

Genetic Characterization of a Wheat Association Mapping Panel Relevant to Brazilian Breeding Using a High-Density Single Nucleotide Polymorphism Array

Greg Mellers,^{*4} Jorge González Aguilera,^{†1} Nick Bird,^{‡2} Ana Lidia Variani Bonato,[†] Sandro Bonow,^{†,3} Eduardo Caierão,[†] Luciano Consoli,[†] Flávio Martins Santana,[†] James Simmonds,[‡] Andrew Steed,[‡] Gisele Abigail Montan Torres,[†] Cristobal Uauy,[‡] Tally I. C. Wright,^{*} Pedro Luiz Scheeren,[†] Paul Nicholson,[‡] and James Cockram^{*,1,5}

^{*}The John Bingham Laboratory, NIAB, 93 Lawrence Weaver Road, Cambridge, CB3 0LE, United Kingdom, [†]EMBRAPA-Trigo, Caixa Postal 3081, 99050970, Passo Fundo, RS, Brazil, and [‡]John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, United Kingdom

ORCID IDs: 0000-0002-4979-4185 (F.M.S.); 0000-0002-9814-1770 (C.U.); 0000-0002-1014-6463 (J.C.)

ABSTRACT Bread wheat (*Triticum aestivum* L.) is one of the world's most important crops. Maintaining wheat yield gains across all of its major production areas is a key target toward underpinning global food security. Brazil is a major wheat producer in South America, generating grain yields of around 6.8 million tons per year. Here, we establish and genotype a wheat association mapping resource relevant to contemporary Brazilian wheat breeding programs. The panel of 558 wheat accessions was genotyped using an Illumina iSelect 90,000 single nucleotide polymorphism array. Following quality control, the final data matrix consisted of 470 accessions and 22,475 polymorphic genetic markers (minor allele frequency $\geq 5\%$, missing data $< 5\%$). Principal component analysis identified distinct differences between materials bred predominantly for the northern Cerrado region, compared to those bred for southern Brazilian agricultural areas. We augmented the genotypic data with 26 functional Kompetitive Allele-Specific PCR (KASP) markers to identify the allelic combinations at genes with previously known effects on agronomically important traits in the panel. This highlighted breeding targets for immediate consideration – notably, increased *Fusarium* head blight resistance via the *Fhb1* locus. To demonstrate the panel's likely future utility, genome-wide association scans for several phenotypic traits were undertaken. Significant (Bonferroni corrected $P < 0.05$) marker-trait associations were detected for *Fusarium* kernel damage (a proxy for type 2 *Fusarium* resistance), identifying previously known quantitative trait loci in the panel. This association mapping panel represents an important resource for Brazilian wheat breeding, allowing future genetic studies to analyze multiple agronomic traits within a single genetically diverse population.

KEYWORDS

Single nucleotide polymorphism (SNP) genotyping quantitative trait locus (QTL) genome-wide association scan (GWAS) genetic diversity polyploid crop breeding *Fusarium* head blight (FHB)

Brazil is the second largest producer of wheat in South America, with a grain production of 6.8 million tons in 2016 (<http://www.fao.org/faostat/en/#home>). The majority of wheat cultivation in Brazil takes place in temperate climate zones of the southern states of Rio Grande do Sul and Paraná. However, wheat is also cultivated in the Cerrado climatic zone of Minas Gerais, Goiás, Mato Grosso and Mato Grosso do Sul in central Brazil. The development of new wheat varieties for cultivation in Brazil is undertaken by public institutes as well as private companies. Various studies have shown that the genetic gains in grain production arising from these breeding activities over the last

40 years has been between 1.2% (irrigated wheat, Cerrado region; Cargnin *et al.* 2008) and 2.9% (temperate region, Rio Grande do Sul; Follmann *et al.* 2017) per year. Such genetic gains have been achieved against a background of numerous biotic and abiotic pressures. Of particular note are two diseases of the wheat ear that are widespread in Brazil's wheat producing regions: (A) *Fusarium* head blight (FHB), predominantly caused in southern Brazil by the fungus *Fusarium graminearum* Schw. (Teleomorph *Gibberella zeae* (Schw.) Petch.) (Astolfi *et al.* 2012), and (B) wheat blast, caused by the fungus *Magnaporthe oryzae* Triticum pathotype, which first emerged in

the Brazilian state of Paraná in 1985 (Igarashi *et al.* 1986) and has since spread to other countries in both South America (Cruz and Valent 2017) and Asia (Malaker *et al.* 2016; Islam *et al.* 2016). Similarly, a suite of three abiotic pressures are generally acknowledged to be critical for Brazilian wheat production (Scheeren *et al.* 2008): (i) drought - a major limiting factor for rain-fed wheat cropping areas in the Cerrado, a region previously highlighted as a new frontier for increasing Brazilian wheat production (Poersch-Bortolon *et al.* 2016), (ii) increasing temperature - forecasted to reduce global wheat production by 6% for every degree Celsius increase in mean global temperature (Asseng *et al.* 2015), and (iii) aluminum toxicity - a widespread problem on the acidic soils of Brazil, where toxic aluminum ions inhibit root growth (De Sousa 2006), resulting in knock-on effects on crop growth and productivity, including an exacerbation of the effects of drought and heat stresses. Collectively, these principal biotic and abiotic factors represent key breeding targets for the protection of future Brazilian wheat yields.

Despite rapid advances over the last 40 years, there is evidence that genetic gains in Brazilian wheat yield are beginning to stagnate (*e.g.*, Cargnin *et al.* 2008), as is the case in many other countries around the world (Mackay *et al.* 2011; Brisson *et al.* 2010). Molecular and genetic characterization of the germplasm underpinning these gains in Brazilian grain yield could help breeding programs maintain, or even increase, such gains in the future. Indeed, the advent of association mapping techniques in cereal species has meant that it is possible to exploit both historic and *de novo* phenotypic data available for collections of both varieties and breeders' lines to undertake quantitative trait locus (QTL) analyses via genome-wide association scans (GWAS) (Cockram and Mackay 2018), for example in wheat (Bentley *et al.* 2014) and barley (Cockram *et al.* 2010a). Such analyses rely on the availability of genome-wide genetic marker datasets, commonly consisting of thousands of single nucleotide polymorphism (SNP) markers, with various SNP arrays (*e.g.*, Cavanagh *et al.* 2013; Wang *et al.* 2014; Allen *et al.* 2017) and sequencing-based genotyping technologies (reviewed by Nguyen *et al.* 2019), recently having been applied in wheat. Previous GWAS studies have identified genetic loci controlling FHB (*e.g.*, Wang *et al.* 2017) and awning (*e.g.*, DeWitt *et al.* 2019) (for further examples, see also Discussion). As a complementary approach to the use of high-density genotyping approaches, the availability of genetic markers diagnostic or highly predictive of allelic state at genetic loci controlling major agronomic traits in wheat (*e.g.*, those listed on CerealsDB, <https://www.cerealsdb.uk.net/cerealgenomics/>), and subsequently published by Rasheed *et al.* 2016) allow known loci to be rapidly characterized in breeder-relevant germplasm, including association

mapping panels (*e.g.*, Li *et al.* 2016). Despite the potential for such marker-informed approaches, relatively limited genome-wide genotypic investigation of Brazilian wheat germplasm has been undertaken to date (Schuster *et al.* 2009). Published studies investigating genetic diversity in Brazilian wheat have predominantly used relatively small numbers of varieties and genetic markers - for example, a study of 36 varieties with 23 simple sequence repeat (SSR) markers (Schuster *et al.* 2009). A notable exception is a recent study of 211 Brazilian varieties genotyped with an Axiom 30,000 feature WhtBrd-1 Array, and analyzed for genetic diversity (Scherlosky *et al.* 2018). However, no genetic association mapping panel specifically focused on use in Brazilian wheat breeding has been characterized both phenotypically and genotypically to date, nor has such a panel been characterized using functional genetic markers.

The aim of this study was to (1) assemble a wheat association mapping panel of relevance to Brazilian wheat breeding, (2) genotypically characterize the panel using a high-density SNP array, (3) augment the genome-wide SNP datasets via the addition of genotypic data for functional SNP markers, (4) demonstrate the utility of the association mapping panel by conducting GWAS for exemplar phenotypic traits, and (5) make these resources publicly available. To this end, we assembled a panel of 470 wheat varieties and breeders' lines and genotyped it with a 90,000 feature SNP array. The panel was assessed for its genetic diversity, and the potential utility of the resource for quantitative trait locus (QTL) analysis demonstrated via preliminary genome wide association mapping of FHB resistance and awn presence/absence. To help further inform future exploitation of the panel for wheat breeding and genetic analyses a set of 26 functional SNP markers tagging known genetic loci were genotyped. Here, we make these germplasm, SNP and genetic resources publicly available, providing useful resources for wheat research and development in Brazil and beyond. The results are discussed in the context of current and future Brazilian wheat breeding targets.

MATERIALS AND METHODS

Wheat germplasm

In total, 558 bread wheat accessions were collated from the collections maintained by Brazilian Agricultural Research Corporation (EMBRAPA)-Trigo, Brazil. The panel represents wheat germplasm released between 1852 and 2013, and was selected to be broadly representative of the genetic diversity used within Brazilian wheat breeding programs over the last 75 years at EMBRAPA-Trigo (Passo Fundo, Brazil). After quality control filtering, the panel consists of 470 accessions with 364 varieties and 64 breeders' lines, as well as 42 synthetic lines (Supplementary Table 1), with 327 of the accessions of Brazilian origin.

