Title

Identification, Characterization, and Deployment of an FT-A2 Allele for Increased Grain Number and of its Interactor bZIPC1 in Wheat (Triticum L.)

Author

Glenn, Priscilla Denise

Publication Date

2022

Peer reviewed|Thesis/dissertation

Identification, Characterization, and Deployment of an FT-A2 Allele for Increased Grain Number and of its Interactor bZIPC1 in Wheat (*Triticum* L.)

By

PRISCILLA DENISE GLENN

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Horticulture and Agronomy

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Jorge Dubcovsky, Chair

Charlie Brummer

Neelima Sinha

Allen Van Deynze

Committee in Charge

2022

TABLE OF CONTENTS

ABSTRACT

Identification, characterization and deployment of an FT-A2 allele for increased grain number and of its interactor bZIPC1 in wheat (*Triticum* L.*)*

Chapter one of this thesis describes the discovery and validation of a natural polymorphism in the A genome copy of the *FLOWERING LOCUS T2* (*FT2)*, a gene which increases spikelet number per spike (SNS) in wheat. A previous study indicated that loss-of-function mutations in *FT2* increased SNS, but also decreased fertility, negating the positive effects on grain yield. Fortunately, we found a natural amino acid change in the FT-A2 protein at position 10 from aspartic acid (D) to alanine (A) (henceforth, D10A), that was associated with increased SNS and no negative effects on fertility. We used a high-density genetic map to delimit the candidate gene region to 28 genes, and then determined that among them only FT-A2 had a non-synonymous polymorphism (D10A) consistent in all studied mapping populations. We concluded that *FT-A2* was likely the cause for the increased SNS and validated its effect in a hexaploid spring and winter wheat population.

Investigating the frequency of the A10 allele, we found that while it was present in 56% of the common wheat accessions, it was present in less than 1% of the durum accessions analyzed. This rapid increase from durum wheat to common wheat suggests that the A10 allele had undergone positive selection in common wheat and may be a useful target for improving grain yield in durum and common wheat.

Chapter two of this thesis investigates the effects of the introgression of the *FT-A2* A10 allele into four high-yielding durum wheat varieties. We found that all lines with the A10 allele had significantly higher SNS with no significant decreases in fertility. However, due to variety x *FT-A2* interactions, grain number per spike (GNS) was significantly increased in only three of

the four varieties. Unfortunately, the increases in SNS and GNS were offset by a significant decrease in kernel weight in all four varieties, resulting in no significant differences in spike yield or grain yield per plot. There was a significant negative correlation between GNS and thousand kernel weight (r = -0.42, $P = 0.0006$), which we hypothesize reflects source limitations in our environment that resulted in incomplete filling of the extra grains. Incorporation of the *FT-A2* A10 allele into high bio-mass genotypes which are less source-limited and the evaluation of sister isogenic lines in different environments will be necessary to test the usefulness of this allele in durum breeding programs.

Chapter three of this thesis describes the identification and characterization of *bZIPC1*, a bZIP-containing transcription factor from the C-subfamily, which we identified utilizing a yeast-two hybrid screen with the FT2 protein as bait. Combined loss-of-functions mutations in *bZIPC-A1* and *bZIPC-B1* (*bzipc1*) in tetraploid wheat resulted in drastic reduction in SNS with a limited effect on heading date. Expression studies revealed that genes previously known to affect SNS were not significantly affected by the *bzipc1* mutations, suggesting that *bZIPC1* may affect SNS through a different pathway. Investigating the natural variation in the *bZIPC-B1* region, revealed three major haplotypes (H1-H3), with the H1 haplotype showing significantly higher SNS, GNS, and spike yield than both the H2 and H3 haplotypes, as well as an increased frequency from the ancestral tetraploids to the modern durum and common wheat varieties. We developed markers of the two non-synonymous SNPs that differentiate the H1 haplotype from H2 and H3 so the H1 haplotype can be identified, introgressed, and deployed in wheat breeding programs.

ACKNOWLEDGMENTS

First off, my most heartfelt gratitude goes out to all those throughout these years who have believed in, supported, and inspired me to pursue and accomplish my doctorate. Without their collective belief, this incredible journey would have been drastically altered.

I want to give special thanks to my professor, Dr. Jorge Dubcovsky. I am ever so grateful he welcomed me into his lab. Under his mentorship and guidance, my scientific skills and writing have flourished. In addition, my participation in the WheatCAP under his leadership has provided ample opportunities for professional development. I also thank Dr. Daniel Woods for his co-mentorship during the final years of my dissertation as we discovered the *bZIPC1* gene. My wet-lab skillset and critical thinking abilities expanded exponentially thanks to his tutelage. A special shout out as well to my previous mentor, Dr. Jeffrey Demuth (UT Arlington), who accepted a young and hopeful plant geneticist into his beetle *Tribolium* lab and encouraged me to visit UC Davis.

I am also grateful to the members of the Dubcovsky lab, who were always willing to discuss and troubleshoot ideas. In particular, I want to thank Saarah Kuzay for her mentorship during my first years at UC Davis. She welcomed me into the lab with open arms and was always willing to guide by example. In addition, I cannot forget about Joshua Hegarty and his assistance in the field and the many wonderful trips to Tulelake, CA or my birthday buddy, Stephen Bolus, who brought the lab to life. I am also greatly appreciative of my many assistants throughout the years who went above and beyond to assist me in my studies.

Finally, I would like to thank my family and friends. To my Davis friends and roommates, who made sure I didn't stay inside all day every day. I will miss you all as I move onto the next chapter of my life. To the Davis Newman Center and all those who passed through

its doors, you gave me a home away from home I could always return to for which I am forever grateful for. A heartfelt thanks to Brian Doupnik, who reminded me to eat, breath, and take breaks during my dissertation writing. A loving thanks to my siblings, Nicole, Preston, and Julia Glenn who made me laugh and always have my back, no matter how far away they are. And last but not least, an enormous thank you to my parents, Jeff and Theresa Glenn who raised me by example to always do my best and to live in a manner pleasing to the Lord. Thank you for your love and support and for always calling me your baby girl. I am proud to call myself your daughter.

INTRODUCTION

Wheat domestication began approximately 10,000 years ago, with the domestication of wild einkorn and emmer wheat in the Diyarbakir region in southeastern Turkey (Dubcovsky & Dvorak, 2007). Gene exchanges between the northern domesticated emmer and the southern emmer then resulted in a secondary center of domesticated emmer in southern Levan. Northeast expansion of this domesticated emmer resulted in sympatry with *Ae. tauschii*, originating common hexaploid wheat. During this process of domestication, the BRITTLE RACHIS allele (*Br)* was replaced with non-brittle rachis allele (*br)* and the hulled grains were replaced with naked grains that facilitate threshing (via recessive mutations at the *TENACIOUS GLUME (tg)* loci and a dominant mutation at the *Q* locus) (Dubcovsky & Dvorak, 2007) . Since then, wheat has spread across the globe and is produced in a variety of climates, markets, and production systems with breeders selecting upon both quality and yield.

This thesis focuses on improving grain yield, as current gain rates (0.9% per year) need to be increased to match the projected human population growth by 2050 (Ray et al., 2013). Wheat contributes around 20% of the protein and calories consumed worldwide, with demand for wheat expected to increase by 50% (wheat.org) in the next decades. If breeders are going to meet this expectation, genetic gains need to be accelerated. The identification, characterization, and deployment of alleles contributing to the improvement of different grain yield components can assist in increasing wheat grain yield potential in diverse environments.

This task is complicated though by complex gene networks controlling yield and influences from the environment (Reynolds et al., 2012). Grain yield has a lower heritability than its comprised traits (Fethi & Gazzah Mohamed, 2010), so a systematic approach to increase our understanding of grain yield is to dissect the gene networks regulating each of the individual

yield components. Yield components include traits such as spikes per plant (tiller number per plant), spikelet number per spike (SNS), kernel weight, grain number per spikelet (fertility), and plants per m^2, which together, are equivalent to the total yield (kg/ha).

SNS is determined during a short period in the early stages of plant reproductive development and is not influenced by later environmental variability, resulting in high heritability ($h^2 > 0.8$), which is why SNS is the yield component trai of focus for this thesis. The spikelets are arranged on the spike's main axis as two opposite rows and the final number is determined with the initiation of the terminal spikelet at the tip of the spike at developmental stage Zadok 3.0 (Zadoks et al., 1974, Pinthus & Millet, 1978). Wheat MADS box genes of the Squamosa class (*VERNALIZATION1 (VRN1), FRUITFULL2 (FUL2)* and *FRUITFULL3 (FUL3*) have been shown to be essential for this transition and for the formation of spikelets in general (Li et al., 2019). In addition, wheat MADS-Box genes from the SVP clade were also found to have a significant effect on SNS (Li et al., 2021).

The spikelet meristem is indeterminate and produces a variable number of florets (Hanif & Langer, 1972), joined alternatively on opposite sides of the axis and encompassed by two small bracts called glumes. The fertile florets within each spikelet then develop and determine the total number of grains produced by each spikelet affecting the final GNS (Guo & Schnurbusch, 2015; Screenivasulu & Schnurbush T, 2012). Early in development, a wheat spikelet produces up to 12 floret primordia (Guo et al., 2016; Guo & Schnurbusch, 2015). However, over 70% of the florets will abort throughout spikelet development (Guo et al., 2016; Guo & Schnurbusch, 2015) leaving either 2-3 grains/spikelet in durum or 3-5 grains/spikelet in common wheat (Shitsukawa et al., 2009).

Previous studies have indicated that in some cases, such as with *ft2*-null mutants, increases in SNS can be associated with floret fertility, negating any positive effects on GNS (Shaw et al., 2019). However, when fertility remains constant, increases in SNS are positively correlated with GNS as with the *FT-A2* A10 allele (Glenn et al., 2022). In addition, it has been recently shown that variation at the *GRAIN NUMBER INCREASE 1* (*GNI1*) gene can be used to increase fertility in wheat and barley (Sakuma et al., 2019). Mutations that reduced the activity of this homeodomain leucin zipper class I (*HD-ZIP I*) gene have been shown to contribute to the increased number of fertile florets per spikelet. Thus, by increasing SNS and maintaining or increasing fertility one can potentially increase GNS (Sakuma et al., 2019).

Increasing GNS comes with its own caveats, as GNS often shows a negative correlation with kernel weight, limiting the translation of the effects of increases in GNS to increases in grain yield, as was the case with lines carrying *da1* mutants (Mora-Ramirez et al., 2021). One possible strategy to reduce this negative correlation would be to intercross lines with high GNS with lines with high biomass, that have the resources to fill the extra grains (Bustos et al., 2013; Kuzay et al., 2019, 2022; Simmonds et al., 2016).

Thesis Goals and Rationale

Overall, my thesis work focuses on the identification, characterization, and deployment of genes and their alleles which impact SNS. Previous studies revealed that *ft2*-null mutants increased SNS, but deployment was hindered by its negative impact on fertility. Given that breeders select for beneficial alleles, we explored the natural variation in *FT2* and identified the *FT-A2* D10A polymorphism that showed a rapid increase in frequency in commercial hexaploid wheat varieties. From there, we initiated a fine-mapping project to determine and narrow the

candidate region for QTLs. We also conducted a yeast-two hybrid screen to identify other genes that encode proteins that interact with FT2 and impact SNS.

As part of my effort to increase our understanding of the SNS pathway and introgress alleles for improved SNS, I identified and validated the effect of the natural D10A polymorphism in *FT-A2* on SNS (Chapter 1, presented in the original published format), the effect of the A10 allele introgressed into high-yielding varieties on yield and yield-component traits (Chapter 2), and identified and characterized a novel SNS gene, *bZIPC1* that encodes a protein that interacts with FT2 (Chapter 3). I believe that this work will help elevate our understanding of the genetic pathway impacting SNS and accelerate genetic gains to improve wheat grain yield potential.

INTRODUCTION REFERENCES

Bustos, D. v., Hasan, A. K., Reynolds, M. P., & Calderini, D. F. (2013). Combining high grain number and weight through a DH-population to improve grain yield potential of wheat in high-yielding environments. *Field Crops Research*, *145*, 106–115.

Dubcovsky, J., & Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, *316*, 1862–1866.

Fethi, B., & Gazzah Mohamed, E. (2010). Epistasis and genotype-by-environment interaction of grain yield related traits in durum wheat. *Journal of Plant Breeding and Crop Science*, *2*, 24–29.

Glenn, P., Zhang, J., Brown-Guedira, G., Dewitt, N., Cook, J. P., Li, K., Akhunov, E., & Dubcovsky, J. (2022). Identification and characterization of a natural polymorphism in *FT-A2* associated with increased number of grains per spike in wheat. *Theoretical and Applied Genetics*, *135*, 679–692.

Guo, Z., & Schnurbusch, T. (2015). Variation of floret fertility in hexaploid wheat revealed by tiller removal. *Journal of Experimental Botany*, *66*, 5945–5958.

Guo, Z., Slafer, G. A., & Schnurbusch, T. (2016). Genotypic variation in spike fertility traits and ovary size as determinants of floret and grain survival rate in wheat. *Journal of Experimental Botany*, *67*, 4221–4230.

Hanif, M., & Langer, R. H. M. (1972). The vascular system of the spikelet in wheat (*Triticum aestivum*). *Annals of Botany*, *36*, 721–727.

Kuzay, S., Lin, H., Li, C., Chen, S., Woods, D. P., Zhang, J., Lan, T., von Korff, M., & Dubcovsky, J. (2022). *WAPO-A1* is the causal gene of the 7AL QTL for spikelet number per spike in wheat. *PLOS Genetics*, *18*, e1009747.

Kuzay, S., Xu, Y., Zhang, J., Katz, A., Pearce, S., Su, Z., & Fraser, M. (2019). Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theoretical and Applied Genetics*, *132*, 2689–2705.

Li, C., Lin, H., Chen, A., Lau, M., Jernstedt, J., & Dubcovsky, J. (2019). Wheat *VRN1, FUL2* and *FUL3* play critical and redundant roles in spikelet development and spike determinacy. *Development*, *146*, dev1753981.

Li, K., Debernardi, J. M., Li, C., Lin, H., Zhang, C., Jernstedt, J., von Korff, M., Zhong, J., & Dubcovsky, J. (2021). Interactions between SQUAMOSA and SHORT VEGETATIVE PHASE MADS-box proteins regulate meristem transitions during wheat spike development. *The Plant Cell*, *33*, 3621–3644.

Mora-Ramirez, I., Weichert, H., von Wirén, N., Frohberg, C., de Bodt, S., Schmidt, R. C., & Weber, H. (2021). The *da1* mutation in wheat increases grain size under ambient and elevated CO2 but not grain yield due to trade-off between grain size and grain number. *Plant-Environment Interactions*, *2*, 61–73.

Pinthus, M. J., & Millet, E. (1978). Interactions Among Number of Spikelets, Number of Grains and Grain Weight in the Spikes of Wheat (Triticum aestivum L.). *Annals of Botany*, *42*, 839–848.

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE*, *8*, 66428.

Reynolds, M., Foulkes, J., Furbank, R., Griffiths, S., King, J., Murchie, E., Parry, M., & Slafer, G. (2012). Achieving yield gains in wheat. *Plant, Cell, & Environment*, *35*, 1799–1823.

Sakuma, S., Golan, G., Guo, Z., Ogawa, T., Tagiri, A., Sugimoto, K., Bernhardt, N., Brassac, J., Mascher, M., Hensel, G., Ohnishi, S., Jinno, H., Yamashita, Y., Ayalon, I., Peleg, Z., Schnurbusch, T., & Komatsuda, T. (2019). Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proceedings of the National Academy of Sciences*, *116*, 5182–5187.

Screenivasulu, N., & Schnurbush T. (2012). A genetic playground for enhancing grain number in cereals. *Trends in Plant Science*, *17*, 91–101.

Shaw, L. M., Lyu, B., Turner, R., Li, C., Chen, F., Han, X., Fu, D., & Dubcovsky, J. (2019). *FLOWERING LOCUS T2* regulates spike development and fertility in temperate cereals. *Journal of Experimental Botany*, *70*, 193–204.

Shitsukawa, N., Kinjo, H., Takumi, S., & Murai, K. (2009). Heterochronic development of the floret meristem determines grain number per spikelet in diploid, tetraploid and hexaploid wheats. *Annals of Botany*, *104*, 243–251.

Simmonds, J., Scott, P., Brinton, J., Mestre, T. C., Bush, M., del Blanco, A., Dubcovsky, J., & Uauy, C. (2016). A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theoretical and Applied Genetics*, *129*, 1099–1112.

Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, *14*, 415–421.

CHAPTER 1

Identification and characterization of a natural polymorphism in *FT-A2* associated with increased number of grains per spike in wheat

Theoretical and Applied Genetics (2022) Vol 135, p. 679-692

Theoretical and Applied Genetics (2022) 135:679–692
https://doi.org/10.1007/s00122-021-03992-y

ORIGINAL ARTICLE

Identification and characterization of a natural polymorphism in *FT-A2* associated with increased number of grains per spike in wheat

Priscilla Glenn[1] · Junli Zhang[1] · Gina Brown-Guedira[2] · Noah DeWitt[3] · Jason P. Cook[4] · Kun Li[1,5] · Eduard Akhunov[6] · Jorge Dubcovsky[1,5]

Received: 14 September 2021 / Accepted: 2 November 2021 / Published online: 26 November 2021

Abstract

Key message **We discovered a natural *FT-A2* allele that increases grain number per spike in both pasta and bread wheat with limited effect on heading time.**

Abstract Increases in wheat grain yield are necessary to meet future global food demands. A previous study showed that loss-of-function mutations in *FLOWERING LOCUS T2* (*FT2*) increase spikelet number per spike (SNS), an important grain yield component. However, these mutations were also associated with reduced fertility, offsetting the beneficial effect of the increases in SNS on grain number. Here, we report a natural mutation resulting in an aspartic acid to alanine change at position 10 (D10A) associated with significant increases in SNS and no negative effects on fertility. Using a high-density genetic map, we delimited the SNS candidate region to a 5.2-Mb region on chromosome 3AS including 28 genes. Among them, only *FT-A2* showed a non-synonymous polymorphism (D10A) present in two different populations segregating for the SNS QTL on chromosome arm 3AS. These results, together with the known effect of the *ft-A2* mutations on SNS, suggest that variation in *FT-A2* is the most likely cause of the observed differences in SNS. We validated the positive effects of the A10 allele on SNS, grain number, and grain yield per spike in near-isogenic tetraploid wheat lines and in an hexaploid winter wheat population. The A10 allele is present at very low frequency in durum wheat and at much higher frequency in hexaploid wheat, particularly in winter and fall-planted spring varieties. These results suggest that the *FT-A2* A10 allele may be particularly useful for improving grain yield in durum wheat and fall-planted common wheat varieties.

Introduction

Wheat is a global crop of major economic value and nutritional importance as it provides around 20% of the calories and protein consumed by the human population (http://www.fao.org/faostat/en/#data/FBS). However, with ever-changing environmental conditions and the rising human population, it is critical to increase wheat grain yield to meet future demands. Yield is a multifaceted trait that can be partitioned into several yield components, including spikes per unit of area, spikelet number per spike (SNS), grains per spikelet, and grain weight. Several genes have been identified that affect these grain yield components (Kuzay et al. 2019; Li et al. 2019; Poursarebani et al. 2015; Sakuma et al. 2019; Shaw et al. 2013; Simmonds et al. 2016; Wang et al. 2019).

Many of the genes affecting SNS also have strong effects on flowering time that can limit their use in variety development. Flowering before or after the optimum flowering time can result in yield penalties due to reduced fertility or increased risks of frost or heat damage, respectively. The

Communicated by Aimin Zhang.

✉ Jorge Dubcovsky
jdubcovsky@ucdavis.edu

[1] Department of Plant Sciences, University of California, Davis, CA 95616, USA

[2] USDA-ARS Plant Science Research, Raleigh, NC 27695, USA

[3] Department of Crop and Soil Sciences, North Carolina State University, Raleigh, NC 27695, USA

[4] Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT, USA

[5] Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

[6] Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA

vernalization gene *VRN1* is a good example of a gene affecting both flowering time and SNS. The *vrn1*-null mutant significantly increases SNS by delaying the transition of the inflorescence meristem to a terminal spikelet, but also delays the transition of the vegetative meristem to inflorescence meristem, resulting in a very late heading time (Li et al. 2019). Another good example is the main wheat photoperiod gene *PHOTOPERIOD1* (*PPD1*), which shows a strong correlation between heading date and SNS in lines carrying different dosages of *PPD1* loss-of-function mutations ($R^2 = 0.74$) (Shaw et al. 2013). A correlation between heading date and SNS has also been observed in genes regulated by *PPD1* such as the *FLOWERING LOCUS T1* gene (*FT1*) (Brassac et al. 2021; Finnegan et al. 2018; Isham et al. 2021; Lv et al. 2014).

FT1 encodes a mobile protein that travels through the phloem and carries environmental signals from the leaves to the shoot apical meristem (SAM), where it forms a complex with 14–3-3 and FD-like proteins (Florigen Activation Complex) (Taoka et al. 2011). This complex binds to the promoter of the meristem identity gene *VERNALIZATION1* (*VRN1*), promoting its expression and the transition from the vegetative to the reproductive phase in wheat (Li et al. 2015). Induction of *FT1* also results in the upregulation of *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1-1* (*SOC1*), *LEAFY* (LFY) and genes in the gibberellin (GA) pathway that are essential for spike development and stem elongation (Pearce et al. 2013). A deletion of *FT-B1* in hexaploid wheat delays the transition to reproductive growth and increases SNS (Finnegan et al. 2018).

In addition to *FT1*, wheat has at least five *FT-like* paralogs designated as *FT2* to *FT6* (Lv et al. 2014), which have some overlapping functions but also varying degrees of subfunctionalization (Halliwell et al. 2016; Lv et al. 2014). *FT2* is the most similar paralog to *FT1* (78% protein identity), but the two genes exhibit marked differences in transcription and protein interaction profiles. Whereas the FT1 protein interacts with five out of the six wheat 14-3-3 proteins tested so far, FT2 failed to interact with any of these members of the Florigen Activation Complex (Li et al. 2015). The two genes also differ in their temporal and spatial transcription profiles. *FT1* transcript levels in the leaves are upregulated earlier than *FT2* when plants are grown at room temperature, but only *FT2* is induced when plants are grown for a long period at 4 °C (vernalization) (Shaw et al. 2019). Interestingly, *FT2* is the only member of the wheat *FT-like* gene family that is expressed directly in the shoot apical meristem (SAM) and in the developing spike (Lv et al. 2014), in addition to leaves and elongating stems (Fig. S1).

Loss-of-function mutations in *FT2*, identified in a sequenced mutant population of the tetraploid wheat variety Kronos (Krasileva et al. 2017), resulted in limited differences in heading time but significantly increased SNS (Shaw et al. 2019). Similar increases in SNS were observed in *ft-B2* natural mutants detected in hexaploid wheat (Gauley et al. 2021). The loss-of function mutation in the A-genome copy of *FT2* (*FT-A2*) in Kronos was associated with significantly larger increases in SNS (10–15%) than the mutation in the B-genome copy (*FT-B2*, 2–5%). This difference in SNS was associated with much higher transcript levels of *FT-A2* relative to *FT-B2* in all tissues and developmental stages (Fig. S1). The increases in spikelet number in the double *ft-A2 ft-B2* mutant (henceforth *ft2*-null) were significantly larger than in the single *ft-A2* mutant confirming that the *FT-B2* gene still has a residual effect on SNS in spite of its lower transcript levels.

The increase in SNS in the *ft-A2* mutant was associated with reduced fertility, offsetting the potential positive effects of the increase in SNS on total grain yield (Shaw et al. 2019). This effect was observed in growth chambers under optimal conditions suggesting that is not an indirect effect of altered flowering time. We hypothesized that strong selection in cultivated wheat for grain yield might have selected an *FT-A2* variant with a positive effect on SNS, but without the associated negative effect on fertility. Analysis of natural variation in *FT-A2* revealed an aspartic acid to alanine change at position 10 (D10A) that was rare in tetraploid wheat but frequent in modern common wheat varieties, suggesting positive selection for the new allele. In this study, we characterized the effect of the D10A polymorphism on wheat heading time, SNS, grain number, and spike yield in different wheat classes and performed a high-density genetic mapping of the SNS QTL that identified *FT-A2* as the most likely candidate gene.

Material and methods

Analysis of the exome capture data generated by the WheatCAP project using the assay developed by NimbleGen (Krasileva et al. 2017) and deposited in the Wheat T3 database (https://wheat.triticeaetoolbox.org/) revealed the existence of an A to C SNP within the *FT-A2* coding region that resulted in the D10A polymorphism. We studied the effect of this SNP on heading time, SNS, grain number, and spike yield in two segregating populations in tetraploid and hexaploid wheat.

Biparental mapping population in tetraploid wheat (*Triticum turgidum* ssp. *durum*)

The tetraploid mapping population included 163 BC_1F_2 lines from the cross Kronos *2/Gredho (designated KxG hereafter). Kronos (PI 576168, *FT-A2* D10 allele) is a semi-dwarf (*Rht-B1b*), spring wheat with reduced photoperiod sensitivity (*Ppd-A1a*), whereas Gredho (PI 532239, *FT-A2* A10

allele) is a tall (*Rht-B1a*), photoperiod sensitive (*Ppd-A1b*) spring landrace from Oman. We planted the KxG population as headrows in the 2015–2016 season at the UC Experimental Field Station in Davis, CA, with each row including on average five individual plants.

Near-isogenic lines of the *FT-A2* A10 allele from Gredho into Kronos

We also evaluated the effect of the *FT-A2* alleles in two sets of near-isogenic lines (NILs). For the first set, we selected *FT-A2* heterozygous BC_1F_2 and BC_1F_3 lines from the cross Kronos *2/Gredho and selected two sets of homozygous BC_1F_{3-4} homozygous A10 and D10 sister lines using the *FT-A2* marker (H2-14 and H2-23). These lines were semidwarf and carried the *Ppd-A1a* allele for reduced photoperiod sensitivity and the *Vrn-A1* allele for spring growth habit. We used the BC_1F_{3-5} grains produced by these plants for two field experiments, one at the University of California, Davis (UCD) and the other one at Tulelake (California northern intermountain region). Both field experiments were organized in a complete randomized design with plants as experimental units. Three to five spikes were measured per plant and averaged for 10 plants per genotype at the UC Davis experiment. In the Tulelake experiment, 23–27 spikes per genotype were randomly collected and used as experimental units in the statistical analyses.

In parallel, we backcrossed the A10 allele into Kronos for three additional generations (Kronos *5/Gredho) and then selected BC_4F_2 NILs homozygous for the A10 and D10 alleles using the *FT-A2* molecular marker. The BC_4F_3 seed was increased in the greenhouse in 2020, and the BC_4F_4 grains were used for field experiments at UCD and Tulelake in 2021. These experiments used small plots (four 1-m rows, 1.1 m^2) as experimental units, organized in a randomized complete block design with 12 blocks. Grains of the BC_4F_4 Kronos isogenic line with the A10 allele were deposited in the National Small Grain Collection as PI 699107.

Biparental mapping population in hexaploid winter wheat

The hexaploid population included 358 F_5-derived recombinant inbred lines (RILs) from the cross between soft-red winter wheat lines LA95135 (CL-850643/PIONEER-2548//COKER9877/3/FL-302/COKER-762) x SS-MVP57 (FFR555W/3/VA89-22-52/TYLER//REDCOAT*2/GAINES). LA95135 is semidwarf (*Rht-D1b*) and photoperiod sensitive (*Ppd-D1b*), whereas SS-MVP57 is tall (*Rht-D1a*) and has reduced photoperiod sensitivity (*Ppd-D1a*) (DeWitt et al. 2021). This winter wheat population was previously genotyped and phenotyped as 1-m rows in the field at Raleigh, NC, and Kinston, NC, during the 2017–2018 season, and in Raleigh, Kinston, and Plains, GA, in the 2018–2019 season (DeWitt et al. 2021). These locations will be referred to as Raleigh (Ral), Kinston (Kin), and Plains (Pla) followed by the harvest year (18 or 19).

FT-A2 marker development and allelic frequencies

We targeted the *FT-A2*, D10A SNP at position 124,172,909 bp (RefSeq v1.0) on chromosome 3A with a Cleaved Amplified Polymorphic Sequence (CAPS) marker. Primers FT-A2-D10A forward and reverse (Table S1) amplify a fragment of 705 bp. After digestion with the restriction enzyme *Apa*I, the fragment amplified from the D10 allele remained undigested, whereas the fragment amplified from the A10 allele was digested into two fragments of 448 and 257 bp.

We used this marker to determine the frequency of the D10A mutation in 89 *T. urartu*, 82 *T. turgidum* ssp. *dicoccoides,* 32 *T. turgidum* ssp. *dicoccon,* 417 *T. turgidum* ssp. *durum,* and 705 *T. aestivum* accessions summarized in Supplementary Appendix S1. Among the hexaploid lines, we included a collection of 238 landraces and varieties (He et al. 2019) and a set of 126 winter wheats (T3/Wheat) genotyped by exome capture and with data for the *FT-A2* D10A polymorphism. We also used the *FT-A2* marker to genotype a panel of 242 spring wheats with reduced photoperiod sensitivity (Zhang et al. 2018) and a panel of 99 varieties and modern breeding lines from the Montana State University wheat breeding program (Supplementary Appendix S1). Based on the planting season used in the area where the spring varieties were developed, they were divided into those developed under spring planting (hereafter "DuS") or under fall planting (hereafter "DuF"). A previous study has shown that DuS and DuF groups are genetically differentiated using the 90 K SNP array (Zhang et al. 2018) (Supplementary Appendix S1).

High resolution genetic map

The high-resolution map of the KxG population was developed in two phases. In the first phase, we identified two BC_1F_3 plants from the KxG BC_1F_2 head rows, H2 and D12, which were heterozygous for *FT-A2* candidate region. From these heterozygous lines, we generated large segregating Heterogeneous Inbred Families (HIF) populations to identify recombination events within the *FT-A2* candidate region. For phenotypic screenings, these recombinants were space-planted at least three inches apart in a completely randomized design. Additional markers in the candidate gene region were developed for 11 genes on both sides of *FT-A2* covering a region of ~ 10 Mb using the exome capture sequence data from Kronos and Gredho (Table S1).

Statistical analysis

In the tetraploid biparental population, we analyzed the effect of the *FT-A2* alleles using a 3 × 2 factorial ANOVA. This analysis included the genotypic variation at *PPD-A1* and *RHT-B1* as additional factors, since both genes are known to have pleiotropic effects on heading time and yield components. In the hexaploid winter wheat population, we analyzed the effect of *FT-A2* in a 4 × 2 factorial ANOVA including the segregating genes *PPD-D1*, *RHT-D1*, and *WHEAT ORTHOLOG OF APO1* (*WAPO-A1*), which was previously shown to affect SNS (Kuzay et al. 2019). Analysis of Variance was conducted with the "Anova" function in R package "car" (Fox et al. 2019) with type 3 sum of squares.

Yeast two-hybrid assays

Modified Gateway (Invitrogen) bait/prey vectors pLAW10 and pLAW11 (Cantu et al. 2013) and yeast strain Y2HGold (Clontech, Mountain View, CA, USA) were used in the yeast two-hybrid assays. pLAW10 is the Gateway version of pGBKT7 (GAL4 DNA-binding domain, BD), and pLAW11 is the Gateway version of pGADT7 (GAL4 activation domain, AD). For all Gateway compatible cloning, pDONR/Zeo (Life Technologies, Grand Island, NY, USA) was used to generate the entry vectors. All constructs were verified by sequencing. Yeast two-hybrid assays were performed according to the manufacturer's instructions (Clontech). Transformants were selected on SD medium lacking leucine (− L) and tryptophan (− W) plates and re-plated on SD medium lacking –L, –W, histidine (− H), and adenine (− A) to test the interactions.

Results

Natural variation in *FT-A2*

We used exome capture data deposited in the T3/Wheat database (https://triticeaetoolbox.org/wheat/) to explore the natural polymorphisms in *FT-A2*. We identified an A to C SNP at position 124,172,909 in chromosome arm 3AS of the Chinese Spring (CS) RefSeq v1.0, which resulted in an amino acid change at position 10 of the FT-A2 protein from aspartic acid (D) to alanine (A) (henceforth, D10 and A10 alleles). In the analyzed accessions of *T. urartu*, *T. turgidum* ssp. *dicoccoides*, and *T. turgidum* ssp. *dicoccon*, we detected only the D10 allele (Table 1). D10 was also the only allele detected in all the other grass species we analyzed including *Lolium perenne* (AMB21802), *Oryza sativa* (XP_021310907), *Zea mays* (NP_001106251), and *Panicum virgatum* (APP89655), indicating that D10 is the ancestral allele. In this study, we describe the change from the ancestral to the derived allele (D10A) rather than relative to the Chinese Spring (CS) reference genome that carries the derived A10 allele,

We also screened a collection of 417 *T. turgidum* ssp. *durum* accessions with a CAPS marker for the D10A polymorphism (see "Material and methods") and found that only 0.7% carried the A10 allele (Table 1). Two of the three accessions with the A10 allele were from Oman (PI 532239 = 'Gredho' and PI 532242, 'Musane and Byaza') and the other one was from Turkey (PI 167718), suggesting that the A10 allele is almost absent from modern Western durum germplasm.

We detected a higher frequency of the A10 allele (56.5%) among 705 *T. aestivum* ssp. *aestivum* lines (Table 1). This overall frequency was similar to that detected in a worldwide collection of landraces and varieties combining winter and

Table 1 Frequency of the *FT-A2* alleles in different germplasm collections

Species	Ploidy	No. acc	A10%	D10%	A10	D10
T. urartu	2x	89	0.0%	100.0%	0	89
T. turgidum ssp. *dicoccoides*	4x	82	0.0%	100.0%	0	82
T. turgidum ssp. *dicoccon*	4x	32	0.0%	100.0%	0	32
T. turgidum ssp. *durum*	4x	417	0.7%	99.3%	3	414
T. aestivum Exome capture[a]	6x	238	59.7%	40.3%	142	96
T. aestivum US winter wheats[b]	6x	126	81.7%	18.3%	103	23
T. aestivum Spring DUF[c]	6x	149	58.4%	41.6%	87	62
T. aestivum Spring DUS[d]	6x	192	34.4%	65.6%	66	126

[a]He et al. (2019)

[b]T3/wheat

[c]Zhang et al. (2018)

[d] Zhang et al. (2018) + 99 breeding lines from MT

spring lines (59.7%) (He et al. 2019). We also analyzed the frequency of the D10A polymorphisms in two collections with known growth habit and found a higher frequency of the A10 allele among the winter lines (81.7%) than among the spring lines (44.9%, Table 1). Among the 341 spring wheat lines genotyped with the *FT-A2* marker, we found that varieties developed under fall-planting (DuF or long cycle) had a significantly higher frequency of the A10 allele (58.4%) than those developed under spring-planting (DuS or short cycle, 34.4%, χ^2 $P < 0.001$, Table 1). A complete list of the accessions used in these calculations is available in Supplementary Appendix 1, and a summary of the frequencies is presented in Table 1.

Effect of the D10A polymorphism in tetraploid wheat

To test the effect of the D10A polymorphism on SNS, we used the diagnostic CAPS marker to screen 163 BC_1F_2 plants from the KxG population segregating for this polymorphism. We also genotyped this population with markers for the segregating *RHT-B1* (Guedira et al. 2010) and *PPD-A1* (Wilhelm et al. 2009) genes, which can also affect SNS. Plants were grown in the field in the 2015–2016 season in Davis, CA, and were phenotyped for individual plant height (HT), days to heading (DTH), and spikelet number per spike (SNS, Table 2).

The three-way factorial ANOVAs including *FT-A2, RHT-B1,* and *PPD-A1* as factors showed significant effects for SNS, HT, and DTH and no significant interactions for any of the traits. As expected, *RHT-B1* showed the strongest effect on plant height and *PPD-A1* on heading time, although both genes affected both traits (Table 2). The strongest effect on SNS was detected for *PPD-A1*, but a significant effect was also detected for *FT-A2* (Table 2), with plants homozygous for A10 showing 6.4% higher SNS than those homozygous for D10 allele (Table 2). The differences in SNS between the *FT-A2* alleles were larger in the late flowering plants homozygous for the photoperiod sensitive allele from Gredho (2.3 spikelets/spike) than in the early flowering plants homozygous for the Kronos allele for reduced photoperiod sensitivity (1.0 spikelets per spike), but the interaction was not significant.

