plant biology

Plant Biology ISSN 1435-8603

STØTTET AF
Promilleafgiftsfonden for landbrug

REVIEW ARTICLE

Recent advances in the genetics underlying wheat grain protein content and grain protein deviation in hexaploid wheat

C. Paina & P. L. Gregersen
Department of Agroecology, Aarhus University, Slagelse, Denmark

Keywords
Genomic prediction; grain protein content; grain protein deviation; nitrogen; senescence; wheat.

Correspondence
C. Paina and P. L. Gregersen, Department of Agroecology, Aarhus University, Forsøgsvej 1, 4200 Slagelse, Denmark.
E-mail: cristiana.paina@agro.au.dk and per.gregersen@agro.au.dk

Editor
J. Bechteler

Received: 13 December 2022;
Accepted: 27 April 2023

doi:10.1111/plb.13550

ABSTRACT

Wheat is one of the most important global crops and selection for better performance has been ongoing since ancient times. As a quantitative trait controlled by the interplay of several genomic loci and under the strong influence of the environment, grain protein content (GPC) is of major interest in breeding programs. Here, we review the most recent contributions to the genetics underlying wheat GPC and grain protein deviation (GPD, representing the relationship between grain protein content and yield), together with the performance of genomic prediction models characterizing these traits. A total of 364 significant loci related to GPC and GPD are positioned on the hexaploid wheat genome, highlighting genomic regions where significant independent QTL overlap, with special focus on two regions located on chromosomes 3A and 5A. Some of the corresponding homoeologous sequences co-locate with significant independent QTL reported on the B and D subgenomes. Overlapping independent QTL from different studies are indicative of genomic regions exhibiting stability across environments and genotypes, with promising candidates for improving grain quality.

BACKGROUND

Grain protein content (GPC) is a primary focus for wheat breeding programs globally. It is a highly important quality trait controlled by multiple quantitative trait loci (QTL) interacting with each other, with the environment and influenced by agricultural management practices (Groos *et al.* 2003; Blanco *et al.* 2006; Nigro *et al.* 2019). The GPC of mature grains ranges between 8% and 20% and translates into flour quality. Breeding efforts over the past few decades have seen a decline in GPC because of its often observed negative association with grain yield (Oury & Godin 2007; Laidig *et al.* 2017). However, some studies in recent years have reported genotypes and populations where this negative association is broken (DePauw *et al.* 2007; Liu *et al.* 2019; Fatiukha *et al.* 2020) and genomic regions positively affecting GPC but without negative effect on yield, as well as genomic regions positively affecting both traits (Thorwarth *et al.* 2019; Ruan *et al.* 2021). These results support the possibility of improving GPC without compromising yield. Thus, within breeding programs, interest has grown into identifying the genetic elements that improve GPC while maintaining yield across environments.

Genotype (g), environment (e) and their interaction (g × e) have significant effects on GPC. Consequently, numerous studies have reported variations in QTL effects for QTL detected in different environments and under diverse genetic backgrounds (Huang *et al.* 2003; Huang *et al.* 2004; Kumar *et al.* 2007; Kuchel *et al.* 2007a,b; Maccaferri *et al.* 2008; McIntyre *et al.* 2010). This translates into the main limitation of previous studies: that the identified QTL is often specific to the plant material and the environments tested and/or the agricultural management practice used (Kumar *et al.* 2018b). Furthermore, most studies, even though performed in multiple environments, are based on relatively small population sizes. In contrast, Juliana *et al.* (2019) genotyped more than 44,000 bread wheat lines, with more than 3,400 lines tested on different continents and considered 50 traits, including traits related to yield and stress resilience and GPC. The design of this study allowed for several traits to be investigated and related at the same time, while the large number of lines increased confidence and support for the results. Another limitation of several studies is related to the ploidy level of the genotypes. Where the focus is on diploid and tetraploid wheat, the D subgenome remains uncovered when it comes to translating results to hexaploid wheat (Rapp *et al.* 2018; Kumar *et al.* 2018b; Giancaspro *et al.* 2019; Liu *et al.* 2019; Nigro *et al.* 2019; Fatiukha *et al.* 2020; Ruan *et al.* 2021). In the last few years, however, a number of studies based on hexaploid wheat have provided significant contributions to uncovering the genetic elements underlying GPC (Goel *et al.* 2019; Michel *et al.* 2019; Chen *et al.* 2020; Muqaddasi *et al.* 2020; Yang *et al.* 2020; Lou *et al.* 2021; Sandhu *et al.* 2021b; Zhang *et al.* 2021; Pu *et al.* 2022; Sandhu *et al.* 2021c). Recent years have seen significant advances in wheat genomics. In spite of the large genome size, with high sequence similarity of the three subgenomes and a large amount of repetitive elements, valuable genomic resources for hexaploid wheat have become available to assist in identifying relevant trait-related genes (Avni *et al.* 2017; Zimin *et al.* 2017; https://www.wheatinitiative.org/iwgsc; Maccaferri *et al.* 2019; Walkowiak *et al.* 2020; Gaurav *et al.* 2021).

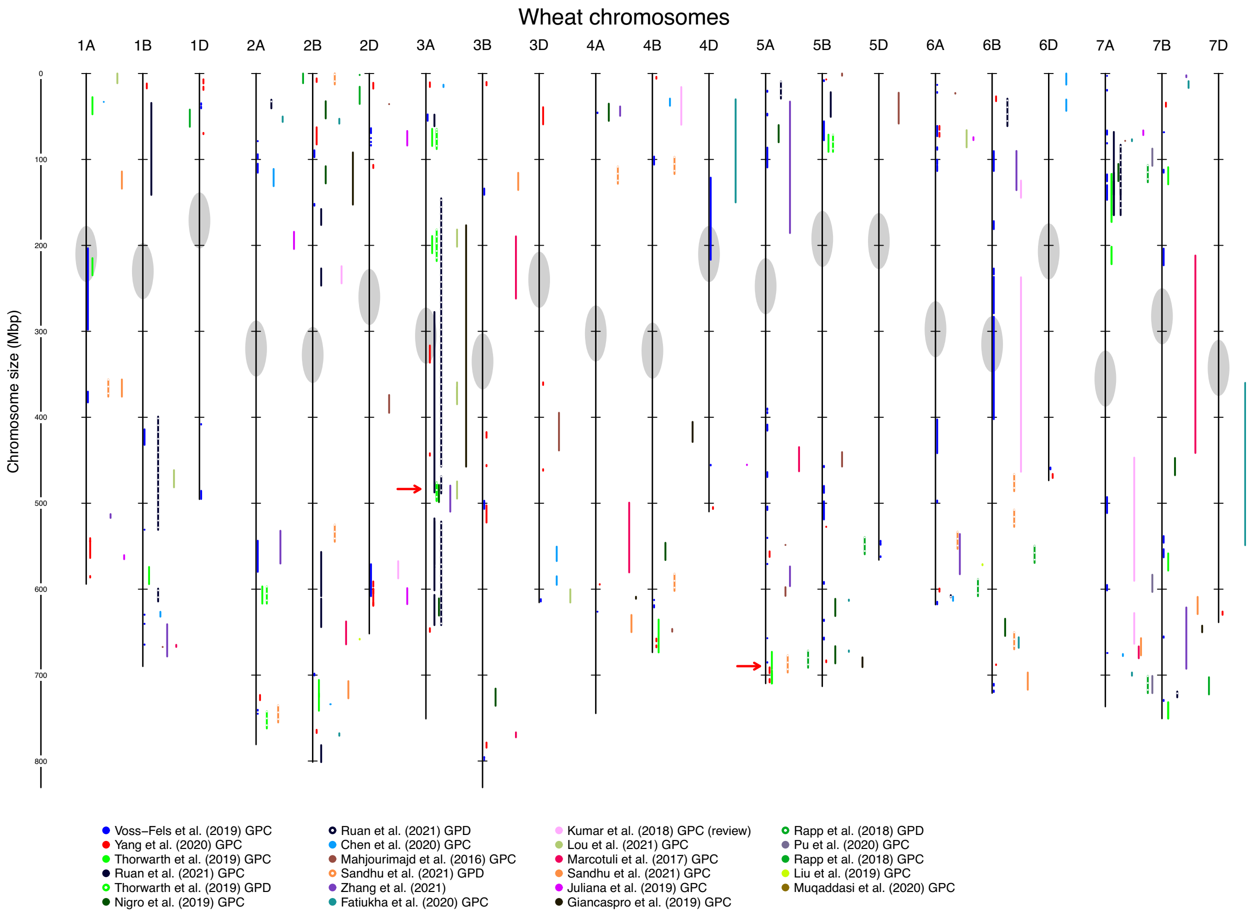

Fig. 1. Tentative physical chromosome positions on the hexaploid wheat genome (IWGSC RefSeq v 1.0) of published QTL for grain protein content (GPC) and grain protein deviation (GPD) in hexa- and tetraploid wheat. Physical positions were mainly based on information on flanking DNA markers of QTL. In cases where only one marker was available, the QTL is marked ±10,000,000 bp of the marker position. For studies conducted in tetraploid wheat, information on marker position was computed based on the expected synteny with hexaploid wheat. The studies are ordered based on number of QTL, to the right of the chromosomes. From the review of Kumar *et al.* (2018b), only the selected nine stable QTL are included in this figure. GPD-specific QTL are marked with dotted lines. The two red arrows at 3A and 5A indicate the two genomic regions discussed in detail in this study.

