

A Low Resolution Epistasis Mapping Approach To Identify Chromosome Arm Interactions in Allohexaploid Wheat

Nicholas Santantonio,^{*,1} Jean-Luc Jannink,^{*,†} and Mark Sorrells^{*}

^{*}Cornell University, Plant Breeding and Genetics Section, School of Integrated Plant Sciences, College of Agriculture and Life Sciences, Ithaca, NY 14853 and [†]USDA ARS, Robert W. Holley Center for Agriculture & Health, Ithaca, NY 14853

ORCID IDs: 0000-0002-4351-4023 (N.S.); 0000-0003-4849-628X (J.-L.J.); 0000-0002-7367-2663 (M.S.)

ABSTRACT Epistasis is an important contributor to genetic variance. In inbred populations, pairwise epistasis is present as additive by additive interactions. Testing for epistasis presents a multiple testing problem as the pairwise search space for modest numbers of markers is large. Single markers do not necessarily track functional units of interacting chromatin as well as haplotype based methods do. To harness the power of multiple markers while minimizing the number of tests conducted, we present a low resolution test for epistatic interactions across whole chromosome arms. Epistasis covariance matrices were constructed from the additive covariances of individual chromosome arms. These covariances were subsequently used to estimate an epistatic variance parameter while correcting for background additive and epistatic effects. We find significant epistasis for 2% of the interactions tested for four agronomic traits in a winter wheat breeding population. Interactions across homeologous chromosome arms were identified, but were less abundant than other chromosome arm pair interactions. The homeologous chromosome arm pair 4BL/4DL showed a strong negative relationship between additive and interaction effects that may be indicative of functional redundancy. Several chromosome arms appeared to act as hubs in an interaction network, suggesting that they may contain important regulatory factors. The differential patterns of epistasis across different traits demonstrate that detection of epistatic interactions is robust when correcting for background additive and epistatic effects in the population. The low resolution epistasis mapping method presented here identifies important epistatic interactions with a limited number of statistical tests at the cost of low precision.

KEYWORDS

epistasis
allopolyploidy
multiple test
correction
heritability
mapping

Epistasis is the interaction of alleles, or variants, at two or more loci. Early observations of epistasis by Bateson (1909) were mostly qualitative, noting that certain loci could mask the effects at other loci. Quantitative epistasis was first suggested and defined by Fisher (1919) who coined the term 'epistasy'. Statistically, epistasis is the deviation from an

additive expectation of two or more loci, often described as a change in the slope of one locus conditional on the genotype at another locus (Fisher 1919). Variance due to quantitative epistasis has been shown to be an important contributor to the genetic variance in populations of model organisms such as *Arabidopsis* (Malmberg *et al.* 2005; Kusterer *et al.* 2007), as well as crop species such as maize (Stuber and Moll 1971; Melchinger *et al.* 1986; Lamkey *et al.* 1995; Wolf and Hallauer 1997; Lukens and Doebley 1999) and rice (Yu *et al.* 1997; Li *et al.* 2008; Shen *et al.* 2014). Significant epistasis has also been reported in allopolyploid crops like cotton (Lee *et al.* 1968) and wheat (Crossa *et al.* 2010; Jiang *et al.* 2017). Epistasis across subgenomes may be indicative of interactions between homeologous loci, analogous to dominance in diploids, and a possible contributor to that adaptation of these crops to a wide landscape (Wendel 2000; Adams and Wendel 2005; Chen 2010, 2013). However, there is still little direct evidence that epistasis between

Copyright © 2019 Santantonio *et al.*

doi: <https://doi.org/10.1534/g3.118.200646>

Manuscript received August 6, 2018; accepted for publication November 5, 2018; published Early Online November 19, 2018.

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

Supplemental material available at Figshare: <https://doi.org/10.25387/g3.7311797>.

¹Corresponding author: Cornell University, Plant Breeding and Genetics Section, School of Integrated Plant Sciences, College of Agriculture and Life Sciences, 240 Emerson Hall, Ithaca, NY 14853 E-mail: ns722@cornell.edu

homeologous loci is a large contributor to the total genetic variance in allopolyploids (Santantonio *et al.* 2018a,b).