Effect of the *FT-A2* alleles in Kronos near-isogenic lines

To analyze the effect of the D10A polymorphism independently of the variability generated by other major genes, we evaluated two sets of near-isogenic lines in field experiments in 2020 (BC_1F_{3-5} sister lines) and 2021 ($BC_4F_{2:4}$ sister lines, see "Material and methods") at UCD and Tulelake. In the 2020 experiment at UCD, lines with the A10 allele showed large and significant increases in SNS (13.8%), grain number per spike (GNS, 31.7%), grains per spikelet (16.1%, also referred to as fertility), and grain yield per spike (33.0%) relative to the sister lines homozygous for the D10 allele (Table 3). The results from this experiment were consistent between two independent pairs of BC_1F_{3-5} sister lines (H2-14 and H2-23, Table 3). The 2020 experiment in Tulelake (Northern California, spring planting) using BC_1F_{3-5} sister lines from family H2-14 also showed a significant increase in SNS (4.0%), but the increases in GNS, grains per spikelet, and grain yield per spike were not significant (Table 3).

For the 2021 UCD experiment using 1.1-m^2 small plots as experimental units (12 replications), $BC_4F_{2:4}$ lines with the A10 allele headed on average 0.8 d later than the sister lines with the D10 allele ($P = 0.0252$) and showed significant increases in SNS (5.7%, $P = 0.0011$) and GNS (6.3%, $P = 0.0168$, Table 3). In this experiment, we did not detect

Table 2 Effects of *FT-A2*, *PPD-A1* and *RHT-B1* on plant height (HT), days to heading (DTH), and spikelet number per spike (SNS)

		Plant height (HT, cm)	Days to heading (DTH)	Spikelet no./spike (SNS)
FT-A2	Kronos (D10)	113.4 ± 3.2	130.5 ± 0.9	25.1 ± 0.5
LSmean ± s.e.m	Gredho (A10)	118.3 ± 2.2	130.6 ± 0.6	26.7 ± 0.3
Three-way ANOVA	*P* value	ns	ns	*
PPD-A1	Kronos	108.1 ± 2.3	120.8 ± 0.6	22.6 ± 0.4
LSmean ± s.e.m	Gredho	121.6 ± 2.6	141.0 ± 0.7	29.6 ± 0.4
Three-way ANOVA	*P* value	***	***	***
Rht-B1	Kronos	97.1 ± 2.5	131.5 ± 0.7	26.2 ± 0.5
LSmean ± s.e.m	Gredho	131.4 ± 2.8	130.2 ± 0.8	25.8 ± 0.4
Three-way ANOVA	*P* value	***	*	ns

Three-way ANOVA with *P* values of the main effects and least-square means (LSmeans)

ns not significant, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$

All the interactions were non-significant

Table 3 Comparisons of near-isogenic lines with the *FT-A2* A10 and D10 alleles in field experiments at UC Davis and Tulelake in 2020 and 2021. All % changes are relative to D10

	Allele	*N*	SNS	GN	Grains/spikelet	GW mg	Yield/spike g	Yield/plot g
Davis 2020								
H2-14	D10	10[a] (54 spikes)	20.27	59.22	2.92	55.81	3.31	NA
H2-14	A10	10[a] (38 spikes)	21.92	70.77	3.23	56.78	4.07	NA
		A10% change	8.1**	19.6***	10.6**	1.8 ns	22.9**	
H2-23	D10	10[a] (39 spikes)	19.36	56.31	2.89	54.47	3.07	NA
H2-23	A10	10[a] (38 spikes)	23.11	81.06	3.51	54.21	4.41	NA
		A10% change	19.4***	44.0***	21.5***	−0.6 ns	43.6***	
Tulelake 2020								
H2-14	D10	27 spikes	17.15	44.11	2.57	38.26	1.69	NA
H2-14	A10	23 spikes	17.83	46.48	2.61	40.49	1.88	NA
		A10% change	4.0***	5.4 ns	1.4 ns	5.8 ns	11.2 ns	
Davis 2021								
$BC_4F_{2:4}$	D10	12[b] (96 spikes)	18.52	67.85	3.67	60.34	4.09	1254
$BC_4F_{2:4}$	A10	12[b] (96 spikes)	19.58	72.15	3.69	55.61	4.01	1251
		A10% change	5.7 **	6.3 *	0.5 ns	−7.8 ***	−2.0 ns	−0.2 ns
Tulelake 2021								
$BC_4F_{2:4}$	D10	12[b] (96 spikes)	15.39	41.48	2.69	58.34	2.43	746
$BC_4F_{2:4}$	A10	12[b] (96 spikes)	16.02	44.02	2.75	59.45	2.62	854
	Null	12[b] (96 spikes)	16.54	40.39	2.44	54.07	2.20	789
		A10% change	4.1**	6.1**	2.1 ns	1.9 ns	7.8**	14.5 ns
		ft2-null % change	7.5***	–2.6 ns	−9.4***	−7.3***	−9.5**	5.8 ns

[a]Experimental units were 1 m rows, with 3–5 spikes measured per row

[b]Experimental units were 4 row plots (1.1 m^2), with 8 spikes measured per plot

significant differences in grains per spikelet ($P = 0.7919$). We observed a negative correlation between average GNS and grain weight across the 24 plots ($R = -0.61$) and a significant negative effect of the A10 allele on kernel weight (– 7.8%, $P = 0.0002$). The negative effect on grain weight offset the positive effect of the A10 allele on grain number resulting in non-significant differences in grain weight per spike or per plot (Table 3).

For the 2021 Tulelake experiment using 1.1-m^2 small plots (12 replications), we included the Kronos lines with truncation mutations in *FT-A2* and *FT-B2* (*ft2*-null, henceforth) developed before (Shaw et al. 2019) in addition to the $BC_4F_{2:4}$ Kronos lines with the D10 and A10 alleles. The lines with the A10 allele showed highly significant increases in SNS (4.1%), GNS (6.1%), and grain yield per spike (7.7%) that were of similar magnitude to the ones observed in the 2020 Tulelake experiment (Table 3). The null line also showed a significant increase in SNS (7.5%) relative to the wildtype Kronos (D10), but the negative impact of the *FT2* loss-of-function mutations in grains per spikelet (−9.4%) and grain weight (−7.3%) resulted in a significant reduction in grain yield per spike (−9.3%, Table 3). No significant differences in grain yield per plot were detected among the three lines.

The A10 allele has a positive effect on SNS and spike yield in winter wheat

To analyze the effect of the D10A *FT-A2* alleles in winter wheat, we used phenotypic data available from 358 F_5-derived RILs from the cross between soft-red winter wheat lines LA95135 and SS-MVP57 (DeWitt et al. 2021) and genotypic data for the *FT-A2* marker developed in this study. This population was also segregating for *PPD-D1*, *RHT-D1*, and *WAPO-A1*, which were included as factors together with *FT-A2* in a 4 × 2 factorial ANOVA.

Plants carrying the *FT-A2* allele A10 (SS-MVP57) headed on average 1.7 days later ($P < 0.001$, Fig. 1a) and had 0.6 more spikelet per spike (5.1% increase, $P < 0.001$, Fig. 1b) than plants carrying the D10 allele (LA95135). The differences in SNS were significant in all tested locations. The A10 allele was also associated with a significant increase in GNS in the overall ANOVA ($P < 0.001$), but the separate analyses of the two locations showed significant differences only at Pla19 (4.4 more grains per spike, $P < 0.001$, Fig. 1c). No significant effects were detected on fertility (Fig. 1d). A significant increase in spike yield was associated with the A10 allele in the overall ANOVA (average 4.6%,

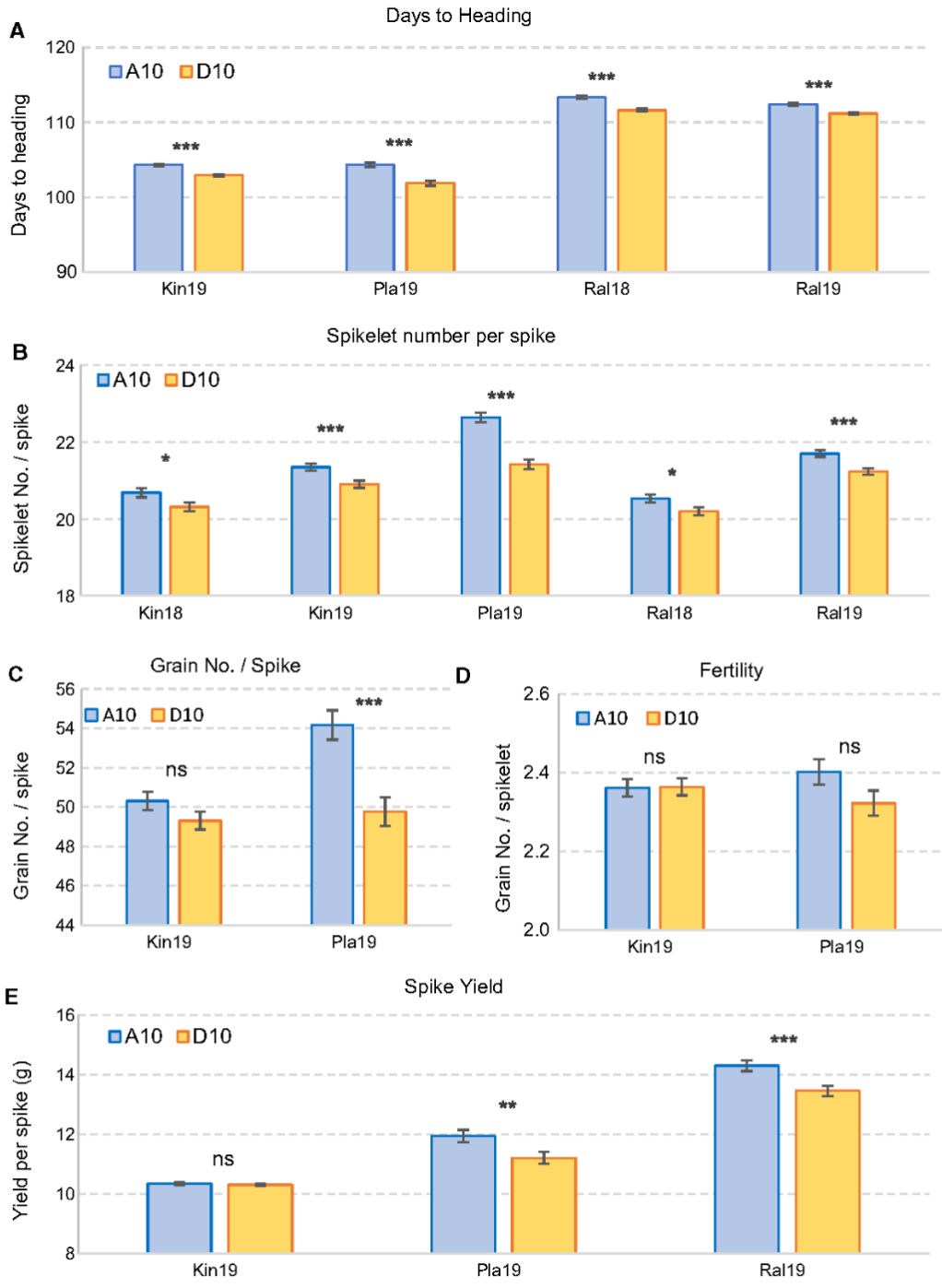


Fig. 1 Comparison between *FT-A2* A10 (SS-MVP57) and D10 (LA95135) alleles in winter wheat. **a** Days to heading. **b** Spikelet number per spike. **c** Grain number per spike. **d** Grain number per spikelet (fertility). **e** Average spike yield. Bars are least square means from a factorial ANOVA including *PPD*-D1, *RHT-D1* and *WAPO-A1* as factors. Error bars are s.e.m. *ns* not significant, $*P=0.05$, $**P=0.01$, $***P=0.001$

$P < 0.001$), and two out of the three tested locations were significant in the analyses by location ($P < 0.001$, Fig. 1e).

High resolution mapping of the SNS QTL on chromosome 3AS

The previous results showed that the haplotypes associated with the *FT-A2* D10 and A10 alleles have a significant effect on SNS. To narrow down the candidate gene region and explore the linkage between the *FT-A2* D10A polymorphism and the differences in SNS, we generated a high-density map of the 3AS chromosome region in tetraploid wheat using a total of 3,161 BC_1F_3, BC_1F_4, and BC_1F_5 plants derived from the KxG population. These plants were screened in separate batches over three years using flanking markers 3A-117.83 and 3A-127.82 (numbers indicate coordinates in RefSeq v1.0 in Mb). Within this 9.9-Mb region including *FT-A2* (124.17 Mb), we identified 76 recombination events corresponding to a genetic distance of 1.58 cM (6.26 Mb per cM). One of these recombination events (H2-6-#14-5) was detected in the progeny test of primary recombinant H2-#6, which explains the presence of two close recombination events in this line (Table 4).

In addition to the molecular marker for the *FT-A2* D10A SNP and the two flanking markers, we developed eight more KASP and CAPS markers in the candidate region (Table S1) and used them to genotype plants carrying recombination events in the region. The lines with the 10 closest recombination events to *FT-A2* are presented in Table 4 together with the results of the field progeny tests for SNS. Progenies of the lines H2-#6 and H2-14#17-2 heterozygous for *FT-A2* showed significant differences in SNS ($P < 0.01$) between lines homozygous for the two parental alleles, whereas progeny tests for the eight lines homozygous for *FT-A2* did not show significant difference in SNS between parental alleles in the heterozygous flanking regions (Table 4). Average SNS were as expected, with the lines homozygous for the A10 allele having 1.3 more spikelets on average than the lines homozygous for the D10 allele.

The phenotype of the critical recombinant line #18-5 with the closest distal recombination event to *FT-A2* was validated in a separate experiment in Davis in 2021 (Table S2). In this experiment, control lines showed highly significant differences in SNS ($P < 0.0001$) confirming that the differences in SNS were detectable in this experiment. By contrast, there was no significant difference between the sister lines with and without the recombination event #18-5, with both lines showing SNS values similar to the control line with the Gredho allele (Table S2). Taken together, these results confirmed that the causal gene for the 3AS QTL for SNS was proximal to the marker located at CS RefSeq v1.0 coordinate 120,227,651 (Table 4).

Table 4 Critical recombinant BC_1F_5 from Davis 2019–2020 field seasons

Marker	Chr. 3AS CS	H2	H2-6	H2-14				H2-23		D12 11-1	
		#6	#14-5	#17-2	#1-3	#18-5	#47-1	#47-5	#53-4	#71-1	#73-1
3A-117.83	117,828,272	H	H	H	H	H	K	H	H	H	K
3A-120.23	120,227,651	H	G	H	K	H	K	K	G	K	K
3A-121.48	121,482,459	H	G	H	K	G	K	K	G	K	K
3A-121.65	121,646,195	H	G	H	K	G	K	K	G	K	K
3A-122.54	122,540,617	H	G	H	K	G	K	K	G	K	K
FT-A2-L4	122,542,102	H	G	H	K	G	K	K	G	K	K
FT-A2 SNP D10A	124,172,909	H	G	H	K	G	K	K	G	K	K
SNS PHENOTYPE		H	G	H	K	G	K	K	G	K	K
FT-A2-R1	125,094,949	H	G	H	K	G	K	K	G	K	K
3A-125.40	125,402,254	H	G	H	K	G	K	K	G	K	K
3A-126.57	126,567,437	K	K	H	K	G	K	K	G	K	K
3A-127.82	127,821,835	K	K	K	K	G	H	K	G	K	H
Number of plants in PT		34	83	71	72	79	70	74	75	80	81
SNS Avg. Gredho allele (G)		22.1	23.9	23.5	22.4	24.2	22.4	22.7	24.1	22.3	23.1
SNS Avg. Kronos allele (K)		21.6	23.2	21.7	22.5	23.0	21.9	22.4	24.3	21.7	22.4
P values K vs G		3e-05	NS	0.004	NS	NS	NS	NS	NS	NS	NS

All lines except recombinant H2 #6 were evaluated in the 2020 field season. Comparisons of SNS for statistical significance are only between sister lines segregating for the heterozygous region

Table 5 Spikelet number per spike (SNS) evaluation of BC_1F_6 homozygous sister lines from recombinant line H2-18 #28-4 in Tulelake 2020

Marker	Chr.3AS CS	H2-18 #28-4-1	H2-18 #28-4-3	H2-14 #17-2-18	H2-14 #17-2-22
3A-117.82	117,828,272	G	G	K	G
3A-120.2	120,227,651	G	G	K	G
3A-121.4	121,482,459	G	G	K	G
3A-121.64	121,646,195	G	G	K	G
3A-122.540	122,540,617	G	G	K	G
FT-A2-L4	122,542,102	G	G	K	G
FT-A2 SNP D10A	124,172,909	G	G	K	G
SNS PHENOYPE		G	G	K	G
FT-A2-R1	125,094,949	G	G	K	G
3A-125.4	125,402,254	K	G	K	G
3A-126.5	126,567,437	K	G	K	G
3A-127.8	127,821,835	K	G	K	K
Number of plants		40	42	43	40
SNS Avg		17.68	17.87	16.94	17.94
P values		0.287		1.78E-09	

Sister line #28-4-#1 carried a proximal Kronos chromosome segment and sister line #28-4-#3 a proximal Gredho chromosome segment. Lines #17-2-18 (*FT-A2* D10) and #17-2-22 (*FT-A2* A10) were included as controls

Later, we identified an additional line (BC_1F_4 H2-18 #28-4) with a closer recombination event to *FT-A2* in the proximal region between *FT-A2-R1* and *3A-125.4*, which we planted in a separate field experiment at Tulelake in the spring of 2020. This experiment included homozygous sister lines #28-4-1 and #28-4-3 fixed for either the Kronos or Gredho alleles in the segregating proximal region (Table 5) and as controls sister lines derived from plant #17-2 (Table 4) that were either homozygous for the *FT-A2* D10 (#17-2-18) or A10 allele (#17-2-22, Table 5). These two lines showed highly significant differences in SNS ($P < 0.0001$, Table 5) confirming that it was possible to detect differences between the two *FT-A2* alleles in this experiment. By contrast, there was no significant difference between the H2-18 #28-4 recombinant sister lines, confirming that the candidate gene was still linked to *FT-A2* (Table 5). Based on this result, we established a closer proximal flanking marker (3A-125.4) and reduced the candidate region for the 3AS QTL to a 5.2 Mb interval between coordinates 120,227,651 and 125,402,254 (Table 5).

Genes in the candidate gene region for the 3AS QTL for SNS

The annotated Chinese Spring reference genome region (RefSeq v1.1) between the two flanking markers defined in the previous section encompasses 28 high-confidence genes (including flanking genes *TraesCS3A02G141000* and *TraesCS3A02G143700*). The exome capture data revealed non-synonymous SNPs between Kronos and Gredho in only three out of the 28 genes, including the D10A polymorphism in *FT-A2*. The other two genes are described briefly below.

TraesCS3A02G142200 encodes a leucine-rich repeat receptor-like protein kinase, so it is difficult to predict its potential effects. The predicted R872H amino acid change in Kronos (RefSeq v1.1 3AS 121,646,195) is in a conserved region close to the end of the protein (893 amino acids) and has a BLOSUM62 score of 0, predictive of a low probability of changes in protein structure or function. The R872H polymorphism was not detected in the parental lines LA95135 and SS-MVP57 of the hexaploid winter wheat populations segregating for the 3AS SNS QTL.

TraesCS3A02G143600 encodes a short peptide (104 amino acids) with a polymorphism in Kronos that generates a premature stop codon (S59*, RefSeq v1.1 3AS 125,094,949 C to A). However, the predicted protein in Gredho also seems to be truncated since it is much shorter (104 amino acids) than the orthologous protein in wild emmer (XP_037404892.1, 483 amino acids) or *T. urartu* (EMS53367.1, 348 amino acids). In addition, the 104 amino acids in Gredho showed no similarity to other plant proteins in the GenBank nr database in species outside

the genus *Triticum*, suggesting that *TraesCS3A02G143600* encodes a non-functional protein in both Kronos and Gredho. Similar to R872H, the S59* polymorphism was not detected in winter lines LA95135 and SS-MVP57.

The QTLs in the KxG and LA95135 x SS-MVP57 segregating populations are co-located and affect the same traits, so we hypothesized that they should have mutations in the same gene. Therefore, we prioritized genes carrying mutations in both populations. The predicted R872H amino acid change in *TraesCS3A02G142200* was polymorphic in the KxG population but not in the winter wheat population, and the same was true for the S59* premature stop codon in *TraesCS3A02G143600*. By contrast, the D10A polymorphism in FT-A2 (*TraesCS3A02G143100*) was detected in both mapping populations.

The R872H and S59* polymorphisms were found in tetraploid wheat but were not detected in any of the sequenced hexaploid wheats in the wheat PanGenome project (Walkowiak et al. 2020). By contrast, *FT-A2* was polymorphic in the same group of varieties, with CDC Landmark, Lancer, and Spelt carrying the D10 allele and CS, Julius, Jagger, CDC Stanley, ArinaLRFor, Mace, Norin 61, and SY Mattis carrying the *FT-A2* A10 allele. The previous observations, together with the known effect of *FT2* mutations on SNS (Shaw et al. 2019), suggest that *FT-A2* is the most likely candidate gene in this region. However, we cannot rule out the possibility of polymorphisms shared between the two populations in the regulatory regions of other candidate genes.

Effect of the D10A polymorphism on FT-A2 interactions with 14-3-3 proteins

Previous results have shown positive interactions between FT1 and six of the seven 14-3-3 proteins tested, whereas FT-A2 did not interact with any of the 14-3-3 proteins (Li et al. 2015). This was a puzzling result because all other four FT-like genes showed positive interactions with at least one 14-3-3 protein. Since the original study was done using only the FT-A2 D10 allele, we decided to explore the effect of the A10 allele. In this study, both the FT-A2-D10 and FT-A2-A10 proteins failed to interact with any of the six tested 14-3-3 proteins, whereas the FT1 positive control showed a strong interaction signal (Fig. S2). No autoactivation was observed in the negative controls. Given the lack of interactions between both FT-A2 alleles and any of the tested 14-3-3 protein, we have initiated Y2H screens to test if there are other protein partners of FT-A2.

Discussion

Candidate gene and causal polymorphism

In this study, we focused on SNS for the high-density mapping because this trait has a higher heritability ($h > 0.8$) than other yield components (Kuzay et al. 2019; Zhang et al. 2018). Spikelet number per spike is determined early after the transition from the vegetative to the reproductive phase, when the spike meristem transitions into a terminal spikelet (Li et al. 2019). This limits the influence of later environmental variability on SNS relative to GNS or grain weight, which are affected by fertility, grain abortions, and conditions affecting grain filling until the end of the season.

The high heritability of SNS helped us to Mendelize the QTL (using large progeny tests) and to generate a high-density map of the SNS QTL on chromosome arm 3AS. We established a 5.2 Mb candidate gene region on chromosome arm 3AS including 28 annotated high-confidence genes in CS, including three with non-synonymous polymorphisms between Kronos (D10) and Gredho (A10): *TraesCS3A02G143100* (D10A), *TraesCS3A02G142200* (R872H), and *TraesCS3A02G143600* (S59*). The last two polymorphism were detected in the Kronos x Gredho population but not in the LA95135 (D10) and SS-MVP57 (A10) suggesting that they are unlikely candidate genes for the SNS QTL. In additions, the S59* and R872H polymorphisms were not detected among the varieties sequenced in the wheat pangenome (Walkowiak et al. 2020), which suggests that they originated in durum wheat and that the A10 mutation occurred in a haplotype different from the one present in modern durum wheat varieties (S59*-R872H haplotype).

After the elimination of these two genes, *FT-A2* is the only gene in the candidate region that has a non-synonymous polymorphism (D10A) linked to the differences in SNS in both mapping populations. Although we cannot rule out the possibility of polymorphisms in regulatory regions of other candidate genes affecting SNS in both populations, the genetic data presented here, together with the known effect of the loss-of-function mutations in *FT-A2* on SNS (Shaw et al. 2019), point to *FT-A2* as the most likely candidate gene for the SNS QTL.

The D10A amino acid change in FT-A2 has a BLOSUM 62 score of -2 and is located in a conserved region of the protein, suggesting a high probability of an effect on either protein structure or function. To test if any other polymorphisms in *FT-A2* were associated with the D10A polymorphism, we compared the available exons, introns, 5' upstream region (5000 bp) and 3' downstream region (2000 bp) of *FT-A2* in genomic sequences of hexaploid wheat (Walkowiak et al. 2020). We did not find any additional SNPs that differentiate the varieties with the D10 allele (CDC Landmark, Lancer and Spelta) from those carrying the A10 allele (CS, Julius, Jagger, CDC Stanley, ArinaLRFor, Mace, Norin 61, and SY Mattis) in the analyzed regions. Although we cannot completely rule out the possibility of polymorphisms located in regulatory regions outside the investigated region, the available evidence points to D10A as the most likely causal polymorphism. A conclusive test of this hypothesis will

require the editing of the A124,172,909C, but this is not simple because this is a transversion, and currently available plant gene editors are not efficient to edit transversions. New prime editing technologies (Anzalone et al. 2019) may solve this problem once they become more efficient in plants (Lin et al. 2020).

Differential recombination rates within the candidate gene region

The distribution of recombination events (RE) in the 10-Mb region between the flanking markers used in this study was not uniform. In the 2.4 Mb distal to the candidate gene region (117.8 to 120.2 Mb, 14 genes), we detected 56 RE resulting in an average of 23.3 RE/Mb or 4.0 RE/gene. In the 2.4 Mb proximal to the candidate region (125.4 to 127.8 Mb, 13 genes), we detected 20 RE resulting in a frequency of 8.3 RE /Mb or 1.5 RE/gene. Surprisingly, not a single RE was detected in the 5.2-Mb central candidate region (120.2 to 125.4 Mb, 28 genes), despite being twice as large and including twice the number of genes as the flanking regions. Recombination events occur mainly in gene regions (Darrier et al. 2017), so we would have expected to find 39 of the 76 RE within the candidate region if RE were distributed proportionally to the number of genes. The same number would be expected if RE were distributed proportionally to the physical length of the interval.

To explore if this lack of recombination in the central region was caused by a structural rearrangement, we used the sequenced genome of the tetraploid variety Svevo (Maccaferri et al. 2019) that showed the same SNPs as the Kronos exome capture across the candidate gene region. Since Gredho showed very few polymorphisms with CS across the candidate gene region, we compared the genomes of CS (A10) and Svevo (D10) in this region. In Svevo, we found orthologs to the 28 high confidence genes present in CS, with the exception of *TraesCS3A02G142500* that was present in the correct position and strand in Svevo (100% identical over all its length) but was not annotated. All the genes were in the same orientation in CS and Svevo, and the total length of the region was similar in both species (5.2 Mb), suggesting that no major structural rearrangements occurred in the candidate gene region.

Finally, we did a BLAST comparison of all the Svevo genes to a Kronos scaffold assembly from the Earlham Institute, UK, and were able to detect 27 of the 28 genes with 100% identity. The only exception was *TRITD3Av1G056250* (ortholog of CS *TraesCS3A02G142600*), for which we only detected the B-genome homeolog in Kronos. These results suggest the Kronos genome is not very different from Svevo in this region. We currently do not know the cause of the reduced recombination frequency between 121.5 and 125.1 Mb in the KxG population, but since no pseudomolecule assembly of Kronos or Gredho is available, we cannot rule out the possibility of structural rearrangements in this region in one of these two varieties.

Effect of *FT-A2* D10A polymorphism on heading time and fertility

Wheat varieties are selected to flower within a narrow time window to maximize grain productivity. This limits the introgression of loss-of-function alleles that have beneficial effects on SNS but generate large delays in heading time, such as those in *VRN1* (Li et al. 2019) or *PPD1* (Shaw et al. 2013). By contrast, the *FT-A2* A10 allele has a positive effect on SNS and limited effect on heading time. Even when loss-of-function mutations in *ft-A2* and *ft-B2* were combined in Kronos, the delay in heading time was only 2–4 days (Shaw et al. 2019). In this study, the D10A polymorphism showed small effects on DTH in different genetic backgrounds, ranging from a non-significant difference in the initial Kronos x Gredho population (Table 2), a marginally non-significant difference of 0.8 d ($P=0.053$) in the 2021 field experiment comparing Kronos isogenic lines, and an average difference of 1.7 d in the winter wheat population (Fig. 1A).

An important limitation for the utilization of the *ft-A2* loss-of-function mutation for wheat improvement was its negative effect on fertility (Shaw et al. 2019), which offset its positive effect on SNS, as confirmed in the 2021 Tulelake experiment in this study (Table 3). This motivated our initial search for *FT-A2* natural variants that separated the positive effects on SNS from the negative effects on fertility. Results presented in this study show that the positive effect of the A10 polymorphism on SNS were translated into positive effects on GNS in both the winter wheat population (Fig. 1e) and in the spring NILs (Table 3). In addition, this allele was not associated with negative effects on the number of grains per spikelet in any of the studied populations, suggesting that the A10 allele has no negative effect on fertility. These results provide a good example of the value of using natural variants selected by breeders to identify mutations that optimize specific traits with limited negative pleiotropic effects.

FT-A2 effects on SNS, GNS, grain weight and spike yield

It was encouraging to see that the positive effect of the A10 allele on SNS and GNS was expressed in both winter (Fig. 1) and spring wheats (Table 3), and among the latter in both spring and fall planted spring wheats. However, the magnitude of the increases in SNS, GNS, and spike yield associated with the A10 allele varied among experiments, suggesting that the effects of this *FT-A2* polymorphisms on these traits are modulated by the environment. We also observed variable effects of the A10 polymorphisms on grain weight.

Whereas no significant effects were detected for this trait in the experiments performed at UCD in 2020 and at Tulelake in 2020 and 2021, we detected a significant reduction in grain weight in the field experiment performed at UCD in 2021, which offset the gains in GNS (Table 3).

Similar observations have been reported for *WAPO-A1*, the causal gene of a wheat SNS QTL on the long arm of chromosome 7AL (Kuzay et al. 2019). Increases in SNS associated with the favorable *Wapo-A1b* allele were translated into significant increases in grain yield only when the favorable *WAPO-A1* allele was present in high-yielding/high-biomass genetic backgrounds and the plants were grown in a favorable environment. When the *Wapo-A1b* allele was present in poorly adapted varieties or when the lines were grown under water-limiting conditions, the plants did not have enough resources to fill the extra grains, resulting in a negative correlation between grain number and grain weight that limited the gains in grain yield (Kuzay et al. 2019). A study with elite CIMYT lines also highlighted the importance of genetic-by-environment interactions on the trade-offs between grain number and grain weight (Quintero et al. 2018). We hypothesize that environmental differences between our 2020 and 2021 field trials may have contributed to the observed differences in grain weight, in spite of the positive effects of the A10 allele on SNS and GNS detected in both years (Table 3). We also hypothesize that the introgression of the *FT-A2* A10 allele into more productive durum wheat varieties than Kronos may result in significant increases in total grain yield.

FT-A2 as a candidate gene for previously published SNS QTLs on chromosome arm 3AS

A QTL for DTH (*Qncb.HD-3A*) was previously mapped on chromosome 3A within a 400 Mb interval including *FT-A2* (DeWitt et al. 2021) in the LA95135 x SS-MVP47 population. We found in this study that LA95135 carries the D10 allele and SS-MVP47 the A10 allele, and after genotyping the population with the *FT-A2* marker, we found that the A10 allele was associated not only with a slight delay in heading time but also with higher SNS, GN, and grain yield per spike (Fig. 1). The similar pleiotropic effects of the SNS QTL in the winter wheat population and the KxG population, together with the overlapping mapping regions, suggest that the *FT-A2* D10A polymorphism may have contributed to the *Qncb.HD-3A* identified in the LA95135 × SS-MVP47 population.

An additional QTL for DTH was identified in the Avalon x Cadenza population (U.K.) on chromosome arm 3AS around the peak marker BS00021976 (169 Mb RefSeq v1.0) (Martinez et al. 2021). This QTL interval (60 Mb at each side of BS00021976) includes 536 annotated genes, among which the authors proposed *FT-A2* as a candidate of particular interest. Using our *FT-A2* marker, we established that both Avalon and Cadenza carry the A10 allele, so we conclude that the D10A polymorphism is not the cause for the observed QTL for DTH on 3AS in this population. Martinez et al. (2021) suggested that differences in *FT-A2* transcript levels may contribute to the differences in DTH, but more precise mapping of this QTL will be necessary to support this hypothesis.

Several QTLs for grain yield components have been reported in different regions of chromosome 3AS in a recombinant inbred chromosome line from the cross between cultivar Cheyenne and a substitution of chromosome 3A of Wichita in Cheyenne (CNN(Wichita-3A)) (Ali et al. 2011; Campbell et al. 2003; Dilbirligi et al. 2006). QTLs for grain yield and grain number per square meter were mapped in a region between markers *barc86* and *barc67* (54.4 to 464.3 Mb RefSeq v1.0, "Region 2") which encompasses the *FT-A2* locus. However, both Cheyenne and CNN(Wichita-3A) have the A10 allele of *FT-A2* (Supplementary Appendix S1), suggesting that a different gene (or a different polymorphism in *FT-A2*) was the cause of this QTL.

FT-A2 allele frequencies and breeding applications

The *FT-A2* alleles show contrasting frequencies in durum and common wheat, with the A10 allele present in less than 1% of the durum accessions and in 56% of the common wheat varieties analyzed in this study (Table 1). We currently do not know if the A10 allele originated in the few durum accessions carrying this allele in Oman and Turkey, or if these represent later introgressions from hexaploid to tetraploid wheat. Independently of its origin, the frequency of the A10 allele increased rapidly since its introgression or origin in common wheat, suggesting that this allele was favored by common wheat breeders.

The low frequency of the A10 allele in durum wheat could be a result of infrequent gene flow from hexaploid wheat to tetraploid wheat. However, it can also be the result of selection for larger grains and indirect selection for reduced GNS in environments showing a negative correlation between these two traits. Similar to *FT-A2,* the *Wapo-A1a* allele for low SNS is almost fixed in durum wheat, whereas the *Wapo-A1b* allele for high-SNS is found at high frequencies in hexaploid wheat (Kuzay et al. 2019). We interpret this similar asymmetric distribution of *WAPO-A1* and *FT-A2* alleles for SNS in common and durum wheat as indirect support for the hypothesis that selection for larger grains may have resulted in indirect selection for reduced SNS in durum wheat.

Among hexaploid spring wheats, we also observed significant differences in the distribution of the *FT-A2* alleles, with a larger frequency of the A10 allele among spring varieties

developed under a long growing cycle (DuF, 58.4%) than among those developed under a short growing cycle (DuS, 34.4%). We speculate that longer cycles may provide more resources to fill the extra grains associated with the A10 allele, facilitating the translation of the difference in SNS into differences in grain yield. This in turn may result in a positive selection pressure for the A10 allele in the fall-planted programs. This idea is indirectly supported by the high frequency of the A10 allele among the US winter wheat varieties (Table 1, 81.7%). Additional experiments with D10 and A10 NILs in different genetic backgrounds tested in different spring-planted and fall-planted locations will be necessary to test this hypothesis.

The high frequency of the A10 allele in the winter wheats and fall-planted spring wheats provides additional evidence that this allele has positive effects in the regions where these cultivars are grown. However, as the frequency of the A10 allele increases, the number of varieties that can benefit from its introgression decreases. By contrast, the A10 allele is almost absent from modern durum wheat breeding programs, providing an opportunity to benefit a large proportion of these varieties. To facilitate the testing and introgression of the A10 allele into durum wheat breeding programs, we deposited the Kronos NIL with the A10 allele in the NSGC (PI 699107). Kronos is a modern durum wheat variety with excellent pasta quality, which makes it a better donor parent than Gredho.

Our preliminary results suggest that the A10 allele may be more beneficial in fall planted than in the spring planted durum wheat programs, but additional experiments are necessary to test this hypothesis. It will be also interesting to investigate the combined effect of the A10 allele with alleles from other genes that also result in increases in SNS such as *Wapo-A1b* (Kuzay et al. 2019) and the *Elf3* allele from *T. monococcum* (Alvarez et al. 2016).

In summary, the genetic information provided in this study, together with the previous mutant information, provides strong evidence that *FT-A2* is the causal gene for the differences in SNS, GNS, and spike yield associated with this region on chromosome arm 3AS. The identification of the likely causal polymorphism (D10A) and the development of a perfect marker for this polymorphism in this study can accelerate the deployment of this favorable allele in wheat breeding programs worldwide.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-021-03992-y.