A wide range of population types have been used to study GPC, including crosses and families, commercial varieties, germplasm collections, hybrid and inbred populations, association populations and populations of synthetic hexaploid wheat. Different approaches have been combined to increase detection power and accuracy, including genome-wide association studies (GWAS) and testing of candidate genes; synteny-based approaches; and tests for significant association of specific metabolites (Nigro *et al.* 2019; Thorwarth *et al.* 2019).

The last review focusing on GPC in wheat was published in 2018, in which Kumar *et al.* (2018b) compiled a comprehensive overview of advances related to the genetic elements underlying wheat GPC, including the genetic relationship between GPC and grain yield. The study presented GPC QTL identified on all chromosomes, with a total of 325 main effect QTL and 42 epistatic QTL reported before 2017, with over 80 QTL showing stability across different environments. Kumar *et al.* (2018b) highlighted nine stable QTL with associated markers reported across the studies, located on chromosomes 2AS, 2BS, 2BL, 2DL, 3BS, 4AL, 4B, 6BS and 7A. Since then, further studies have made valuable contributions and are included in the present review, with a focus on genomic regions and genetic elements that are stable across environments. Figure 1 integrates information about the genomic localization of GPC QTL from recent studies using tetra- and hexaploid wheat.

GPC, GRAIN YIELD AND GPD

Breeding for increased GPC is affected by the negative correlation with yield that is generally reported (Oury & Godin 2007; Laidig *et al.* 2017), which has materialized in varieties with higher grain yield but lower GPC. Data from Danish national trials between 1995 and 2018 (Fig. 2) showed that yield increased while GPC decreased, with newly bred wheat varieties having higher yield at the expense of GPC (www.sortinfo.dk). This is probably an issue, particularly for highly productive wheat cultivation areas such as Northern Europe, where plant breeding has been able to push wheat yields to high levels, even under restricted nitrogen fertilizer regimes. However, some varieties deviate from the general trend (Fig. 2). Some of these observed deviations from the regression line appear to be stable, and thus genetically controlled, indicating the possibility of

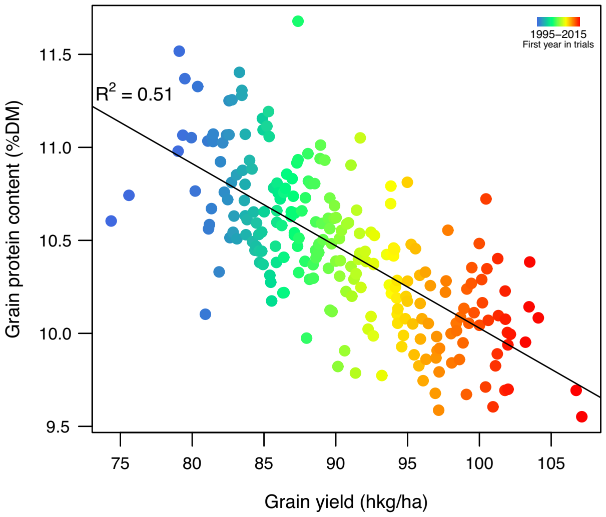

Fig. 2. The relationship between grain yield and grain protein content of wheat cultivars in Danish national trials, between the years 1995 and 2018. Each spot represents one cultivar, with the mean of yield and grain protein content across years and locations. Only cultivars that were in the trials for a minimum of 3 years are included. The colour scale indicates the first year of appearance in the Danish national trials for each cultivar. Source: www.sortinfo.dk.

improving both GPC and grain yield simultaneously in breeding programs (Sears 1998).

Grain protein deviation (GPD) represents the residuals of the regression of protein content on grain yield and was proposed as an indicator of the closeness of the relationship between yield and GPC for a specific variety (Monaghan *et al.* 2001). A positive, high GPD is indicative of a genotype combining good yield and higher GPC. Studies employing GPD highlight its potential to identify and facilitate selection of genotypes with good yield and higher GPC (Monaghan *et al.* 2001; Oury & Godin 2007; Rapp *et al.* 2018; Thorwarth *et al.* 2018; Thorwarth *et al.* 2019; Mosleth *et al.* 2020). Identification of GPD QTL and associated markers is of particular interest due to their potential to increase GPC without negatively affecting yield components. Mosleth *et al.* (2020) reported high genetic variation for GPD, with differences in GPD stability across environments observed for different genotypes. Some genotypes showed stable positive or negative GPD, while others displayed unstable GPD across different environments. Genotype was the main component accounting for GPD variance; however, genotype–site/year interaction and even nitrogen level also played a role.

Several of the identified GPC QTL have a negative effect on yield. However, in some tetraploid wheat trials, no significant correlation between GPC and yield (1000 kernel weight) was observed (Fatiukha *et al.* 2020). The lack of negative association between these two traits in the plant material used in this study suggests that the significant genomic regions identified could be used to improve GPC without negatively affecting yield. Moreover, additive QTL, with positive effects on both GPC and yield, have also been identified, for example on chromosome 2A (Thorwarth *et al.* 2018). These are interesting candidates for the simultaneous improvement of both traits. The significant loci are of high interest, especially when supported by other, independent studies, as revealed by overlapping significant genomic regions, as seen in Fig. 1. In general, hybrid genotypes have superior performance across multiple traits. At a given yield level, hybrids had higher GPC (high GPD) than inbred lines, indicating their potential for overcoming the negative correlation between grain yield and protein content (Thorwarth *et al.* 2018). Hybrids have also shown earlier heading and a longer senescence phase compared to the parents, resulting in a longer grain filling period (Gimenez *et al.* 2021). This finding adds support to the significance of post-anthesis nitrogen uptake previously reported to be associated with GPD (Bogard *et al.* 2010).

GPC, GPD, NITROGEN AND SENESCENCE

Grain protein content (GPC) is strongly influenced by application of nitrogen fertilizer, as well as by the nitrogen uptake capacity and nitrogen utilization efficiency (Garnett *et al.* 2015; Cormier *et al.* 2016). An appropriate balance between post-anthesis nitrogen uptake and remobilization during processes related to senescence has been linked to GPC (Bogard *et al.* 2010). GPD was also shown to be tightly linked with post-anthesis nitrogen uptake (Bogard *et al.* 2010) and a positive association between GPC QTL and nitrogen use efficiency (NUE) has been reported (Zhang *et al.* 2021). Increased nitrogen fertilizer application in the field has been widely used to increase GPC (Belete *et al.* 2018; Sultana *et al.* 2021). However, due to the negative environmental impact of this practice, the development and use of varieties with a positive GPD, independent of external nitrogen applications, is crucial. Plant senescence has a large impact on GPC, but also on grain yield, with the early onset of senescence being correlated to higher GPC and lower grain yield, while delayed senescence is associated with lower GPC and higher yield (Sultana *et al.* 2021). However, genotype specific variations have been described (Sultana *et al.* 2021), suggesting the possibility for selecting gene combinations favouring both GPC and yield. Senescence and NUE are tightly connected and depend on nitrogen availability in the environment. It is well-known that low nitrogen conditions accelerate senescence, while higher nitrogen has the opposite effect (Sultana *et al.* 2021). During senescence, up to 90% of the nitrogen in the plant is remobilized to the grain (Gaju *et al.* 2014). NUE seems to link the negative correlation of GPC and grain yield. Increasing the rate of amino acid remobilization from senescing vegetative organs to the grain has led to higher GPC and lower yield, while a delayed senescence and longer photosynthesis period has resulted in higher grain yield and lower GPC (Benbella & Paulsen 1998a,b). On the other hand, it has also been suggested that final grain yield depends on the grain filling capacity rather than on the length of the photosynthesis period (Borrill *et al.* 2015). Delaying the onset of senescence in genotypes with an increased grain filling capacity should be further explored for improving GPC.

The first gene reported to be involved in increasing GPC was *GPC-B1* (*NAM-B1*), located on the short arm of chromosome 6B, encoding a *NAC* (*No Apical Meristem [NAM], Arabidopsis Transcription Activation Factor [ATAF], cup-shaped cotyledon [CUC]*) transcription factor. It was first identified in wild emmer wheat as a gene for high GPC (Avivi 1977), increasing nitrogen remobilization and grain filling efficiency and accelerating leaf and whole plant senescence (Uauy *et al.* 2006a,b). It was observed that modern wheat cultivars contain a non-functional mutation in the *NAM-B1* gene, which lowers GPC and the amount of minerals, like iron (Fe) and zinc (Zn), and also delays senescence (Uauy *et al.* 2006b). Furthermore, experiments showed that silencing all of the homoeologous *GPC-1* genes resulted in delayed senescence and lower GPC, with reduced Zn and Fe content of the grain (Tabbita *et al.* 2013; Avni *et al.* 2014; Tabbita *et al.* 2017). The connection between GPC and grain mineral content has been observed in other studies, through significant positive correlations between GPC and grain minerals, especially Fe, Mn and Zn (Bhatta *et al.* 2018; Kumar *et al.* 2018a). Moreover, Kumar *et al.* (2018a) identified several marker–trait associations and epistatic interactions, nine of which showed association with three traits of high interest in breeding programs: GPC, Fe content and grain yield. However, there are also reports of *GPC-B1* introgression having no significant effect on GPC (Pan *et al.* 2020). This strongly suggests that the function of *GPC-B1* is dependent on other genetic and/or environmental factors.

