syntenic to two rice MQTLs located on chromosome 3 (MQTL3.3 and MQTL3.4).

Number of orthologous gene models in individual MQTL varied from 3 in MQTL1A.3 to 19 genes in MQTL2D.2. Five ortho-MQTLs (viz., 2A.1, 2D.2, 4D.8, 5D.1 and 7D.1) each involving more than 10 gene models syntenic to rice genome are presented in Fig 5. Gene models present in the two wheat MQTLs (viz., 2D.2 and 4D.8) located on chromosome 2D were found to be orthologous to rice MQTL4.5 and MQTL3.4, respectively, in the reverse order (Fig 5b, c). From this, it can be inferred that genomic regions in MQTL2D.2 and MQTL4D.8 were inverted once in the evolutionary time and conserved in the rice genome. MQTL5D.1 with 13 gene models was corresponding to rice MQTL3.7. Wheat genes present in MQTL5D.1 was found to be collinear to one rice MQTL3.7 (Fig 5d). We also reported a disruption in collinearity between the genes available from the wheat (MQTL7D.1) and rice (MQTL6.1) MQTLs (Fig 5e). Wheat MQTL2A.1 with 14 genes was found to be syntenic to the rice MQTL4.2 (Fig 5a).

Gene models available in MQTL regions and their expression analysis

A total of 108 MQTLs were anchored to the physical map of wheat reference genome. Physical positions of the two MQTLs (viz., MQTL6D.3 and MQTL6D.4) could not be deduced, as these MQTLs were flanked by markers with no sequence information available. Gene mining within 108 MQTL regions permitted the identification of as many as 2533 gene models. Maximum number of gene models (276) were

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

identified for MQTLs located on the chromosome 2A, while the minimum number of the gene models (6 gene models) were available from the MQTLs located on chromosome 6B (Table 2; S9 Table).

Table 2. Number of initial QTLs, projected QTLs, gene models identified in the different chromosomes

Chromosome	Initial number of QTLs	Projected QTLs	Unique gene models	Gene models with >2 TPM expression
1A	206	113	262	52
1B	289	175	119	47
1D	158	86	116	25
2A	121	61	276	47
2B	177	64	88	17
2D	100	32	140	38
3A	90	35	93	14
3B	105	48	79	21
3D	34	22	104	26
4A	124	69	112	12
4B	142	69	117	31
4D	49	42	242	69
5A	93	44	94	18
5B	100	29	28	8
5D	94	50	143	38
6A	82	28	204	43
6B	94	30	6	2
6D	62	32	54	18
7A	140	2	47	7
7B	88	43	118	25
7D	110	47	91	22
Total	**2458**	**1121**	**2533**	**580**

Expression analysis of available gene models detected 556 gene models with >2 transcripts per million (TPM) expression, which included 94 gene models with >5 TPM and 6 gene models with >10 TPM expression in grains and its component tissues such as seed coat, starchy endosperm, whole endosperm, aleurone layer and grain transfer cells in (Fig 8; S10 Table). Of these 556 gene models, 117 gene models showed constitutive expression in all the tissues studied. While, the remaining 439 gene models showed the differential expression (tissue specific). Further, 44 differentially expressed gene models with more than 5 TPM expression were involved in the improvement of grain quality (Fig 7). Based on their expression pattern, gene models could be divided into two classes, i.e., genes in class-I showed expression in the seed, while genes in the class-II exhibited expression in the different sub-tissues. Further, functional characterization of these genes revealed their involvement in transcriptional and translational regulation of various genes, signalling mechanism, metabolism, cellular development and transfer, etc.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In recent study, transcriptome analysis in the hexaploid wheat and its diploid progenitors identified the differentially expressed genes (DEGs) associated with carbohydrate metabolism (Kaushik et al. 2020). Seven DEGs (Table 3) identified from Kaushik et al. (2020) study were common to the genes characterized in the present study. Of these seven genes, six genes were downregulated (Table 3) and one gene, *TraesCS7B02G194000.1* was upregulated. Among the downregulated genes, two genes were encoding for NB-ARC domain containing proteins and one gene for the F-box domain (Table 3). While the functions of the three downregulated genes were not characterized.

Table 3. Differentially expressed genes detected from the transcriptomic studies (Kaushik et al. 2020) which co-localized with MQTLs identified in the current study

Sl. No.	MQTL	Gene stable ID	Gene position (bp)	GO term	Interpro Description	Expression
1	MQTL1A.1	TraesCS1A02 G041100	22612227-22614080	ADP binding	NB-ARC	Down-regulated
2	MQTL2A.2	TraesCS2A02 G042500	17025906-17033628	ADP binding	NB-ARC	
3	MQTL2B.3	TraesCS2B02 G430500	619391291-619391622	-	-	
4	MQTL4A.5	TraesCS4A02 G316100	606335446-606340512	protein binding	F-box domain	
5	MQTL6A.3	TraesCS6A02 G365500	593732776-593733555	-	-	
6	MQTL7A.2	TraesCS7A02 G078900	44355885-44362888	-	-	
7	MQTL7B.5	TraesCS7B02 G194000	333406094-333422302	catalytic activity	Alpha/beta hydrolase fold-1	Upregulated

Cloned and characterized wheat genes associated with grain protein content (GPC) (*GPC-B1/NAM-B1*), HMW glutenin (*Glu-B1-1b*, *Glu-1D-1d*, *1Dx2t* and *Glu-1By9*) and starch synthase enzymes (*TaSSI*, *TaSSIIa*, *TaGBSSIa* and *TaSSIVb*) were identified in the 12 MQTL regions (S11 Table). Two MQTLs (6B.1 and 6B.2) present on chromosome 6B are associated with the GPC. MQTLs lying on the chromosome 1B (1B.2, 1B.3 and 1B.4) and 1D (1D.4 and 1D5) are act as a precursor for the HMW. Genes for the starch synthase enzyme present on the 1D (MQTL1D.2) and 7D (MQTL7D.2, MQTL7D.3, MQTL7D.4 & MQTL7D.5) chromosomes are responsible for the amylopectin biosynthesis via α-1,4-glycosidic linkages. Present study has also identified the 165 genes associated with catechol oxidase activity and metal ion binding (related PPO), Zn-transporter and zinc-binding site (Zn and Fe content), Small hydrophilic plant seed protein, Amino acid transporter and seed storage helical domain (seed storage protein), sweet-sugar transporters, UDP-glucuronosyl/UDP-glucosyltransferase and Sugar/inositol transporter (starch content) etc. Among them, 44 genes had the expression value of more than two TPM and were regarded as the putative CGs (Table 4). All the

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

six CGs, *TraesCS3D02G095700*, *TraesCS3D02G096000*, *TraesCS4B02G017500*, *TraesCS4D02G016000*, *TraesCS4D02G016100* and *TraesCS6D02G000200* having high expression value of >10 TPM were showed to be involved in improving the seed quality through nutrient reservoir activity.

Table 4. High-confidence candidate genes with >5 TPM expression identified in the present study

Sl. No.	MQTL No.	Gene stable ID	Gene start (bp)	Gene end (bp)	GO term name	Interpro Description	Expression (in TPM)
1	MQTL 1A.2	TraesCS1A02 G060100	41357284	41362423	peptidase activity	Peptidase C13, legumain	>2
2	MQTL 1A.7	TraesCS1A02 G062500	42670658	42671830		seed storage helical domain superfamily	>5
3	MQTL 1A.7	TraesCS1A02 G062700	42695416	42698084		seed storage helical domain superfamily	>5
4	MQTL 1A.7	TraesCS1A02 G062900	42721261	42721834		seed storage helical domain superfamily	>5
5	MQTL 1A.5	TraesCS1A02 G224000	394325250	394328538		Peptidase M24	>5
6	MQTL 1A.5	TraesCS1A02 G224200	394482047	394482797	cytosol	Small hydrophilic plant seed protein	>5
7	MQTL 1A.5	TraesCS1A02 G224600	394587930	394588284	cytosol	Small hydrophilic plant seed protein	>5
8	MQTL 1B.6	TraesCS1B02 G274300	481540499	481542564	nucleus	NAC domain	>2
9	MQTL 1B.1	TraesCS1B02 G407800	634934360	634934858		seed storage helical domain	>2
10	MQTL 1D.4	TraesCS1D02 G317300	412223432	412224112	IgE binding	seed storage helical domain	>5
11	MQTL 1D.2	TraesCS1D02 G346100	433899291	433905345	metal ion binding	Peptidase M16, zinc-binding site	>2
12	MQTL 2A.1	TraesCS2A02 G028400	12990103	12992210	transferase activity, transferring hexosyl groups	UDP-glucuronosyl/UDP-glucosyltransferase	>2
13	MQTL 2A.1	TraesCS2A02 G028500	13002991	13004793	transferase activity, transferring hexosyl groups	UDP-glucuronosyl/UDP-glucosyltransferase	>5
14	MQTL 2A.6	TraesCS2A02 G081100	36700022	36701587	transferase activity, transferring hexosyl groups	UDP-glucuronosyl/UDP-glucosyltransferase	>2
15	MQTL 2B.3	TraesCS2B02 G430900	619754863	619757940	serine-type carboxypeptidase activity	Peptidase S10, serine carboxypeptidase	>2
16	MQTL 2D.4	TraesCS2D02 G382400	486390272	486392935	peptidase activity	Peptidase C13, legumain	>5
17	MQTL 3A.4	TraesCS3A02 G408200	653303151	653306203	integral component of membrane	SWEET sugar transporter	>2