Acknowledgements We thank Josh Hegarty for his help with the field experiments and Youngjun Mo for the development of Kronos x Gredho population. We also thank Xiaoqin Zhang for her help with the introgression of the *A10* allele into Kronos and Saarah Kuzay for the phenotypic and genotypic data for the Kronos x Gredho population. We thank Mohammed Guedira for help with development and field evaluation of the LA95135 X SS-MPV57 population and Alina Akhunova, Mary Guttieri and Brian Ward for early access to the genotypic data for the winter wheat varieties.

Author contribution statement PG conducted most of the experimental work and wrote the first version of the manuscript. JZ contributed experimental work and many of the statistical analysis. KL contributed the Y2H experiments. GBG and ND contributed the LA95135 × SS-MVP57 population and the corresponding genotypic and phenotypic data. JC contributed the frequency of the D10A polymorphism in Montana spring wheat breeding lines and EA in US winter wheat lines. JD initiated and coordinated the project, contributed to data analyses, and supervised PG. All authors reviewed the manuscript and provided suggestions.

Funding This project was supported by the Agriculture and Food Research Initiative Competitive Grants 2017-67007-25939 (WheatCAP) from the USDA National Institute of Food and Agriculture and by the Howard Hughes Medical Institute. Priscilla Glenn acknowledges support from NSF Graduate Research Program Fellowship Grant 2036201.

Data availability All data and materials described in this paper are available from the corresponding author upon request. The *FT-A2* introgression in Kronos is deposited in the National Small Grains Collection (PI 699107). PI accession numbers are provided for all germplasm used when available. The datasets retrieved and analyzed during the current study are available in the T3/Wheat exome capture database (https://wheat.triticeaetoolbox.org/).

Code availability Not applicable.

Declarations

Conflict of interest The authors declare no conflict of interests or competing interests.

References

Ali ML, Baenziger PS, Al Ajlouni Z, Campbell BT, Gill KS, Eskridge KM, Mujeeb-Kazi A, Dweikat I (2011) Mapping QTL for agronomic traits on wheat chromosome 3A and a comparison of recombinant inbred chromosome line populations. Crop Sci 51:553–566

Alvarez MA, Tranquilli G, Lewis S, Kippes N, Dubcovsky J (2016) Genetic and physical mapping of the earliness per se locus *Eps-A*^{m} 1 in *Triticum monococcum* identifies *EARLY FLOWERING* 3 (*ELF*3) as a candidate gene. Funct Integr Genomic 16:365–382

Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 576:149–157

Brassac J, Muqaddasi QH, Plieske J, Ganal MW, Röder MS (2021) Linkage mapping identifies a non-synonymous mutation in *FLOWERING LOCUS T* (*FT-B1*) increasing spikelet number per spike. Sci Rep-Uk 11:1585

Campbell BT, Baenziger PS, Gill KS, Eskridge KM, Budak H, Erayman M, Dweikat I, Yen Y (2003) Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat. Crop Sci 43:1493–1505

Cantu D, Yang B, Ruan R, Li K, Menzo V, Fu D, Chern M, Ronald PC, Dubcovsky J (2013) Comparative analysis of the defense response interactomes of rice and wheat. BMC Genomics 14:166

Darrier B, Rimbert H, Balfourier F, Pingault L, Josselin AA, Servin B, Navarro J, Choulet F, Paux E, Sourdille P (2017) High-resolution mapping of crossover events in the hexaploid wheat genome suggests a universal recombination mechanism. Genetics 206:1373–1388

DeWitt N, Guedira M, Lauer E, Murphy JP, Marshall D, Mergoum M, Johnson J, Holland JB, Brown-Guedira G (2021) Characterizing the oligogenic architecture of plant growth phenotypes informs genomic selection approaches in a common wheat population. BMC Genomics 22:402

Dilbirligi M, Erayman M, Campbell BT, Randhawa HS, Baenziger PS, Dweikat I, Gill KS (2006) High-density mapping and comparative analysis of agronomically important traits on wheat chromosome 3A. Genomics 88:74–87

Finnegan EJ, Ford B, Wallace X, Pettolino F, Griffin PT, Schmitz RJ, Zhang P, Barrero JM, Hayden MJ, Boden SA, Cavanagh CA, Swain SM, Trevaskis B (2018) Zebularine treatment is associated with deletion of FT-B1 leading to an increase in spikelet number in bread wheat. Plant Cell Environ 41:1346–1360

Fox J, Weisberg S (2019) An {R} companion to applied regression, 3rd edn. Sage, Thousand Oaks. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Gauley A, Boden SA (2021) Stepwise increases in *FT1* expression regulate seasonal progression of flowering in wheat (*Triticum aestivum*). New Phytol 229:1163–1176

Guedira M, Brown-Guedira G, Van Sanford D, Sneller C, Souza E, Marshall D (2010) Distribution of *Rht* genes in modern and historic winter wheat cultivars from the Eastern and Central USA. Crop Sci 50:1811–1822

Halliwell J, Borrill P, Gordon A, Kowalczyk R, Pagano ML, Saccomanno B, Bentley AR, Uauy C, Cockram J (2016) Systematic investigation of *FLOWERING LOCUS T*-like Poaceae gene families identifies the short-day expressed flowering pathway gene, *TaFT3* in wheat (*Triticum aestivum* L.). Front Plant Sci 7

He F et al (2019) Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. Nat Genet 51:896–904

Isham K, Wang R, Zhao W, Wheeler J, Klassen N, Akhunov E, Chen J (2021) QTL mapping for grain yield and three yield components in a population derived from two high-yielding spring wheat cultivars. Theor Appl Genet 134:2079–2095

Krasileva KV et al (2017) Uncovering hidden variation in polyploid wheat. Proc Natl Acad Sci U S A 114:E913–E921

Kuzay S et al (2019) Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. Theor Appl Genet 132:2689–2705

Li C, Lin H, Dubcovsky J (2015) Factorial combinations of protein interactions generate a multiplicity of florigen activation complexes in wheat and barley. Plant J 84:70–82

Li C, Lin H, Chen A, Lau M, Jernstedt J, Dubcovsky J (2019) Wheat *VRN1*, *FUL2* and *FUL3* play critical and redundant roles in spikelet development and spike determinacy. Development 146:dev175398

Lin QP, Zong Y, Xue CX, Wang SX, Jin S, Zhu ZX, Wang YP, Anzalone AV, Raguram A, Doman JL, Liu DVR, Gao CX (2020) Prime genome editing in rice and wheat. Nat Biotechnol 38:582–585

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

Maccaferri M et al (2019) Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet 51:885–895

Martinez AF, Lister C, Freeman S, Ma J, Berry S, Wingen L, Griffiths S (2021) Resolving a QTL complex for height, heading, and grain yield on chromosome 3A in bread wheat. J Exp Bot 72:2965–2978

Pearce S, Vanzetti LS, Dubcovsky J (2013) Exogenous gibberellins induce wheat spike development under short days only in the presence of *VERNALIZATION1*. Plant Physiol 163:1433–1445

Poursarebani N et al (2015) The genetic basis of composite spike form in barley and "Miracle-Wheat." Genetics 201:155–165

Quintero A, Molero G, Reynolds MP, Calderini DF (2018) Trade-off between grain weight and grain number in wheat depends on GxE interaction: A case study of an elite CIMMYT panel (CIMCOG). Eur J Agron 92:17–29

Sakuma S et al (2019) Unleashing floret fertility in wheat through the mutation of a homeobox gene. Proc Natl Acad Sci USA 116:5182–5187

Shaw LM, Turner AS, Herry L, Griffiths S, Laurie DA (2013) Mutant alleles of Photoperiod-1 in wheat (*Triticum aestivum* L.) that confer a late flowering phenotype in long days. PLoS ONE 8:e79459

Shaw LM, Lyu B, Turner R, Li C, Chen F, Han X, Fu D, Dubcovsky J (2019) *FLOWERING LOCUS T2* regulates spike development and fertility in temperate cereals. J Exp Bot 70:193–204

Simmonds J, Scott P, Brinton J, Mestre TC, Bush M, Del Blanco A, Dubcovsky J, Uauy C (2016) A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. Theor Appl Genet 129:1099–1112

Taoka K et al (2011) 14-3-3 proteins act as intracellular receptors for rice *Hd3a* florigen. Nature 476:332–325

Walkowiak S et al (2020) Multiple wheat genomes reveal global variation in modern breeding. Nature 588:277–283

Wang W, Pan QL, Tian B, He F, Chen YY, Bai GH, Akhunova A, Trick HN, Akhunov E (2019) Gene editing of the wheat homologs of *TONNEAU1*-recruiting motif encoding gene affects grain shape and weight in wheat. Plant J 100:251–264

Wilhelm EP, Turner AS, Laurie DA (2009) Photoperiod insensitive *Ppd-A1a* mutations in tetraploid wheat (*Triticum durum* Desf.). Theor Appl Genet 118:285–294

Zhang JL, Gizaw SA, Bossolini E, Hegarty J, Howell T, Carter AH, Akhunov E, Dubcovsky J (2018) Identification and validation of QTL for grain yield and plant water status under contrasting water treatments in fall-sown spring wheats. Theor Appl Genet 131:1741–1759

CHAPTER 2

Integration of the *FT-A2* A10 allele for high spikelet number per spike into elite durum wheat varieties and its effect on yield components.

Priscilla Glenn[1], Xiaoqin Zhang[1,2], Jorge Dubcovsky[1,2]

[1] Department of Plant Sciences, University of California, Davis, CA 95616, USA.

[2] Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.

Priscilla Glenn: ORCID 0000-0002-2200-7241
Jorge Dubcovsky: ORCID 0000-0002-7571-4345

AUTHOR CONTRIBUTION STATEMENT

PG collected most of the experimental data and wrote the first version of this manuscript. XZ began the introgression of the *A10* alleles into varieties. JD initiated and coordinated the project, contributed to data analyses, and supervised PG. All authors reviewed the manuscript and provided suggestions.

ACKNOWLEDGEMENTS

This project was supported by the Agriculture and Food Research Initiative Competitive Grants 2017-67007-25939 (WheatCAP) from the USDA National Institute of Food and Agriculture and by the Howard Hughes Medical Institute. We thank Josh Hegarty for his help designing, managing, and harvesting the field.

2.1 ABSTRACT

Previous results revealed that knock-out mutants in *FLOWERING LOCUS T2* (*FT2*) significantly increased spikelet number per spike (SNS), but also decreased fertility. Fortunately, a natural mutation discovered in *FT-A2* that resulted in a change from Aspartic Acid to Alanine at position 10 (D10A), was associated with increased SNS and grain number per spike (GNS) with no significant impacts on fertility. The favorable A10 allele was present in less than 1% of the durum accessions analyzed while in 56% of the common wheat accessions, suggesting that it had undergone positive selection in common wheat. To determine if the A10 allele can be used to increase yield in durum wheat, we introgressed the *FT-A2* A10 allele into four high-yielding genetic backgrounds and evaluated them in replicated field trials. We found that the lines with the A10 allele had significantly higher SNS than the respective sister lines with the D10 allele in all four genetic backgrounds and no significant differences in fertility. The A10 allele was also associated with a higher number of grains per spike, but the differences were significant only for three of the four varieties. Unfortunately, the increases in grain number were offset by significant decreases in kernel weight in all four varieties, resulting in no significant differences in spike weight or grain yield per plot. We hypothesize that the significant negative correlation between grain number and individual kernel weight (r = -0.42, P = 0.0006) reflects source limitations to fill the extra grains in these varieties in this environment. Incorporation of the *FT-A2* A10 allele into high biomass varieties and evaluation of sister isogenic lines in different environments will be necessary to test the usefulness of this allele in durum breeding programs.

2.2 INTRODUCTION

Previous results revealed that knock-out mutants in *FLOWERING LOCUS T2 (FT2)* significantly increased spikelet number per spike (SNS) in wheat with minor effects on days to headings (2-4 days) (Shaw et al., 2019). However, the introgression of these loss-of-function alleles also resulted in reduced fertility that offset and sometimes reverted the positive effect of increased SNS on grain number per spike (GNS) (Shaw et al., 2019). Fortunately, we discovered a natural mutation in *FT-A2* that resulted in a change from Aspartic Acid to Alanine at position 10 (D10A). This non-synonymous mutation was associated with increased SNS and grain number per spike (GNS) with no significant impacts on fertility (Glenn et al., 2022).

The A10 allele was not detected in the diploid ancestor of the A genome (*T. urartu*) nor in the ancestral tetraploids *T. turgidum* subsp. *dicoccoides* and *T. turgidum* subsp. *dicoccon* indicating that D10 is the ancestral allele and A10 the derived one. Interestingly, these *FT-A2* alleles show contrasting frequencies in durum and common wheat, with the A10 allele present in less than 1% of the durum accession and in 56% of the common wheat varieties analyzed (Glenn et al., 2022). The high frequency of the A10 allele in common wheat and its positive effect on grain yield components suggests that this allele has undergone positive selection. The positive effect of the *FT-A2* A10 allele on SNS and GNS in hexaploid wheat was previously shown in a segregating population from the cross between winter lines LA95135 x SS-MBP57 (DeWitt et al., 2021). Reanalysis of the data from that population using the *FT-A2* marker showed no impact on fertility (Glenn et al., 2022). Although TKW was not recorded in that study, average spike yield was significantly increased in two of the three locations (Glenn et al., 2022), indicating that either TKW was not diminished or the impact on yield of the increased GNS was not completely offset by a decrease in TKW.

The A10 allele is present at a extremely low frequency in modern commercial durum wheat varieties, but we previously introgressed this allele into the commercial durum variety Kronos (PI 576168) and tested it in the field in different environments. Although the presence of the A10 allele was consistently associated with increases in SNS, the differences in GNS, TKW, and spike yield varied among experiments (Glenn et al., 2022). These results suggest that the effects of the *FT-A2* polymorphism on these traits was modulated by the environment. The importance of genetic-by-environment interactions between grain number and kernel weight has been highlighted as well in a study of elite CIMMYT lines (Quintero et al., 2018) and a similar observation was reported in *WAPO-A1,* the causal gene of a wheat SNS QTL on chromosome 7AL (Kuzay et al., 2019, 2022). Results indicated that grain yield can be increased by the introgressing of the favorable *Wapo-A1b* allele. However, these increases were only significant when this allele was present in high yielding varieties or high biomass varieties. In addition, if the lines were grown under stressful/non-optimal conditions, the plants did not have enough resources to fill the extra grains associated with the increased SNS, thus resulting in a decrease in TKW and negating any net positive effect on yield (Kuzay et al., 2019).

To further explore the potential of the *FT-A2* A10 allele in durum wheat, we introgressed it into three more modern and highly productive durum wheat varieties including the UC Davis varieties Miwok and Desert Gold and the high yielding breeding line UC1771. We increased seeds of these three pairs of near isogenic lines, together with the Kronos *FT-A2* near isogenic lines, planted them in a field experiment with 8 replications at UC Davis and characterized them for grain yield and different grain yield components. This chapter presents and discusses the results of this experiment.

2.3 MATERIALS & METHODS

Plant Materials

Durum variety Kronos (PI 576168) was developed by Arizona Plant Breeders in 1992 and was the dominant variety planted in California because of its renowned pasta quality (goldenstategrains.com). However, it later became susceptible to stripe rust and was replaced by resistant varieties with similar quality and higher yield. One such line is Desert King (PVP 2005-00187), which was developed by the UC Davis wheat breeding program to be resistant to the current races of leaf rust, stripe rust, and barley yellow dwarf virus with superior yields than Kronos. UC-Desert Gold is the result of the introgression of three loci in Desert King using marker-assisted backcrossing. These genes include the low cadmium allele for the *Cdu1* gene (Wiebe et al., 2010) that was transferred from the UC breeding line UC1848 and two QTL for high yellow pigment content derived from durum variety 'Kofa' (Zhang et al., 2008). Miwok (PVP 2014-00031) was also developed by the UC Davis wheat breeding program from the cross WWW D6523 / UC1308. This variety carries the *Cdu1* allele for reduced cadmium content and has high yield potential. Miwok is grown mainly in the Imperial and San Joaquin Valleys of California. Finally, UC1771 is a durum breeding line developed by the UC Davis wheat breeding program from the cross Kofa/UC1113//Platinum. This is a high yielding line with excellent pasta quality that is being considered as a potential variety release. More extensive descriptions of the UC Davis released varieties can be found at https://fsp.ucdavis.edu/seed-catalog/wheat-varieties.

The durum variety Gredho (PI 532239) from Oman, was used as the donor of the *FT-A2* A10 allele. The F_1 plants were backcrossed to their respective recurrent parents for four generations using the *FT-A2* molecular marker to select the plants carrying the A10 allele (Glenn

et al., 2022). In the B_4F_2 generation, we selected near isogenic sister lines (NILs) homozygous for the A10 and D10 alleles using the same molecular marker.

The $B_4F_{2:3}$ seed of Desert Gold, Miwok, and UC1771 was increased in Tulelake 2021 and the $BC_4F_{2:4}$ grains were used for yield trails at UCD in 2022. Kronos $B_4F_{2:3}$ seed was increased in Davis 2021 was used to plant a $BC_4F_{2:4}$ trial in Tulelake 2021 as reported in Glenn et al 2022. The $BC_4F_{2:5}$ grains were then used for the 2022 UCD trials.

Yield trials were organized in a split-plot design with eight replications, with varieties as the main plots and the *FT-A2* alleles as subplots randomized within each main plot. Average kernel weight from the previous generation was used to plant all plots at the same seed rate of 1.0 M seeds / acre. Plots were 1.37 meters wide and ranged from 3.05 to 3.29 meters in length. Yield was calculated using the calculated plot area and expressed in kg/ha. Four spikes were measured per plot for spikelet number per spike (SNS), grain number per spike (GNS), fertility (grain number / SNS), thousand kernel weight (TKW in g), and spike yield (weight in grams of the grains only).

2.4 RESULTS

In this field experiment the plots were provided abundant irrigation and fertilization which resulted in very high grain yields. These yields varied from 9,353 kg/ha in Kronos to 11,460 kg/ha in UC1171 and were intermediate for Desert Gold (10,298 kg/ha) and Miwok (10,254 kg/ha). These results suggest that plant growth was not limited by major abiotic or biotic stresses.

We first explored the correlations among the different traits that are summarized in Table 1. We observed highly significant ($P < 0.001$) positive correlations among SNS, GNS and spike

weight (r = 0.59 to 0.76). However, the correlations between these three parameters and total grain yield were weaker (r = 0.111 to 0.285) and either less significant ($P < 0.05$ for GNS and spike weight) or not significant (for SNS). This may be explained, at least in part, by the highly significant and negative correlation observed between GNS and TKW (r = -0.417, $P < 0.001$) which contributed to a negative correlation between SNS and TKW as well (r = -0.325). Fertility was not correlated with TKW or SNS but it showed a positive and significant correlation with GNS (r = 0.597) and Spike Yield (r = 0.284), but not with Plot Yield (NS).

Taken together, these results suggest that increases in both SNS and Fertility contributed to the increases in GNS, but also that increases in grain number were associated with a decrease in individual kernel weight.

Table 1. Correlation between yield component traits. Negative correlations are red. ns = not significant * = P < 0.05, ** = P < 0.01, *** = P < 0.001

CORREL	SNS	GN	FERT	TKW	Spike Yield	Plot Yield
SNS		0.771	-0.047	-0.325	0.582	0.111
GN	***		0.597	-0.417	0.762	0.284
Fert	ns	***		-0.253	0.458	0.285
TKW	*	***	ns		0.262	-0.033
Spike Yield	***	***	***	*		0.247
Plot Yield	ns	*	*	ns	*	

We then performed a randomized complete block design split-plot ANOVA for the six traits across the four varieties summarized in Table 2. We found significant differences in *FT-A2* for SNS, GNS and TKW, but not for Fertility, Spike Yield or Plot Yield. The A10 allele was associated with positive effects on SNS (+13.5%) and GNS (+10.4%), but with negative effects of the same magnitude in TKW (-9.8%). These opposite effects offset each other, contributing to the absence of significant differences in Spike Yield and Plot Yield.

Table 2. ANOVAs for the RCBD- Split Plot field experiment at UC Davis 2022 across the four varieties.

P values	Block	Variety	*FT-A2*	Variety x *FT-A2*	Avg. A10	Avg. D10	% change
SNS	0.3491	<0.0001	<0.0001	0.1055	22.77	20.07	13.5%
GNS	0.8422	<0.0001	<0.0001	0.0034	70.74	64.05	10.4%
Fertility	0.8859	<0.0001	0.1361	0.1495	3.11	3.19	-2.5%
TKW	0.9632	<0.0001	<0.0001	0.0395	53.96	59.84	-9.8%
Spike weight	0.8933	<0.0001	0.9505	0.3308	3.822	3.827	-0.1%
Plot Weight	<0.0001	<0.0001	0.5956	0.0567	10,312	10,369	-0.5%

Some of the traits showed significant interactions between *FT-A2* and genotype so we also analyzed the effects of the A10 and D10 alleles for each of the four varieties separately. Regarding SNS, all the varieties show a significant increase in this trait for the A10 allele, resulting in a non-significant Variety x *FT-A2* interaction (*P* = 0.1055). There were, however, some differences in the magnitude of the differences, which varied from 2.0 (8.7%) additional spikelets in Desert Gold to 3.3 additional spikelets in Miwok (16.4%, **Figure 1A** and **Table 3**).

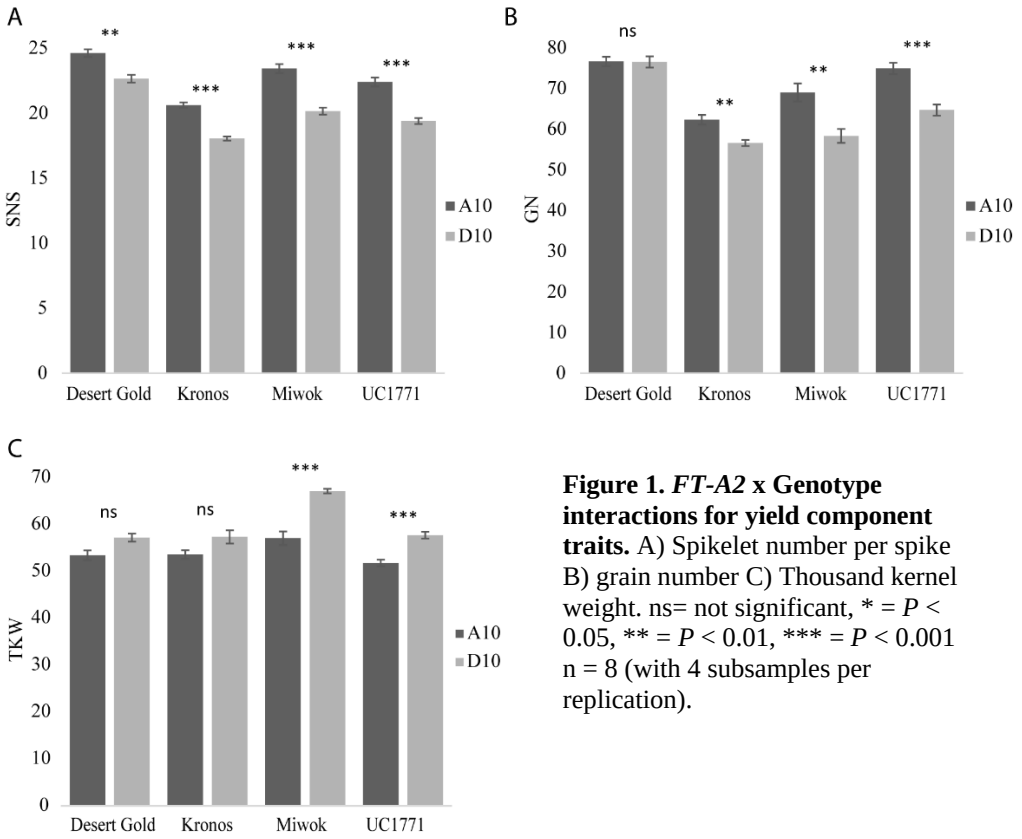


Figure 1. *FT-A2* x Genotype interactions for yield component traits. A) Spikelet number per spike B) grain number C) Thousand kernel weight. ns= not significant, * = *P* < 0.05, ** = *P* < 0.01, *** = *P* < 0.001 n = 8 (with 4 subsamples per replication).

Table 3. Effect of the *FT-A2* A10 allele on yield and yield components in different genetic backgrounds. Plot Yield was measured per plot. All other parameters are the averages of four spikes per plot.

A. Desert Gold ($BC_4F_{2:4}$)	SNS	Plot (kg/ha)	Fertility	GNS	TKW (g)	Spike Yield (g)
A10	24.63	10082.05	3.12	76.66	53.41	4.10
D10	22.66	10513.95	3.38	76.53	57.18	4.38
Diff A10-D10	1.97	-431.90	-0.26	0.13	-3.76	-0.27
%inc/D10	8.7%	-4.1%	-7.7%	0.2%	-6.6%	-6.2%
ANOVA *P*	0.0048	0.0821	0.0604	0.9530	0.0721	0.2213
B. Kronos ($BC_4F_{2:5}$)						
A10	20.63	9396.24	3.03	62.34	53.57	3.35
D10	18.06	9309.93	3.13	56.59	57.35	3.27
Diff A10-D10	2.56	86.31	-0.11	5.75	-3.78	0.08
%inc/D10	14.2%	0.9%	-3.4%	10.2%	-6.6%	2.6%
ANOVA *P*	<0.0001	0.6844	0.3688	0.0184	0.1137	0.6427
C. Miwok ($BC_4F_{2:4}$)						
A10	23.44	10124.09	2.94	69.03	57.06	3.95
D10	20.16	10383.11	2.89	58.34	67.11	3.93
Diff A10-D10	3.28	-259.02	0.05	10.69	-10.05	0.02
%inc/D10	16.4%	-2.5%	1.7%	18.3%	-15.0%	0.6%
ANOVA *P*	0.0001	0.2139	0.6093	0.0039	0.0004	0.9062
D. UC1771 ($BC_4F_{2:4}$)						
A10	22.41	11647.78	3.34	74.94	51.80	3.88
D10	19.41	11272.98	3.34	64.72	57.72	3.74
Diff A10-D10	3.00	374.80	0.01	10.22	-5.92	0.15
%inc/D10	15.5%	3.3%	0.2%	15.8%	-10.3%	3.9%
ANOVA *P*	<0.0001	0.1715	0.9376	0.0004	0.0006	0.2554

Grain number per spike showed a significant interaction between *FT-A2* and genotype (*P* = 0.0034, Table 2) indicating different responses across varieties. Indeed, whereas Kronos, Miwok, and UC1771 showed significant increase in GNS associated with the A10 allele (10.2-18.3%), Desert Gold showed no significant increase in GNS despite its significant increase in SNS. This is likely the result of a combination between the smaller increase in SNS in Desert Gold relative to the other varieties and its larger decrease in Fertility (-7.7%, NS, Figure 1B and Table 3).

Kernel weight also showed a significant *FT-A2* x genotype interaction ($P = 0.0395$, Table 2). Even though the *FT-A2* A10 allele was associated with significant decreases in TKW in all four varieties, the differences were significant in Miwok (-15.0%) and UC1771 (-10.3%) but not in Desert Gold (-6.6%) and Kronos (-6.6%) (Table 3, Figure 1C), contributing to the significant interaction. Consistent with the overall ANOVAS, the separate analyses by variety for Fertility, Spike Yield and Plot Yield failed to reveal any significant differences between *FT-A2* alleles.

Based on the significant interactions between traits and varieties and the differences among varieties described above, we decided to investigate the correlation between traits within varieties. The correlation between SNS and GNS was significant for Kronos, Miwok, and UC1771 ($P = 0.011$, <0.001, and <0.001) but not for Desert Gold (Figure 2A). This may be the result of the slightly reduced fertility associated with *FT-A2* A10 allele in this genotype (Table 3).

The negative correlation between GNS and TKW was only significant in Miwok and UC1771 ($P = 0.002$ and 0.009, Figure 2B) while the negative correlation between SNS and TKW was significant in Kronos, Miwok, and UC1771 ($P = 0.045$, < 0.001, 0.002, Figure 2C). Taken together, these results indicate that the introgression of the *FT-A2* A10 allele in these durum wheat varieties resulted in increases in SNS and GNS that were offset by decreases of the same magnitude in TKW, and that the relative magnitude of these effects varied among varieties.

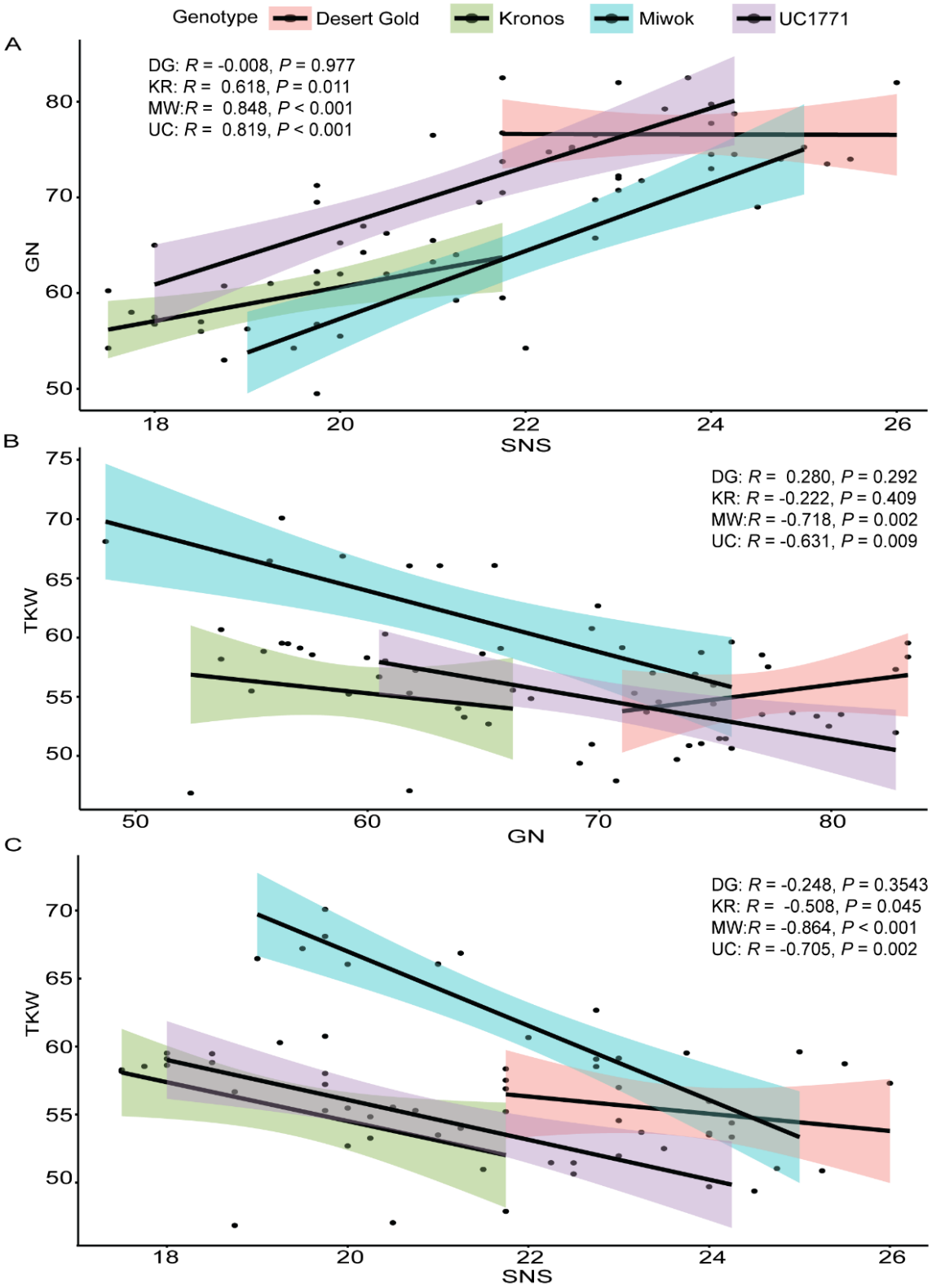


Figure 2. Correlation between yield components traits by varieties. A) Spikelet number per spike x Grain number B) Grain number x Thousand kernel weight C) Spikelet number per spike x Thousand kernel weight.

2.5 DISCUSSION

Wheat breeders select for yield and quality. Of these two, we focused on yield since current rates of increase (0.9%) are insufficient to match the projected growth of the human population by 2050 (Ray et al., 2013). Breeding for yield, however, is complicated due to its multifaceted nature. Yield is determined by multiple genes, along with their epistatic interactions and environmental influences (Reynolds et al., 2012). Total grain yield, however, can be divided into several yield components including total number of spikes/tillers, grains per spike (which can be further separated into spikelet number per spike and grains per spikelet), and thousand kernel weight.

To contribute to our understanding of the genetic networks controlling these grain yield components, our laboratory works on the identification of genes contributing to the regulation of SNS and its effects on other yield components. As part of these efforts, we have previously identified *FT-A2* as an active gene regulating SNS. Loss-of-function mutations in both *FT2* homoeologs result in significant increases in SNS (Shaw et al., 2019), suggesting that *FT2* acts as a promoter of the transition of the inflorescence meristem into a terminal spikelet. According to this hypothesis, the mutants delay this transition, providing more time for the inflorescence meristem to form additional lateral spikelets.

Although the loss-of-function mutations have a positive effect on SNS, they also have a negative effect on fertility, resulting in no gains in GNS (Shaw et al., 2019). These results indicate that *FT-A2* plays additional roles in spikelet or floral development that contribute to fertility and limits the utilization of the loss-of-function mutations in breeding applications. This limitation led to the search for *FT2* natural variants associated with increases in SNS but without

pleiotropic effects on fertility, and to the discovery and characterization of the D10A mutations described in Chapter 1 of this thesis (Glenn et al., 2022).

This study showed that the beneficial A10 allele was not present in the ancestral wheat tetraploid species and was in extremely low frequency (0.7%) in a collection of 417 commercial durum wheat varieties. By contrast the A10 allele was found in more than half of the 705 hexaploid accessions analyzed in the study (Glenn et al., 2022). This rapid increase in gene frequency in the ~10,000 years since the origin of hexaploid wheat suggests that the A10 allele may have a beneficial effect on grain yield or in adaptation to particular environments.

We currently do not know if the A10 allele originated in tetraploid wheat or hexaploid wheat. One possibility is that the A10 allele originated in hexaploid wheat, where its frequency increased rapidly due to positive selection, and was transferred later to a few tetraploid wheats by some rare crosses between hexaploid and tetraploid wheat. The other alternative is that A10 originated in durum wheat and was transferred from there to hexaploid wheat, and that different selection pressures resulted in its increase in hexaploid wheat (based on a positive effect on grain yield), but to its decrease in tetraploid wheat due to selection for large grains and indirectly for fewer grains and reduced SNS. If the first hypothesis is correct, it will implicate a good opportunity to improve grain yield in durum wheat. This motivated our interest in testing the effect of the A10 allele in different durum wheat backgrounds on grain yield components and total grain yield.

Effect of genetic background on *FT-A2* effects on SNS and fertility

It was encouraging to see that the *FT-A2* A10 allele significantly increased SNS in each of the elite varieties it was introgressed into, although the magnitude of the effect of the *FT-A2*

alleles on SNS varied almost 70% among the four genetic backgrounds (from 2.0 to 3.3 additional spikelets). The lowest increase in SNS was observed in Desert Gold, which is the variety that already has the highest number of SNS among the isogenic lines carrying the wildtype D10 allele. Desert Gold has 2.5 more spikelets than Miwok (P =0.003), 3.3 higher than UC1771 ($P < 0.001$) and 4.6 higher than Kronos ($P < 0.001$). We hypothesize that the presence of additional genes increasing SNS in Desert Gold relative to the other genetic backgrounds may have interfered or limited the effect of the *FT-A2* A10 allele on SNS.

Desert Gold showed not only the lowest increases in SNS associated with the A10 allele, but also the highest reduction in fertility associated with this allele (although still not significant). We hypothesize that the reduced effect on SNS together with the higher reduction in fertility contributed to the absence of significant increases in GNS in Desert Gold.

Taken together, these results suggest that the effects of *FT-A2* on SNS and fertility may be modulated by the genetic background.

Effect of *FT-A2* on GNS and TKW

We found that in Kronos, Miwok and UC1771 the A10 allele was associated with significant increases in GNS, but also that those increases did not result in gains in Spike Yield or total Plot Yield due to associated negative effects on kernel weight of a similar magnitude. We discuss below two non-mutually exclusive hypotheses that may explain these results.