Field trials and phenotyping

Field trials for assessment of FHB resistance in the 2011, 2012 and 2013 seasons were undertaken at Embrapa-Trigo, Passo Fundo, Brazil (28°13'46" S, 52°24'07" W). Each genotype was planted, without replication, in a single plot consisting of 3 rows of 3 meters length, with 20 cm between rows. Following common protocols for assessment of FHB, the trial was divided into three sections according to the maturity group of each line: 'early', 'medium' and 'late' flowering, and entries randomized within each block. To aid FHB infection, plots were watered using mist irrigation for ten minutes six times per day (three times in the morning, and three times in the evening) in order to guarantee humidity during the whole night. Additionally, *Fusarium* infected wheat grain was spread between the rows of the

Copyright © 2020 Mellers *et al.*

doi: <https://doi.org/10.1534/g3.120.401234>

Manuscript received March 13, 2020; accepted for publication April 24, 2020; published Early Online April 29, 2020.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at figshare: <https://doi.org/10.25387/g3.12077124>.

¹Present address: Universidade Federal de Mato Grosso do Sul - Campus de Chapadão do Sul, Rodovia MS 306 km 105, Chapadão do Sul, MS, CEP 79560-000, Brazil

²Present address: KWS-UK Ltd., 56 Church Street, Thriplow, Hertfordshire, SG8 7RE, United Kingdom

³Present address: Embrapa Clima Temperado, Caixa Postal 403, 96010971, Pelotas, RS, Brazil

⁴Deceased

⁵Corresponding author: John Bingham Laboratory, NIAB, 93 Lawrence Weaver Road, Cambridge, CB3 0LE, United Kingdom. E-mail: james.cockram@niab.com

trial 60 days after sowing. Percentage of *Fusarium* damaged kernels (%FDK), a proxy of disease spread within the ear (type 2 resistance), was assessed by sampling 1,000 seeds after harvest and manually identifying the proportion of *Fusarium* infected seeds. The number of genotypes was variable between years (Supplementary Table 2, intersection of accessions across years illustrated in Supplementary Figure 1) hence a between year analysis was performed. This took the best linear unbiased predictions (BLUPs) for genotypes in a design that was unbalanced across years from a model fitted to the FDK data with genotype and year as random effects, implemented using the package lme4 (Bates *et al.* 2015) in R, and subsequently termed the ‘meta-analysis’. The model used to fit phenotypic data across years was as follows:

$$y_{ij} = \mu + G_i + T_j + e_{ij}$$

where y_{ij} is the %FDK of a plot, μ is the overall mean, G_i and T_j are random effects of genotype and year respectively, and e_{ij} is the residual error of y_{ij} . In the absence of weights for a generalized binomial mixed effect model, and due to the positive skew in the % FDK data, an initial transformation of the data were performed using box-cox transformation producing a lambda value of -1 for transformation. Final per genotype BLUPs were used as input for subsequent GWAS analyses. For the phenotype awn presence/absence, a binary assessment (1/0) of presence or absence was used as an input to subsequent analyses. Using Genstat (VSN International 2017a), generalized heritability (H^2) was estimated on a line mean basis following Cullis *et al.* (2006):

$$H^2 = 1 - \frac{V_{tt}}{(2\sigma_G^2)}$$

where σ_G^2 was taken as the genetic variance and V_{tt} was the average variance for difference between BLUPs (VSN International 2017b).

DNA extraction and genotyping

Genomic DNA was extracted from leaf material harvested from a single individual per accession using previously described protocols (Palotta *et al.* 2003). DNA quality and quantity were assessed using agarose gel electrophoresis and Qubit (Thermo Fisher), respectively. DNA was diluted to a concentration of 60-200 ng/ μ l using PCR grade water (Sigma Aldrich) and submitted under subcontract to Bristol University or the Institute of Biological, Environmental and Rural Sciences (IBERS), Aberystwyth University for genotyping. Genotyping was performed using a 90k feature Illumina iSelect SNP array developed by Wang *et al.* (2014). SNPs were called using GenomeStudio v2011.1 (Illumina) using a no call threshold of 0.1. Clusters were called in the Polyploid genotyping package with a minimum of 15 individuals per cluster and a 0.07 cluster distance with DBSCAN. SNP loci with greater than four clusters were removed. All remaining SNP loci that could not be automatically clustered were manually curated. Of the 67,774 SNPs successfully assayed on the array, 22,625 were polymorphic and passed quality control thresholds of minor allele frequency (MAF) $\geq 5\%$ and missing data $< 5\%$. Individual genotypes were then only kept with $< 5\%$ missing data. Genetic markers were subsequently filtered again ($< 5\%$ MAF), resulting in a total of 22,475 SNPs for downstream analyses (Supplementary Table 3). Inter-plate variation was assessed within GenomeStudio using replicate genotypes. Inter-machine variation between providers was visually assessed for fluorescence variance in GenomeStudio. Where required, missing data were imputed using softImpute v1.4 (Mazumder *et al.* 2010) in R v3.3.3. Markers were first centered

and scaled with softImpute before imputing missing data. Additionally, a subset of 26 previously published KASP markers for genes that underpin economically important wheat traits (open access on CerealsDB and also as published by Rasheed *et al.* 2016) were genotyped using established protocols (Trick *et al.* 2012).

Population structure analysis

All analyses were undertaken using the software, R v3.3.3 (R Core Team 2015). Of the 470 accessions, the synthetic hexaploid wheat accessions were removed due to their different origin, and the 22,625 genetic markers were re-filtered again on $< 5\%$ MAF using a custom script. Principal component analysis (PCA) was conducted using the base package, ‘stats’, function ‘prcomp’. For PCA, the marker data were thinned using a custom script: where two markers had a pairwise correlation coefficient (r) > 0.75 , only a single marker from the comparison was kept. PCA plots were generated using the R packages ggplot2 (Wickham 2016) and ggpubr (Kassambara 2019). Intra-chromosomal linkage disequilibrium (LD) between genetically mapped loci was calculated by the squared correlation coefficient (r^2) using the function ‘LD.Measures’ implemented by the package LDcorSV (Desrousseaux *et al.* 2017) in R. Marker data from the 427 accessions were first thinned at a threshold of $r < 1$ using a custom script to remove co-located markers, leaving 8,527 markers with genetic map positions from the Wang *et al.* (2014) linkage map. Results were visualized per chromosome by plotting r^2 values against genetic distance and using a locally weighted scatterplot smoothing (LOWESS) to fit a line to the data, implemented by the ‘lowess’ function in R. For the A and B genomes, a smoothing span of 5% of data points was used, while for the D genome a smoothing span of 20% of data points was used.

Genome wide association scans

GWAS was undertaken with the package, GAPIT (Lipka *et al.* 2012) using a Mixed Linear Model (MLM) algorithm as

$$y = \beta + Z\mu + e$$

where y is the vector of observed phenotypes, X is the incidence matrix of fixed effects β is a vector of coefficients of the fixed effects (here the mean and one SNP), Z is the incidence matrix mapping for individual genetic effects, μ is the vector of individual genetic effects with $\text{Var}(\mu) = K\sigma_g^2$ where K is the kinship matrix estimated from the markers e is an $n \times n$ matrix of residual effect such that $\text{Var}(e) = I\sigma_e^2$. These analyses excluded the synthetic wheat lines due to their highly differentiated genetic structure. Correction for kinship was undertaken using the default kinship matrix constructed in GAPIT, from a subset of 4,313 high quality SNPs parsed from the full quality-controlled data matrix as follows: missing data $< 5\%$, MAF $> 5\%$ and thinned by LD at $r < 0.75$ using a custom script. Skimming by r was undertaken to reduce the effect of uneven LD and SNP density on the calculation of kinship (Speed *et al.* 2012), an issue that is particularly acute in wheat. Missing data in the SNP set was imputed using iterative soft-threshold singular value decomposition, implemented in the R package, softImpute. Mapped SNPs published by Wang *et al.* (2014) were used for GWAS (6,215). Bonferroni-adjusted p values at a significance threshold of 0.05 or 0.01 were used for all genome wide association scans ($P < 8.05 \times 10^{-6}$ or 1.60×10^{-6} , denoted by blue and red threshold lines, respectively). The wheat reference genome assembly (IWGSC RefSeq v1.0, cv. Chinese Spring, IWGSC, 2018), was used to determine the physical positions of significant ($p_{adj} < 0.05$) markers of interest. SNP flanking DNA