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

18	MQTL 3A.3	TraesCS3A02 G446200	686762258	686767701	serine-type carboxypeptidase activity	Peptidase S10, serine carboxypeptidase	>2
19	MQTL 3B.2	TraesCS3B02 G059300	31058376	31059986	hydrolase activity	Peptidase C26	>2
20	MQTL 3B.2	TraesCS3B02 G059500	31066580	31068297	hydrolase activity	Peptidase C26	>2
21	MQTL 3D.3	TraesCS3D02 G095700	48543331	48543795	serine-type endopeptidase inhibitor activity	Cereal seed allergen/trypsin and alpha-amylase inhibitor	>10
22	MQTL 3D.3	TraesCS3D02 G096000	48732052	48732516	serine-type endopeptidase inhibitor activity	Cereal seed allergen/trypsin and alpha-amylase inhibitor	>10
23	MQTL 3D.3	TraesCS3D02 G096500	48836328	48842240	integral component of membrane	Amino acid/polyamine transporter I	>2
24	MQTL 3D.1	TraesCS3D02 G514700	598160861	598163274	transferase activity, transferring hexosyl groups	UDP-glucuronosyl/UDP-glucosyltransferase	>2
25	MQTL 4B.1	TraesCS4B02 G017500	13052500	13054709	nutrient reservoir activity	Cupin 1	>10
26	MQTL 4B.1	TraesCS4B02 G017600	13077784	13080425	nutrient reservoir activity	Cupin 1	>5
27	MQTL 4B.1	TraesCS4B02 G020000	14117806	14125905	zinc ion binding	Peptidase M1, alanine aminopeptidase	>2
28	MQTL 4B.6	TraesCS4B02 G196600	423357718	423362900		Alpha/beta hydrolase fold-1	>2
29	MQTL 4D.6	TraesCS4D02 G016000	7182216	7204066	nutrient reservoir activity	Cupin 1	>10
30	MQTL 4D.6	TraesCS4D02 G016100	7203800	7206343		Cupin 1	>10
31	MQTL 4D.6	TraesCS4D02 G017700	7826751	7834962	zinc ion binding	Peptidase M1, alanine aminopeptidase	>2
32	MQTL 4D.2	TraesCS4D02 G051100	26933336	26934958		Alpha/beta hydrolase fold-1	>2
33	MQTL 4D.5	TraesCS4D02 G336100	494050540	494057878	integral component of membrane	Sugar/inositol transporter	>2
34	MQTL 4D.5	TraesCS4D02 G337300	494829139	494833871	serine-type carboxypeptidase activity	Peptidase S10, serine carboxypeptidase	>2
35	MQTL 4D.8	TraesCS4D02 G350300	502576795	502579734	integral component of membrane	Peptidase S54, rhomboid	>2
36	MQTL 4D.8	TraesCS4D02 G350700	502747671	502751833		Alpha/beta hydrolase fold-1	>2
37	MQTL 5D.4	TraesCS5D02 G381000	451173705	451179099		Peptidase S1, PA clan	>2
38	MQTL	TraesCS5D02	45154202	45154741		Peptidase S1, PA clan	>2

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

	5D.4	G381200	4	7			
39	MQTL 6A.4	TraesCS6A02 G370300	595182823	595186561	hydrolase activity, acting on carbon-nitrogen bonds	Acetamidase/Formamidase	>5
40	MQTL 6D.4	TraesCS6D02 G000200	52444	52899	serine-type endopeptidase inhibitor activity	Cereal seed allergen/trypsin and alpha-amylase inhibitor	>10
41	MQTL 6D.1	TraesCS6D02 G091300	57107880	57116429	serine-type endopeptidase activity	Peptidase S9, prolyl oligopeptidase	>2
42	MQTL 7B.5	TraesCS7B02 G193800	332923582	332926431	integral component of membrane	Diacylglycerol acyltransferase	>2
43	MQTL 7D.2	TraesCS7D02 G121200	75330927	75337847	serine-type peptidase activity	Peptidase S9, prolyl oligopeptidase, catalytic domain	>2
44	MQTL 7D.2	TraesCS7D02 G121900	75930917	75945541	serine-type endopeptidase activity	Peptidase S8/S53 domain	>2

Discussion

Discovery of molecular markers and advancements in QTL mapping strategies have promoted the plant researchers across the globe to underpin the complex genetic architecture of quality traits in wheat (Quraishi et al. 2017). Results of these intensive studies has accrued a large number of QTLs controlling the quality traits in wheat. Majority of these QTLs are minor QTLs (with low PVE value) and have larger CI, making them unsuitable for the marker assisted plant breeding (Kumar Saini et al. 2021). Further, it has been observed that, QTLs identified from one population may not be effective for a breeding programme involving the other population (Yang et al.). Downsides of these QTLs promoted us to reanalyze all of these loci together, with the goal of fine-tuning this vast amount of data, so that breeders and researchers can make greater use of these QTLs.

Meta-QTL analysis has the great capability to compile the information from multiple QTL mapping studies involving diverse environments with different genetic backgrounds to identify stable and reliable MQTLs with reduced CI (Welcker et al.). MQTL analysis have already conducted for the major crops including the rice, wheat, maize, etc. (Yang et al.; Li et al. 2013; Quraishi et al. 2017). Meta-QTL analysis for agronomic and quality traits in wheat has identified the six MQTLs for GPC and eight MQTLs for baking quality from the initial pool of 155 QTLs involving eight genetic studies published before 2013 (Quraishi et al. 2017). Further, seven MQTLs associated with Zn and Fe content were identified from initial set of 50 QTLs involving seven

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

independent segregating populations. These studies have employed a small number of initial QTLs for the meta-analysis. However, the results of MQTL analysis have been found to be significantly and positively correlated with the quality of the results of meta-analysis (Yang et al.; Quraishi et al. 2017). Furthermore, as molecular genetics and QTL mapping methodologies are advancing, novel QTLs are adding to the database on a regular basis, therefore we must keep up with this pace in order to integrate new QTLs into more stable and reliable MQTLs.

Thus, to supplement the previous studies (Quraishi et al. 2017), in the present study, we conducted the meta-analysis for the QTLs reported during 2013-20 by projecting very large number of initial QTLs (1986 QTLs) on the high-density consensus map and identified 110 MQTLs each involving at least two initial QTLs. In addition to Zn, Fe, GPC and baking quality (studied by Quraishi et al. 2017), the present study also considered the other important quality parameters such as starch, PPO, Ax content, bread making properties, etc.

From the breeding perspectives, it is noteworthy to identify the most stable and reliable MQTLs involving a greater number of initial QTLs identified across the different populations and environments. Eleven such MQTLs involving more than 20 initial QTLs were detected in the present study. One MQTL located on chromosome 1B (MQTL1B.1) included as many as 85 initial QTLs, which is much higher than earlier study of (Quraishi et al. 2017). Further, to reduce the confidence interval, meta-QTL analysis relies on a large number of initial QTLs and heterogeneous populations (Berman and Parker 2002). Present study aggregated a vast QTL information from multiple genetic backgrounds and efficiently reduced the CI of the QTLs, thereby improved the accuracy of candidate gene prediction from the important QTL regions. The average CI of MQTL was 18.84-fold less than that of initial QTLs, and there were significant differences among different chromosomes. Some of the MQTLs with reduced CI identified in the present study may lay the foundation for molecular cloning and functional characterization of genes associated with quality traits in wheat. Further, this may offer the possibilities for the marker assisted gene transfer across the populations and contribute to the wheat quality improvement.

Validating MQTLs with GWAS-MTAs

Genome wide association studies (GWAS) is a powerful tool which allows the investigation of highly complex characters by exploiting the recent and historical recombinant events present in the natural populations and permits the high-resolution mapping (Bush and Moore 2012). High-throughput, low-cost sequencing technologies have helped to identify MTAs for many quality traits in natural populations of wheat (Suliman et

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

al.; Godoy et al. 2018). In the present study, 39.81% (43/108) of the identified MQTLs were verified by the GWAS-MTAs, which is in agreement with the earlier reports on the comparison of MQTLs with MTAs (Yang et al.; Aduragbemi and Soriano 2021). It is indicated from the previous studies (Yang et al.; Aduragbemi and Soriano 2021) and the present study, that, only a part of the MQTL region could be verified by the MTAs. Considering this, we propose the following two hypothesis: (i) none of the approaches whether meta-QTL analysis or GWAS approach included all the genetic variation present in the crop species, and (ii) genetic materials used in the two approaches were entirely different from each other. MQTLs could be considered more stable and consistent, if they are validated by more than one GWA study and also included the large number of initial QTLs from independent experiments. Five such MQTLs (1A.1, 1B.2, 3B.4, 5B.2 and 6B2) with more than 20 initial QTLs were verified with multiple MTAs derived from different GWA studies. These MQTL regions may be further investigated to explore the underlying complex genetics of the quality traits.

More interestingly, five genes identified in the wheat multi-locus GWA study (Yang et al.) were co-localized with four MQTLs detected in the current study. These genes are known to take active participation in the various biological processes to improve the grain quality in the wheat. For instance, gene, *TraesCS1A02G040600* encodes the cupin superfamily proteins such as phospho-glucose isomerase (PGI), which in turn takes part in the non-enzymatic protein storage in plant seeds (Dunwell and Gane 1998)**.** Genes, *TraesCS3D02G095700* and *TraesCS3D02G096000* encode for the wheat allergens and trypsin & alpha-amylase inhibitor, respectively. Wheat allergens such as seed storage proteins (glutelin and prolamins) causing celiac disease may reduce the nutritional value of wheat seeds. Creating the null mutants for such genes through RNAi or CRISPR-CAS9 may reduce such allergic compounds present in the seeds (Zhang et al. 2014). Alpha-amylase inhibitors prevent the hydrolysis of storage starch granules present in the endosperm and there by maintain the integrity of starch in the seeds (Ali and Elozeiri 2017).

Ortho-MQTL mining: revealing conserved genomic regions between wheat and rice

Identification of ortho-MQTLs among the related crop species may provide a better understanding of genes controlling various quality characters with similar evolutionary lineage and conserved functions. Further, this may provide potential opportunities to transfer information from on species to another (Khahani et al. 2020; Kumar Saini et al. 2021). Ortho-MQTL analysis have been sparingly conducted in wheat, with only three studies are available for different studies including, nitrogen-use efficiency and related traits (Quraishi et al. 2017; Kumar Saini et al. 2021) and thermotolerance (Kumar et al. 2020). However, to our knowledge, no study

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

has reported the ortho-MQTLs for the quality related traits in wheat and other cereals so far. This is the first report, where we identified the 23 ortho-MQTLs associated with quality traits conserved between wheat and rice. Functional characterization of gene models present in the wheat ortho-MQTLs revealed their associations with the sugar-phosphate transportation, seed storage protein, lipid storage activities, etc. Wheat gene, *TraesCS1B02G248000* present on the ortho-MQTL1B.4 is belonging to the NAC domain family and is associated with improving the GPC in wheat seeds. Corresponding orthologous rice gene, *BGIOSGA019745* located on MQTL5.6 belongs to the same gene family and is associated with improving the seed size (Mathew et al. 2016). Further, *TraesCS2B02G430300* gene present on the ortho-MQTL2B.3 is belonging to hydrophobic seed protein domain which accumulates hydrophobic proteins (HPS) on the seed surface. HPS are the abundant seed constituent and a potentially hazardous allergen that causes asthma in persons allergic to the crop dust (Gijzen et al. 1999). Phenotypic or genetic screening may be devised to select plants with reduced amounts of HPS on the seed surface and may offers new avenue for lessening the health hazard of seed dust exposure. Orthologue for this gene, *BGIOSGA016898* is present on the rice MQTL4.5 and is belong to the same class of protein domain. Exploring of such conserved genomic regions among related species may unravel the common regulatory pathways and may assist the utilization of genomic resources from the model species, which may greatly contribute to the rapid crop improvement.