Epistasis has also been shown to be an important contributor to evolution (Doebley *et al.* 1995; Lukens and Doebley 1999; Carlborg *et al.* 2006; Phillips 2008; Hansen 2013; Doust *et al.* 2014). There has been considerable effort over the past several decades to incorporate these non-additive genetic factors into the genotype to phenotype map. More recently these effects have been incorporated into whole genome prediction models (Vitezica *et al.* 2013; Martini *et al.* 2016; Jiang and Reif 2015; Akdemir and Jannink 2015; Wolfe *et al.* 2016; Akdemir *et al.* 2017; Jiang *et al.* 2017).

In practice, detecting epistatic interactions is difficult. The pairwise search space is large even for modest numbers of markers. For example, a population genotyped with 100 markers would require 4,950 tests for pairwise epistasis. With advances in genotyping technologies, the number of DNA markers available is typically much larger, in the tens to hundreds of thousands, and more recently in the millions. In this study, 11,604 markers were available, which would result in approximately 67 million tests for pairwise epistasis. A 0.05 genome-wide Bonferroni significance threshold for all pairwise epistasis tests in this study would then be 7.4×10^{-10} .

Several methods have been proposed to reduce the multiple testing problem. Epistasis is partitioned in part to the additive variance, particularly when allele frequencies differ from 0.5 at either locus (Hill *et al.* 2008). Therefore, genome-wide scans can be used to first identify variants with a significant additive effect, then test only all pairwise variants identified in the scan (Carlson *et al.* 2004). This can greatly reduce the number of epistatic tests performed, while increasing the likelihood that epistasis will be identified. Other methods include relaxing the multiple test correction threshold (Benjamini and Hochberg 1995), or reducing the marker pairs tested based on some criteria such as biological function (Ritchie 2011; Cowman and Koyutürk 2017; Crawford *et al.* 2017).

The multiple test correction problem is not the only challenge to identifying epistatic interactions. Allele frequency, linkage disequilibrium and the number of alleles at a given locus can all reduce the efficacy of pairwise marker epistasis detection. Low allele frequencies at either locus reduce the epistatic effect, partitioning it to the additive instead (Hill *et al.* 2008). Less than perfect linkage disequilibrium between the markers and causal mutations also reduces the apparent effect size, limiting detection much as it does for additive effects (Carlson *et al.* 2004). Single nucleotide polymorphism (SNP) markers are typically considered bi-allelic, despite the potential for numerous alleles at a single locus in the population. The impact of these factors can be reduced by using multiple linked markers to determine haplotypes. Haplotypes have been shown to be powerful in the detection of additive and interaction effects by accurately tracking larger segments of DNA in high or perfect linkage disequilibrium (LD), and allowing multiple alleles at every locus (Lin and Zeng 2006; Zhang *et al.* 2012; Jiang *et al.* 2018). While allele frequencies are typically reduced using haplotypes (*i.e.*, the frequency of two alleles will be higher than the frequency of three alleles), the added power from accurately tracking relevant LD blocks make these methods attractive.

Haplotypes do not need to be assigned directly to gain an advantage from using multiple markers to identify regions associated with complex traits. Regional heritability mapping (Nagamine *et al.* 2012; Riggio and Pong-Wong 2014) has been used to identify additive effects of rare and common variants in humans (Nagamine *et al.* 2012; Shiralil *et al.* 2016) as well as plants species like eucalyptus (Resende *et al.* 2017) and cassava (Okeke *et al.* 2018). These methods employ the estimation of additive covariance between individuals based on markers in a given region of chromatin, and are used in a mixed model to estimate the genetic variance attributable to the region. Variance components can

then be tested to determine if they are greater than zero using a likelihood ratio test.

We propose a method to greatly reduce the number of statistical tests while taking advantage of multiple markers to determine importance of epistatic interactions across chromosome arms of an allohexaploid wheat population. This method is similar to the “divide and conquer” method of Akdemir and Jannink (2015), but models interactions across chromosomes instead of local epistasis. Epistatic covariances can be formed using the Hadamard product of component additive or dominance covariance matrices (Henderson 1985; Jiang and Reif 2015; Martini *et al.* 2016). Additive by additive epistatic interactions between disjoint sets of related (*i.e.*, linked) markers can be modeled by first calculating an additive covariance for each marker set, \mathbf{K}_i and \mathbf{K}_j , and using $\mathbf{K}_i \odot \mathbf{K}_j$ as the covariance estimate of the epistatic term between these sets. We define marker sets by the chromosome arm to which they belong, and estimate the epistatic variance component between the two arms using restricted maximum likelihood (REML) while correcting for background additive and epistatic effects.