The simplest hypothesis is that the observed negative correlation between GNS and TKW was the result of interactions between 'source' (e.g. biomass) and 'sink' (e.g. grains) organs. If our varieties are source-limited, they would be unable to fill the additional grains generated by the A10 allele, leading to smaller and/or wrinkled grains. The varieties tested here have a high

yield potential, as demonstrated by the yields obtained in this experiment, but we still do not know if their yield potential is source- or sink-limited. The negative correlation between grain number and weight may also explain the lower SNS observed in durum wheat compared with common wheat varieties, which correlates with larger grains in durum wheat. If farmers and breeders selected for larger seeds in durum wheat, they would have indirectly selected for reduced SNS in source-limited plants or environments.

An alternative explanation for the limited effects of the A10 allele on Spike Yield or Plot Yield and its low frequency in tetraploid wheat could be that the A10 allele itself has a negative effect on grain development in tetraploid wheat. Gene networks controlling grain size are complex and include multiple genes, so it is possible that strong epistatic interactions among these genes result in the selection of different allele combinations in tetraploid and hexaploid wheat. Under this hypothesis, for *FT-A2* to have a positive effect on tetraploid wheat, it would need to be transferred together with other compatible alleles. This hypothesis may also explain the different frequencies of the A10 allele among different hexaploid classes. We have observed higher frequencies of A10 among the winter wheats (81.7%) than among the spring wheat (44.9%), and within the spring wheats among the accessions developed under fall planting conditions (Glenn et al., 2022), which suggest that the beneficial effects of the A10 allele may be modulated by the genetic background and/or the environment.

The breeding utility of *FT-A2* A10 allele

The lack of significant effects of the A10 introgression on Spike Yield or Plot Yield was a disappointing but not completely surprising result. Although it showed that a simple backcross introgression of the A10 in our durum breeding program may not have the desired positive

impact on grain yield, we cannot rule out positive effects in other genetic backgrounds or environments. Given the high frequency of the A10 allele among the hexaploid winter wheats, it would be interesting to test the effect of the A10 allele introgression in winter durum wheat varieties grown in their local environments.

If the limitations in source were the cause of the negative correlation detected in our experiment between GNS and TKW, future experiments could attempt to introgress the A10 into high-biomass durum varieties or varieties known by other methods to be sink limited. An alternative approach could include the introgression of alleles or genes to increase the source, once those are identified in wheat. However, if the A10 itself interacts poorly with other genes involved in grain development in tetraploid wheat, a possible strategy will be to combine *FT-A2* with other genes from hexaploid wheat known to have a positive effect on grain size.

An interesting combination will be between the A10 allele and the loss-of-function mutants in the A genome homoeolog of the E3 ubiquitin ligase *GRAIN WEIGHT2* (*gw-A2*). The *gw-A2* mutation was shown to be associated with consistent increase in grain size and weight in both tetraploid and hexaploid wheat (Simmonds et al., 2016). Another interesting combination would be the *FT-A2* A10 allele with the natural *GS-A3* allele associated with increased grain size in hexaploid wheat (Yang et al., 2019; Zhang et al., 2014).

In addition, it would be interesting to generate lines combining the A10 allele with alleles from other genes that also result in increases in SNS such as *WAPO-A1b* (Kuzay et al., 2019, 2022), *Elf3* allele from *T. monococcum* (Alvarez et al., 2016), or *COL5* (Zhang et al., 2022) to test their epistatic interactions. Introgression of the *FT-A2* A10 allele (or other genes that increase SNS) may also provide a useful tool to determine in which varieties the yield potential is source or sink limited.

In summary, this study provides valuable information for the utilization of the A10 allele in durum wheat programs, by suggesting that simple backcross introgression of this allele in fall-planted spring durum varieties may have a limited impact on grain yield in spite of the positive effects on SNS and GNS. It also points to missing research areas required for a more complete evaluation of the potential value of the A10 allele to improve grain yield in tetraploid wheat.

CHAPTER 2 REFERENCES

Alvarez, M. A., Tranquilli, G., Lewis, S., Kippes, N., & Dubcovsky, J. (2016). Genetic and physical mapping of the earliness per se locus *Eps-A^m* 1 in *Triticum monococcum* identifies *EARLY FLOWERING 3 (ELF3)* as a candidate gene. *Functional and Integrative Genomics*, *16*, 365–382.

DeWitt, N., Guedira, M., Lauer, E., Murphy, J. P., Marshall, D., Mergoum, M., Johnson, J., Holland, J. B., & Brown-Guedira, G. (2021). Characterizing the oligogenic architecture of plant growth phenotypes informs genomic selection approaches in a common wheat population. *BMC Genomics*, *22*, 1–18.

Glenn, P., Zhang, J., Brown-Guedira, G., Dewitt, N., Cook, J. P., Li, K., Akhunov, E., & Dubcovsky, J. (2022). Identification and characterization of a natural polymorphism in *FT-A2* associated with increased number of grains per spike in wheat. *Theoretical and Applied Genetics*, *135*, 679–692.

Kuzay, S., Lin, H., Li, C., Chen, S., Woods, D. P., Zhang, J., Lan, T., von Korff, M., & Dubcovsky, J. (2022). *WAPO-A1* is the causal gene of the 7AL QTL for spikelet number per spike in wheat. *PLOS Genetics*, *18*, e1009747.

Kuzay, S., Xu, Y., Zhang, J., Katz, A., Pearce, S., Su, Z., & Fraser, M. (2019). Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theoretical and Applied Genetics*, *132*, 2689–2705.

Quintero, A., Molero, G., Reynolds, M. P., & Calderini, D. F. (2018). Trade-off between grain weight and grain number in wheat depends on GxE interaction: A case study of an elite CIMMYT panel (CIMCOG). *European Journal of Agronomy*, *92*, 17–29.

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE*, *8*, 66428.

Reynolds, M., Foulkes, J., Furbank, R., Griffiths, S., King, J., Murchie, E., Parry, M., & Slafer, G. (2012). Achieving yield gains in wheat. *Plant, Cell, & Environment*, *35*, 1799–1823.

Shaw, L. M., Lyu, B., Turner, R., Li, C., Chen, F., Han, X., Fu, D., & Dubcovsky, J. (2019). *FLOWERING LOCUS T2* regulates spike development and fertility in temperate cereals. *Journal of Experimental Botany*, *70*, 193–204.

Simmonds, J., Scott, P., Brinton, J., Mestre, T. C., Bush, M., del Blanco, A., Dubcovsky, J., & Uauy, C. (2016). A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theoretical and Applied Genetics*, *129*, 1099–1112.

Wiebe, K., Harris, N. S., Faris, J. D., Clarke, J. M., Knox, R. E., Taylor, G. J., & Pozniak, C. J. (2010). Targeted mapping of *Cdu1*, a major locus regulating grain cadmium concentration in durum wheat (*Triticum turgidum* L. var *durum*). *Theoretical and Applied Genetics*, *121*, 1047–1058.

Yang, J., Zhou, Y., Zhang, Y., Hu, W., Wu, Q., Chen, Y., Wang, X., Guo, G., Liu, Z., Cao, T., & Zhao, H. (2019). Cloning, characterization of *TaGS3* and identification of allelic variation associated with kernel traits in wheat (*Triticum aestivum* L.). *BMC Genetics*, *20*, 98.

Zhang, W., Chao, S., Manthey, F., Chicaiza, O., Brevis, J. C., Echenique, V., & Dubcovsky, J. (2008). QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. *Theoretical and Applied Genetics*, *117*, 1361–1377.

Zhang, X., Jia, H., Li, T., Wu, J., Nagarajan, R., Lei, L., Powers, C., Kan, C.-C., Hua, W., Liu, Z., Chen, C., Carver, B. F., & Yan, L. (2022). *TaCol-B5* modifies spike architecture and enhances grain yield in wheat. *Science*, *376*, 180–183.

Zhang, Y., Liu, J., Xia, X., & He, Z. (2014). *TaGS-D1*, an ortholog of rice *OsGS3*, is associated with grain weight and grain length in common wheat. *Molecular Breeding*, *34*, 1097–1107.

CHAPTER 3

Identification of bZIPC1 as a protein interactor of FT2 that affects spikelet number per spike in wheat

Priscilla Glenn[1], Daniel Woods[1], Junli Zhang[1], Jorge Dubcovsky[1,2]

[1] Department of Plant Sciences, University of California, Davis, CA 95616, USA.

[2] Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.

Priscilla Glenn: ORCID 0000-0002-2200-7241
Daniel Woods: ORCID 0000-0002-1498-5707
Junli Zhang: ORCID 0000-0001-6625-2073
Jorge Dubcovsky: ORCID 0000-0002-7571-4345

AUTHOR CONTRIBUTION STATEMENT

PG and DW collected most of the experimental data. PG wrote the first version of this manuscript. JZ contributed much of the statistical analysis. JD initiated and coordinated the project, contributed to data analyses, and supervised PG. All authors reviewed the manuscript and provided suggestions.

ACKNOWLEDGEMENTS

This project was supported by the Agriculture and Food Research Initiative Competitive Grants 2017-67007-25939 (WheatCAP) from the USDA National Institute of Food and Agriculture and by the Howard Hughes Medical Institute.

3.1 ABSTRACT

Loss-of-function mutations and natural variation in the wheat gene *FLOWERING LOCUS T2* (*FT2*) have previously been shown to affect spikelet number per spike (SNS). However, while other FT-like wheat proteins genes interact with bZIP-containing transcription factors from the A-group, FT2 does not interact with any of them. In this study, we used a yeast-two-hybrid (Y2H) screen with FT2 as bait and identified a grass-specific bZIP-containing transcription factor from the C-group, designated here as bZIPC1. Within the C-group, we identified four clades including wheat proteins that show Y2H interactions with different sets of FT- and CEN-like encoded proteins. Combined loss-of-function mutations in *bZIPC-A1* and *bZIPC-B1* (*bzipc1*) in tetraploid wheat resulted in a drastic reduction in SNS with a limited effect on heading date. Transcript levels of genes previously shown to affect SNS showed no significant differences between *bzipc1* and wildtype, suggesting that *bZIPC1* may affect SNS through a different pathway. Analysis of natural variation in the *bZIPC-B1* (*TraesCS5B02G444100*) region revealed three major haplotypes (H1-H3), with the H1 haplotype showing significantly higher SNS, grain number per spike and kernel weight per spike than both the H2 and H3 haplotypes. The favorable effect of the H1 haplotype was also supported by its increased frequency from the ancestral tetraploids to the modern durum and common wheat varieties. We developed markers for the two non-synonymous SNPs that result in N151K and V166M amino acid changes that differentiate the *bZIPC-B1b* allele in the H1 haplotype (KM) from the ancestral *bZIPC-B1a* allele present in all other haplotypes (VN). These diagnostic markers will be useful to accelerate the deployment of the favorable *bZIPC-B1b* allele in pasta and bread wheat breeding programs.

3.2 INTRODUCTION

Wheat contributes to almost 20% of the world caloric intake, with approximately 772 million tons produced annually across the globe (fao.org, marketing year of 2020/21). Current rates of yield increase (0.9%) are insufficient to match the projected growth of the human population by 2050 (Ray et al., 2013), which has generated a renewed interest in understanding and improving this trait. However, yield is multifaceted, determined by multiple genes and complex epistatic interactions among them and with the environment that complicate its study (Reynolds et al., 2012). To facilitate its genetic dissection, total grain yield can be divided into several yield components, including number of spikes per surface unit, grains per spike (GNS) and thousand kernel weight (TKW). The number of grains per spike can be further divided into spikelet number per spike (SNS) and grains per spikelet (also known as fertility).

Of the different yield components, SNS has the highest heritability (Zhang et al., 2018) which facilitates its genetic characterization. The number of spikelets per spike is determined when the inflorescence meristem stops producing lateral meristems and transitions into a terminal spikelet. This transition occurs early after the initiation of the reproductive phase, limiting the influence of the environment later in the growing season and resulting in a higher heritability than other grain yield components that are affected by the environment throughout the growing season. Increases in SNS have been associated with increases in grain yield in productive genotypes grown in optimum environments (Boden et al., 2015; Dobrovolskaya et al., 2014; Hai et al., 2008; Kuzay et al., 2019; Wolde et al., 2019). However, in suboptimal environments or in source limited genotypes the increases in SNS are no longer associated with yield increases due to reductions in fertility and/or kernel weight (Kuzay et al., 2019).

In wheat, SNS has been shown to be strongly influenced by flowering time, with genotypes flowering earlier usually having lower SNS than those flowering later (Shaw et al., 2013). Flowering time in wheat is mainly regulated by the vernalization (long exposures to low temperatures providing the competency to flower) and photoperiod pathways (requirement for long days), which converge on the regulation of *FT1* (Distelfeld et al., 2009), the wheat homolog of *Arabidopsis thaliana* (*A. thaliana*) *FLOWRING LOCUS T* (*FT*) and *Oryza sativa* (rice) *HEADING DATE 3* (*HD3*) (Corbesier et al., 2007; Distelfeld et al., 2009; Tamaki et al., 2007). The FT encoded protein, often referred to as florigen, has been shown to be a mobile signal transported through the plant phloem from the leaves to the shoot apical meristem (SAM) in both *A. thaliana* and rice (Corbesier et al., 2007; Tamaki et al., 2007). In wheat, *FT1* is also a strong promoter of flowering that is expressed in the leaves and not in the SAM, so it is also assumed to be a mobile signal in wheat (Lv et al., 2014; Yan et al., 2006). Natural variation in *FT1* in wheat and barley has been associated with differences in both heading time and SNS (Brassac et al., 2021; Chen et al., 2020; Isham et al., 2021; Nitcher et al., 2014; Yan et al., 2006; Zhang et al., 2022). In the absence of *FT1* expression (e.g., under short days in photoperiod sensitive wheats or under continuous cold temperatures), the SAM transitions to the reproductive phase but spike development is arrested, and the spikes fail to emerge from the leaf sheaths or head extremely late.

In rice, it has been shown that HD3 activates the expression of MADS-box meristem identity genes by forming complexes with 14-3-3 and FD-like proteins (Florigen Activation Complex), which bind to the MADS-box gene promoters (Taoka et al., 2011). In wheat, FT1 has also been shown to interact with the bZIP transcription factor FDL1 and a variety of 14-3-3 proteins which are thought to act as a bridge facilitating the interaction between FT and FD-like

proteins (Li et al., 2015; Li & Dubcovsky, 2008). Once FT1 arrives at the SAM, the floral activation complex plays a critical role in accelerating the transition from the vegetative to the reproductive phase by activating floral homeotic genes such as *VERNALIZATION1* (*VRN1*) (Li et al., 2015). In addition, FT1 plays a central role in early spike development and elongation by upregulating the *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1), LEAFY (LFY),* and the genes in the gibberellin (GA) biosynthetic pathway (Pearce et al., 2013). These genes are essential for the elongation of the stem and emergence of the spike, which is called heading (Pearce et al., 2013). FT1 also interacts with *TEOSINTE BRANCHED1* (*TBI*), which impacts the formation of paired spikelets but has limited effects on wheat heading time (Dixon et al., 2018).

The expansion of the *FT*-like gene family in monocot and eudicot species has led to the diversification of *FT* functions (Ballerini & Kramer, 2011; Chardon & Damerval, 2005; Higgins et al., 2010; Lv et al., 2014). For example, in beet (*Beta vulgaris* ssp. *vulgaris*), the regulation of flowering time is controlled by two *FT* paralogs with antagonistic functions. *BvFT2* is functionally conserved with *FT* and essential for flowering, while *BvFT1* represses flowering and its down-regulation is crucial for the vernalization response (Pin et al., 2010). In potato an FT-like protein StSP6A was identified as a major component of potato tuber formation (Teo et al., 2017); whereas in *Brachypodium distachyon* (*B. distachyo*n), *FTL9* (closely related to *FT3/FT5* in wheat) is essential for short day (SD) vernalization, which provides un-vernalized winter plants competency to flower after exposure to SD (Woods et al., 2019). In wheat, twelve different *FT*-like genes have been identified (Halliwell et al., 2016; Lv et al., 2014) and examples of sub-functionalization have been also described. For example, *FT3* and *FT5* have a more critical role in the induction of flowering under SD than other LD expressed *FT*-like genes

(Casao et al., 2011; Faure et al., 2007; Kikuchi et al., 2009; Lv et al., 2014; Zikhali et al., 2017). In addition, *FT2* was shown to have a more predominant influence on SNS than on heading time (Glenn et al., 2022; Shaw et al., 2019).

The functional differences in *FT2* are particularly interesting, because this gene is the closest paralog of *FT1* (78% identical at the protein level) and the only *FT-like* gene in wheat that has been shown to be transcribed directly in the SAM and developing spike (Shaw et al., 2019). Loss-of-function mutations in both homoeologs of *FT2* (henceforth *ft2*) result in small differences in heading time but significant increases in SNS. However, the *ft2* mutant also shows reduced fertility precluding its utilization in breeding applications (Shaw et al., 2019). Fortunately, a natural variant of *FT-A2*, resulting in an Aspartic Acid (D) change to Alanine (A) at position 10 (henceforth, D10A), has been recently associated with increases in SNS, grain number per spike and total spike weight indicating no negative impacts on fertility (Glenn et al., 2022). The beneficial A10 allele was almost absent in tetraploid wheat but was present in more than half of the analyzed common wheat varieties suggesting that this allele has been favored by selection for improved grain yield in common wheat (Glenn et al., 2022).

The functional differentiation of *FT1* and *FT2* was paralleled by differences in their protein-protein interactions. Whereas FT1 was able to interact in yeast-two-hybrid (Y2H) assays with five of the six tested 14-3-3 proteins, FT2 showed no interactions with any of them (Li et al., 2015). The two proteins also differed in their interactions with FDL proteins. Specifically, FT1 showed positive Y2H interactions with FDL2 and FDL6, in contrast FT2 showed a positive interaction with FDL13 (Li & Dubcovsky, 2008). However, later it was found that FDL13 was an alternative splice form of FDL15 with a retained intron and a premature stop codon (Li et al.,

2015), and the complete FDL15 did not interact with FT2. In summary, no protein interactors of FT2 were identified before this study.

Here, we explored the proteins FT2 can interact with by performing Y2H screens and identified a bZIP-containing transcription factor from the grass-specific C-group that we designated as bZIPC1. We show that bZIPC1 and three other members of the C-group can interact with FT- and CEN-like encoded wheat proteins, a function that was previously shown only for bZIP members of the A-group (FDL2, FDL6 and FDL15) (Li et al., 2015). Functional characterization of bZIPC1 mutants revealed that this gene impacts spikelet number per spike but has limited effect on heading time. Finally, we characterized natural variation in this gene and identified variants associated with differences in SNS that may be of value for wheat breeding applications.

3.3 RESULTS

Identification of FT2 interactors by yeast-two-hybrid (Y2H) screening

To identify proteins that interact with FT2, we performed Y2H screens (details described in the Material and Methods section) using the wheat Chinese Spring FT-A2 protein (TraesCS3A02G143100) as bait and a *B. distachyo*n photoperiod cDNA library developed from harvesting whole plants across a diurnal cycle in long and short photoperiods (Cao et al., 2011) as prey. In the three screens performed against the photoperiod library (transformation efficiencies of $4.5x10^5$, $4.6X10^6$ and $4.1x10^6$), we detected 121 positive colonies corresponding to 26 different genes (Table S1). Among them the one represented by the largest number of colonies (67) was *Bradi1g35230* (annotated as a kinectin-related protein with motor activity). The second most abundant gene from the photoperiod library screens was *Bradi1g05480*, which was detected in 29 colonies. This gene was annotated as an uncharacterized bZIP transcription

factor. We decided to prioritize this bZIP for functional analyses because it was the only interactor detected in all three Y2H screens and also because of the known protein-protein interactions between FT1 and other distantly related bZIP proteins (Li et al., 2015). The complete list of detected interactors identified in the Y2H screens, and their functional annotation is presented in Table S1.

Phylogenetic analyses of bZIPC-like genes

The large family of bZIP domain transcription factors includes 14 subfamilies (Agarwal et al., 2019; Guedes Corrêa et al., 2008; Peviani et al., 2016). Reciprocal BLASTs between Bradi1g05480 and *A. thaliana* did not reveal a clear one-to-one orthologous relationship but were sufficient to place Bradi1g05480 within the C-group (Fig. 1), which includes *A. thaliana* genes *bZIP63* (At5g28770), *bZIP10* (At4g02640), and *bZIP25* (At3g54620). To reflect Bradi1g05480 phylogenetic relationship within the C-group of this large and diverse family of bZIP transcription factors, we decided to incorporate the letter identifying the bZIP group within the gene name and designated this gene as *bZIPC1*.

To understand the evolutionary history within the bZIP C-group and identify the closest genes to *Bradi1g05480* in other species, we selected 58 bZIP-C class proteins spanning flowering plant diversification from BLASTN searches (Fig. 1). Relationships among these bZIP-C class proteins were estimated using Bayesian phylogenetic methods on an alignment of the highly conserved bZIP DNA binding motif and other conserved regions within the protein (Fig. S1). Using a Bayesian phylogenetic analysis, we estimated a >0.95 posterior probability for the *bZIPC1* clade, which comprised only grass species (Fig. 1). The closest clade to *bZIPC1*, which we are designating as *bZIPC2*, also comprised only grass species. (Fig. 1). The nearest grass relative, *Joinvillea ascendens*, showed a single sequence associated with the grass *bZIPC1*

and *bZIPC2* clades, suggesting that these two clades are the result of a grass specific duplication event, which is consistent with the known whole genome duplication event at the base of the grasses (Paterson et al., 2004).

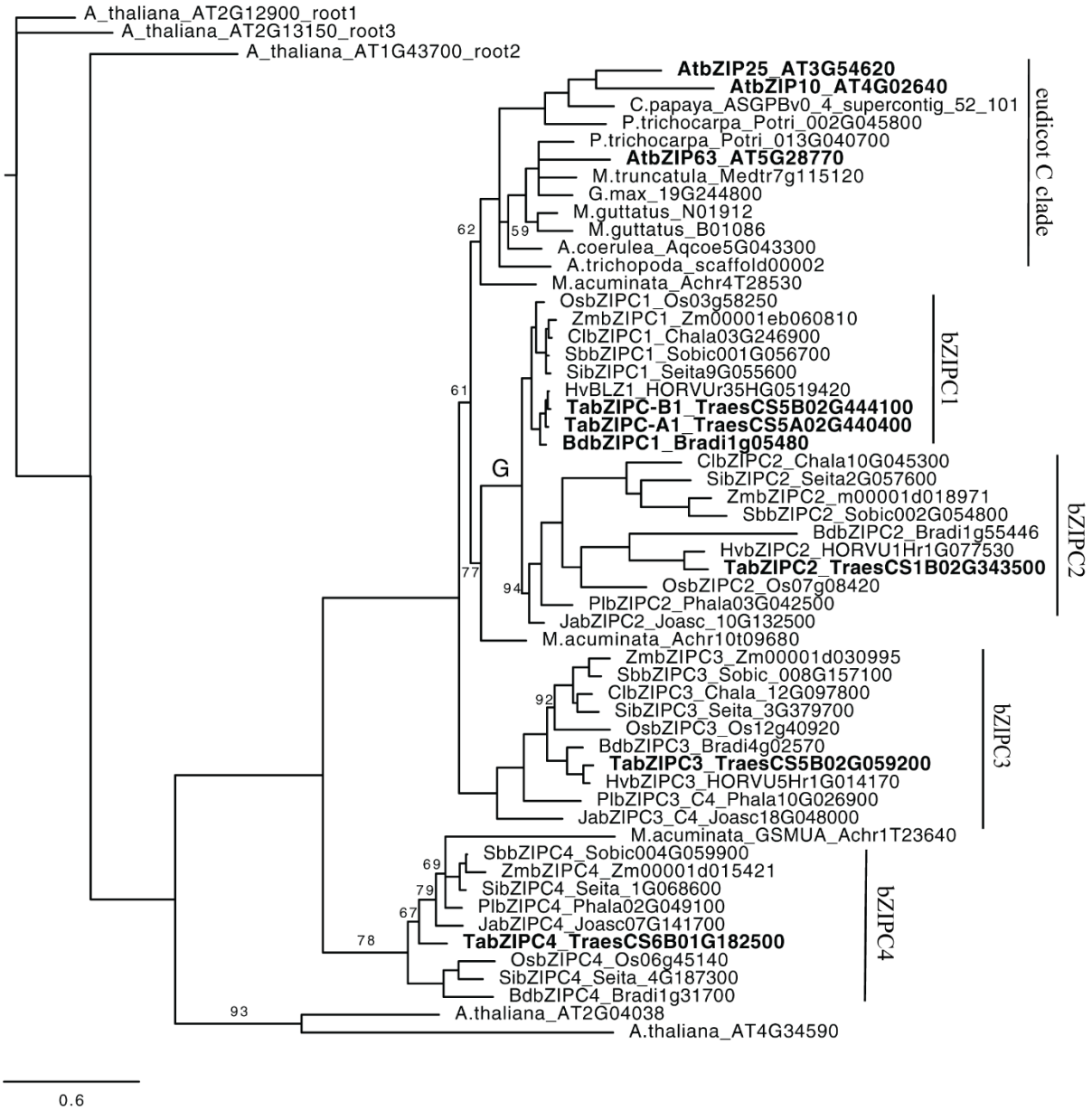


Fig. 1: Bayesian consensus phylogram of 58 group C-group bZIP domain containing proteins based on 169 alignable amino acids. Posterior probabilities (PP) are indicated only for the branching points supported at PP values <98. All other branching points are supported at PP values > 98. "G" above the branch indicates a grass-specific duplication. Scale bar indicates substitutions per site, excluding the root node which is shrunken for viewability. Genes discussed in the manuscript are labeled in bold font. At=*Arabidopsis thaliana*, Os=*Oryza sativa*, Zm=*Zea mays*, Cl=*Chasmanthium laxum*, Sb=*Sorghum bicolor*, Si=*Setaria italica*, Hv=*Hordeum vulgare*, Ta=*Triticum aestivum*, Bd= *Brachypodium distachyon*, Pl=*Pharus latifolius*, Ja=*Joinvillea ascendens*.

The bZIP C-group includes two additional monocot-specific clades, designated here as *bZIPC3* and *bZIPC4*, which are more distantly related to *bZIPC1* and *bZIPC2* (Fig. 1). The closest *A. thaliana* bZIP members of the C-group to these four monocot-specific clusters, *bZIP10*, *bZIP25* and *bZIP63*, cluster together with other eudicot members of this group suggesting independent duplications events in the *A. thaliana* and the grass lineages (Fig. 1). Thus, while *bZIPC1* shares sequence similarity to *bZIP63*, *bZIP25* and *bZIP10* from *A. thaliana*, these *bZIPC* duplicated genes have independent evolutionary histories in the grasses and the Brassicales.

Interactions between bZIPC-like and FT-like proteins

The *B. distachyon* clones identified in the Y2H library were not full-length genes based on BLASTN and BLASTP comparisons. To validate the interaction with full length bZIPC1, we conducted direct Y2H assays using the full-length wheat bZIPC1 protein (TraesCS5A02G440400.1, 392 amino acids). Since there are two different natural variants in FT2 known to impact SNS (D10 and A10) (Glenn et al., 2022), we tested both protein variants in the direct Y2H assays. We found that bZIPC1 can interact with both FT2 protein variants in yeast when used as either bait or prey, confirming the initial Y2H screen results (Table 1; Fig. 2*)*.

To explore if bZIPC1 can interact with other FT-like proteins, we also performed pairwise Y2H assays between bZIPC1 and FT-A1 (TraesCS7A02G115400), FT-B3 (TraesCS1B02G351100), FT-B5 (TraesCS4B02G379100), CEN-B2 (TraesCS2B02G310700), CEN-A4 (TraesCS4A02G409200) and CEN-A5 (TraesCS5A02G128600). We found that in addition to FT2, bZIPC1 can interact with FT3 and weakly with FT5, however we did not observe positive interactions with any of the other FT-like genes tested (Table 1, Fig. 2).

Table 1: Y2H screen results between bZIPC1 and paralogs to FT-like proteins.

Bait pDEST32	Prey pDEST 22			
	bZIPC1	**bZIPC2**	**bZIPC3**	**bZIPC4**
FT1	NO	WEAK	NO	WEAK
FT2-D10	YES	WEAK	NO	NO
FT2-A10	YES	NA	NA	NA
FT3	YES	WEAK	YES	NO
FT5	WEAK	YES	WEAK	NO
CEN2	NO	YES	NO	NO
CEN4	NO	WEAK	NO	WEAK
CEN5	WEAK	NO	YES	NO

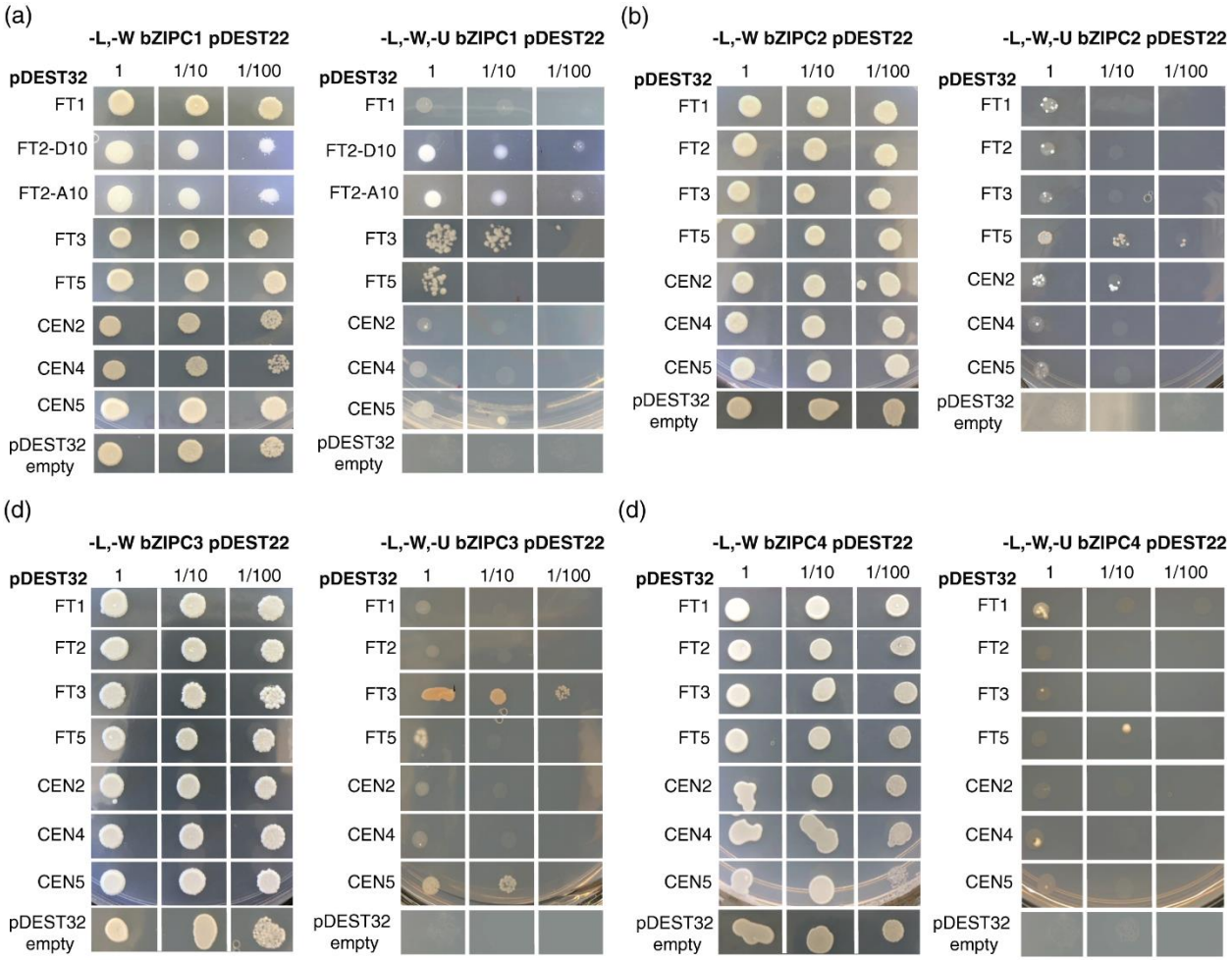


Fig. 2: Yeast-two-hybrid (Y2H) interactions. a) bZIPC1 (b) bZIPC2, (c) bZIPC3, (d) bZIPC4 and FT1, FT2-D10, FT2-A10 (bZIPC1 only), FT3, FT5, CEN2, CEN4, CEN5 and empty pDEST32 (auto-activation test). Left panel: SD medium lacking Leucine and Tryptophan (-L-W) to select for yeast transformants containing both bait and prey. Right panel: interaction on -L-W-U medium (no L, W, and U (Uracil). Dilution factors = 1, 1:10 and 1:100.

bZIPC-B2 was synthesized and tested alongside *bZIPC-A1* in a large set of Y2H pairwise tests. We found that bZIPC2 interacted strongly with FT5 and CEN2 while interacting only weakly with the other genes, excluding CEN5, which showed no detectable interaction. bZIPC3 only interacted with FT2 and CEN5 and weakly with FT5; whereas the bZIPC4 protein only interacted weakly with FT1 and CEN4. (Table 1, Fig. 2).

Loss of function mutations in bZIPC1 are associated with reduced spikelet number per spike (SNS)

To determine the function of *bZIPC1*, we combined loss-of-function mutants in the two homoeologs of this gene in tetraploid wheat Kronos. We identified one tilling mutant line with a loss-of-function mutation in the A genome homoeolog *bZIPC-A1* (*TraesCS5A02G440400*) and three in the B genome homoeolog *bZIPC-B1* (*TraesCS5B02G444100*). The G to A mutation in *bZIPC-A1* (CS RefSeq v.1, chromosome 5A position 621,763,602) in the Kronos mutant K3308 results in a premature stop codon at position 97 (Q97*). In *bZIPC-B1*, surprisingly, we identified two independent mutants with identical C to T mutations (K3532 and K2991) in chromosome 5B position 616,654,824 at the acceptor splicing site at the end of intron one (Fig. 3a). To confirm that this mutation does impact the splicing of the *bZIPC-B1* gene, we sequenced the mRNA products of both *bZIPC-B1* mutant alleles in K2991 and K3532. We found that the splice site mutations resulted in the second exon (94 bp) being spliced out, generating a reading frame shift resulting in a premature stop codon (Fig. 3b). We also identified a premature stop codon in Kronos mutant line K2038, where a C to T mutation in chromosome 5B at position 616,655,991 results in a premature stop codon (W117*).

