# RESEARCH



# Dissecting the genetic basis of fruiting efficiency for genetic enhancement of harvest index, grain number, and yield in wheat



Dipendra Shahi<sup>1</sup>, Jia Guo<sup>2</sup>, Md Ali Babar<sup>3\*</sup>, Sumit Pradhan<sup>3</sup>, Muhsin AVCl<sup>3</sup>, Jordan McBreen<sup>3</sup>, Zhao Liu<sup>4</sup>, Guihua Bai<sup>5</sup>, Paul St. Amand<sup>5</sup>, Amy Bernardo<sup>5</sup>, Matthew Reynolds<sup>6</sup>, Gemma Molero<sup>6</sup>, Sivakumar Sukumaran<sup>6,7,8</sup>, John Foulkes<sup>9</sup> and Jahangir Khan<sup>10</sup>

# Abstract

**Background** Grain number (GN) is one of the key yield contributing factors in modern wheat (*Triticum aestivum*) varieties. Fruiting efficiency (FE) is a key trait for increasing GN by making more spike assimilates available to reproductive structures. Thousand grain weight (TGW) is also an important component of grain yield. To understand the genetic architecture of FE and TGW, we performed a genome-wide association study (GWAS) in a panel of 236 US soft facultative wheats that were phenotyped in three experiments at two locations in Florida and genotyped with 20,706 single nucleotide polymorphisms (SNPs) generated from genotyping-by-sequencing (GBS).

**Results** FE showed significant positive associations with GN, grain yield (GY), and harvest index (HI). Likewise, TGW mostly had a positive correlation with GY and HI, but a negative correlation with GN. Eighteen marker-trait associations (MTAs) for FE and TGW were identified on 11 chromosomes, with nine MTAs within genes. Several MTAs associated with other traits were found within genes with different biological and metabolic functions including nuclear pore complex protein, F-box protein, oligopeptide transporter, and glycoside vacuolar protein. Two KASP markers showed significant mean differences for FE and TGW traits in a validation population.

**Conclusions** KASP marker development and validation demonstrated the utility of these markers for improving FE and TGW in breeding programs. The results suggest that optimizing intra-spike partitioning and utilizing marker-assisted selection (MAS) can enhance GY and HI.

Keywords Fruiting efficiency, Harvest index, KASP marker, GWAS, Marker-trait association

\*Correspondence:

- Md Ali Babar
- mababar@ufl.edu
- <sup>1</sup> School of Plant, Environmental and Soil Sciences, Louisiana State Agricultural Center, Louisiana State University, Baton Rouge, LA 70803, USA

<sup>2</sup> Inari Agriculture, 1281 Win Hentschel Blvd w1108, West Lafayette, IN 47906, USA

<sup>3</sup> Department of Agronomy, University of Florida, 3105 McCarty Hall B, Gainesville, FL 32611, USA

<sup>4</sup> Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA

 $^{\rm 5}$  USDA-ARS Hard Winter Wheat Genetics Research Unit, Manhattan, KS 66506, USA

<sup>6</sup> International Maize and Wheat Improvement Center (CIMMYT), Carretera México-Veracruz, Km. 45, El Batán, Texcoco 56237, México

 <sup>7</sup> Queensland Alliance for Agriculture and Food Innovation (QAAFI), The University of Queensland, Hermitage Research Facility, Warwick, Queensland 4370, Australia

<sup>8</sup> Present Address: Department of Plant & Environmental Sciences & Advanced Plant Technology Program, Clemson University, Clemson, SC 29634, USA

<sup>9</sup> Division of Plant and Crop Sciences, School of Biosciences, University of Nottingham, Leicestershire LE12 5RD, UK

<sup>10</sup> PARC-Balochistan Agricultural Research and Development Center, Quetta 87300, Pakistan



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

# Background

Wheat is one of the most widely grown crops worldwide in terms of both area and quantity of production [1]. It is a critical component of global food security, fulfilling approximately 20% of the protein and calorie demand of the world population [2, 3]. At present, the average genetic gain in wheat is lower than 1% per annum; however, 2% of annual yield gain is required [4] to meet the future demand of a 9.5 billion population by 2050 [5]. As cultivable land is already declining in many areas due to soil erosion and degradation, it is crucial to improve the genetic yield potential of crops if we are to avoid further loss of natural ecosystems [6, 7]. A better understanding of the genetic relationship between yield and associated traits may facilitate further breakthroughs in developing high-yielding wheat varieties.

Grain number per unit area (GN) is one of the important determinants of grain yield in wheat [5, 8, 9]. Substantial improvement of GN is essential for achieving large genetic gains in wheat yield as grain filling is limited mainly by the sink capacity in modern wheat cultivars under optimal conditions [5, 10]. Therefore, exploring the physiology and genetics involved in GN determination is necessary for grain yield improvement in wheat. GN has a positive relationship with spike partitioning index (SPI) and can be improved by selecting for higher spike dry weight or SPI [5, 7, 11, 12]. A higher proportion of biomass to the spike during anthesis, which is largley due to reduced height in the past, has resulted in the survival of more fertile florets and greater GN [7, 13, 14]. As the modern varieties are mostly within the optimum stature (0.7-1 m) [15, 16], new ways to improve the grain number need to be studied.

In addition to SPI driving greater spike biomass at anthesis, GN may vary among genotypes due to another important partitioning trait, namely fruiting efficiency (FE). FE is the ratio of the grain number at maturity and spike dry weight at anthesis and can be used as an alternative way to explain GN determination. FE refers to the efficiency of a plant in converting the resources that are allocated for growing spikes before anthesis into grains [7, 17, 18]. Increased allocation of assimilates during spike growth to reproductive structures is associated with increased survival of distal florets and grain formation before anthesis [7, 12, 17], and with decreased allocation to structural components of a spike (rachis, glumes, etc.). Therefore, improved FE is a result of efficient intraspike partitioning through which assimilates are diverted towards florets rather than spike structural components [7, 19]. Wide genetic variation for intra-spike partitioning traits such as glume partitioning index (GPI), lemma partitioning index (LmPI), and palea partitioning index (PPI) has been reported [19, 20]. Although more attention has been given to SPI for effectively increasing harvest index (HI) in wheat, FE could be employed as a complementary trait for improving GN and HI [7, 9, 21-24]. A positive association between FE and GN has been reported in many studies [8, 9, 25]. Due to a wide range of variation among genotypes [7, 8, 18, 24, 26], along with responsiveness to selection, FE is a novel physiological trait with huge potential for improving GN, grain yield (GY), and harvest index (HI). As we examine FE and its influence on GN, it is relevant to consider a possible trade-off between SPI and FE [8, 26] because some studies have shown a negative relationship between these traits [7, 26, 27] although it could not be confirmed in some other studies [7, 8]. However, wheat cultivars that combined both high FE and SPI have been reported [19, 20]. Likewise, some studies reported a negative relationship between FE and grain weight (GW) [16, 19, 28], but they were unrelated in other studies [22]. Improvement in GN through FE would be discouraged if GW is highly reduced due to the trade-off between FE and GW [7, 16]. Hence, it is important to study the possible trade-off between FE and GW in wheat. It should also be mentioned that measuring FE is a time-consuming task that requires data collection at two stages: measuring spike dry weight at anthesis and counting grain number at maturity [7, 24]. FE measurement is also a destructive process, both of which make it impractical to be applied in progeny selection [7, 29]. Identifying genetic loci controlling FE and associated markers will make it possible to select FE through marker-assisted selection (MAS) to improve HI and GY in breeding programs.