Common wheat is an important allohexaploid crop with three subgenomes, A, B and D, resulting from hybridization events approximately 500 thousand and 10 thousand years ago. Due to the allopolyploid nature of wheat, we were interested in identifying interactions across homeologous loci. Interactions at homeologous loci are analogous to dominance effects in diploid hybrids, and could be used to fix favorable homeoallelic interactions in inbred lines (Wendel 2000; Adams and Wendel 2005; Birchler *et al.* 2010; Chen 2010, 2013). Of the 21 chromosomes of wheat, chromosome arms pairs include $\binom{3}{2} 14 = 42$ homeologous pairs, $\binom{14}{2} 3 = 273$ within subgenome pairs, and $\binom{14}{2} 6 = 546$ across subgenome arm pairs.

Each chromosome arm of the wheat genome was sequenced independently using flow cytometry to assist in the assembly of the large complex genome (International Wheat Genome Sequencing Consortium 2014). The lone exception was chromosome 3B, which was sequenced and assembled in its entirety before the other chromosomes of wheat (Paux *et al.* 2008; Choulet *et al.* 2014). Therefore, assigning markers to a chromosome arm is feasible, but their position along that arm may not be well defined if the number of scaffolds is large, as was the case with the first wheat survey sequence (International Wheat Genome Sequencing Consortium 2014). Using markers across an entire chromosome arm known to be homeologous to another chromosome arms may therefore be a better strategy than attempting to assign single homeologous marker pairs (Santantonio *et al.* 2018b). If interactions are detected across homeologous regions, this may provide evidence of beneficial homeoallelic interactions indicative of inter-genomic heterosis.

We demonstrate the low resolution epistasis mapping methodology using the CNLM winter wheat dataset from Santantonio *et al.* (2018a), and show that epistasis can be detected between homeologous and non-homeologous chromosome arms.

MATERIALS AND METHODS

Plant Materials

Details of the CNLM population used in this study can be found in Santantonio *et al.* (2018a). Briefly, the dataset consists of 8,692 observations of 1,447 soft winter wheat breeding lines evaluated for four traits, grain yield (GY), plant height (PH), test weight (TW) and heading date (HD), in 26 environments across 10 years in an unbalanced design. The population was genotyped with 11,604 genotyping by

sequencing (GBS) markers distributed throughout the genome, albeit with fewer markers on the D subgenome relative to the A and B subgenomes.

Chromosome centromere positions

Chromosome centromere positions were provided by the International Wheat Genome Sequencing Consortium (IWGSC) for all chromosomes except 3B (IWGSC 2018, personal communication). Those positions were assigned by determining where chromosome arm library reads aligned to the final assembly. Each chromosome arm was sequenced independently using flow-cytometry to remove the chromosome arm from a series of aneuploid stocks, each containing an extra arm. The lone exception was chromosome 3B, which was sequenced in its entirety, so no centromere position was available for the 3B chromosome. Centromere start and stop addresses provided by IWGSC are shown in Supplemental Table S1.

While restriction sites are expected to be uniformly distributed throughout the genome, methylation of cytosine is not. One of the restriction enzymes used to generate GBS libraries, *MspI*, is sensitive to DNA methylation, digesting unmethylated DNA at a much higher rate than methylated DNA (McClelland *et al.* 1994). Methylation is an important regulator of chromatin structure, where euchromatin tends to contain few methylation sites relative to heterochromatin (Keshet *et al.* 1986). Therefore restriction sites in heterochromatin with high levels of methylation, such as at the centromere, are less likely to be retained as GBS markers because digestion is less likely to happen at these sites. This means that the GBS markers can be used to roughly assign a centromere position using the density of GBS markers along the chromosome.

