



# Haplotype Loci Under Selection in Canadian Durum Wheat Germplasm Over 60 Years of Breeding: Association With Grain Yield, Quality Traits, Protein Loss, and Plant Height

## OPEN ACCESS

### Edited by:

Agata Gadaleta,  
Università degli Studi di Bari, Italy

### Reviewed by:

Ivan A. Matus,  
Instituto de Investigaciones  
Agropecuarias (INIA), Chile  
Shichen Wang,  
Texas A&M University, United States

### \*Correspondence:

Amidou N'Diaye  
amidou.ndiaye@usask.ca  
Curtis J. Pozniak  
curstis.pozniak@usask.ca

†Retired

### Specialty section:

This article was submitted to  
Plant Breeding,  
a section of the journal  
Frontiers in Plant Science

**Received:** 10 August 2018

**Accepted:** 15 October 2018

**Published:** 05 November 2018

### Citation:

N'Diaye A, Haile JK, Nilsen KT,  
Walkowiak S, Ruan Y, Singh AK,  
Clarke FR, Clarke JM and Pozniak CJ  
(2018) Haplotype Loci Under  
Selection in Canadian Durum Wheat  
Germplasm Over 60 Years  
of Breeding: Association With Grain  
Yield, Quality Traits, Protein Loss,  
and Plant Height.  
Front. Plant Sci. 9:1589.  
doi: 10.3389/fpls.2018.01589

Amidou N'Diaye<sup>1\*</sup>, Jemanesh K. Haile<sup>1</sup>, Kirby T. Nilsen<sup>1</sup>, Sean Walkowiak<sup>1</sup>, Yuefeng Ruan<sup>2</sup>, Asheesh K. Singh<sup>3</sup>, Fran R. Clarke<sup>2†</sup>, John M. Clarke<sup>1</sup> and Curtis J. Pozniak<sup>1\*</sup>

<sup>1</sup> Department of Plant Sciences, Crop Development Centre, University of Saskatchewan, Saskatoon, SK, Canada,

<sup>2</sup> Agriculture and Agri-Food Canada, Swift Current Research and Development Centre, Swift Current, SK, Canada,

<sup>3</sup> Department of Agronomy, Iowa State University, Ames, IA, United States

Durum wheat was introduced in the southern prairies of western Canada in the late nineteenth century. Breeding efforts have mainly focused on improving quality traits to meet the pasta industry demands. For this study, 192 durum wheat lines were genotyped using the Illumina 90K Infinium iSelect assay, and resulted in a total of 14,324 polymorphic SNPs. Genetic diversity changed over time, declining during the first 20 years of breeding in Canada, then increased in the late 1980s and early 1990s. We scanned the genome for signatures of selection, using the total variance Fst-based outlier detection method (Lositan), the hierarchical island model (Arlequin) and the Bayesian genome scan method (BayeScan). A total of 407 outliers were identified and clustered into 84 LD-based haplotype loci, spanning all 14 chromosomes of the durum wheat genome. The association analysis detected 54 haplotype loci, of which 39% contained markers with a complete reversal of allelic state. This tendency to fixation of favorable alleles corroborates the success of the Canadian durum wheat breeding programs over time. Twenty-one haplotype loci were associated with multiple traits. In particular, *hap\_4B\_1* explained 20.6, 17.9 and 16.6% of the phenotypic variance of pigment loss, pasta b\* and dough extensibility, respectively. The locus *hap\_2B\_9* explained 15.9 and 17.8% of the variation of protein content and protein loss, respectively. All these pleiotropic haplotype loci offer breeders the unique opportunity for further improving multiple traits, facilitating marker-assisted selection in durum wheat, and could help in identifying genes as functional annotations of the wheat genome become available.

**Keywords:** haplotype, loci under selection, durum wheat, quality traits, grain yield, protein loss, plant height

## INTRODUCTION

Durum wheat (*Triticum turgidum* L. ssp. *durum* Desf. Husn.,  $2n = 4x = 28$ ; genome AABB) is an important crop in Canada, grown on an average of approximately 2 million hectares and comprising about 25% of total wheat area (Canada, 2018). Nearly all of Canada's wheat is produced in the western prairie provinces of Saskatchewan, Alberta and Manitoba, with a relatively small area in British Columbia and eastern Canada (McCallum and DePauw, 2008). Durum wheat was introduced into western Canada in the late nineteenth century (see Dexter, 2008 for a detailed history of Canadian durum wheat breeding) and planned hybridization and targeted selection started in 1928 (Clarke J.M. et al., 2010). However, the first variety developed in Canada, Stewart 63, was not released until 1963 (Dexter, 2008).

The improvement of quality traits, such as yellow pigment and gluten strength, was a major focus for durum breeding to satisfy the requirements of the global pasta industry. Canadian durum wheat is classified into four Canada Western Amber Durum (CWAD) wheat milling grades defined by the Canadian Grain Commission (Dexter and Edwards, 1998). Only varieties that meet a set of requirements for a grade are registered. Specifications for new varieties continue to evolve in response to the feedback of CWAD customers (Dexter, 2008). Durum breeding in Canada has made steady genetic progress to improve yield and agronomic traits. This was done concomitantly with improvements in end-use quality attributes such as grain protein concentration, yellow pigment concentration and gluten strength, while improving or maintaining resistance to disease (Clarke J.M. et al., 2010). Grain protein concentration and gluten strength are crucial factors in pasta manufacturing and cooking quality (Feillet and Dexter, 1996). These and other quality trait targets have indirectly driven durum wheat breeders to design hybridization programs within narrow limits, using a similar set of standard cultivars as donors of these quality traits. In particular, high grain protein concentration is a requirement for durum cultivar registration in Canada and this likely limited grain yield gain (Clarke J.M. et al., 2010) due to the generally negative relationship between grain yield and protein concentration (Clarke et al., 2009).

When it is necessary to bring new diversity for particular traits into the breeding programs, it is imperative to return the elite materials to a new state of equilibrium as quickly as possible. Efficient means to identify both core essential adaptation and quality traits in addition to new traits being introgressed can facilitate this process. Starting in the late 1990s, molecular markers became an important tool for Canadian wheat breeding programs. However, the lack of tightly linked diagnostic markers, QTL  $\times$  environmental interactions and prevalence of QTL background effects have limited the application of marker assisted breeding for some traits (Randhawa et al., 2013). Advances in high-throughput genotyping platforms at a low cost have now made it possible to consider using empirical LD patterns to conduct genome-wide scans to link markers with phenotypes of interest. Additionally, the recent availability of high quality genome assemblies for tetraploid wheat (Avni et al., 2017) could be a valuable

source to facilitate the identification of loci and genes of interest.

Association mapping (AM) is increasingly being adopted as a complementary method to bi-parental linkage mapping to identify genotype–phenotype associations. Several AM studies have been conducted to dissect the genetic basis of durum grain yield (Maccaferri et al., 2011; Mengistu et al., 2016; Kidane et al., 2017; Sukumaran et al., 2018), grain and semolina quality traits (Fiedler et al., 2017), semolina and pasta color (N'Diaye et al., 2017), agronomic and morphological traits (Hu et al., 2015) and grain cadmium concentration and yellow color loss during pasta manufacturing (Pozniak et al., 2012) in durum wheat. Recently, the haplotype-based AM approach was suggested as an efficient method for investigating the genetic basis of traits of interest in durum wheat by detecting more loci (N'Diaye et al., 2017), capturing epistatic interactions and reducing type I error rate (Morris and Kaplan, 2002; Zhao et al., 2007; Hamblin and Jannink, 2011). All of these genome-wide association studies in durum wheat populations were performed using the whole set of polymorphic SNPs, but not loci under selection over the course of durum wheat breeding. However, the detection of selection signatures is gaining ground in modern population genetics (Fariello et al., 2013). The growing availability of large-scale genotyping data has facilitated the identification of regions targeted by natural and/or artificial selection in wild and domesticated populations of plants, animals and humans (Nielsen et al., 2007). The search for molecular signatures aims to uncover the evolutionary past of species, understand their functional or adaptive importance, and detect associations between these genomic regions and traits of interest (López et al., 2015). It has become an efficient approach in biomedical sciences to identify genes related to disease resistance (Tishkoff et al., 2001; Barreiro et al., 2008; Albrechtsen et al., 2010; Fumagalli et al., 2010; Cagliani et al., 2011), adaptation to climate (Lao et al., 2007; Sturm, 2009; Rees and Harding, 2012), or altitude (Bigham et al., 2010; Simonson et al., 2010). In livestock species, where artificial selection has been carried out by humans since domestication, it contributes to mapping traits of commercial interest (Guan et al., 2016; Liu et al., 2016; Taye et al., 2017).

Therefore, the main objectives of this study were to scan the genome of elite Canadian durum wheat lines tested in the pre-registration trial for selection footprints and relate these genomic regions to phenotypes. The availability of such loci targeted by selection would be a valuable resource for developing markers and/or investigating gene candidates controlling the traits of interest we analyzed, in particular grain yield, protein loss and pasta quality traits.

## MATERIALS AND METHODS

### Plant Materials and Phenotypic Data Analyses

One hundred and ninety-two durum wheat lines, including the 169 lines from our previous study (N'Diaye et al., 2017) from the official Canadian durum cultivar registration trial

(Durum wheat Cooperative Test) were used for this project (**Supplementary Table S1**). Phenotypic data and trials were described in previous reports (Clarke J.M. et al., 2010; Pozniak et al., 2012; Haile et al., 2018). Briefly, the trials were run under the auspices of science/industry groups responsible for recommending cultivars for registration by the Canadian Food Inspection Agency and grown each year at 10–12 locations in western Canada and one in the United States. Candidate lines were tested for 1–3 years, those that did not meet the merit requirements being withdrawn, with 3 years of data required for registration of cultivars. Each trial included four or five check cultivars. The checks AC Avonlea, AC Morse, AC Navigator and Strongfield were in the trials for the years since 1999, and Commander since 2001. Trials were arranged in lattice designs with four replications. For the end-use quality traits, we evaluated gluten index described by Haile et al. (2018), semolina pigment, semolina b\*, pasta b\* and pigment loss as previously described (N'Diaye et al., 2017). We also included alveograph measures (Haile et al., 2018): dough tenacity (P), dough extensibility (L), and deformation energy (W). Protein loss was estimated as the difference between grain protein and semolina protein concentration.

The historical and unbalanced phenotypic data were analyzed using SAS version 9.4 PROC HP MIXED with three models due to the different data structures to calculate best linear unbiased predictions (BLUPs). For grain yield concentration which was measured on composites of locations within years (175 lines), year, location, replication, and genotype, and their interactions were considered random; for grain protein and yellow pigment concentrations (186 lines), years, locations, genotypes, and interactions were random; for gluten strength and color traits measured on yearly composites (170 lines), genotypes and years were random. The analyses included all genotypes tested (up to approximately 300 depending on trait) in the registration trials, not just those used in the present study, to provide a better estimate of random variances and covariances (Clarke F.R. et al., 2010; Pozniak et al., 2012).

## SNP Genotyping and Data Curation

DNA extraction and genotyping with the 90K iSelect assay chip were carried out as reported in our previous paper (N'Diaye et al., 2017). A total of 14,324 polymorphic SNPs were scored and missing calls were imputed using the RF regression procedure (Breiman, 2001) as implemented in the randomForest R package (Liaw and Wiener, 2002; R Development Core Team, 2013). After removing SNPs having MAF < 0.05, a total of 11,323 SNPs were kept for analyses.

Because relatively few semi-dwarf lines in the panel were selected for very high pigment, presenting the possibility of spurious associations, the lines were also genotyped with *Rht-B1b*, an allele known to confer semi-dwarf growth habit in wheat (Ellis et al., 2002). In order to distinguish the association signals from *Rht-B1b*, pairwise LD ( $r^2$ ) was performed between all 4B association signals and this gene using MIDAS software (Gaunt et al., 2006).

## Population Structure and Genetic Diversity Analysis

First, duplicate SNPs were removed using an in-house Ruby script, as previously described (N'Diaye et al., 2017). Then, a subset of 4,235 highly polymorphic ( $0.32 \leq \text{PIC} \leq 0.45$ ) SNPs were selected for the clustering analysis (Kabbaj et al., 2017). The population structure among the 192 breeding lines was investigated using the discriminant analysis of principal components (DAPC) as implemented in the Adegenet R package (Jombart, 2008; Jombart et al., 2010; Jombart and Ahmed, 2011). The number of clusters (K) for the DAPC was estimated from the lowest value of the Bayesian information criterion (BIC) according to Jombart et al. (2010).

The genetic differentiation between populations derived from the DAPC was assessed by pairwise  $F_{st}$  values, and the variance between and within populations was calculated using the analysis of molecular variance (AMOVA) as implemented in the Arlequin 3.5 software (Excoffier and Lischer, 2010). Significance levels for variance components and  $F_{st}$  statistics were estimated based on 10,000 permutations. To analyze the changes in diversity over time, the breeding lines were assigned to temporal groups (decades) according to the time of entry into the Durum wheat Cooperative Test and the genetic diversity ( $\pi$ ) was computed according to Nei and Li (1979), using Arlequin 3.5 software (Excoffier and Lischer, 2010).

## Selection Signatures

The subset of 4,235 highly polymorphic SNPs was used to scan the genome for signatures of selection, using three approaches: the total variance  $F_{st}$ -based outlier detection method implemented in Lositan (Antao et al., 2008), the hierarchical island model with 100,000 simulations using Arlequin (Excoffier and Lischer, 2010) and the Bayesian genome scan approach implemented in BayeScan (Foll and Gaggiotti, 2008). For the latter, we performed 20 pilot runs of 50,000 iterations, followed by 100,000 iterations on a sample size of 5000 and thinning interval of 10.

For Lositan, markers were considered under divergent selection if the  $F_{st}$  values were higher than 99% of the neutral distribution. For Arlequin, loci were considered being under selection if the observed  $F_{st}$  values were higher than expected on the basis of neutral variation, and showing  $F_{st}$  out of the 99% quantile based on coalescent simulations (Beaumont and Nichols, 1996). We identified markers under selection with BayeScan by comparing posterior probabilities and threshold values obtained from the FDR (FDR  $q$ -values < 0.05). As a result, the  $F_{st}$  cut-off for declaring outliers was 0.20 for Arlequin and Lositan, and 0.15 for BayeScan (**Supplementary Figure S1**).