Genome-wide association study (GWAS) is one of the most popular methods for deciphering genetic architectures of complex traits in plants [30]. With advancements in next-generation sequencing (NGS) technologies, lowcost and high-density genome-wide single nucleotide polymorphisms (SNPs) make it possible for most breeding programs to use the genomic tools in routine breeding selection to accelerate crop improvement [30, 31]. To date, only limited research on understanding the genetic basis of FE in wheat has been reported. The purposes of this study are to (i) evaluate associations between FE and HI, GY and GN in a panel of US soft wheat, (ii) study dry matter partitioning amongst spike structural components in a small subset of accessions from the panel, (iii) use GWAS to identify QTLs and molecular markers associated with FE, and (iv) validate KASP markers developed from the QTL information in a new population.

# Results

#### Genetic variation and heritability

A combined analysis of variance (ANOVA) displayed significant genotypic variations for all the measured traits (Table S1) in Experiment I. Significant effects of environments and genotype-by-environment interactions on FE and TGW were observed. The distribution of TGW values from different environments are presented in Fig. 1. The mean phenotypic values ranged from 33.69 (E1\_Q17) to 49.56 (E1\_Q18) grains  $g^{-1}$  of spike dry weight at A +7d for FE [32]. and 31.53 (E1\_C18) to 38.84 g for the TGW. TGW had a higher broad-sense heritability of 0.75 compared to 0.25 for FE [32].

In Experiment II, ANOVA showed significant variation among different intra-spike partitioning traits (Table S2). Significant environmental effects were observed for all traits, and genotype-by-environment interaction effects were significantly different for most traits (Table S2). FE ranged from 36.19 to 46.46 grains  $g^{-1}$  (Table 1). Likewise, the glume partitioning index (GPI) varied from 0.21 to 0.24, the lemma partitioning index (LmPI) from 0.24 to 0.25, and the palea partitioning index (API) remained constant at 0.13. The awn partitioning index (API) ranged from 0.18 to 0.24 (Table 1). The sterile floret partitioning index (FPI) and the rachis partitioning index (RPI) ranged from 0.03 to 0.04 and 0.15 to 0.18, respectively. The broad-sense heritability for intra-spike partitioning

 
 Table 1
 Summary of phenotypic adjusted means and heritability (Experiment II)

Traits	E2_C18	E2_C19	E2_Q18	E2_Combined	H <sup>2</sup>
FE	36.19	46.46	45.46	40.88	0.18
GPI	0.22	0.21	0.24	0.22	0.88
LmPl	0.25	0.25	0.24	0.25	0.7
API	0.18	0.24	0.18	0.23	0.89
PPI	0.13	0.13	0.13	0.13	0.69
FPI	0.04	0.03	0.04	0.03	0.33
RPI	0.18	0.17	0.15	0.17	0.87

*FE* fruiting efficiency in grains g<sup>-1</sup> of spike dry weight at anthesis + 7 days, *GPI* glume partitioning index, *LmPI* lemma Partitioning index, *API* awn partitioning index, *PPI* palea partitioning index, *FPI* floret partitioning index, *RPI* rachis partitioning index. E2\_C18, E2\_C19, E2\_Q18, and E2\_Combined refers to Citra 2018, Citra 2019, Quincy 2018 and combined dataset respectively for experiment II

traits was generally high, with the exception of FPI. API showed the highest heritability of 0.89, followed by GPI (0.88) and RPI (0.87). FE had an  $H^2$  value of 0.18 comparable to our association studies. These results are summarized in Table 1.



Fig. 1 Distribution of TGW, thousand grain weight values. E1\_C18, E1\_Q17, E1\_Q18, and E1\_Combined refers to Citra 2018, Quincy 2017, Quincy 2018, and combined dataset for experiment I

**Table 2**Pearson's correlation coefficient between phenotypictraits using Best Linear Unbiased Estimates (Experiment I)

	Traits	н	GY	GN	TGW
E1_C18	TGW	0.26**	0.24**	-0.34**	1
	FE	0.51***	0.27**	0.52***	-0.36**
E1_Q17	TGW	0.15*	0.03	-0.31**	1
	FE	0.48***	0.48***	0.54***	-0.23**
E1_Q18	TGW	-0.08	-0.04	-0.08	1
	FE	0.47***	0.38**	0.47***	-0.17*
E1_Combined	TGW	0.10	0.21**	-0.07	1
	FE	0.27**	0.36**	0.43***	-0.12

*HI* harvest index, *GY* grain yield in kg ha<sup>-1</sup>, *GN* Grain number m<sup>-2</sup>, *FE* fruiting efficiency in grains g<sup>-1</sup> of spike dry weight at anthesis + 7 days, *TGW* thousand grain weight in g. E1\_C18, E1\_Q17, E1\_Q18, and E1\_Combined refers to Citra 2018, Quincy 2017, Quincy 2018 and combined dataset respectively for the experiment I. \*, \*\*, \*\*\* denotes significant at 0.05, 0.01 and 0.001 significance levels, respectively

## Phenotypic correlations among traits

FE showed significantly positive correlations with GY  $(r=0.27^{**} \text{ to } 0.48^{***})$ , GN  $(r=0.47^{***} \text{ to } 0.54^{**})$ , and HI  $(r=0.35^{**} \text{ to } 0.51^{***})$  [32], and negative correlation with TGW  $(r=-0.12 \text{ to } 0.26^{**})$  (Table 2). Likewise, TGW showed significant positive correlations with HI  $(r=0.15^{*} \text{ to } 0.26^{***})$  and GY  $(r=0.21^{**} \text{ to } 0.24^{**})$ , but negative correlation with GN  $(r=-0.31^{**} \text{ to } -0.34^{**})$  and FE  $(-0.17^{*} \text{ to } -0.36^{**})$  (Table 2). Path coefficient analysis showed both FE and TGW had significant positive effects on HI (Table 3). Furthermore, FE had a positive indirect effect on HI through GN whereas TGW had negative indirect effects on GN.

FE also had negative correlations with various intra-spike portioning traits including GPI (r = -0.25 to  $-0.73^*$ ), API (r = -0.16 to  $-0.64^*$ ), FPI (r = -0.08 to  $-0.75^*$ ), and RPI (r = -0.15 to -0.44) (Fig. 2). GPI showed positive correlations with FPI (r = 0.02 to  $0.75^*$ ) and RPI (r = 0.29 to  $0.82^*$ ). Likewise, LmPI had positive correlations with RPI (r = 0.01 to 0.30). Path coefficient analysis indicated that GPI, API, and LmPI had negative direct effects on FE in all datasets, except for E2\_C18. In contrast, PPI had positive direct effects on FE (Table S3), while FPI showed negative direct effects on FE in all datasets, except for E2\_Combined.