The *NAC-S* gene, corresponding to another *NAC* transcription factor, was found to be positively correlated with GPC and delayed senescence, without significant negative effects on yield (Zhao *et al.* 2015). Overexpression of this gene led to increased nitrogen content and GPC, without negatively affecting yield. There are no recent reports on the successful exploitation of the *NAC-S* gene in breeding for higher GPC.

Six candidate genes involved in nitrogen metabolism (*alanine aminotransferase* [*AlaAT*], *NADH-dependent glutamate synthase* [*NADH – GOGAT*], *nitrate reductase* [*NR*], *nitrate transporter 2* [*NRT2*], *glutamine synthetase 2* [*GS2*] and *nitrite reductase* [*NIR*]) showed significant associations with GPC and GPD in tetraploid wheat, with no effect on grain yield (Nigro *et al.* 2019). These are of particular interest as they represent perfect candidates to be used in breeding programs without negative effects on yield. Significant effects of environment, genotype and environment × genotype interaction have been recorded. The strong effect of the environment translated into a wide range of GPC values across seven tested environments. Also, the correlation values of GPC means between some of the environments, as low as 0.38 and up to 0.83, are consistent with the strong influence of the environment on the phenotypic expression of GPC.

The putative amino acid transporter *AAP6* was shown to function as a GPC regulator in rice (Peng *et al.* 2014). Identified through homology cloning in wheat, *AAP6* with its favoured allele located on chromosome 3B, was preferentially expressed in the early stages of developing grains and was significantly correlated with high GPC, leading to the conclusion that it may also be involved in GPC regulation in wheat (Jin *et al.* 2018). The introgression of a *high molecular weight glutenin subunit* (*HMW-GS*) allele led to an increase in wheat GPC and even yield, and induced changes in the abundance of other grain storage proteins (Cao *et al.* 2021). The increased expression of the microRNA-resistant alleles of a class III *homeodomain-leucine-zipper transcription factor HOMEOBOX DOMAIN-2* (*HB-2*) located on chromosome 1 of subgenomes A and D, was shown to increase GPC through increasing amino acid supply during grain development and independent of the senescence pathway (Dixon *et al.* 2022). These results highlight different mechanisms converging on GPC, together with the possibility of identifying GPC regulators in different pathways. The near future should also see the identification of GPC and GPD regulators from other pathways, including phytohormones and sugar metabolism. Additional studies should further investigate the effects of these genes in different wheat genotypes and different environments. It would be interesting to test the simultaneous effects of introgressing not only one, but two, three or more of these proposed GPC increasing genes.

LINKAGE *VERSUS* PLEIOTROPIC EFFECTS

A number of studies have reported QTL for GPC in the same genomic regions with QTL for other traits of interest. This raises the question about linkage or pleiotropic effects in these specific genomic regions. In the case of the *GPC-B1* gene, the introduction of the functional allele in wheat genotypes not only accelerated senescence, but also increased GPC, pointing towards pleiotropic effects (Hagenblad *et al.* 2012; Asplund *et al.* 2013; Avni *et al.* 2014; Chapman *et al.* 2021). Furthermore, seven GPC QTL were found to co-localize with nitrogen-related candidate genes. Functional validation using near-isogenic lines segregating for the two genes *GS2* and *Fd-GOGAT (ferredoxin-dependent glutamate synthase)* located on chromosomes 2B and 2A, respectively, supports their involvement in GPC (Nigro *et al.* 2020). While no pleiotropic effect on yield was observed, this strongly suggests pleiotropic effects on GPC and nitorgen metabolism, making these genes candidates of interest for molecular breeding.

On chromosome 6A, two markers had significant association with both GPC and grain starch content, but with opposite effects on these two traits (Muqaddasi *et al.* 2020), suggestive of potential pleiotropic effects of the genes located in this region. The QTL for GPC and grain yield on chromosomes 3A and 7A also had antagonistic effects, again strongly suggesting pleiotropy (Thorwarth *et al.* 2019). The mode of action for a QTL on chromosome 2A could not be clearly established, but due to the positive effect on both traits, this QTL remains of high interest (Thorwarth *et al.* 2019). Pleiotropy was suggested for a QTL on chromosome 2D, where the same marker showed significant association with different traits, while linkage was suggested on 6A, where distinct markers located in the same region were associated with different traits (Ladejobi *et al.* 2019).

OVERLAPPING GPC AND GPD LOCI ON CHROMOSOMES 3A AND 5A

A total of 41 QTL presented in this review are related to GPD (Fig. 1). Here, we focus on the genomic regions on chromosomes 3A (478.6–488.7 Mb) and 5A (681.4–697.7 Mb and 704.9–708.4 Mb). At least two independent GPD QTL locate in each of these regions, as reported by different authors (Fig. 1, Table S1), which overlap with GPC QTL from other studies. On chromosome 3A, 148 genes are located in the highlighted region including five million bp flanking regions. The five million flanking bp of the two regions on 5A overlap, resulting in a larger hot spot covering 455 genes (Table S1).

The overlapping independent QTL hot spot on chromosome 3A covers a genomic region of approximately ten million bp, plus a flanking region of five million base pairs on each side, resulting in a region of 20 Mb. Lou *et al.* (2021), Fatiukha *et al.* (2020) and Nigro *et al.* (2019) reported GPC QTL in this region, while Ruan *et al.* (2021) and Thorwarth *et al.* (2019) reported GPD QTL. Out of the 148 genes located in this region, 85 genes had homoeologous sequences located on B and/or D in regions where significant independent QTL have been reported in different studies (Table S1). Forty of these were located within the reported QTL intervals, while the remaining 45 were located in the flanking regions. These genes represent good candidates for GPC and GPD. Interestingly, 52 out of the 85 genes had homoeologues located in significant GPC regions only on chromosome 3D, with 30 of these located within QTL intervals and 22 genes in the flanking regions. Among the genes located in this genomic hot spot are: *late embryogenesis abundant* (*LEA*)*-like* proteins and *dehydrins*, which have been shown to have an important role in abiotic stresses (Yu *et al.* 2018). Also, regulators such as transcription factors *MYB* and *MTB1*, *RE1-silencing transcription factor-like*, *AP2/ERF* and *B3 domain-containing transcriptional activators*, genes involved in protein metabolism, like *E3 ubiquitin protein ligases*, *threonine synthase*, *phytochrome-associated serine/threonine-protein phosphatase*, *EEF1A lysine methyltransferase* and *aspartic proteinase NANA*. *Mitogen-activated protein kinases* (*MAPK*) also locate here. *MAPK*s have been shown to regulate plant growth and development, but also responses to environmental stresses (Zhan *et al.* 2017). *Germins* were also reported to be involved in plant development processes as well as biotic and abiotic stresses (Patnaik & Khurana 2001). A number of sequences were annotated as *BRUTUS-like* (*BTS*). *BTS* genes are involved in drought stress responses (Selote *et al.* 2018) and regulate Fe homeostasis (Matthiadis & Long 2016; Rodríguez-Celma *et al.* 2019). *IST1-LIKE* genes have been shown to be essential for normal plant growth, repression of spontaneous cell death and post embryonic lethality in *Arabidopsis* (Buono *et al.* 2016). Another class of candidate genes located here are the transporters, such as *ABC transporters*, *two-pore potassium channel* and *chloride channel protein CLC*.

A strong candidate in the chromosome 3A genomic region is the high affinity *nitrate transporter NRT2.5* gene. Its homoeologous sequences are located in the flanking regions of significant GPC loci on both 3B and 3D. The *NRT2.5* gene was shown to have nitrate transport activity in *Xenopus oocytes* (Kotur *et al.* 2012), while overexpression of its homoeologous gene on 3B led to a significant increase in post-anthesis nitrogen uptake and grain yield (Li *et al.* 2020). The *NRT2.5* gene was linked to a metaQTL for yield (Yang *et al.* 2021). Another strong candidate in this region is *glutamate synthase NADH-GOGAT* (Quraishi *et al.* 2011; Yang *et al.* 2019), which has also been linked to a metaQTL for yield (Yang *et al.* 2021), to NUE (Saini *et al.* 2021) and was validated as underlying a GPC QTL (Nigro *et al.* 2020).