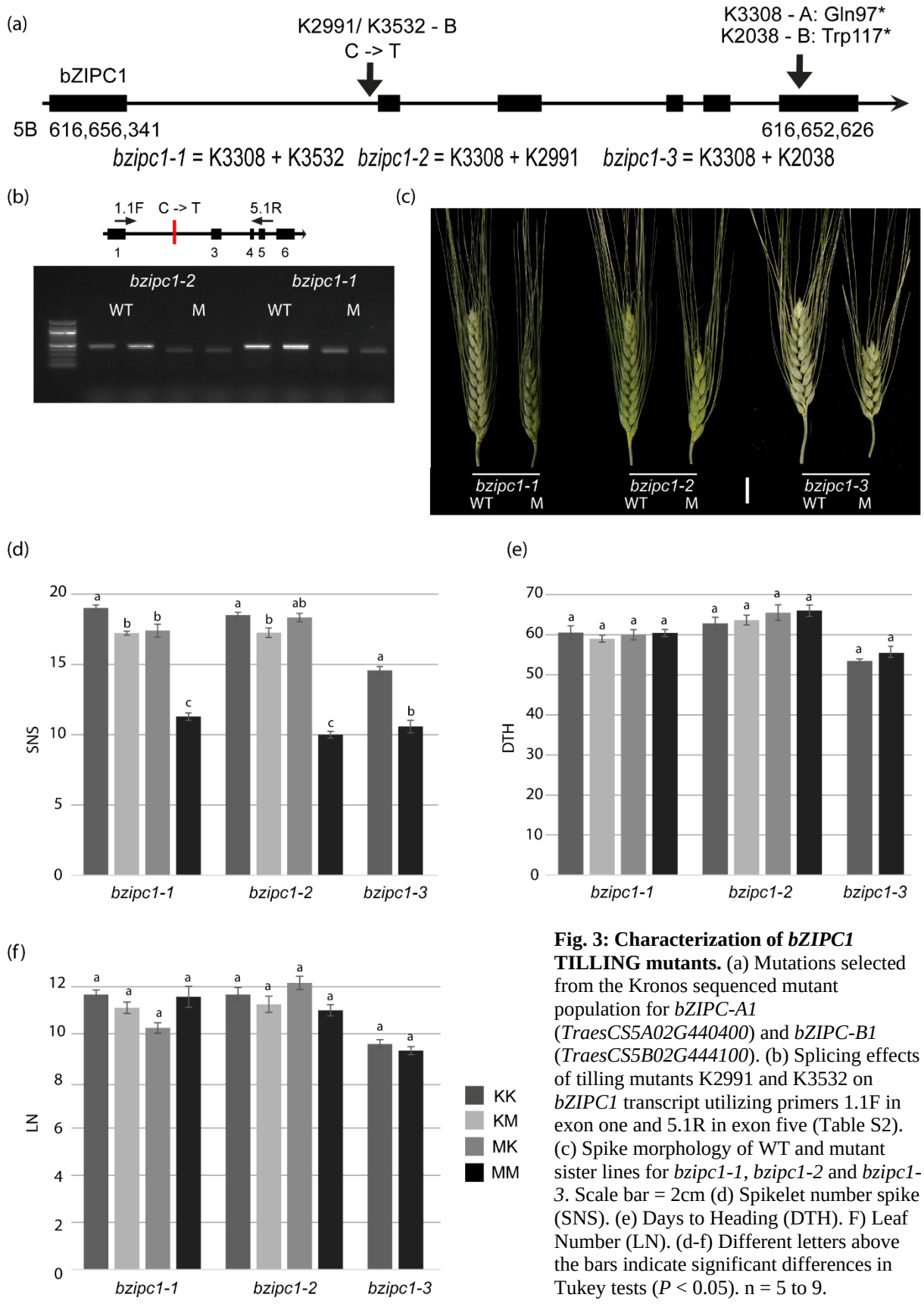


Fig. 3: Characterization of *bZIPC1* TILLING mutants. (a) Mutations selected from the Kronos sequenced mutant population for *bZIPC-A1* (*TraesCS5A02G440400*) and *bZIPC-B1* (*TraesCS5B02G444100*). (b) Splicing effects of tilling mutants K2991 and K3532 on *bZIPC1* transcript utilizing primers 1.1F in exon one and 5.1R in exon five (Table S2). (c) Spike morphology of WT and mutant sister lines for *bzipc1-1*, *bzipc1-2* and *bzipc1-3*. Scale bar = 2cm (d) Spikelet number spike (SNS). (e) Days to Heading (DTH). F) Leaf Number (LN). (d-f) Different letters above the bars indicate significant differences in Tukey tests ($P < 0.05$). n = 5 to 9.

We designed KASP markers for each of the EMS induced mutations (Table S2) to trace the mutations during crosses and backcrosses. We made crosses between (K3308) (A genome premature stop) and the three different B genome mutations and designated the different combined homozygous mutant alleles as *bzipc1-1* (K3308 + K3532), *bzipc1-2* (K3308+ K2991), and *bzipc1-3* (K3308 + K2038) (Fig. 3a). Plants heterozygous for the mutant alleles in both homoeologs in the F_1 hybrids and homozygous in their progeny were confirmed in the three mutant combinations using the KASP markers (Table S2).

Visual comparison of the spikes (Fig. 3c) showed no obvious differences in the spikelets between the combined mutants and the WT. However, all three double mutants had significantly reduced SNS in two independent experiments compared to the WT, with an average reduction of 5.8 spikelets in the greenhouse experiment and 6.9 in the growth chamber ($P < 0.001$, Fig. 3d). We found no significant differences in days to heading (DTH) or leaf number (LN) in both experiments, indicating that *bZIPC1* has a limited effect on both the timing of the transition between the vegetative and reproductive phase stages and the duration of the elongation phase (Fig. 3e-3f). The single *bzipc-A1* and *bzipc-B1* mutants showed no significant differences with the wild-type for SNS, DTH, and LN phenotypes (Fig. 3d-f), suggesting that both homoeologs have redundant functions. The F_2 results are presented in Fig. 3d-f, and their validation using F_3 plants in Fig. S2.

Expression of *bZIPC1* and *FT2* overlap in the developing spikelet

We took advantage of published RNAseq data in Chinese Spring from five different tissues at three developmental stages (Choulet et al., 2014) to evaluate the expression levels of

the different bZIPC1 homoeologs. In CS, *bZIPC1* transcripts were detected across all tissues for the three homoeologs (Fig. 4a).

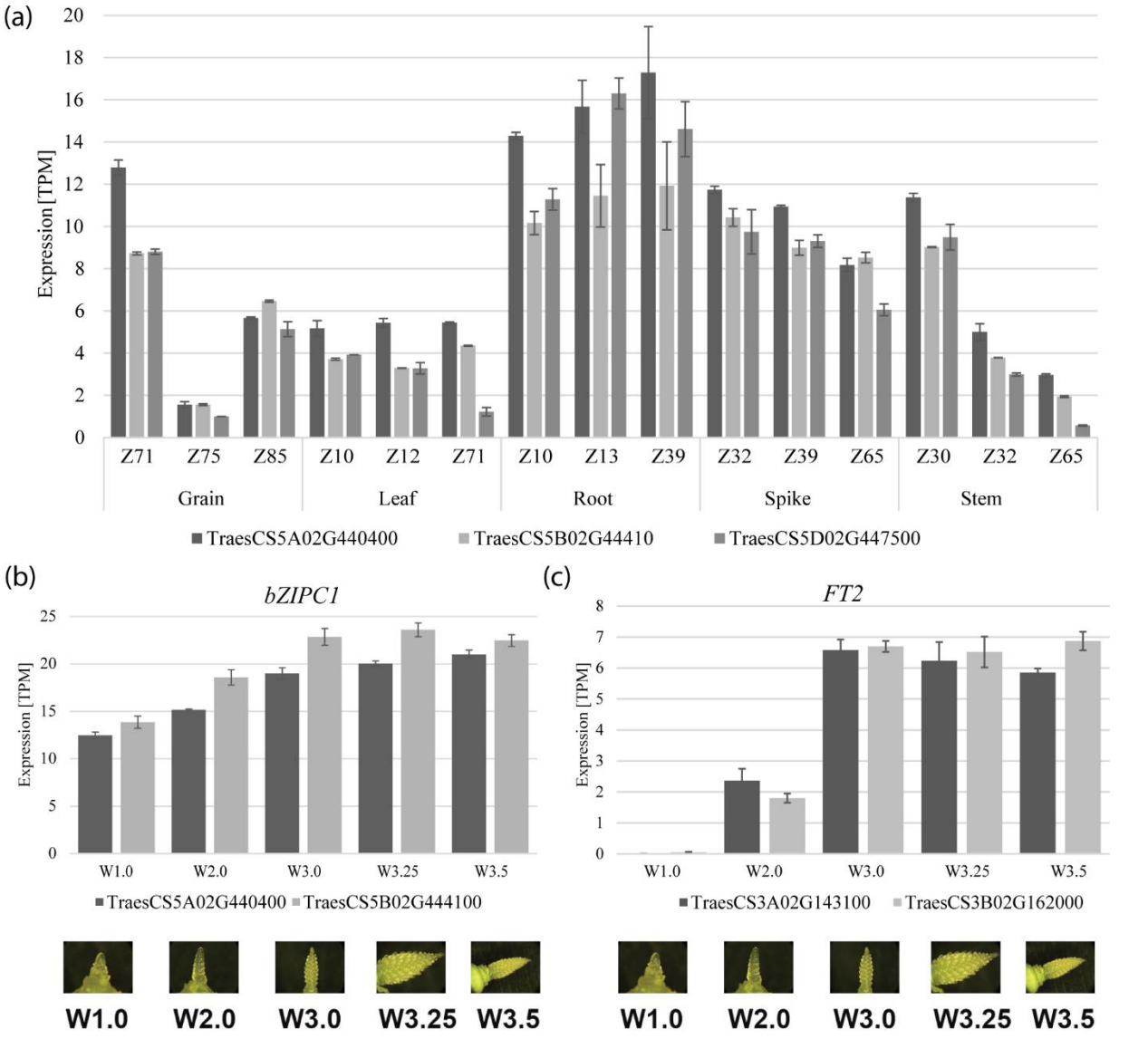


Fig. 4: Relative expression of *bZIPC1* across tissues. (a) Transcripts per million (TPM) of the three *bZIPC1* homoeologs from previously published Chinese Spring RNAseq expression across five tissues and three developmental stages (Choulet et al., 2014). *TraesCS5A02G44044 = bZIPC-A1*, *TraesCS5B02G444100 = bZIPC-B1*, and *TraesCS5D02G447500 = bZIPC-D1*. (b) Expression of *bZIPC-A1* and *bZIPC-B1* in tetraploid wheat Kronos across the early stages of spike development (n = 4). (c) Expression of *FT-A2* and *FT-B2* in tetraploid wheat Kronos across the early stages of spike development (n = 4).

Root tissue had the highest expression levels of *bZIPC1*, which was fairly consistent across the three developmental stages included in the study. The spike tissue also showed high

and relatively consistent expression levels of all three *bZIPC1* homoeologs across the three timepoints, which were all collected after the initiation of the elongation phase (Zadoks scale Z32, Z39 and Z65 (Zadoks et al., 1974). *bZIPC1* showed a relatively constant low expression level in the leaves, while in the stem the expression was comparable to that in the spike at stage Z30, although it decreased at later stages (Z32 and Z65). In the grain, *bZIPC1* was expressed at stage Z71 but then decreased greatly in Z75 before rising again slightly in Z85.

To study the expression of *bZIPC1* earlier in spike development, we took advantage of an RNAseq study in tetraploid wheat Kronos including samples collected at the vegetative (W1.0), double ridge (W2.0), glume primordia (W3.0), lemma primordia (W3.25), and floret primordia (W3.5) stages (VanGessel et al., 2022), with W values based on the Waddington spike developmental scale (Waddington et al., 1983). Overall expression was similar between the A and B genome, with expression levels increasing from W1.0 to W3.0 and remaining constant through W3.5 (Fig. 4b). *FT2* expression levels in the same RNAseq samples were extremely low in W1.0 but then rose significantly in W2.0 to W3.0 before stabilizing (Fig. 4c). Thus, *bZIPC1* and *FT2* are expressed in the same organ at the same developmental stages.

Previously, *FT2* has been shown to have different transcription profiles under LD and SD. Under LD, *FT2* expression in the leaves of seven-week-old plants was relatively high with a peak close to 3-fold *ACTIN* at Zeitgeber time 8 (ZT8= 8 h after the lights were turned on), while under SD *FT2* expression was low throughout the diurnal cycle (Shaw et al., 2019). To explore if *bZIPC1* expression changes between LD and SD are similar to *FT2*, we performed a real-time quantitative reverse transcription PCR (qRT-PCR) analysis on *bZIPC1* in leaf samples collected every four hours throughout the day from tetraploid wheat Kronos plants grown under SD (8h light and 16 h darkness) and LD conditions (16 h light and 8 h darkness). Transcript profiles of

bZIPC1 under LD and SD were more similar to each other than in *FT2* (Shaw et al. 2019). Transcript levels at both photoperiods were similar at ZT0, decreased during the day and increased during the night, although they reached higher levels in the last night time point under the long nights of SD than under the shorter nights of LD (Fig. 5a and b).

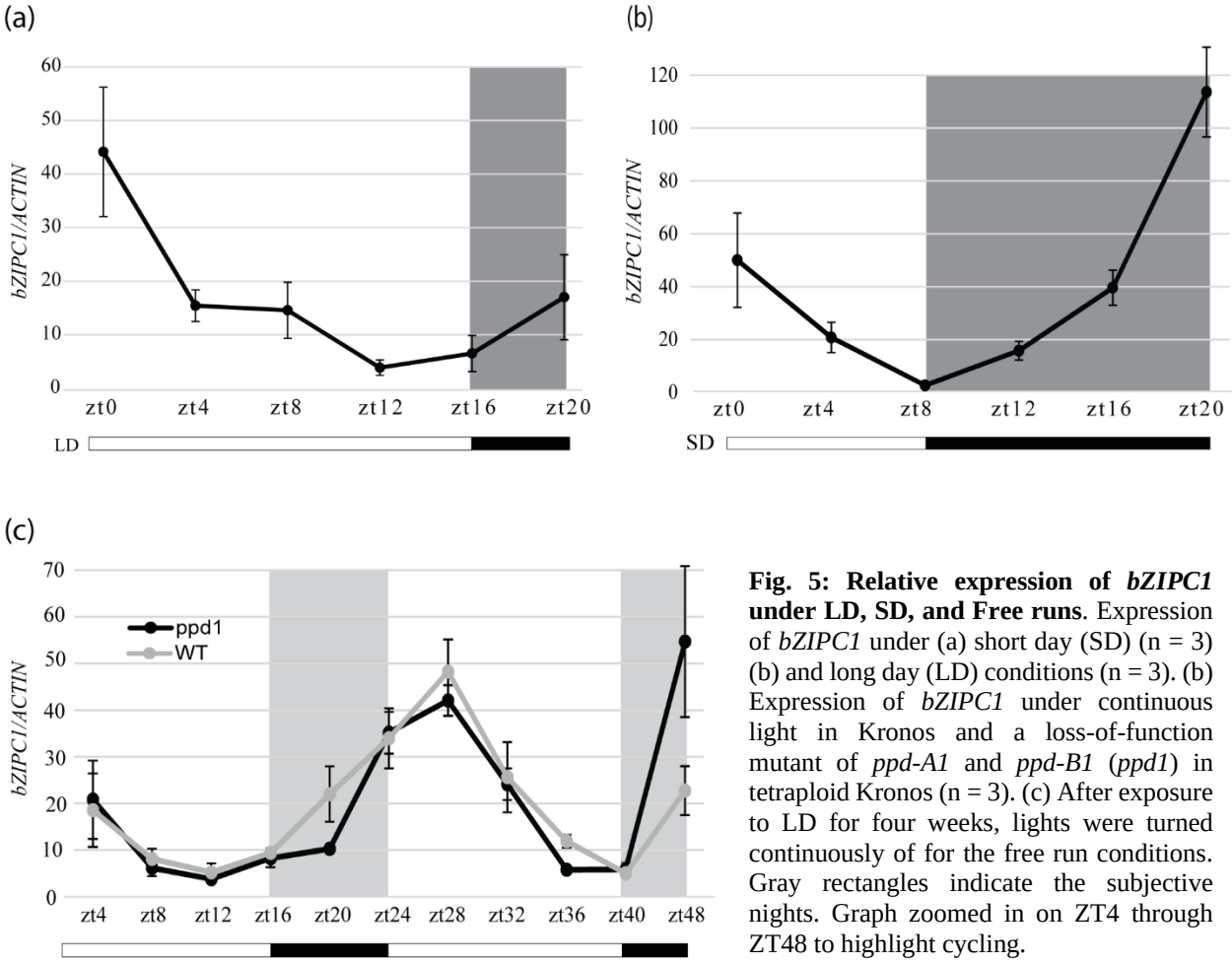


Fig. 5: Relative expression of *bZIPC1* under LD, SD, and Free runs. Expression of *bZIPC1* under (a) short day (SD) (n = 3) (b) and long day (LD) conditions (n = 3). (b) Expression of *bZIPC1* under continuous light in Kronos and a loss-of-function mutant of *ppd-A1* and *ppd-B1* (*ppd1*) in tetraploid Kronos (n = 3). (c) After exposure to LD for four weeks, lights were turned continuously of for the free run conditions. Gray rectangles indicate the subjective nights. Graph zoomed in on ZT4 through ZT48 to highlight cycling.

To determine if the diurnal expression cycling of *bZIPC1* is clock regulated as has been shown in other *bZIPC*-like genes in *A. thaliana* (Frank et al., 2018), we collected leaf samples every four hours from Kronos plants under continuous light for 48 hours. Plants were grown under LD for four weeks and then transferred to continuous light (free running). *bZIPC1* showed circadian oscillations with a peak at ZT28 h after the initiation of the continuous light period

(Fig. 5c). We also included in the study samples of a loss-of-function mutant of *PHOTOPERIOD 1* (combined *ppd-A1* and *ppd-B1*, henceforth *ppd1*) in the same Kronos background, since this gene is a known positive regulator of *FT2* under LD. However, there were no significant differences between the WT and *ppd1* mutants at any time point, suggesting that while *bZIPC1* is regulated by the circadian clock, its transcript levels are not impacted by *PPD1* (Fig. 5c), which agrees with the limited differences between LD and SD described above (Fig. 5 a and b).

Effect of the *bzipc1* mutant on the transcript levels of genes involved in spike development

Since the *bzipc1* mutants exhibited a significant reduction in SNS, we compared the expression of genes previously known to be involved in the control of SNS in two of the three mutants (*bzipc1-1* and *bzipc1-2*) and their respective wildtype sister lines. Specifically, we tested the expression of *VRN1, FUL2, FUL3, WAPO1, LFY, PPD1* (= *PRR37*), *FT2, SEP1-6* (ortholog of rice *MADS34*), *CEN2, CEN4, CEN5, PRR73, FLC2* (*MADS51*), and *AG1* (ortholog of rice *MADS58*) via qRT-PCR at the double ridge (W2.5) and glume primordia (W3.0) developmental stages. We did not find significant differences for any of these genes that were consistent between loss-of-function mutants *bzipc1-1* and *bzipc1-2*, (Table 2, Fig. 6a). These results suggest that *bZIPC1* is not controlling SNS through the modulation of expression of these genes at the tested stages. However, we cannot currently rule out effects on the expression of these genes at other spike developmental stages.

Table 2: Transcript levels of genes known to be involved in spike development in *bzipc1* mutant and wild type (WT) sister lines. Transcript levels were calculated using the $2^{\Delta\text{-}Ct}$ method using *IF4A* as endogenous control.

		bzipc1-2			*bzipc1-1*		
Gene	Stage	Mutant	WT	*P* value	Mutant	WT	*P* value [a]
LFY	W2.5 [b]	0.002713	0.004204	0.06	0.002814	0.002562	ns
	W3.0	0.002235	0.003983	0.07	0.00505	0.005461	ns
FUL2	W2.5	0.008892	0.011719	ns	0.014926	0.005968	0.07
	W3.0	0.016907	0.045001	*	0.021374	0.019443	ns
VRN1	W2.5	0.093737	0.100668	ns	0.13067	0.099647	ns
	W3.0	0.073094	0.130994	0.09	0.154238	0.114678	ns
MADS34	W2.5	0.00408	0.009783	*	0.005794	0.003841	ns
	W3.0	0.010519	0.017487	0.08	0.01408	0.01284	ns
CEN2	W2.5	0.029902	0.04925	ns	0.043705	0.02868	0.07
	W3.0	0.012279	0.021746	ns	0.026776	0.021089	ns
CEN5	W2.5	0.044676	0.041554	ns	0.081628	0.021158	*
	W3.0	0.020352	0.028069	ns	0.045873	0.033324	ns
CEN4-A	W2.5	0.015671	0.011644	ns	0.013048	0.006435	ns
	W3.0	0.007858	0.012705	ns	0.015146	0.009866	ns
FUL3	W2.5	0.050587	0.04894	ns	0.039389	0.023576	ns
	W3.0	0.041085	0.051989	ns	0.066782	0.073974	ns
MADS51	W2.5	5.6E-05	0.00012	ns	5.49E-05	6.54E-05	ns
	W3.0	0.000334	0.000223	ns	0.00023	0.00089	ns
MADS58	W2.5	7.91E-05	0.000166	ns	0.000165	0.000118	ns
	W3.0	0.000173	0.000467	ns	0.000305	0.000301	ns
WAPO1	W2.5	0.000177	0.000219	ns	0.000153	0.000132	ns
	W3.0	5.58E-05	0.00011	ns	6.41E-05	7.91E-05	ns
PPD1	W2.5	0.001311	0.001881	ns	0.000881	0.001719	ns
	W3.0	0.000602	0.000851	ns	0.000472	0.00053	ns
PRR73	W2.5	0.004928	0.005259	ns	0.005948	0.004962	ns
	W3.0	0.003086	0.004676	ns	0.003809	0.004424	ns

[a] ns = not significant, * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, n = 4.

[b] W2.5= double ridge and W3.0 = glume primordia spike developmental stages in Waddington scale.

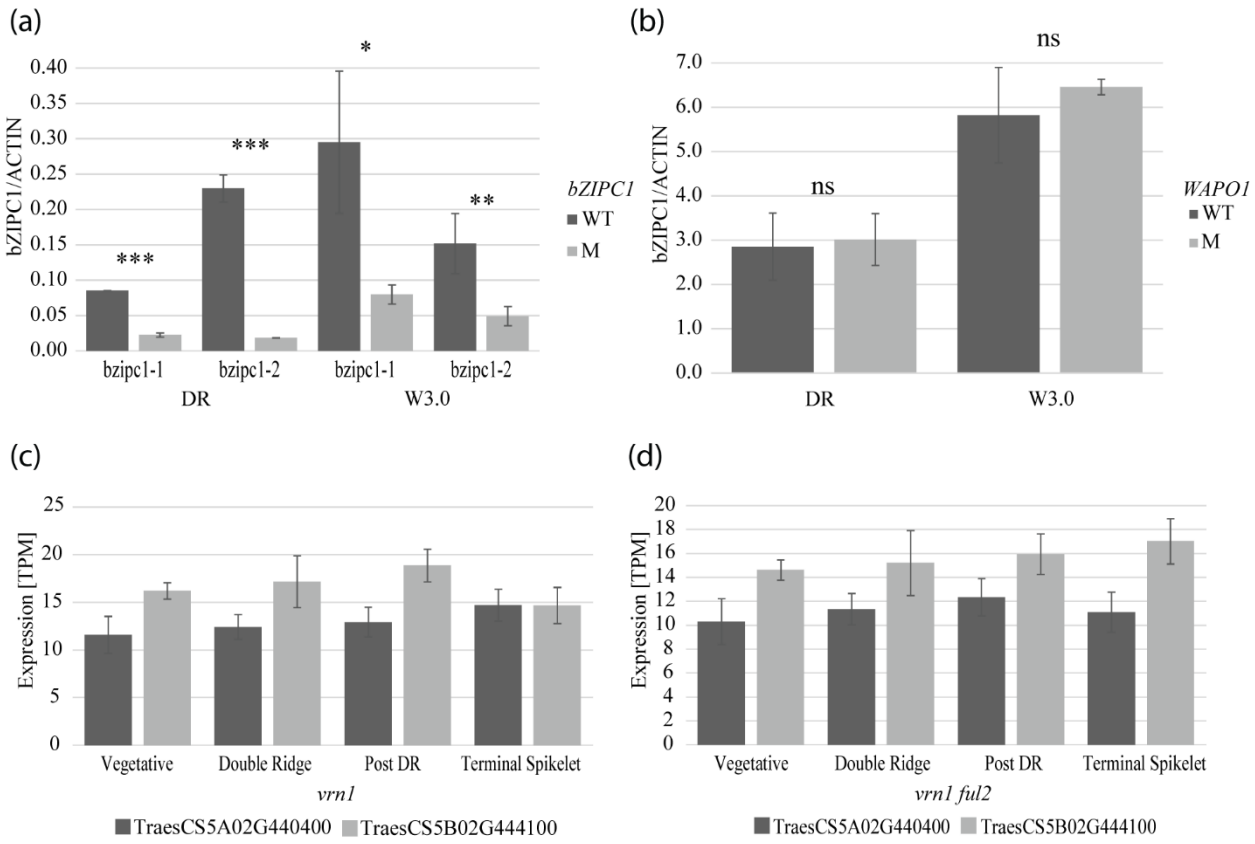


Fig. 6. Effect of mutations in *bzipC1*, *wapo1*, *vrn1* and *ful2* on the expression of *bZIPC1*. Transcript levels of *bZIPC1* during meristem development determined by qRT-PCR in (a) Kronos wildtype and *bzipc1-1* and *bzipc1-2* sister mutant lines (n=4) and (b) Kronos wildtype and *wapo-a1* mutant sister lines (n=3). ns = not significant, * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, n = 4. (c-d) *bZIPC1* transcripts per million from previously published RNA-Seq study (Li et al. 2019) in (c) Kronos and *vrn1 ful2* mutant sister lines, and (d) Kronos and *vrn1* mutant sister lines.

Transcript levels of *bZIPC1* homoeologs in the *bzipc1-1* and *bzipc1-2* combined mutants were at least 3-fold lower than in the WT (qRT-PCR, Fig. 6a), suggesting reduced stability of the mRNAs carrying the premature stop codon, possibly via the nonsense mediated decay pathway (Nickless et al., 2017).

To further explore if genes involved in SNS development were involved in the transcriptional regulation of *bZIPC1*, we measured the combined *bZIPC1* mRNA levels from the A and B genome in null mutants from genes previously shown to impact SNS such as *wapo1* (Kuzay et al., 2022), *vrn1* and *vrn1 ful2* mutants (Kuzay et al., 2022; Li et al., 2020). In both

cases, *bZIPC1* expression was not significantly altered in the mutant genetic backgrounds (Fig. 6b-d). Thus, *bZIPC1* transcript levels do not appear to be modulated by these genes.

Given the limited impact of *bZIPC1* mutations on the transcript levels of genes known to affect spike development, we initiated a Quant-Seq experiment to explore other pathways affected by *bZIPC1*. We extracted RNA from developing spikes at double ridge (W2.5) and glume primordia (W3.0) stages from two independent loss-of-function *bzipc1* mutants (*bzipc1-1* and *-2*). As biological replications, we used 4 pools of 4 dissected developing spikes at the same developmental stage for each genotype at each developmental stage. These future transcriptomic results should allow us to identify the transcriptomic changes associated with the mutant phenotype, and the genes and pathways regulated by *bZIPC1*.

Natural variation in *bZIPC1*

The significant differences in SNS observed in the *bzipC1* mutants motivated us to look at the natural variation in both *bZIPC1* homoeologs in a set of 55 tetraploid and hexaploid wheat accessions from the Wheat T3 database (Blake et al., 2016) sequenced by exome capture (Fig. 7). We found two non-synonymous SNPs in *bZIPC-B1* (*TraesCS5B02G444100*) but none in *bZIPC-A1* (*TraesCS5A02G440400*) or *bZIPC-D1* (*TraesCS5D02G447500*) in hexaploid wheat, so we focused our haplotype analysis in a 1.3 Mb region on chromosome 5B flanking *bZIPC-B1* (CS RefSeq v1.1 615,695,210 to 617,038,639).

A cluster analysis based on the 242 SNPs detected in the exome capture data of the 55 accessions in this region revealed four haplotypes designated here as H1 to H4 (Fig. 6). The H1 haplotype was the most frequent (72.7%) among these accessions and was more closely related to H2 than to the other haplotypes. Haplotypes H3 and H4 were related to each other, but since

H4 included only a single tetraploid accession (280-1-Yr15) used for the introgression of *Yr15* from *T. turgidum* subsp. *dicoccoides* (Yaniv et al., 2015), we did not characterize it further.

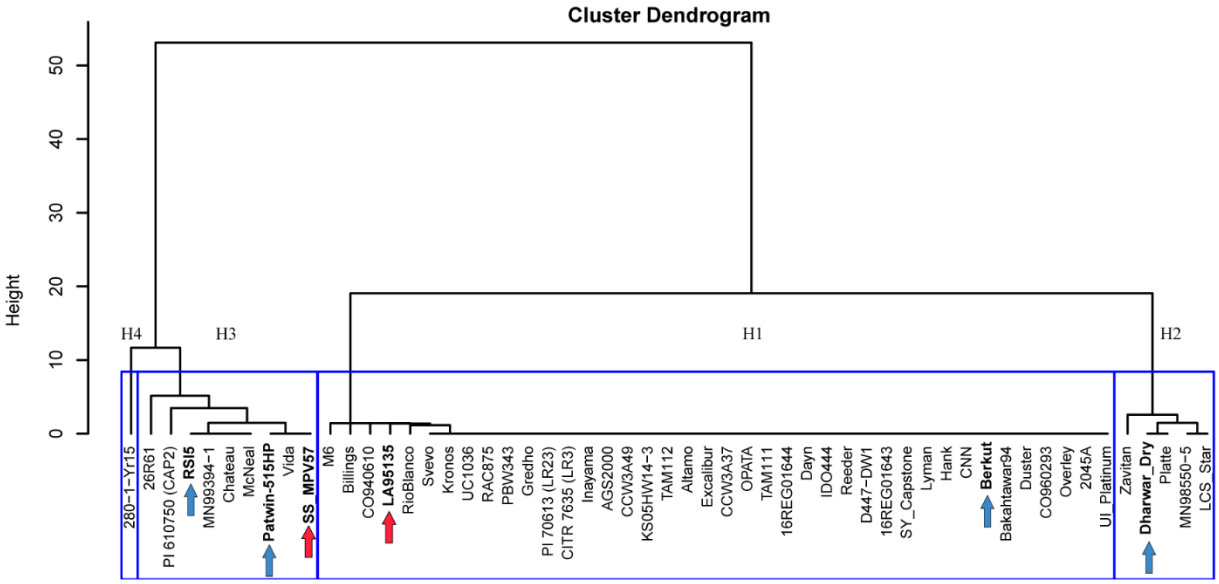


Fig. 7: Haplotypes in the *bZIPC1* region. Cluster analysis based on 242 SNPs detected in chromosome 5B (CS RefSeq v1.1 615,695,210 to 617,038,639 bp) in exome capture data extracted from T3/Wheat (https://wheat.triticeaetoolbox.org/). Blue arrows indicate spring lines from haplotypes H2 and H3 used as parental lines in crosses Berkut (H1). Red arrows indicate winter lines used as parental lines in the soft red winter mapping population (Dewitt et al 2022).

We identified one historical recombination event between H1 and H3 at position 615,736,949-615,951,042 in line 26R61, with *bZIPC1* located within the H3 haplotype region. Another historical recombination event was detected between H1 and H2 at position 617,038,639-617,049,421 in tetraploid lines Kronos and Gredho. Common wheat winter line LA95135 showed a recombination event at the same position but between H1 and an unknown haplotype in the distal region. In these four lines with historic recombination events in the region, the *bZIPC-B1* gene is located within the H1 haplotype.

We then compared the four haplotypes for polymorphisms within the *bZIPC-B1* coding region and identified one synonymous SNP at position 616,652,860 that differentiated H3 from all other three haplotypes, and two non-synonymous SNPs at positions 616,654,272 (N151K)

and 616,654,229 (V166M) that differentiate the H1 haplotype (Table 3). The linked amino acids K151 and M166 (KM henceforth) were found only in the H1 haplotypes whereas the N151 and V166 linked changes (NV henceforth) were found in the other three *bZIPC-B1* haplotypes, *bZIPC-A1*, *bZIPC-D1*, and *Hordeum vulgare*, suggesting that NV is the ancestral state and KM (H1) is a derived state. In addition, while we observed variation at position 166 (Table 3), the N151 amino acid was conserved in *B. distachyon*, rice, *Setaria*, and *Panicum* supporting the previous hypothesis.

Table 3: Distribution of *bZIPC-B1* non -synonymous SNPs among wheat haplotypes, wheat homoeologs *bZIPC-A1* (A) and *bZIPC-D1* (D), and other grass species. *bZIPC-B1* (*TraesCS5B02G444100*) is located in the minus strand. DNA changes are described based on the plus strand.

RefSeqV1.1 5B	H1	H2/3/4	H1	H2/3/4	A	B	Barley	Brachy	Rice	Setaria	Panicum
616,654,229	T	C	M166	V166	V	V	V	E	S	S	E
616,654,272	C	A	K151	N151	N	N	N	N	N	N	N

Analysis of the N151K and V166M polymorphisms in the sequenced genomes of tetraploid wheat revealed that both Svevo (Maccaferri et al., 2019) and Kronos ('Kronos EI v1', https://opendata.earlham.ac.uk/opendata/data/Triticum_turgidum/EI/v1/) carry the H1 haplotype, whereas Zavitan (*T. turgidum* subsp. *dicoccoides*) has the H2 haplotype (Avni et al., 2017). The H1 haplotype was also frequent among the sequenced hexaploid wheats in the PanGenome project (Walkowiak et al., 2020), including Chinese Spring, Lancer, SY Mattis, Stanley, Jagger, and CDC Landmark. A comparison of the promoter region including the 1,500 bp upstream of the start codon revealed no polymorphisms among these accessions, supporting the hypothesis that H1 is a relatively recent haplotype. The ancestral NV combination was detected in the genomes of Norin61, Spelt, Mace, and Rubigus carrying the H3 haplotype and in the genomes of

ArinaLrFor, Julius, Cadenza, Paragon, and Clair carrying the H2 haplotypes. A comparison of the *bZIPC-B1* promoter region (1500 bp) between Chinese Spring (H1) and Norin 61 (H3) revealed 19 SNPs and two indels, whereas the same region showed 9 SNPs and 4 indel between the H1 and H2 haplotypes, confirming the closer relationship between H1 and H2 relative to H3.

To explore the variation of the *bZIPC-B1* haplotypes in time, we screened a collection of 63 *T. turgidum subsp. dicoccoides*, 77 *T. turgidum subsp. dicoccon*, and 328 *T. turgidum subsp. durum* with KASP markers for the K151N and M166V SNPs (see "Material and methods"). We found that 35 % of the *T. turgidum* subsp. *dicoccoides* accessions and 4 % of the *T. turgidum subsp. dicoccon* accessions have the derived KM allele (Table 4). For the *T. turgidum subsp. durum*, we split the group into old landraces and more modern cultivars and breeding lines. Of the 175 landraces, we found that 37% had the derived KM allele while of the 153 cultivars and breeding lines, 54% carried the KM allele (Table 4), indicating an increased frequency of the derived allele in the modern cultivated durum wheats.

A similar analysis of exome capture data from 456 *T. aestivum* subsp. *aestivum* accessions (He et al., 2019) revealed that 73 % of the accessions have the derived *bZIPC-B1* KM allele and only 27% the ancestral haplotype (Table 4). Taken together, these results suggest positive selection for the KM derived allele. To explore this possibility further, we tested the effect of the *bZIPC-B1* KM allele in four segregating populations.

Table 4: *bZIPC-B1* allele distribution in wild and cultivated tetraploid and hexaploid wheat.

	NV (ancestral)	KM (derived)	NV %	KM %
T. turgidum subsp. *dicoccoides*	41	22	65%	35%
T. turgidum subsp. *dicoccon*	74	3	96%	4%
T. turgidum subsp. *durum* Landraces	110	65	63%	37%
T. turgidum subsp. *durum* cvs. & breeding lines	70	83	46%	54%
T. aestivum subsp. *aestivum*	125	331	27%	73%

Effect of *bZIPC1* alleles on grain yield components

The haplotype analysis revealed that the group of varieties with the H1 haplotype included the spring wheat variety Berkut, which is the central parent in the spring wheat Nested Association Mapping (NAM) population (Blake et al., 2019). We previously characterized populations generated from crosses between Berkut and spring common wheat accessions Dhar-war (H2), RS15 (H3) and PATWIN-515HP (H3) for SNS and other grain yield components in multiples locations and under two watering regimes (Zhang, Gizaw, Bossolini, et al., 2018).

These populations were previously genotyped with the Illumina 90K SNP chip (Jordan et al., 2015), and among these SNPs we identified IWB56221 (RefSeq v1.1 5B 616,652,623) as the closest diagnostic marker for the H1 haplotype (H1= T, H2, H3 and H4 = C). Using this marker, we performed *t*-tests between the two alleles for the different traits in the different environments (Table 5).

The comparison between the H1 and H3 haplotypes in the crosses Berkut x PATWIN-515HP and Berkut x RSI5 showed that the H1 haplotype (=KM *bZIPC-B1* allele) was associated with significant increases relative to H3 (NV) in the number of total and fertile spikelets per spike (3.1 to 6.1%), grain number per spike (11.0 to 11.5% increase), and kernel weight per spike (10.6% to 11.9%, Table 5) in multiple locations. The H1 haplotype was also associated with significant increase in grain number per spike (7.3%) and kernel weight per spike (10.6%, Table 5) relative to the H2 haplotype (NV) in the cross Berkut x Dhar-war. Although the number of total and fertile spikelets per spike was not significant in this population, it showed the same trend as in the other two populations, with higher values in H1 (1.1 to 1.7%) relative to H2.

Table 5: Effect of marker IWB56221 tightly linked to *bZIPC-B1* on total and fertile spikelet number per spike (SNS), grain number per spike (GNS) and kernel weight per spike (KWS) in spring wheat NAM populations. Only environments showing significant effects are presented (complete information is available in Supplemental Appendix 2). Imp = Imperial Valley, California, Dav= Davis, Sacramento Valley, California, 15= 2015, 16=2016, I= normal irrigation, and D= without last irrigation. Data is from Zhang et al. 2018.