#### PCA biplot analysis

The principal component analysis (PCA) biplot further supported the association between FE and other traits (HI, GN, TGW, API, FPI, GPI, RPI) from correlation analysis. In the first PCA, HI was close to GY, and FE was close to GN. The first two PCs explained 50.4 to 55.1% and 20.2 to 28.5% of the variation, respectively (Fig. 3). **Table 3** Direct and indirect effects on HI identified through path coefficient analysis (Experiment I)

E1_C18				
Traits	Direct effect	GN	FE	TGW
GN	0.434***	0.434	0.257	-0.199
FE	0.494***	0.226	0.494	-0.216
TGW	0.595***	-0.145	-0.150	0.595
E1_Q17				
Traits	Direct effect	GN	FE	TGW
GN	0.426***	0.426	0.186	-0.115
FE	0.341***	0.232	0.341	-0.090
TGW	0.376***	-0.130	-0.082	0.376
E1_Q18				
Traits	Direct effect	GN	FE	TGW
GN	0.336***	0.336	0.148	0.000
FE	0.31***	0.159	0.317	0.00
TGW	0.001	-0.026	-0.054	0.001
E1_Combined				
Traits	Direct effect	GN	FE	TGW
GN	0.353***	0.353	0.101	-0.030
FE	0.215***	0.165	0.215	-0.030
TGW	0.160**	-0.066	-0.040	0.160

*HI* harvest index, *GN* Grain number m<sup>-2</sup>, *FE* fruiting efficiency in grains g<sup>-1</sup> of spike dry weight at anthesis + 7 days, *TGW* thousand grain weight in g. E1\_C18, E1\_Q17, E1\_Q18, E1\_Combined refers to Citra 2018, Quincy 2017, Quincy 2018 and combined dataset respectively for experiment I. \*, \*\*, \*\*\*\* denotes significant at 0.05, 0.01 and 0.001 significance levels, respectively

For intra-spike partitioning traits, the first two PCs explained 28.6 to 53% and 21.1 to 29.6% of the variation, respectively (Fig. 4). Furthermore, FPI, GPI, and RPI were clustered together, away from a cluster of LmPI and PPI whereas FE and API didn't have any cluster.

# Genome-wide association study and gene annotation

Population structure and LD decay analysis using 20,706 SNPs were performed earlier [33]. To summarize, a total of 7,935 (38.32%), 7,496 (36.20%), and 5,275 (25.48%) SNPs were found in A, B, and D genomes, respectively. The PC analysis showed an admixture among genotypes with three clusters. The LD decay was found to be 3.4 Mbp for whole genome.

Thirteen MTAs for FE were identified on 10 chromosomes (1D, 2A, 2B, 3A, 3D, 5D, 6A, 6D, 7A, and 7D) (Fig. 5, Fig. 6, Figure S1a) with phenotypic variation explained (PVE) ranging from 10.34 to 17.56% (Table 4). Additionally, 5 MTAs on chromosomes 2A, 2B, 3A, 5B, and 6D (Fig. 5, Fig. 6, Figure S1b) were significant with PVE ranging from 9.62 to 14.92% for TGW (Table 4). Two MTAs for FE were repeated in two environments.

Functional gene annotation of the significant MTAs in the IWGSC RefSeq v1.0 reference genome identified



**Fig. 2** Principal component analysis biplot of measured traits using best linear unbiased estimates (BLUEs). HI, harvest index; GY, grain yield; GN, Grain number m<sup>-2</sup>; FE, fruiting efficiency; TGW, thousand grain weight. E1\_C18, E1\_Q17, E1\_Q18, E1\_Combined refers to Citra 2018, Quincy 2017, Quincy 2018, and combined dataset respectively for experiment I

several putative candidate genes, and their physical locations. Several MTAs were within genes with different biological and metabolic functions including zinc finger, F-box protein, nuclear pore complex protein, Oligopeptide transporter, and Fasciclin-like arabinogalactan protein (Table S4). Likewise, five genes that are in proximity (within 200 kb) of significant SNPs for FE were also found (Table S5).

# **KASP** marker validation

A total of 7 KASP markers were developed and used for validation. Of these, one each newly developed KASP markers showed significant mean difference in FE and TGW traits, respectively, between two allelic groups in a validation population (Fig. 7 and Table S6). The VEP search in Ensembl showed that the two KASP markers were associated with un-identified protein or intergenic sequences. The KASP markers for both TGW and FE also showed significant phenotypic mean difference in other traits, particularly the TGW marker showed significant phenotypic mean difference in grain yield (Table S6).

## Discussion

The yield increase in the 'Green Revolution' era was mostly owing to an increase in HI; however, there has been little progress in improving HI in recent decades. The HI of currently used wheat cultivars remains at about 0.45 to 0.51 in spring wheat and 0.5–0.55 in winter wheat, which is lower than the theoretical upper limit of 0.62 [9, 34, 35]; therefore, there is still scope for improvement



Fig. 3 Principal component analysis biplot of measured traits using best linear unbiased estimates (BLUEs). FE, fruiting efficiency; GPI, glume partitioning index; LmPI, lemma Partitioning index; API, awn partitioning index; PPI, palea partitioning index; FPI, floret partitioning index; RPI, rachis partitioning index. E2\_C18, E2\_C19, E2\_Q18, and E2\_Combined refers to Citra 2018, Citra 2019, Quincy 2018 and combined dataset respectively for experiment II

of HI in wheat and its stabilization across seasons and environments. Some recent studies have demonstrated that increased biomass had a positive association with grain yield, but a negative association with HI [34, 36]. To improve HI through converting increased biomass towards grain yield, it is essential to study novel partitioning traits that enable discrimination between "useful" and "non-useful" biomass [5]. The number of grains (sink strength) is a major trait for GY and HI improvement. There is strong evidence that during grain filling, wheat yield (grain weight per spike and spikes per unit area) is sink limited because carbon accumulation is limited by the total storage capacity of the grains. Therefore, improving the grain number per unit area is an important target in the genetic improvement of HI and yield potential. Additionally, grain number has more plasticity than grain weight and, therefore, is more responsive to genetic and environmental changes than grain weight [15, 37]. Understanding the genetic mechanisms of component traits that control grain number is pertinent for further improvement of GY and HI. Fruiting efficiency has been proposed as one of the novel promising traits that could be exploited to further increase the GN and HI [24].

The current study found a significant variation of FE in the US facultative soft wheat, which indicated the scope for improvement of this particular trait through breeding. This result agrees with several other studies [7, 19, 24, 38, 39]. Positive associations of wheat FE with GN,



**Fig. 4** Pearson's correlation coefficient between phenotypic traits using best linear unbiased estimates (Experiment II). FE, fruiting efficiency in grains g<sup>-1</sup> of spike dry weight at anthesis + 7 days; GPI, glume partitioning index; LmPI, lemma Partitioning index; API, awn partitioning index; PPI, palea partitioning index; FPI, floret partitioning index; RPI, rachis partitioning index. E2\_C18, E2\_C19, E2\_Q18, E2\_Combined refers to Citra 2018, Citra 2019, Quincy 2018 and combined dataset respectively for the experiment II

HI, and GY which have been observed in the current study and other recent studies [19, 24] were supported by PCA analysis that clustered these traits together. The positive physiological relationship between GN and FE has been reported [7, 8, 15, 24], with fruiting efficiency measuring the efficiency with which resources are used to set grains [7]. Higher FE indicates a declined abortion rate and increased survival of floret primordia during the stem elongation phase before anthesis [7, 15, 24], which eventually increases the likelihood of fertile florets to set grains [15, 37]. The low heritability of FE in our study suggested the complex nature and high environmental influence. The heritability estimate can be improved by including more replications, consistent environments,

and precise measurement. However, various studies have demonstrated reasonable heritability [39–41] and a positive response to the selection of the FE trait [7]. Additionally, positive transgressive segregation reported for FE [7, 8, 39] suggested its applicability in further improvement of GN through selecting genotypes for higher FE. Previous studies have demonstrated that HI could be improved by (a) increasing SPI; and (b) reducing dry matter partitioning towards competing plant parts specifically internode 2 and/or internode 3 [20]. Reducing second or third internode length to increase HI might have to involve height reduction, however, further reducing already optimum wheat plant height (0.7–1 m) may penalize wheat yield [15, 16]. This could limit the extensive use of stem



Significant SNP Distribution

Fig. 5 Marker-trait associations across three genomes in US soft wheat association panel

partitioning traits up to a specific point. Thus, improving FE could be a pragmatic way to increase GN and HI through the improvement of the spike partitioning index and the internode partitioning index in wheat. Pyramiding QTLs linked to these partitioning traits [33] in high FE genotypes could further boost harvest index and grain yield.