The two regions of interest in the distal region of the long arm of chromosome 5A, having an overlapping flanking region, cover over 36 Mb. Based on the available studies, only 16 genes had homoeologous sequences on 5B and/or 5D located in regions reported as significantly associated with GPC and GPD, with one gene having a homoeologue only on 5D and six genes having homoelogues on 5B. Eleven homoeologous genes were located within significant QTL, while five were in the flanking regions (Table S1). Thorwarth *et al.* (2019), Voss-Fels *et al.* (2019) and Yang *et al.* (2020) reported loci associated with GPC in this genomic hot spot, while Ruan *et al.* (2021) and Sandhu *et al.* (2021) reported loci associated with GPD. Several sequences in this region correspond to uncharacterized or hypothetical proteins. However, there are also transcription regulator sequences (*homeobox protein BEL1* homologue, *bHLH* transcription factors, *scarecrow-like protein 6*, *RAX2*, *GATA transcription factor 2*, *CCR4-NOT transcription complex subunit 1-like* sequences), ubiquitin and protein processing related sequences (*E3 ubiquitin ligase*, *ARABIDILLO 1*) and transporter sequences (*ABC transporters*, *sugar transport protein MST1-like*, *aquaporins*). Although located in the overlapping flanking region, an interesting candidate here is the low affinity *nitrate transporter NRT1* sequence. This gene may be a link between nitrogen uptake and transport (Kong *et al.* 2021), yield (Yang *et al.* 2021) and potentially GPC, given its location close to several independent GPC and GPD QTL. Another potential candidate is the *ethylene-insensitive EIN2* gene, which was reported to be involved in ethylene signalling pathways regulating plant development, senescence and grain size (Jun *et al.* 2004; Zhao *et al.* 2021). The *EIN2* gene could be the link between senescence and GPC. In the same overlapping flanking region are genes involved in flowering. *VRN2-1* and *VRN2-2* genes are repressors of flowering and downregulated by vernalization (Yan *et al.* 2004), while *FT5-1* and *FT5-2* genes are part of the *FT* flowering promoter family (Lv *et al.* 2014). Bogard *et al.* (2011) reported that the onset of flowering was a significant driving factor for leaf senescence during the grain filling period in a winter wheat population segregating for flowering time, observing colocalization of QTL for the duration of post-anthesis leaf senescence with QTL for the time of anthesis. A more recent study based on spring wheat observed a clear allelic effect of flowering time genes on biomass, harvest index and on grain yield (Dreisigacker *et al.* 2021). The results of these studies suggest a clear connection between heading and flowering time, senescence, grain yield and GPC. The distal region of the long arm of the 5A chromosome could be a hub intersecting and regulating these processes. It could be speculated that since each of these genes has a well-established distinct function, linkage and a common epigenetic regulator seems more likely in this region.

GPC AND GENOMIC SELECTION

Different studies have reported a wide range of GPC heritability values, due to heritability being highly dependent and specific to population and environment. Broad sense heritability for GPC has been reported to range from 0.07 (Kumar *et al.* 2018a) to 0.91 (Nigro *et al.* 2019), depending on genotype, location, and number of field trials, but also the type of computational analysis conducted. As expected, heritability values dropped when assessed across environments. For example, Nigro *et al.* (2019) observed heritability values of 0.76–0.91 in specific environments, and 0.6 in a study across seven environments. Although the majority of studies reported a strong and significant effect of the environment, heritability values for GPC are in general moderate to high, increasing the potential of tools like genomic prediction and genomic selection. The prediction accuracy of different studies varies widely, ranging from negative values indicating a very poor prediction ability, to very high positive values. Performance depends on the modelling but also on the composition of the training population. Although the accuracy of different studies varies, most are moderate to high, suggesting good potential for employing these tools to improve GPC and GPD. Semagn *et al.* (2022) reported the highest prediction accuracies for GPC, ranging between 0.60–0.97; they also reported no significant differences in prediction results between conventional and organic farming (Semagn *et al.* 2022).

A number of studies have tested genomic predictions for GPC with very promising results and confirming genomic selection as a method for improving breeding efforts for this trait (Kristensen *et al.* 2018; Rapp *et al.* 2018; Ladejobi *et al.* 2019; Michel *et al.* 2019; Thorwarth *et al.* 2019; Muqaddasi *et al.* 2020). Sandhu *et al.* (2021b) evaluated different models for employing genomic selection in a NAM spring wheat population, suggesting that prediction accuracy of both GPC and grain yield improved by including secondary traits such as spectral reflectance indices and plant growth stage in multi-trait genomic selection models. As suggested, the increase in accuracy is probably due to the fact that the secondary traits included in the model account for some of the environmental effect, a good correlation with the secondary traits and higher heritability of the secondary traits included in the model.

Models for GPC prediction have recently been proposed based on satellite images and a partial least square algorithm (Tan *et al.* 2020) or by combining remote sensing imagery and meteorological data (Xu *et al.* 2020). Good predictions and good results were obtained by predicting traits controlled by multiple small effect loci (Ladejobi *et al.* 2019). Combining metabolic and genomic marker data in prediction models using a population of hybrids led to the conclusion that the best prediction values were achieved using genomic data alone (Thorwarth *et al.* 2019).

In cases with a high positive correlation of GPD and protein content and a high positive correlation of protein yield with grain yield, selecting for a high GPD would mainly increase the protein content, while a selection based on protein yield would mainly improve grain yield. Hence, a combination of the two indices through modelling balances this selection and allows for simultaneously selection for high yield and high GPC (Rapp *et al.* 2018). Most identified QTL had small effects and were specific for only one germplasm set, thus limiting the potential of marker-assisted selection for trait improvement, but highlighting the complex inheritance of GPC and GPD, as well as the significant influence of the plant material and environment. When high and significant genotypic variation is present, genome-wide prediction is a promising option. Genome-wide prediction yields acceptable prediction abilities within the single datasets, but prediction abilities are usually strongly reduced when using different panels as training and as prediction set, highlighting the strong influence of the plant material.

PERSPECTIVES

The results from recent years align with previous observations about the presence of significant loci involved in GPC on all wheat chromosomes, as indicated by the genomic regions highlighted in Fig. 1. This figure comprises 364 significant GPC loci, identified using tetraploid (Marcotuli *et al.* 2017; Rapp *et al.* 2018; Giancaspro *et al.* 2019; Liu *et al.* 2019; Nigro *et al.* 2019; Fatiukha *et al.* 2020; Ruan *et al.* 2021) or hexaploid (Mahjourimajd *et al.* 2016; Thorwarth *et al.* 2018; Juliana *et al.* 2019; Thorwarth *et al.* 2019; Chen *et al.* 2020; Muqaddasi *et al.* 2020; Yang *et al.* 2020; Lou *et al.* 2021; Zhang *et al.* 2021; Sandhu *et al.* 2021c) wheat. Genomic regions identified as significant by more than one independent study are of particular interest, suggesting loci with more stable involvement in GPC and GPD and less affected by environmental factors, especially when these are based on unrelated genotypes and tested in different environments. Such loci overlap in several genomic regions (Fig. 1). Two such regions are highlighted in this study, on chromosomes 3A and 5A, along with potential candidate genes. Homoeologous sequences locate on the B and D subgenomes in regions where significant associations with GPC and GPD were also reported. Future studies based on hexaploid wheat will help elucidate the involvement of genetic elements from the D subgenome, as well as the interplay between the genetic elements located on the three subgenomes. Especially in order to identify associations significant and stable across environments, studies should ideally involve thousands of replicated lines tested in several different environments; however, due to high cost, this is often difficult to achieve and even smaller studies can contribute valuable information.

Gene expression studies, as well as studies based on mutants, could further narrow down the list and contribute towards elucidating the genetic elements involved in GPC and GPD, as well as their genotype and environment specificity. It cannot be ruled out that a slightly different combination of genes and/or alleles control GPC and GPD in wheat lines with different genetic backgrounds and tested in different environments. In practice, wheat varieties are developed to perform in specific climate regions, and performance in a different environment is often altered, usually negatively. Therefore, the involvement of environment-specific genetic elements has to be considered.

Significant genetic variation is observed for wheat GPC and genotypes with high GPC and good yield parameters, stable across multiple environments, have been described. Different studies report relatively high heritability for GPC and good performance for genomic prediction models. This indicates that the genetic resources and genomic tools for GPC improvement are available. With the increasing accuracy and decreasing price of the sequencing technologies, genomic selection is becoming an affordable and preferred tool in breeding. For correlated traits with lower heritability values, multi-trait multi-location modelling is a promising approach, even under strong genotype × environment interactions (Sandhu *et al.* 2021c).

Considering the increasing number of transcriptome studies, gene co-expression network analysis (Langfelder & Horvath 2007, 2008; Oldham *et al.* 2012; Lemoine *et al.* 2021) could be a tool to highlight known and identify yet unknown genetic elements and connections involved in determining GPC and the tightly related traits, such as senescence, NUE and even flowering time. Transcriptome sequencing and differential gene expression studies focusing on genes located in genomic regions highlighted by GWAS have good potential to confirm and narrow the list of candidate genes. As seen with the *NAM-B1*/*GPC-B1* gene, a single gene variation can have a significant effect and be of importance. Moreover, recent studies have shown the potential of proteomics to further understand and improve the protein content and protein composition of wheat grain (Afzal *et al.* 2021; El Hassouni *et al.* 2023).