Candidate genes available in MQTL regions

MQTLs are considered as the potential targets for mining candidate genes (CGs) associated with the traits in question. Further, MQTL regions have been shown to have high correlation with gene density in genome as revealed by earlier reports on meta-QTL studies conducted in different crops including rice, wheat, and maize (Yang et al.; Swamy et al. 2011; Quraishi et al. 2017). In wheat, previous studies have reported 15,772 gene models for yield, baking quality and grain protein content (Quraishi et al. 2017), 324 gene models for Fusarium head blight (Venske et al. 2019), 228 gene models for drought (Kumar et al. 2020), and 237 gene models (Yang et al.) for yield related traits. In the present study, gene mining within 108 MQTLs identified 2533 gene models in wheat, at least some of them are known to be associated with quality traits.

Gene expression analysis is the study of how genes are transcribed to produce functional gene products, such as RNA or protein. *In silico* expression analysis permitted the identification of 556 gene models with more than 2 TPM expression. An extensive survey of literature revealed their involvement in transcriptional and translation regulation of different genes, signalling mechanism, metabolism, cellular

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

development and transfer etc. (Table. S7). In the present study, ABC transporter genes were identified on chromosomes 1A, 2A, 2B, 4A, 4B, 6D and 7B, which may give opportunity to develop nutritionally enriched, high-yielding *wheat* cultivars. ABC type transporters are known to regulate seed weight, by transcriptional regulation and modulation of the transport of glutathione conjugates in chickpea (Basu et al. 2019). Along with seed weight, they may enhance the quality component traits, as reported in chickpea (Basu et al. 2019). Ammonium transporter genes available in MQTL6A.5 may play a crucial role in the transport of seed nitrogen (N) and metabolic processes for seed N accumulation, seed yield and N use efficiency (NUE) (McDonald and Ward 2016). They take part in increasing the level of specific amino acids, total soluble protein, and specific high-quality proteins in the seeds. Overall, the present study identified at least 165 gene models associated with catechol oxidase activity and metal ion binding (related PPO), Zn-transporter and zinc-binding site (Zn and Fe content), small hydrophilic plant seed protein, amino acid transporter and seed storage helical domain (seed storage protein), sweet-sugar transporters, UDP-glucuronosyl/UDP-glucosyltransferase, sugar/inositol transporter (starch content) etc. which are known to be associated with quality related traits in different crops. 44 of these 165 gene models with functions previously reported to be important for quality related traits in different crops were selected and recommended for future basic studies involving cloning and functional characterization.

Gene expression analysis of hexaploid bread wheat through RNA-seq and the GeneChip®Wheat Genome array identified numerous DEGs related to starch biosynthesis, nutrient reservoir activity, carbohydrate metabolism and seed storage protein synthesis (Kaushik et al. 2020). Seven DEGs identified in (Kaushik et al. 2020) study were co-located with some MQTLs identified on following 7 chromosomes 1A, 2A, 2B, 4A, 6A, 7A and 7B in the present study. Of these seven genes, six genes were downregulated and one gene, *TraesCS7B02G194000.1* was upregulated. Among the downregulated genes, two genes encode for NB-ARC domain containing proteins and one gene for the F-box domain (Table 3). While the functions of the three downregulated genes have not been characterized. Up-regulated gene, *TraesCS7B02G194000* encode the alpha/beta hydrolase enzyme which take part in the multitude of biochemical processes, including bioluminescence, fatty acid and polyketide biosynthesis and metabolism (M 2000).

To strengthen the location of MQTLs identified in this study, a search for co-localization of quality related genes was also carried out. This enabled detection of functionally characterized wheat genes such as, *GPC-B1/NAM-B1*, *Glu-B1-1b*, *Glu-1D-1d*, *1Dx2t*, *Glu-1By9*, *TaSSI*, *TaSSIIa*, *TaGBSSIa* and *TaSSIVb* in different MQTL regions. MQTL6B.1 and MQTL6B.2 were identified to be associated with the GPC, which may also take part in the nutrient remobilization from leaves to developing grains (Distelfeld et al. 2014). Genomic

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

regions underlying the five wheat MQTLs, (viz., 1B.2, 1B.3, 1B.4, 1D.4 and 1D.5) carried some known genes which are the known precursors of HMW glutenin, a prime deciding factor for determining the dough elasticity and bread-making quality in wheat (Anjum et al. 2007). Genes for the starch synthase enzyme available from MQTL1D.2, 7D.2, 7D.3, 7D.4 and 7D.5 regions are responsible for the amylopectin biosynthesis via α-1,4-glycosidic linkages. Amylopectin-A is the unique chemical compounds present in the wheat which trigger the low-density lipoproteins, which take part in the transportation and delivery of fatty acids, triacylglycerol, and cholesterol in many plant organs (Horstmann et al. 2017).

Wheat homologues of known rice genes available in MQTL regions

Comparative genomics study of wheat with model grasses such as rice has revolutionized the molecular genetics and contributed to the wheat improvement by identifying the linkage blocks, gene rearrangements and conserved regions in the wheat (Kumar et al. 2020). Functionally characterized rice genes are known to have similar function in the wheat and their homologues have been identified to be co-located with MQTLs for the various traits (Hanif et al. 2016). For instance, in the present study, three wheat homologues of one rice genes (*OsSWEET4*) were identified in different MQTLs located on homoeologous group 2 chromosomes, these wheat homologues may be engaged in the embryo nourishment by supplying the sufficient amount of nutrients to the developing embryo in the wheat (Yang et al. 2018).

Using comparative genomics, homologues of these genes in wheat can be characterized, and functional markers for these genes can be developed and validated. A meta-analysis of QTLs associated with grain weight in tetraploid wheat, for instance, resulted in the identification of one important locus, mQTL-GW-6A on chromosome 6A (Avni et al. 2018). Within this MQTL region, authors discovered and characterized a wheat homologue of the rice gene, *OsGRF4* (Avni et al. 2018). This suggests that combining a MQTL study with a well-annotated genome can result in the rapid identification of CGs underlying traits of interest. Manipulation and integration of these genes in the breeding programme may contribute to the enhanced wheat quality.

Conclusion

Meta-analysis of QTLs and comparative genomic approaches helped us to dissect the complex genetic architecture underlying various quality parameters such as GPC, total starch, Zn, Fe, PPO, baking and bread making properties in the wheat. Identified MQTLs and corresponding high-confidence CGs may enhance the marker assisted breeding for quality improvement in wheat. Molecular cloning and functional characterization

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of these CGs may further improve the understanding of wheat genetics. Ortho-MQTLs identified in this study may help to understand the common evolutionary pathways underlying the quality traits. Breeders across the globe can make use of most promising MQTLs (viz., 2D.2, 5A.2, 5D.4, and 7B.3) associated with multiple quality traits and CGs in their study to improve the quality related traits in wheat.

Figures

Fig 1. Salient features of the initial QTLs.

(a) number of QTLs available from different chromosomes, (b) sizes of the populations utilized for mapping, (c) number of different types of populations, (d) different types of markers utilized for mapping, (e) methods employed for mapping, (f) PVE, (g) LOD, and (h) peak positions of the QTLs.

Fig 2. Consensus map, "Wheat_Reference_GeneticMap-2021" showing the marker density (number of markers/cM), chromosome length (cM) and number of markers in each chromosome.

Fig 3. Characteristic features of the MQTLs.

(a) distribution of MQTLs identified on different wheat chromosomes, (b) trait-wise distribution of the MQTLs, (c) average CIs of initial QTLs and MQTLs and fold reduction in CI values after meta-QTL analysis, (d) different number of QTLs involved in MQTLs, (e) MQTLs associated with different number of traits.

Fig 4. Distribution of different MQTLs on wheat chromosomes. Explanation for different colours utilized to represent the MQTLs are given at the bottom of the figure.

Fig 5. The syntenic region of MQTLs between the wheat and rice. The genomic position, chromosome number, and common genes between the wheat and rice are indicated.

Fig 6. Circular diagram representing the conserved regions between wheat and rice MQTLs.

Fig 7. Expression pattern of 44 high-confidence candidate genes in the different relevant tissues.

Fig 8. Circular diagram showing the different features of QTLs and MQTLs

(A) number of QTLs for each trait in each MQTL (i.e., Ax, PPO, SV, DRP, nutritional traits, processing quality traits) and number of traits in each MQTL, **(B)** chromosome wise distribution of initial number of QTLs, projected QTLs, average confidence interval (CI) of MQTLs and average CI of initial QTLs,

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(C) number of initial QTLs projected on each of the MQTL, **(D)** MQTLs colocalized with known wheat genes, GWAS-MTAs and functionally characterized rice genes, **(E)** Gene density distribution in the MQTL regions, **(F)** expression pattern of candidate gene models with >2 TPM in the 8 different tissues (i.e., whole grain, whole endosperm, starchy endosperm, aleurone layer, seed coat, starchy endosperm + seed coat, transfer cells and aleurone layer + starchy endosperm), **(G)** syntenic region between wheat and rice MQTLs. Outer circle represents the physical maps of wheat and rice chromosomes (wheat chromosomes 1A to 7D and rice chromosomes R1 to R9).

Supporting information

S1 Table. List of identified QTLs from the 50 (63 populations) independent studies and the related information (QTLs highleted with yellow colour have not been used for MQTL analysis)

S2 Table. Summary of the individual QTL mapping studies utilized in the present study

S3 Table. Traits grouped into different major trait categories

S4 Table. Wheat_Reference_GeneticMap-2021

S5 Table. Details on identified MQTLs detected in the present study

S6 Table. MQTLs co-located with MTAs identified in the earlier GWA studies

S7 Table. Wheat homologues of known rice genes identified in different MQTLs located on sub-genomes A, B and D

S8 Table. Ortho-MQTLs identified between wheat from rice

S9 Table. Details on gene models identified in the present study

S10 Table. List of important gene models and their expression pattern in different tissues (MQTLs in black had >2 TPM, MQTLs in red had >5 TPM, and MQTLs in blue had >10 TPM expressions

S11 Table. Functionally characterized wheat genes co-localized with MQTLs identified in the current study

Declarations

Funding

No external funding was received for this study.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Conflicts of interest

The authors of this manuscript declare no conflicts of interest.

Availability of data and material

Relevant data are included in this paper and its associated Supplementary Information (SI).