Utilization of FE in the breeding program may need to consider two potential trade-offs: spike weight at anthesis and grain weight [24, 39]. A negative association has been reported between FE and SPI in some studies [7, 18, 26, 27]. The negative correlation between SPI and FE suggests that selection for high FE may reduce SPI. However, different authors suggested the possibility of identifying genotypes with high SPI and FE [5, 7, 19, 20]. Marker-assisted selection may facilitate combining the genes that control these traits. In the current study, a negative relationship between TGW and FE suggests an increase in FE resulted in a decrease in grain weight, consistent with most previous studies [15, 16, 19, 24, 28, 39] except that González et al. (2014) [22] did not find

such an association. A clear trade-off occurs only if grain weight at each floral position is reduced as a result of an increase in FE, which would reduce GY [7, 42]. However, some studies suggested that the reduction in grain weight owing to increased FE usually occurred at distal smaller florets that were infertile in low FE genotypes and the proximal florets remained unaffected, thereby the reduction in the average grain weight is not a trade-off, but due to increased fertile florets [7, 10, 15, 39]. Improving vascular connections within the rachilla can reduce resistance and improve floret fertility of more distal florets, therefore, minimizing such trade-offs [36, 43].

Among 13 MTAs for FE, five MTAs were in genes (Table 4, Table S4). These putative candidate genes, along with nearby genes (Table S5), can be used to develop selectable markers for further improvement of FE. Since the heritability of FE was found to be relatively low, marker-assisted selection would play an important role in the improvement of this trait. Out of 5 MTAs identified for TGW, 4 MTAs were found within genes (Table 4, Table S4). Previous studies identified QTLs/



Fig. 6 Overview of significant markers trait associations identified on each chromosome for phenotypic traits obtained from GWAS. FE, fruiting efficiency; TGW, thousand grain weight

 $\label{eq:stable} \begin{array}{l} \textbf{Table 4} & \textbf{Summary of significant marker-trait associations for FE} \\ \textbf{and TGW} \end{array}$ 

Traits	SNPs	Chromosomes	PVE
FE	13	1A, 2A, 2B, 3A, 3D, 5D, 6A, 6D, 7A, 7D	10.34–17.56
TGW	5	2A, 2B, 5B, 6D	9.62–14.92

FE fruiting efficiency, TGW thousand grain weight

MTAs responsible for TGW on 2A [44–46], 2B [33, 46], 3A [33, 47, 48], and 6D [46]. An MTA S6D\_24028817 was in a gene *TraesCS6D02G049500* which is annotated as an Oligopeptide transporter protein (Table S4). Oligopeptide transporters (OPTs) are membrane-localized proteins with various biological processes, such as substrate transportation [49]. An MTA S2B\_644162831 found within gene *TraesCS2B02G450900* was annotated as a Nucleoporin protein (Table S4). Nucleoporin proteins have different functions such as nucleocytoplasmic



Fig. 7 Boxplot representing KASP markers showed significant allelic effect in the diversity population. The KASP markers were identified using a studentized t-test (*P*-value < 0.05) in validation panel

transport, cell differentiation, cell signaling, and gene expression [50–52]. Likewise, an MTA S6D\_11346685 was found in a gene *TraesCS1B02G401200* which annotates for the F-box family protein. This protein plays a role in the developmental and physiological processes such as spike development, and pollen recognition in plants [53, 54]. These genes, along with those in close proximity to significant SNPs (Table S5) and their associated functions denote their potential for further manipulation to improve TGW. In addition, we found significant SNPs associated with FE which had an allelic effect in the same direction on other traits along with FE.

Negative associations were observed between FE and intra-spike partitioning traits including awn, glume, sterile floret, and rachis partitioning. These non-grain sinks compete with florets for assimilates, resulting in lower FE. The negative relation between RPI and FE could be taken as an example of such competition. Both GPI and sterile FPI have a significant negative correlation with FE. Although the physiological basis for that relationship is not clear, high negative indirect effects through rachis PI seem to be a driving force in both cases. The highly negative direct effect of awn PI on FE shows that the formation of awns and fertile florets competes for the same source of assimilates in a spike, and investment of dry mater on awns ultimately reduces FE. Awns can increase spike surface area by up to 50% [55], and contribute to total photosynthesis [56, 57]; whereas awnletted wheat may produce significantly more grains, therefore, awn elimination could potentially enhance grain numbers [24, 58–60]. Since grain yield in wheat is mostly sink limited during grain filling, photosynthetic contribution by awns towards grain yield/number might not be enough to counterbalance the extent of assimilate investment towards their production and physiological maintenance.

Higher FE could be achieved through a preferential allocation of spike DM to the florets with a proportional reduction in partitioning to non-productive sinks or other structural parts (rachis, glumes, awns). It can be assumed that competition for assimilates leads to the shortage of assimilates, resulting in floret primordia abortion [61]. Higher availability of assimilates to florets improves the probability to turn labile floret primordia into fertile florets, which results in increased FE and GN [24]. In addition, increasing spike sugar content [19, 27, 62] and modifying plant signaling responses [17, 19, 20] also increase FE and GN. In summary, FE can be increased through decreasing intra-spike dry matter partitioning to awn, rachis, and glume, and potentially increasing partitioning to the lemma and palea within the spikelet morphological components. This optimized partitioning would eventually have a positive impact in reaching the theoretical HI limit of above 60%. Two KASP markers, one significantly contributing to TGW and the other to FE, could be used for future breeding programs intended to select elite lines for grain partitioning traits. Additional research is needed to fine mapping region encompassing these KASP markers to determine the underlying genes and their subsequent functions.

# Conclusions

Results showed that FE is strongly associated with GN, HI, and grain yield. Increasing FE would most likely contribute to higher GN. Reducing glume, awn, and rachis partitioning and increasing partitioning of lemma and palea would further increase FE and GN. GWAS analysis suggested that the genetic control of FE is complex and controlled by several putative QTL regions. GWAS analysis successfully identified 18 markers significantly associated with FE and TGW. Nine MTAs for FE and TGW were identified to be within genes which are potential targets of selection in marker-assisted breeding to capitalize genetic variation for the trait. KASP marker development and validation would facilitate transfer of these QTL into adapted cultivars through a marker-assisted breeding to improve GY and HI.

#### Methods

# Trait evaluation in field experiments

Three experiments were conducted in two locations in northern and north-central Florida. The genetic materials and environments involved in the first experiment (Experiment I) have been described in Shahi et. al (2022) [32]. In brief, a panel of 236 facultative soft wheat elite lines and varieties were planted in the Plant Science Research and Education Unit (PSREU) in Citra, Florida in 2017–2018, and in the North Florida Research and Education Center (NFREC) in Quincy in 2016–2017 and 2017–2018 using an augmented design with three repeated checks, SS8641 (PI 674197), AGS2000 (PI 656845), and Jamestown (PI 653731) to control spatial variability. These check cultivars are popular commercial soft wheat varieties grown throughout the southern and south-eastern US. The checks were replicated twice in each block. The weather conditions of experiments are provided in Table S7.