The results of the several studies focusing on GPC and yield strongly suggest that the genetic background and specific combination of the responsible genetic elements have a great role in determining positive or negative associations between these two traits and thus a positive or negative GPD. Studies on hybrids indicate that the combination of different genetic backgrounds results in improved yield with a higher and positive GPD, clearly pointing towards the possibility of identifying favourable genetic combinations, which will allow the simultaneous improvement of both yield and GPC. Hybrids used in combination with genomic selection, in the context of advances in wheat genomic tools, seems a promising strategy for the development of high-yielding wheat varieties with high GPC.

ACKNOWLEDGEMENTS

We gratefully acknowledge Lars Bonde Eriksen for processing the data used as source for Fig. 2 (www.sortinfo.dk), Fiona Hay (Aarhus University) and anonymous reviewers for valuable comments that improved the manuscript. This study was supported by Promilleafgiftsfonden for Landbrug, Denmark.

AUTHOR CONTRIBUTIONS

Cristiana Paina: conceptualization, methodology, formal analyses, investigation, data curation, writing original draft, writing – review and editing. Per L. Gregersen: conceptualization, methodology, formal analyses, investigation, data curation, visualization, writing – review and editing, funding acquisition.

CONFLICT OF INTEREST

The authors declare no competing interests.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Annotation of sequences co-localizing with significant GPC/GPD QTL. Sheet '3A&5A_sequences_annotation' presents the genes located in the selected significant regions on chromosomes 3A and 5A, with their position on the chromosome (as per the IWGSC RefSeq v 1.1 gene annotation available at https://plants.ensembl.org/Triticum_aestivum) and corresponding annotation obtained by Blastx search against the non-redundant protein database from NCBI (queried May 2022). A significant QTL region was defined as the physical location on the chromosome of a reported QTL flanked by five million base pairs. The Blastx parameters are presented for the best hit. For sequences which had an undescribed best hit (hypothetical protein, predicted protein) but had another significant hit with a specific annotation available, the latter is presented in the following columns, together with the Blastx search result parameters. The value "1" in the Transcript ID column indicates genes with only one available transcript. The value "2" indicates genes with at least two transcripts; for these, the Blastx search parameter values are presented for transcript with id ".2". Significant QTL and their position are inserted at the corresponding location on the chromosomes. Sheets '3BD_homoeologous sequences annot' and '5BD_homoeologous sequences annot' present the annotation of sequences located on subgenomes B and D which are homoeologous or potentially homoeologous to those located in the regions of selected significant QTL of chromosomes 3A and 5A, respectively, and which co-localize with QTL reported on the B and D subgenomes including a five million base pairs flanking region. Homoeologous sequences reported by Ensembl are marked in bold. Sequences located within the five million flanking base pairs are presented in orange, while sequences located within the region of the QTL are presented in red. Blue cell colour of the B and D chromosome field indicates homoeologous sequences which were found co-localizing with significant QTL on only one of the subgenomes, B or D.

REFERENCES

Afzal M., Sielaff M., Curella V., Neerukonda M., El Hassouni K., Schuppan D., Tenzer S., Longin C.F.H. (2021) Characterization of 150 wheat cultivars by LC-MS-based label-free quantitative proteomics unravels possibilities to design wheat better for baking quality and human health. *Plants*, **10**, 424.

Asplund L., Bergkvist G., Leino M., Westerbergh A., Weih M. (2013) Swedish spring wheat varieties with the rare high grain protein allele of NAM-B1 differ in leaf senescence and grain mineral content. *PLoS One*, **8**, 1–11.

Avivi L. (1977) High grain protein-content in wild wheat. *Canadian Journal of Genetics and Cytology*, **19**, 569–570.

Avni R., Nave M., Barad O., Baruch K., Twardziok S.O., Gundlach H., Hale I., Mascher M., Spannagl M., Wiebe K., Jordan K.W., Golan G., Deek J., Ben-Zvi B., Ben-Zvi G., Himmelbach A., MacLachlan R.P., Sharpe A.G., Fritz A., Ben-David R., Budak H., Fahima T., Korol A., Faris J.D., Hernandez A., Mikel M.A., Levy A.A., Steffenson B., Maccaferri M., Tuberosa R., Cattivelli L., Faccioli P., Ceriotti A., Kashkush K., Pourkheirandish M., Komatsuda T., Eilam T., Sela H., Sharon A., Ohad N., Chamovitz D.A., Mayer K.F.X., Stein N., Ronen G., Peleg Z., Pozniak C.J., Akhunov E.D., Distelfeld A. (2017) Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, **357**, 93–97. https://doi.org/10.1126/science.aan0032

Avni R., Zhao R., Pearce S., Jun Y., Uauy C., Tabbita F., Slade A.J., Dubcovsky J., Distelfeld A. (2014) Functional characterization of GPC-1 genes in hexaploid wheat. *Planta*, **239**, 313–324.

Belete F., Dechassa N., Molla N.A., Tana T. (2018) Effect of split application of different N rates on productivity and nitrogen use efficiency of bread wheat (*Triticum aestivum* L.). *Agriculture and Food Security*, **7**, 1–10.

Benbella M., Paulsen G. (1998a) Efficacy of treatments for delaying senescence of wheat leaves: I. Senescence under controlled conditions. *Agronomy Journal*, **90**, 329–332.

Benbella M., Paulsen G. (1998b) Efficacy of treatments for delaying senescence of wheat leaves: II. Senescence and grain yield under field conditions. *Agronomy Journal*, **90**, 332–338.

Bhatta M., Baenziger P., Waters B., Poudel R., Belamkar V., Poland J., Morgounov A. (2018) Genome-wide association study reveals novel genomic regions associated with 10 grain minerals in synthetic hexaploid wheat. *International Journal of Molecular Sciences*, **19**, 3237.

Blanco A., Simeone R., Gadaleta A. (2006) Detection of QTLs for grain protein content in durum wheat. *Theoretical and Applied Genetics*, **112**, 1195–1204.

Bogard M., Allard V., Brancourt-Hulmel M., Heumez E., Machet J.M., Jeuffroy M.H., Gate P., Martre P., Le Gouis J. (2010) Deviation from the grain protein concentration-grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *Journal of Experimental Botany*, **61**, 4303–4312.

Bogard M., Jourdan M., Allard V., Martre P., Perretant M.R., Ravel C., Heumez E., Orford S., Snape J., Griffiths S., Gaju O., Foulkes J., Le Gouis J. (2011) Anthesis date mainly explained correlations between post-anthesis leaf senescence, grain yield, and grain protein concentration in a winter wheat population segregating for flowering time QTLs. *Journal of Experimental Botany*, **62**, 3621–3636.

Borrill P., Fahy B., Smith A., Uauy C. (2015) Wheat grain filling is limited by grain filling capacity rather than the duration of flag leaf photosynthesis: a case study using NAM RNAi plants. *PLoS One*, **10**, e0134947.

Buono R.A., Paez-Valencia J., Miller N.D., Goodman K., Spitzer C., Spalding E.P., Otegui M.S. (2016)

Role of SKD1 regulators LIP5 and IST1-LIKE1 in endosomal sorting and plant development. *Plant Physiology*, **171**, 251–264.

Cao H., Duncan O., Islam S., Zhang J., Ma W., Millar A.H. (2021) Increased wheat protein content via introgression of an HMW glutenin selectively reshapes the grain proteome. *Molecular & Cellular Proteomics*, **20**, 100097. https://doi.org/10.1016/j.mcpro.2021.100097

Chapman E.A., Orford S., Lage J., Griffiths S. (2021) Delaying or delivering: Identification of novel NAM-1 alleles which delay senescence to extend wheat grain fill duration. *Journal of Experimental Botany.*, **72**, 7710–7728.

Chen H., Bemister D., Iqbal M., Strelkov S., Spaner D. (2020) Mapping genomic regions controlling agronomic traits in spring wheat under conventional and organic managements. *Crop Science*, **60**, 2038–2052.

Cormier F., Foulkes J., Hirel B., Gouache D., Moenne-Loccoz Y., Le Gouis J. (2016) Breeding for increased nitrogen-use efficiency: a review for wheat (*T. aestivum* L.). *Plant Breeding*, **135**, 255–278.

DePauw R.M., Knox R.E., Clarke F.R., Wang H., Fernandez M.R., Clarke J.M., McCaig T.N. (2007) Shifting undesirable correlations. *Euphytica*, **157**, 409–415.