Code availability

Not applicable

Author's contribution (SG, DKS, GS, PH, PK, MS, MJT, AS)

AS, SG, and DKS conceived and planned this study. SG, PH, PK, and MJT performed the literature search, retrieved data, developed consensus map, conducted meta-analysis. GS, DKS and MS helped SG in the analysis and interpretation of results and in writing of the first draft of the manuscript. AS, DKS and GS critically revised and edited the manuscript. All authors have read and agree to the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

Acuña-Galindo MA, Mason RE, Subramanian NK, Hays DB (2015) Meta-analysis of wheat QTL regions associated with adaptation to drought and heat stress. Crop Sci 55:477–492. https://doi.org/10.2135/CROPSCI2013.11.0793

Aduragbemi A, Soriano JM (2021) Unravelling consensus genomic regions conferring leaf rust resistance in wheat via meta-QTL analysis. bioRxiv 2021.05.11.443557. https://doi.org/10.1101/2021.05.11.443557

Ali AS, Elozeiri AA (2017) Metabolic Processes During Seed Germination. Adv Seed Biol. https://doi.org/10.5772/INTECHOPEN.70653

Anjum FM, Khan MR, Din A, et al (2007) Wheat Gluten: High Molecular Weight Glutenin Subunits—Structure, Genetics, and Relation to Dough Elasticity. J Food Sci 72:R56–R63. https://doi.org/10.1111/J.1750-3841.2007.00292.X

Avni R, Oren L, Shabtay G, et al (2018) Genome Based Meta-QTL Analysis of Grain Weight in Tetraploid Wheat Identifies Rare Alleles of GRF4 Associated with Larger Grains. Genes (Basel) 9:. https://doi.org/10.3390/GENES9120636

Barrero JM, Cavanagh C, Verbyla KL, et al (2015) Transcriptomic analysis of wheat near-isogenic lines identifies PM19-A1 and A2 as candidates for a major dormancy QTL. Genome Biol 2015 161 16:1–18. https://doi.org/10.1186/S13059-015-0665-6

Basu U, Upadhyaya HD, Srivastava R, et al (2019) ABC Transporter-Mediated Transport of Glutathione Conjugates Enhances Seed Yield and Quality in Chickpea. Plant Physiol 180:253–275. https://doi.org/10.1104/PP.18.00934

Berman NG, Parker RA (2002) Meta-analysis: Neither quick nor easy. BMC Med Res Methodol 2002 21 2:1–9. https://doi.org/10.1186/1471-2288-2-10

Borrill P, Ramirez-Gonzalez R, Uauy C (2016) expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform. Plant Physiol 170:2172–2186. https://doi.org/10.1104/PP.15.01667

Bush WS, Moore JH (2012) Chapter 11: Genome-Wide Association Studies. PLoS Comput Biol 8:. https://doi.org/10.1371/JOURNAL.PCBI.1002822

Cavanagh CR, Chao S, Wang S, et al (2013) Genome-wide comparative diversity uncovers multiple targets of

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

selection for improvement in hexaploid wheat landraces and cultivars. Proc Natl Acad Sci U S A 110:8057–8062. https://doi.org/10.1073/PNAS.1217133110

Cui F, Zhang N, Fan XL, et al (2017) Utilization of a Wheat660K SNP array-derived high-density genetic map for high-resolution mapping of a major QTL for kernel number. Sci Rep 7:. https://doi.org/10.1038/S41598-017-04028-6

Darvasi A, Soller M (1997) A simple method to calculate resolving power and confidence interval of QTL map location. Behav Genet 27:125–132. https://doi.org/10.1023/A:1025685324830

Darzi-Ramandi H, Vahid Shariati J, Tavakol E, et al (2017) Detection of consensus genomic regions associated with root architecture of bread wheat on groups 2 and 3 chromosomes using QTL meta-analysis. Aust J Crop Sci 11:777–785. https://doi.org/10.21475/AJCS.17.11.07.PNE312

Distelfeld A, Avni R, Fischer AM (2014) Senescence, nutrient remobilization, and yield in wheat and barley. J Exp Bot 65:3783–3798. https://doi.org/10.1093/JXB/ERT477

Doerge RW (2002) Mapping and analysis of quantitative trait loci in experimental populations. Nat Rev Genet 3:43–52. https://doi.org/10.1038/NRG703

Dunwell JM, Gane PJ (1998) Microbial relatives of seed storage proteins: Conservation of motifs in a functionally diverse superfamily of enzymes. J Mol Evol 46:147–154. https://doi.org/10.1007/PL00006289

Endelman JB, Plomion C (2014) LPmerge: An R package for merging genetic maps by linear programming. Bioinformatics 30:1623–1624. https://doi.org/10.1093/BIOINFORMATICS/BTU091

Fofana B, Humphreys DG, Rasul G, et al (2009) Mapping quantitative trait loci controlling pre-harvest sprouting resistance in a red x white seeded spring wheat cross. Euphytica 165:509–521. https://doi.org/10.1007/S10681-008-9766-6

Gijzen M, Miller SS, Kuflu K, et al (1999) Hydrophobic Protein Synthesized in the Pod Endocarp Adheres to the Seed Surface. Plant Physiol 120:951. https://doi.org/10.1104/PP.120.4.951

Gillies SA, Futardo A, Henry RJ (2012) Gene expression in the developing aleurone and starchy endosperm of wheat. Plant Biotechnol J 10:668–679. https://doi.org/10.1111/J.1467-7652.2012.00705.X

Godoy J, Gizaw S, Chao S, et al (2018) Genome-wide Association Study of Agronomic Traits in a Spring-

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Planted North American Elite Hard Red Spring Wheat Panel. Crop Sci 58:1838–1852. https://doi.org/10.2135/CROPSCI2017.07.0423

Goffinet B, Gerber S (2000) Quantitative trait loci: A meta-analysis. Genetics 155:463–473. https://doi.org/10.1093/GENETICS/155.1.463

GP W, K K, VJ L (2013) A model based criterion for gene expression calls using RNA-seq data. Theory Biosci 132:159–164. https://doi.org/10.1007/S12064-013-0178-3

Hanif M, Gao F, Liu J, et al (2016) TaTGW6-A1, an ortholog of rice TGW6, is associated with grain weight and yield in bread wheat. Mol Breed 36:1–8. https://doi.org/10.1007/S11032-015-0425-Z

Horstmann SW, Lynch KM, Arendt EK (2017) Starch Characteristics Linked to Gluten-Free Products. Foods 2017, Vol 6, Page 29 6:29. https://doi.org/10.3390/FOODS6040029

Huang XQ, Cloutier S, Lycar L, et al (2006) Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (Triticum aestivum L.). Theor Appl Genet 113:753–766. https://doi.org/10.1007/S00122-006-0346-7

Kaushik M, Rai S, Venkadesan S, et al (2020) Transcriptome Analysis Reveals Important Candidate Genes Related to Nutrient Reservoir, Carbohydrate Metabolism, and Defence Proteins during Grain Development of Hexaploid Bread Wheat and Its Diploid Progenitors. Genes 2020, Vol 11, Page 509 11:509. https://doi.org/10.3390/GENES11050509

Khahani B, Tavakol E, Shariati V, Fornara F (2020) Genome wide screening and comparative genome analysis for Meta-QTLs, ortho-MQTLs and candidate genes controlling yield and yield-related traits in rice. BMC Genomics 21:. https://doi.org/10.1186/S12864-020-6702-1

Kumar A, Saripalli G, Jan I, et al (2020) Meta-QTL analysis and identification of candidate genes for drought tolerance in bread wheat (Triticum aestivum L.). Physiol Mol Biol Plants 26:1713. https://doi.org/10.1007/S12298-020-00847-6

Kumar Saini D, Srivastava P, Pal N, Gupta PK (2021) Meta-QTLs, Ortho-MetaQTLs and Candidate Genes for Grain yield and Associated Traits in Wheat (Triticum aestivum L.). https://doi.org/10.21203/RS.3.RS-430452/V1

Li H-Z, Gao X, Li X-Y, et al (2013) Evaluation of Assembly Strategies Using RNA-Seq Data Associated with

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Grain Development of Wheat (Triticum aestivum L.). PLoS One 8:e83530. https://doi.org/10.1371/JOURNAL.PONE.0083530

Liu S, Hall MD, Griffey CA, McKendry AL (2009) Meta-Analysis of QTL Associated with Fusarium Head Blight Resistance in Wheat. Crop Sci 49:1955–1968. https://doi.org/10.2135/CROPSCI2009.03.0115

Liu Y, Salsman E, Wang R, et al (2020) Meta-QTL analysis of tan spot resistance in wheat. Theor Appl Genet 133:2363–2375. https://doi.org/10.1007/S00122-020-03604-1

M H (2000) Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. Curr Protein Pept Sci 1:209–235. https://doi.org/10.2174/1389203003381405

Mann G, Diffey S, Cullis B, et al (2009) Genetic control of wheat quality: Interactions between chromosomal regions determining protein content and composition, dough rheology, and sponge and dough baking properties. Theor Appl Genet 118:1519–1537. https://doi.org/10.1007/S00122-009-1000-Y

Marone D, Russo MA, Laidò G, et al (2013) Genetic basis of qualitative and quantitative resistance to powdery mildew in wheat: from consensus regions to candidate genes. BMC Genomics 14:562. https://doi.org/10.1186/1471-2164-14-562

Mathew IE, Das S, Mahto A, Agarwal P (2016) Three Rice NAC Transcription Factors Heteromerize and Are Associated with Seed Size. Front Plant Sci 0:1638. https://doi.org/10.3389/FPLS.2016.01638

McCartney CA, Somers DJ, Lukow O, et al (2006) QTL analysis of quality traits in the spring wheat cross RL4452 x “AC domain.” Plant Breed 125:565–575. https://doi.org/10.1111/J.1439-0523.2006.01256.X

McDonald TR, Ward JM (2016) Evolution of Electrogenic Ammonium Transporters (AMTs). Front Plant Sci 0:352. https://doi.org/10.3389/FPLS.2016.00352

Mohler V, Albrecht T, Mrva K, et al (2014) Genetic analysis of falling number in three bi-parental common winter wheat populations. Plant Breed 133:448–453. https://doi.org/10.1111/PBR.12177

Nuttall JG, O’Leary GJ, Panozzo JF, et al (2017) Models of grain quality in wheat—A review. F Crop Res 202:136–145. https://doi.org/10.1016/J.FCR.2015.12.011

Pearce S, Huttly AK, Prosser IM, et al (2015) Heterologous expression and transcript analysis of gibberellin biosynthetic genes of grasses reveals novel functionality in the GA3ox family. BMC Plant Biol 2015 151 15:1–19. https://doi.org/10.1186/S12870-015-0520-7

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Pfeifer M, Kugler KG, Sandve SR, et al (2014) Genome interplay in the grain transcriptome of hexaploid bread wheat. Science (80-) 345:. https://doi.org/10.1126/SCIENCE.1250091