A. Berkut x PATWIN-515HP

	Fertile SNS		SNS		GNS		KWS		
	Imp15_D	Imp15_I	Imp15_D	Imp15_I	Dav15_D	Imp15_I	Dav15_D	Imp15_D	Imp15_I
Avg Berkut	18.44	18.21	21.79	21.21	71.92	57.71	2.72	1.17	1.85
Avg PATWIN-515HP	17.49	17.05	21.10	20.53	65.71	50.81	2.49	1.05	1.61
t-Test	0.0006	0.0001	0.016	0.017	0.040	0.008	0.011	0.06	0.002
Diff	0.95	1.16	0.69	0.67	6.21	6.90	0.23	0.12	0.24
% Diff	5.4%	6.8%	3.3%	3.3%	9.4%	13.6%	9.4%	11.6%	14.7%
Avg. Diff.		1.05		0.68		6.55			0.20
% change		6.1%		3.3%		11.5%			11.9%

N= 69

B. Berkut x RSI5

		Fertile SNS		SNS		GNS		KWS
	Dav16_D	Imp15_D	Imp15_I	Imp15_D	Imp15_I	Dav15_I	Imp15_I	Imp15_I
Avg. Berkut	17.67	17.34	17.43	19.65	19.73	55.62	53.23	1.88
Avg. RSI5	17.15	16.64	16.71	19.27	18.94	50.93	47.16	1.70
t-Test	0.0344	0.0155	0.0616	0.0712	0.0795	0.0954	0.0072	0.0134
Diff	0.52	0.70	0.72	0.37	0.80	4.69	6.07	0.18
% Diff	3.0%	4.2%	4.3%	1.9%	4.2%	9.2%	12.9%	10.6%
Avg.			0.65		0.58		5.38	0.18
Avg.			3.8%		3.1%		11.0%	10.6%

N= 64

C. Berkut x Dhar-war

	GNS	KWS
	Dav15_D	Dav15_D
Avg. Berkut	63.42	1.88
Avg.Dhar-war	59.09	1.70
t-Test	0.035	0.013
Diff	4.33	0.18
% Diff	7.3%	10.6%

N= 734

We also explored the effect of the H1 haplotype relative to H3 in the winter wheat population LA95135 (H1) x SS-mvp57 (H3) evaluated in five different environments in North Carolina, USA (DeWitt et al., 2021). In addition to *bZIPC-B1*, this population segregates for

another four genes affecting SNS including: *PPD-D1*, *RHT-D1*, *FT-A2* (Glenn et al., 2022), and *WAPO-A1* (Kuzay et al., 2019). We inferred the *bZIPC-B1* genotype from the consensus of two SNPs flanking this gene at 616,367,516 and 616,955,782, and used it together with the genotypes of the other four genes in a factorial ANOVA including the five genes as factors, all the possible two-way interactions among genes, and environments as blocks.

The combined ANOVA across environments, explained 43.4% of the variation in SNS and revealed highly significant differences in SNS for all five genes including *bZIPC-B1* ($P < 0.0001$, Table 6). *PPD-D1* and *WAPO-A1* showed the strongest effects, whereas *bZIPC-B1* had the weakest effect, with the Berkut allele showing on average a 1.4% increase in SNS (0.3 spikelets, Table 6). The differences in SNS for *bZIPC-B1* were significant in three out of the five environments when the ANOVAs were performed separately by environment. None of the two-way interactions involving *bZIPC-B1* with the other four genes were significant (only *RHT-D1* interactions with *PPD-D1* and *WAPO-A1* were significant, Table 6).

In addition, we analyzed the effect of *bZIPC-B1* on fertility, grain number per spike (GNS), days to heading (DTH) and spike grain yield. We found no significant differences for fertility associated with *bZPIC1* except for a marginally significant interaction with *PPD-D1*, so it was not surprising that the significant differences in SNS associated with *bZIPC1* were paralleled by significant differences in GNS. We also detected significant differences in fertility for *PPD-D1*, *RHT-D1*, and *WAPO-A1* and for GNS for *FT-A2*, and *RHT-D1* and for the *RHT*-D1 x *WAPO-A1* interaction (Table 6). Lines which carried the *bZIPC1* H1 haplotype on average had 1.31 more grains than the H3 haplotype (2.6%, Table 6).

Table 6: Factorial ANOVA for SNS, GN, Fertility, DTH, and Spike Yield including 5 genes (factors), two alleles (levels), and environments (blocks).

	SNS	Fertility	GNS	DTH	Spike Yield
Env	<.0001	0.6608	0.0001	<.0001	<.0001
WAPO-A1	<.0001	0.0339	0.1460	0.1501	0.2597
PPD-D1	<.0001	0.0080	0.2935	<.0001	0.0769
FT-A2	<.0001	0.0772	<.0001	<.0001	<.0001
RHT-D1	<.0001	<.0001	<.0001	<.0001	<.0001
bZIPC1	**<.0001**	**0.5311**	**0.0431**	**<.0001**	**0.0107**
Two-way interactions					
*PPD-D1*FT-A2*	0.6972	0.9239	0.9780	0.2954	0.4778
*PPD-D1*bZIPC1*	0.5463	**0.0414**	0.1157	0.6935	0.0511
*FT-A2*bZIPC1*	0.1060	0.1056	0.2489	0.0813	0.1682
*PPD-D1*RHT-D1*	**0.0236**	0.1088	0.3442	0.8272	0.6681
*FT-A2*RHT-D1*	0.0964	0.5089	0.2742	0.0650	0.7315
*bZIPC1*RHT-D1*	0.2871	0.6282	0.5759	0.1394	0.4894
*PPD-D1*WAPO-A1*	0.1071	0.6038	0.9115	0.9042	0.101
*FT-A2*WAPO-A1*	0.4229	0.9551	0.6043	0.2298	0.9133
*bZIPC1*WAPO-A1*	0.6649	0.4243	0.5732	**0.0016**	0.2449
*RHT-D1*WAPO-A1*	**<.0001**	0.2099	**0.0333**	0.8634	0.0539
No. of environments	5	2	2	4	3
Avg. bZIPC1 LA95135 (H1)	21.23	2.38	51.65	107.52	12.0919912
Avg. bZIPC1 SS-MVP57 (H3)	20.94	2.36	50.34	107.98	11.7403914
% change	1.41%	0.76%	2.60%	-0.43%	2.99%

On average, the H1 haplotype showed a significant 0.5-day delay in heading date relative to H3, which is consistent with the H1 positive effect on SNS. There was also a significant interaction between *bZIPC-B1* and *WAPO-A1* ($P = 0.0016$) so we analyzed the four possible simple effects. We found that the H1 haplotype was significantly different from H3 (0.93 d, $P < 0.0001$) within the *Wapo-A1b* allele (associated with high SNS and late heading) but not significant within the *Wapo-A1c* allele (0.04 d, $P = 0.12$ Fig. 8). The differences between the *WAPO-A1* allele were significant within the *bZIPC-B1* H1 ($P = 0.0151$) haplotype but not within the H3 haplotype ($P = 0.9982$). These results suggest that *bZIPC1* and *WAPO1* may affect heading time through a common pathway.

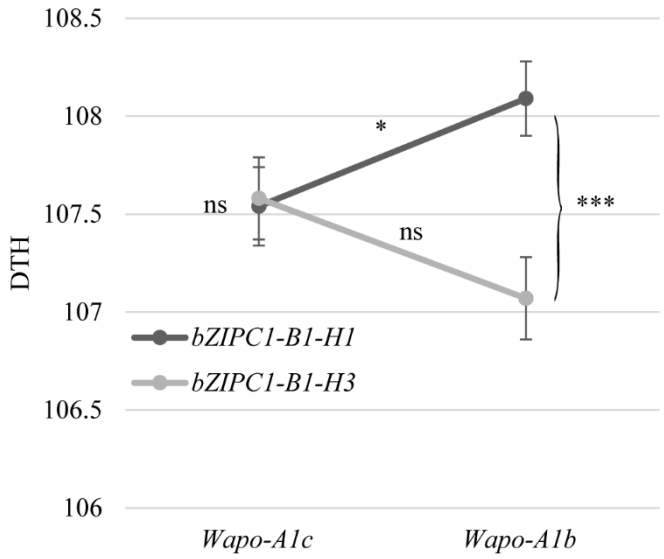


Fig. 8: *bZIPC-B1* x *WAPO-A1* interaction for days to heading (DTH). ns= not significant, * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$

Finally, we detected significant differences in spike grain yield for *bZIPC-B1*, *FT-A2* and *RHT-D1*, but none of the two-way interactions for this trait were significant (Table 6). On average, the spike grain yield associated with the *bZIPC-B1* H1 haplotype was 0.35 g greater than the H3 haplotype (2.99 % increase, Table 6), which is a promising result.

It is also interesting to note, that there were only two significant interactions for SNS between these genes (*RHT-D1* with *PPD-D1* and *WAPO-A1)*, and only a spattering of two-way interactions amongst them for the other yield-component traits, indicating that these SNS genes may operate mainly as additive factors in the determination of SNS.

Taken together, these results indicate that the *bZIPC-B1* KM allele (H1 haplotype) is associated with beneficial effects on SNS, GNS, and kernel weight per spike relative to the ancestral VN allele, providing a plausible explanation for the rapid increased in the frequency of the derived KM allele during wheat domestication and wheat improvement (Table 4).

3.4 DISCUSSION

bZIP transcription factors are a large and functionally diversified family

The basic leucine zipper (bZIP) transcription factors (TF) are a large and diverse family present in eukaryotes from *Saccharomyces cerevisiae* to *Homo sapiens* (Agarwal et al., 2019). They are characterized by two conserved domains. The first one is a DNA-binding domain rich in basic amino acids, which is followed by a leucine zipper consisting of several heptad repeats of hydrophobic amino acids (Hurst, 1995). This last region favors the formation of homodimers or heterodimers, resulting in multiple combinations with unique effects on transcriptional regulation (Deppmann et al., 2006). In plants, the bZIP transcription factors have a variety of roles ranging from hormonal responses, light signaling, nitrogen/carbon metabolism, biotic and abiotic stress, flower induction, development, and seed storage and maturation (Agarwal et al., 2019; Alonso et al., 2009; Alves et al., 2013; Baena-González et al., 2007; Choi et al., 2000; Fujita et al., 2005; Jakoby et al., 2002; Schütze et al., 2008).

In *A. thaliana,* 78 bZIP TFs have been identified and divided into 10 subfamilies (A to I and S) based on their similar motifs. The bZIP proteins from the C-group, which is the focus of this study, preferentially form heterodimers with other members of the C-group as well as with members of the S1-group (Ehlert et al., 2006). Known as the “C/S1 network”, this network has been shown to be involved in amino acid metabolism, stress response, and energy homeostasis (Alonso et al., 2009; Dröge-Laser & Weiste, 2018; Feng et al., 2021).

C-group bZIP transcription factors have diversified in the grasses

In *A. thaliana* the C-group includes four members (AtbZIP-9, -10, -25, and -63) which participate in a wide range of functions. For example, AtbZIP63 has been demonstrated to regulate the starvation response (Mair et al., 2015) and to play an important role in fine-tuning of

ABA-mediated, sugar dependent abiotic stress responses (Matiolli et al., 2011). bZIP10 was identified to modulate basal defense and cell death in *A. thaliana* (Kaminaka et al., 2006) and to work with glutathione to induce various heat shock proteins (Kumar & Chattopadhyay, 2018)

Searching for interactions with TaFT2, our screening of the *B. distachyon* cDNA libraries isolated the uncharacterized bZIP transcription factor, Bradi1g05480, which is closely related to *A. thaliana bZIP63* (At5g28770), *bZIP10* (At4g02640), and *bZIP25* (At3g54620) of the bZIP C-group. In the monocots, the C-group includes four main clusters, two of which include only grass species. To generate a more comprehensive nomenclature for the large family of wheat bZIP transcription factors that better reflects their evolutionary relationships, we propose to incorporate the group name to the bZIP gene name and designate the wheat *bZIP* genes from C-group as *bZIPC1*, *bZIPC2*, *bZIPC3*, and *bZIPC4* (Table 7).

Table 7: Proposed nomenclature for bZIP wheat genes from C-group

Gene name	Homoeologs	Gene ID (RefSeq v1.1)	Synonyms and barley orthologs
bZIPC1	*bZIPC-A1*	*TraesCS5A02G440400*	*TaSPH* *HvBLZ1*
	bZIPC-B1	*TraesCS5B02G444100.1*	
	bZIPC-D1	*TraesCS5D02G447500*	
bZIPC2	*bZIPC-A2*	*TraesCS1A02G329900*	*TaSPA* *HvBLZ2*
	bZIPC-B2	*TraesCS1B02G343500*	
	bZIPC-D2	*TraesCS1D02G332200*	
bZIPC3	*bZIPC-A3*	*TraesCS5A02G057500*	
	bZIPC-B3	*TraesCS5B02G059200*	
	bZIPC-D3	*TraesCS5D02G068800*	
bZIPC4	*bZIPC-A4*	*TraesCS6A02G154600*	
	bZIPC-B4	*TraesCS6B02G182500*	
	bZIPC-D4	*TraesCS6D02G144400*	

Functional characterization of *bZIPC1* and *bZIPC2* genes

The bZIPC1 cluster includes previously studied proteins such as wheat TaSHP (Boudet et al., 2019), barley BLZ1 (Vicente-Carbajosa et al., 1998) and maize OHP1 (Pysh et al., 1993), whereas the *bZIPC2* cluster includes wheat TaSPA (Albani et al., 1997) barley BLZ2 (Oñate et al., 1999), and maize OPAQUE2 (Pysh et al., 1993). The bZIPC3 and bZIPC4 subgroups are more distantly related and less studied.

The first wheat protein described for the bZIP C-group, the STORAGE PROTEIN ACTIVATOR (TaSPA = bZIPC2), was first shown to be involved in the regulation of grain storage proteins (Albani et al., 1997) and later in starch accumulation (Guo et al., 2020). This protein is orthologous to the barley protein BLZ2, which interacts in Y2H with the bZIP protein BLZ1 (Oñate et al., 1999). Since both *BLZ1* and *BLZ2* were shown to be involved in the regulation of seed storage proteins (Vicente-Carbajosa et al., 1998), the wheat ortholog of *BLZ1*, *TraesCS5B02G44410.1*, was cloned and named SPA HETERODIMERIZING PROTEIN (TaSHP) (Boudet et al., 2019). Although this name was based on the known interaction of the barley orthologs BLZ1 and BLZ2, this interaction was not validated in wheat in the previous study.

The maize ortholog of *bZIPC2*, *OPAQUE2* (*O2*), was initially identified as a positive regulator of grain storage protein (Pysh et al., 1993), and a similar function has been proposed for the rice ortholog *RISBZ1* (*OsbZIP58*) (Kawakatsu et al., 2009). Similarly, the orthologs of *bZIPC1* in barley (BLZ1) and maize (OPAQUE2 HETERODIMERIZING PROTEINs OHP1 and OHP1b) have been shown to bind to the promoters of the genes encoding storage proteins promoting their activation (Vicente-Carbajosa et al., 1998; Zhang et al., 2015). *TaSHP* (= *bZIPC1*), however, has been shown to act as a repressor of glutenin storage proteins synthesis in

wheat (Boudet et al., 2019). These results suggest that bZIP genes are involved in the regulation of storage proteins and can act both as transcriptional activators or repressors, likely depending on complex interactions with other bZIP genes in the different genetic backgrounds.

In addition to their role in the regulation of grass storage proteins, *bZIPC1* and *bZIPC2* genes have been shown to have additional functions. The rice ortholog of *bZIPC1*, *OsbZIP33* (Os03g58250), is reported to be involved in the resistance of abiotic stresses via ABA-dependent stress signal transduction (Rabbani et al., 2003; Zeller et al., 2009), with transgenic rice plants overexpressing *OsbZIP33* showing significantly higher survival rates than their wild-type counterparts after dehydration treatment (Chen et al., 2015). Similarly, the maize ortholog of *bZIPC1*, ZmbZIP41, was identified as related to drought in a drought re-watering transcriptome dataset but no functional characterization was completed (Cao et al., 2019). The significant reduction in SNS detected in our *bzipc1* wheat mutants demonstrates a previously unknown role of this gene in the regulation of spike development.

Given the additional functions of the *bZIPC1* genes described in the literature, it would be interesting to test the effect of the wheat *bzipc1* mutants on the relative abundance of different classes of grain storage proteins and on survivability compared to their wild-type sister lines in drought and re-watering experiments. In addition, given that the expression of *bZIPC1* was greatest in the roots, further analyses of root architecture and biomass in the wheat *bzipc1* mutants and wildtype sister lines are worth pursuing in future studies.

Functional characterization of *bZIPC3* and *bZIPC4* genes

The *bZIPC3* and *bZIPC4* genes have been less characterized than *bZIPC1* and *bZIPC2*, to which they are more distantly related (Fig. 1). The wheat genes *bZIPC-B3*

(*TraesCS5B02G059200*) was upregulated in response to cold in a study of wheat responses to freezing temperatures, (Tian et al., 2022), but no additional characterization has been completed for this gene or its homoeologs (Table 7). The maize ortholog of *bZIPC3, ZmbZIP11* (*Zm00001d030995*), was upregulated during the induced systemic resistance activation by *Trichoderma atroviride* (Agostini et al., 2019), and the rice orthologous protein OsbZIP88 (Os12g40920) was shown to form heterodimers with Group-S1 member, OsbZIP71, which is an important regulator in ABA-mediate drought and salt stress response in rice (Liu et al., 2014). Overall, the few bZIPC3 genes characterized so far affect traits encompassing a broad spectrum of stress related features.

None of the three wheat *bZIPC4* homoeologs (Table 7) has been functionally characterized, but the *bZIPC-D4* (*TraesCS6D02G144400*) gene showed higher transcript levels after drought and heat treatments (Liu et al., 2015). The maize *bZIPC4* ortholog (*ZmbZIP104*) has been shown to be regulated by *OPAQUE2* in a study of the maize endosperm transcriptome (Zhan et al., 2018). The rice *bZIPC4* ortholog (*RISBZ5* = OsbZIP52) has been proposed to function as a negative regulator in cold and drought stress environments, with rice plants overexpressing *RISBZ5* showing significant increases in cold and drought stress sensitivity (Liu et al., 2014)

Although additional functional characterization studies are necessary to understand the different roles of the *bZIPC3* and *bZIPC4* genes, the available information for the four bZIP C-group clades is indicative of some degree of sub-functionalization. This sub-functionalization is reflected in our Y2H results (Table 1), which showed that each of the four wheat bZIP proteins of the C-group interacts with a different set of *FT*- and *CEN*-like encoded proteins, supporting the diversification of these clades.

bZIPC1 interacts with FT2 and regulates SNS via an unknown mechanism

In spite of its sequence similarity with *FT1*, *FT2* shows some unique characteristics that set it apart from other *FT*-like genes. *FT2* is the only *FT*-like gene in wheat that is expressed directly in the developing spike and the only one that encodes a protein that does not interacts with any of the known bZIP proteins of the A-group (FDL2, FDL6, and FDL15) or the known 14-3-3 proteins that interact with the other FT-like proteins. In addition, it is the only one that shows significant effects on SNS with limited effects on heading time. These unique characteristics motivated us to look for its protein interactors.

We identified bZIPC1 as a novel interactor of FT2, and since we were not aware of previous reports of C-group bZIP proteins interacting with FT-like proteins, we decided to pursue its functional characterization. We hypothesized that knock-out mutants in *bZIPC1* would limit the ability of *FT2* to accelerate the formation of the terminal spikelet and would also result in an increase in SNS. Surprisingly, all three *bzipc1* mutant alleles in tetraploid wheat showed a significant decrease in SNS.

Although this clearly demonstrates that our initial hypothesis was false, the mechanism to explain why the loss of function of *bZIPC1* results in reduced SNS remains elusive. The qRT-PCR analysis of developing spikes in the wildtype and *bzipc1* mutants showed no consistent difference in expression in *FT2* suggesting that the *bZIPC1* effect on SNS was not mediated by the transcriptional regulation of *FT2*. Since bZIPC1 can also interact with FT3, we cannot currently rule out the hypothesis that the *bzipc1* mutant effect on SNS was mediated by FT3. Combined *ft3 bzipc1* will be required to test this hypothesis.

Finally, an alternative possibility is that the bZIPC1 protein interaction with FT2 results in a non-functional complex that competes with other FT2 interactors required for FT2 function.

bZIP proteins from the C-group are known to form heterodimers with other members of the C-group and with members of the S1-group generating a complex network of potential competitive interactions (Ehlert et al., 2006). Under this scenario, the *bzipc1* mutant could result in reduced competition with an interactor required for FT2 function, increased FT2 activity, and reduced SNS. Testing the interaction between bZIPC1 and other bZIP proteins from the C- and S1-groups could be an initial step to test this hypothesis.

The absence of significant differences in expression between the *bzipc1* mutants and the wildtype for several spike development genes suggests that these genes are likely not part of the mechanism by which *bZIPC1* regulates SNS. The tested genes with known effects on wheat SNS that showed no expression differences included *PHOTOPERIOD1 (PPD1)* (Shaw et al., 2013), *FLOWERING LOCUS T1 (FT1)* (Lv et al., 2014), *WHEAT ORTHOLOG OF APO1 (WAPO-A1)* (Kuzay et al., 2019, 2022), and the meristem identity genes *VRN1*, *FUL2* and *FUL3* (Li et al., 2019). Therefore, these genes are unlikely targets of a *bZIPC1* mediated transcriptional activation at the tested developmental stages (W2.5 and W3.0). These developmental stages were selected because they precede the formation of the terminal spikelet, and also because expression analyses showed that *bZIPC1* and *FT2* expression overlapped at these stages. However, we cannot completely rule out the possibility that the expression of these genes is affected by *bZIPC1* at a different time point of spike development.

Furthermore, no significant differences in *bZIPC1* expression were observed between the Kronos wildtype and the *vrn1*, *vrn1 ful2, wapo1,* and *ppd1* mutants in the same backgrounds suggesting that *bZIPC1* transcription is not regulated by these genes. The expression of *bZIPC1* in the developing spike at the same stages as *FT2* in the same set of RNASeq samples indicates that these two genes are expressed in the same organ at the same time before the formation of the

terminal spikelet. These results support the hypothesis that the bZIPC1-FT2 interactions observed in yeast may be biologically relevant.

We show that bZIPC1 oscillates throughout the day under continuous light indicating that this gene is regulated by the circadian clock. This agrees with results from other species, where bZIP genes from the C-group have also been shown to be clock regulated (e.g. Frank et al., 2018). In *A. thaliana* the cycling of *bZIP63* plays an important role in the regulation of *PRR7* (*PPD1* homolog) in response to sugar levels and can itself impact clock function (Frank et al., 2018).

In contrast with *FT2*, whose transcript levels are significantly downregulated under SD, *bZIPC1* showed similar profiles under SD and LD, although higher transcript levels were detected in the last time point sampled during the long nights of SD than under LD. Consistent with the previous results, loss-of-function mutation in the main wheat photoperiodic gene *PPD1* did not significantly affect the transcriptional profile of *bZPC1* under continuous light.

Another possibility is that *bZIPC1* impacts SNS via stress or sugar related pathways. The *bZIPC1* paralogs in *A. thaliana* have been implicated in both abiotic stress and sugar responses (Frank et al., 2018; Kaminaka et al., 2006; Kumar & Chattopadhyay, 2018; Matiolli et al., 2011). For example, *bZIP63* impacts the overall energy metabolism in plants (Frank et al., 2018; Matiolli et al., 2011). Moreover, overexpression of the *bZIPC1* gene in rice, *OsbZIP33*, impacts ABA-dependent stress signal transduction (Rabbani et al., 2003; Zeller et al., 2009) and the *bZIPC1* gene in maize, *ZmbZIP41*, was also identified in a drought-stress transcriptome dataset (Cao et al., 2019). In addition, the bZIP C/S1 network has been shown to be involved in stress response and energy homeostasis in *A. thaliana* (Ehlert et al., 2006). Under starvation conditions, the C/S1 network reprograms metabolic gene expression to support survival while, in prevailing

energy supply, controls plant growth, development and stress responses by fine tuning carbon and nitrogen responses (Dröge-Laser & Weiste, 2018). Our experimental plants were grown under a non-stressed environment, but it is possible that the *bzipc1* mutants were interfering with regulation of plant growth and development. It will be interesting to test if bZIPC1 is able to interact with any S1 group proteins. Additionally, the planned transcriptomic work between *bzipc1* versus wild type should reveal if any stress related pathways are affected by the *bzipc1* mutation.

Natural variation in bZIPC1 and potential applications to plant breeding

Analysis of natural variation in the *bZIPC-B1* region revealed three major haplotypes (H1-H3) with bZIPC1 protein in the H1 haplotype being distinguished from H2/H3 by two SNPs resulting in amino acid changes K151N and M166V. Based on comparison with other grasses we determined that the bZIPC1 NV combination corresponds to the ancestral allele, and that the frequency of the derived KM genotype was low in *T. turgidum* subsp. *dicoccoides*, *T. turgidum* subsp. *dicoccon*. Among the durum accessions, the frequency of the derived KM allele increased from 37% in the durum landraces to over 50% in the cultivated and breeding durum lines. In addition, the frequency increased to 73% in common wheat lines. These changes suggest that the H1 haplotype might have been favored by positive selection, and that there is still a significant number of varieties (27% in hexaploid wheat) that can potentially benefit from the incorporation of the H1 allele.

The K151N and M166V mutations are located outside the conserved basic region and leucine zipper domains and have relatively high BLOSUM62 scores (0 and 1 respectively), predicting a limited effect on protein structure and/or function of the individual mutations.

However, a more detailed structural analysis of the proteins will be required to predict the effect of the combined mutations. It is also possible that the multiple SNPs and indels we detected between the H1 and H2/H3 haplotypes in the *bZIPC1* promoter regions contribute to differences in expression and in phenotype between haplotypes. A detailed comparative study of the expression profiles of the different haplotypes in an isogenic background during spike development will be necessary to see if differences in expression contribute to the phenotypic differences observed among the *bZIPC-B1* alleles.

Finally, we cannot rule out the possibility that selection for other genes tightly linked to *bZIPC-B1* within the H1 haplotype contributed to its increased frequency in hexaploid wheat. We have initiated a fine mapping project to narrow down the candidate region of the H1 haplotype and the number of potential candidate genes for the reduced SNS phenotype. Given the effect of the *bZIPC1* null mutants on SNS, *bZIPC-B1* stands as a strong gene candidate for the observed differences in SNS.

Overall, the H1 haplotype showed significantly higher SNS, grain number per spike, and kernel weight per spike than both the H2 and H3 haplotypes in multiple locations and in spring and winter wheat. These significant increases in kernel weight per spike in all three spring wheat populations and the winter wheat population are encouraging, because they show that the increases in grain number were not completely offset by a decrease in kernel weight. Replicated field experiments using near isogenic lines for the bZIPC1 haplotypes in highly productive genotypes grown in optimum environments will be necessary to test if the positive impacts observed in the yield components is translated into positive effects on total grain yield.

The magnitude of the *bZIPC-B1* H1 haplotype effect on these yield components varied across environments, similarly to what has been reported previously for other wheat genes

affecting SNS such as *FT-A2* (Glenn et al., 2022) and *WAPO-A1* (Kuzay et al., 2019, 2022). Therefore, before adoption of this gene into breeding programs, it would be prudent to test the effects of the *bZIPC-B1* haplotypes into local genetic backgrounds and environments. Fortunately, the H1 haplotype seems to be frequent in wheat germplasm from different parts of the world, which together with the haplotype analysis presented here, will facilitate the evaluation of its effects if SNP information is available. The two molecular markers developed in this study are diagnostic for the beneficial H1 haplotype and the associated KM polymorphism and can be used to differentiate H1 from the other three haplotypes carrying the NV polymorphism. Therefore, breeders can screen first their germplasm with these two markers to determine which lines have the H1 haplotype and which ones can benefit by its introgression. The sequences of these markers are publicly available to facilitate the deployment of this potentially beneficial *bZIPC-B1* allele.

3.5 MATERIAL AND METHODS

Plant materials

We identified the *bZIPC1* orthologs in the durum wheat cultivar Kronos using the sequences from Chinese Spring *bZIPC-A1* (*TraesCS5B02g444100*) and *bZIPC-B1* (*TraesCS5A02G440400*) and identified loss of function mutations in the sequenced Kronos mutant population (Krasileva et al., 2017). We then combined mutations in the A and B homoeologs from each gene to generate three loss-of-function lines designated as *bzipc1-1*, *bzipc1-2* and *bzipc1-3*. We backcrossed each of the three mutant alleles to Kronos twice and when possible phenotyped the mutant and segregating wild-type allele for bZIPC1.

Plant growth conditions

bZIPC1 mutant lines were initially grown in growth chambers at 16 hour long-day with temperatures oscillating between 22 and 18 °C during the day and night respectively. After heading and phenotyping, plants were moved to a greenhouse for drying and seed increases. Seed increases were grown in long-day greenhouses given supplemental lighting during the winter.

Phylogenetic analysis

Phylogenetic analyses of bZIP C-group genes were performed using BdbZIPC1, AtbZIP63, AtbZIP10 and AtbZIP25 as seed sequences for BLAST searches using Phytozome and NCBI as described previously (Woods et al., 2011). Amino acid sequences were aligned using MUSCLE (Edgar, 2004), before manual alignment of amino acid sequences in Mesquite (Maddison & Maddison, 2007; Fig. S1). Unalignable regions were pruned from the analysis in order to minimize noise in the phylogenetic signal. Maximum-likelihood analyses were conducted using SeaView version 5.0.4 using the LG model (Gouy et al., 2010). Bayesian phylogenetic analyses used MrBayes v3.2.6 x64 (Ronquist & Huelsenbeck, 2003) on the FARM cluster at the University of California Davis with 10 million generations, using the GTR+G model as determined by Mr. Model Test version 2.3 (Nylander, 2004). After convergence had been reached, the first 25% of trees were removed as burn-in and clade credibility values estimated using MrBayes. Trees were imaged using FigTree version 1.4.4 and further annotated in Adobe Illustrator 2021.

Statistical analysis

Analysis of Variance was conducted with the "Anova" function in R package "car" (Fox & Weisberg, 2020) with type 3 sum of squares and LS Means to accommodate unbalanced designs.

Marker Development

We used the available exome capture sequence from Berkut and PATWIN-515HP (Wheat/T3) to develop KASP markers for the bZIPC1 SNPs, N151K and V166M. We also created KASP markers for the tilling mutant alleles. (Table S2).

Yeast two hybrid assays

The full-length coding region of *FT2* (TraesCS3A02G143100) was cloned from Chinese Spring as described in Li et al 2015. The *FT2* coding region was recombined into pDONRzeo using Life Technologies BP Clonase following the manufacturer's protocol. *FT2* in pDONRzeo was subsequently recombined into the pDEST32 and pDEST22 yeast destination vectors using Life Technologies LR Clonase II following the manufacturer's protocol. Clones were verified by sequencing at each cloning step to ensure sequence integrity. All direct assays and library screens were performed using the MaV203 yeast strain as described in the ProQuest manual (Invitrogen). Before screening the cDNA libraries, we confirmed that FT2 was not auto-activated when used as prey or bait by testing it against the empty vector.

We screened a cDNA "photoperiod" library previously developed in *B. distachyon* (Cao et al., 2011). The photoperiod library was generated from shoots of two-week-old plants collected over 24 h grown under long (20 h light) and short day (8 h light) photoperiods (Cao et

al., 2011). For the screen, transformants were selected on plates with SD medium lacking leucine (L) and tryptophan (W) and were replica plated on SD medium lacking L, W and uracil (U) to identify putative FT2 interactors. Yeast colony PCR was done using the Phire polymerase following manufacturer's instructions (Thermo Fisher).

To confirm the Y2H screen results, we cloned the full-length coding sequence from Bradi1g05480 using a pooled cDNA mixture derived from young leaf and developing spikes from early double ridge (W2.0) to floret primordia present (W3.5) developmental stages (Waddington et al., 1983). The *bZIPC1* coding region was amplified using primers at the start and end of the coding region of *bZIPC1* and cloned into the pJET vector following the manufacturer's protocol (ThermoFisher). Primers are described in Table S2. Once *bZIPC1* was in pJET, this was used as a template to amplify the *bZIPC1* coding region using bZIPC1 BP primers (Table S2) and recombining into pDONRzeo using Life Technologies BP clonase following the manufacturer's protocol. The pDONRzeo vector containing the desired *bZIPC1* coding region was recombined into pDEST22 and pDEST32 using Life Technologies LR Clonase II following the manufacturer's protocol. Clones were verified by sequencing at each cloning step. We also cloned the wheat *bZIPC1* paralogs *bZIC3* and *bZIPC4*, and genes *CEN4*, *CEN5*, *FT2*, *FT3* and *FT5* using the same cloning strategy described for *bZIPC1* (Table S2), whereas *bZIPC2* was synthesized and cloned. *bZIPC1* and all genes used in the directed Y2H assays were confirmed to not auto-activate when either used as prey or bait. *FT1* and *CEN2* were previously cloned (Li et al., 2015) and these were transferred into pDEST22 and pDEST32 using the same strategy described for *bZIPC1*.

Gene expression studies

Plants were grown in 4.5 L pots placed in CONVIRON growth chambers under 16 h light at 22℃ (330 mol intensity) and 8 h darkness at 17℃ for 25–30 days. To compare the transcript levels of *bZIPC1* and several genes known to affect SNS between *bzipc1-1* and *bzipc1-2* mutant plants and their corresponding wildtype sister lines, we pooled 4 developing meristems when they reached the W2.0 and W3.0 stages (Waddington et al., 1983). RNA extraction and expression analyses were done as described previously (Shaw et al., 2020). Real-time PCR primers were utilized that amplify both the A and B genome homoeologs from the following genes: *FUL2*, *VRN1*, *SEP1-6* (=*MADS34*), *FLC2* (*MADS51*), *AG1* (*MADS58*) (Li et al., 2021), *WAPO1* (Kuzay et al., 2022), *CEN2*, *CEN4-A*, *CEN5* (Li et al., 2019), *PPD1*, *PRR72* (Chen et al., 2014), *FUL3* (Li et al., 2019), and *bZIPC1* (Table S2). *INITIATION FACTOR 4A* (*IF4A*) and *ACTIN* were used as endogenous controls for the qRT-PCR experiments.