Dixon L.E., Pasquariello M., Badgami R., Levin K.A., Poschet G., Ng P.Q., Orford S., Chayut N., Adamski N.M., Brinton J., Simmonds J., Steuernagel B., Searle I.R., Uauy C., Boden S.A. (2022) MicroRNA-resistant alleles of HOMEOBOX DOMAIN-2 modify inflorescence branching and increase grain protein content of wheat. *Science Advances*, **8**, eabn5907. https://doi.org/10.1126/sciadv.abn5907

Dreisigacker S., Burgueño J., Pacheco A., Molero G., Sukumaran S., Rivera-Amado C., Reynolds M., Griffiths S. (2021) Effect of flowering time-related genes on biomass, harvest index, and grain yield in CIMMYT elite spring bread wheat. *Biology*, **10**, 855. https://doi.org/10.3390/biology10090855

El Hassouni K., Afzal M., Steige K.A., Sielaff M., Curella V., Neerukonda M., Tenzer S., Schuppan D., Longin C.F.H., Thorwarth P. (2023) Multiomics based association mapping in wheat reveals genetic architecture of quality and allergenic related proteins. *International Journal of Molecular Sciences*, **24**, 1485.

Fatiukha A., Filler N., Lupo I., Lidzbarsky G., Klymiuk V., Korol A.B., Pozniak C., Fahima T., Krugman T. (2020) Grain protein content and thousand kernel weight QTLs identified in a durum × wild emmer wheat mapping population tested in five environments. *Theoretical and Applied Genetics*, **133**, 119–131.

Gaju O., Allard V., Martre P., Le Gouis J., Moreau D., Bogard M., Hubbart S., Foulkes M.J. (2014) Nitrogen partitioning and remobilization in relation to leaf senescence, grain yield and grain nitrogen concentration in wheat cultivars. *Field Crops Research*, **155**, 213–223.

Garnett T., Plett D., Heuer S., Okamoto M. (2015) Genetic approaches to enhancing nitrogen-use efficiency (NUE) in cereals: challenges and future directions. *Functional Plant Biology*, **42**, 921–941.

Gaurav K., Arora S., Silva P., Sánchez-Martín J., Horsnell R., Gao L., Brar G.S., Widrig V., John Raupp W., Singh N., Wu S., Kale S.M., Chinoy C., Nicholson P., Quiroz-Chávez J., Simmonds J., Hayta S., Smedley M.A., Harwood W., Pearce S., Gilbert D., Kangara N., Gardener C., Forner-Martínez M., Liu J., Yu G., Boden S.A., Pascucci A., Ghosh S., Hafeez A.N., O'Hara T., Waites J., Cheema J., Steuernagel B., Patpour M., Justesen A.F., Liu S., Rudd J.C., Avni R., Sharon A., Steiner B., Kirana R.P., Buerstmayr H., Mehrabi A.A., Nasyrova F.Y., Chayut N., Matny O., Steffenson B.J., Sandhu N., Chhuneja P., Lagudah E., Elkot A.F., Tyrrell S., Bian X., Davey R.P., Simonsen M., Schauser L., Tiwari V.K., Kutcher H.R., Hucl P., Li A., Liu D.-C., Mao L., Xu S., Brown-Guedira G., Faris J., Dvorak J., Luo M.-C., Krasileva K., Lux T., Artmeier S., Mayer K.F.X., Uauy C., Mascher M., Bentley A.R., Keller B., Poland J., Wulff B.B.H. (2021) Population genomic analysis of Aegilops tauschii identifies targets for bread wheat improvement. *Nature Biotechnology*, **40**, 422–431. https://doi.org/10.1038/s41587-021-01058-4

Giancaspro A., Giove S.L., Zacheo S.A., Blanco A., Gadaleta A. (2019) Genetic variation for protein content and yield-related traits in a durum population derived from an inter-specific cross between hexaploid and tetraploid wheat cultivars. *Frontiers in Plant Science*, **10**, 1509.

Gimenez K., Blanc P., Argillier O., Pierre J.B., Le Gouis J., Paux E. (2021) Dissecting bread wheat heterosis through the integration of agronomic and physiological traits. *Biology*, **10**, 907.

Goel S., Singh K., Singh B., Grewal S., Dwivedi N., Alqarawi A., Abd-Allah E.F., Ahmad P., Singh N.K. (2019) Analysis of genetic control and QTL mapping of essential wheat grain quality traits in a recombinant inbred population. *PLoS One*, **14**, e0200669.

Groos C., Robert N., Bervas E., Charmet G. (2003) Genetic analysis of grain protein content, grain yield and thousand kernel weight in bread wheat. *Theoretical and Applied Genetics*, **106**, 1032–1040.

Hagenblad J., Asplund L., Balfourier F., Ravel C., Leino M. (2012) Strong presence of the high grain protein content allele of NAM-B1 in Fennoscandian wheat. *Theoretical and Applied Genetics*, **125**, 1677–1686.

Huang X.Q., Cöster H., Ganal M.W., Röder M.S. (2003) Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, **106**, 1379–1389.

Huang X.Q., Kempf H., Ganal M.W., Röder M.S. (2004) Advanced backcross QTL analysis in progenies derived from a cross between a German elite winter wheat variety and a synthetic wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, **109**, 933–943.

Jin X., Feng B., Xu Z., Fan X., Liu J., Liu Q., Zhu P., Wang T. (2018) TaAAP6-3B, a regulator of grain protein content selected during wheat improvement. *BMC Plant Biology*, **18**, 71. https://doi.org/10.1186/s12870-018-1280-y

Juliana P., Poland J., Huerta-Espino J., Shrestha S., Crossa J., Crespo-Herrera L., Toledo F.H., Govindan V., Mondal S., Kumar U., Bhavani S., Singh P.K., Randhawa M.S., He X., Guzman C., Dreisigacker S., Rouse M.N., Jin Y., Pérez-Rodríguez P., Montesinos-López O.A., Singh D., Mokhlesur R.M., Marza F., Singh R.P. (2019) Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics. *Nature Genetics*, **51**, 1530–1539.

Jun S.-H., Han M.-J., Lee S., Seo Y.S., Kim W.T., An G. (2004) OsEIN2 is a positive component in ethylene signaling in rice. *Plant and Cell Physiology*, **45**, 281–289. https://doi.org/10.1093/pcp/pch033

Kong L., Zhang Y., Du W., Xia H., Fan S., Zhang B. (2021) Signaling responses to N starvation: focusing on wheat and filling the putative gaps with findings obtained in other plants. A review. *Frontiers in Plant Science*, **12**, 656696. https://doi.org/10.3389/fpls.2021.656696

Kotur Z., Mackenzie N., Ramesh S., Tyerman S.D., Kaiser B.N., Glass A. (2012) Nitrate transport capacity of the *Arabidopsis thaliana* NRT2 family members and their interactions with AtNAR2.1. *New Phytologist*, **194**, 724–731.

Kristensen P., Jahoor A., Andersen J., Cericola F., Orabi J., Janss L., Jensen J. (2018) Genome-wide association studies and comparison of models and cross-validation strategies for genomic prediction of quality traits in advanced winter wheat breeding lines. *Frontiers in Plant Science*, **9**, 69.

Kuchel H., Williams K., Langridge P., Eagles H.A., Jefferies S.P. (2007a) Genetic dissection of grain yield in bread wheat. II. QTL-by-environment interaction. *Theoretical and Applied Genetics*, **115**, 1015–1027.

Kuchel H., Williams K.J., Langridge P., Eagles H.A., Jefferies S.P. (2007b) Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theoretical and Applied Genetics*, **115**, 1029–1041.

Kumar A., Jain S., Elias E.M., Ibrahim M., Sharma L.K. (2018b) An overview of QTL identification and marker-assisted selection for grain protein content in wheat. In: Sengar R.S., Singh A. (Eds), *Eco-friendly agro-biological techniques for enhancing crop productivity*. Springer, Singapore, pp 245–274.

Kumar J., Saripalli G., Gahlaut V., Goel N., Meher P.K., Mishra K.K., Mishra P.C., Sehgal D., Vikram P., Sansaloni C., Singh S., Sharma P.K., Gupta P.K. (2018a) Genetics of Fe, Zn, β-carotene, GPC and yield traits in bread wheat (*Triticum aestivum* L.) using a multi-locus and multi-trait GWAS. *Euphytica*, **214**, 219.

Kumar N., Kulwal P.L., Balyan H.S., Gupta P.K. (2007) QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. *Molecular Breeding*, **19**, 163–177.

Ladejobi O., Mackay I.J., Poland J., Praud S., Hibberd J.M., Bentley A.R. (2019) Reference genome anchoring of high-density markers for association mapping and genomic prediction in European winter wheat. *Frontiers in Plant Science*, **10**, 1278.

Laidig F., Piepho H.P., Rentel D., Drobek T., Meyer U., Huesken A. (2017) Breeding progress, variation, and correlation of grain and quality traits in winter rye hybrid and population varieties and national on-farm progress in Germany over 26 years. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, **130**, 981–998.

Langfelder P., Horvath S. (2007) Eigengene networks for studying the relationships between co-expression modules. *BMC Systems Biology*, **1**, 54.

Langfelder P., Horvath S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.

Lemoine G., Scott-Boyer M., Ambroise B., Perin O., Droit A. (2021) GWENA: gene co-expression networks analysis and extended modules characterization in a single Bioconductor package. *BMC Bioinformatics*, **22**, 267.