## Haplotype Blocks and Association Study

Haplotype blocks were built as previously described (N'Diaye et al., 2017), with little modifications. Only markers that showed strong evidence of directional selection were clustered into haplotype blocks based on the durum wheat consensus map (Maccaferri et al., 2014), using an in-house python script. These

haplotype loci were used to perform association studies in order to relate them to the traits of interest (e.g., gluten index, semolina pigment, grain yield, protein content, protein loss, and plant height) that have been scored over years. The haplotype-trait association analyses were carried out using a Mixed Linear Model (MLM) (Yu et al., 2006) with either the Kinship matrix alone (MLM-K) or the Q matrix from the DAPC plus Kinship (MLM-QK) as random effect, using TASSEL software version 3 (Bradbury et al., 2007). The Q-Q (quantile-quantile) plot profiles (**Supplementary Figure S2**) showed that MLM-K better controlled the *P*-value inflation while MLM-QK led to overcorrections. Therefore, all association analyses were carried out using the MLM-K model. A false discovery rate (FDR) of 5% was applied and haplotype loci with *P*-value  $\leq 0.05$  were declared significant. The allelic effect of haplotypes was estimated as the difference between the mean value of the lines carrying these haplotypes and the mean value of the entire

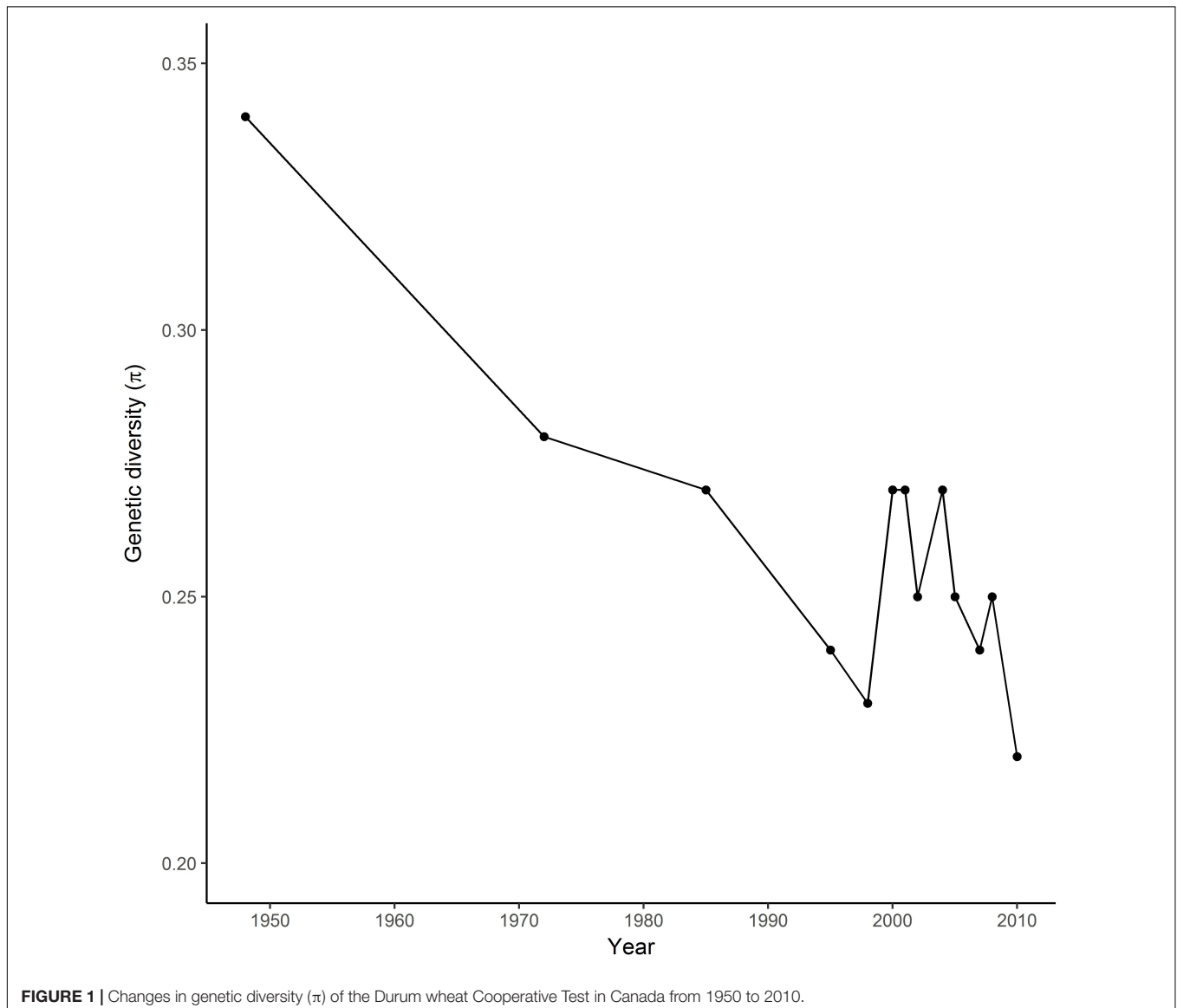
population for each trait, as previously described (N'Diaye et al., 2017).

## RESULTS

### Genetic Diversity and Population Stratification

SNP markers used for the analysis were distributed across all 14 chromosomes of the durum wheat genome. Genetic diversity in the registration trials changed over time, declining during the first 20 years of breeding in Canada when the germplasm shifted from introduced cultivars in the 1940s and 1950s to locally-bred lines in the 1970s (**Figure 1**). Diversity increased in the late 1980s and early 1990s, then started a decrease in the 2005s.

The genetic relationship between the 192 durum lines as revealed by the DAPC is illustrated in **Figure 2**. Based on the



lowest BIC, the lines were clustered into four sub-populations. However, the pairwise  $F_{st}$  analysis showed only moderate ( $0.05 < F_{st} \leq 0.15$ ) genetic differentiation between the four sub-populations (Table 1). The AMOVA revealed that only 9.8% of the genetic variation is found between sub-populations, whereas 85.2% of the genetic variation resides between individuals (Table 2).

## Loci Under Selection

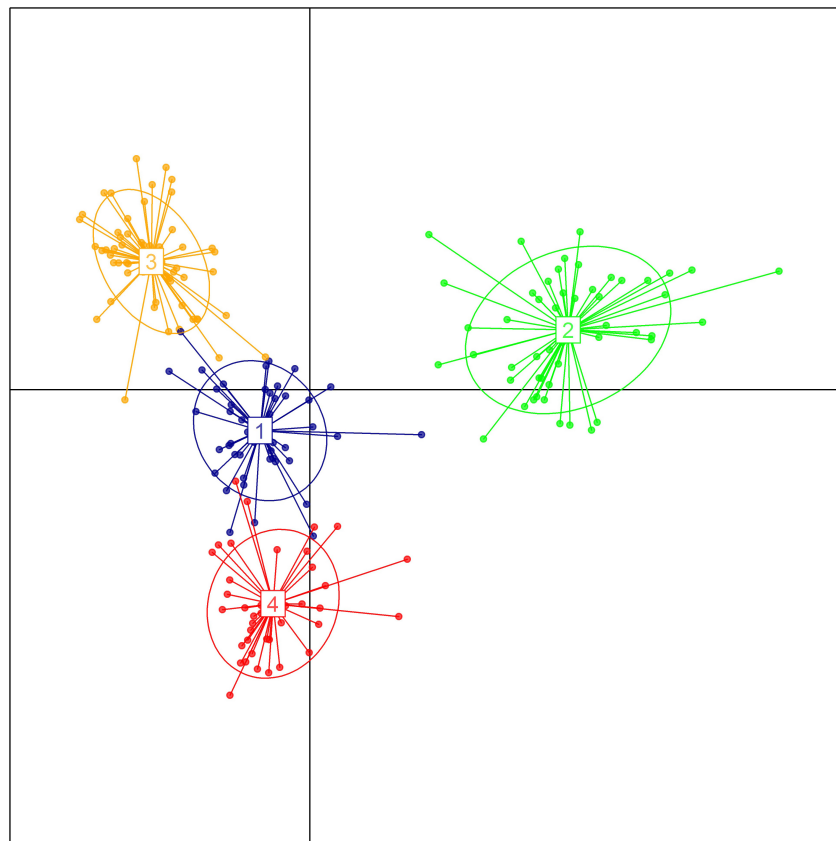
The  $F_{st}$  distribution and threshold (0.15 for BayeScan, 0.2 for both Arlequin and Lositan) to declare loci being under selection are shown on Supplementary Figure S1. From the total of 4,235 most informative SNPs, 407 appeared to be under selection, spanning all 14 chromosomes of the durum wheat genome (Figure 3). Lositan detected 403 outliers, including all 397 markers from Arlequin (Figure 4). BayeScan identified only 144 outliers, of which 4 SNPs were undetected by Lositan and BayeScan. Twenty-three percent (95/407) of markers under selection showed a complete reversal of allelic state in response to selection over time. However, the time at which the allelic state changed was different depending on the marker. For example, the switch of allele's frequency occurred in the early 1960s (e.g., BobWhite\_c8016\_301, BS00009060\_51, and BS00087544\_51), in the mid-1980s (BS00091561\_51,

BS00097263\_51, and CAP7\_c11156\_108) or in the late 2000s (e.g., Excalibur\_c29304\_176, IACX6346, and Ra\_c11263\_2353) during the breeding program. Of the markers that showed a complete reversal of allelic state, 11% are fixed (allele frequency = 1) in the population (Supplementary Figure S3).

The distribution pattern of markers under selection was different between chromosomes (Figure 3), outliers spanned several genomic regions on the entire chromosome (e.g., 1B, 2B, and 7A) or were clustered at only few localized regions on the chromosome (e.g., 2A and 4A). The markers under selection were clustered into 84 LD-based haplotype loci, containing 1–28 SNPs (Supplementary Table S2). The number of haplotype loci varied among chromosomes, ranging from 3 (chromosomes 2A, 3B) to 11 (chromosome 7A). Thirty-six (30/84) percent of the haplotype loci under selection harbored markers with a complete reversal of allelic state in response to selection over time; the proportion of SNPs with a complete reversal of allelic ranging from 20 (*hap\_5A\_5*) to 100% (e.g., *hap\_1B\_5*).

## Phenotypes Analysis

Wide phenotypic variation was observed among lines in the breeding panel for all of the traits (Table 3 and Supplementary Figure S4). Sample means were significantly ( $P$ -value  $\leq 0.05$ ) different between sub-populations for all of the traits, except



**FIGURE 2 |** Population structure of the breeding panel as revealed by discriminant analysis of principal components. The axes represent the first two Linear Discriminants. Each color represents a sub-population.

**TABLE 1** | Pairwise *F*<sub>st</sub> between subpopulations, SP1 (*n* = 45), SP2 (*n* = 50), SP3 (*n* = 57), and SP4 (*n* = 40).

Sub-populations	SP1	SP2	SP3	SP4
SP2	0.11			
SP3	0.06	0.15		
SP4	0.07	0.10	0.09	

**TABLE 2** | Analysis of molecular variance (AMOVA).

Sources of variation	d. f.	Sum of squares	Variance components	Variation (%)
Populations (P)	3	22,270.31***	70.30	9.80
Individuals within P	188	129,003.06***	35.85	5.00
Individuals	192	117,895.00***	610.85	85.20
Total	383	269,168.37	717.00	

\*\*\*The source of variation was highly significant at  $P \leq 0.001$ .

grain yield. Many traits were significantly ( $P$ -value  $\leq 0.05$ ) correlated to each other (Figure 5). In particular, pasta b\* was strongly positively correlated with pigment loss, semolina pigment and semolina b\*,  $r = 0.80$ ,  $0.66$ , and  $0.68$ , respectively. Grain protein and protein loss showed significant ( $P$ -value  $< 0.05$ ) positive correlation,  $r = 0.82$ .

## Haplotype-Traits Association Analyses

The association analysis detected 54 haplotype loci, of which 39% (21/54) were associated with at least two traits. For quality traits (Table 4), 49 loci were detected, spanning all chromosomes. In particular, *hap\_1B\_2* on chromosome 2B was associated with dough extensibility, gluten index, dough tenacity and deformation energy while *hap\_4B\_2* (chromosome 4B) was associated with pigment loss, pasta b\*, protein loss, semolina pigment and protein content. The phenotypic variations explained by the pleiotropic loci varied depending on the traits. For example, *hap\_4B\_1* explained 20.6, 17.9, 16.6, and 12.1% of the phenotypic variance of pigment loss, pasta b\*, dough extensibility and gluten index, respectively, whereas *hap\_4B\_2* explained 6.8% (semolina pigment, protein content) to 16.8% (protein loss) of the phenotypic variation.

For protein content, six loci were detected, spanning chromosomes 1B, 2B, 4B, and 7A (2 loci) and explained 3.5 (*hap\_7A\_6*) to 15.9% (*hap\_2B\_9*) of the phenotypic variation (Table 4).

Five loci, located on chromosomes 1B, 2A, 2B, 4B, 7A, and 7B, were associated with protein loss (Table 4). In particular, *hap\_2B\_9*, *hap\_2A\_3*, and *hap\_1B\_10* explained 17.8, 15.3, and 13.4% of the phenotypic variance, respectively.

Semolina pigment showed associated loci on all 14 chromosomes of the durum wheat genome. In particular, haplotype loci *hap\_3A\_5*, *hap\_7A\_3*, *hap\_1B\_5*, *hap\_2A\_3*, and *hap\_5A\_3* explained 23.4, 21.2, 15.8, 15.4, and 15.3% of the phenotypic variation, respectively.

For pigment loss, three loci were detected, *hap\_1A\_4*, *hap\_4B\_1*, and *hap\_4B\_2*, explaining 6.7, 20.6, and 16.8% of the phenotypic variation.

A total of five loci, *hap\_1A\_1*, *hap\_1B\_2*, *hap\_2B\_7*, *hap\_4B\_1*, and *hap\_7A\_6* were associated with gluten index. The locus *hap\_4B\_1* gave the highest allelic effect (21.3) and explained 12.1 % of gluten index variation.

Five loci, located on chromosomes 1A, 4B (2 loci), 6A, and 7B, were significantly ( $P$ -value  $\leq 0.05$ ) associated with pasta b\*, explaining 5.4 (*hap\_6A\_3*) to 17.9% (*hap\_4B\_1*) of the phenotypic variation.