Powell W (1990) Environmental and Genetical Aspects of Pollen Embryogenesis. 45–65. https://doi.org/10.1007/978-3-642-61499-6_2

Quraishi UM, Pont C, Ain QU, et al (2017) Combined genomic and genetic data integration of major agronomical traits in bread wheat (Triticum aestivum L.). Front Plant Sci 8:. https://doi.org/10.3389/FPLS.2017.01843/FULL

S G, J S, M L, et al (2009) Meta-QTL analysis of the genetic control of ear emergence in elite European winter wheat germplasm. Theor Appl Genet 119:383–395. https://doi.org/10.1007/S00122-009-1046-X

Shariatipour N, Heidari B, Richards CM (2021) Meta-analysis of QTLome for grain zinc and iron contents in wheat (Triticum aestivum L.). Euphytica 217:. https://doi.org/10.1007/S10681-021-02818-8

Somers DJ, Isaac P, Edwards K (2004) A high-density microsatellite consensus map for bread wheat (Triticum aestivum L.). Theor Appl Genet 2004 1096 109:1105–1114. https://doi.org/10.1007/S00122-004-1740-7

Sorrells ME, Rota M La, Bermudez-Kandianis CE, et al (2003) Comparative DNA Sequence Analysis of Wheat and Rice Genomes. Genome Res 13:1818. https://doi.org/10.1101/GR.1113003

Sosnowski O, Charcosset A, Joets J (2012) Biomercator V3: An upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. Bioinformatics 28:2082–2083. https://doi.org/10.1093/BIOINFORMATICS/BTS313

Suliman S, Alemu A, Abdelmula A, et al Genome-wide association analysis uncovers stable QTLs for yield and quality traits of spring bread wheat (Triticum aestivum) across contrasting environments. Elsevier

Swamy BM, Vikram P, Dixit S, et al (2011) Meta-analysis of grain yield QTL identified during agricultural drought in grasses showed consensus. BMC Genomics 12:. https://doi.org/10.1186/1471-2164-12-319

Tyagi S, Gupta PK (2012) Meta-analysis of QTLs Involved in Pre-harvest Sprouting Tolerance and Dormancy in Bread Wheat. Triticeae Genomics Genet. https://doi.org/10.5376/TGG.2012.03.0002

Tyagi S, Mir RR, Balyan HS, Gupta PK (2015) Interval mapping and meta-QTL analysis of grain traits in common wheat (Triticum aestivum L.). Euphytica 201:367–380. https://doi.org/10.1007/S10681-014-1217-Y

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Venske E, dos Santos RS, Farias D da R, et al (2019) Meta-analysis of the QTLome of fusarium head blight resistance in bread wheat: Refining the current puzzle. Front Plant Sci 10:. https://doi.org/10.3389/FPLS.2019.00727/FULL

Veyrieras JB, Goffinet B, Charcosset A (2007) MetaQTL: A package of new computational methods for the meta-analysis of QTL mapping experiments. BMC Bioinformatics 8:. https://doi.org/10.1186/1471-2105-8-49

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. https://doi.org/10.1111/PBI.12183

Welcker C, Sadok W, Dignat G, et al A common genetic determinism for sensitivities to soil water deficit and evaporative demand: meta-analysis of quantitative trait loci and introgression lines of maize. academic.oup.com

Winfield MO, Allen AM, Burridge AJ, et al (2016) High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. Plant Biotechnol J 14:1195–1206. https://doi.org/10.1111/PBI.12485

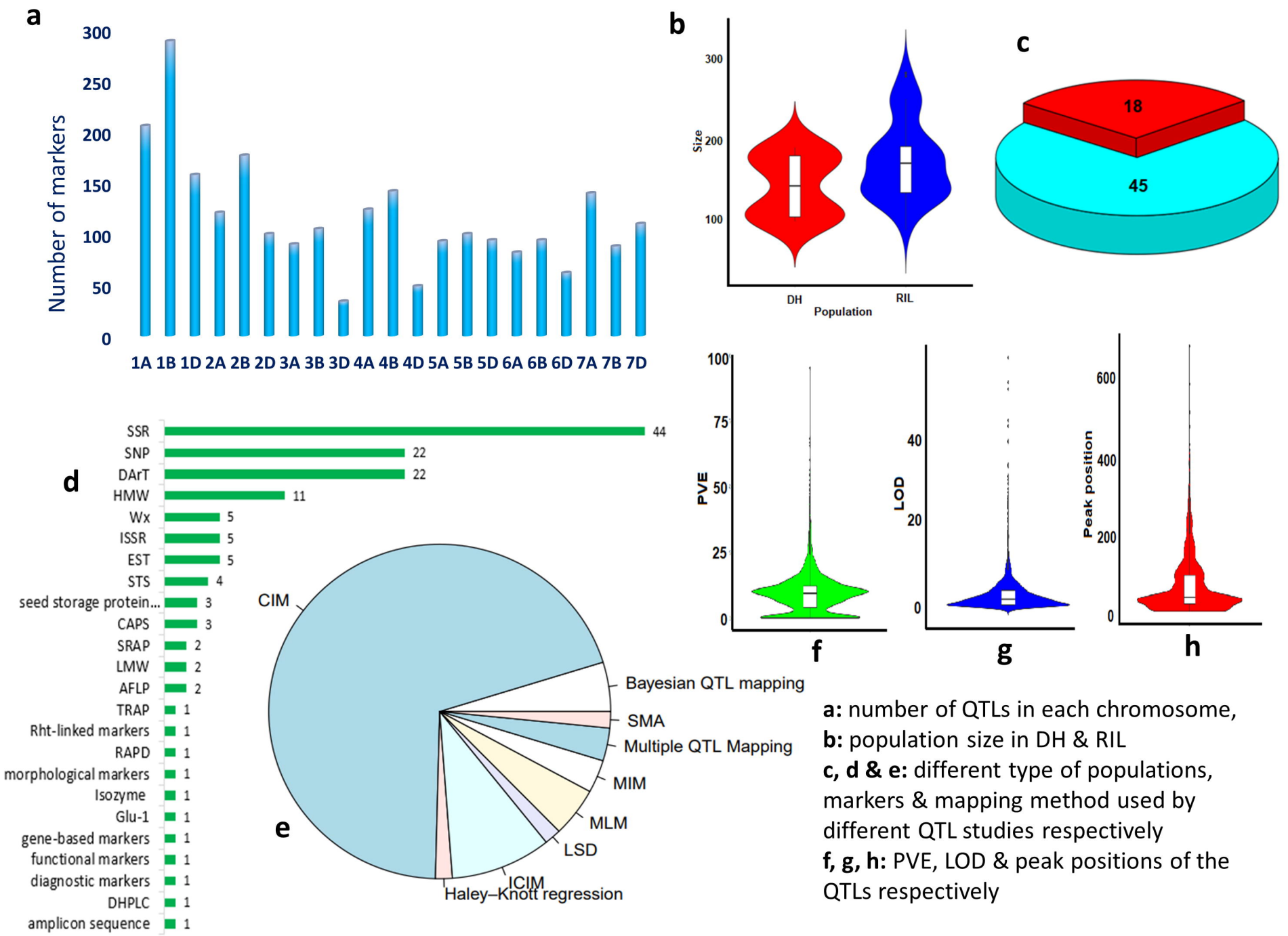
Yang J, Luo D, Yang B, et al (2018) SWEET11 and 15 as key players in seed filling in rice. New Phytol 218:604–615. https://doi.org/10.1111/NPH.15004

Yang Y, Aduragbemi A, Wei D, et al Large-Scale Integration of Meta-QTL and Genome-Wide Association Study Discovers the Genomic Regions and Candidate Genes for Yield and Yield-related Traits in Bread Wheat. https://doi.org/10.21203/rs.3.rs-342038/v1

Youlin P, Xia Z, Ting Z, et al (2021) Meta-Analysis of Qtl Associated With Starch Pasting Viscosity in Rice (Oryza Sativa L.). Bangladesh J Bot 50:269–276. https://doi.org/10.3329/BJB.V50I2.54082

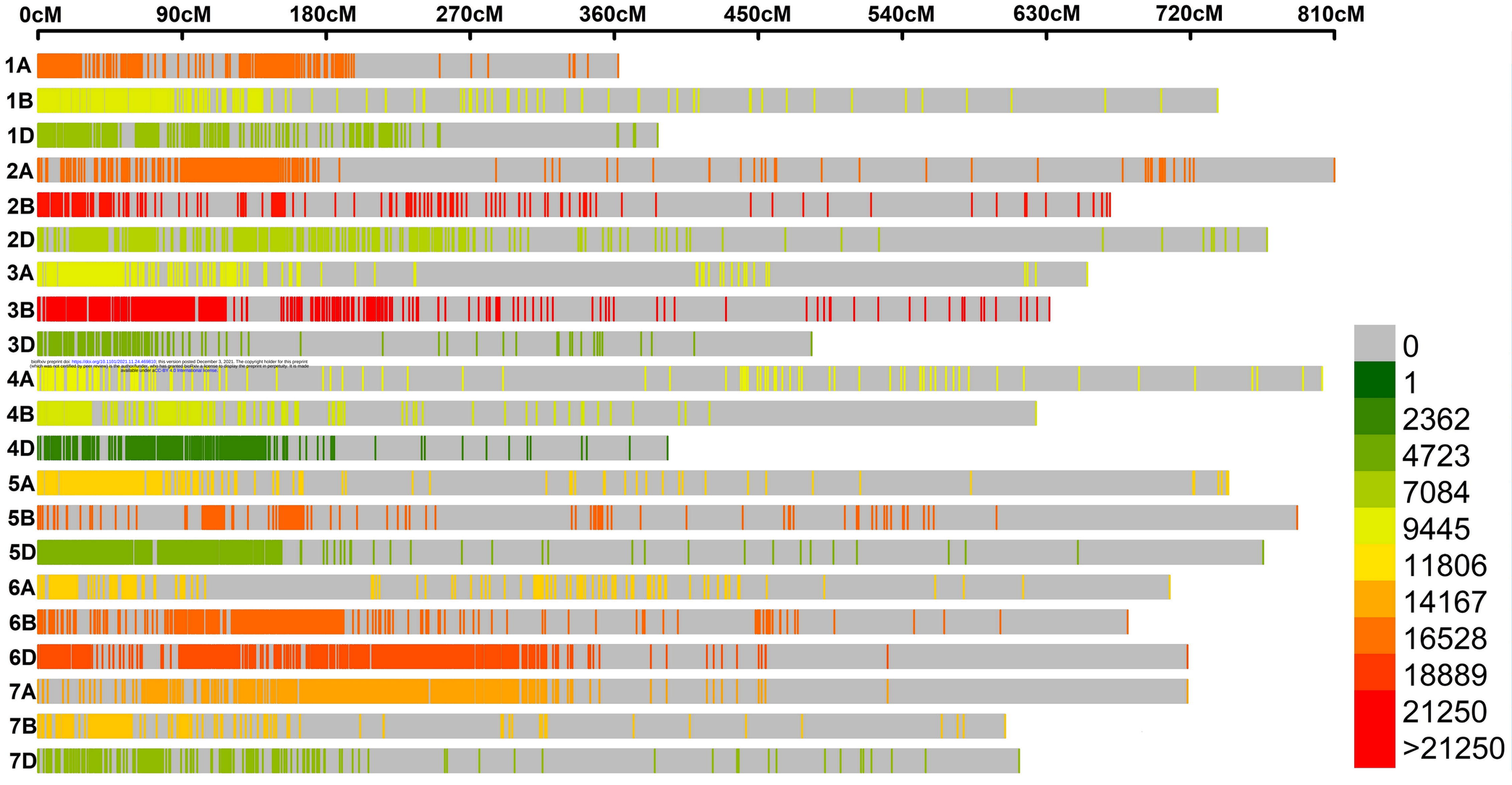
Zhang F, Wen Y, Guo X (2014) CRISPR/Cas9 for genome editing: progress, implications and challenges. Hum Mol Genet 23:R40–R46. https://doi.org/10.1093/HMG/DDU125