Li W., He X., Chen Y., Jing Y., Shen C., Yang J., Teng W., Zhao X., Hu W., Hu M., Li H., Miller A.J., Tong Y. (2020) A wheat transcription factor positively sets seed vigour by regulating the grain nitrate signal. *New Phytologist*, **225**, 1667–1680.

Liu J., Huang L., Wang C., Liu Y., Yan Z., Wang Z., Xiang L., Zhong X., Gong F., Zheng Y., Liu D., Wu

B. (2019) Genome-wide association study reveals novel genomic regions associated with high grain protein content in wheat lines derived from wild emmer wheat. *Frontiers in Plant Science*, **10**, 464.

Lou H., Zhang R., Liu Y., Guo D., Zhai S., Chen A., Zhang Y., Xie C., You M., Peng H., Liang R., Ni Z., Sun Q., Li B. (2021) Genome-wide association study of six quality-related traits in common wheat (*Triticum aestivum* L.) under two sowing conditions. *Theoretical and Applied Genetics*, **134**, 399–418.

Lv B., Nitcher R., Han X., Wang S., Ni F., Li K., Pearce S., Wu J., Dubcovsky J., Fu D. (2014) Characterization of FLOWERING LOCUS T1 (FT1) gene in Brachypodium and wheat. *PLoS ONE*, **9**, e94171. https://doi.org/10.1371/journal.pone.0094171

Maccaferri M., Harris N.S., Twardziok S.O., Pasam R.K., Gundlach H., Spannagl M., Ormanbekova D., Lux T., Prade V.M., Milner S.G., Himmelbach A., Mascher M., Bagnaresi P., Faccioli P., Cozzi P., Lauria M., Lazzari B., Stella A., Manconi A., Gnocchi M., Moscatelli M., Avni R., Deek J., Biyiklioglu S., Frascaroli E., Corneti S., Salvi S., Sonnante G., Desiderio F., Marè C., Crosatti C., Mica E., Özkan H., Kilian B., De Vita P., Marone D., Joukhadar R., Mazzucotelli E., Nigro D., Gadaleta A., Chao S., Faris J.D., Melo A.T.O., Pumphrey M., Pecchioni N., Milanesi L., Wiebe K., Ens J., MacLachlan R.P., Clarke J.M., Sharpe A.G., Koh C.S., Liang K.Y.H., Taylor G.J., Knox R., Budak H., Mastrangelo A.M., Xu S.S., Stein N., Hale I., Distelfeld A., Hayden M.J., Tuberosa R., Walkowiak S., Mayer K.F.X., Ceriotti A., Pozniak C.J., Cattivelli L. (2019) Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics*, **51**, 885–895. https://doi.org/10.1038/s41588-019-0381-3

Maccaferri M., Sanguineti M.C., Corneti S., Ortega J.L., Salem M.B., Bort J., DeAmbrogio E., del Moral L.F., Demontis A., El-Ahmed A., Maalouf F., Machlab H., Martos V., Moragues M., Motawaj J., Nachit M., Nserallah N., Ouabbou H., Royo C., Slama A., Tuberosa R. (2008) Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics*, **178**, 489–511.

Mahjourimajd S., Taylor J., Rengel Z., Khabaz-Saberi H., Kuchel H., Okamoto M., Langridge P. (2016) The genetic control of grain protein content under variable nitrogen supply in an Australian wheat mapping population. *PLoS One*, **11**, e0159371.

Marcotuli I., Gadaleta A., Mangini G., Signorile A., Zacheo S., Blanco A., Simeone R., Colasuonno P. (2017) Development of a high-density SNP-based linkage map and detection of QTL for beta-glucans, protein content, grain yield per spike and heading time in durum wheat. *International Journal of Molecular Sciences*, **18**, 1329.

Matthiadis A., Long T.A. (2016) Further insight into BRUTUS domain composition and functionality. *Plant Signaling and Behavior*, **11**, e1204508.

McIntyre C.L., Mathews K.L., Rattey A., Chapman S.C., Drenth J., Ghaderi M., Reynolds M., Shorter R. (2010) Molecular detection of genomic regions associated with grain yield and yield-related components in an elite bread wheat cross evaluated under irrigated and rainfed conditions. *Theoretical and Applied Genetics*, **120**, 527–541.

Michel S., Loschenberger F., Ametz C., Pachler B., Sparry E., Burstmayr H. (2019) Simultaneous selection for grain yield and protein content in genomics-assisted wheat breeding. *Theoretical and Applied Genetics*, **132**, 1745–1760.

Monaghan J., Snape J., Chojecki A., Kettlewell P. (2001) The use of grain protein deviation for identifying wheat cultivars with high grain protein concentration and yield. *Euphytica*, **122**, 309–317.

Mosleth E., Lillehammer M., Pellny T., Wood A., Richeb A., Hussain A., Griffiths S., Hawkesford M.J., Shewry P.R. (2020) Genetic variation and heritability of grain protein deviation in European wheat genotypes. *Field Crops Research*, **255**, 107896.

Muqaddasi Q., Brassac J., Ebmeyer E., Kollers S., Korzun V., Argillier O., Stiewe G., Plieske J., Ganal M., Röder M.S. (2020) Prospects of GWAS and predictive breeding for European winter wheat's grain protein content, grain starch content, and grain hardness. *Scientific Reports*, **10**, 12541.

Nigro D., Fortunato S., Giove S., Mazzucotelli E., Gadaleta A. (2020) Functional validation of glutamine synthetase and glutamate synthase genes in durum wheat near isogenic lines with QTL for high GPC. *International Journal of Molecular Sciences*, **21**, 9253.

Nigro D., Gadaleta A., Mangini G., Colasuonno P., Marcotuli I., Giancaspro A., Giove S.L., Simeone R., Blanco A. (2019) Candidate genes and genome-wide association study of grain protein content and protein deviation in durum wheat. *Planta*, **249**, 1157–1175.

Oldham M., Langfelder P., Horvath S. (2012) Network methods for describing sample relationships in genomic datasets: application to Huntington's disease. *BMC Systems Biology*, **6**, 63.

Oury F., Godin C. (2007) Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes? *Euphytica*, **157**, 45–57.

Pan W., Kidwell K., McCracken V., Bolton R., Allen M. (2020) Economically optimal wheat yield, protein and nitrogen use component responses to varying N supply and genotype. *Frontiers in Plant Science*, **10**, 1790.

Patnaik D., Khurana P. (2001) Germins and germin like proteins: an overview. *Indian Journal of Experimental Biology*, **39**, 191–200.

Peng B., Kong H., Li Y., Wang L., Zhong M., Sun L., Gao G., Zhang Q., Luo L., Wang G., Xie W., Chen J., Yao W., Peng Y., Lei L., Lian X., Xiao J., Xu C., Li X., He Y. (2014) OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice. *Nature Communications*, **5**, 4847. https://doi.org/10.1038/ncomms5847

Pu Z.E., Ye X.L., Li Y., Shi B.X., Guo Z., Dai S.F., Ma J., Liu Z.H., Jiang Y.F., Li W., Jiang Q.T., Chen G.Y., Wei Y.M., Zheng Y.L. (2022) Identification and validation of novel loci associated with wheat quality through a genome-wide association study. *Journal of Integrative Agriculture*, **21**, 3131–3147.

Quraishi U.M., Abrouk M., Murat F., Pont C., Foucrier S., Desmaizieres G., Confolent C., Rivière N., Charmet G., Paux E., Murigneux A., Guerreiro L., Lafarge S., Le Gouis J., Feuillet C., Salse J. (2011) Cross-genome map based dissection of a nitrogen use efficiency ortho-metaQTL in bread wheat unravels concerted cereal genome evolution. *The Plant Journal*, **65**, 745–756. https://doi.org/10.1111/j.1365-313x.2010.04461.x

Rapp M., Lein V., Lacoudre F., Lafferty J., Muller E., Vida G., Bozhanova V., Ibraliu A., Thorwarth P., Piepho H.P., Leiser W.L., Würschum T., Longin C. (2018) Simultaneous improvement of grain yield and protein content in durum wheat by different phenotypic indices and genomic selection. *Theoretical and Applied Genetics*, **131**, 1315–1329.

Rodríguez-Celma J., Connorton J.M., Kruse I., Green R.T., Franceschetti M., Chen Y.T., Cui Y., Ling H.Q., Yeh K.C., Balk J. (2019) Arabidopsis BRUTUS-LIKE E3 ligases negatively regulate iron uptake by targeting transcription factor FIT for recycling. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 17584–17591.

Ruan Y., Yu B., Knox R., Zhang W., Singh A., Cuthbert R., Fobert P., DePauw R., Berraies S., Sharpe A., Fu B.X., Sangha J. (2021) Conditional mapping identified quantitative trait loci for grain protein concentration expressing independently of grain yield in Canadian durum wheat. *Frontiers in Plant Science*, **12**, 642955.

Saini D.K., Chopra Y., Pal N., Chahal A., Srivastava P., Gupta P.K. (2021) Meta-QTLs, ortho-MQTLs and candidate genes for nitrogen use efficiency and root system architecture in bread wheat (*Triticum aestivum* L.). *Physiology and Molecular Biology of Plants*, **27**, 2245–2267.