Measurement of traits HI, GY, GN, and FE has already been described in Shahi et al. (2022) [32]. Days to anthesis was taken for each plot as the days from planting to when 50% of plants flowered [63]. At 7 days after anthesis (A+7d, Zadoks scale: GS70), wheat tillers were cut at ground level from a  $0.25 \text{ m}^2$  area of each plot. The sample was oven-dried at 60 °C for 72 h and then dried biomass was weighted and converted to  $g m^{-2}$ . Spikes were separated from stem and leaves, weighed, and converted to spike dry matter  $m^{-2}$  at A+7d. Traits such as GN, GY, TGW, and HI were recorded at physiological maturity (Zadoks scale: GS90). Days to physiological maturity were taken for each plot when 50% of the peduncle turned yellow. At physiological maturity, plants from a 0.25 m<sup>2</sup> plot area were harvested and threshed. Grain in each plot was weighed and converted to  $g m^{-2}$ . Thousand-grain weight (TGW) was obtained by weighing 1,000 kernels counted in a seed counter (Seedburo Equipment Co., Chicago, IL). GN was calculated by dividing grain yield  $m^{-2}$  by mean grain weight (TGW/1000). HI was measured as the ratio of grain yield  $m^{-2}$  to above-ground dry biomass  $m^{-2}$ . Likewise, GY was measured in kg ha<sup>-1</sup> as the total grain weight per plot divided by the plot area adjusted to 12% moisture. FE was calculated as a ratio of GN  $\mathrm{m}^{-2}$  at maturity to spike dry weight  $m^{-2}$  at A + 7d and expressed as grain numbers per gram of spike dry weight at A + 7d.

A sub-set of 10 genotypes from the panel used in Experiment I were evaluated for non-grain spike partitioning (Experiment II) at PSREU, Citra, Florida in the 2017-2018 and 2018-2019 growing seasons, and at NFREC, Quincy, Florida in 2017-2018 growing season (similar seed rate and plot size). The experiments used a randomized complete block design (RCBD) with 3 replications and the same agronomic and management practices as described in Experiment I. Seven spikes were randomly selected from each plot at harvest and separated into different spike parts (awns, rachis, glume, lemma, palea, and infertile floret). The dry weight (DW) of total spikes and each of the spike parts and grains were also measured to calculate the partitioning of each part. The awn partitioning index (API) was calculated as (DW of awns)/ (DW of total spike—grain DW). Similarly, rachis, glume, lemma, palea, and infertile floret partitioning indices were calculated as the DW of the respective spike part divided by the DW of the non-grain spike calculated as the difference of total spike DW and grain DW.

A validation panel consisting of two diversity populations of 178 facultative lines and 59 spring lines was planted during 2018–2019 season to validate KASP markers (Experiment III). The populations were planted at the Plant Science Research and Education Unit (PSREU) in Citra, Florida. A randomized augmented design with three repeated checks ("AGS 2000", "SS8641", and "Jamestown") was used for the experiment.

#### Genotyping and KASP marker development

Detailed description of the methodologies for genotyping have been provided previously [32]. In brief, high-quality DNA was extracted from freeze-dried, powdered leaf tissue obtained from 2-week-old plants using modified CTAB (cetyltrimethylammonium bromide) protocol and genotyping-by-sequencing (GBS) library was constructed using two restriction enzymes, *MspI* and *PstI*-HF [64]. SNPs were called using the TASSEL v5.0 GBS v2.0 discovery pipeline [65] and the IWGSC RefSeq v1.1 reference genome [66].

Significant SNPs identified from GWAS were selected for the development of KASP markers. Primers for KASP were designed based on the sequences flanking the SNPs in the Chinese Spring reference genome RefSeq v1.0 (IWGSC 2018) using PolyMarker (http://www.polym arker.info/), which considers the polyploid nature of multiple homologs. Designed primer sequence was searched in entire genome to remove possible duplication and only the markers that showed clear cluster separation between two alleles were selected. The KASP assays were carried out in a GeneAmp<sup>™</sup> PCR System 9700 Fast Thermal Cycler at USDA-ARS Central Small Grain Genotyping Lab, Manhattan, Kansas. A 5 µL of KASP PCR mix composed of 1.94 µL of 2×PACE Genotyping Master Mix (https://3crbio. com/), 0.06  $\mu$ L primer mix, and 3  $\mu$ L DNA at 10–20 ng/ $\mu$ l. Amplification of PCR began with a denaturation of 94 °C for 15 min, followed by 10 touch-down PCR cycles at 94 °C for 20 s and 67 °C for 60 s with -1.0 °C/cycle, and then 32 cycles at 94 °C for 20 s and 57 °C for 60 s.

#### Phenotypic data analysis

Analysis of variance (ANOVA) was performed using the "lme4" package [67] in R software (v3.5.1, R Development Core Team). Best linear unbiased estimates (BLUEs) were calculated for each environment separately as well as across all environments. In the analysis, the genotypic effects were considered fixed, while all the other effects were regarded as random. All traits were adjusted using days to anthesis as a covariate. The following model applied to individual environments for experiment I and III.

 $Y_{ijkl} = \mu + B_i + ID_j + G_k + C_l + \epsilon_{ijkl}$ 

For combined data across environments, the following model was used.

where Y is the trait of interest;  $\mu$  is the effect of the mean; G corresponds to the un-replicated genotypes; C is the effect of the replicated checks on each block;  $E_i$  is the effect of the *i*th environment, and I is the effect of the identifier of the checks.  $I \times E_i$ ,  $G \times E_i$ , and  $C \times E_i$  are the effects of the check identifier by environment, genotype by environment, and check by environment interactions, respectively;  $B_k(E_i)$  is the block effect nested within each environment; and  $\varepsilon$  is the standard normal error [68].

To estimate BLUEs for experiment II (RCBD), the following model was used for individual environments.

$$Y_{ijk} = \mu + +Gj + B_k + \varepsilon_{ijk}$$

The following model was used for combined environments.

$$Y_{ijk} = \mu + E_i + G_j + G \times E_i + B_k(E_i) + \varepsilon_{ijk}$$

where the phenotypic response  $(Y_{ijk})$  is a function of the overall mean  $(\mu)$ , j<sup>th</sup> genotype  $(G_i)$ , i<sup>th</sup> environment, genotype-environment interaction  $(G \times E_i)$ ,  $B_k(E_i)$  is the block effect nested within each environment, and  $\epsilon$  is the standard normal error.

Broad-sense heritability was calculated using the following formula:

$$H^{2} = \frac{\sigma^{2}G}{\sigma^{2}G + \frac{\sigma^{2}_{G \times E}}{n} + \frac{\sigma^{2}e}{nr}}$$

where  $H^2$  is the broad-sense heritability estimate,  $\sigma^2 G$  is the genetic variance,  $\sigma^2_{GXE}$  is the genotype by environmental variance,  $\sigma^2 e$  is the residual variance, n is the number of environments, and r is the number of replications.

Pearson's correlation among traits was computed using the "corrplot" package in R [69]. A PCA biplot was created with the "factoextra" package in a R [70]. Additionally, path coefficient analysis was conducted in R using the "lavaan" package [71]. In the first analysis, GN, FE, and TGW were used as predictors whereas HI was designated as a response. To estimate the direct and indirect effects of different traits on FE, glume partitioning index (GPI), lemma partitioning index (LmPI), awn partitioning index (API), palea partitioning index (PPI), floret partitioning index (FPI), and rachis partitioning index (RPI) were designated as predictors whereas FE was designated as a response.