3.6 SUPPLEMENTAL DATA

Table S1. List of detected interactors identified in the yeast two hybrid (Y2H) screens. Total of 26 unique genes. *Bradi1g35230* was detected in 67 colonies. *Bradi1g05480* was detected

Screen	Gene No.	Gene ID	Gene Description
P1		*Bradi1g05480*	bZIP transcription factor; bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor
P2		*Bradi1g05480*	bZIP transcription factor; bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	BZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2	1	*Bradi1g05480*	BZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2		*Bradi1g05480*	bZIP transcription factor RISBZ2
P3		*Bradi1g05480*	bZIP transcription factor; bZIP transcription factor RISBZ2
P3		*Bradi1g05480*	bZIP transcription factor RISBZ2
P3		*Bradi1g05480*	bZIP transcription factor RISBZ2
P3		*Bradi1g05480*	bZIP transcription factor RISBZ2
P2	2	*Bradi1g14170*	dormancy-associated protein 1; Dormancy/auxin associated protein (Auxin_repressed);Component of nitrate treatment specific coexpression subnetwork
P2	3	*Bradi1g15730*	purple acid phosphatase 18;Differentially upregulated in above ground in 20d 18lgt 6dark condition in DS.Allvs1Hr comparison
P2	4	*Bradi1g18780*	PROTEIN EARLY RESPONSIVE TO DEHYDRATION 15; Coexpressed along with genes transcribed in shoot tissue during day; protein EARLY RESPONSIVE TO DEHYDRATION 15
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2	5	*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451

P2	5	*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity;uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P3		*Bradi1g35230*	Motor activity; uncharacterized LOC100838451
P2	6	*Bradi1g47367*	CHD3-type chromatin-remodeling factor PICKLE; HELICASE PROTEIN MOM1-RELATED; Differentially upregulated in above ground in 20d 18lgt 6dark condition in DS.Allvs1Hr comparison
P2	7	*Bradi2g07070*	Homologous to Rice Flowering Time Locus T -Like1
P2	8	*Bradi2g09120*	arachidonic acid binding; syntaxin-22
P2	9	*Bradi2g11080*	MYB transcrition factor; ranscription factor MYB61

P2	10	*Bradi2g11940*	LATE EMBRYOGENESIS ABUNDANT 3 (LEA3) FAMILY PROTEIN;protein SENESCENCE-ASSOCIATED GENE 21, mitochondrial
P2	11	*Bradi2g12720*	SUMO-specific protease activity; Differentially upregulated in above ground in 20d 18lgt 6dark condition in DS.Allvs1Hr comparison; ubiquitin-like-specific protease ESD4
P2	12	*Bradi2g30400*	PROTEIN PLASTID TRANSCRIPTIONALLY ACTIVE 16;Coexpressed along with genes transcribed in shoot tissue during day; protein plastid transcriptionally active 16, chloroplastic
P2	13	*Bradi2g31070*	proactivator polypeptide-like 1;SAPOSIN-RELATED
P2	14	*Bradi3g23520*	F-box/kelch-repeat protein SKIP11
P2	15	*Bradi3g24760*	protein PHYTOCHROME KINASE SUBSTRATE 1;PROTEIN PHYTOCHROME KINASE SUBSTRATE 3
P1	16	*Bradi3g48170*	Plant protein of unknown function (DUF247); Highly expressed in root;
P2	17	*Bradi3g48800*	uncharacterized protein
P2	18	*Bradi3g52680*	Zeaxanthin 7,8-dioxygenase / Zeaxanthin 7,8(7',8')-cleavage dioxygenase;9-cis-epoxycarotenoid dioxygenase NCED1, chloroplastic
P3	19	*Bradi3g53327*	Coexpressed with genes downregulated during drought stress
P2	20	*Bradi3g56560*	ADP-ribosylation factor GTPase-activating protein AGD12; Rac GTPase activator activity
P3	21	*Bradi4g19510*	SKP1-like protein 1; SKP1, subfamily Skp1 family, putative, similar to ASK2 (ARABIDOPSIS SKP1-LIKE 2)
P3	22	*Bradi4g20197*	probable beta-D-xylosidase 7; PERIPLASMIC BETA-GLUCOSIDASE-RELATED // SUBFAMILY NOT NAMED
P2	23	*Bradi4g26366*	Permease of the major facilitator superfamily
P2	24	*Bradi4g39260*	isoflavone reductase homolog; (+)-pinoresinol reductase / Pinoresinol/lariciresinol reductase
P3	25	*Bradi5g01720*	peptidyl-prolyl cis-trans isomerase Pin1; phosphoserine binding, phosphothreonine binding
P2	26	*Bradi5g15050*	F-Box Protein; Differentially downregulated in above ground in 20d 18lgt 6dark condition in CS.Allvs1Hr comparison

Table S2. List of markers utilized throughout the study.

Marker	**Position**	**Primers**			**Allele**	
KASP		**FAM**	**VIC**	**COM**	**Kronos**	**Mut**
Kr-3308[1]	621,763,602	gtacaacgccatgctcaagC	gtacaacgccatgctcaagT	ggagcacgccgCagTattc	FAM	VIC
Kr-2991/3532[1]	616,654,824	tcatgtcattttcgtccgcaG	tcatgtcattttcgtccgcaA	CacaATCgttcattccctaccA	FAM	VIC
Kr-2038[2]	616,655,991	agctgacggtcgtaccctC	agctgacggtcgtaccctT	atgctcaagcagaagctg	FAM	VIC
					Berkut	**UC1419**
bZIPC1-4229[2]	616,654,229	ctgTtggcgcaagtgggA	ctgTtggcgcaagtgggG	ccccattcccagtgatctcG	FAM	VIC
bZIPC1-4272[2]	616,654,272	gcttttggacacatattaggCaaG	gcttttggacacatattaggCaaT	gacaTcccacttgcgccaA	FAM	VIC
CAPS		**5B-1.1F**	**5B-5.1R**	**Restriction enzyme**	**Kronos**	**Mut**
5B-Splicing[3]	NA	catgctcaagcagaagctg	gccttaacagtgaggagttttc	NA	High	Low
Real-time PCR		**bZIPC1-AB-qPCR-F1**	**bZIPC1-AB-qPCR-R1**			
bZIPC1	NA	cggaggaggaagcaatcc	ggtgaagggcatgctgag			

1 94 °C for 15 min, 10 cycles of: 94 °C 20 s, -0.8 °C touch down from 65 to 57 °C 1 min, and 30 cycles of: 94 °C 20s, 57 °C 1 min.

1 94 °C for 15 min, 10 cycles of: 94 °C 20 s, -0.6 °C touch down from 61 to 55 °C 1 min, and 30 cycles of: 94 °C 20s, 55 °C 1 min.

3 94 °C for 5 min, 45 cycles of: 94 °C 30 s, 55 °C 30 s, 72 °C 30 s, and a final cycle of: 72 °C for 5 min

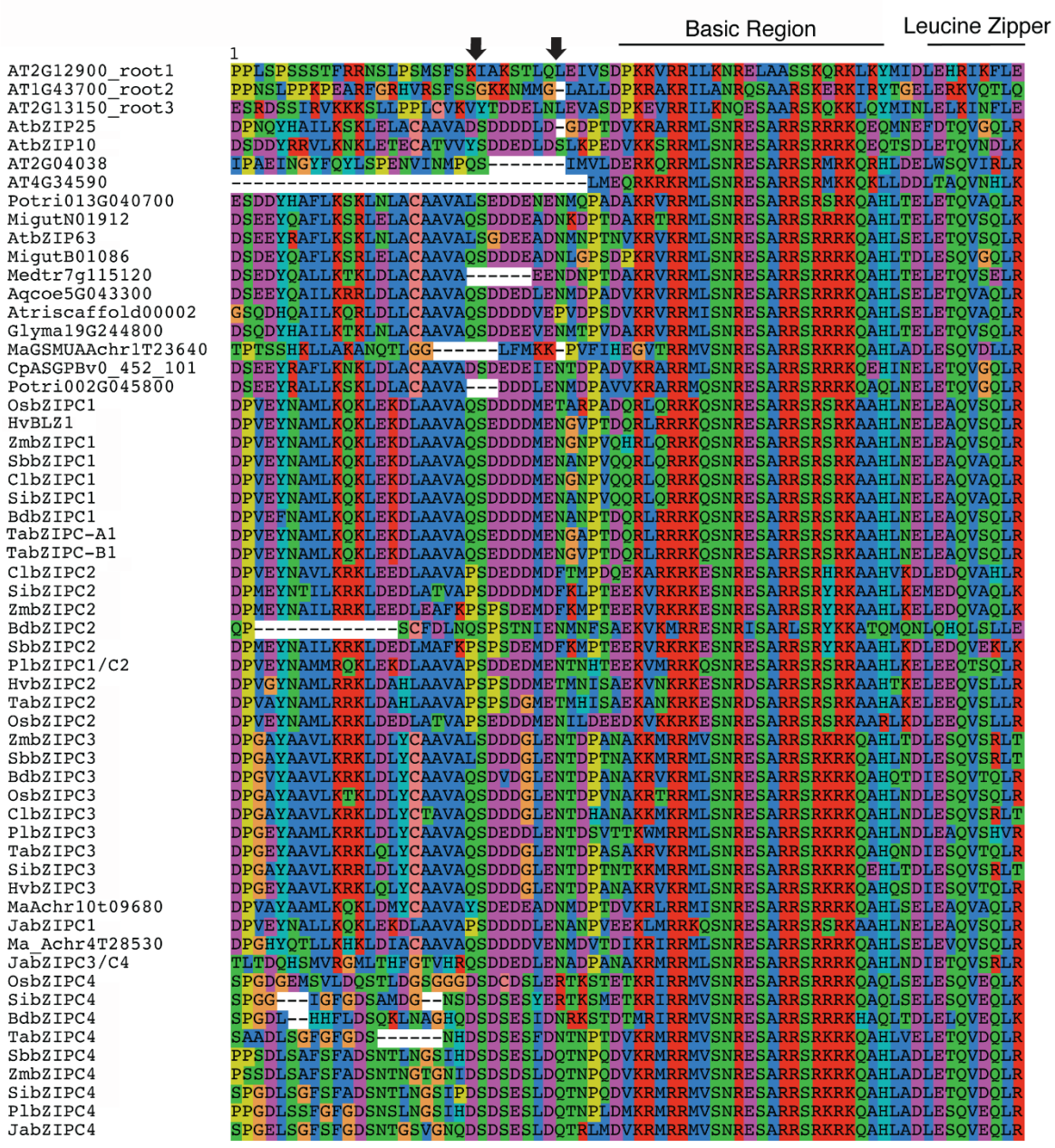
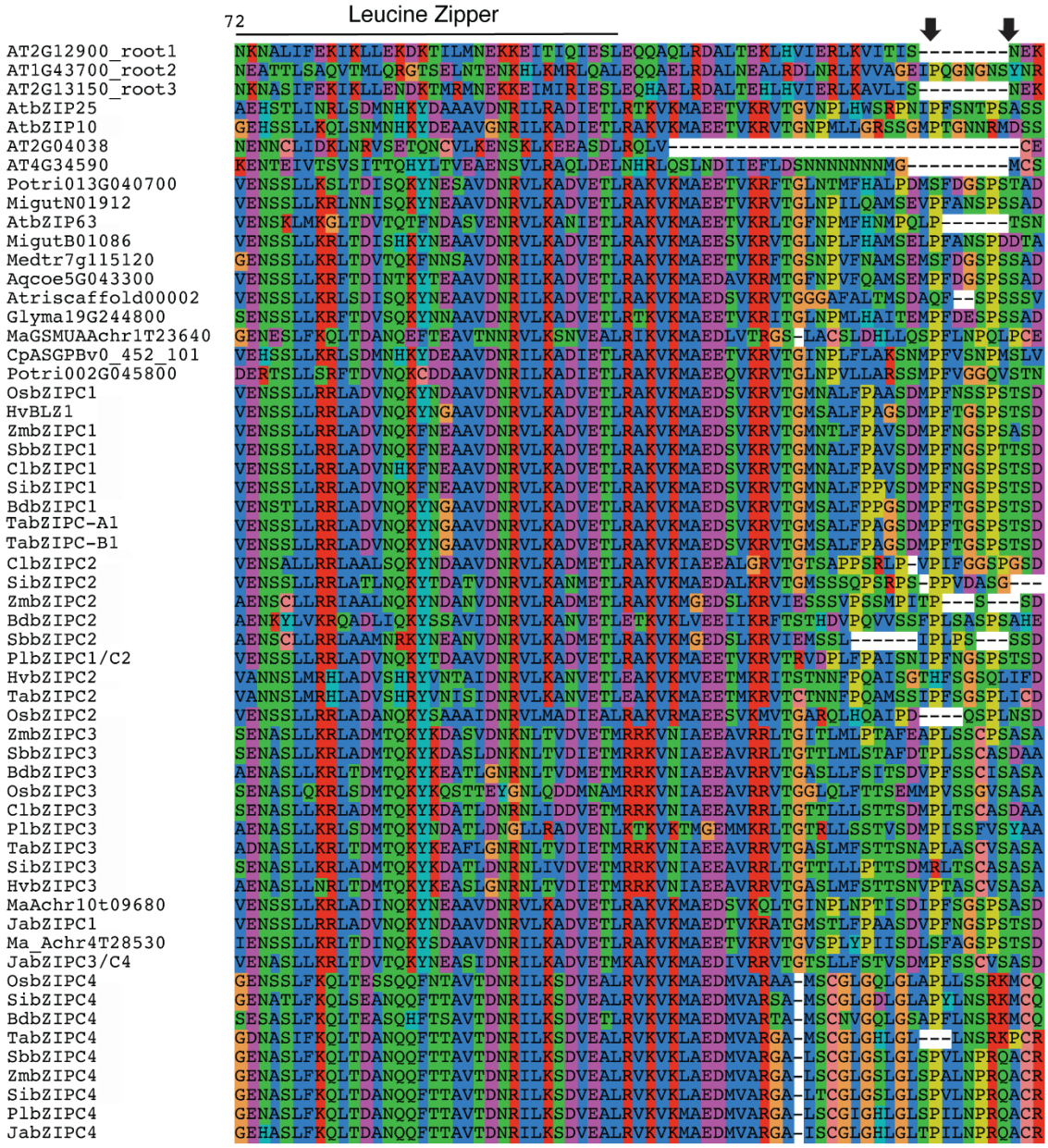
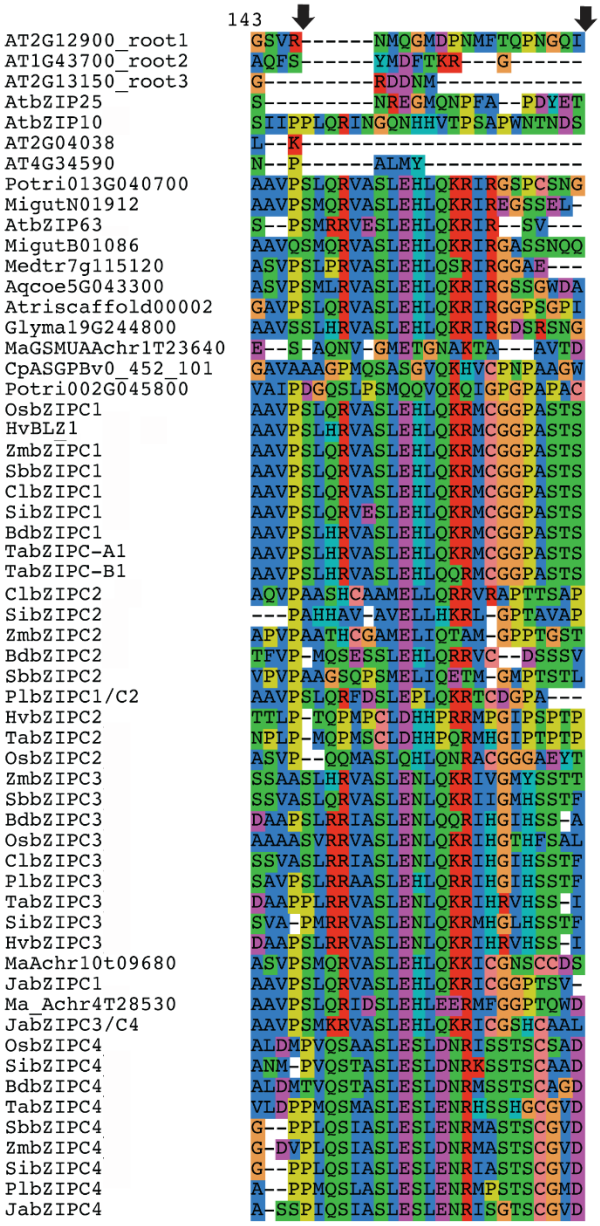


Figure S1. Amino acid sequence alignment of bZIP C subfamily proteins used for the phylogenetic analysis in Figure 1. The alignment was visualized using SeaView (Gouy et al., 2010). Arrows indicate places in which regions of the bZIPC1 protein were deleted to include only alignable regions within the protein across all taxa and paralogs


72
Leucine Zipper
AT2G12900_root1
AT1G43700_root2
AT2G13150_root3
AtbZIP25
AtbZIP10
AT2G04038
AT4G34590
Potri013G040700
MigutN01912
AtbZIP63
MigutB01086
Medtr7g115120
Aqcoe5G043300
Atriscaffold00002
Glyma19G244800
MaGSMUAAchr1T23640
CpASGPBv0_452_101
Potri002G045800
OsbZIPC1
HvBLZ1
ZmbZIPC1
SbbZIPC1
ClbZIPC1
SibZIPC1
BdbZIPC1
TabZIPC-A1
TabZIPC-B1
ClbZIPC2
SibZIPC2
ZmbZIPC2
BdbZIPC2
SbbZIPC2
PlbZIPC1/C2
HvbZIPC2
TabZIPC2
OsbZIPC2
ZmbZIPC3
SbbZIPC3
BdbZIPC3
OsbZIPC3
ClbZIPC3
PlbZIPC3
TabZIPC3
SibZIPC3
HvbZIPC3
MaAchr10t09680
JabZIPC1
Ma_Achr4T28530
JabZIPC3/C4
OsbZIPC4
SibZIPC4
BdbZIPC4
TabZIPC4
SbbZIPC4
ZmbZIPC4
SibZIPC4
PlbZIPC4
JabZIPC4


143
AT2G12900_root1
AT1G43700_root2
AT2G13150_root3
AtbZIP25
AtbZIP10
AT2G04038
AT4G34590
Potri013G040700
MigutN01912
AtbZIP63
MigutB01086
Medtr7g115120
Aqcoe5G043300
Atriscaffold00002
Glyma19G244800
MaGSMUAAchr1T23640
CpASGPBv0_452_101
Potri002G045800
OsbZIPC1
HvBLZ1
ZmbZIPC1
SbbZIPC1
ClbZIPC1
SibZIPC1
BdbZIPC1
TabZIPC-A1
TabZIPC-B1
ClbZIPC2
SibZIPC2
ZmbZIPC2
BdbZIPC2
SbbZIPC2
PlbZIPC1/C2
HvbZIPC2
TabZIPC2
OsbZIPC2
ZmbZIPC3
SbbZIPC3
BdbZIPC3
OsbZIPC3
ClbZIPC3
PlbZIPC3
TabZIPC3
SibZIPC3
HvbZIPC3
MaAchr10t09680
JabZIPC1
Ma_Achr4T28530
JabZIPC3/C4
OsbZIPC4
SibZIPC4
BdbZIPC4
TabZIPC4
SbbZIPC4
ZmbZIPC4
SibZIPC4
PlbZIPC4
JabZIPC4

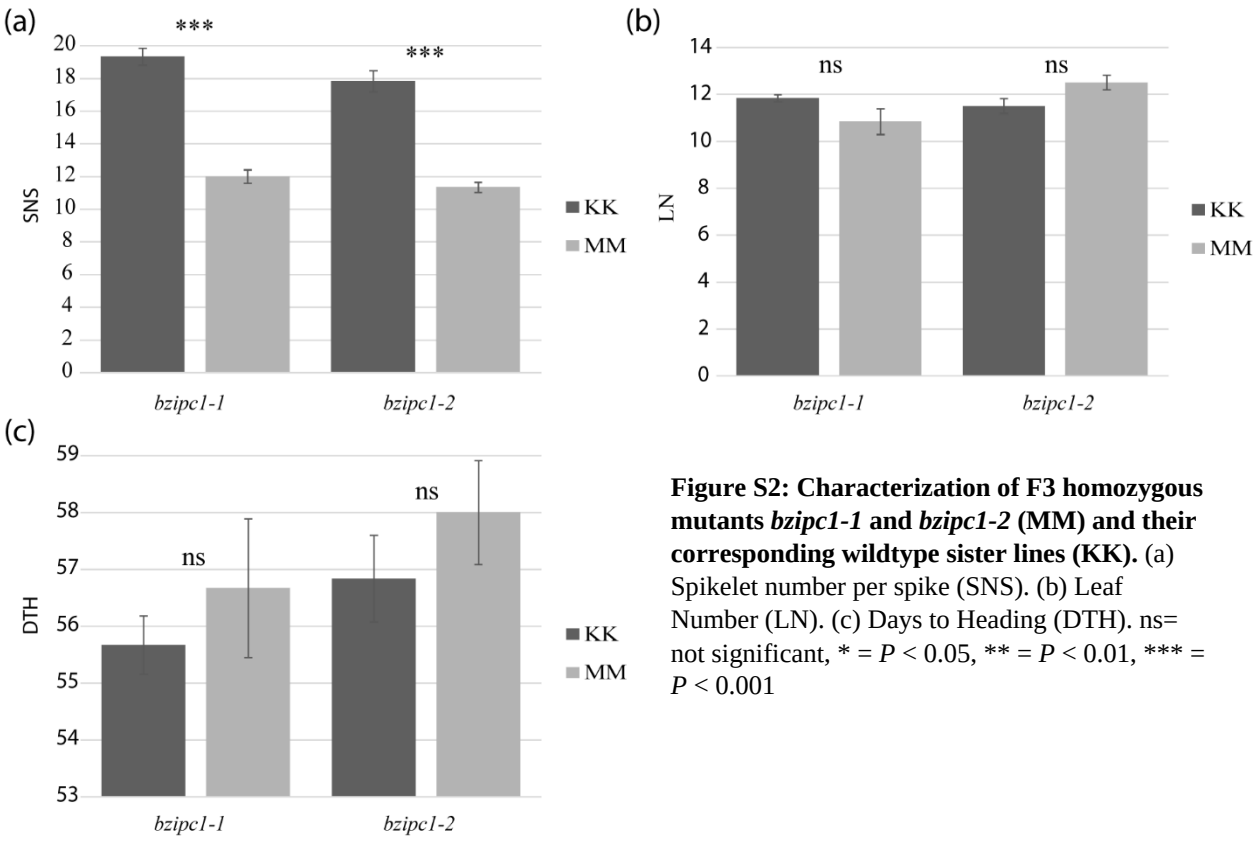


Figure S2: Characterization of F3 homozygous mutants *bzipc1-1* and *bzipc1-2* (MM) and their corresponding wildtype sister lines (KK). (a) Spikelet number per spike (SNS). (b) Leaf Number (LN). (c) Days to Heading (DTH). ns= not significant, * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$

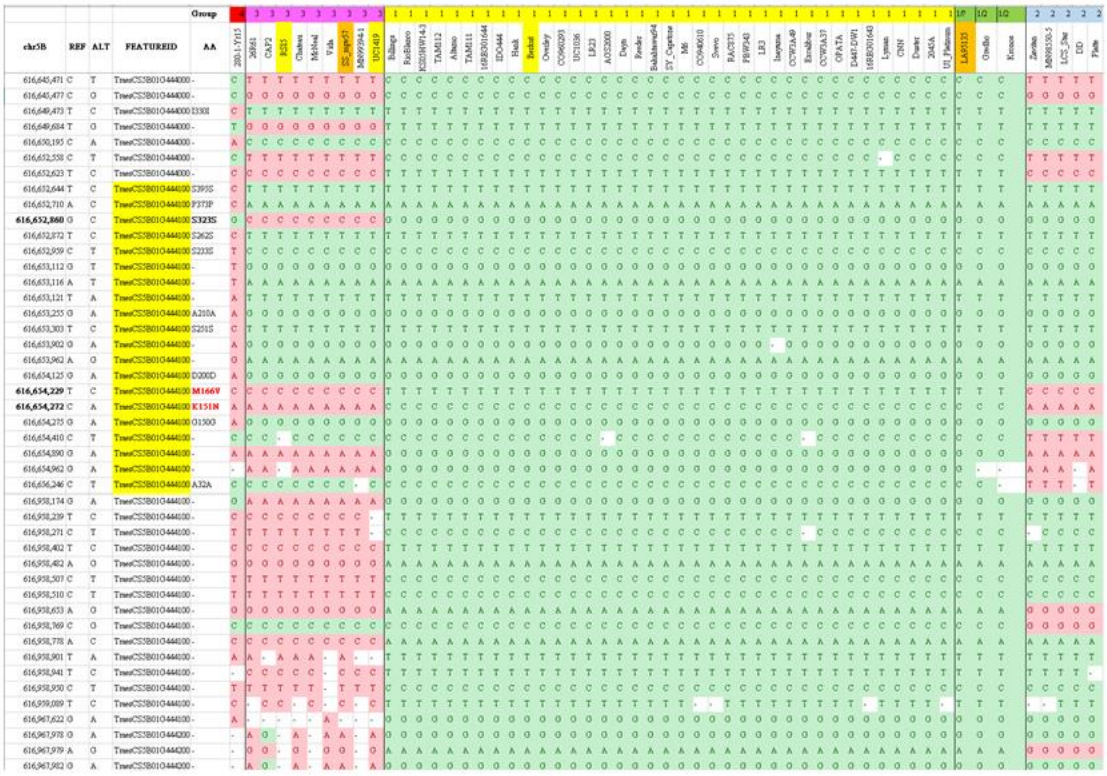


Figure S3: Section of the haplotype analysis including *bZIPC1* (*TraesCS5B02G444100*).

CHAPTER 3 REFERENCES

Agarwal, P., Baranwal, V. K., & Khurana, P. (2019). Genome-wide analysis of bZIP transcription factors in wheat and functional characterization of a *TabZIP* under abiotic stress. *Scientific Reports*, *9*, 1–18.

Agostini, R. B., Postigo, A., Rius, S. P., Rech, G. E., Campos-Bermudez, V. A., & Vargas, W. A. (2019). Long-lasting primed state in maize plants: Salicylic acid and steroid signaling pathways as key players in the early activation of immune responses in silks. *Molecular Plant-Microbe Interactions*, *32*, 90–106.

Albani, D., Hammond-Kosack, M. C. U., Smith, C., Conlan, S., Colot, V., Holdsworth, M., & Bevan, M. W. (1997). The wheat transcriptional activator SPA: A seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *The Plant Cell*, *9*, 171–184. https://about.jstor.org/terms

Alonso, R., Oñ ate-Sá nchez, L., Weltmeier, F., Ehlert, A., Diaz, I., Dietrich, K., Vicente-Carbajosa, J., & Drö ge-Laser, W. (2009). A pivotal role of the basic leucine zipper transcription factor bZIP53 in the regulation of *Arabidopsis* seed maturation gene expression based on heterodimerization and protein complex formation. *The Plant Cell*, *21*, 1747–1761.

Alves, M. S., Dadalto, S. P., Gonçalves, A. B., de Souza, G. B., Barros, V. A., & Fietto, L. G. (2013). Plant bZIP transcription factors responsive to pathogens: A Review. *International Journal of Molecular Science*, *14*, 7815–7828.

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., … Distelfeld, A. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, *357*, 93–97.

Baena-González, E., Rolland, F., Thevelein, J. M., & Sheen, J. (2007). A central integrator of transcription networks in plant stress and energy signalling. *Nature*, *448*, 938–942.

Ballerini, E. S., & Kramer, E. M. (2011). In the light of evolution: A reevaluation of conservation in the *CO-FT* regulon and its role in photoperiodic regulation of flowering time. *Frontiers in Plant Science*, *2*.

Blake, N. K., Pumphrey, M., Glover, K., Chao, S., Jordan, K., Jannick, J.-L., Akhunov, E. A., Dubcovsky, J., Bockelman, H., & Talbert, L. E. (2019). Registration of the Triticeae-CAP spring wheat nested association mapping population. *Journal of Plant Registrations*, *13*, 294–297.

Blake, V., Birkett, C., Matthews, D. E., Hane, D. L., Bradbury, P., & Jannink, J.-L. (2016). The Triticeae Toolbox: combining phenotype and genotype data to advance small-grains breeding. *The Plant Genome*, *9*.

Boden, S. A., Cavanagh, C., Cullis, B. R., Ramm, K., Greenwood, J., Jean Finnegan, E., Trevaskis, B., & Swain, S. M. (2015). *Ppd-1* is a key regulator of inflorescence architecture and paired spikelet development in wheat. *Nature Plants*, 14016.

Boudet, J., Merlino, M., Plessis, A., Gaudin, J.-C., Dardevet, M., Perrochon, S., Alvarez, D., Risacher, T., Martre, P., & Ravel, C. (2019). The bZIP transcription factor SPA Heterodimerizing Protein represses glutenin synthesis in *Triticum aestivum*. *The Plant Journal*, *97*, 858–871.

Brassac, J., Muqaddasi, Q. H., Plieske, J., Ganal, M. W., & Röder, M. S. (2021). Linkage mapping identifies a non-synonymous mutation in *FLOWERING LOCUS T (FT-B1)* increasing spikelet number per spike. *Scientific Reports*, *11*, 1585.

Cao, L., Lu, X., Zhang, P., Wang, G., Wei, L., & Wang, T. (2019). Systematic analysis of differentially expressed maize *ZmbZIP* genes between drought and rewatering transcriptome reveals bZIP family members involved in abiotic stress responses. *International Journal of Molecular Sciences*, *20*, 4069.

Cao, S., Siriwardana, C. L., Kumimoto, R. W., & Holt, B. F. (2011). Construction of high quality Gateway™ entry libraries and their application to yeast two-hybrid for the monocot model plant *Brachypodium distachyon*. *BMC Biotechnology*, *11*, 53.

Casao, M. C., Karsai, I., Igartua, E., Gracia, M. P., Veisz, O., & Casas, A. M. (2011). Adaptation of barley to mild winters: A role for *PPDH2*. *BMC Plant Biology*, *11*, 164.

Chardon, F., & Damerval, C. (2005). Phylogenomic analysis of the PEBP gene family in cereals. *Journal of Molecular Evolution*, *61*, 579–590.

Chen, A., Li, C., Hu, W., Lau, M. Y., Lin, H., Rockwell, N. C., Martin, S. S., Jernstedt, J. A., Lagarias, J. C., & Dubcovsky, J. (2014). PHYTOCHROME C plays a major role in the acceleration of wheat flowering under long-day photoperiod. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 10037–10044.

Chen, H., Dai, X. J., & Gu, Z. Y. (2015). OsbZIP33 is an ABA-dependent enhancer of drought tolerance in rice. *Crop Science*, *55*, 1673–1685.

Chen, Z., Cheng, X., Chai, L., Wang, Z., Du, D., Wang, Z., Bian, R., Zhao, A., Xin, M., Guo, W., Hu, Z., Peng, H., Yao, Y., Sun, Q., & Ni, Z. (2020). Pleiotropic QTL influencing spikelet number and heading date in common wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, *133*, 1825–1838.

Choi, H. I., Hong, J. H., Ha, J. O., Kang, J. Y., & Kim, S. Y. (2000). ABFs, a family of ABA-responsive element binding factors. *Journal of Biological Chemistry*, *275*, 1723–1730.

Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., Pingault, L., Sourdille, P., Couloux, A., Paux, E., Leroy, P., Mangenot, S., Guilhot, N., le Gouis, J., Balfourier, F., Alaux, M., Jamilloux, V., Poulain, J., Durand, C., … Feuillet, C. (2014). Structural and functional partitioning of bread wheat chromosome 3B. *Science*, *345*.

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., & Coupland, G. (2007). FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science*, *316*, 1030–1033.

Deppmann, C. D., Alvania, R. S., & Taparowsky, E. J. (2006). Cross-Species annotation of basic leucine zipper factor interactions: Insight into the evolution of closed interaction networks. *Molecular Biology and Evolution*, *23*, 1480–1492.

DeWitt, N., Guedira, M., Lauer, E., Murphy, J. P., Marshall, D., Mergoum, M., Johnson, J., Holland, J. B., & Brown-Guedira, G. (2021). Characterizing the oligogenic architecture of plant growth phenotypes informs genomic selection approaches in a common wheat population. *BMC Genomics*, *22*, 402.

Distelfeld, A., Li, C., & Dubcovsky, J. (2009). Regulation of flowering in temperate cereals. *Current Opinion in Plant Biology*, *12*, 178–184.

Dixon, L. E., Greenwood, J. R., Bencivenga, S., Zhang, P., Cockram, J., Mellers, G., Ramm, K., Cavanagh, C., Swain, S. M., & Boden, S. A. (2018). *TEOSINTE BRANCHED1* regulates inflorescence architecture and development in bread wheat (*Triticum aestivum*). *The Plant Cell*, *30*, 563–581.

Dobrovolskaya, O., Pont, C., Sibout, R., Martinek, P., Badaeva, E., Murat, F., Chosson, A., Watanabe, N., Prat, E., Gautier, N., Gautier, V., Poncet, C., Orlov, Y. L., Krasnikov, A. A., Bergès, H., Salina, E., Laikova, L., & Salse, J. (2014). *FRIZZY PANICLE* drives supernumerary spikelets in bread wheat. *Plant Physiology*, *167*, 189–199.

Dröge-Laser, W., & Weiste, C. (2018). The C/S1 bZIP network: A regulatory hub orchestrating plant energy homeostasis. *Trends in Plant Science*, *23*, 422–433.

Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*, 113.

Ehlert, A., Weltmeier, F., Wang, X., Mayer, C. S., Smeekens, S., Vicente-Carbajosa, J., & Dröge-Laser, W. (2006). Two-hybrid protein-protein interaction analysis in Arabidopsis protoplasts: Establishment of a heterodimerization map of group C and group S bZIP transcription factors. *The Plant Journal*, *46*, 890–900.

Faure, S., Higgins, J., Turner, A., & Laurie, D. A. (2007). The *FLOWERING LOCUS T*-like gene family in barley (*Hordeum vulgare*). *Genetics*, *176*, 599.

Feng, Y., Wang, Y., Zhang, G., Gan, Z., Gao, M., Lv, J., Wu, T., Zhang, X., Xu, X., Yang, S., & Han, Z. (2021). Group-C/S1 bZIP heterodimers regulate *MdIPT5b* to negatively modulate drought tolerance in apple species. *The Plant Journal*, *107*, 399–417.

Fox, J., & Weisberg, S. (2020). *Using car and effects Functions in Other Functions*. Computer Program.

Frank, A., Matiolli, C. C., Viana, A. J. C., Hearn, T. J., Kusakina, J., Belbin, F. E., Wells Newman, D., Yochikawa, A., Cano-Ramirez, D. L., Chembath, A., Cragg-Barber, K., Haydon, M. J., Hotta, C. T., Vincentz, M., Webb, A. A. R., & Dodd, A. N. (2018). Circadian entrainment in *Arabidopsis* by the sugar-responsive transcription factor bZIP63. *Current Biology*, *28*, 2597-2606.e6.

Fujita, Y., Fujita, M., Satoh, R., Maruyama, K., Parvez, M. M., Seki, M., Hiratsu, K., Ohme-Takagi, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2005). AREB1 is a transcription activator of novel ABRE-Dependent ABA signaling that enhances drought stress tolerance in *Arabidopsis*. *The Plant Cell*, *17*, 3470–3488.

Glenn, P., Zhang, J., Brown-Guedira, G., Dewitt, N., Cook, J. P., Li, K., Akhunov, E., & Dubcovsky, J. (2022). Identification and characterization of a natural polymorphism in *FT-A2* associated with increased number of grains per spike in wheat. *Theoretical and Applied Genetics*, *135*, 679–692.

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView Version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, *27*, 221–224.

Guedes Corrêa, L. G., Riaño-Pachón, D. M., Guerra Schrago, C., Vicentini dos Santos, R., Mueller-Roeber, B., & Vincentz, M. (2008). The role of bZIP transcription factors in green plant evolution: Adaptive features emerging from four founder genes. *PLoS ONE*, *3*, e2944.

Guo, D., Hou, Q., Zhang, R., Lou, H., Li, Y., Zhang, Y., You, M., Xie, C., Liang, R., & Li, B. (2020). Over-expressing *TaSPA-B* reduces prolamin and starch accumulation in wheat (*Triticum aestivum* L.) grains. *International Journal of Molecular Sciences*, *21*, 3257.

Hai, L., Guo, H., Wagner, C., Xiao, S., & Friedt, W. (2008). Genomic regions for yield and yield parameters in Chinese winter wheat (*Triticum aestivum* L.) genotypes tested under varying environments correspond to QTL in widely different wheat materials. *Plant Science*, *175*, 226–232.

Halliwell, J., Borrill, P., Gordon, A., Kowalczyk, R., Pagano, M. L., Saccomanno, B., Bentley, A. R., Uauy, C., & Cockram, J. (2016). Systematic Investigation of *FLOWERING LOCUS T*-like Poaceae gene families identifies the short-day expressed flowering pathway gene *TaFT3* in wheat (*Triticum aestivum* L.). *Frontiers in Plant Science*, *7*, 857.

He, F., Pasam, R., Shi, F., Kant, S., Keeble-Gagnere, G., Kay, P., Forrest, K., Fritz, A., Hucl, P., Wiebe, K., Knox, R., Cuthbert, R., Pozniak, C., Akhunova, A., Morrell, P. L., Davies, J. P., Webb, S. R., Spangenberg, G., Hayes, B., … Akhunov, E. (2019). Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nature Genetics*, *51*, 896–904.

Higgins, J. A., Bailey, P. C., & Da, L. (2010). Comparative genomics of flowering time pathways using *Brachypodium distachyon* as a model for the temperate grasses. *PLoS ONE*, *5*, e10065.

Hurst GC. (1995). *Transcription Factors 1: bZIP Protein Profile 2 proteins*. 101–168.

Isham, K., Wang, R., Zhao, W., Wheeler, J., Klassen, N., Akhunov, · Eduard, & Chen, · Jianli. (2021). QTL mapping for grain yield and three yield components in a population derived from two high-yielding spring wheat cultivars. *Theoretical and Applied Genetics*, *134*, 2079–2095.

Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., & Parcy, F. (2002). bZIP transcription factors in *Arabidopsis*. *Trends in Plant Science*, *7*, 106–111.