Zheng T, Hua C, Li L, et al (2021) Integration of meta-QTL discovery with omics: Towards a molecular breeding platform for improving wheat resistance to Fusarium head blight. Crop J 9:739–749. https://doi.org/10.1016/J.CJ.2020.10.006



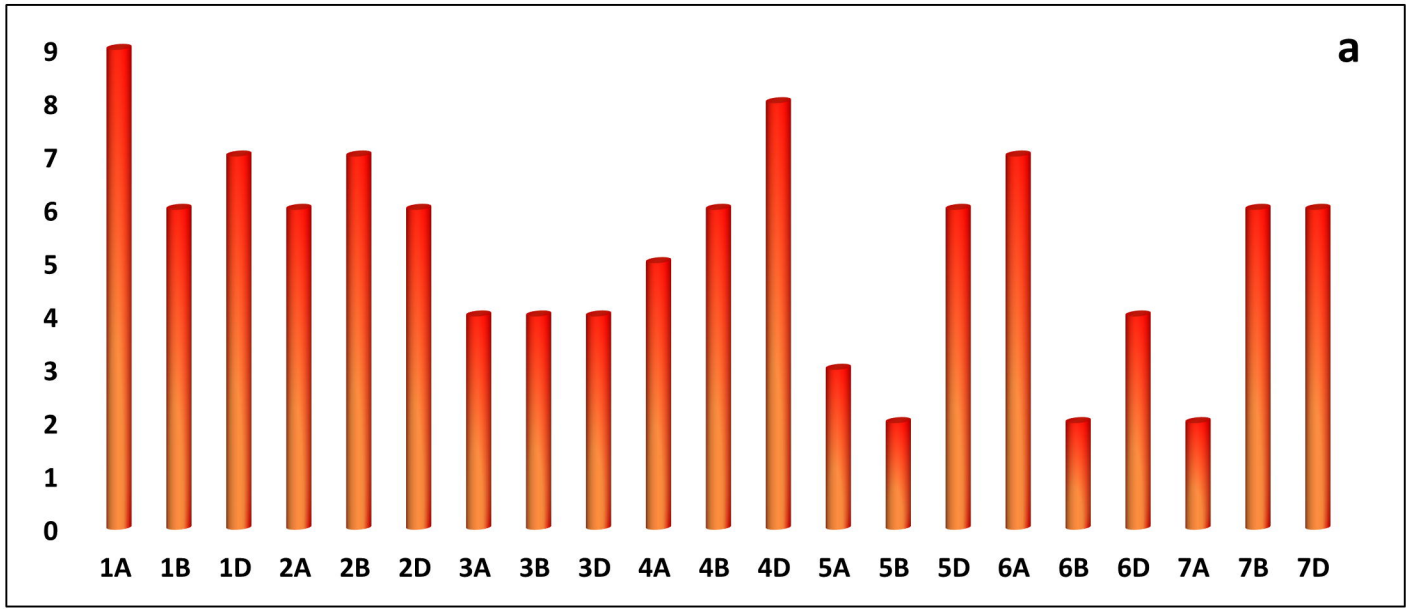
a: number of QTLs in each chromosome,
b: population size in DH & RIL
c, d & e: different type of populations, markers & mapping method used by different QTL studies respectively
f, g, h: PVE, LOD & peak positions of the QTLs respectively

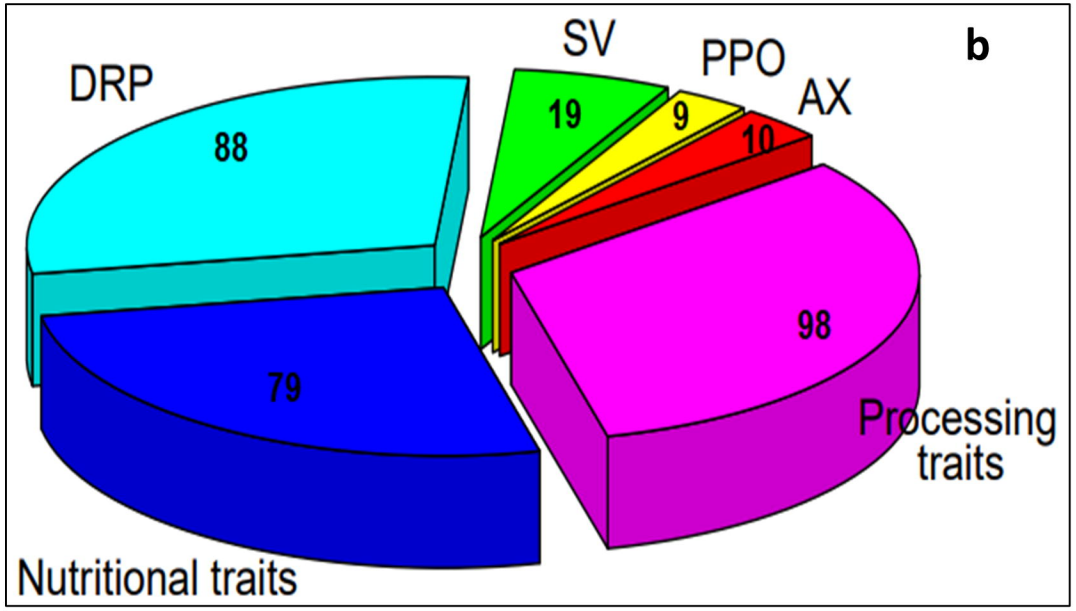
The number of markers within 1cM window size

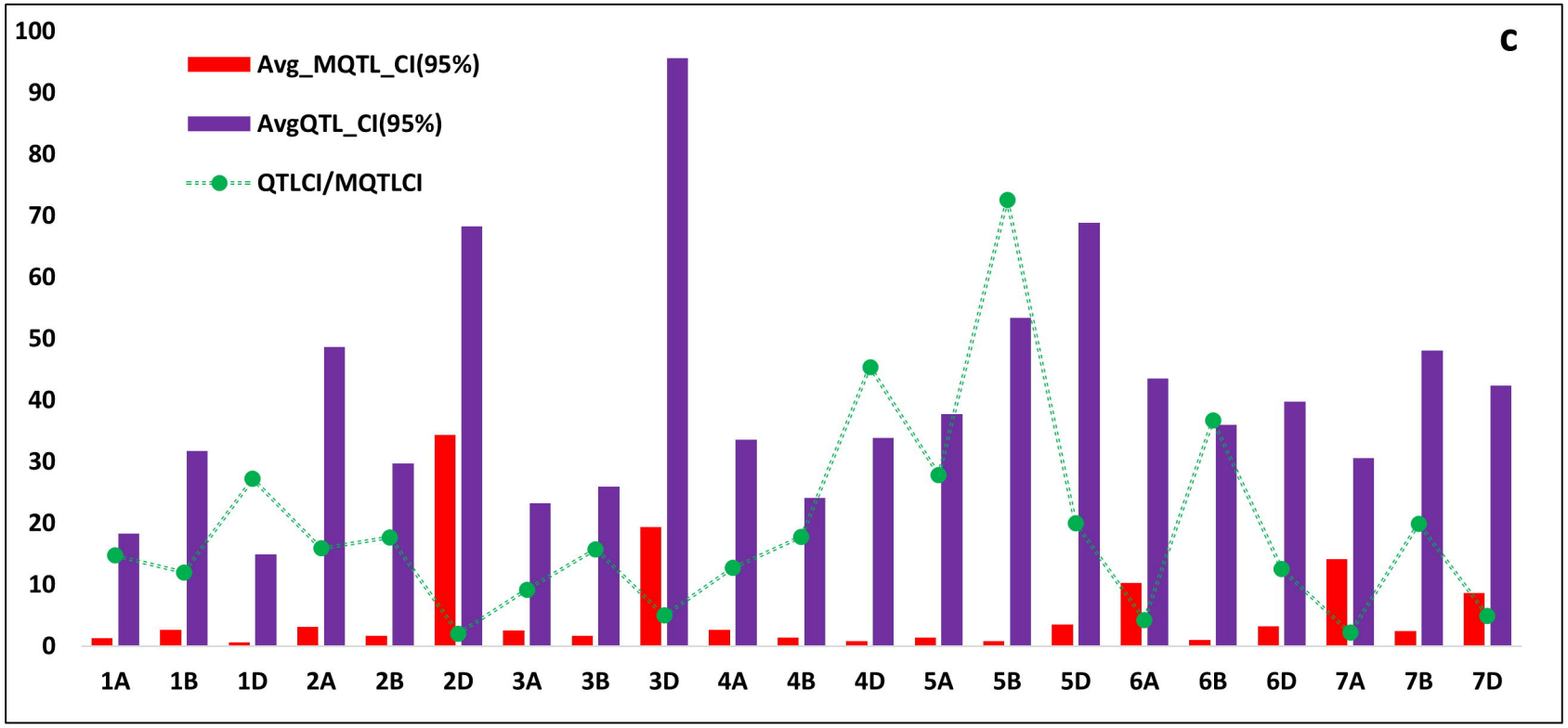


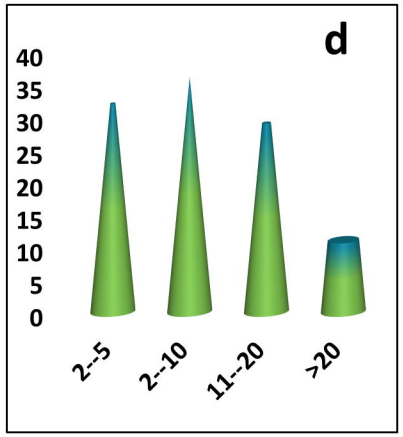
bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-4.0 International license.