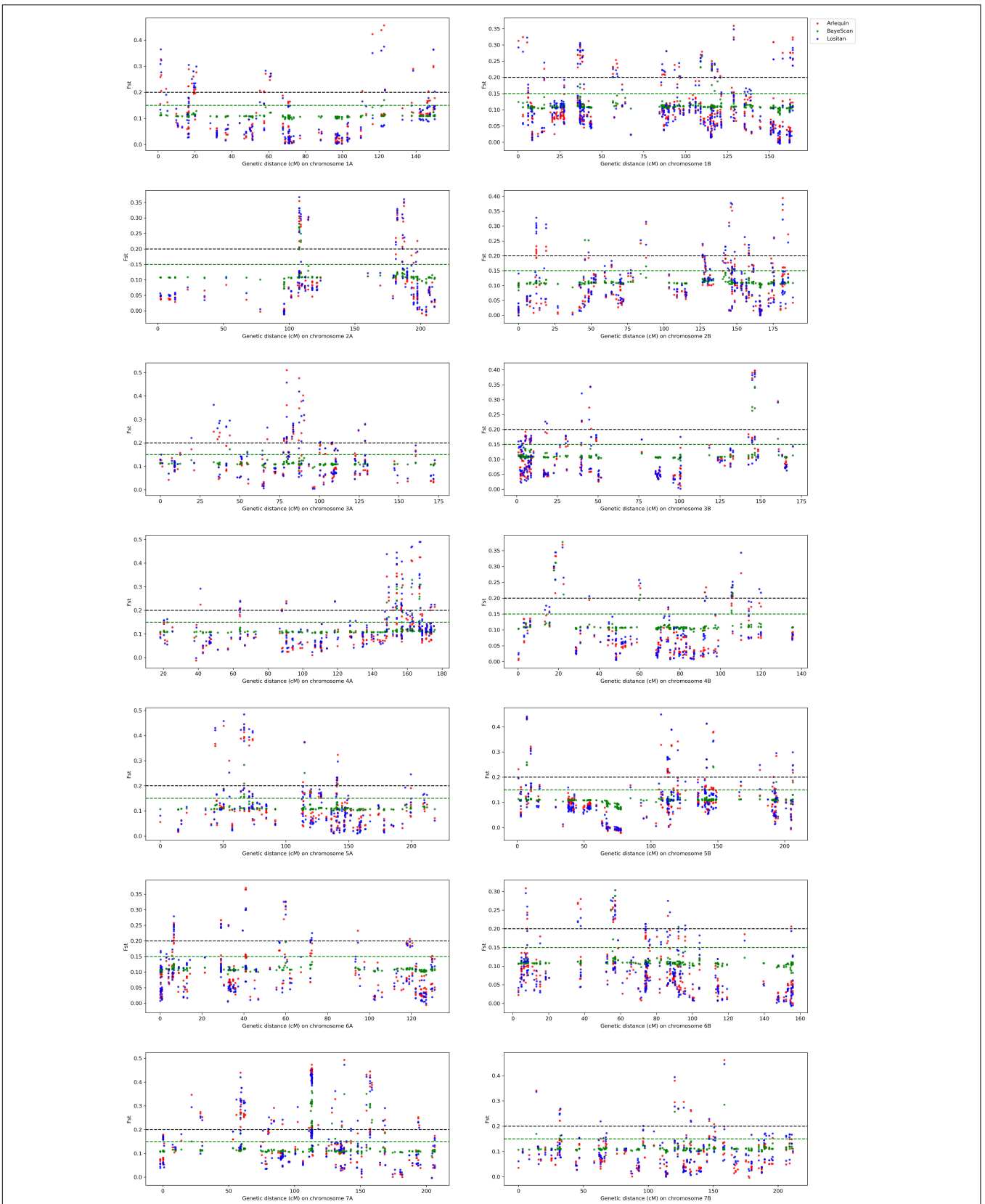
A total of 11 haplotype loci were associated with plant height and/or grain yield (Table 5). These loci spanned chromosomes 2A, 4A, 4B, 5A, 7A, and 7B for plant height; and 1B, 2B, 4A, and 4B for grain yield. In particular, the loci *hap\_4A\_5* and *hap\_4B\_1* controlled both plant height and grain yield. The locus *hap\_4A\_5*, which showed the highest allelic effect (190.03) on grain yield, explained 43.9 and 38.5% of the phenotypic variation of plant height and grain yield, respectively. On the other hand, *hap\_4B\_1* was strongly associated ( $r^2 = 0.90$ ) with the dwarfing gene *Rht-B1b* and explained 30.0% of the phenotypic variation of plant height.

Thirty-nine (21/54) percent of haplotype loci associated with the traits we investigated contained markers with a complete reversal of allelic state (Supplementary Table S3). In particular, *hap\_1B\_5* and *hap\_2A\_1* consisted of 100% (7/7) and 91% (10/11) markers with a complete reversal of allelic state, respectively. For each haplotype locus, the time at which the switch in allelic state occurred varied among SNPs, regardless of the number of traits the haplotype locus was associated with. For example, for *hap\_1B\_5* associated with only semolina pigment, the switch in allelic state occurred in the early 1960s (Tdurum\_contig85180\_99), in the late 1980s (BobWhite\_c17644\_456) and in the 2000s (IACX6346) (Supplementary Figure S5). For *hap\_2A\_1* that was associated with plant height and dough extensibility, the switch in allelic state occurred in the 1970s (BS00097263\_51), in the 1985s (e.g., CAP7\_c239\_267, CAP7\_c11156\_108), and in the 1990s (wsnp\_Ex\_rep\_c66615\_64916512) (Supplementary Figure S6).

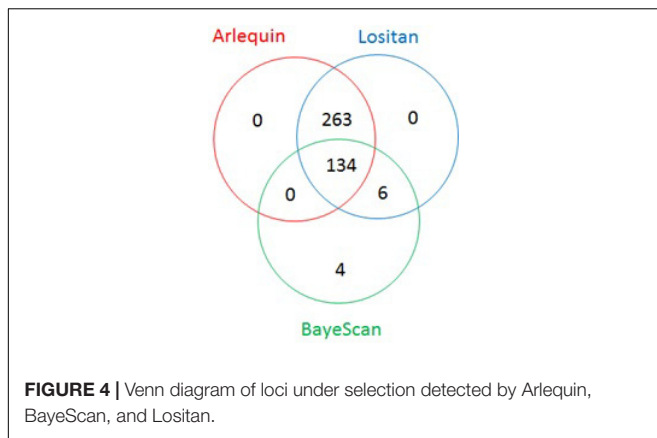
## DISCUSSION

### Population Stratification and Genetic Diversity Over Time

The presence of genetic structure within a population can lead to spurious association signals (Marchini et al., 2004; Kang et al., 2008; Myles et al., 2009; Mezouk et al., 2011; Cappa et al., 2013). Therefore, the analysis of the actual population structure of the durum breeding panel was intended to limit the FDR in the association analysis. The discriminant analysis of principal components (Jombart et al., 2010) clustered the 192 lines into four sub-populations. Nonetheless, the genetic differentiation between the four sub-populations was moderate ( $0.05 < F_{st} \leq 0.15$ ), whereas 85.2% of the genetic variation resided between individuals. Therefore, the association analysis accounted only for kinship (MLM-K model). The population structure is in agreement with known differences in pedigree, breeding program sources and era of testing in the trials, as described in our previous study (N'Diaye et al., 2017).



**FIGURE 3 |** Genome-wide analysis of loci under selection using Arlequin (red), BayeScan (green), and Lositan (blue). The horizontal black dash line indicates the threshold for selection with Arlequin and Lositan while the horizontal green dash line shows the selection threshold for BayeScan.



Similarly, many other studies about genetic diversity changes in durum wheat germplasm over time revealed that most (up to 91%) of the variations were found between individuals within groups compared to that between groups (Henkrar et al., 2016).

Durum breeding began in Canada in the early 1950s, with the first locally-bred cultivar registered in 1963. The Canadian wheat industry has strict processing quality requirements for registration of durum cultivars, so the majority of breeding crosses involve closely-related local materials, with relatively little use of either introduced modern cultivars or landraces except where absolutely necessary to introduce a new trait. Genetic diversity in the registration trials changed over time, declining during the first 20 years of breeding in Canada when the germplasm shifted from introduced cultivars in the 1940s and 1950s to locally-bred lines in the 1970s (Figure 1). Diversity increased in the late 1980s and early 1990s coinciding with introgression of higher yellow pigment concentration, gluten strength and low grain cadmium concentration (Clarke J.M. et al., 2010), and with major expansion of Canadian durum breeding programs due to increased funding. The sources of these traits were improved cultivars from CIMMYT, Germany, Italy, and United States. Further cycles of crossing and selection then reduced diversity observed during the past decade to a level similar to the 1970s. In contrast, a relatively stable diversity over 25 years was observed in the CIMMYT Elite Spring Wheat Yield Trial (Dreisigacker et al., 2012), probably reflecting consistent usage of diverse parents to meet requirements for adaptation to global production environments. Similarly, the impact of traits improvement on allelic changes over time has been described in Canadian bread wheat cultivars (Fu and Somers, 2011).

The impacts of modern plant breeding on crop genetic diversity has been a major concern of many scientists in search for implementing efficient conservation strategies and a better utilization of germplasms (see Duvick, 1984; Fu, 2007, 2015 for review). In particular, some temporal patterns of genetic diversity have been observed (Rauf et al., 2010; van de Wouw et al., 2010b,c). The high selection pressure in modern plant breeding has reduced crops genetic diversity over time (Duvick, 1984; Hallauer and Darrah, 1985; Bowman, 2003; Gepts, 2006;

Fu and Dong, 2015). In particular, allelic reduction at specific loci and allele loss over time have been reported in the Canadian hard red spring wheat germplasm (Fu et al., 2005). However, no loss of allele's number was found in European winter wheat varieties over time (Huang et al., 2007). Similarly, a meta-analysis of 44 diversity studies did not reveal evidence of genetic erosion in released varieties (van de Wouw et al., 2010a). The discrepancy between genetic diversity trends in released cultivars and varieties was attributed to differences in breeding goals and methods that have been applied over time (Gepts and Hancock, 2006; Rauf et al., 2010). Many other sources that could affect the crops genetic diversity have been reported, including use of different genetic diversity measures, sampling bias and arbitrary grouping of cultivars to represent specific breeding periods (Reif et al., 2005; Fu, 2007; Aremu, 2011).

## Loci Under Selection

Loci under directional selection are expected to have lower intra-population variability and larger inter-population variability than neutral loci. Thus, loci under directional selection can be unveiled by patterns in heterozygosity and/or differences in  $F_{st}$  values (Eveno et al., 2008; Perez-Figueroa et al., 2010; Kirk and Freeland, 2011; Konijnendijk et al., 2015; Lin et al., 2017). Evidence for selection was investigated using, the total variance  $F_{st}$ -based outlier detection method implemented in Lositan (Antao et al., 2008), the hierarchical island model implemented in Arlequin (Excoffier and Lischer, 2010) and the Bayesian genome scan approach implemented in BayeScan (Foll and Gaggiotti, 2008). The number of outliers varied according to the method. Lositan detected 403 outliers, including all 397 markers from Arlequin. BayeScan identified the less number (144) of outliers of which 4 SNPs were undetected by Lositan and BayeScan. Comparison of several outlier detection methods showed that these methods differ in their type I (false positives) and type II (false negatives) error rates (Perez-Figueroa et al., 2010; Narum and Hess, 2011). Thus, combining the properties of different outlier detection methods could reduce the percentage of false positives and strengthen the candidate status of the identified outlier loci (Vasemagi et al., 2005). A similar approach has been reported for many crops, including wheat (Gao et al., 2017; Ren et al., 2017), rice (Xia et al., 2014), and apple (Khan et al., 2014). In general, researchers feel more comfortable when loci under selection are detected by at least two different methods. However, among the four outliers detected by only BayeScan, one marker representing the haplotype locus *hap\_3A\_3*, showed strong association with semolina pigment and explained 4.2% of the phenotypic variance (Table 4). Studies where only one method was used to scan for loci under selection have been reported elsewhere (Mäkinen et al., 2008; Jiao et al., 2012; Liao et al., 2013). Therefore, it seems reasonable to take into consideration any outlier detected by a single method even if it might not be ranked as a prime candidate.

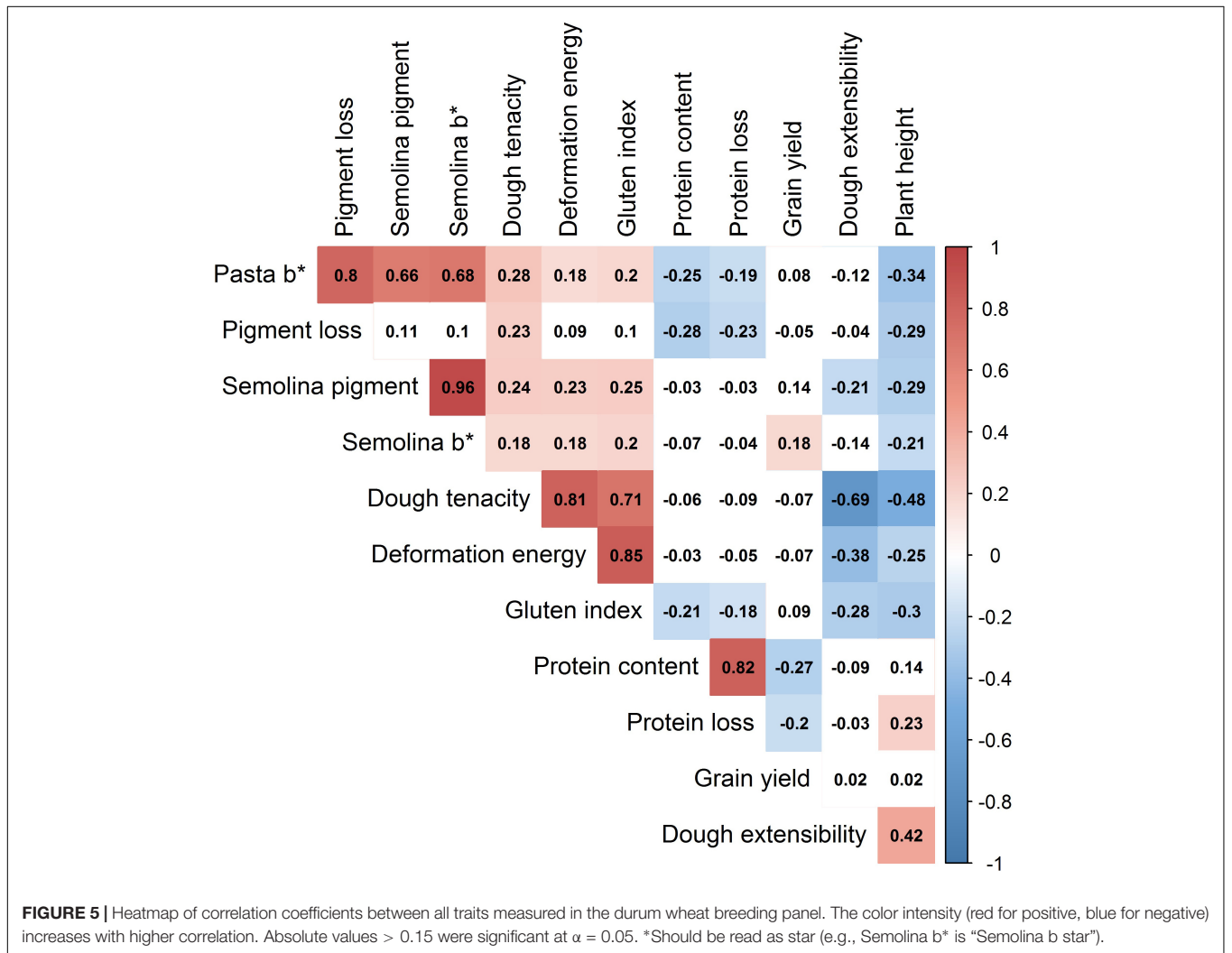
Markers under selection showed different distribution patterns, spanning several genomic regions or clustered at