To determine the centromere position of 3B, we employed kernel density estimation using the `density()` function of the 'stats' package in R to determine the smoothed density of GBS marker positions. We then assigned the 3B centromere interval to the chromosome positions flanking the second position for which the derivative of the density was zero. We performed this operation for all chromosomes to determine the efficacy of this method for determining the centromere position (Supplementary Figure S1). The positions provided by the IWGSC were used for all chromosomes other than our estimate for the 3B centromere position for all additional analyses.

Regional epistasis mapping

The low resolution epistasis mapping approach employed here uses markers from two defined regions, i and i' , to calculate additive covariance between individuals (VanRaden 2008, method I) based on those regions (*i.e.*, \mathbf{K}_i and $\mathbf{K}_{i'} \forall i \neq i'$). The Hadamard product of these additive covariance matrices can be used to produce the pairwise additive by additive epistatic relationship, $\mathbf{K}_{i \times i'} = \mathbf{K}_i \odot \mathbf{K}_{i'}$, between these two regions (Henderson 1985; Jiang and Reif 2015; Martini *et al.* 2016). In this study, we defined regions as the short (S) and long (L) arms of each chromosome, where $i \in \{1AS, 1AL, 1BS, \dots, 7BL, 7DS, 7DL\}$. Variance components for each region and their respective interaction were estimated by fitting the following nested models

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{g}_{G^-} + \mathbf{Z}\mathbf{g}_{I^-} + \boldsymbol{\varepsilon} \quad (1)$$

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{g}_{G^-} + \mathbf{Z}\mathbf{g}_{I^-} + \mathbf{Z}\mathbf{g}_{A_i} + \mathbf{Z}\mathbf{g}_{A_{i'}} + \boldsymbol{\varepsilon} \quad (2)$$

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{g}_{G^-} + \mathbf{Z}\mathbf{g}_{I^-} + \mathbf{Z}\mathbf{g}_{A_i} + \mathbf{Z}\mathbf{g}_{A_{i'}} + \mathbf{Z}\mathbf{g}_{A_i \times A_{i'}} + \boldsymbol{\varepsilon} \quad (3)$$

where \mathbf{y} is the phenotype vector, $\mathbf{1}$ is a vector of ones, μ is the population mean, \mathbf{X} is the environment incidence matrix, $\boldsymbol{\beta}$ is the

vector of fixed environmental effects, and \mathbf{Z} is the genotype incidence matrix. Residuals were assumed to be normally distributed such that $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$. Chromosome arm additive effects were assumed to be $\mathbf{g}_{A_i} \sim \mathcal{N}(0, \sigma_{a_i}^2\mathbf{K}_i)$ and $\mathbf{g}_{A_{i'}} \sim \mathcal{N}(0, \sigma_{a_{i'}}^2\mathbf{K}_{i'})$, while the chromosome arm interaction effect was assumed to be $\mathbf{g}_{A_i \times A_{i'}} \sim \mathcal{N}(0, \sigma_{a_i \times a_{i'}}^2\mathbf{K}_{i \times i'})$. Background additive, \mathbf{g}_{G^-} , and epistatic, \mathbf{g}_{I^-} , effects were included to account for population structure. The covariances of the background effects were calculated as described in Santantonio *et al.* (2018a, equation 5), but with markers belonging to region i and i' omitted from the calculation.

Sequential nested likelihood ratio tests were used to determine if the additive (model 2 vs. model 1) and interaction (model 3 vs. model 2) variance estimates of the chromosome arms were greater than zero. From the Neyman-Pearson lemma (Neyman and Pearson 1933), the likelihood ratio test statistic is defined as $D = -2(\log \mathcal{L}_{alternative} - \log \mathcal{L}_{null})$, where $D \sim \chi_{df_{H_1} - df_{H_0}}^2$, and is uniformly most powerful (UMP).

Best linear unbiased predictors (BLUPs) of each region were subsequently used to look for patterns between additive and interaction effects for the chromosome arm pair. The pairwise product of the additive chromosome arm BLUPs was then compared to the chromosome arm interaction BLUPs, in a manner analogous to the Additive \times Additive single locus model (Hill *et al.* 2008; Santantonio *et al.* 2018b). Negative associations should indicate a less-than-additive model, whereas positive relationships would demonstrate a greater than additive epistatic effect.