Jordan, K. W., Wang, S., Lun, Y., Gardiner, L.-J., MacLachlan, R., Hucl, P., Wiebe, K., Wong, D., Forrest, K. L., Sharpe, A. G., Sidebottom, C. H., Hall, N., Toomajian, C., Close, T., Dubcovsky, J., Akhunova, A., Talbert, L., Bansal, U. K., Bariana, H. S., … Akhunov, E. (2015). A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. *Genome Biology*, *16*, 48.

Kaminaka, H., Näke, C., Epple, P., Dittgen, J., Schütze, K., Chaban, C., Holt, B. F., Merkle, T., Schäfer, E., Harter, K., & Dangl, J. L. (2006). bZIP10-LSD1 antagonism modulates basal defense and cell death in *Arabidopsis* following infection. *EMBO Journal*, *25*, 4400–4411.

Kawakatsu, T., Yamamoto, M. P., Touno, S. M., Yasuda, H., & Takaiwa, F. (2009). Compensation and interaction between RISBZ1 and RPBF during grain filling in rice. *The Plant Journal*, *59*, 908–920.

Kikuchi, R., Kawahigashi, H., Ando, T., Tonooka, T., & Handa, H. (2009). Molecular and functional characterization of PEBP genes in barley reveal the diversification of their roles in flowering. *Plant Physiology*, *149*, 1341–1353.

Krasileva, K. v., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J., Ramirez-Gonzalez, R. H., Wang, X., Borrill, P., Fosker, C., Ayling, S., Phillips, A. L., Uauy, C., & Dubcovsky, J. (2017). Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, E913–E921.

Kumar, D., & Chattopadhyay, S. (2018). Glutathione modulates the expression of heat shock proteins via the transcription factors BZIP10 and MYB21 in Arabidopsis. *Journal of Experimental Botany*, *69*, 3729–3743.

Kuzay, S., Lin, H., Li, C., Chen, S., Woods, D. P., Zhang, J., Lan, T., von Korff, M., & Dubcovsky, J. (2022). *WAPO-A1* is the causal gene of the 7AL QTL for spikelet number per spike in wheat. *PLOS Genetics*, *18*, e1009747.

Kuzay, S., Xu, Y., Zhang, J., Katz, A., Pearce, S., Su, Z., & Fraser, M. (2019). Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theoretical and Applied Genetics*, *132*, 2689–2705.

Li, C., & Dubcovsky, J. (2008). Wheat FT protein regulates VRN1 transcription through interactions with FDL2. *The Plant Journal*, *55*, 543–554.

Li, C., Lin, H., Chen, A., Lau, M., Jernstedt, J., & Dubcovsky, J. (2019). Wheat *VRN1*, *FUL2* and *FUL3* play critical and redundant roles in spikelet development and spike determinacy. *Development*, *146*, dev1753981.

Li, C., Lin, H., & Dubcovsky, J. (2015). Factorial combinations of protein interactions generate a multiplicity of florigen activation complexes in wheat and barley. *The Plant Journal*, *84*, 70.

Li, K., Debernardi, J. M., Li, C., Lin, H., Zhang, C., Jernstedt, J., Korff, M. von, Zhong, J., & Dubcovsky, J. (2021). Interactions between SQUAMOSA and SHORT VEGETATIVE PHASE MADS-box proteins regulate meristem transitions during wheat spike development. *The Plant Cell*, *33*, 3621–3644.

Li, Y., Qian, Q., Zhou, Y., Yan, M., Sun, L., Zhang, M., Fu, Z., Wang, Y., Han, B., Pang, X., Chen, M., & Li, J. (2020). *BRITTLE CULM1*, which encodes a COBRA-like protein, affects the mechanical properties of rice plants. *The Plant Cell*, *15*, 2020–2031.

Liu, C., Mao, B., Ou, S., Wang, W., Liu, L., Wu, Y., Chu, C., & Wang, X. (2014). OsbZIP71, a bZIP transcription factor, confers salinity and drought tolerance in rice. *Plant Molecular Biology*, *84*, 19–36.

Liu, Z., Xin, M., Qin, J., Peng, H., Ni, Z., Yao, Y., & Sun, Q. (2015). Temporal transcriptome profiling reveals expression partitioning of homoeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum* L.). *BMC Plant Biology*, *15*, 152.

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*, e494171.

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., … Cattivelli, L. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics*, *51*, 885–895.

Maddison, W., & Maddison, D. (2007). *Mesquite: a modular system for evolutionary analysis*. *Version 2.0*. http://mesquiteproject.org

Mair, A., Pedrotti, L., Wurzinger, B., Anrather, D., Simeunovic, A., Weiste, C., Valerio, C., Dietrich, K., Kirchler, T., Agele, T. N. ¨, Jesús, J. J., Carbajosa, V., Hanson, J., Baena-Gonzálezgonz´gonzález, E., Chaban, C., Weckwerth, W., Dröge, W., Dröge-Laser, D., & Teige, M. (2015). SnRK1-triggered switch of bZIP63 dimerization mediates the low-energy response in plants. *ELife*.

Matiolli, C. C., Tomaz, J. P., Duarte, G. T., Prado, F. M., Vieira, E., Bem, D., Silveira, A. B., Gauer, L., Gustavo, L., Corrêa, G., Duarte Drumond, R., Carvalho Viana, A. J., Mascio, P. di, Meyer, C., & Vincentz, M. (2011). The Arabidopsis bZIP gene AtbZIP63 is a sensitive integrator of transient abscisic acid and glucose signals. *Plant Physiology*, *157*, 692–705.

Nickless, A., Bailis, J. M., & You, Z. (2017). Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell & Bioscience*, *7*, 26.

Nitcher, R., Pearce, S., Tranquilli, G., Zhang, X., & Dubcovsky, J. (2014). Effect of the Hope *FT-B1* allele on wheat heading time and yield components. *Journal of Heredity*, *105*, 666–675.

Nylander, J. (2004). *MrModeltest v2. Program*. https://github.com/nylander/MrModeltest2

Oñate, L., Vicente-Carbajosa, J., Lara, P., Díaz, I., & Carbonero, P. (1999). Barley BLZ2, a seed-specific bZIP protein that interacts with BLZ1 in vivo and activates transcription from the GCN4-like motif of *B-hordein* promoters in barley endosperm. *Journal of Biological Chemistry*, *274*, 9175–9182.

Paterson, A. H., Bowers, J. E., & Chapman, B. A. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 9903–9908.

Pearce, S., Vanzetti, L. S., & Dubcovsky, J. (2013). Exogenous gibberellins induce wheat spike development under short days only in the presence of *VERNALIZATION1*. *Plant Physiology*, *163*, 1433–1445.

Peviani, A., Lastdrager, J., Hanson, J., & Snel, B. (2016). The phylogeny of C/S1 bZIP transcription factors reveals a shared algal ancestry and the pre-angiosperm translational regulation of S1 transcripts. *Scientific Reports*, *6*, 30444.

Pin, P. A., Benlloch, R., Bonnet, D., Wremerth-Weich, E., Kraft, T., Gielen, J. J. L., & Nilsson, O. (2010). An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science*, *330*, 1397–1400.

Pysh, L. D., Aukerman, M. J., & Schmidt', R. J. (1993). OHPI: A maize basic domain/leucine zipper protein that interacts with Opaque2. *The Plant Cell*, *5*, 227–236.

Rabbani, M. A., Maruyama, K., Abe, H., Ayub Khan, M., Katsura, K., Ito, Y., Yoshiwara, K., Seki, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2003). Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiology*, *133*, 1755–1767.

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE*, *8*, 66428.

Reynolds, M., Foulkes, J., Furbank, R., Griffiths, S., King, J., Murchie, E., Parry, M., & Slafer, G. (2012). Achieving yield gains in wheat. *Plant, Cell, & Environment*, *35*, 1799–1823.

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*, 1572–1574.

Schütze, K., Harter, K., & Chaban, C. (2008). Post-translational regulation of plant bZIP factors. *Trends in Plant Science*, *13*, 247–255.

Shaw, L. M., Li, C., Woods, D. P., Alvarez, M. A., Lin, H., Lau, M. Y., Chen, A., & Dubcovsky, J. (2020). Epistatic interactions between *PHOTOPERIOD1*, *CONSTANS1* and *CONSTANS2* modulate the photoperiodic response in wheat. *PLOS Genetics*, *16*, e1008812.

Shaw, L. M., Lyu, B., Turner, R., Li, C., Chen, F., Han, X., Fu, D., & Dubcovsky, J. (2019). *FLOWERING LOCUS T2* regulates spike development and fertility in temperate cereals. *Journal of Experimental Botany*, *70*, 193–204.

Shaw, L. M., Turner, A. S., Herry, L., Griffiths, S., & Laurie, D. A. (2013). Mutant alleles of *Photoperiod-1* in Wheat (*Triticum aestivum* L.) that confer a late flowering phenotype in long days. *PLoS ONE*, *8*, e79459.

Tamaki, S., Matsuo, S., Hann, L. W., Yokoi, S., & Shimamoto, K. (2007). Hd3a protein is a mobile flowering signal in rice. *Science*, *316*, 1033–1036.

Taoka, K. I., Ohki, I., Tsuji, H., Furuita, K., Hayashi, K., Yanase, T., Yamaguchi, M., Nakashima, C., Purwestri, Y. A., Tamaki, S., Ogaki, Y., Shimada, C., Nakagawa, A., Kojima, C., & Shimamoto, K. (2011). 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature*, *476*, 332–335.

Teo, C. J., Takahashi, K., Shimizu, K., Shimamoto, K., & Taoka, K. I. (2017). Potato tuber induction is regulated by interactions between components of a tuberigen complex. *Plant and Cell Physiology*, *58*, 365–374.

Tian, Y., Peng, K., Lou, G., Ren, Z., Sun, X., Wang, Z., Xing, J., Song, C., & Cang, J. (2022). Transcriptome analysis of the winter wheat *Dn1* in response to cold stress. *BMC Plant Biology*, *22*, 277.

VanGessel, C., Hamilton, J., Tabbita, F., Dubcovsky, J., & Pearce, S. (2022). Transcriptional signatures of wheat inflorescence development. *BioRxiv*.

Vicente-Carbajosa, J., Oñ, L., Lara, P., Diaz, I., & Carbonero, P. (1998). Barley BLZ1: a bZIP transcriptional activator that interacts with endosperm-specific gene promoters. *The Plant Journal*, *13*, 629–640.

Waddington, S. R., Cartwright, P. M., & Wall, P. C. (1983). A quantitative scale of spike initial and pistil development in barley and wheat. *Annals of Botany*, *51*, 119–130.

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., … Pozniak, C. J. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature*, *588*, 277–283.

Wolde, G. M., Mascher, M., & Schnurbusch, T. (2019). Genetic modification of spikelet arrangement in wheat increases grain number without significantly affecting grain weight. *Molecular Genetics and Genomics*, *294*, 457–468.

Woods, D., Dong, Y., Bouche, F., Bednarek, R., Rowe, M., Ream, T., & Amasino, R. (2019). A florigen paralog is required for short-day vernalization in a pooid grass. *ELife*, *8*, e42153.

Woods, D., Hope, C. L., & Malcomber, S. T. (2011). Phylogenomic analyses of the *BARREN STALK1/LAX PANICLE1 (BA1/LAX1)* genes and evidence for their roles during axillary meristem development. *Molecular Biology and Evolution*, *28*, 2147–2159.

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S., & Dubcovsky, J. (2006). The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proceedings of the National Academy Science of the U.S.A*, *103*, 19581–19586.

Yaniv, E., Raats, D., Ronin, Y., Korol, A. B., Grama, A., Bariana, H., Dubcovsky, J., Schulman, A. H., & Fahima, T. (2015). Evaluation of marker-assisted selection for the stripe rust resistance gene *Yr15*, introgressed from wild emmer wheat. *Molecular Breeding*, *35*, 43.

Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, *14*, 415–421.

Zeller, G., Henz, S. R., Widmer, C. K., Sachsenberg, T., Rä Tsch, G., Weigel, D., & Laubinger, S. (2009). Stress-induced changes in the *Arabidopsis thaliana* transcriptome analyzed using whole-genome tiling arrays. *The Plant Journal*, *58*, 1068–1082.

Zhan, J., Li, G., Ryu, C.-H., Ma, C., Zhang, S., Lloyd, A., Hunter, B. G., Larkins, B. A., Drews, G. N., Wang, X., & Yadegari, R. (2018). Opaque-2 regulates a complex gene network associated with cell differentiation and storage functions of maize endosperm. *The Plant Cell*, *30*, 2425–2446.

Zhang, J., Gizaw, S. A., Bossolini, E., Hegarty, J., Howell, T., Carter, A. H., Akhunov, E., & Dubcovsky, J. (2018). Identification and validation of QTL for grain yield and plant water status under contrasting water treatments in fall-sown spring wheats. *Theoretical and Applied Genetics*, *131*, 1741–1759.

Zhang, J., Gizaw, S. A., Eligio Bossolini, Hegarty, J., Howell, T., Carter, A. H., Akhunov, E., & Dubcovsky, J. (2018). Identification and validation of QTL for grain yield and plant water status under contrasting water treatments in fall-sown spring wheats. *Theoretical and Applied Genetics*, *131*, 1741–1759.

Zhang, X., Jia, H., Li, T., Wu, J., Nagarajan, R., Lei, L., Powers, C., Kan, C., Hua, W., Liu, Z., Chen, C., Carver, B. F., & Yan, L. (2022). *TaCol-B5* modifies spike architecture and enhances grain yield in wheat. *Science*, *183*, 180–183.

Zhang, Z., Yang, J., & Wu, Y. (2015). Transcriptional regulation of Zein gene expression in maize through the additive and synergistic action of opaque2, prolamine-box binding factor, and O2 heterodimerizing proteins. *The Plant Cell*, *27*, 1162–1172.

Zikhali, M., Wingen, L. U., Leverington-Waite, M., Specel, S., & Griffiths, S. (2017). The identification of new candidate genes *Triticum aestivum FLOWERING LOCUS T3-B1 (TaFT3-B1)* and *TARGET OF EAT1 (TaTOE1-B1)* controlling the short-day photoperiod response in bread wheat. *Plant Cell and Envrionment*, *40*, 2678–2690.

CONCLUSION

Our results suggest that the *FT-A2* D10A polymorphism is associated with increases in spikelet number per spike (SNS) and grain number per spike (GNS). While the preceding *ft2*-null mutants decreased fertility, we found that in all four of our introgressed durum varieties, the sister lines with the *FT-A2* A10 allele, had significantly higher SNS with no significant decreases in fertility. The consistency in fertility and improved SNS contributed to a positive correlation between SNS and GNS, leading to a significantly higher GNS overall. Interestingly, we found that the magnitude of increased SNS by the A10 allele, significantly differed between varieties with the highest SNS variety, Desert Gold, increasing the least.

However, there was also a significant negative correlation between GNS and thousand kernel weight (TKW). While the reduction in TKW differed in magnitude depending on the variety as well, overall, it negated the positive effects from increased GNS, leading to no differences in spike yield or plot yield. This negative correlation limits the value of deploying the *FT-A2* A10 allele into durum wheat. However, it is still possible that in different genetic backgrounds with higher biomass or in different environments this allele can have a positive contribution to grain yield.

In spite of the importance of FT2 in the regulation of SNS, nothing is currently known about its interactions with other proteins. While *FT2* is the closest *FT*-like paralog to *FT1* with 78% protein identity, unlike FT1, it did not interact with any of the previously tested 14-3-3 nor FD-like proteins. So, utilizing a yeast-2-hybrid (Y2H) screen we discovered the first protein interactor of FT2, a bZIP-containing transcription factor from the C-group, which we dubbed *bZIPC1*. Interaction studies between bZIPC1 and FT2 revealed that bZIPC1 interacted with FT2

proteins carrying both the D10 and A10 alleles. We also discovered that in addition to FT2, bZIPC1 can interact with FT3 and weakly with FT5.

To understand the evolutionary history of the C-group proteins, we performed a Bayesian phylogenetic analysis utilizing proteins spanning flowering plant diversification and identified four clades, dubbed bZIPC1 – C4. We then tested if the wheat paralogs within these clades could interact with FT-like proteins as well and found that each of the C-group clades interacted with a different set of FT- and CEN-like encoded proteins. These different protein-protein interaction profiles are consistent with the different functions reported for the bZIP proteins of the different C-group clades.

To explore the function of *bZIPC1* in wheat, we developed knock-out mutants of both homoeologs in tetraploid wheat (*bzipc-A1 bzipc-B1*). Previous *ft2*-null mutants had increased SNS and given the interaction between FT2 and bZIPC1 proteins, we expected a similar phenotype for the combined *bzipc1* mutant. Surprisingly however, the double *bzipc1* mutants instead significantly reduced SNS, with a limited effect on heading date. These truncated *bzipc1* mutants also showed significantly reduced *bZIPC1* expression compared to their WT sister lines. However, we could not find significant differences in expression levels for any of the known SNS developmental genes tested in our study, indicating that bZIPC1 may be influencing SNS through a different pathway. We could not find any reports of *bZIP* genes from the C-group affecting grass inflorescence development, indicating that the effect of *bZIPC1* on wheat spike development likely represents a novel result.

We then investigated the natural variation within bZIPC1 and found four haplotypes in the bZIPC1 region. The bZIPC1 protein associated with the derived H1 haplotype had two amino acid changes which differentiated it from the other haplotypes. The H1 haplotype was associated

with increased SNS, GNS and Spike Yield, which is encouraging, since it shows that the increases in GNS were not completely offset by a decrease in TKW. In addition, the frequency of the H1 haplotype increased from 37% in the durum landraces to over 50% in the breeding durum lines, and to 73% in common wheat lines, suggesting positive selection and supporting a positive effect of this gene on grain yield. These results suggest that both durum and common varieties can potentially benefit from the incorporation of the H1 allele.

In conclusion, this thesis identified, characterized, and/or introgressed two genes, *FT-A2* and *bZIPC1*, which impacts SNS, and provides valuable alleles for breeders to select upon in the quest for increasing GNS and overall yield.

FUTURE DIRECTIONS

FT-A2 A10 allele

Current evidence points to the *FT-A2* D10A SNP as the most likely causal polymorphism (Glenn et al., 2022). A conclusive test of this hypothesis would require the editing of position 124,172,909 from C to A (alanine to cytosine). Unfortunately, this is a transversion, and current plant gene editors are not efficient in editing transversions. New prime editing technologies (Anzalone et al., 2019) may help solve this problem once they become more efficient in plants (Lin et al., 2020). Until then, a potentially informative experiment could include the generation of transgenic plants expressing different combinations of the D10 and A10 coding regions, each combined with both natural promoters. Characterization of SNS in multiple independent transgenic plants generated for each of the four possible combinations will inform us if it is the D10A polymorphism, the promoter region, or both the allele and promoter which are causing the observed differences.

Regardless of the causal polymorphisms, the isogenic lines developed in this study allowed us to test the combined effect of the promoter and coding region SNPs on grain yield components in durum wheat. While the *FT-A2* A10 allele showed a positive effect on SNS and GNS, it did not have a significant effect on total grain yield in our four varieties. For future studies, it would be interesting to test if the introgression of the A10 allele in durum varieties that are not source limited can result in increases in grain yield. In addition, we observed a higher frequency of the A10 allele among winter wheats (Glenn et al., 2022), suggesting winter varieties may benefit more than spring varieties from the A10 introgression. To test this hypothesis, it would be interesting to test the effect of the A10 allele in winter durum wheat varieties.

It is also possible that the limited effect of A10 on grain yield in durum wheat may be determined by its epistatic interactions with other gene combinations fixed in durum wheat. With this in mind, it will be interesting to combine the *FT-A2* A10 allele with other known genes which may counteract its negative correlation with thousand kernel weight (TKW).

Some studies have suggested that wheat grain yield is more affected by variation in GNS than by TKW (Feng et al., 2018; Lynch et al., 2017). If this is the case, one route to maximize the impact of SNS on GNS is by increasing floret fertility. Previous studies have demonstrated that the locus *Grain Number Increase 1* (*GNI1*) is an important contributor to floret fertility with mutants carrying the impaired *GN1-A1* 105Y allele having one more grain per spikelet on average and higher grain yields than the WT lines in both durum and common wheat (Sakuma et al., 2019). This is an encouraging result, since it suggests that in some backgrounds the increases in gran number can be translated into increases in grain yield. A similar positive effect of an increase in SNS and GNS on grain yield was reported for the WAPO-A1 allele (Kuzay et al. 2019). Thus, combining the *FT-A2* A10 allele with the *GNI1* 105Y allele may further increase GNS and increase grain yield in genotypes that are not source-limited.

Another alternative is to mitigate the negative impact of the *FT-A2* A10 allele on TKW by combining it with loci with known positive effects on TKW. Both *GRAIN WEIGHT2 (GW2)* and *DA1* (DA means "large" in Chinese) are negative regulators of grain size. It was found that plants homozygous for *gw2-a1* mutant alleles had increased grain width, weight, and length, (Simmonds et al., 2016; Yang et al., 2012) while promoter differences in *DA1* reduced expression levels in haplotype, *TaDA1-A-Hapl,* and increased the size and weight of wheat grains (Liu et al., 2020). *Hapl* has become the dominant haplotype in modern cultivars, suggesting that it underwent positive selection in breeding programs for higher yield (Liu et al.,

2020). Interestingly, *da1* mutants increased grain size but reduced SNS and GNS, leaving grain yield unchanged (Mora-Ramirez et al., 2021). In addition, DA1 and GW2 physically interact and have an additive effect on kernel weight (Liu et al., 2020), thus if we are able to combine the *FT-A2* A10 allele with *gw-A2* and *TaDA1-A-Hapl*, we could potentially increase GNS, maintain TKW, and increase yield in genotypes that are not source-limited.

Lastly, if plants are source-limited it may prove beneficial to combine the *FT-A2* A10 allele with other genes that increase biomass. One such route is through changes in the genes affecting plant height. Genotypic differences in plant height are genetically correlated with variation in grain number and grain yield (Rebetzke & Richards, 2000) due to an increased harvest index with no overall change in above-ground biomass. It is generally believed this is due to reduced competition between developing stems and spikes for limited carbon resources (Rebetzke & Richards, 2000; Tang et al., 2021). Currently, the Gibberellin (GA)-insensitive dwarfing genes, *Rht-B1b (Rht1)* and *Rht-D1b* (*Rht2)* are utilized throughout global breeding programs to reduce plant height and increase grain yield (Evans, 1998), including Kronos which has the semi-dwarf allele, *Rht1*. However, they are also associated with GA-insensitivity reduced vigor and coleoptile length that often results in poor establishment and slow early growth (Richards, 1992). Studies have shown though, that GA-sensitive height-reducing genes, such as *Rht8, Rht9, Rht11, Rht12, Rht13,* and *Rht14*, have the potential to develop shorter wheats with both improved seedling emergence and vigor (Rebetzke & Richards, 2000). Given the genetic background influences on the magnitude of the effect of *FT-A2* it would be interesting to see how it behaves in semi-dwarf backgrounds determined by GA-sensitive *Rht* genes.

Another route to increase biomass is by manipulation of genes regulating plant growth. For example, the GROWTH-REGULATING FACTOR 4 (GRF4) transcription factor has been

shown to promote and integrate nitrogen assimilation, carbon fixation, and growth (S. Li et al., 2018). It is balanced, however, by an antagonist regulatory relationship with the growth-repressing DELLA proteins (DELLAs) (Li et al., 2018). DELLAs are generally destroyed via the gibberellin (GA) pathway (Harberd et al., 2009; Xu et al., 2014). However, in semi-dwarf *Rht1* and *Rht2* varieties, the DELLAs are resistant to GA-stimulated degradation and accumulate, conferring the plants semi-dwarfism stature (Peng et al., 1999; Zhang et al., 2014).

Unfortunately, this accumulation of DELLAs reduces GRF4 activity and enhances insensitivity to nitrogen and reduced nitrogen-use-efficiency (Gooding et al., 2012). In an effort to counter this, recent studies have identified two SNPs in $GRF4^{ngr2}$ which prevents miRNA-cleavage of *GRF4* mRNA and increases the abundance of GRF4 (Li et al., 2018). This tips the GRF4-DELLA balance to favor GRF4 leading to increased carbon and nitrogen assimilation, biomass, and leaf and stem width without losing the benefits from semi-dwarfism (Li et al., 2018; Serrano-Mislata et al., 2017). The combination of the $GRF4^{ngr2}$ SNPs and the *FT-A2* allele may thus prove beneficial in improving above-ground biomass and harvest index.

Finally, it would be interesting to test the epistatic interactions between *FT-A2* and other genes affecting SNS in wheat, to determine their combined effects. Several genes affecting SNS have been identified recently, including *FT-A2* (Glenn et al., 2022), *WAPO-A1* (Kuzay et al., 2022), *COL5* (Zhang et al., 2022), and *ELF3* (Alvarez et al., 2016). In the winter population segregating for five genes to influencing SNS, *FT-A2* showed no significant interactions with *WAPO-A1*, *bZIPC-B1*, *RHT-D1* and *PPD-D1* suggesting additive effects. However, *FT-A2* is likely to interact with other genes because its effects varied in different backgrounds. For example, Desert Gold, which had the highest average SNS among the tested durum varieties, had the smallest increase associated with the *FT-A2* allele. Thus, studies combining the different

alleles will be necessary to determine the optimum allele combinations to maximize SNS. One potential use of these lines with increased SNS and GNS, would be as research stocks to test the effects of genes impacting biomass on total grain yield.

Overall, it would be interesting to characterize the genetic interaction of *FT-A2* with other yield-related genes in an effort to translate the positive effects of the A10 allele on SNS and GNS into real increases in total grain yield in both durum and common wheat.

bZIPC1

The discovery of *bZIPC1* has revealed exciting new information about the bZIP C-group including its impact on SNS through a different pathway. However, it has also raised new questions that will require further experiments.

Studies in wheat, *A. thaliana*, and rice have reported that the FT protein is a key player in the mobile flowering signal (florigen) (Corbesier et al., 2007; Li et al., 2015; Tamaki et al., 2007). Two FT monomers and two DNA-binding bZIP transcription factors (FD and FD PARALOG) interact with a dimeric 14-3-3 protein bridge to form the 'florigen activation complex' (FAC) (Taoka et al., 2011). However, in wheat, FT2 was the only FT- or FD-like protein to not interact with any of the six tested 14-3-3 proteins from the FAC, regardless of the D10A allele (Glenn et al., 2022; Li et al., 2015). Since then, we have identified bZIPC1 as the first protein interactor with FT2. But given that bZIPC1 is a bZIP transcription factor like FD, it will be informative to determine if the wheat bZIPC1 protein can interact with any of the known 14-3-3 proteins, and if the combined bZIPC1-14-3-3 can interact with FT2.

It has also been demonstrated that bZIP C-group members in *A. thaliana* can form heterodimers with other members of the C-group as well as with members of the S1-group

(Ehlert et al., 2006). Thus, future studies should test the interaction of bZIPC1 with the other wheat C-group proteins in addition to wheat S1-group proteins.

Another area that requires further studies is the co-localization of FT2 and bZIPC1 in the wheat developing spike. Although we have shown that the genes encoding these two proteins are expressed in the developing spike at the same developmental stages, we currently do not know which region of the developing spike they are expressed in. We have begun *in-situ* hybridization work to determine the exact place and time of expression of these two genes within the developing spike. For example, it would be interesting to know if these two genes are expressed in the inflorescence meristem (IM) close to the time of its transition to the terminal spikelet.

Given that *ft2*-null mutants increased SNS (Shaw et al., 2019), we were surprised that the *bzipc1*-null mutants drastically decreased SNS. Thus, to better tease apart the *FT2-bZIPC1* relationship and mechanisms, we have initiated crosses between the double *bzipc1* mutants and the *ft2* mutant. It will be informative to see the epistatic interactions of these genes on SNS.

In addition, to increase our understanding of bZIPC1 potential downstream targets, we have initiated a Quant-Seq experiment utilizing the *bzipc1* mutants to identify the transcriptomic changes associated with the mutant phenotype and the genes and pathways regulated by *bZIPC1*.

Lastly, while *bZIPC1-B1* is a strong candidate gene for the H1 haplotype's effect on SNS, it is also possible that it is only tightly linked to a different causal gene. To narrow the candidate region and potential genes, we have initiated a fine mapping project. Regardless of the causal gene, the complete H1 candidate region can be integrated into elite varieties using flanking markers to start evaluating the effect of the H1 haplotype on different grain yield-components in different genetic backgrounds and environments.

In conclusion, the discovery of the *bZIPC1* gene and its associated H1 haplotype with positive effect on SNS, GNS and spike yield opens new basic and practical areas of research on wheat spike development and grain yield.

FUTURE DIRECTION REFERENCES

Alvarez, M. A., Tranquilli, G., Lewis, S., Kippes, N., & Dubcovsky, J. (2016). Genetic and physical mapping of the earliness per se locus *Eps-A*^{m}*1* in *Triticum monococcum* identifies *EARLY FLOWERING 3* (*ELF3*) as a candidate gene. *Functional and Integrative Genomics*, *16*, 365–382.

Anzalone, A. v., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., & Liu, D. R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, *576*, 149–157.

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., & Coupland, G. (2007). FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science*, *316*, 1030–1033.

Ehlert, A., Weltmeier, F., Wang, X., Mayer, C. S., Smeekens, S., Vicente-Carbajosa, J., & Dröge-Laser, W. (2006). Two-hybrid protein-protein interaction analysis in Arabidopsis protoplasts: Establishment of a heterodimerization map of group C and group S bZIP transcription factors. *The Plant Journal*, *46*, 890–900.

Evans, L. T. (1998). *Feeding the ten billion: plants and population growth*. Cambridge University Press.

Feng, F., Han, Y., Wang, S., Yin, S., Peng, Z., Zhou, M., Gao, W., Wen, X., Qin, X., & Siddique, K. H. M. (2018). The effect of grain position on genetic improvement of grain number and thousand grain weight in winter wheat in North China. *Frontiers in Plant Science*, *9*, 129.

Glenn, P., Zhang, J., Brown-Guedira, G., Dewitt, N., Cook, J. P., Li, K., Akhunov, E., & Dubcovsky, J. (2022). Identification and characterization of a natural polymorphism in *FT-A2* associated with increased number of grains per spike in wheat. *Theoretical and Applied Genetics*, *135*, 679–692.

Gooding, M. J., Addisu, M., Uppal, R. K., Snape, J. W., & Jones, H. E. (2012). Effect of wheat dwarfing genes on nitrogen-use efficiency. *Journal of Agricultural Science*, *150*, 3–22.

Harberd, N. P., Belfield, E., & Yasumura, Y. (2009). The angiosperm gibberellin-GID1-DELLA growth regulatory mechanism: how an “inhibitor of an inhibitor” enables flexible response to fluctuating environments. *Plant Cell*, *21*, 1328–1339.

Kuzay, S., Lin, H., Li, C., Chen, S., Woods, D. P., Zhang, J., Lan, T., von Korff, M., & Dubcovsky, J. (2022). *WAPO-A1* is the causal gene of the 7AL QTL for spikelet number per spike in wheat. *PLOS Genetics*, *18*, e1009747.

Li, C., Lin, H., & Dubcovsky, J. (2015). Factorial combinations of protein interactions generate a multiplicity of florigen activation complexes in wheat and barley. *The Plant Journal*, *84*, 70.

Li, S., Tian, Y., Wu, K., Ye, Y., Yu, J., Zhang, J., Liu, Q., Hu, M., Li, H., Tong, Y., Harberd, N. P., & Fu, X. (2018). Modulating plant growth–metabolism coordination for sustainable agriculture. *Nature, 560*, 595–600.

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A. v., Raguram, A., Doman, J. L., Liu, D. R., & Gao, C. (2020). Prime genome editing in rice and wheat. *Nature Biotechnology*, *38*, 582–585.

Liu, H., Li, H., Hao, C., Wang, K., Wang, Y., Qin, L., An, D., Li, T., & Zhang, X. (2020). *TaDA1*, a conserved negative regulator of kernel size, has an additive effect with *TaGW2* in common wheat (*Triticum aestivum* L.). *Plant Biotechnology Journal*, *18*, 1330–1342.

Lynch, J. P., Doyle, D., McAuley, S., McHardy, F., Danneels, Q., Black, L. C., White, E. M., & Spink, J. (2017). The impact of variation in grain number and individual grain weight on winter wheat yield in the high yield potential environment of Ireland. *European Journal of Agronomy*, *87*, 40–49.

Mora-Ramirez, I., Weichert, H., von Wirén, N., Frohberg, C., de Bodt, S., Schmidt, R. C., & Weber, H. (2021). The *da1* mutation in wheat increases grain size under ambient and elevated CO2 but not grain yield due to trade-off between grain size and grain number. *Plant-Environment Interactions*, *2*, 61–73.

Peng, J., Richards, D. E., Hartley, N. M., Murphy, G. P., Devos, K. M., Flintham, J. E., Beales, J., Fish, L. J., Worland, A. J., Pelica, F., Sudhakar, D., Christou, P., Snape, J. W., Gale, M. D., & Harberd, N. P. (1999). 'Green revolution' genes encode mutant gibberellin response modulators. *Nature*, *400*, 256–261.

Rebetzke, G. J., & Richards, R. A. (2000). Gibberellic acid-sensitive dwarfing genes reduce plant height to increase kernel number and grain yield of wheat. *Australian Journal of Agricultural Research*, *51*, 235–246.

Richards, R. A. (1992). The effect of dwarfing genes in spring wheat in dry environments. II. Growth, water use and water-use efficiency. *Australian Journal of Agricultural Research*, *43*, 529–539.

Sakuma, S., Golan, G., Guo, Z., Ogawa, T., Tagiri, A., Sugimoto, K., Bernhardt, N., Brassac, J., Mascher, M., Hensel, G., Ohnishi, S., Jinno, H., Yamashita, Y., Ayalon, I., Peleg, Z., Schnurbusch, T., & Komatsuda, T. (2019). Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proceedings of the National Academy of Sciences*, *116*, 5182–5187.

Serrano-Mislata, A., Bencivenga, S., Bush, M., Schiessl, K., Boden, S., & Sablowski, R. (2017). *DELLA* genes restrict inflorescence meristem function independently of plant height. *Nature Plants*, *3*, 749–754.

Shaw, L. M., Lyu, B., Turner, R., Li, C., Chen, F., Han, X., Fu, D., & Dubcovsky, J. (2019). *FLOWERING LOCUS T2* regulates spike development and fertility in temperate cereals. *Journal of Experimental Botany*, *70*, 193–204.

Simmonds, J., Scott, P., Brinton, J., Mestre, T. C., Bush, M., del Blanco, A., Dubcovsky, J., & Uauy, C. (2016). A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theoretical and Applied Genetics*, *129*, 1099–1112.

Tamaki, S., Matsuo, S., Hann, L. W., Yokoi, S., & Shimamoto, K. (2007). Hd3a protein is a mobile flowering signal in rice. *Science*, *316*, 1033–1036.

Tang, T., Botwright Acuña, T., Spielmeyer, W., & Richards, R. A. (2021). Effect of gibberellin-sensitive *Rht18* and gibberellin-insensitive *Rht-D1b* dwarfing genes on vegetative and reproductive growth in bread wheat. *Journal of Experimental Botany*, *72*, 445–458.

Taoka, K. I., Ohki, I., Tsuji, H., Furuita, K., Hayashi, K., Yanase, T., Yamaguchi, M., Nakashima, C., Purwestri, Y. A., Tamaki, S., Ogaki, Y., Shimada, C., Nakagawa, A., Kojima, C., & Shimamoto, K. (2011). 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature*, *476*, 332–335.

Xu, H., Liu, Q., Yao, T., & Fu, X. (2014). Shedding light on integrative GA signaling. *Current Opinion Plant Biology*, *21*, 89–95.

Yang, Z., Bai, Z., Li, X., Wang, P., Wu, Q., Yang, L., Li, L., & Li, X. (2012). SNP identification and allelic-specific PCR markers development for *TaGW2*, a gene linked to wheat kernel weight. *Theoretical and Applied Genetics*, *125*, 1057–1068.

Zhang, C., Gao, L., Sun, J., Jia, J., & Ren, Z. (2014). Haplotype variation of Green Revolution gene *Rht-D1* during wheat domestication and improvement. *Journal of Integrative Plant Biology*, *56*, 774–780.

Zhang, X., Jia, H., Li, T., Wu, J., Nagarajan, R., Lei, L., Powers, C., Kan, C. C., Hua, W., Liu, Z., Chen, C., Carver, B. F., & Yan, L. (2022). *TaCol-B5* modifies spike architecture and enhances grain yield in wheat. *Science*, *376*, 180–183.