**TABLE 3** | Comparison of traits means between sub-populations (SP1, SP2, SP3, and SP4).

Traits	Population	SP1	SP2	SP3	SP4	Size <sup>1</sup>
Dough tenacity (P)	78.98 (± 15.00)	85.73 <sup>a</sup>	78.81 <sup>b</sup>	71.61 <sup>c</sup>	85.59 <sup>a</sup>	168/38/47/45/38
Deformation energy (W)	204.23 (± 41.44)	221.82 <sup>a</sup>	204.47 <sup>b</sup>	177.16 <sup>c</sup>	218.37 <sup>ab</sup>	168/38/47/45/38
Dough extensibility (L)	100.23 (± 15.48)	97.10 <sup>a</sup>	95.94 <sup>b</sup>	111.20 <sup>a</sup>	95.70 <sup>b</sup>	168/38/47/45/38
Gluten index	63.93 (± 15.17)	69.45 <sup>a</sup>	65.20 <sup>a</sup>	55.69 <sup>b</sup>	66.57 <sup>a</sup>	169/38/48/45/38
Semolina pigment	8.34 (± 1.12)	8.27 <sup>b</sup>	8.84 <sup>a</sup>	7.77 <sup>c</sup>	8.60 <sup>ab</sup>	184/42/48/54/40
Pasta b*	64.75 (± 1.83)	64.73 <sup>b</sup>	64.27 <sup>b</sup>	64.55 <sup>b</sup>	65.62 <sup>a</sup>	168/38/47/45/38
Semolina b*	34.22 (± 1.59)	34.04 <sup>ab</sup>	34.65 <sup>a</sup>	33.68 <sup>b</sup>	34.53 <sup>a</sup>	167/38/46/45/38
Pigment loss	-0.07 (± 1.36)	0.32 <sup>a</sup>	-0.86 <sup>b</sup>	0.10 <sup>a</sup>	0.59 <sup>a</sup>	168/38/47/45/38
Protein content	13.65 (± 0.27)	13.63 <sup>b</sup>	13.78 <sup>a</sup>	13.59 <sup>b</sup>	13.58 <sup>b</sup>	173/39/50/46/38
Protein loss	0.95 (± 0.23)	0.91 <sup>b</sup>	1.06 <sup>a</sup>	0.93 <sup>b</sup>	0.88 <sup>b</sup>	173/39/50/46/38
Plant height	89.11 (± 5.04)	88.83 <sup>b</sup>	88.78 <sup>b</sup>	91.69 <sup>a</sup>	86.71 <sup>b</sup>	173/39/50/46/38
Grain yield	4088.91 (± 99.41)	4101.94 <sup>a</sup>	4098.36 <sup>a</sup>	4076.34 <sup>a</sup>	4078.34 <sup>a</sup>	175/39/50/46/38

<sup>1</sup>Number of lines in the population (entire panel), SP1, SP2, SP3, and SP4, respectively. Values with the same appended letter are not significantly different according to the least significant difference test at  $p < 0.05$  (for each trait). \*Should be read as star (e.g., Pasta b\* is "Pasta b star").



only a few localized regions on the chromosome, reflecting linkage and/or pleiotropy. It is now well documented that genes are not evenly distributed across the genome and

QTL might be clustered in 'hot regions'. The clustering of QTL for different traits have been reported for many crops, including rice (Cai and Morishima, 2002; Crowell et al., 2016;

**TABLE 4 |** Haplotypes loci significantly ( $P$ -value  $\leq 0.05$ ) associated with quality traits as revealed by the Mixed Linear Model with Kinship matrix (MLM-K).

Haplotypes	Position <sup>a</sup>	Traits	$P$ -value	$R^2$ (%)	Allelic effect <sup>b</sup>
hap_1A_1	1A (1.3–1.7)	Dough extensibility	8.71E-03	9.6	15.94
		Deformation energy	1.42E-02	8.8	30.68
		Semolina pigment	2.11E-02	8.5	1.67
		Gluten index	2.57E-02	7.8	11.7
hap_1A_3	1A (58.4–61.2)	Semolina pigment	2.99E-02	7.0	0.39
hap_1A_4	1A (116.4–123.2)	Pigment loss	5.35E-02	6.7	–0.2
hap_1A_6	1A (149.5–149.5)	Pasta b*	4.36E-02	3.8	3.54
hap_1B_1	1B (0.3–5.5)	Protein content	2.24E-02	9.8	1.78
		Dough extensibility	1.82E-02	9.5	22.26
hap_1B_2	1B (15.7–15.7)	Dough extensibility	2.68E-04	8.3	2.81
		Gluten index	1.80E-03	6.0	2.8
		Dough tenacity	1.23E-02	3.8	9.73
		Deformation energy	2.42E-02	3.1	6.35
hap_1B_3	1B (35.7–38.8)	Semolina pigment	3.51E-02	12.7	1.67
hap_1B_5	1B (86.1–96.2)	Semolina pigment	2.52E-02	15.8	1.75
hap_1B_6	1B (109–109.8)	Semolina pigment	2.70E-02	5.2	0.27
hap_1B_7	1B (115.7–119.1)	Semolina pigment	2.01E-02	13.8	2.39
		Protein content	2.66E-02	12.2	2.25
hap_1B_10	1B (158–164.1)	Protein loss	2.52E-02	13.4	–0.54
hap_2A_1	2A (107.7–108.9)	Dough extensibility	1.50E-02	13.8	28.58
hap_2A_3	2A (181.2–187.7)	Semolina pigment	3.04E-02	15.4	3.17
		Protein loss	3.13E-02	15.3	–0.72
hap_2B_5	2B (126.6–131.9)	Semolina pigment	2.08E-02	8.5	1.67
hap_2B_6	2B (142.2–146.8)	Semolina pigment	4.23E-02	9.9	1.67
hap_2B_7	2B (155.5–158.3)	Gluten index	1.53E-02	7.5	15.95
		Semolina pigment	4.66E-02	6.4	0.72
hap_2B_9	2B (181.6–185.3)	Protein loss	1.44E-03	17.8	–0.32
		Protein content	3.84E-03	15.9	2.32
		Deformation energy	2.26E-02	13.0	54.18
		Semolina pigment	1.56E-02	4.7	0.04
hap_3A_3	3A (54.7–54.7)	Semolina pigment	5.20E-02	5.3	0.03
hap_3A_4	3A (64.2–67.3)	Semolina pigment	4.66E-02	23.5	2.11
hap_3A_5	3A (77.8–90.4)	Semolina pigment	3.53E-02	9.4	2.19
hap_3A_6	3A (124.6–128.7)	Semolina b*	7.94E-03	11.9	1.67
hap_3B_2	3B (144.8–146.4)	Semolina pigment	2.19E-02	5.5	1.86
hap_4A_2	4A (64.1–64.1)	Semolina pigment	4.23E-02	3.6	1.67
hap_4A_3	4A (90.6–90.6)	Dough tenacity	3.28E-03	5.3	8.71
hap_4A_4	4A (118.5–118.5)	Dough extensibility	2.86E-02	2.9	2.75
		Semolina pigment	7.44E-03	8.0	2.86
hap_4A_6	4A (173.3–175.8)	Semolina b*	2.74E-02	6.8	3.64
		Pigment loss	1.36E-03	20.6	–0.81
hap_4B_1	4B (17.7–22.5)	Pasta b*	5.06E-03	17.9	3.98
		Dough extensibility	9.50E-03	16.6	24.32
		Gluten index	4.52E-02	12.1	21.3
		Pigment loss	2.33E-06	16.8	–1.1
hap_4B_2	4B (59.8–60.4)	Pasta b*	3.04E-05	13.3	2.29
		Protein loss	7.72E-04	11.0	–0.53
		Semolina pigment	1.70E-02	6.8	1.28
		Protein content	3.57E-03	6.8	0.93
		Semolina pigment	2.52E-02	9.1	1.67
hap_4B_4	4B (105.5–110.2)	Semolina pigment	1.43E-02	4.8	0.23
hap_5A_1	5A (43.8–43.8)	Semolina pigment	3.86E-02	5.7	0.59
hap_5A_2	5A (50.5–54.9)	Semolina pigment	5.88E-03	15.3	2.86
hap_5A_3	5A (64.2–73.6)	Semolina pigment			

(Continued)

TABLE 4 | Continued

Haplotypes	Position <sup>a</sup>	Traits	P-value	R <sup>2</sup> (%)	Allelic effect <sup>b</sup>
hap_5A_5	5A (140.6–141.4)	Semolina pigment	6.86E-03	11.2	1.86
hap_5B_1	5B (7.2–10.1)	Semolina pigment	2.13E-02	7.5	0.27
		Dough extensibility	4.85E-02	5.9	1.26
hap_5B_4	5B (181.5–181.5)	Semolina pigment	1.17E-02	5.0	0.15
hap_6A_3	6A (39.5–41)	Semolina pigment	2.11E-03	9.7	1.67
		Pasta b*	3.15E-02	5.4	3.54
hap_6B_2	6B (36.2–37.8)	Semolina pigment	1.58E-02	8.0	1.3
hap_6B_3	6B (54.6–57)	Semolina pigment	3.86E-02	9.3	2.51
hap_6B_5	6B (92.3–96)	Dough tenacity	2.03E-02	6.0	3.01
hap_7A_1	7A (21.9–21.9)	Semolina pigment	1.67E-02	4.6	0.36
		Protein loss	4.28E-02	3.5	−0.01
hap_7A_3	7A (55.9–62.5)	Semolina pigment	2.59E-02	21.3	2.86
hap_7A_4	7A (82.4–84.4)	Semolina pigment	2.99E-02	8.8	1.67
hap_7A_5	7A (90.9–90.9)	Semolina pigment	4.87E-02	3.4	1.67
hap_7A_6	7A (102.4–102.4)	Gluten index	1.53E-02	3.6	0.01
		Protein content	1.53E-02	3.5	0.14
		Dough tenacity	4.70E-02	2.4	6.57
		Deformation energy	5.02E-02	2.3	1.06
hap_7A_7	7A (112.2–118)	Semolina pigment	3.60E-02	9.4	1.67
hap_7A_9	7A (154.6–158.8)	Semolina pigment	1.81E-02	12.3	1.67
		Deformation energy	3.78E-02	11.9	58.48
hap_7A_11	7A (193.9–194.6)	Semolina pigment	6.52E-03	5.7	1.6
		Protein content	1.50E-02	5.0	1.27
		Semolina b*	4.23E-02	3.9	1.62
hap_7B_3	7B (120.4–127.4)	Semolina pigment	1.62E-02	11.6	2.51
		Pasta b*	3.48E-02	10.3	3.54
hap_7B_4	7B (132.8–133)	Semolina pigment	1.67E-02	4.6	0.33

<sup>a</sup>Position in centiMorgan on the chromosome. <sup>b</sup>Allelic effect: The allelic effect of haplotypes was estimated as the difference between the mean value of the lines carrying these haplotypes and the mean value of the entire population for each trait. \*Should be read as star (e.g., Semolina b\* is "Semolina b star").

TABLE 5 | Haplotype loci significantly ( $P$ -value  $\leq 0.05$ ) associated with plant height and grain yield as revealed by the Mixed Linear Model with Kinship matrix (MLM-K).

Haplotypes	Position <sup>a</sup>	Trait	P-value	R <sup>2</sup> (%)	Allelic effect <sup>b</sup>
hap_1B_8	1B (128.8–128.8)	Grain yield	3.28E-02	5.2	28.66
hap_2A_1	2A (107.7–108.9)	Plant height	6.07E-03	15.0	−9.71
hap_2B_2	2B (19–19)	Grain yield	4.45E-03	6.5	5.95
hap_2B_7	2B (155.5–158.3)	Grain yield	2.29E-04	13.4	76.52
hap_4A_5	4A (148.3–167.6)	Plant height	1.04E-02	43.9	−3.15
		Grain yield	4.03E-02	38.5	190.03
hap_4A_6	4A (173.3–175.8)	Grain yield	9.49E-04	11.4	5.34
hap_4B_1	4B (17.7–22.5)	Plant height	6.44E-06	30.0	−9.09
		Grain yield	2.25E-02	14.4	65.21
hap_4B_2	4B (59.8–60.4)	Plant height	1.01E-06	17.4	−8.03
hap_5A_1	5A (43.8–43.8)	Plant height	4.74E-02	2.3	−0.29
hap_7A_10	7A (168.6–168.6)	Plant height	1.61E-03	6.0	−2.73
hap_7B_5	7B (147–147)	Plant height	4.90E-02	3.6	−3.48

<sup>a</sup>Position in centiMorgan on the chromosome. <sup>b</sup>Allelic effect: The allelic effect of haplotypes was estimated as the difference between the mean value of the lines carrying these haplotypes and the mean value of the entire population for each trait.

Singh et al., 2017), wheat (Zhang et al., 2007, 2015), eggplant (Portis et al., 2014), cabbage (Lv et al., 2016), and Brassica (Basnet et al., 2015). The existence of such hot regions could offer an appealing opportunity to select

for multiple traits, especially when traits are positively correlated.

Thirty-six (30/84) percent of the haplotype loci under selection harbored markers with a complete reversal of allelic

state. This tendency towards fixation of favorable alleles in our material corroborates the success of the Canadian durum wheat breeding programs over time.

## Genome-Wide Association Studies

Previous studies showed that haplotype-based analysis increases the power of QTL detection (Liu et al., 2008; Hamblin and Jannink, 2011; Gawenda et al., 2015; Contreras-Soto et al., 2017) and explains a larger proportion of the QTL variance (Grapes et al., 2006; Hayes et al., 2007; N'Diaye et al., 2017) as compared to single-marker analysis. Therefore, all of the 407 SNP markers under selection were clustered into 84 LD-based haplotype loci, as reported in our previous manuscript (N'Diaye et al., 2017).