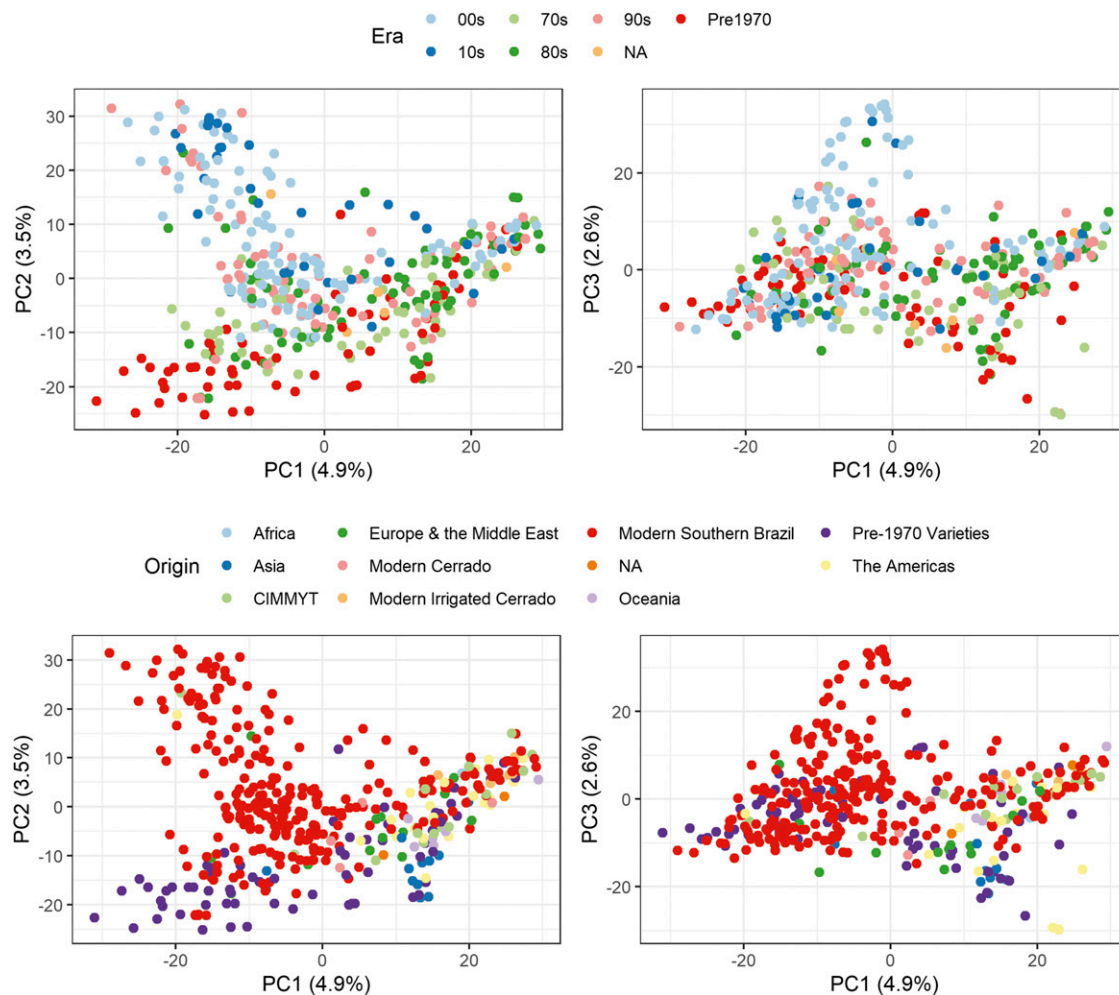


Figure 1 Principal component analysis of the genotyped association mapping panel (n accessions = 427) highlights major divisions in the material when considered by the era of variety release (top), and origin (bottom). Analysis of principle component 1 (PC1) vs. PC2 shows a division between the more modern germplasm released in the decades 2000, 2010 and those released before this time. Second, separate breeding targets for the northerly (modern Cerrado) and southerly wheat growing environments of Brazil result in segregation of material, with the more northerly material showing greater similarity to international material.

sequences were acquired from GrainGenes (Carollo *et al.* 2005. Available at <https://wheat.pw.usda.gov/GG3/>) and TriticeaeToolbox (Blake *et al.* 2016. Available at <https://triticeaetoolbox.org/>), and BLASTn (Altschul *et al.* 1990) was used to determine marker positions in the wheat physical map (e -value threshold $< 1e^{-20}$).

Data availability

SNP, KASP and phenotypic data are available in Supplementary Tables 1 through to 4, available at figshare. File Supplementary Table 1 contains details of the AM panel. File Supplementary Table 2 contains the phenotype data used for GWAS. File Supplementary Table 3 contains the 90k SNP genotypic data. File Supplementary Table 4 contains the KASP SNP data. Seed for the AM panel can be requested from Embrapa-Trigo: contact cnpt.chpd@embrapa.br. Supplemental material available at figshare: <https://doi.org/10.25387/g3.12077124>.

RESULTS

Genome-wide genetic markers

After quality control, of the 67,774 SNPs successfully assayed on the array, 22,475 polymorphic SNPs remained for downstream analyses.

All genetic markers were ordered using the wheat consensus genetic map (Wang *et al.* 2014), allowing 18,895 SNPs to be assigned a genetic map position (Supplementary Table 3). Analysis of SNP frequencies across the A, B and D sub-genomes of wheat found the B genome to possess the highest number of SNPs (9,716), while the D genome had the lowest (1,863). By chromosome, the number of SNPs ranged from 1789 (chromosome 5B) to 95 (4D). Markers were further thinned at a threshold of $r < 0.95$ to remove redundant SNPs in very high LD (as commonly applied in large genome temperate cereal crops, *e.g.*, Bentley *et al.* 2014), producing a thinned total of 8,721 SNPs for 470 individuals of which 6,215 had assigned genetic map positions (Wang *et al.* 2014).

Population structure and linkage disequilibrium

The geographic origin of the accessions of the association mapping panel was attributed to one of ten categories by the wheat breeders at EMBRAPA Trigo, Brazil including: Africa, the Americas, Asia, CIMMYT collections, Europe and the Middle East, southern Brazil, northern Brazil (*i.e.*, Cerrado region, irrigated and non-irrigated), Oceania, and pre-1970 (*i.e.*, before the Green Revolution). A subset of

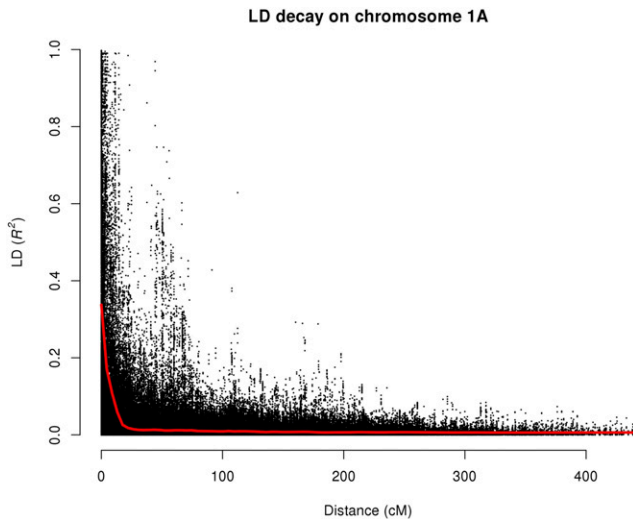


Figure 2 Intra-chromosomal linkage disequilibrium (LD) in the association mapping panel, as measured by r^2 between genetically mapped marker pairs. A LOWESS curve is fitted to the data, indicated in red. Wheat chromosome 1A is shown here, with LD plots for all other chromosomes provided in Supplementary Figure 2.

4,313 SNPs parsed from the full Illumina iSelect dataset (missing data <5%, MAF \geq 5% and thinned by LD at $r < 0.75$) was used to investigate the genetic substructure of the panel. Principal component (PC) analysis found the first three principal components to capture 11% of the total variance (Figure 1). Plotting PC1 against PC2 and overlaying information on era of variety release and geographic origin resolved some broad trends. The most striking was a central cluster of modern southern Brazilian lines released predominantly in the 2000s and 2010s largely separated from all other accessions via the PC2 axis. These modern materials were flanked on each side by a cluster predominantly represented by pre-1970s germplasm to one side, and a cluster which contained accessions from various origins on the other - including all of the accessions from the Brazilian Cerrado, and most of the international material, in particular all accessions from Asia, Oceania, the Americas and CIMMYT, as well as most of the remaining pre-1970s lines. The more modern material showed little clustering with year of release (Figure 1) but instead broadly grouped by region of cultivation in Brazil with southerly and northerly material predominantly clustering separately. Analysis of the kinship matrix generated using the same subset of 4,313 SNPs supports the broad groupings identified by principal component analysis (Supplementary Figure 2). Linkage disequilibrium for mapped markers was found to decay at a relatively slow rate, similar to that observed in previous wheat studies (e.g., Maccaferri *et al.* 2015) (Figure 2; Supplementary Figure 3).

Using diagnostic KASP markers to survey allelic variation for key agronomic traits

Twenty-six published KASP markers diagnostic for allelic state, or closely linked to previously identified loci (*i.e.*, predictive of allelic state) controlling a range of agronomically relevant traits were used to genotype 349 accessions of the AM panel (Figure 3. Supplementary Table 4). Using diagnostic markers for allelic state at the two major loci controlling plant height (*RHT-B1* and *RHT-D1*; KASP markers wMAS000001 and wMAS000002, respectively), 59% of accessions were classified as being 'semi-dwarf' in stature, while the remaining

lines that lacked semi-dwarfing alleles at either of the two loci were classified as 'tall'. Mutant *Rht-B1b* or *Rht-D1b* alleles were responsible for the semi-dwarf phenotype in 75% and 25% of the accessions carrying dwarfing alleles, respectively. Genotyping ten markers assessing eight disease resistance loci found resistance alleles to be present at relatively low frequency, with eye spot resistance alleles conferred by *Pch1* representing the most infrequent (1%; wMAS000023). Resistance alleles for *Fusarium* head blight conferred by the major resistance locus *Fhb1* were predicted in 3% of lines (wMAS000009, wMAS000009). Resistance alleles to stem rust conferred by *Stem rust 36* (*Sr36*) were predicted to be present at 3% (wMAS000015). The most frequent resistance alleles were identified for the soil borne mosaic virus resistance locus *Sbm1* (15%; wMAS000016) and the yellow rust resistance gene *Lr34* (25%; wMAS000003 and wMAS000004). Four loci controlling grain quality traits were assessed: (1) via the two-locus haplotypes derived from the assays targeting the *Pina-D1* and *Pinb-D1* loci (wMAS000018 and wMAS000019), 55% accessions were classified as having 'hard' grain (allele combination *Pina-D1b* + *Pinb-D1b*) and 45% as 'soft' grain (*Pina-D1a* + *Pinb-D1b* or *Pina-D1b* + *Pinb-D1a*). (2) Alleles contributing to the composition of high molecular weight glutenin subunits (HMW-GS) in the grain were predicted via assays for the *Glu-A1* (wMAS000012, wMAS000013) and *Glu-D1* (wMAS000014) loci. The two-SNP haplotypes generated allowed classification of *Glu-A1* alleles as 'null' (16%), '1Ax1' (24% of lines assayed) or '1Ax2*' (59%), ordered here in increasing positive effect on bread making quality. Genotyping of the SNP linked to *Glu-D1* predicted 51% of lines to be classified as '5+10', indicative of improved bread making quality, while the remaining 49% of lines were predicted to carry allele(s) for the '2+12 and others' gluten class associated with weaker gluten. (3) Via assessment of the KASP marker assaying a SNP linked to the *Gpc-B1* locus (wMAS000017), the allele conferring high grain protein content was predicted to be present in 1% of accessions. (4) Assessment of allelic state at the sucrose synthase *Sus2-2B* locus (wMAS000021) found 8% of lines to carry the T586C_hapH allele associated with higher thousand kernel weight. Finally, four major flowering time loci were assayed for a subset of known alleles using ten KASP markers: the photoperiod response loci *PPD-D1* and *PPD-A1*, and the vernalization response loci *VRN-A1* and *VRN-B1*. At *PPD-D1*, assessment for the presence of the 2,089 bp promoter deletion first characterized in the variety Ciano67 identified photoperiod insensitive alleles in 48% of lines (wMAS000024). The photosensitive insensitive allele conferred by the insertion of a Mariner transposable element within *PPD-D1* intron-1 was identified at low frequency (3%; wMAS000025) in the lines assayed. The *PPD-D1* 5 bp exon 7 deletion conferring photoperiod insensitivity was identified in 40% of lines (wMAS000026). Collectively, *PPD-D1* photoperiod insensitive alleles were present in 91% of the accessions screened. Allelic variation at the T ('winter' allele) to C ('spring' allele) SNP in exon-4 of the *API*-like gene underlying the *VRN-A1* locus found 19% of the accessions tested to carry 'spring' alleles (wMAS000034), while the assay targeting the 'spring' *Vrn-A1a* allele conferred by a deletion within the promoter was present in 23% of lines (wMAS000035). Based on these two assays, a total of 42% of lines were predicted to carry 'spring' alleles at *VRN-A1*. The SNP previously reported to be associated with 'spring' and 'winter' alleles at *VRN-A1* (Yan *et al.* 2004) and subsequently converted to KASP marker wMAS000033 found 17% of lines to carry 'spring' alleles. Finally, genotyping the *VRN-B1* locus found 'spring' *Vrn-B1b* alleles to be present in 8% of lines (wMAS000037).