 $Y_{iik} = \mu + I + G + C + E_i + I \times E_i + G \times E_i + C \times E_i + B_k(E_i) + \varepsilon_{iik}$ 

#### Genome-wide association study and gene annotation

The compressed mixed linear model (CMLM) of the genome association predicted integrated tool (GAPIT) in R [72] was used to identify the associations between phenotypic and genotypic data. A GWAS was performed using adjusted means (BLUEs) from the dataset collected from Citra (2017-2018) [E1\_C18], Quincy (2016-2017) [E1 Q17] and Quincy (2017-2018) [E1 Q18], and combined data (E1\_Combined) in Experiment I. Significant MTAs were claimed at a threshold of FDR (p < 0.10). We used "cmplot" package in R to generate SNP density plots, Manhattan plots, and Q-Q plots. IWGSC wheat reference genome RefSeq v1.0 [65] was used to find candidate genes associated with significant MTAs. Their annotations were determined using the Variant Effect Predictor (VEP) tool in the Ensemble Plants (http:// plants.ensembl.org/Triticum\_aestivum/ Tools/VEP).

In the validation analysis, a student's t-test was utilized. However, if non-normal residuals were observed, the kruskal–wallis test was employed. Differences in means between line groups with two or combinations of allelic types were computed. A KASP marker was deemed effective if there was significant mean difference between the contrasting allelic groups.

#### Abbreviations

ANOVA	Analysis of Variance
API	Awn partitioning index
BLUE	Best Linear Unbiased Estimation
CMLM	Compressed Mixed Linear Model
FPI	Floret partitioning index
FE	Fruiting efficiency
GAPIT	Genome Association Predicted Integrated Toc
GWAS	Genome -Wide Association Studies
GPI	Glume partitioning index
GN	Grain Number
GY	Grain Yield
HI	Harvest Index
KASP	Kompetitive allele specific PCR
LmPI	Lemma Partitioning index
LD	Linkage Disequilibrium
MTA	Marker-Trait Association
PPI	Palea partitioning index
PVE	Percent Variation Explained
QTL	Quantitative Trait Loci
RPI	Rachis partitioning index
SPI	Spike Partitioning Index
TGW	Thousand grain weight

## **Supplementary Information**

The online version contains supplementary material available at https://doi. org/10.1186/s12870-025-06072-1.

Additional file 1: Figure S1a Manhattan plot (left) and quantile-quantile plots (right) showing genome-wide SNP loci associated with fruiting efficiency (FE) ordered on C18, Q17, Q18, and Combined. The horizontal line in Manhattan plot represents the expected value with a uniform suggestive genome wide significance threshold [-FDR  $\leq$  0.10]. Figure S1b Manhattan plot (left) and quantile-quantile plots (right) showing genome-wide SNP loci associated with Thousand Grain Weight (TGW) ordered on C18, Q17, Q18, and Combined. The horizontal line in Manhattan plot

represents the expected value with a uniform suggestive genome wide significance threshold [-FDR  $\leq 0.10$ ].

Additional File 2: Table S1 Summary of ANOVA results testing the effects of genotype (G), environment (E), and genotype-by-environment interaction (GxE) in phenotypic traits evaluated in combined analysis of experiment I. The table includes mean square values and significance level of each term. Table S2 Summary of ANOVA results testing the effects of genotype (G), environment (F), and genotype-by-environment interaction ( $G \times F$ ) in phenotypic traits evaluated in combined analysis of experiment II. The table includes mean square values and significance level of each term. Table S3 Direct and indirect effects of different traits on fruiting efficiency (FE) identified through path coefficient analysis (Experiment II). Table S4 Summary of all significant markers and their functional annotations associated with two traits. Table S5 Information on genes at proximity to significant SNPs for phenotypic traits. Table S6 Significant KASP markers detected in validation study (previously identified by an association mapping study). Table S7 Weather table showing Tave (monthly average temperature) and Ppt (monthly precipitation in mm). Association mapping panel was planted two seasons in Citra (2017/ 2018) and Quincy (2016/2017, 2017/2018).

#### Acknowledgements

We extend our acknowledgement to Dr. Jerry Johnson (University of Georgia, Georgia, USA), Dr. Stephen Harrison (Louisiana State University), and Dr. Esten Mason (Colorado State University) for providing seeds of the association panel. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The USDA is an equal opportunity provider and employer.

#### **Clinical trial number**

Not applicable.

#### Author's contributions

M.A.B., M.P.R., J.F., G.M., and S.S. planned and designed the research; D.S. & J.G. performed the research and analyzed the data; D.S. wrote the paper; Z.L., G.B., P.S.A., A.B. performed genotyping, processed data, SNP calling; S.P., J.K., J.G., M.A.V., and J.M. collected data in the field; M.A.B., M.P.R., G.B., and J.F. edited the paper. All authors approved the manuscript.

#### Funding

The research was funded by NIFA-IWYP (National Institute of Food and Agriculture – International Wheat Yield Partner) Award # 2017–67007-25929 and by the Agriculture and Food Research Initiative Competitive Grant 2022–68013-36439 (WheatCAP) from the USDA National Institute of Food and Agriculture.

#### Data availability

The datasets and R code used during the current study are available in the https://doi.org/10.6084/m9.figshare.26046508.v1.

#### Declarations

#### Ethics approval and consent to participate

The plant field trial experiments in the current study complied with the relevant institutional, national, and international guidelines and legislation. The appropriate permissions and/or licenses for collection of plant or seed specimens were obtained.

#### **Consent for publication**

Not Applicable.

#### **Competing interests**

The authors declare no competing interests.

Received: 25 July 2024 Accepted: 7 January 2025 Published online: 24 January 2025