Thirty-six (30/84) percent of the haplotype loci under selection failed to be associated with any of the traits that we investigated. Because these loci showed strong evidence of directional selection, the lack of an association signal could suggest that they might be controlling traits that we do not have data for, and we didn't analyze here. Therefore, these loci could still be of interest for further investigations when more phenotypic data becomes available.

A total of 49 haplotype loci, spanning all 14 chromosomes of the durum wheat genome, were associated with quality traits. In particular, *hap\_4B\_1* located on chromosome 4B explained 20.6, 17.9, 16.6, and 12.1% of the phenotypic variance of pigment loss, pasta b\*, dough extensibility and gluten index, respectively. This result is congruent with many studies that reported major QTL for yellow pigment on chromosome 4B (Mares and Campbell, 2001; Pozniak et al., 2007; Blanco et al., 2011; Roncallo et al., 2012; N'Diaye et al., 2017). The haplotype locus *hap\_3A\_5* explained 23.4% of the phenotypic variation of semolina pigment, contrasting with the minor QTL of pigment color reported by other studies in durum wheat (Reimer et al., 2008; Roncallo et al., 2012). However, major QTL explaining up to 20% of the variation of flour color have been reported on chromosome 3A in hexaploid wheat (Parker et al., 1998; Mares and Campbell, 2001; Crawford and Francki, 2013). The locus *hap\_7A\_3* explained 21.3% of the phenotypic variation of semolina pigment; coincident with the existence of a major QTL of yellow pigment on chromosome 7A in durum wheat (Patil et al., 2008; Zhang and Dubcovsky, 2008; Zhang et al., 2008). These conflicting results clearly demonstrate the complex inheritance of yellow pigment in wheat.

Protein content is a key specification for wheat because of its nutritive value and its impact on many processing properties, such as water absorption and gluten strength (Delcour et al., 2010; Kaur et al., 2015; Kaushik et al., 2015). Six haplotype loci were associated with protein content, spanning different chromosomes, of which chromosome 6B that had been reported to harbor a major QTL for grain protein (Joppa et al., 1997; Olmos et al., 2003; Uauy et al., 2006; Patil et al., 2009; Suprayogi et al., 2009; Conti et al., 2011). Nonetheless, in our study the highest proportion (15.9%) of variance explained came from the locus *hap\_2B\_9*, located on chromosome 2B. Depending on the populations, the QTL explaining the largest variance of protein content in durum wheat was found on different chromosomes, such as 2A (Blanco et al., 2006), 5B (Golabadi et al., 2012), and 5A

(Marcotuli et al., 2017). Similarly, in hexaploid wheat, the number of QTL controlling protein content were reported on different chromosomes (Börner et al., 2002; Sun et al., 2008; Wang et al., 2012; Li et al., 2013; Terasawa et al., 2016; Tiwari et al., 2016; Zou et al., 2017). These results suggest the complex inheritance of grain protein and the strong influence of the underlying population and the environment.

High semolina protein content determines end-use products quality such as texture, appearance and firmness. However, because of its negative correlation with grain yield (Blanco et al., 2006; Würschum et al., 2016), selecting for high grain yield has indirectly resulted in lines with lower protein content (De Vita et al., 2007; Acreche and Slafer, 2009). To simultaneously improve grain yield and protein content in durum wheat, an index selection method has recently been developed (Rapp et al., 2018). Moreover, it has been reported that protein content substantially decreases (up to 25% loss of protein) during milling mainly with regard to the milling methods (Prabhasankar and Haridas Rao, 2001; Ramberg and McAnalley, 2002; Heshe et al., 2016; Oghbaei and Prakash, 2016). Because the concentrations of protein components follow a gradient (higher concentration in the outer layer) along the different parts of the grain (Tosi et al., 2011; He et al., 2013; Li et al., 2016; Savill et al., 2018) and that this gradient is in part determined by genetic factors (He et al., 2013), one might wonder if protein loss during milling could have some genetic basis. To the best of our knowledge, no investigation for a possible genetic basis of protein loss has been carried out. In this study, five haplotype loci were associated with protein loss, explaining 3.5 (*hap\_7A\_1*) to 17.8% (*hap\_2B\_9*) of the phenotypic variance. The haplotype locus *hap\_2B\_9*, explaining the highest proportion (17.8%) of protein loss variance, also explained 15.9% of the variation of protein content. This result is not surprising given the strong positive ( $r = 0.82$ ) correlation between these two traits; the higher protein content the bigger the protein loss. Similarly, Savill et al. (2018) reported that any increase in the gradient of protein concentration may result in an even greater amount of protein being lost during milling. In fact, during milling some endosperm tissues remain adhered to the bran layers and because protein is concentrated in the outer layers of the endosperm, a proportionally greater amount of protein relative to starch is lost during the production of white flour. Therefore, there is a need to select for genotypes having high protein content but low protein loss during milling. The availability of many haplotype loci for both protein content and protein loss offers the opportunity to screen different allele combinations to work this challenge out.

A total of 11 haplotype loci were associated with plant height and/or grain yield, explaining low (2.3%) to high (43.9%) proportion of the phenotypic variance depending on the trait. For grain yield, of the six haplotype loci that were detected, *hap\_4A\_5* located on chromosome 4A, explained 38.5% of the phenotypic variance. Yield is the most important and genetically complex trait in wheat, being controlled by a large number of small effect QTL across all chromosomes (Wu et al., 2012). Many QTL and marker-trait associations for yield and yield-related traits (e.g., spike length, spikelet per spike and 1000 kernel weight) have been reported on all chromosomes of wheat

(Marza et al., 2006; Heidari et al., 2011; Ma et al., 2012; Li et al., 2015; Marcotuli et al., 2017). The major yield QTL were reported on different chromosomes depending on the study. In particular, a comprehensive QTL analysis in a durum wheat population across 16 environments detected 16 QTL of grain yield, including two major QTL on 2B and 3B, explaining 21.5 and 13.8% of the variance, respectively (Maccaferri et al., 2008). Nonetheless, of the many QTL regions affecting yield and yield-related traits, two strong QTL hotspots were identified on chromosomes 2A and 2B in a durum wheat panel of 208 lines (Sukumaran et al., 2018). In contrast, the most important QTL affecting yield in a durum wheat population evaluated in different environments were reported on chromosome 3B and 3A, explaining 20.7 and 18.0 % of the phenotypic variation, respectively (Roncallo et al., 2017). A number of QTL controlling grain yield have also been reported on different chromosomes of common wheat (Quarrie et al., 2006; Zhang et al., 2010; Wang et al., 2012; Simmonds et al., 2014; Li et al., 2015; Tahmasebi et al., 2016; Sehgal et al., 2017). In particular, one and two major yield QTL were reported on chromosomes 3A (Mengistu et al., 2012) and 3B (Bennett et al., 2012), respectively.

Among the seven haplotype loci associated with plant height, *hap\_4A\_5* (chromosome 4A) and *hap\_4B\_1* (chromosome 4B) explained 43.9 and 30.0% of the variance, respectively. The *Rht-B1b* gene was mapped on chromosome 4B. These results are consistent with those of many studies that reported major QTL for plant height on chromosomes group 4 in both durum and hexaploid wheat (Gao et al., 2015; Guo et al., 2017; Iannucci et al., 2017; Shi et al., 2017). Moreover, many QTL and marker-trait associations for plant height have been reported, spanning all chromosomes of wheat (Griffiths et al., 2012; Bentley et al., 2014; Zanke et al., 2014). Plant height is one of the most studied traits in wheat because it determines the general architecture of the plant and has strong effect on lodging stability, harvest index and ultimately grain yield. Because tall plants are more susceptible to lodging (Berry et al., 2003), plant height reduction has been a major target for wheat breeding programs for many years. As a result, major dwarfing genes such as *Rht-D1b*, *Rht-B1b*, *Rht8*, and *Rht12* have been incorporated in new wheat varieties to reduce plant height without reducing grain yield potential (Flintham et al., 1997; Korzun et al., 1998; Ellis et al., 2002; Borojevic and Borojevic, 2005; Kowalski et al., 2016).

The haplotype locus *hap\_4A\_5* was associated with both plant height and grain yield, explaining 43.9 and 30.0% of the phenotypic variation, respectively. The influence of plant height on grain yield is well known and pleiotropic loci controlling these two traits have been reported in durum wheat (Maccaferri et al., 2008) and common wheat (McCartney et al., 2005; Gao et al., 2015; Cabral et al., 2018).

Overall, the presence of numerous pleiotropic haplotype loci, such as *hap\_2B\_9* (protein content, protein loss, deformation energy), *hap\_4A\_5* (plant height, grain yield) and *hap\_4B\_1* (pigment loss, pasta b\*, dough tenacity, gluten index), offers breeders the appealing opportunity for improving multiple traits. Indeed, pinpointing the genes controlling these traits and identifying causal mutations would be greatly resourceful to the wheat community. However, tracking down causative genes is

not trivial and might take several years, especially when dealing with many complex traits. For sake of information, we retrieved the genes and their annotations (**Supplementary Table S4**), using the Chinese spring reference genome (Appels et al., 2018).

Thirty-nine percent of haplotype loci associated with the traits we investigated contained markers with a complete reversal of allelic state, suggesting a tendency of fixation of favorable alleles in our plant material. However, according to the theory of selection limits in finite population such as plant breeding populations, artificial selection is expected to increase the frequency of favorable alleles, with the possibility of other less desirable or selectively neutral alleles being fixed simply by chance (Robertson, 1960). Long-term improvement of a trait influences genes associated with the trait and nearby chromosomal regions via linkage (Hedrick, 2000). Therefore, separating the loci directly controlling the improved trait from those derived from hitchhiking could be challenging (Fu and Somers, 2011). Thus, for each haplotype locus, we only kept the alleles series offering the largest allelic effect on the phenotype.

The time at which the switch in allelic state occurred varied among SNPs within each haplotype locus, probably due to the involvement of different genes in the genetic control of the targeted traits. The selection response of complex traits is the result of simultaneous changes in allele's frequencies across many genetic variants and large effect loci are more likely to be fixed rapidly (Johansson et al., 2010). However, because of the relatively small size of our population (192 lines) and the possible involvement of many genes with small effects, it is challenging to predict how fast the favorable alleles will be fixed. Similarly, many years (1845–2004) of improvement of different traits in Canadian hard red spring wheat introduced significant changes in allele frequencies at different loci across the whole genome (Fu and Somers, 2011).

Our results clearly demonstrate the impact of artificial selection on the dynamics of the durum wheat genome, in terms of allelic changes at selected loci over time. The long-term breeding efforts have impacted different chromosomal regions as many genes were targeted by the selection of complex traits.

## CONCLUSION

Eighty-four LD-based haplotype loci showed strong evidence of selection over 60 years of breeding in Canadian durum wheat germplasm. The distribution pattern of these loci was different between chromosomes, outliers spanned several genomic regions on the entire chromosome (e.g., 1B, 2B, and 7A) or were clustered at only few localized regions on the chromosome (e.g., 2A and 4A). The association analysis detected 54 haplotype loci, of which 39% contained markers with a complete reversal of allelic state. This tendency towards fixation of favorable alleles corroborates the success of the Canadian durum wheat breeding programs over time. All of the loci under selection uncovered by this study could be helpful in the identification of genes related to many traits of interest as functional annotations of the wheat genome become available, and to facilitate marker-assisted selection in durum wheat.

## AUTHOR CONTRIBUTIONS

AN designed the experiment, performed all analyses and wrote the manuscript. JKH contributed to the initial draft and edited the manuscript. KTN and SW edited the manuscript. AKS, FRC, JMC, and YR collected phenotypic data and edited the manuscript. CJP supervised the project, designed the experiment, collected phenotypic data and edited the manuscript.

## FUNDING

This research was conducted as part of the Canadian Triticum Applied Genomics (CTAG<sup>2</sup>) project. We are grateful for the financial support of Genome Canada, Genome Prairie, Western Grains Research Foundation, Saskatchewan Wheat Development Commission, and the Saskatchewan Ministry of Agriculture.

## ACKNOWLEDGMENTS

We acknowledge the technical assistance of Shawn Yates at the Agriculture and Agri-Food Canada Research and Development Centre, Swift Current for curation of the database from which the breeding lines data were drawn. The data from these trials were conducted under the auspices of the Prairie Recommending Committee for Wheat, Rye and Triticale and we acknowledge them for permission to use the data, and the many people who conducted the field trials. We acknowledge the coordinators of these trials, including Dr. D. Leisle, Agriculture and Agri-Food Canada, Winnipeg (retired) from 1961 to 1994, Dr. J. M. Clarke, from 1995 to 2007, Dr. A.K. Singh from 2008 to 2012, and Dr. R.M. DePauw in 2013, all from Agriculture and Agri-Food Canada. We also acknowledge those responsible for end-use quality analysis of these trials from the Grain Research Laboratory of the Canadian Grain Commission. We are also thankful to Krystalee Wiebe and Justin Coulson for support

## REFERENCES

- Acreche, M. M., and Slafer, G. A. (2009). Variation of grain nitrogen content in relation with grain yield in old and modern Spanish wheats grown under a wide range of agronomic conditions in a Mediterranean region. *J. Agric. Sci.* 147, 657–667. doi: 10.1017/S0021859609990190
- Albrechtsen, A., Moltke, I., and Nielsen, R. (2010). Natural selection and the distribution of identity-by-descent in the human genome. *Genetics* 186, 295–308. doi: 10.1534/genetics.110.113977
- Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., and Luikart, G. (2008). LOSITAN: a workbench to detect molecular adaptation based on a F<sub>st</sub> -outlier method. *BMC Bioinformatics* 9:323. doi: 10.1186/1471-2105-9-323
- Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191
- Aremu, C. O. (2011). Genetic diversity: a review for need and measurements for intraspecies crop improvement. *J. Microbiol. Biotech. Res.* 1, 80–85.
- Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., et al. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* 357, 93–97. doi: 10.1126/science.aan0032
- Barreiro, L. B., Laval, G., Quach, H., Patin, E., and Quintana-Murci, L. (2008). Natural selection has driven population differentiation in modern humans. *Nat. Genet.* 40, 340–345. doi: 10.1038/ng.78
- Basnet, R. K., Duwal, A., Tiwari, D. N., Xiao, D., Monakhos, S., Bucher, J., et al. (2015). Quantitative trait locus analysis of seed germination and seedling vigor in *Brassica rapa* reveals QTL hotspots and epistatic interactions. *Front. Plant Sci.* 6:1032. doi: 10.3389/fpls.2015.01032
- Beaumont, M. A., and Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 263, 1619–1626. doi: 10.1098/rspb.1996.0237
- Bennett, D., Reynolds, M., Mullan, D., Izanloo, A., Kuchel, H., Langridge, P., et al. (2012). Detection of two major grain yield QTL in bread wheat (*Triticum aestivum* L.) under heat, drought and high yield potential environments. *Theor. Appl. Genet.* 125, 1473–1485. doi: 10.1007/s00122-012-1927-2
- Bentley, A. R., Scutari, M., Gosman, N., Faure, S., Bedford, F., Howell, P., 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. doi: 10.1007/s00122-014-2403-y

in generation of the molecular data used in this project. We are grateful to Brook Byrns for his help in retrieving the genes annotation information. Special thanks go to François Roewer-Despres for his valuable advices and helpful assistance in Python programming.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.01589/full#supplementary-material>

**FIGURE S1** | Scatter plots showing the thresholds to declare loci being under selection, 0.2 (Arlequin, Lositan) and 0.15 (BayeScan).