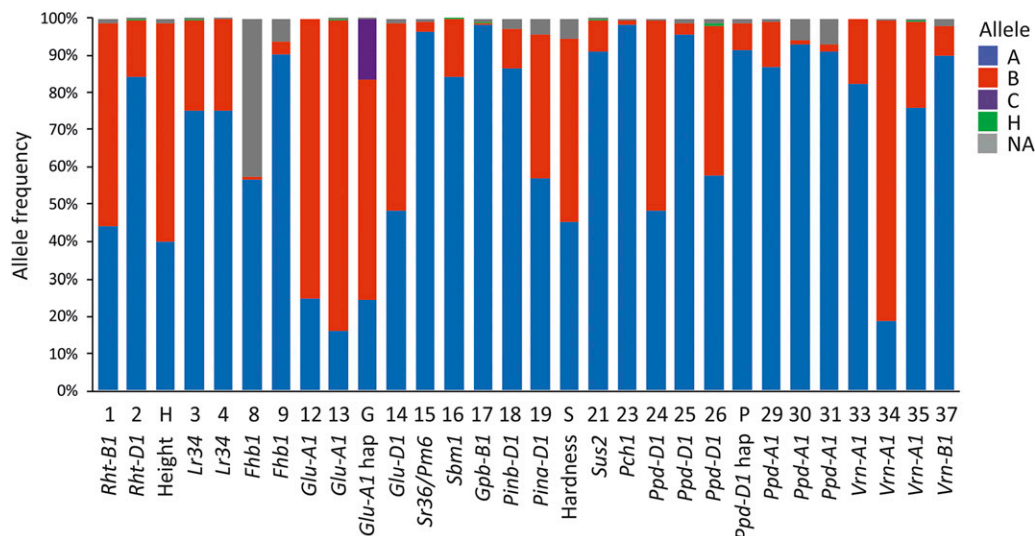


Figure 3 Summary of genotyping of 26 functional KASP markers in a subset of the association mapping panel ($n = 349$). KASP assay number is indicated on the x-axis, coded as 1 = wMAS000001, 2 = wMAS000002, etc. SNP allele calls are color coded: for disease resistance genetic loci, allele A (blue) represents the susceptible allele and allele B (red) the resistant allele. For all other SNPs, alleles A and B represent the wild type and mutant alleles, respectively. Allele NA = missing data/null allele call, allele H = heterozygous. Additionally, for some traits the genotypic results from a subset of KASP markers combine to predict overall phenotype, coded as: H = plant height (allele A = tall, B = semi-dwarf, surmised from KASP markers 1 and 2); G = *Glu-A1* allele (A = allele 1, B = allele 2*, C = allele null, surmised from KASP markers 12 and 13); S = seed hardness (A = soft, B = hard, surmised from KASP markers 18 and 19); P = *Ppd-D1* haplotype (A = photoperiod sensitive, B = photoperiod insensitive, surmised from KASP markers 24-26).

predict overall phenotype, coded as: H = plant height (allele A = tall, B = semi-dwarf, surmised from KASP markers 1 and 2); G = *Glu-A1* allele (A = allele 1, B = allele 2*, C = allele null, surmised from KASP markers 12 and 13); S = seed hardness (A = soft, B = hard, surmised from KASP markers 18 and 19); P = *Ppd-D1* haplotype (A = photoperiod sensitive, B = photoperiod insensitive, surmised from KASP markers 24-26).

Genome wide association scans

To validate this panel for future use toward genetic analysis of traits relevant to Brazilian wheat production, phenotypic data were collected for two traits for subsequent GWAS. A subset of the full panel ($n = 149$) was phenotyped for %FDK in 2011, 2012 and 2013, as well as for awn presence/absence ($n = 199$) (Supplementary Table 2). For %FDK a meta-analysis was performed between years for GWAS. The overall mean of the resulting BLUPs for %FDK was 5.46 with a standard deviation of 0.69 and the data followed a normal distribution (Supplementary Table 2). For the qualitative awn trait, the number of genotypes with and without awns was 175 and 24, respectively. Of the traits studied, both showed significant marker trait associations (Figure 4). For awns presence/absence ($n = 199$) two markers on chromosome 5A were found to be highly significant ($p_{\text{adj}} < 0.01$), BobWhite_c8266_227 at 703.91cM and RAC875_c8642_231 at 709.71cM. For the meta-analysis of %FDK, which had a generalized heritability of 0.68, four significant markers ($p_{\text{adj}} < 0.05$) were identified on chromosome 6B, between 218.86 and 226.64 cM (Table 1).

DISCUSSION

The analyses in this study centered on wheat germplasm used in Brazilian breeding programs since 1935. These genotypes represent the diversity of lines employed in breeding at Embrapa, and include varieties released for cultivation across all Brazilian wheat growing regions. In addition, lines were also included from global germplasm that had specific ongoing utility to Brazilian germplasm, or which may have relevance in the near future. Our aim was to provide a comprehensive survey of genetic variation of relevance to the Brazilian wheat gene pool, understand how this total variability is distributed in the more recently developed genotypes, and to demonstrate the utility of the resources generated for genetic analysis of agronomic traits.

Population structure

The broad trend toward a three-cluster separation shown by PCA was supported by the results of the kinship matrix (Figure 1; Supplementary

Figure 2). These clusters can be broadly grouped according to the age and geographic origin of the accessions, with modern Brazilian germplasm from southern regions and the northern Cerrado region predominantly grouping within largely separate PCA clusters - although some mixing is evident of southern germplasm within the PCA cloud containing the majority of the northern Brazil germplasm. The large environmental differences between the Cerrado and Southern Brazilian agricultural environments necessitate largely independent breeding priorities, and our analyses of population structure indicate this may have led to the genetic differentiation of varieties bred for cultivation in these two regions. Wheat production in the Cerrado is typified by low yield potential when irrigation is not used, and with relatively high yield potential when irrigated (>8 tons per ha). A lack of *Fusarium* disease prevalence means that a broad range of international material, particularly lines adapted to drought environments and CIMMYT material, is suitable for breeding for improved yield in this more northerly Brazilian environment. In contrast, the major *Fusarium* disease pressure experienced in the southern Brazilian wheat growing regions have thus far limited the use of some international material for breeding of varieties targeting the more temperate climates. The potential long term effects of breeding for these two distinct agricultural environments is likely reflected in the PCA (Figure 1), in which material bred for southerly cultivation sits broadly independent of most other material in this study; including both the germplasm from international sources as well as the breeding material of the Cerrado region.

Genetic control of *Fusarium* disease resistance

Meta-analysis of three years of field data on the percentage of *Fusarium* damaged kernels found the generalized heritability of the trait to be 0.68. The peak SNPs identified by GWAS were located on chromosome 6B. Previous work in wheat has identified several major QTL, including two well defined regions for *Fusarium* resistance (Buerstmayr *et al.* 2009). The first of these, *Fusarium head blight 1* (*Fhb1*) on chromosome 3B, controls resistance predominantly through limiting the spread of disease through the rachis from an infected floret (type 2 resistance) via the action of a pore-forming