Sandhu K., Lozada D., Zhang Z., Pumphrey M., Carter A. (2021) Deep learning for predicting complex traits in spring wheat breeding program. *Frontiers in Plant Science*, **11**, 613325.

Sandhu K., Mihalyov P., Lewien M., Pumphrey M., Carter A. (2021b) Combining genomic and phenomic information for predicting grain protein content and grain yield in spring wheat. *Frontiers in Plant Science*, **12**, 613300.

Sandhu K., Mihalyov P., Lewien M., Pumphrey M., Carter A. (2021c) Genomic selection and genome-wide association studies for grain protein content stability in a nested association mapping population of wheat. *Agronomy*, **11**, 2528.

Sears R.G. (1998) Improving grain protein concentration and grain yield in USA hard winter wheat. In: Fowler D.B., Geddes W.E., Johnston A.M., Preston K.R. (Eds), *Wheat protein production and marketing, proceedings of the wheat protein production symposium*. University of Saskatchewan press, Saskatoon, Saskatchewan, Canada, pp 63–67.

Selote D., Matthiadis A., Gillikin J.W., Sato M.H., Long T.A. (2018) The E3 ligase BRUTUS facilitates degradation of VOZ1/2 transcription factors. *Plant, Cell & Environment*, **41**, 2463–2474.

Semagn K., Iqbal M., Crossa J., Jarquin D., Howard R., Chen H., Bemister D.H., Beres B.L., Randhawa H., N'Diaye A., Pozniak C., Spaner D. (2022) Genome-based prediction of agronomic traits in spring wheat under conventional and organic management systems. *Theoretical and Applied Genetics*, **135**, 537–552.

Sultana N., Islam S., Juhasz A., Ma W. (2021) Wheat leaf senescence and its regulatory gene network. *Crop Journal*, **9**, 703–717.

Tabbita F., Lewis S., Vouilloz J., Ortega M., Kade M., Abbate P., Barneix A.J. (2013) Effects of the Gpc-B1 locus on high grain protein content introgressed into Argentinean wheat germplasm. *Plant Breeding*, **132**, 48–52.

Tabbita F., Pearce S., Barneix A. (2017) Breeding for increased grain protein and micronutrient content in wheat: ten years of the GPC-B1 gene. *Journal of Cereal Science*, **73**, 183–191.

Tan C., Zhou X., Zhang P., Wang Z., Wang D., Guo W., Yun F. (2020) Predicting grain protein content of field-grown winter wheat with satellite images and partial least square algorithm. *PLoS One*, **15**, e0228500.

Thorwarth P., Liu G., Ebmeyer E., Schacht J., Schachschneider R., Kazman E., Reif J.C., Würschum T., Longin C. (2019) Dissecting the genetics underlying the relationship between protein content and grain yield in a large hybrid wheat population. *Theoretical and Applied Genetics*, **132**, 489–500.

Thorwarth P., Piepho H., Zhao Y., Ebmeyer E., Schacht J., Schachschneider R., Kazman E., Reif J.C., Würschum T., Longin C.F.H. (2018) Higher grain yield and higher grain protein deviation underline the potential of hybrid wheat for a sustainable agriculture. *Plant Breeding*, **137**, 326–337.

Uauy C., Brevis J., Dubcovsky J. (2006a) The high grain protein content gene Gpc-B1 accelerates senescence and has pleiotropic effects on protein content in wheat. *Journal of Experimental Botany*, **57**, 2785–2794.

Uauy C., Distelfeld A., Fahima T., Blechl A., Dubcovsky J. (2006b) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science*, **314**, 1298–1301.

Voss-Fels K.P., Stahl A., Wittkop B., Lichthardt C., Nagler S., Rose T., Chen T.W., Zetzsche H., Seddig S., Baig M.M., Ballvora A., Frisch M., Ross E., Hayes B.J., Hayden M.J., Ordon F., Leon J., Kage H., Friedt W., Stützel H., Snowdon R.J. (2019) Breeding improves wheat productivity under contrasting agrochemical inpot levels. *Nature Plants*, **5**, 706–714.

Walkowiak S., Gao L., Monat C., Haberer G., Kassa M.T., Brinton J., Ramirez-Gonzalez R.H., Kolodziej M.C., Delorean E., Thambugala D., Klymiuk V., Byrns B., Gundlach H., Bandi V., Siri J.N., Nilsen K., Aquino C., Himmelbach A., Copetti D., Ban T., Venturini L., Bevan M., Clavijo B., Koo D.-H., Ens J., Wiebe K., N'Diaye A., Fritz A.K., Gutwin C., Fiebig A., Fosker C., Fu B.X., Accinelli G.G., Gardner K.A., Fradgley N., Gutierrez-Gonzalez J., Halstead-Nussloch G., Hatakeyama M., Koh C.S., Deek J., Costamagna A.C., Fobert P., Heavens D., Kanamori H., Kawaura K., Kobayashi F., Krasileva K., Kuo T., McKenzie N., Murata K., Nabeka Y., Paape T., Padmarasu S., Percival-Alwyn L., Kagale S., Scholz U., Sese J., Juliana P., Singh R., Shimizu-Inatsugi R., Swarbreck D., Cockram J., Budak H., Tameshige T., Tanaka T., Tsuji H., Wright J., Wu J., Steuernagel B., Small I., Cloutier S., Keeble-Gagnère G., Muehlbauer G., Tibbets J., Nasuda S., Melonek J., Hucl P.J., Sharpe A.G., Clark M., Legg E., Bharti A., Langridge P., Hall A., Uauy C., Mascher M., Krattinger S.G., Handa H., Shimizu K.K., Distelfeld A., Chalmers K., Keller B., Mayer K.F.X., Poland J., Stein N., McCartney C.A., Spannagl M., Wicker T., Pozniak C.J. (2020) Multiple wheat genomes reveal global variation in modern breeding. *Nature*, **588**, 277–283. https://doi.org/10.1038/s41586-020-2961-x

Xu X., Teng C., Zhao Y., Du Y., Zhao C., Yang G., Jin X., Song X., Gu X., Casa R., Chen L., Li Z. (2020) Prediction of wheat grain protein by coupling multisource remote sensing imagery and ECMWF data. *Remote Sensing*, **12**, 1349.

Yan L., Helguera M., Kato K., Fukuyama S., Sherman J., Dubcovsky J. (2004) Allelic variation at the VRN-1 promoter region in polyploid wheat. *Theoretical and Applied Genetics*, **109**, 1677–1686. https://doi.org/10.1007/s00122-004-1796-4

Yang Y., Amo A., Wei D., Chai Y., Zheng J., Qiao P., Cui C., Lu S., Chen L., Hu Y.G. (2021) 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. *Theoretical and Applied Genetics*, **134**, 3083–3109.

Yang Y., Chai Y., Zhang X., Lu S., Zhao Z., Wei D., Chen L., Hu Y.G. (2020) Multi-locus GWAS of quality traits in bread wheat: Mining more candidate genes and possible regulatory network. *Frontiers in Plant Science*, **11**, 1091.

Yang J., Wang M., Li W., He X., Teng W., Ma W., Zhao X., Hu M., Li H., Zhang Y., Tong Y. (2019) Reducing expression of a nitrate-responsive bZIP transcription factor increases grain yield and N use in wheat. *Plant Biotechnology Journal*, **17**, 1823–1833. https://doi.org/10.1111/pbi.13103

Yu Z., Wang X., Zhang L. (2018) Structural and functional dinamics of dehydrins: a aplant protector protein under abiotic stress. *International Journal of Molecular Sciences*, **19**, 3420.

Zhan H., Yue H., Zhao X., Wang M., Song W., Nie X. (2017) Genome-wide identification and analysis of MAPK and MAPKK gene families in bread wheat (*Triticum aestivum* L.). *Genes*, **8**, 284.

Zhang J., She M., Yang R., Jiang Y., Qin Y., Zhai S., Balotf S., Zhao Y., Anwar M., Alhabbar Z., Juhász A., Chen J., Liu H., Liu Q., Zheng T., Yang F., Rong J., Chen K., Lu M., Islam S., Ma W. (2021) Yield-related QTL clusters and the potential candidate genes in two wheat DH populations. *International Journal of Molecular Science*, **22**, 11934.

Zhao D., Derkx A., Liu D., Buchner P., Hawkesford M. (2015) Overexpression of a NAC transcription factor delays leaf senescence and increases grain nitrogen concentration in wheat. *Plant Biology*, **17**, 904–913.

Zhao H., Yin C., Ma B., Chen S., Zhang J. (2021) Ethylene signaling in rice and Arabidopsis: new regulators and mechanisms. *Journal of Integrative Plant Biology*, **63**, 102–125. https://doi.org/10.1111/jipb.13028

Zimin A.V., Puiu D., Hall R., Kingan S., Clavijo B.J., Salzberg S.L. (2017) The first near-complete assembly of the hexaploid bread wheat genome, Triticum aestivum. *GigaScience*, **6**, 1–7. https://doi.org/10.1093/gigascience/gix097