#### References

- 1. FAO. World Food and Agriculture Statistical Yearbook 2022.Rome. 2022. https://openknowledge.fao.org/handle/20.500.14283/cc2211en
- Bhatta M, Morgounov A, Belamkar V, Wegulo SN. Genome-wide association study for multiple biotic stress resistance in synthetic hexaploid wheat. Int J Mol Sci. 2019;20(15):3667.
- Pradhan S, Babar MA, Bai G, Khan J, Shahi D, Avci M, et al. Genetic dissection of heat-responsive physiological traits to improve adaptation and increase yield potential in soft winter wheat. BMC Genomics. 2020;21(1):1–15. https://doi.org/10.1186/s12864-020-6717-7.
- Lopes MS, Reynolds MP. Stay-green in spring wheat can be determined by spectral reflectance measurements (Normalized Difference Vegetation Index) independently from phenology. J Exp Bot. 2012. https://doi.org/ 10.1093/jxb/ers071.
- Reynolds M, Foulkes J, Furbank R, Griffiths S, King J, Murchie E, et al. Achieving yield gains in wheat. Plant Cell Environ. 2012;35(10):1799–823. https://doi.org/10.1111/j.1365-3040.2012.02588.x.
- Albajes R, Cantero-Martínez C, Capell T, Christou P, Farre A, Galceran J, et al. Building bridges: An integrated strategy for sustainable food production throughout the value chain. Mol Breed. 2013. https://doi.org/10. 1007/s11032-013-9915-z.
- Slafer GA, Elia M, Savin R, García GA, Terrile II, Ferrante A, et al. Fruiting efficiency: An alternative trait to further rise wheat yield. Food Energy Secur. 2015;4(2):92–109. https://doi.org/10.1002/fes3.59.
- García GA, Serrago RA, González FG, Slafer GA, Reynolds MP, Miralles DJ. Wheat grain number: Identification of favourable physiological traits in an elite doubled-haploid population. Field Crops Res. 2014;168:126–34. https://doi.org/10.1016/j.fcr.2014.07.018.
- Foulkes MJ, Slafer GA, Davies WJ, Berry PM, Sylvester-Bradley R, Martre P, et al. Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. J Exp Bot. 2011;62(2):469–486. https://doi.org/10.1093/jxb/erg300.
- Miralles DJ, Slafer GA. Sink limitations to yield in wheat: How could it be reduced? J Agric Sci. 2007;145(2):139–49. https://doi.org/10.1017/S0021 859607006752.
- Ferrante A, Savin R, Slafer GA. Floret development of durum wheat in response to nitrogen availability. J Exp Bot. 2010;61(15):4351–9. https:// doi.org/10.1093/jxb/erq236.
- Ferrante A, Savin R, Slafer GA. Floret development and grain setting differences between modern durum wheats under contrasting nitrogen availability. J Exp Bot. 2013;64(1):169–84. https://doi.org/10.1093/jxb/ers320.
- Ferrante A, Savin R, Slafer GA. Is floret primordia death triggered by floret development in durum wheat? J Exp Bot. 2013;64(10):2859–69. https:// doi.org/10.1093/jxb/ert129.
- Dreccer MF, Wockner KB, Palta JA, McIntyre CL, Borgognone MG, Bourgault M, et al. More fertile florets and grains per spike can be achieved at higher temperature in wheat lines with high spike biomass and sugar content at booting. Funct Plant Biol. 2014;41(5):482–95. https://doi.org/ 10.1071/FP13232.
- 15. Ferrante A, Savin R, Slafer GA. Relationship between fruiting efficiency and grain weight in durum wheat. Field Crops Res. 2015;177:109–16. https://doi.org/10.1016/j.fcr.2015.03.009.
- Terrile II, Miralles DJ, González FG. Fruiting efficiency in wheat (Triticum aestivum L): Trait response to different growing conditions and its relation to spike dry weight at anthesis and grain weight at harvest. Field Crops Res. 2017;201:86–96. https://doi.org/10.1016/j.fcr.2016.09.026.
- González FG, Terrile II, Falcón MO. Spike fertility and duration of stem elongation as promising traits to improve potential grain number (and yield): Variation in modern Argentinean wheats. Crop Sci. 2011;51(4):1693–702.
- Abbate PE, Pontaroli AC, Lázaro L, Gutheim F. A method of screening for spike fertility in wheat. J Agric Sci. 2013;151(3):322–330. https://doi.org/10. 1017/S0021859612000068. https://doi.org/10.2135/cropsci2010.08.0447.
- Rivera-Amado C, Trujillo-Negrellos E, Molero G, Reynolds MP, Sylvester-Bradley R, Foulkes MJ. Optimizing dry-matter partitioning for increased spike growth, grain number and harvest index in spring wheat. Field Crops Res. 2019;240:154–67. https://doi.org/10.1016/j.fcr.2019.04.016.
- Sierra-Gonzalez A, Molero G, Rivera-Amado C, Babar MA, Reynolds MP, Foulkes MJ. Exploring genetic diversity for grain partitioning traits to enhance yield in a high biomass spring wheat panel. Field Crops Res. 2021;260:107979. https://doi.org/10.1016/j.fcr.2020.107979.

- Ferrante A, Cartelle J, Savin R, Slafer GA. Yield determination, interplay between major components and yield stability in a traditional and a contemporary wheat across a wide range of environments. Field Crops Res. 2017;203:114–27. https://doi.org/10.1016/j.fcr.2016.12.028.
- González FG, Aldabe ML, Terrile II, Rondanini DP. Grain weight response to different post-flowering source:sink ratios in modern high-yielding Argentinean wheats differing in spike fruiting efficiency. Crop Sci. 2014;54(1):297–309. https://doi.org/10.2135/cropsci2013.03.0157.
- Alonso MP, Mirabella NE, Panelo JS, Cendoya MG, Pontaroli AC. Selection for high spike fertility index increases genetic progress in grain yield and stability in bread wheat. Euphytica. 2018;214:1–12. https://doi.org/10. 1007/s10681-018-2193-4.
- Garcia AL, Savin R, Slafer GA. Fruiting efficiency differences between cereal species. Field Crops Res. 2019;231:68–80. https://doi.org/10.1016/j. fcr.2018.11.005.
- Guo Z, Slafer GA, Schnurbusch T. Genotypic variation in spike fertility traits and ovary size as determinants of floret and grain survival rate in wheat. J Exp Bot. 2016;67(14):4221–30. https://doi.org/10.1093/jxb/ erw200.
- Ferrante A, Savin R, Slafer GA. Differences in yield physiology between modern, well adapted durum wheat cultivars grown under contrasting conditions. Field Crops Res. 2012;136:52–64. https://doi.org/10.1016/j.fcr. 2012.07.015.
- Dreccer MF, van Herwaarden AF, Chapman SC. Grain number and grain weight in wheat lines contrasting for stem water soluble carbohydrate concentration. Field Crops Res. 2009;112(1):43–54. https://doi.org/10. 1016/j.fcr.2009.02.006.
- Gaju O, Reynolds MP, Sparkes DL, Foulkes MJ. Relationships between large-spike phenotype, grain number, and yield potential in spring wheat. Crop Sci. 2009;49(3):961–973. 10.2135
- Pedro A, Savin R, Parry MAJ, Slafer GA. Selection for high grain number per unit stem length through four generations from mutants in a durum wheat population to increase yields of individual plants and crops. Field Crops Res. 2012;129:59–70. https://doi.org/10.1016/j.fcr.2012.01.016.
- Kaler AS, Purcell LC. Estimation of a significance threshold for genomewide association studies. BMC Genomics. 2019;20:1–8. https://doi.org/10. 1186/s12864-019-5992-7.
- Thomson MJ. High-throughput SNP genotyping to accelerate crop improvement. Plant Breed Biotechnol. 2014;3:195–212. https://doi.org/10. 9787/pbb.2014.2.3.195.
- Shahi D, Guo J, Pradhan S, Khan J, Avci M, Khan N, et al. Multi-trait genomic prediction using in-season physiological parameters increases prediction accuracy of complex traits in US wheat. BMC Genomics. 2022;23(1). https://doi.org/10.1186/s12864-022-08487-8.
- Shahi D, Guo J, Babar MA, Pradhan S, Khan J, et al. Deciphering the genetic basis of novel traits that discriminate useful and non-useful biomass to enhance harvest index in wheat. The plant genome. 2024;e20512, https://doi.org/10.1002/tpg2.20512.
- Pradhan S, Babar MA, Robbins K, Bai G, Mason RE, Khan J, et al. Understanding the genetic basis of spike fertility to improve grain number, harvest index, and grain yield in wheat under high temperature stress environments. Front Plant Sci. 2019;10:1–13. https://doi.org/10.3389/fpls. 2019.01481.
- Reynolds M, Foulkes MJ, Slafer GA, Berry P, Parry MAJ, Snape JW, Angus WJ. Raising yield potential in wheat. J Exp Bot. 2009;60(7):1899–918. https://doi.org/10.1093/jxb/erp016.
- Aisawi KAB, Reynolds MP, Singh RP, Foulkes MJ. The physiological basis of the genetic progress in yield potential of CIMMYT spring wheat cultivars from 1966 to 2009. Crop Sci. 2015;55(4):1749–64. https://doi.org/10.2135/ cropsci2014.09.0601.
- Slafer GA, Savin R, Sadras VO. Coarse and fine regulation of wheat yield components in response to genotype and environment. Field Crops Res. 2014;157:71–83. https://doi.org/10.1016/j.fcr.2013.12.004.
- Gonzalez-Navarro OE, Griffiths S, Molero G, Reynolds MP, Slafer GA. Variation in developmental patterns among elite wheat lines and relationships with yield, yield components and spike fertility. Field Crops Res. 2016;196:294–304. https://doi.org/10.1016/j.fcr.2016.07.019.
- Elía M, Savin R, Slafer GA. Fruiting efficiency in wheat: physiological aspects and genetic variation among modern cultivars. Field Crops Res. 2016;191:83–90. https://doi.org/10.1016/j.fcr.2016.02.019.