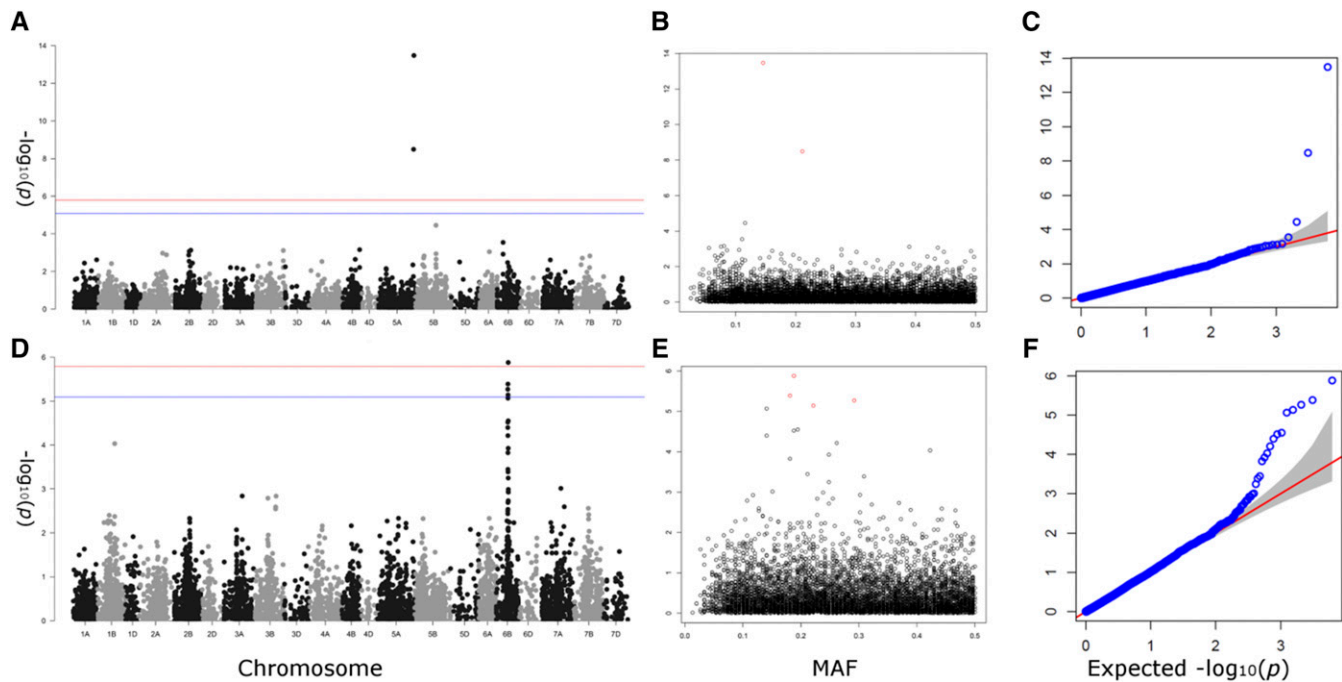


Figure 4 Genome wide association scans for the awn presence/absence (A, B, C) and % *Fusarium* damaged kernels (D, E, F). Manhattan plots of kinship corrected association analyses (A, D), minor allele frequency (MAF) against significance (B, E) and observed-expected significance (C, F). The potentially higher number of structurally driven significant loci ($P < 0.05$) in % *Fusarium* damaged kernels likely derive from the fewer lines ($n = 149$) in this analysis. However, the highly significant ($P < 0.01$) region is consistent with the previously mapped region.

toxin-like (PFT) gene (Rawat *et al.* 2016). The second, *Fhb2* on 6B, has been suggested to reinforce the cell wall and sequester deoxynivalenol (DON) toxins reducing pathogen spread within the ear (Dhokane *et al.* 2016). The peak markers identified in this Brazilian association mapping panel (covering a region of 279.36 Mb) (Table 1) span the chromosome 6B region previously identified via fine-mapping of *Fhb2* (Cuthbert *et al.* 2007) which spans a region of 102.3 Mb (anchored between 217.96 to 320.26 Mb on 6B by markers gwm644 and gwm133, respectively). This result agrees with previous data, where spread of the disease within the ear was found to be regulated by *Fhb2* (Buerstmayr *et al.* 2009), with resistance alleles at this locus reducing percentage of FDK. Indeed, a QTL conferring resistance to both *Fusarium* disease and FDK identified via marker gwm88 has previously been reported to be located to the *Fhb2* locus using a biparental mapping population constructed from the Brazilian varieties Frontana and Mini Manó, which show moderate and low resistance to *Fusarium*, respectively (Ánges *et al.* 2014). BLASTn analysis finds gwm88 to be located on chromosome 6B at 430.08 Mb, close to the location of the most significant SNP identified here by GWAS (BS00103275_51, 453.93 Mb). Hence, the association mapping panel genotyped here is sufficient with low numbers of phenotyped accessions ($n = 149$) to identify loci relevant to *Fusarium* disease resistance, and shows that existing Brazilian germplasm, such as the variety Trigo-BR24, potentially carry type 2 resistance alleles at *Fhb2*. GWAS for FHB resistance has been previously conducted using various germplasm sources, including accessions predominantly from the USA (Arrunda *et al.* 2016; Liu *et al.* 2019; Tessmann *et al.* 2019; Wang *et al.* 2017), Europe (Kollers *et al.* 2013; Miedaner *et al.* 2011) and Asia (Li *et al.* 2016; Wu *et al.* 2019), as well as panels derived from accessions around the world (Hu *et al.* 2020). While *Fhb2* has previously been reported using bi-parental populations, our work is the first time that *Fhb2* has been identified by GWAS as far as

we are aware. The information on which FHB resistance genes are present, or absent, in the association mapping panel used in our study will inform future Brazilian wheat breeding approaches. The ability to identify QTL by GWAS depends on various factors, including marker allele frequencies in the panel. Thus, while the GWAS undertaken for FHB and awning demonstrates the general utility of the panel for use in GWAS, successful QTL identification for other phenotypes will of course depend on a trait-by-trait and QTL-by-QTL basis. Indeed, we note that the inability to detect *Fhb1* via GWAS in our panel was likely due to the low frequency of resistance alleles at this locus in the association mapping panel, which was below the 5% MAF threshold for inclusion of SNPs in the association scans. As an infection of the wheat ear, FHB directly impacts grain yield (e.g., Schuman and D'Arcy 2006; Zhang *et al.* 2011a; Reis & Carmona 2013). Additionally, FHB affects grain quality due to the accumulation of mycotoxins, predominantly DON. As these mycotoxins are harmful to humans and animals, grain contamination above threshold levels has a direct impact on grain use and trade. Future work in this association mapping panel to extend the available phenotypic data should allow for the identification of additional *Fusarium* resistance loci in the future. Combined with using marker-assisted tracking of novel and known *Fusarium* resistance loci (such as *Fhb1*, for which we find resistance alleles to only very rarely be present in the panel, <3%), this has the potential to enhance breeding of future Brazilian wheat varieties with enhanced disease resistance and grain quality. A locus conferring type 1 FHB susceptibility has previously been reported closely linked to the semi-dwarf *Rht-D1b* allele on chromosome 4D (Srinivasachary *et al.* 2008). While semi-dwarfing alleles were found to be more commonly conferred by *Rht-B1b* alleles in the panel, the *Rht-D1b* semi-dwarfing allele was nevertheless observed at a frequency of 16%. GWAS did not detect a susceptibility locus linked to *RHT-D1*, and further investigation

■ **Table 1** Details of significant ($p_{adj} < 0.05$) single nucleotide polymorphisms (SNPs) identified by genome-wide association scans (GWAS) for awn presence/absence and resistance to *Fusarium* disease, as assessed via percentage of *Fusarium* damaged kernels (FDK). [†]Wang et al. (2014). [‡]IWGSC Refseq v1.0 (IWGSC, 2018). Chr = chromosome. PVE = percentage variation explained. N = number of lines. MAF = minor allele frequency

Trait	Marker	Chr [†]	Genetic pos (cM) [†]	Physical pos (Mb) [‡]	PVE	p-value	MAF
Awn	BobWhite_c8266_227	5A	703.91	698.51	8.5	3.23×10^{-9}	0.21
Awn	RAC875_c8642_231	5A	709.71	509.60	14.8	3.34×10^{-14}	0.15
FDK	Kukri_c25377_240	6B	218.86	174.57	11.4	5.37×10^{-6}	0.19
FDK	IAAV2161	6B	221.38	408.24	11.7	4.08×10^{-6}	0.18
FDK	Excalibur_c23462_677	6B	222.92	438.93	11.1	7.23×10^{-6}	0.22
FDK	BS00103275_51	6B	226.64	453.93	13.0	1.31×10^{-6}	0.18

would be needed to clarify the allelic status of the 4D FHB susceptibility locus in this panel.

Genetic control of awning

GWAS for awning identified a major genetic locus on the long arm of chromosome 5A at approximately 704 cM. In wheat, three dominant inhibitors of awning have previously been identified: *Hd* (Hooded, chromosome 4A), *B1* (*Tipped 1*, 5A) and *B2* (*Tipped 2*, 6B) (McIntosh et al. 1998; Kato et al. 1998, Yoshioka et al. 2017). While *Hd* and *B2* remain only coarsely mapped, the *B1* locus has previously been mapped to a 7.5 cM interval on the long arm of chromosome 5A, distal to the 4AL/5AL translocation breakpoint (Mackay et al. 2014), which was further fine mapped to a 2.2 cM interval (Yoshioka et al. 2017). More recently, a C2H2 zinc finger transcription factor has been identified as a candidate gene for *B1* using GWAS in an association mapping panel and by mapping in a biparental populations (DeWitt et al. 2019; Wang et al. 2019), located at 698.529 Mb in the reference wheat genome sequence assembly. The significant SNP for awning identified here by GWAS (BobWhite_c8266_227, 698.508 Mb, Table 1) is located ~21 kb from the C2H2 zinc finger candidate gene, and has previously been identified as the most predictive marker for awning in a UK Multiparent Advanced Generation Inter-Cross (MAGIC) population (Mackay et al. 2014).

Inspection of the reference genome assembly and gene model annotations for cv. Chinese Spring (IWGSC, 2018) shows that the fine-mapped *B1* region (690.31 - 698.19 Mb, anchored via markers WABM232824 and WMS291, respectively) (Yoshioka et al. 2017) also contains the previously identified genes underlying the cereal *VRN-2* vernalisation locus (*ZCCT1*, ~698.18 Mb on chromosome 5A) (Yan et al. 2004, Cockram et al. 2010b) which control flowering time and seasonal growth habit in response to low temperatures (termed ‘vernalization’). This indicates it could be of interest to explore whether the preference for awned varieties in warmer climates might also be due to linkage to specific seasonal growth habit alleles at *VRN-2*. By anchoring SNPs identified here as highly significant to the wheat reference genome, as well as those previously shown to have significant association with awn length on chromosome 5A (Yoshioka et al. 2017), we find the physical region identified by GWAS here (covering a region of 188.91 Mb. Table 1) spans the fine-mapped 5A region previously identified by Yoshioka et al. (2017), demonstrating the association mapping panel assembled in this work is sufficient with limited awn phenotype data ($n = 199$) to identify target genetic regions.