For the 14 three-way homeologous arm sets, a three-way interaction was included and tested against a model with only the three two-way interaction terms. We did not attempt to run all three-way chromosome arm combinations, as this would have been computationally infeasible, with $\binom{42}{3} = 11,480$ combinations. The Hadamard product of the three additive covariance matrices was used to produce the three-way additive by additive by additive epistatic relationship, $\mathbf{K}_{i \times i' \times i''} = \mathbf{K}_i \odot \mathbf{K}_{i'} \odot \mathbf{K}_{i''}$. The following two models were fit to test the three-way interaction.

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{g}_{G^-} + \mathbf{Z}\mathbf{g}_{I^-} + \mathbf{Z}\mathbf{g}_{A_i} + \mathbf{Z}\mathbf{g}_{A_{i'}} + \mathbf{Z}\mathbf{g}_{A_{i''}} + \mathbf{Z}\mathbf{g}_{A_i \times A_{i'}} + \mathbf{Z}\mathbf{g}_{A_i \times A_{i''}} + \mathbf{Z}\mathbf{g}_{A_{i'} \times A_{i''}} + \boldsymbol{\varepsilon} \quad (4)$$

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{g}_{G^-} + \mathbf{Z}\mathbf{g}_{I^-} + \mathbf{Z}\mathbf{g}_{A_i} + \mathbf{Z}\mathbf{g}_{A_{i'}} + \mathbf{Z}\mathbf{g}_{A_{i''}} + \mathbf{Z}\mathbf{g}_{A_i \times A_{i'}} + \mathbf{Z}\mathbf{g}_{A_i \times A_{i''}} + \mathbf{Z}\mathbf{g}_{A_{i'} \times A_{i''}} + \boldsymbol{\varepsilon} \quad (5)$$

A likelihood ratio test was then used to determine if adding the three-way interaction term significantly improved the model fit beyond the two-way interaction terms.

In summary, this method estimates a random additive effect for each region and a random interaction effect between these regions for all individuals while correcting for the background genetic effects of the remainder of the genome. The likelihood ratio test is used to determine if the variability of the interaction effects is greater than zero. If the interaction effects do significantly deviate from 0, then we reject the null hypothesis that there are no interacting loci between the two regions.

Software

Variance component estimation was accomplished using restricted maximum likelihood (REML) implemented in 'ASReml-R' (Gilmour 1997; Butler 2009). Other computation, analyses and figures were made using base R (R Core Team 2015) implemented in the Microsoft Open R environment 3.3.2 (Microsoft 2017) unless noted otherwise. The 'circlize' R package (Gu *et al.* 2014) was used to make Figures 2

and S4. LaTeX tables were generated using the R package 'xtable' (Dahl 2016).

Data Availability

All data used for this study can be found in Santantonio *et al.* (2018a). Additionally, we provide an example script that uses a custom R package, 'lre', to fit the models for the 4B and 4D chromosome pair (Supplementary File S1.tar.gz). The package relies on the freely available R package 'EMMREML' for solving multi-kernel mixed models (Akdemir and Okeke 2015). Supplemental material available at Figshare: <https://doi.org/10.25387/g3.7311797>.

RESULTS

Centromere positions

Most of the GBS marker density estimates of centromere locations agreed well with the positions provided by the IWGSC (Supplemental Figure S1). Chromosomes 1D and 4A were exceptions. We estimated the 3B centromere to be positioned between 344.4 Mbp and 345.0 Mbp (Supplemental Table S1).

Model fit and p-value distribution

Homeologous chromosome arm pair models each had five random genetic effects and therefore five covariance structures for the two-way interaction models. All models converged, but some variance parameter estimates were often close to the parameter boundary and were considered to be zero. Variance component estimates on the boundary did not occur for the background additive or epistatic effects, but often occurred for one or both of the additive chromosome arm effects or the interaction effect. This resulted in a relatively large number of additive and interaction variance component tests with a p-value of 1. As a result, p-value distributions were heavily skewed toward 0 and 1 (Supplementary Figures S2 and S3). Most chromosome arms had low additive effect p-values, whereas most interaction p-values were high, indicating that the majority of chromosome arm pairs do not have effect interactions large enough to detect.