**FIGURE S2** | Q-Q (quantile-quantile) plot profiles of the association analyses using: **(A)** the Mixed Linear Model with the Kinship matrix alone (MLM-K), and **(B)** with both Q matrix from the discriminant analysis of principal component and Kinship (MLM-QK).

**FIGURE S3** | Distribution of allele's frequency for the 95 SNP markers having a complete reversal of allelic state. Brown color is the overlap between blue bar (A allele) and orange bar (B allele).

**FIGURE S4** | Histograms of the phenotypic traits evaluated in the Canadian durum wheat germplasm.

**FIGURE S5** | Example of changes in allelic states of markers at different time periods for the haplotype locus *hap\_1B\_5*, associated with only semolina pigment.

**FIGURE S6** | Example of changes in allelic states of markers at different time periods for the haplotype locus *hap\_2A\_1*, associated with plant height and dough extensibility.

**TABLE S1** | Lines pedigree and year of introduction into the Durum wheat Cooperative Test.

**TABLE S2** | Haplotype loci under selection in the Durum wheat Cooperative Test.

**TABLE S3** | Haplotype loci associated with traits and containing SNPs with complete reversal of allelic state.

**TABLE S4** | List of genes harboring outliers associated with phenotypic traits in the durum wheat breeding panel.

- Berry, P. M., Sterling, M., Baker, C. J., Spink, J., and Sparkes, D. L. (2003). A calibrated model of wheat lodging compared with field measurements. *Agric. For. Meteorol.* 119, 167–180. doi: 10.1016/S0168-1923(03)00139-4
- Bigham, A., Bauchet, M., Pinto, D., Mao, X., Akey, J. M., Mei, R., et al. (2010). Identifying Signatures of natural selection in tibetan and andean populations using dense genome scan data. *PLoS Genet.* 6:e1001116. doi: 10.1371/journal.pgen.1001116
- Blanco, A., Colasuonno, P., Gadaleta, A., Mangini, G., Schiavulli, A., Simeone, R., et al. (2011). Quantitative trait loci for yellow pigment concentration and individual carotenoid compounds in durum wheat. *J. Cereal Sci.* 54, 255–264. doi: 10.1016/j.jcs.2011.07.002
- Blanco, A., Simeone, R., and Gadaleta, A. (2006). Detection of QTLs for grain protein content in durum wheat. *Theor. Appl. Genet.* 112, 1195–1204. doi: 10.1007/s00122-006-0221-6
- Börner, A., Schumann, E., Fürste, A., Cöster, H., Leithold, B., Röder, M., et al. (2002). Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 105, 921–936. doi: 10.1007/s00122-002-0994-1
- Borojevic, K., and Borojevic, K. (2005). The transfer and history of “Reduced Height Genes” (Rht) in wheat from Japan to Europe. *J. Hered.* 96, 455–459. doi: 10.1093/jhered/esi060
- Bowman, D. T. (2003). Genetic uniformity of the U.S. upland cotton crop since the introduction of transgenic cottons. *Crop Sci.* 43, 515–518. doi: 10.2135/cropsci2003.0515
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/a:1010933404324
- Cabral, A. L., Jordan, M. C., Larson, G., Somers, D. J., Humphreys, D. G., and McCartney, C. A. (2018). Relationship between QTL for grain shape, grain weight, test weight, milling yield, and plant height in the spring wheat cross RL4452/‘AC Domain’. *PLoS One* 13:e0190681. doi: 10.1371/journal.pone.0190681
- Cagliani, R., Riva, S., Fumagalli, M., Biasin, M., Caputo, S. L., Mazzotta, F., et al. (2011). A positively selected APOBEC3H haplotype is associated with natural resistance to HIV-1 infection. *Evolution* 65, 3311–3322. doi: 10.1111/j.1558-5646.2011.01368.x
- Cai, H., and Morishima, H. (2002). QTL clusters reflect character associations in wild and cultivated rice. *Theor. Appl. Genet.* 104, 1217–1228. doi: 10.1007/s00122-001-0819-7
- Canada, S. (2018). *Estimated Areas, Yield, Production, Average Farm Price and Total Farm Value of Principal Field Crops, in Metric and Imperial Units, Annual, CANSIM (Database)*. Ottawa: Statistics Canada.
- Cappa, E. P., El-Kassaby, Y. A., Garcia, M. N., Acuna, C., Borralho, N. M. G., Grattapaglia, D., et al. (2013). Impacts of population structure and analytical models in genome-wide association studies of complex traits in forest trees: a case study in *Eucalyptus globulus*. *PLoS One* 8:e81267. doi: 10.1371/journal.pone.0081267
- Clarke, F. R., Clarke, J. M., Ames, N. A., Knox, R. E., and Ross, R. J. (2010). Gluten index compared with SDS-sedimentation volume for early generation selection for gluten strength in durum wheat. *Can. J. Plant Sci.* 90, 1–11. doi: 10.4141/CJPS09035
- Clarke, J. M., Clarke, F. R., and Pozniak, C. J. (2010). Forty-six years of genetic improvement in Canadian durum wheat cultivars. *Can. J. Plant Sci.* 90, 791–801. doi: 10.4141/cjps10091
- Clarke, F. R., Clarke, J. M., Pozniak, C. J., Knox, R. E., and McCaig, T. N. (2009). Protein concentration inheritance and selection in durum wheat. *Can. J. Plant Sci.* 89, 601–612. doi: 10.4141/CJPS08191
- Conti, V., Roncallo, P. F., Beaufort, V., Cervigni, G. L., Miranda, R., Jensen, C. A., et al. (2011). Mapping of main and epistatic effect QTLs associated to grain protein and gluten strength using a RIL population of durum wheat. *J. Appl. Genet.* 52, 287–298. doi: 10.1007/s13353-011-0045-1
- Contreras-Soto, R. I., Mora, F., de Oliveira, M. A. R., Higashi, W., Scapim, C. A., and Schuster, I. (2017). A genome-wide association study for agronomic traits in soybean using SNP markers and SNP-based haplotype analysis. *PLoS One* 12:e0171105. doi: 10.1371/journal.pone.0171105
- Crawford, A. C., and Francki, M. G. (2013). *Lycopene-ε-cyclase (e-LCY3A)* is functionally associated with quantitative trait loci for flour b\* colour on chromosome 3A in wheat (*Triticum aestivum* L.). *Mol. Breed.* 31, 737–741. doi: 10.1007/s11032-012-9812-x
- Crowell, S., Korniliev, P., Falcão, A., Ismail, A., Gregorio, G., Mezey, J., et al. (2016). Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nat. Commun.* 7:10527. doi: 10.1038/ncomms10527
- De Vita, P., Nicosia, O. L. D., Nigro, F., Platani, C., Riefolo, C., Di Fonzo, N., et al. (2007). Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. *Eur. J. Agron.* 26, 39–53. doi: 10.1016/j.eja.2006.08.009
- Delcour, J. A., Bruneel, C., Derde, L. J., Gomand, S. V., Pareyt, B., Putseys, J. A., et al. (2010). Fate of starch in food processing: from raw materials to final food products. *Annu. Rev. Food Sci. Technol.* 1, 87–111. doi: 10.1146/annurev.food.102308.124211
- Dexter, J. E. (2008). “The history of durum wheat breeding in Canada and summaries of recent research at the Canadian Grain Commission on factors associated with durum wheat processing,” in *Proceedings of the Bosphorus 2008 ICC (International Cereal Congress) International Conference*, Istanbul.
- Dexter, J. E., and Edwards, N. M. (1998). The implications of frequently encountered grading factors on the processing quality of durum wheat. *Assoc. Oper. Millers Bull.* 7165–7171.
- Dreisigacker, S., Shewaryga, H., Crossa, J., Arief, V. N., DeLacy, I. H., Singh, R. P., et al. (2012). Genetic structures of the CIMMYT international yield trial targeted to irrigated environments. *Mol. Breed.* 29, 529–541. doi: 10.1007/s11032-011-9569-7
- Duvick, D. N. (1984). Genetic diversity in major farm crops on the farm and in reserve. *Econ. Bot.* 38, 161–178. doi: 10.1007/bf02858829
- Ellis, M., Spielmeier, W., Gale, K., Rebetzke, G., and Richards, R. (2002). “Perfect” markers for the Rht-B1b and Rht-D1b dwarfing genes in wheat. *Theor. Appl. Genet.* 105, 1038–1042. doi: 10.1007/s00122-002-1048-4
- Eveno, E., Collada, C., Guevara, M. A., Léger, V., Soto, A., Díaz, L., et al. (2008). Contrasting patterns of selection at *Pinus pinaster* Ait. Drought stress candidate genes as revealed by genetic differentiation analyses. *Mol. Biol. Evol.* 25, 417–437. doi: 10.1093/molbev/msm272
- Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x
- Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M., and Servin, B. (2013). Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics* 193, 929–941. doi: 10.1534/genetics.112.147231
- Feillet, P., and Dexter, J. E. (1996). “Quality requirements of durum wheat for semolina milling and pasta production,” in *Monograph on Pasta and Noodle Technology*, eds J. E. Kruger, R. R. Matsuo, and J. W. Dick (St. Paul, MN: AACC Inc.), 95–131.
- Fiedler, J. D., Salsman, E., Liu, Y., Michalak de Jimenez, M., Hegstad, J. B., Chen, B., et al. (2017). Genome-wide association and prediction of grain and semolina quality traits in durum wheat breeding populations. *Plant Genome* 10, 1–12. doi: 10.3835/plantgenome2017.05.0038
- Flintham, J. E., Börner, A., Worland, A. J., and Gale, M. D. (1997). Optimizing wheat grain yield: effects of *Rht* (gibberellin-insensitive) dwarfing genes. *J. Agric. Sci.* 128, 11–25. doi: 10.1017/S002185969603942
- Foll, M., and Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. *Genetics* 180, 977–993. doi: 10.1534/genetics.108.092221
- Fu, Y.-B. (2007). Impact of plant breeding on genetic diversity of agricultural crops: searching for molecular evidence. *Plant Genet. Resour.* 4, 71–78. doi: 10.1079/PGR2006116