KASP markers for the remaining disease traits

Lr34 encodes an ABC-type transporter and confers resistance to a range of wheat fungal pathogens: *Puccinia triticina*, *P. striiformis* f. sp. *tritici*, *P. graminis* f. sp. *tritici* and *Blumeria graminis* (Krattinger et al. 2009).

While stripe rust is not a particularly common pathogen of wheat in Brazil, several novel adult plant resistance loci have recently been identified in Brazilian wheat variety Toropi, as well as the multi-rust resistance locus *Lr46/Yr29* (Rosa et al. 2019). We found resistant alleles at *Lr34* to be relatively common in the AM panel, present in around a quarter of the lines, with KASP genotyping confirming the absence of *Lr34* resistance alleles in the variety Toropi. Resistance to the related fungal wheat pathogen *Puccinia graminis* sp. *tritici*, the causal agent of stem rust, is conferred by multiple genes including *Sr2* and *Sr36*. Both of these resistance genes originated from bread wheat relatives: *Sr2* from *T. turgidum* var. *dicoccum* cv. Yaroslav and *Sr36* from *T. timopheevi*. While stem rust occurs in Brazil, major epidemics are currently rare, despite reports of increased cultivation of susceptible varieties in Southern America (Germán et al. 2006). The low incidence and severity of stem rust in Brazil is mirrored here by the predicted low occurrence of resistance alleles at the two loci assayed, particularly *Sr36*, present in just four Brazilian varieties. Soilborne wheat mosaic virus (SBWMV) is mediated via the soil borne fungus *Polymyxa graminis* Led. with losses of up to 50% reported in infected wheat fields in southern Brazil (Caetano 1982). Despite this, the SBWMV resistance locus *Sbm1* (Bass et al. 2006) was predicted here to be present in 15% of the lines surveyed. A second locus, *Sbm2*, has also been reported to confer increased resistance to SBWMV, although only when present in combination with resistant alleles at *Sbm1* (Bayles et al. 2007). Using a SNP linked to *Sbm1*, The Brazilian variety Embrapa 16 is known to carry SBCMV resistance, mediated by two genetic loci (Barbosa et al. 2001). However, genotyping with the *Sbm1* linked KASP marker predicted Embrapa 16 not to carry a *Sbm1* resistance allele. Therefore, either this variety carries previously uncharacterized SBWMV resistance loci, the *Sbm1* linked marker is not robust in wider germplasm pools, or variation exists at *Sbm1* in accessions of Embrapa 16 held at different genebanks. For eyespot, as this disease is not endemic to Brazil, resistance conferred by *Pch1* on the distal end of chromosome 7D is not a direct breeding target. *Pch1* resistance originated from the wheat relative *Aegilops ventricosa*. As this alien introgression is also associated with linkage drag for undesirable alleles (Pasquariello et al. 2017), it is useful to determine allelic status at *Pch1* in order to select against the underlying alien introgression within a Brazilian breeding context. Here, KASP genotyping predicts the *Pch1* resistance allele to be present in just four accessions, indicating the associated alien introgression does not currently present a significant source of potentially detrimental variation in the germplasm assessed here.

KASP markers for end-use quality traits

Wheat end-use quality is a key target for breeders and growers, with traits such as dough gluten content and grain hardness determining

grain sale price and the products that the grain can be used for. Gluten content is determined by a class of proteins known as the high molecular weight glutenin subunits (HMW-GS), the low molecular weight glutenin subunits (LMW-GS) and the gliadins. Alleles at several genetic loci, including *Glu-A1* and *Glu-D1*, are known to control variation for HMW-GS (e.g., Zhang *et al.* 2011b). Here, of the lines carrying 'null' low glutenin quality alleles at *Glu-A1*, 74% also carried the SNP indicative of low glutenin quality alleles at *Glu-D1*, indicating selection for end-use for biscuit making or animal feed. For lines carrying the high glutenin quality '1Ax2' *Glu-A1* allele, the approximately 50:50 split of high vs. low quality alleles at *Glu-D1* highlights the balance of glutenin subunits selected during breeding for the specific end-use quality desired. Grain endosperm hardness has a major impact on milling properties and is largely conferred by the puroindoline (*Pina* and *Pinb*) genes, with mutation in one or both genes resulting in hard grain. While differences in *Pin* gene expression has also been associated with grain hardness (Nirmal *et al.* 2016), here the two KASP markers employed assayed for wild type (*Pina-D1a*, *Pinb-D1a*) or mutant *Pina-D1b*/*Pinb-D1b* alleles, previously reported to be the most frequently observed mutant alleles in worldwide wheat collections (Lillemo *et al.* 2006). The predominance of grain hardness conferred by the *Pinb-D1b* allele in the AM panel (79% of all hard lines) mirrors previous observations in which 86% of hard wheat lines from the CIMMYT breeding program were found to carry the *Pinb-D1b* allele (Lillemo *et al.* 2006).

KASP markers for phenology traits

Homeologous genes controlling flowering time in response to vernalization (*VERNALIZATION-A1* [*VRN-A1*], *VRN-B1*, *VRN-D1*) and photoperiod (*PHOTOPERIOD-A1* [*PPD-A1*], *PPD-B1*, *PPD-D1*) have a major influence on wheat flowering time, and thus adaptation and grain yield. Recessive wild-type *vrn-A1*, *vrn-B1* and *vrn-D1* alleles confer the vernalization responsive phenotype. Mutations in the promoters, introns and exons of *VRN-1* are known to result in the loss or reduction of the vernalization requirement (reviewed by Bentley *et al.* 2013). Collectively, markers within the *VRN-1* genes diagnostic for the 'spring' alleles *Vrn-B1b* (Diaz *et al.* 2012), *Vrn-A1a* (Yan *et al.* 2004) and the '2147-type' exon-IV mutation (Diaz *et al.* 2012), as well as a linked SNP for a third spring *Vrn-A1* allele, predicted 57% of varieties to be spring, with no lines found to carry spring *Vrn-A1* and *Vrn-B1* alleles according to the polymorphisms assayed. However, it should be noted the KASP assay for allelic variation at the *VRN-D1* locus failed to provide reliable allele calls, and so was excluded. As mutations at *VRN-D1* are reported to have the strongest 'spring' alleles, it is likely that many of the remaining lines would be found to be genetically spring-type if genotyped for allelic state at *VRN-D1*. Considering all three assays for *PPD-D1*, almost all lines investigated (93%) were identified as carrying photoperiod insensitive alleles. This is as expected, as in Brazil wheat is typically grown as part of a cycle of three crop rotations per year, suiting rapidly maturing photoperiod insensitive lines. Of those lines found to lack insensitive alleles at *PPD-D1*, five possessed insensitive alleles at *PPD-A1*. These were all synthetic wheat lines originating from the wheat breeding institute CIMMYT, and carried the 1,117 bp promoter deletion first identified in line GS-105 by Beales *et al.* (2007). The predominance of this photoperiod insensitive allele in the synthetic lines screened indicates it was most likely brought into the bread wheat breeding pool via tetraploid wheat accessions used at CIMMYT during the creation of synthetic wheat lines.

Conclusions

Characterization of genetic variation can increase the efficiency of breeding programs, helping selection of lines that maximize genetic diversity, and the introduction of specific traits and characteristics beneficial to regional agricultural environments. The wheat association mapping panel assembled here is the largest and best characterized of its kind specifically aimed to be relevant to Brazilian wheat breeding. We provide detailed genotypic datasets and analyses using the panel and demonstrate the utility of the resources generated by identifying significant marker-trait association via GWAS. Collectively, the resources generated here will help underpin future breeding programs, and aid future genetic analysis of agronomic traits of importance to Brazilian agricultural environments.

ACKNOWLEDGMENTS

This work was funded by United Kingdom (UK) Biotechnology and Biological Sciences Research Council (BBSRC) grants BB/N00442/1, BB/J01992/1 and BB/P016855/1, by Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) projects 02.11.08.001.00.00 and 12.15.07.002.00.00, and by a Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant PNPd 560516/2010-0 fellowship to JA. We thank Keith Gardner at NIAB for statistical support and Ian Mackay for advice during manuscript revision. In caring memory of Greg Mellers, from all his colleagues in the UK and Brazil.