- Pedró A, Savin R, Slafer GA. Crop productivity as related to single-plant traits at key phenological stages in durum wheat. Field Crops Res. 2012. https://doi.org/10.1016/j.fcr.2012.09.016.
- Martino DL, Abbate PE, Cendoya MG, Gutheim F, Mirabella NE, Pontaroli AC. Wheat spike fertility: inheritance and relationship with spike yield components in early generations. Plant Breed. 2015;134(3):264–70. https://doi.org/10.1111/pbr.12262.
- Calderini DF, Savin R, Abeledo LG, Reynolds MP, Slafer GA. The importance of the period immediately preceding anthesis for grain weight determination in wheat. Euphytica. 2001;119:199–204. https://doi.org/10.1023/A: 1017597923568.
- Bancal P, Soltani F. Source-sink partitioning. Do we need Münch? J Exp Bot. 2002;53(376):1919–1928. https://doi.org/10.1093/jxb/erf037.
- Li F, Wen W, Liu J, Zhang Y, Cao S, He Z, et al. Genetic architecture of grain yield in bread wheat based on genome-wide association studies. BMC Biol. 2019;19:1–19. https://doi.org/10.1186/s12870-019-1781-3.
- 45. Sukumaran S, Reynolds MP, Sansaloni C. Genome-wide association analyses identify QTL hotspots for yield and component traits in durum wheat grown under yield potential, drought, and heat stress environments. Front Plant Sci. 2018;9:1–16. https://doi.org/10.3389/fpls.2018.00081.
- 46. Bhatta M, Morgounov A, Belamkar V, Baenziger PS. Genome-wide association study reveals novel genomic regions for grain yield and yieldrelated traits in drought-stressed synthetic hexaploid wheat. Int J Mol Sci. 2018;19(10):3011. https://doi.org/10.3390/ijms19103011.
- Lozada DN, Mason RE, Babar MA, Carver BF, Guedira GB, Merrill K, et al. Association mapping reveals loci associated with multiple traits that affect grain yield and adaptation in soft winter wheat. Euphytica. 2017;213:1–15. https://doi.org/10.1007/s10681-017-2005-2.
- Sun C, Zhang F, Yan X, Zhang X, Dong Z, Cui D, et al. Genome-wide association study for 13 agronomic traits reveals distribution of superior alleles in bread wheat from the Yellow and Huai Valley of China. Plant biotechnology journal.2017;15(8):953–969.
- Cao J, Huang J, Yang Y, Hu X. Analyses of the oligopeptide transporter gene family in poplar and grape. BMC Genomics. 2011;12:1–16. https:// doi.org/10.1186/1471-2164-12-465.
- Khan AU, Qu R, Ouyang J, Dai J. Role of nucleoporins and transport receptors in cell differentiation. Front Physiol. 2020;11:239. https://doi.org/10. 3389/fphys.2020.00239.
- Wälde S, Kehlenbach RH. The part and the whole: functions of nucleoporins in nucleocytoplasmic transport. Trends Cell Biol. 2010;20:461–9. https://doi.org/10.1016/j.tcb.2010.05.001.
- Chatel G, Fahrenkrog B. Dynamics and diverse functions of nuclear pore complex proteins. Nucleus. 2012;3(2):162–71. https://doi.org/10.4161/ nucl.19674.
- Lechner E, Achard P, Vansiri A, Potuschak T, Genschik P. F-box proteins everywhere. Curr Opin Plant Biol. 2006;9(6):631–8. https://doi.org/10. 1016/j.pbi.2006.09.003.
- Hong MJ, Kim DY, Kang SY, Kim DS, Kim JB, Seo YW. Wheat F-box protein recruits proteins and regulates their abundance during wheat spike development. Mol Biol Rep. 2012;39(10):9681–96. https://doi.org/10. 1007/s11033-012-1833-3.
- Blum A. Photosynthesis and transpiration in leaves and ears of wheat and barley varieties. J Exp Bot. 1985;36(3):432–40. https://doi.org/10.1093/jxb/ 36.3.432.
- Ali M, Hussain M, Khan MI, Ali Z, Zulkiffal M, Anwar J, et al. Source-sink relationship between photosynthetic organs and grain yield attributes during grain filling stage in spring wheat (Triticum aestivum). Int J Agric Biol. 2010;12(3):509–15.
- DeWitt N, Guedira M, Lauer E, Sarinelli M, Tyagi P, Fu DQ, et al. Sequencebased mapping identifies a candidate transcription repressor underlying awn suppression at the B1 locus in wheat. New Phytol. 2020;225(1):326– 39. https://doi.org/10.1111/nph.16152.
- Rebetzke GJ, Bonnett DG, Reynolds MP. Awns reduce grain number to increase grain size and harvestable yield in irrigated and rainfed spring wheat. J Exp Bot. 2016;67(9):2573–86. https://doi.org/10.1093/jxb/erw081.
- Hosseini SM, Poustini K, Siddique KHM, Palta JA. Photosynthesis of barley awns does not play a significant role in grain yield under terminal drought. Crop Pasture Sci. 2012;62(5):489–99. https://doi.org/10.1071/CP11256.
- McKenzie H. Adverse influence of awns on yield of wheat. Can J Plant Sci. 1972;52(1):81–7. https://doi.org/10.4141/cjps72-010.

- Bancal P. Early development and enlargement of wheat floret primordia suggest a role of partitioning within spike to grain set. Field Crops Res. 2009;110(1):44–53. https://doi.org/10.1016/j.fcr.2008.06.014.
- Ghiglione HO, Gonzalez FG, Serrago R, Maldonado SB, Chilcott C, Curá JA, et al. Autophagy regulated by day length determines the number of fertile florets in wheat. Plant J. 2008;55(6):1010–24. https://doi.org/10. 11111/j.1365-313X.2008.03570.x.
- 63. Zadoks JC, Chang TT, Konzak CF. A decimal code for the growth stages of cereals. Weed Res. 1974;14(6):415–21. https://doi.org/10.1111/j.1365-3180.1974.tb01084.x.
- Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, et al. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome. 2012;5(3):103–13. https://doi.org/10.3835/plantgenome2012.06.0006.
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 2007;23(19):2633–5. https://doi.org/10.1093/ bioinformatics/btm308.
- Appels R, Eversole K, Feuillet C, Keller B, Rogers J, Stein N, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science. 2018;361(6403). https://doi.org/10.1126/science.aar7191.
- Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using Ime4. J Stat Softw. 2015;67(1):48. https://doi.org/10.18637/ jss.v067.i01.
- Lozada DN, Ward BP, Carter AH. Gains through selection for grain yield in a winter wheat breeding program. PLoS ONE. 2020;15(4):e0221603. https://doi.org/10.1371/journal.pone.0221603.
- Wei T. Package "corrplot" for R: visualization of a correlation matrix. CRAN. 2017;56:e24.
- Kassambara A, Mundt F. Factoextra: extract and visualize the results of multivariate data analyses. URL http://www.sthda.com/english/rpkgs/ factoextra BugReports. 2017.
- Rosseel Y. lavaan: an R package for structural equation modeling. J Stat Softw. 2012;48:1–36. https://doi.org/10.18637/jss.v048.i02.
- Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, et al. GAPIT: genome association and prediction integrated tool. Bioinformatics. 2012;28(18):2397–9. https://doi.org/10.1093/bioinformatics/bts444.

#### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.