- Fu, Y.-B. (2015). Understanding crop genetic diversity under modern plant breeding. *Theor. Appl. Genet.* 128, 2131–2142. doi: 10.1007/s00122-015-2585-y
- Fu, Y.-B., and Dong, Y.-B. (2015). “Genetic erosion under modern plant breeding: case studies in Canadian crop gene pools,” in *Genetic Diversity and Erosion in Plants: Indicators and Prevention*, eds M. R. Ahuja and S. M. Jain (Cham: Springer International Publishing), 89–104.
- Fu, Y.-B., Peterson, G. W., Richards, K. W., Somers, D., DePauw, R. M., and Clarke, J. M. (2005). Allelic reduction and genetic shift in the Canadian hard red spring wheat germplasm released from 1845 to 2004. *Theor. Appl. Genet.* 110, 1505–1516. doi: 10.1007/s00122-005-1988-6
- Fu, Y.-B., and Somers, D. J. (2011). Allelic changes in bread wheat cultivars were associated with long-term wheat trait improvements. *Euphytica* 179, 209–225. doi: 10.1007/s10681-010-0235-7
- Fumagalli, M., Cagliani, R., Riva, S., Pozzoli, U., Biasin, M., Piacentini, L., et al. (2010). Population genetics of IFIH1: ancient population structure, local selection, and implications for susceptibility to type 1 diabetes. *Mol. Biol. Evol.* 27, 2555–2566. doi: 10.1093/molbev/msq141
- Gao, F., Wen, W., Liu, J., Rasheed, A., Yin, G., Xia, X., et al. (2015). Genome-wide linkage mapping of QTL for yield components, plant height and yield-related physiological traits in the Chinese wheat cross Zhou 8425B/Chinese spring. *Front. Plant Sci.* 6:1099. doi: 10.3389/fpls.2015.01099
- Gao, L., Zhao, G., Huang, D., and Jia, J. (2017). Candidate loci involved in domestication and improvement detected by a published 90K wheat SNP array. *Sci. Rep.* 7:44530. doi: 10.1038/srep44530
- Gaunt, T., Rodriguez, S., Zapata, C., and Day, I. (2006). MIDAS: software for analysis and visualisation of interallelic disequilibrium between multiallelic markers. *BMC Bioinformatics* 7:227. doi: 10.1186/1471-2105-7-227
- Gawenda, I., Thorwarth, P., Günther, T., Ordon, F., and Schmid, K. J. (2015). Genome-wide association studies in elite varieties of German winter barley using single-marker and haplotype-based methods. *Plant Breed.* 134, 28–39. doi: 10.1111/pbr.12237
- Gepts, P. (2006). Plant genetic resources conservation and utilization. *Crop Sci.* 46, 2278–2292. doi: 10.2135/cropsci2006.03.0169gas
- Gepts, P., and Hancock, J. (2006). The future of plant breeding. *Crop Sci.* 46, 1630–1634. doi: 10.2135/cropsci2005-12-0497op
- Golabadi, M., Arzani, A., and Maibody, S. M. (2012). Identification of microsatellite markers associated with grain protein content in durum wheat grown under drought stress at terminal growth stages. *Cereal Res. Commun.* 40, 215–224. doi: 10.1556/crc.40.2012.2.6
- Grapes, L., Firat, M. Z., Dekkers, J. C. M., Rothschild, M. F., and Fernando, R. L. (2006). Optimal haplotype structure for linkage disequilibrium based fine mapping of quantitative trait loci using identity by descent. *Genetics* 172, 1955–1965. doi: 10.1534/genetics.105.048686
- Griffiths, S., Simmonds, J., Leverington, M., Wang, Y., Fish, L., Sayers, L., et al. (2012). Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm. *Mol. Breed.* 29, 159–171. doi: 10.1007/s11032-010-9534-x
- Guan, D., Luo, N., Tan, X., Zhao, Z., Huang, Y., Na, R., et al. (2016). Scanning of selection signature provides a glimpse into important economic traits in goats (*Capra hircus*). *Sci. Rep.* 6:36372. doi: 10.1038/srep36372
- Guo, Y., Du, Z., Chen, J., and Zhang, Z. (2017). QTL mapping of wheat plant architectural characteristics and their genetic relationship with seven QTLs conferring resistance to sheath blight. *PLoS One* 12:e0174939. doi: 10.1371/journal.pone.0174939
- Haile, J. K., N'Diaye, A., Clarke, F., Clarke, J., Knox, R., Rutkoski, J., et al. (2018). Genomic selection for grain yield and quality traits in durum wheat. *Mol. Breed.* 38:75. doi: 10.1007/s11032-018-0818-x
- Hallauer, A. R., and Darrach, L. L. (1985). Compendium of recurrent selection methods and their application. *Crit. Rev. Plant Sci.* 3, 1–33. doi: 10.1080/07352688509382202
- Hamblin, M. T., and Jannink, J. L. (2011). Factors affecting the power of haplotype markers in association studies. *Plant Genome* 4, 145–153. doi: 10.3835/plantgenome2011.03.0008
- Hayes, B. J., Chamberlain, A. J., McPartlan, H., Macleod, I., Sethuraman, L., and Goddard, M. E. (2007). Accuracy of marker-assisted selection with single markers and marker haplotypes in cattle. *Genet. Res.* 89, 215–220. doi: 10.1017/S0016672307008865
- He, J., Penson, S., Powers, S. J., Hawes, C., Shewry, P. R., and Tosi, P. (2013). Spatial patterns of gluten protein and polymer distribution in wheat grain. *J. Agric. Food Chem.* 61, 6207–6215. doi: 10.1021/jf401623d
- Hedrick, P. W. (2000). *Genetics of Populations*, 2nd Edn. Sudbury, MA: Jones and Bartlett Publishers.
- Heidari, B., Sayed-Tabatabaei, B. E., Saeidi, G., Kearsley, M., and Suenaga, K. (2011). Mapping QTL for grain yield, yield components, and spike features in a doubled haploid population of bread wheat. *Genome* 54, 517–527. doi: 10.1139/g11-017
- Henkrar, F., El-Haddoury, J., Ouabbou, H., Nsarellah, N., Iraqi, D., Bendaou, N., et al. (2016). Genetic diversity reduction in improved durum wheat cultivars of Morocco as revealed by microsatellite markers. *Sci. Agric.* 73, 134–141. doi: 10.1590/0103-9016-2015-0054
- Heshe, G. G., Haki, G. D., Woldegiorgis, A. Z., and Gemed, H. F. (2016). Effect of conventional milling on the nutritional value and antioxidant capacity of wheat types common in Ethiopia and a recovery attempt with bran supplementation in bread. *Food Sci. Nutr.* 4, 534–543. doi: 10.1002/fsn3.315
- Hu, X., Ren, J., Ren, X., Huang, S., Sabiel, S. A. I., Luo, M., et al. (2015). Association of agronomic traits with SNP markers in durum wheat (*Triticum turgidum* L. *durum* (Desf.)). *PLoS One* 10:e0130854. doi: 10.1371/journal.pone.0130854
- Huang, X.-Q., Wolf, M., Ganai, M. W., Orford, S., Koebner, R. M. D., and Röder, M. S. (2007). Did modern plant breeding lead to genetic erosion in European winter wheat varieties? *Crop Sci.* 47, 343–349. doi: 10.2135/cropsci2006.04.0261
- Iannucci, A., Marone, D., Russo, M. A., De Vita, P., Miullo, V., Ferragonio, P., et al. (2017). Mapping QTL for root and shoot morphological traits in a durum wheat x *T. dicoccum* segregating population at seedling stage. *Int. J. Genomics* 2017:6876393. doi: 10.1155/2017/6876393
- Jiao, Y., Zhao, H., Ren, L., Song, W., Zeng, B., Guo, J., et al. (2012). Genome-wide genetic changes during modern breeding of maize. *Nat. Genet.* 44, 812–815. doi: 10.1038/ng.2312
- Johansson, A. M., Pettersson, M. E., Siegel, P. B., and Carlborg, Ö. (2010). Genome-wide effects of long-term divergent selection. *PLoS Genet.* 6:e1001188. doi: 10.1371/journal.pgen.1001188
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Jombart, T., and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27, 3070–3071. doi: 10.1093/bioinformatics/btr521
- Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94. doi: 10.1186/1471-2156-11-94
- Joppa, L. R., Du, C., Hart, G. E., and Hareland, G. A. (1997). Mapping gene(s) for grain protein in tetraploid wheat (*Triticum turgidum* L.) using a population of recombinant inbred chromosome lines. *Crop Sci.* 37, 1586–1589. doi: 10.2135/cropsci1997.0011183X003700050030x
- Kabbaj, H., Sall, A. T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., et al. (2017). Genetic diversity within a global panel of durum wheat (*Triticum durum*) landraces and modern germplasm reveals the history of alleles exchange. *Front. Plant Sci.* 8:1277. doi: 10.3389/fpls.2017.01277
- Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., et al. (2008). Efficient control of population structure in model organism association mapping. *Genetics* 178, 1709–1723. doi: 10.1534/genetics.107.080101
- Kaur, A., Singh, N., Kaur, S., Katyaj, M., Viridi, A. S., Kaur, D., et al. (2015). Relationship of various flour properties with noodle making characteristics among durum wheat varieties. *Food Chem.* 188, 517–526. doi: 10.1016/j.foodchem.2015.05.009
- Kaushik, R., Kumar, N., Sihag, M. K., and Ray, A. (2015). Isolation, characterization of wheat gluten and its regeneration properties. *J. Food Sci. Technol.* 52, 5930–5937. doi: 10.1007/s13197-014-1690-2
- Khan, M. A., Olsen, K. M., Sovero, V., Kushad, M. M., and Korban, S. S. (2014). Fruit quality traits have played critical roles in domestication of the apple. *Plant Genome* 7, 1–8. doi: 10.3835/plantgenome2014.04.0018
- Kidane, Y. G., Mancini, C., Mengistu, D. K., Frascaroli, E., Fadda, C., Pè, M. E., et al. (2017). Genome wide association study to identify the genetic base of



- smallholder farmer preferences of durum wheat traits. *Front. Plant Sci.* 8:1230. doi: 10.3389/fpls.2017.01230
- Kirk, H., and Freeland, J. R. (2011). Applications and implications of neutral versus non-neutral markers in molecular ecology. *Int. J. Mol. Sci.* 12, 3966–3988. doi: 10.3390/ijms12063966
- Konijnendijk, N., Shikano, T., Daneels, D., Volckaert, F. A. M., and Raeymaekers, J. A. M. (2015). Signatures of selection in the three-spined stickleback along a small-scale brackish water – freshwater transition zone. *Ecol. Evol.* 5, 4174–4186. doi: 10.1002/ece3.1671
- Korzun, V., Röder, M. S., Ganal, M. W., Worland, A. J., and Law, C. N. (1998). Genetic analysis of the dwarfing gene (Rht8) in wheat. Part I. Molecular mapping of Rht8 on the short arm of chromosome 2D of bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 96, 1104–1109. doi: 10.1007/s001220050845
- Kowalski, A. M., Gooding, M., Ferrante, A., Slafer, G. A., Orford, S., Gasperini, D., et al. (2016). Agronomic assessment of the wheat semi-dwarfing gene Rht8 in contrasting nitrogen treatments and water regimes. *Field Crops Res.* 191, 150–160. doi: 10.1016/j.fcr.2016.02.026
- Lao, O., de Grujter, J. M., van Duijn, K., Navarro, A., and Kayser, M. (2007). Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. *Ann. Hum. Genet.* 71, 354–369. doi: 10.1111/j.1469-1809.2006.00341.x
- Li, H. M., Tang, Z. X., Zhang, H. Q., Yan, B. J., and Ren, Z. L. (2013). Major quality trait analysis and QTL detection in hexaploid wheat in humid rain-fed agriculture. *Genet. Mol. Res.* 12, 1740–1751. doi: 10.4238/2013.May.21.5
- Li, X., Xia, X., Xiao, Y., He, Z., Wang, D., Trethowan, R., et al. (2015). QTL mapping for plant height and yield components in common wheat under water-limited and full irrigation environments. *Crop Pasture Sci.* 66, 660–670. doi: 10.1071/CP14236
- Li, X., Zhou, L., Liu, F., Zhou, Q., Cai, J., Wang, X., et al. (2016). Variations in protein concentration and nitrogen sources in different positions of grain in wheat. *Front. Plant Sci.* 7:942. doi: 10.3389/fpls.2016.00942
- Liao, X., Peng, F., Forni, S., McLaren, D., Plastow, G., and Stothard, P. (2013). Whole genome sequencing of Gir cattle for identifying polymorphisms and loci under selection. *Genome* 56, 592–598. doi: 10.1139/gen-2013-0082
- Liaw, A., and Wiener, M. (2002). Classification and regression by random-forest. *R News* 2, 18–22.
- Lin, Y., Chen, Y., Yi, C., Fong, J. J., Kim, W., Rius, M., et al. (2017). Genetic signatures of natural selection in a model invasive ascidian. *Sci. Rep.* 7:44080. doi: 10.1038/srep44080
- Liu, N., Zhang, K., and Zhao, H. (2008). Haplotype-association analysis. *Adv. Genet.* 60, 335–405. doi: 10.1016/S0065-2660(07)00414-2
- Liu, Z., Ji, Z., Wang, G., Chao, T., Hou, L., and Wang, J. (2016). Genome-wide analysis reveals signatures of selection for important traits in domestic sheep from different ecoregions. *BMC Genomics* 17:863. doi: 10.1186/s12864-016-3212-2
- López, M. E., Neira, R., and Yáñez, J. M. (2015). Applications in the search for genomic selection signatures in fish. *Front. Genet.* 5:458. doi: 10.3389/fgene.2014.00458
- Lv, H., Wang, Q., Liu, X., Han, F., Fang, Z., Yang, L., et al. (2016). Whole-genome mapping reveals novel QTL clusters associated with main agronomic traits of cabbage (*Brassica oleracea* var. *capitata* L.). *Front. Plant Sci.* 7:989. doi: 10.3389/fpls.2016.00989
- Ma, J., Zhang, C.-Y., Yan, G.-J., and Liu, C.-J. (2012). Identification of QTLs conferring agronomic and quality traits in hexaploid wheat. *J. Integr. Agric.* 11, 1399–1408. doi: 10.1016/S2095-3119(12)60139-7
- Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., et al. (2014). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 13, 648–663. doi: 10.1111/pbi.12288
- Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L. A., Salem, M. B., Bort, J., et al. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics* 178, 489–511. doi: 10.1534/genetics.107.077297
- Maccaferri, M., Sanguineti, M. C., Demontis, A., El-Ahmed, A., Garcia del Moral, L., Maalouf, F., et al. (2011). Association mapping in durum wheat grown across a broad range of water regimes. *J. Exp. Bot.* 62, 409–438. doi: 10.1093/jxb/erq287
- Mäkinen, H. S., Cano, J. M., and Merilä, J. (2008). Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. *Mol. Ecol.* 17, 3565–3582. doi: 10.1111/j.1365-294X.2008.03714.x
- Marchini, J., Cardon, L. R., Phillips, M. S., and Donnelly, P. (2004). The effects of human population structure on large genetic association studies. *Nat. Genet.* 36, 512–517. doi: 10.1038/ng1337
- Marcotuli, I., Gadaleta, A., Mangini, G., Signorile, A., Zacheo, S., Blanco, A., et al. (2017). Development of a high-density SNP-based linkage map and detection of QTL for  $\beta$ -glucans, protein content, grain yield per spike and heading time in durum wheat. *Int. J. Mol. Sci.* 18:E1329. doi: 10.3390/ijms18061329
- Mares, D. J., and Campbell, A. W. (2001). Mapping components of flour and noodle colour in Australian wheat. *Aust. J. Agric. Res.* 52, 1297–1309. doi: 10.1071/AR01048
- Marza, F., Bai, G.-H., Carver, B. F., and Zhou, W.-C. (2006). Quantitative trait loci for yield and related traits in the wheat population Ning7840  $\times$  Clark. *Theor. Appl. Genet.* 112, 688–698. doi: 10.1007/s00122-005-0172-3
- McCallum, B. D., and DePauw, R. M. (2008). A review of wheat cultivars grown in the Canadian prairies. *Can. J. Plant Sci.* 88, 649–677. doi: 10.4141/CJPS07159
- McCartney, C. A., Somers, D. J., Humphreys, D. G., Lukow, O., Ames, N., Noll, J., et al. (2005). Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452  $\times$  'AC Domain'. *Genome* 48, 870–883. doi: 10.1139/g05-055
- Mengistu, D. K., Kidane, Y. G., Catellani, M., Frascaroli, E., Fadda, C., Pè, M. E., et al. (2016). High-density molecular characterization and association mapping in Ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. *Plant Biotechnol. J.* 14, 1800–1812. doi: 10.1111/pbi.12538
- Mengistu, N., Baenziger, P. S., Eskridge, K. M., Dweikat, I., Wegulo, S. N., Gill, K. S., et al. (2012). Validation of QTL for grain yield-related traits on wheat chromosome 3a using recombinant inbred chromosome lines. *Crop Sci.* 52, 1622–1632. doi: 10.2135/cropsci2011.12.0677
- Mezmouk, S., Dubreuil, P., Bosio, M., Decousset, L., Charcosset, A., Praud, S., et al. (2011). Effect of population structure corrections on the results of association mapping tests in complex maize diversity panels. *Theor. Appl. Genet.* 122, 1149–1160. doi: 10.1007/s00122-010-1519-y
- Morris, R. W., and Kaplan, N. L. (2002). On the advantage of haplotype analysis in the presence of multiple disease susceptibility alleles. *Genet. Epidemiol.* 23, 221–233. doi: 10.1002/gepi.10200
- Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z. W., Costich, D. E., et al. (2009). Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21, 2194–2202. doi: 10.1105/TPC.109.068437
- Narum, S. R., and Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Mol. Ecol. Resour.* 11, 184–194. doi: 10.1111/j.1755-0998.2011.02987.x
- N'Diaye, A., Haile, J. K., Cory, A. T., Clarke, F. R., Clarke, J. M., Knox, R. E., et al. (2017). Single marker and haplotype-based association analysis of semolina and pasta colour in elite durum wheat breeding lines using a high-density consensus map. *PLoS One* 12:e0170941. doi: 10.1371/journal.pone.0170941
- Nei, M., and Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U.S.A.* 76, 5269–5273. doi: 10.1073/pnas.76.10.5269
- Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C., and Clark, A. G. (2007). Recent and ongoing selection in the human genome. *Nat. Rev. Genet.* 8, 857–868. doi: 10.1038/nrg2187
- Oghbaei, M., and Prakash, J. (2016). Effect of primary processing of cereals and legumes on its nutritional quality: a comprehensive review. *Cogent Food Agric.* 2:1136015. doi: 10.1080/23311932.2015.1136015
- Olmos, S., Distelfeld, A., Chicaiza, O., Schlatter, A. R., Fahima, T., Echenique, V., et al. (2003). Precise mapping of a locus affecting grain protein content in durum wheat. *Theor. Appl. Genet.* 107, 1243–1251. doi: 10.1007/s00122-003-1377-y
- Parker, G. D., Chalmers, K. J., Rathjen, A. J., and Langridge, P. (1998). Mapping loci associated with flour colour in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 97, 238–245. doi: 10.1007/s001220050891