LITERATURE CITED

- Allen, A. M., M. O. Winfield, A. J. Burridge, R. C. Downie, H. R. Benbow *et al.*, 2017 Characterization of a wheat breeders' array suitable for high-throughput SNP genotyping of global accessions of hexaploid wheat (*Triticum aestivum*). *Plant Biotechnol. J.* 15: 390–401. <https://doi.org/10.1111/pbi.12635>
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, 1990 Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Ánges, S.-H., L.-K. Szabolcs, V. Mónica, P. László, P. János *et al.*, 2014 Different influence of QTL linked to Fusarium head blight, Fusarium-damaged kernel, deoxynivalenol contents and associated morphological traits in a Frontana-derived wheat population. *Euphytica* 200: 9–26. <https://doi.org/10.1007/s10681-014-1124-2>
- Arrunda, M. P., P. Brown, G. Brown-Guedira, A. M. Krill, C. Thurber *et al.*, 2016 Genome-wide association mapping of Fusarium head blight resistance in wheat using genotyping-by-sequencing. *Plant Genome* 9: 1–14.
- Asseng, S., F. Ewert, P. Martre, R. P. Rötter, D. B. Lobell *et al.*, 2015 Rising temperatures reduce global wheat production. *Nat. Clim. Chang.* 5: 143–147. <https://doi.org/10.1038/nclimate2470>
- Astolfi, P., M. M. Reynoso, M. L. Ramirez, S. N. Chulze, T. C. A. Alves *et al.*, 2012 Genetic population structure and trichothecene genotypes of *Fusarium graminearum* isolated from wheat in southern Brazil. *Plant Pathol.* 61: 289–295. <https://doi.org/10.1111/j.1365-3059.2011.02515.x>
- Barbosa, M. M., L. R. Goulart, A. M. Prestes, and F. C. Juliatti, 2001 Genetic control of resistance to soilborne mosaic virus in Brazilian cultivars of *Triticum aestivum* L. *Tell. Euphytica* 122: 417–422. <https://doi.org/10.1023/A:1012937116394>
- Bass, C., R. Hendley, M. J. Adams, K. W. Hammond-Kosack, and K. Kanyuka, 2006 The *Sbm1* locus conferring resistance to soil-borne cereal mosaic virus maps to a gene-rice region on 5DL in wheat. *Genome* 49: 1140–1148. <https://doi.org/10.1139/g06-064>
- Bates, D., M. Maechler, B. Bolker, and S. Walker, 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67: 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bayles, R., D. O'Sullivan, V. Lea, S. Freeman, G. Budge, *et al.*, 2007 Controlling soil-borne cereal mosaic virus in the UK by developing resistant wheat cultivars. HGCA project report No. 418.

- Beales, J., A. Turner, S. Griffiths, J. W. Snape, and D. A. Laurie, 2007 A pseudo-response regulator is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 115: 721–733. <https://doi.org/10.1007/s00122-007-0603-4>
- Bentley, A. R., M. Scutari, N. Gosman, S. Faure, F. Bedford *et al.*, 2014 Applying association mapping and genomic selection to the dissection of key traits in elite European wheat. *Theor. Appl. Genet.* 127: 2619–2633. <https://doi.org/10.1007/s00122-014-2403-y>
- Bentley, A. R., E. F. Jensen, I. J. Mackay, H. Hönick, M. Fladung *et al.*, 2013 Flowering time, pp. 1–66 in *Genomics and Breeding for Climate-Resilient Crops*, Vol. 2, edited by Cole, C. Springer-Verlag Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-37048-9_1
- Blake, V. C., C. Birkett, D. E. Matthews, D. L. Hane, P. Bradbury *et al.*, 2016 The Triticeae Toolbox: combining phenotype and genotype data to advance small-grains breeding. *Plant Genome* 9: 1–10. <https://doi.org/10.3835/plantgenome2014.12.0099>
- Brisson, N., P. Gate, D. Gouache, G. Charmet, F. X. Oury *et al.*, 2010 Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. *Field Crops Res.* 119: 201–212. <https://doi.org/10.1016/j.fcr.2010.07.012>
- Buerstmayr, H., T. Ban, and J. A. Anderson, 2009 QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: a review. *Plant Breed.* 128: 1–26. <https://doi.org/10.1111/j.1439-0523.2008.01550.x>
- Caetano, V. R., 1982 Mosaico do trigo transmitido pelosolo ‘wheat soilborne mosaic virus’ Tobamovirus, pp. 563–570 in *Trigo no Brazil*, Vol. 2, edited by Osório, E. A. Fundação Cargill, Campinas, SP.
- Cargnin, A., M. Alves de Souza, and V. Fronza, 2008 Progress in breeding of irrigated wheat for the Cerrado region of Brazil. *Crop Breed. Appl. Biotechnol.* 8: 39–46. <https://doi.org/10.12702/1984-7033.v08n01a06>
- Carollo, V., D. E. Matthews, G. R. Lazo, T. K. Blake, D. D. Hummel *et al.*, 2005 GrainGenes 2.0. An improved resource for the small-grains community. *Plant Physiol.* 139: 643–651. <https://doi.org/10.1104/pp.105.064485>
- Cavanagh, C. R., S. Chao, S. Wang, B. E. Huang, S. Stephen *et al.*, 2013 Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci. USA* 110: 8057–8062. <https://doi.org/10.1073/pnas.1217133110>
- Cockram, J., and I. Mackay, 2018 Genetic mapping populations for conducting high-resolution trait mapping in plants, pp. 109–138 in *Plant Genetics and Molecular Biology. Advances in Biochemical Engineering/Biotechnology*, Vol. 164, edited by R. Varshney, M. Pandey and A. Chitikineni. Springer, Cham.
- Cockram, J., J. White, D. L. Zuluaga, D. Smith, J. Comadran *et al.*, 2010a Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proc. Natl. Acad. Sci. USA* 107: 21611–21616. <https://doi.org/10.1073/pnas.1010179107>
- Cockram, J., R. M. Howells, and D. M. O’Sullivan, 2010b Segmental chromosomal duplications harbouring group IV *CONSTANS*-like genes in cereals. *Genome* 53: 231–240. <https://doi.org/10.1139/G09-101>
- Cruz, C. D., and B. Valent, 2017 Wheat blast disease: danger on the move. *Trop. Plant Pathol.* 42: 210–222. <https://doi.org/10.1007/s40858-017-0159-z>
- Cullis, B. R., A. B. Smith, and N. E. Coombes, 2006 On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11: 381–393. <https://doi.org/10.1198/108571106X154443>
- Cuthbert, P. A., D. J. Somers, and A. Brulé-Babel, 2007 Mapping of *Fhb2* on chromosome 6BS: a gene controlling Fusarium head blight field resistance in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 114: 429–437. <https://doi.org/10.1007/s00122-006-0439-3>
- De Sousa, C. N. A., 2006 Classification of Brazilian wheat cultivars for aluminium toxicity in acid soils. *Plant Breed.* 117: 217–221. <https://doi.org/10.1111/j.1439-0523.1998.tb01929.x>
- Desrousseaux, D., F. Sandron, A. Siberchicot, C. Cierco-Ayrolles, B. Mangin, *et al.*, 2017 *Package ‘Ldcorsv’*.
- DeWitt, N. M., E. Guedira, M. Lauer, P. Sarinelli, P. Tyagi *et al.*, 2019 Sequence-based mapping identifies a candidate transcription repressor underlying awn suppression at the *B1* locus in wheat. *New Phytol.* 225: 326–339. <https://doi.org/10.1111/nph.16152>
- Dhokane, D., S. Karre, A. C. Kushalappa, and C. McCartney, 2016 Integrated metabolo-transcriptomics reveals Fusarium Head Blight candidate resistance genes in wheat QTL-*Fhb2*. *PLoS One* 11: e0155851. <https://doi.org/10.1371/journal.pone.0155851>
- Díaz, A., M. Zikhali, A. S. Turner, P. Isaac, and D. A. Laurie, 2012 Copy number variation affecting the *Photoperiod-b1* and *Vernalization-A1* genes is associated with altered flowering time in wheat (*Triticum aestivum*). *PLoS One* 7: e33234. <https://doi.org/10.1371/journal.pone.0033234>
- Follmann, D. N., A. Cargnelutti Filho, A. D. Lúcio, V. Q. de Souza, M. Caraffa *et al.*, 2017 Genetic progress in homogeneous regions of wheat cultivation in Rio Grande do Sul State, Brazil. *Genet. Mol. Res.* 16: 1.
- Germán, S., A. Barcellos, M. Chaves, M. Kohli, P. Campos *et al.*, 2006 The situation of common rusts the Southern Cone of America and perspectives for control. *Crop Pasture Sci.* 58: 620–630. <https://doi.org/10.1071/AR06149>
- Hu, W., D. Gao, H. Wu, J. Liu, C. Zhang *et al.*, 2020 Genome-wide association mapping revealed syntenic loci *QFhb-4AL* and *QFhb-5DL* for *Fusarium* head blight resistance in common wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 20: 29.
- Igarashi, S., C. M. Utimada, L. C. Igarashi, A. H. Kazuma, and R. S. Lopes, 1986 *Pyricularia* em trigo. I. Ocorrência de *Pyricularia* sp. no estado do Paraná. *Fitopatol. Bras.* 11: 351–352.
- International, V. S. N., 2017a *Genstat for Windows*, Ed. 19th. VSN International, Hemel Hempstead, UK.
- International, V. S. N., 2017b *Genstat Reference Manual (Release 19), Part 3 Procedures*, VSN International, Hemel Hempstead, UK.
- International Wheat Genome Sequencing Consortium (IWGSC), 2018 Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361: 4603.
- Islam, M. T., D. Croll, P. Gladioux, D. M. Soanes, A. Persoons *et al.*, 2016 Emergence of wheat blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*. *BMC Biol.* 14: 84. <https://doi.org/10.1186/s12915-016-0309-7>
- Kassambara, A., 2019 ggpubr: ‘ggplot2’ based publication ready plots. R package version 0.2.2. <https://CRAN.R-project.org/package=ggpubr>.
- Kato, K., H. Miura, M. Akiyama, M. Kuroshima, and S. Sawada, 1998 RFLP mapping of three major genes, *Vrn1*, *Q* and *B1*, on the long arm of chromosome 5A of wheat. *Euphytica* 101: 91–95. <https://doi.org/10.1023/A:1018372231063>
- Kollers, S., B. Rodemann, J. Ling, V. Korzun, E. Ebmeyer *et al.*, 2013 Whole genome association mapping of *Fusarium* head blight resistance in European winter wheat (*Triticum aestivum* L.). *PLoS One* 8: e57500.
- Krattinger, S. G., E. S. Lagudah, W. Spielmeier, R. P. Singh, J. Huerta-Espino *et al.*, 2009 A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323: 1360–1363. <https://doi.org/10.1126/science.1166453>
- Li, T., M. Luo, D. Zhang, D. Wu, L. Li *et al.*, 2016 Effective marker alleles associated with type 2 resistance to *Fusarium* head blight infection in fields. *Breed. Sci.* 66: 350–357. <https://doi.org/10.1270/jsbbs.15124>
- Lillemo, M., F. Chen, X. Xia, M. William, R. J. Peña *et al.*, 2006 Puroindoline grain hardness alleles in CIMMYT bread wheat germplasm. *J. Cereal Sci.* 44: 86–92. <https://doi.org/10.1016/j.jcs.2006.03.004>
- Lipka, A. E., F. Tian, Q. Wang, J. Peiffer, and L. Li, 2012 GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28: 2397–2399. <https://doi.org/10.1093/bioinformatics/bts444>
- Liu, Y., E. Salsman, J. D. Fiedler, J. B. Hegstad, A. Green *et al.*, 2019 Genetic mapping and prediction analysis of FHB resistance in a hard red spring wheat breeding population. *Front. Plant Sci.* 10: 1007. <https://doi.org/10.3389/fpls.2019.0100>
- Maccaferri, M., J. Zhang, P. Bulli, Z. Abate, S. Chao *et al.*, 2015 Genome-wide association study of resistance to stripe rust (*Puccinia striiformis* f. sp. *tritici*) in a worldwide collection of hexaploid spring wheat (*Triticum aestivum* L.) G3 (Bethesda). *Genes Genom. Genet.* 5: 449–465.
- Mackay, I., A. Horwell, J. Garner, J. White, J. McKee *et al.*, 2011 Reanalysis of the historical series of UK variety trials to quantify the contributions of