Size(cM)	Markers	Chromosome
361.8	16729.0	1A
735.7	9308.0	1B
386.5	6431.0	1D
808.6	16687.0	2A
668.6	20378.0	2B
766.5	7539.0	2D
654.4	9466.0	3A
630.8	21253.0	3B
482.4	5231.0	3D
800.9	9840.0	4A
622.4	8975.0	4B
392.7	2225.0	4D
742.3	12507.0	5A
785.3	17169.0	5B
764.1	5286.0	5D
705.8	12602.0	6A
679.6	16987.0	6B
717.0	17993.0	6D
717.0	14456.0	7A
603.2	12998.0	7B
611.9	6017.0	7D
649.4	11908.4	Average









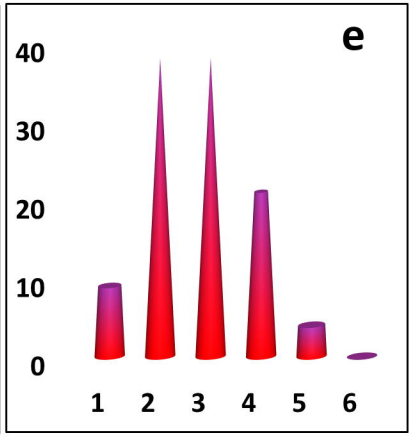
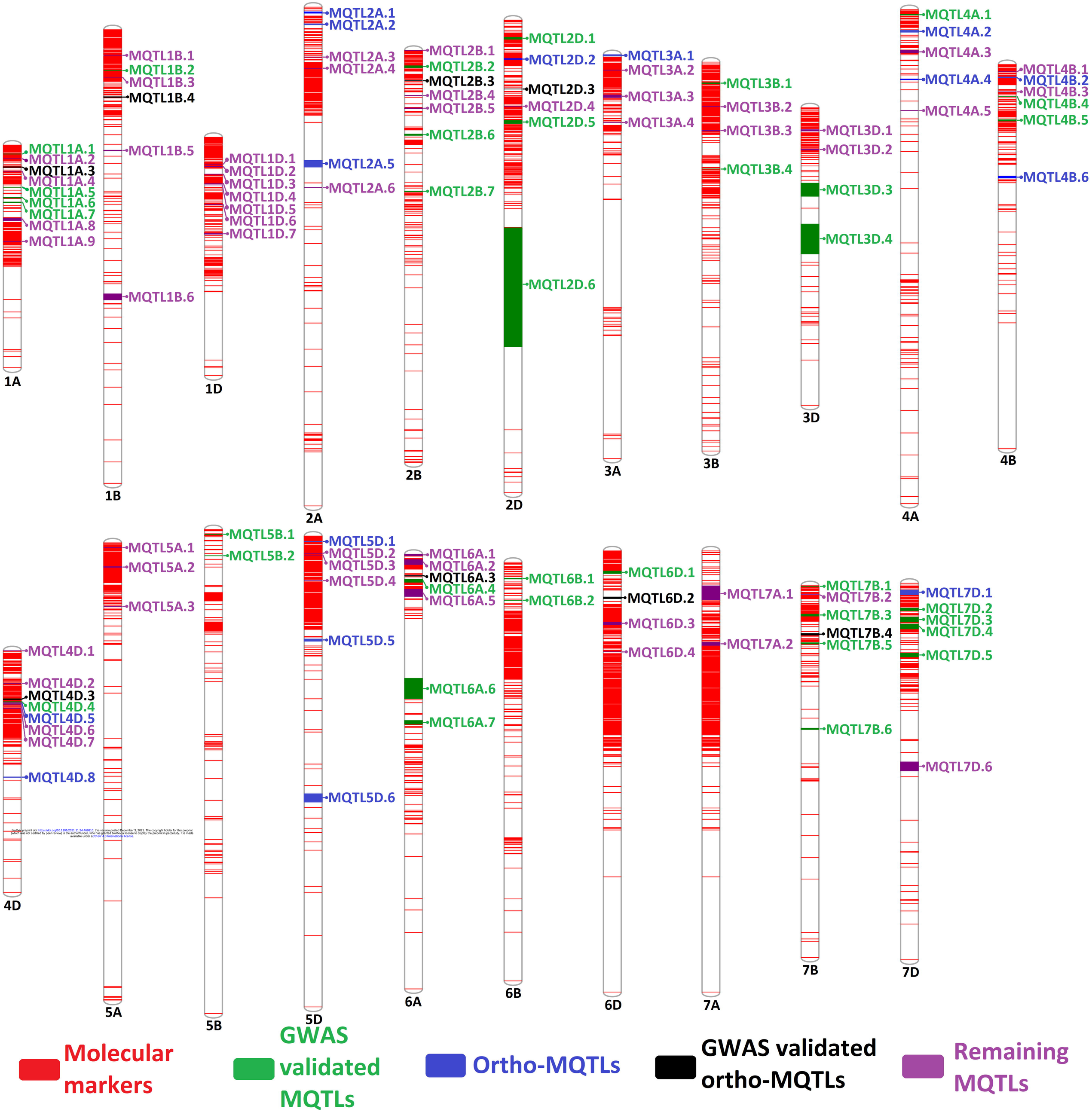
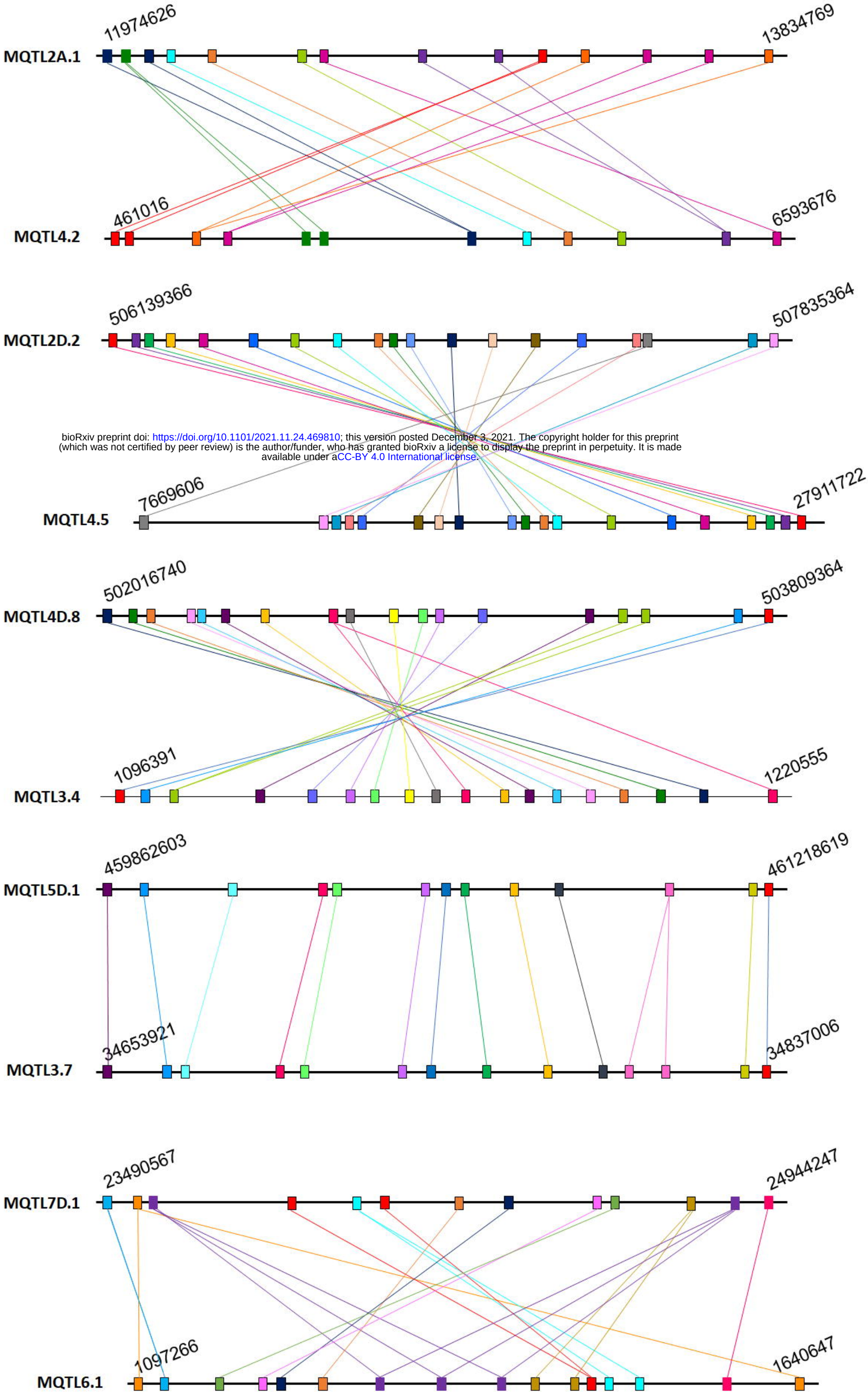