- Patil, R., Oak, M., Tamhankar, S., Sourdille, P., and Rao, V. (2008). Mapping and validation of a major QTL for yellow pigment content on 7AL in durum wheat (*Triticum turgidum* L. sp. *durum*). *Mol. Breed.* 21, 485–496. doi: 10.1007/s11032-007-9147-1
- Patil, R. M., Oak, M. D., Tamhankar, S. A., and Rao, V. S. (2009). Molecular mapping of QTLs for gluten strength as measured by sedimentation volume and mixograph in durum wheat (*Triticum turgidum* L. sp. *durum*). *J. Cereal Sci.* 49, 378–386. doi: 10.1016/j.jcs.2009.01.001
- Perez-Figueroa, A., Garcia-Pereira, M. J., Saura, M., Rolan-Alvarez, E., and Caballero, A. (2010). Comparing three different methods to detect selective loci using dominant markers. *J. Evol. Biol.* 23, 2267–2276. doi: 10.1111/j.1420-9101.2010.02093.x
- Portis, E., Barchi, L., Toppino, L., Lanteri, S., Acciarri, N., Felicioni, N., et al. (2014). QTL mapping in eggplant reveals clusters of yield-related loci and orthology with the tomato genome. *PLoS One* 9:e89499. doi: 10.1371/journal.pone.0089499
- Pozniak, C., Clarke, J., and Clarke, F. (2012). Potential for detection of marker–trait associations in durum wheat using unbalanced, historical phenotypic datasets. *Mol. Breed.* 30, 1537–1550. doi: 10.1007/s11032-012-9737-4
- Pozniak, C. J., Knox, R. E., Clarke, F. R., and Clarke, J. M. (2007). Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. *Theor. Appl. Genet.* 114, 525–537. doi: 10.1007/s00122-006-0453-5
- Prabhasankar, P., and Haridas Rao, P. (2001). Effect of different milling methods on chemical composition of whole wheat flour. *Eur. Food Res. Technol.* 213, 465–469. doi: 10.1007/s002170100407
- Quarrie, S. A., Pekic Quarrie, S., Radosevic, R., Rancic, D., Kaminska, A., Barnes, J. D., et al. (2006). Dissecting a wheat QTL for yield present in a range of environments: from the QTL to candidate genes. *J. Exp. Bot.* 57, 2627–2637. doi: 10.1093/jxb/erl026
- R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Ramberg, J., and McAnalley, B. (2002). From the farm to the kitchen table: a review of the nutrient losses in foods. *Glycosci. Nutr.* 3, 1–12.
- Randhawa, H. S., Asif, M., Pozniak, C., Clarke, J. M., Graf, R. J., Fox, S. L., et al. (2013). Application of molecular markers to wheat breeding in Canada. *Plant Breed.* 132, 458–471. doi: 10.1111/pbr.12057
- Rapp, M., Lein, V., Lacoudre, F., Lafferty, J., Müller, E., Vida, G., et al. (2018). Simultaneous improvement of grain yield and protein content in durum wheat by different phenotypic indices and genomic selection. *Theor. Appl. Genet.* 131, 1315–1329. doi: 10.1007/s00122-018-3080-z
- Rauf, S., da Silva, J. T., Khan, A. A., and Naveed, A. (2010). Consequences of plant breeding on genetic diversity. *Int. J. Plant Breed.* 41, 1–21.
- Rees, J. L., and Harding, R. M. (2012). Understanding the evolution of human pigmentation: recent contributions from population genetics. *J. Investig. Dermatol.* 132(3 Part 2), 846–853. doi: 10.1038/jid.2011.358
- Reif, J. C., Zhang, P., Dreisigacker, S., Warburton, M. L., van Ginkel, M., Hoisington, D., et al. (2005). Wheat genetic diversity trends during domestication and breeding. *Theor. Appl. Genet.* 110, 859–864. doi: 10.1007/s00122-004-1881-8
- Reimer, S., Pozniak, C. J., Clarke, F. R., Clarke, J. M., Somers, D. J., Knox, R. E., et al. (2008). Association mapping of yellow pigment in an elite collection of durum wheat cultivars and breeding lines. *Genome* 51, 1016–1025. doi: 10.1139/g08-083
- Ren, J., Chen, L., Jin, X., Zhang, M., You, F. M., Wang, J., et al. (2017). Solar radiation-associated adaptive SNP genetic differentiation in wild emmer wheat, *Triticum dicoccoides*. *Front. Plant Sci.* 8:258. doi: 10.3389/fpls.2017.00258
- Robertson, A. (1960). A theory of limits in artificial selection. *Proc. R. Soc. Lond. B* 153, 234–249. doi: 10.1098/rspb.1960.0099
- Roncallo, P., Cervigni, G., Jensen, C., Miranda, R. N., Carrera, A., Helguera, M., et al. (2012). QTL analysis of main and epistatic effects for flour color traits in durum wheat. *Euphytica* 185, 77–92. doi: 10.1007/s10681-012-0628-x
- Roncallo, P. F., Akkiraju, P. C., Cervigni, G. L., and Echenique, V. C. (2017). QTL mapping and analysis of epistatic interactions for grain yield and yield-related traits in *Triticum turgidum* L. var. *durum*. *Euphytica* 213:277. doi: 10.1007/s10681-017-2058-2
- Savill, G. P., Michalski, A., Powers, S. J., Wan, Y., Tosi, P., Buchner, P., et al. (2018). Temperature and nitrogen supply interact to determine protein distribution gradients in the wheat grain endosperm. *J. Exp. Bot.* 69, 3117–3126. doi: 10.1093/jxb/ery127
- Sehgal, D., Autrique, E., Singh, R., Ellis, M., Singh, S., and Dreisigacker, S. (2017). Identification of genomic regions for grain yield and yield stability and their epistatic interactions. *Sci. Rep.* 7:41578. doi: 10.1038/srep41578
- Shi, S., Azam, F. I., Li, H., Chang, X., Li, B., and Jing, R. (2017). Mapping QTL for stay-green and agronomic traits in wheat under diverse water regimes. *Euphytica* 213:246. doi: 10.1007/s10681-017-2002-5
- Simmonds, J., Scott, P., Leverington-Waite, M., Turner, A. S., Brinton, J., Korzun, V., et al. (2014). Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 14:191. doi: 10.1186/s12870-014-0191-9
- Simonson, T. S., Yang, Y., Huff, C. D., Yun, H., Qin, G., Witherspoon, D. J., et al. (2010). Genetic evidence for high-altitude adaptation in tibet. *Science* 329, 72–75. doi: 10.1126/science.1189406
- Singh, A., Carandang, J., Gonzaga, Z. J. C., Collard, B. C. Y., Ismail, A. M., and Septiningsih, E. M. (2017). Identification of QTLs for yield and agronomic traits in rice under stagnant flooding conditions. *Rice* 10:15. doi: 10.1186/s12284-017-0154-5
- Sturm, R. A. (2009). Molecular genetics of human pigmentation diversity. *Hum. Mol. Genet.* 18, R9–R17. doi: 10.1093/hmg/ddp003
- Sukumaran, S., Reynolds, M. P., and Sansaloni, C. (2018). Genome-wide association analyses identify QTL hotspots for yield and component traits in durum wheat grown under yield potential, drought, and heat stress environments. *Front. Plant Sci.* 9:81. doi: 10.3389/fpls.2018.00081
- Sun, H., Lü, J., Fan, Y., Zhao, Y., Kong, F., Li, R., et al. (2008). Quantitative trait loci (QTLs) for quality traits related to protein and starch in wheat. *Prog. Nat. Sci.* 18, 825–831. doi: 10.1016/j.pnsc.2007.12.013
- Suprayogi, Y., Pozniak, C. J., Clarke, F. R., Clarke, J. M., Knox, R. E., and Singh, A. K. (2009). Identification and validation of quantitative trait loci for grain protein concentration in adapted Canadian durum wheat populations. *Theor. Appl. Genet.* 119, 437–448. doi: 10.1007/s00122-009-1050-1
- Tahmasebi, S., Heidari, B., Pakniyat, H., and McIntyre, C. L. (2016). Mapping QTLs associated with agronomic and physiological traits under terminal drought and heat stress conditions in wheat (*Triticum aestivum* L.). *Genome* 60, 26–45. doi: 10.1139/gen-2016-0017
- Taye, M., Lee, W., Jeon, S., Yoon, J., Dessie, T., Hanotte, O., et al. (2017). Exploring evidence of positive selection signatures in cattle breeds selected for different traits. *Mamm. Genome* 28, 528–541. doi: 10.1007/s00335-017-9715-6
- Terasawa, Y., Ito, M., Tabiki, T., Nagasawa, K., Hatta, K., and Nishio, Z. (2016). Mapping of a major QTL associated with protein content on chromosome 2B in hard red winter wheat (*Triticum aestivum* L.). *Breed. Sci.* 66, 471–480. doi: 10.1270/jsbbs.16026
- Tishkoff, S. A., Varkonyi, R., Cahinhinan, N., Abbes, S., Argyropoulos, G., Destro-Bisol, G., et al. (2001). Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. *Science* 293, 455–462. doi: 10.1126/science.1061573
- Tiwari, C., Wallwork, H., Arun, B., Mishra, V. K., Velu, G., Stangoulis, J., et al. (2016). Molecular mapping of quantitative trait loci for zinc, iron and protein content in the grains of hexaploid wheat. *Euphytica* 207, 563–570. doi: 10.1007/s10681-015-1544-7
- Tosi, P., Gritsch, C. S., He, J., and Shewry, P. R. (2011). Distribution of gluten proteins in bread wheat (*Triticum aestivum*) grain. *Ann. Bot.* 108, 23–35. doi: 10.1093/aob/mcr098
- Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., and Dubcovsky, J. (2006). A NAC gene regulating senescence improves grain protein, Zn, and Fe content in wheat. *Science* 314, 1298–1301. doi: 10.1126/science.1133649
- van de Wouw, M., Kik, C., van Hintum, T., van Treuren, R., and Visser, B. (2010a). Genetic erosion in crops: concept, research results and challenges. *Plant Genet. Resour.* 8, 1–15. doi: 10.1017/S1479262109990062
- van de Wouw, M., van Hintum, T., Kik, C., van Treuren, R., and Visser, B. (2010b). Genetic diversity trends in twentieth century crop cultivars: a meta analysis. *Theor. Appl. Genet.* 120, 1241–1252. doi: 10.1007/s00122-009-1252-6

- Vasemagi, A., Nilsson, J., and Primmer, C. R. (2005). Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Mol. Biol. Evol.* 22, 1067–1076. doi: 10.1093/molbev/msi093
- Wang, L., Cui, F., Wang, J., Jun, L., Ding, A., Zhao, C., et al. (2012). Conditional QTL mapping of protein content in wheat with respect to grain yield and its components. *J. Genet.* 91, 303–312. doi: 10.1007/s12041-012-0190-2
- Wu, X., Chang, X., and Jing, R. (2012). Genetic insight into yield-associated traits of wheat grown in multiple rain-fed environments. *PLoS One* 7:e31249. doi: 10.1371/journal.pone.0031249
- Würschum, T., Leiser, W. L., Kazman, E., and Longin, C. F. H. (2016). Genetic control of protein content and sedimentation volume in European winter wheat cultivars. *Theor. Appl. Genet.* 129, 1685–1696. doi: 10.1007/s00122-016-2732-0
- Xia, H., Zheng, X., Chen, L., Gao, H., Yang, H., Long, P., et al. (2014). Genetic differentiation revealed by selective loci of drought-responding EST-SSRs between upland and lowland rice in China. *PLoS One* 9:e106352. doi: 10.1371/journal.pone.0106352
- Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702
- Zanke, C. D., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V., et al. (2014). Whole genome association mapping of plant height in winter wheat (*Triticum aestivum* L.). *PLoS One* 9:e113287. doi: 10.1371/journal.pone.0113287
- Zhang, G., Wang, Y., Guo, Y., Zhao, Y., Kong, F., and Li, S. (2015). Characterization and mapping of QTLs on chromosome 2D for grain size and yield traits using a mutant line induced by EMS in wheat. *Crop J.* 3, 135–144. doi: 10.1016/j.cj.2014.11.002
- Zhang, L. Y., Liu, D. C., Guo, X. L., Yang, W. L., Sun, J. Z., Wang, D. W., et al. (2010). Genomic distribution of quantitative trait loci for yield and yield-related traits in common wheat. *J. Integr. Plant Biol.* 52, 996–1007. doi: 10.1111/j.1744-7909.2010.00967.x
- Zhang, W., Chao, S., Manthey, F., Chicaiza, O., Brevis, J. C., Echenique, V., et al. (2008). QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. *Theor. Appl. Genet.* 117, 1361–1377. doi: 10.1007/s00122-008-0869-1
- Zhang, W., and Dubcovsky, J. (2008). Association between allelic variation at the Phytoene synthase 1 gene and yellow pigment content in the wheat grain. *Theor. Appl. Genet.* 116, 635–645. doi: 10.1007/s00122-007-0697-8
- Zhang, X.-Y., Tong, Y.-P., You, G.-X., Hao, C.-Y., Ge, H.-M., Wang, L., et al. (2007). Hitchhiking effect mapping: a new approach for discovering agronomic important genes. *Agric. Sci. China* 6, 255–264. doi: 10.1016/S1671-2927(07)60043-1
- Zhao, K., Aranzana, M. J., Kim, S., Lister, C., Shindo, C., Tang, C., et al. (2007). An *Arabidopsis* example of association mapping in structured samples. *PLoS Genet.* 3:e4. doi: 10.1371/journal.pgen.0030004
- Zou, J., Semagn, K., Iqbal, M., Chen, H., Asif, M., N'Diaye, A., et al. (2017). QTLs associated with agronomic traits in the Attila × CDC Go spring wheat population evaluated under conventional management. *PLoS One* 12:e0171528. doi: 10.1371/journal.pone.0171528

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 N'Diaye, Haile, Nilsen, Walkowiak, Ruan, Singh, Clarke, Clarke and Pozniak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.