- genetic and environmental factors to trends and variability in yield over time. *Theor. Appl. Genet.* 122: 225–238. <https://doi.org/10.1007/s00122-010-1438-y>
- Mackay, I. J., P. Bansept-Basler, T. Barber, A. R. Bentley, J. Cockram *et al.*, 2014 An eight-parent multiparent advanced generation inter-cros population for winter-sown wheat: creation, properties, and validation. *G3 (Bethesda)*. 4: 1603–1610.
- Malaker, P. K., N. C. D. Barma, T. P. Tiwari, W. J. Collis, E. Duveiller *et al.*, 2016 First report of wheat blast caused by *Magnaporthe oryzae* pathotype triticum in Bangladesh. *Plant Dis.* 100: 2330. <https://doi.org/10.1094/PDIS-05-16-0666-PDN>
- Mazumder, R., T. Hastie, and R. Tibshirani, 2010 Spectral regularization algorithms for learning large incomplete matrices. *J. Mach. Learn. Res.* 11: 2287–2322.
- McIntosh, R. A., G. W. Hart, K. M. Devos, M. D. Gale, and W. J. Rogers, 1998 Catalogue of gene symbols for wheat. *Proceedings of the 9th International Wheat Genetics Symposium, Saskatoon* 5: 235.
- Miedaner, T., T. Würschum, H. P. Maurer, V. Korzun, E. Ebmeyer, *et al.*, 2011 Association mapping for Fusarium head blight resistance in European soft winter wheat. *Mol. Breed.* 28: 647–655.
- Nguyen, K. L., A. Grondin, B. Courtois, and P. Ganet, 2019 Next-generation sequencing accelerates crop gene discovery. *Trends Plant Sci.* 24: 263–274. <https://doi.org/10.1016/j.tplants.2018.11.008>
- Nirmal, R. C., A. Furtado, C. Wrigley, and R. J. Henry, 2016 Influence of gene expression on hardness in wheat. *PLoS One* 11: e0164746. <https://doi.org/10.1371/journal.pone.0164746>
- Palotta, M., P. Warner, R. Fox, H. Kuchel, S. Jefferies *et al.*, 2003 Marker assisted wheat breeding in the southern region of Australia. *Proc 10th Int Wheat Genet Symp, Paestum, Italy*, pp 1–6.
- Pasquariello, M., J. Ham, C. Burt, J. Jahier, S. Paillard *et al.*, 2017 The eyespot resistance genes *Pch1* and *Pch2* of wheat are not homoeologs. *Theor. Appl. Genet.* 130: 91–107. <https://doi.org/10.1007/s00122-016-2796-x>
- Poersch-Bortolon, L. B., J. F. Pereira, N. A. Junior, H. H. Gonzáles, G. A. Torres *et al.*, 2016 Gene expression analysis reveals important pathways for drought response in leaves and roots of a wheat cultivar adapted to rainfed cropping in the Cerrado biome. *Genet. Mol. Biol.* 39: 629–645. <https://doi.org/10.1590/1678-4685-gmb-2015-0327>
- Rasheed, A., W. Wen, F. Gao, H. Jin, J. Liu *et al.*, 2016 Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theor. Appl. Genet.* 129: 1843–1860. <https://doi.org/10.1007/s00122-016-2743-x>
- Rawat, N., M. O. Pumphrey, S. Liu, X. Zhang, V. K. Tiwari *et al.*, 2016 Wheat *Fhb1* encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to Fusarium head blight. *Nat. Genet.* 48: 1576–1580. <https://doi.org/10.1038/ng.3706>
- R Core Team, 2015 R: a language and environment for statistical computing. <http://www.R-project.org/>.
- Reis, E. M., and M. A. Carmona, 2013 Integrated disease management of Fusarium head blight, pp. 159–173 in *Fusarium Head Blight in Latin America*, edited by Alconada Magliano, T. M., and S. N. Chulze. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-7091-1_10
- Rosa, S. B., C. M. Zanella, C. W. Hiebert, A. B. Babel, H. Randhawa *et al.*, 2019 Genetic characterization of leaf and stripe rust resistance in the Brazilian cultivar Toropi. *Phytopathology* 109: 1760–1768. <https://doi.org/10.1094/PHYTO-05-19-0159-R>
- Scheeren, P. L., E. Caierão, M. S. Silva, A. J. Nascimento, V. R. Caetano *et al.*, 2008 Challenges to wheat production in Brazil, pp. 167–170 in *International Symposium on Wheat Yield Potential: Challenges to International Wheat Breeding*, edited by M. P. Reynolds, J. Pietragalla, and H. J. Braun. CIMMYT, Mexico City, DF.
- Scherlosky, A., V. S. Marchioro, F. A. F. Franco, A. L. Vrancini, and I. Schuster, 2018 Genetic variability of Brazilian wheat germplasm obtained by high-density SNP genotyping. *Crop Breed. Appl. Biotechnol.* 18: 399–408. <https://doi.org/10.1590/1984-70332018v18n4a59>
- Schuster, I., E. Serra Negra Vieira, G. Jaqueline de Silva, F. de Assis Franco, and V. Sérgio Marchioro, 2009 Genetic variability in Brazilian wheat cultivars assessed by microsatellite markers. *Genet. Mol. Biol.* 32: 557–563. <https://doi.org/10.1590/S1415-47572009005000045>
- Speed, D., G. Hemani, M. R. Johnson, and D. J. Balding, 2012 Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.* 91: 1011–1021. <https://doi.org/10.1016/j.ajhg.2012.10.010>
- Srinivasachary, N., Gosman, A. Steed, J. Simmonds, M. Leverington-Waite *et al.*, 2008 Susceptibility to *Fusarium* head blight is associated with the *Rht-D1b* semi-dwarfing allele in wheat. *Theor. Appl. Genet.* 116: 1145–1153.
- Tessmann, E. W., Y. Dong, and D. A. Van Sanford, *et al.*, 2019 GWAS for Fusarium head blight traits in a soft red winter wheat mapping panel. *Crop Sci.* 59: 1823–1837.
- Trick, M., N. M. Adamski, S. G. Mugford, C. C. Jiang, M. Febrer *et al.*, 2012 Combining SNP discovery for next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC Plant Biol.* 12: 14. <https://doi.org/10.1186/1471-2229-12-14>
- Wang, S., D. Wong, K. Forrest, A. Allen, and S. Chao, 2014 Characterization of polyploid wheat genome diversity using a high-density 90000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12: 787–796. <https://doi.org/10.1111/pbi.12183>
- Wang, R., J. Chen, J. A. Anderson, J. Zhang, W. Zhao *et al.*, 2017 Genome-wide association mapping of Fusarium head blight resistance in spring wheat lines developed in the Pacific Northwest and CIMMYT. *Phytopathol.* 107: 1486–1495.
- Wang, D., K. Yu, D. Jin, L. Sun, J. Chu *et al.*, 2019 *ALI-1*, candidate gene of *B1* locus, is associated with awn length and grain weight in common wheat. [bioRxiv 688085](https://doi.org/10.1101/2019.08.08.688085)
- Wickham, H., 2016 *ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York.
- Wu, L., Y. Zhang, Y. He, P. Jiang, X. Zhang *et al.*, 2019 Genome-wide association mapping of resistance to Fusarium head blight spread and deoxynivalenol accumulation in Chinese elite wheat germplasm. *Phytopathol.* 109: 1208–1216.
- Yan, L., M. Helguera, K. Kato, S. Fukuyama, J. Sherman *et al.*, 2004 Allelic variation at the *VRN-1* promoter region in polyploid wheat. *Theor. Appl. Genet.* 109: 1677–1686. <https://doi.org/10.1007/s00122-004-1796-4>
- Yoshioka, M., J. C. M. Iehisa, and R. Ohno, 2017 Three dominant awnless genes in common wheat: fine mapping, interaction and contribution to diversity in awn shape and length. *PLoS One* 12: e0176148. <https://doi.org/10.1371/journal.pone.0176148>
- Zhang, L., P. G. Luo, Z. L. Ren, and H. Y. Zhang, 2011a Controlling fusarium head blight of wheat (*Triticum aestivum* L.) with genetics. *Adv. Biosci. Biotechnol.* 2: 263–270. <https://doi.org/10.4236/abb.2011.24038>
- Zhang, Y., J. Tang, Y. Zhang, J. Yan, Y. Ziao *et al.*, 2011b QTL mapping for quantities of protein fractions in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 122: 971–987. <https://doi.org/10.1007/s00122-010-1503-6>

Communicating editor: J. Holland