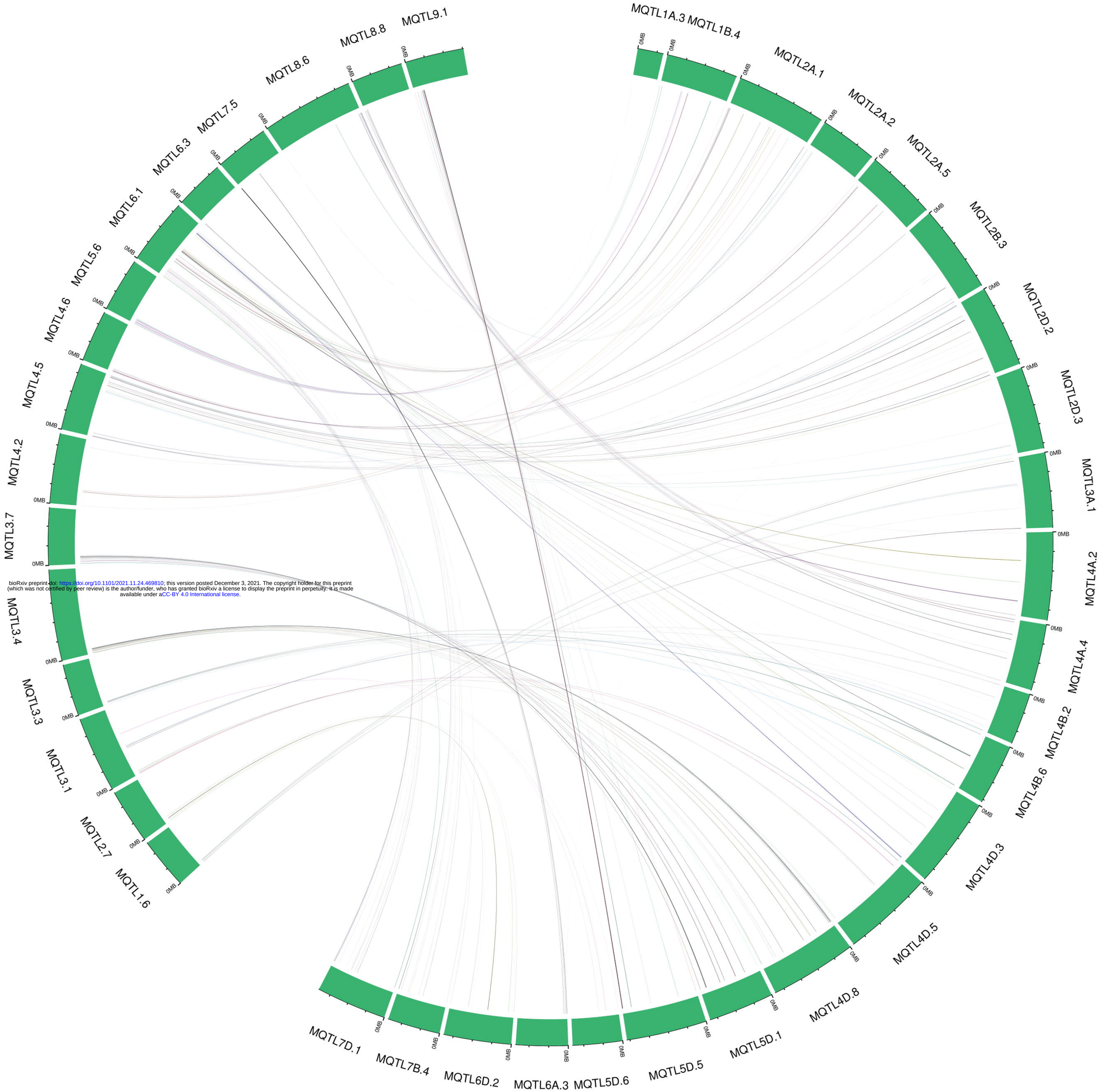
Fig 3. a: MQTLs per chromosome,
b: MQTLs per trait,
c: average CI in MQTL and initial QTLs
d: MQTLs with different number of QTLs
e. MQTLs associated with number of traits



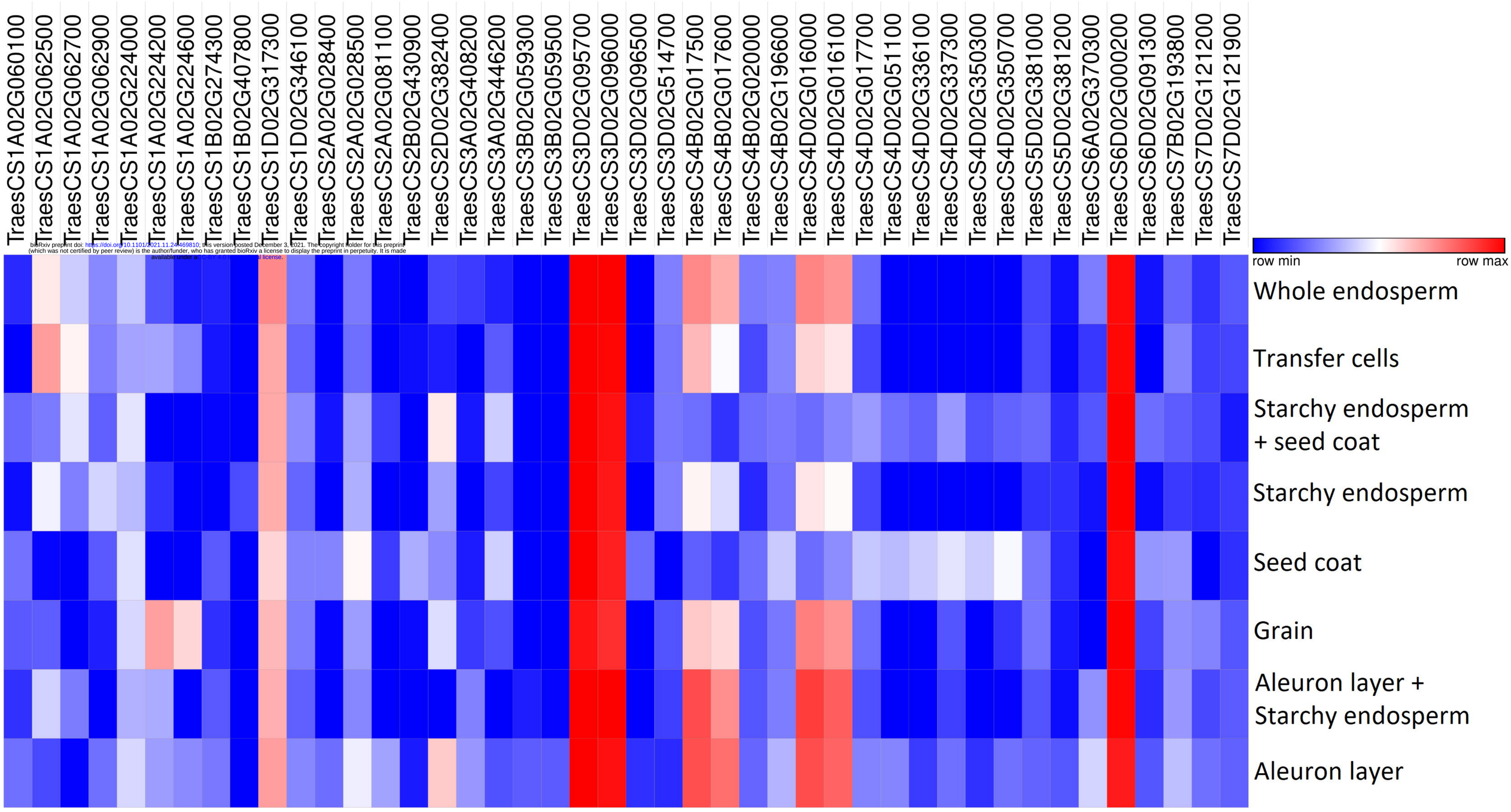
bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



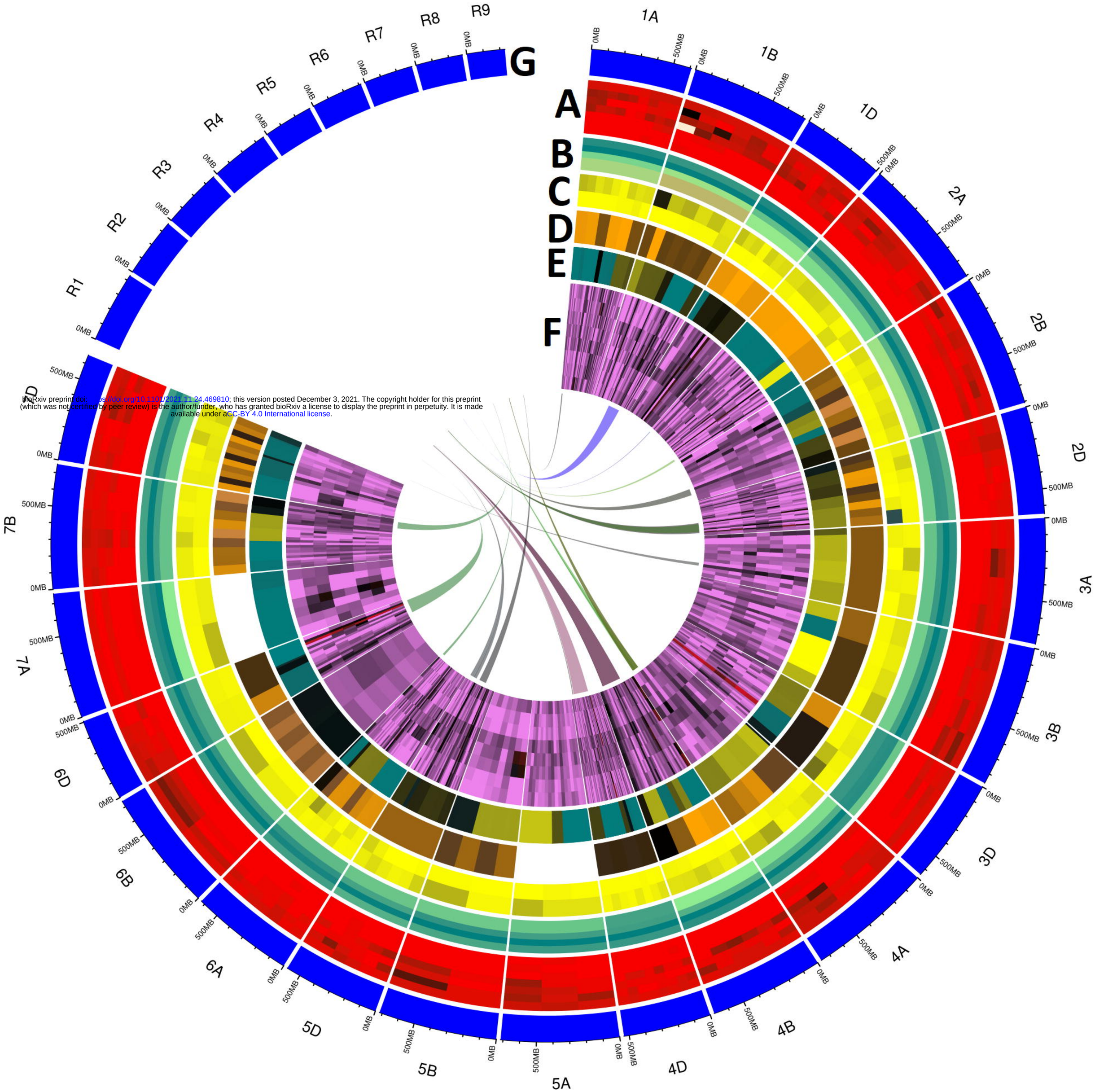
bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.11.24.469810; this version posted December 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.