Homeologous arm tests

The U-shaped distribution of the p-values suggested that when the true variance was very small or zero, the average information algorithm estimated the parameter on the boundary (*i.e.*, 0), and when it was positive, the p-value tended to be low. Larger sample sizes may be necessary to obtain uniform p-value distributions when the null hypothesis is true. We therefore considered all homeologous arm pairs with an interaction variance p-value less than 0.05 that also had positive additive variance component estimates to determine the relationship between additive chromosome arm effects and their interaction.

Seventeen homeologous chromosome arm pairs had significant interaction effects for at least one of the four traits (Table 1 and Supplemental Figure S4). Interactions involving homeologs 4 and 7 were overrepresented, with 14 of the 22 significant interactions identified between one of these two homeologs. Chromosome arm pair tests failed to detect the significant homeologous marker set interactions found on chromosome homeologs 1 and 5 for HD and homeolog 3 for PH (Santantonio *et al.* 2018b). The failure to detect these regions using the chromosome arm test suggests that the associations detected by Santantonio *et al.* (2018b) were spurious, or their signal is being washed out by the abundance of uninformative markers on those chromosome arms. The lack of a two-way arm PH interaction on chromosome arm 3S agrees with the homeologous marker set identified there, where only the three-way homeologous marker set interaction term was significant.

The test for three-way homeologous chromosome arm interactions only revealed three sets of homeologous arms that had a significant three-way interaction at $p < 0.05$ (Table 2). The three-way 3S chromosome arm interaction for PH was found to have a positive three-way arm interaction variance parameter estimate with a p-value of $p = 0.02$, supporting the evidence from Santantonio *et al.* (2018b) that found a significant 3-way interaction on 3S using homeologous markers. The 7L three-way arm interaction term was also found to have a low p-value for TW of $p = 0.006$, confirming another significant three-way homeologous marker interaction found by Santantonio *et al.* (2018b). None of these three-way tests passed a Bonferroni significance threshold.

Many interactions were detected on chromosome arms where no homeologous marker sets were identified with a significant interaction effect (Santantonio *et al.* 2018b). Notably, a strong interaction effect was identified on homeolog 6S for HD, and two regions for GY on 5S and 7L, where no significant homeologous interaction sets were identified. Neither of the interacting pairs for GY had a p-value lower than a homeologous arm Bonferroni correction of $0.05 / 42 = 0.0012$.

Relationships between chromosome arm additive and interaction effects were only considered for the ten chromosome arm pair trait combinations that had all chromosome arm additive and interaction effects with significant non-zero variance components. Of these ten, six had significant correlations between the additive product and the interaction with an absolute value ≥ 0.1 (Table 1). Four of these showed positive relationships, while the other two showed negative relationships. By far the strongest relationship detected was between 4BL and 4DL for PH ($\rho = -0.65$, Figure 1), indicating that individuals with high or low additive values for both arms tended to have genotypic values less than expected by additivity alone. Conversely, the same 4BL/4DL pair had a weak, yet positive relationship for TW ($\rho = 0.14$, Supplemental Figure S5). The 4BS/4DS pair, where the *Rht* genes are known to reside, had a weak, yet significant, positive correlation for PH (Supplemental Figure S6).

All pairwise arm tests

For all $\binom{42}{2} = 861$ pairwise chromosome arm pairs, we only consider those tests that passed a Bonferroni threshold of $0.05 / 861 = 5.8 \times 10^{-5}$ for the interaction term in this section. Seventy-nine chromosome arm interaction variance components were declared significantly greater than zero for at least one trait, representing about 2% of the number tested (Supplementary Tables S2 and S3). Of these, interactions for the PH trait were the most prevalent, representing 49 (62%), of the interactions detected. HD and TW accounted for the remaining 13 (16%) and 17 (22%) interactions, respectively. No chromosome arm interactions were detected for GY at the Bonferroni significance threshold. No interactions were detected for any of the traits involving chromosome arms 1AS, 1DL, 2AS, 2DL, 3DL, 4AS, 5AS, 5BL, 5DL, 6BL, 6DS, and 7BS at this threshold.

There were several chromosome arms that appeared to be interacting with multiple loci (Supplemental Table S4). Of these, several clearly stand out (Figure 2). Chromosome arms 1AL, 2AL, 2DS, 4BS, 4DS, 4DL, 6AS and 7AL were involved in five or more interacting pairs for PH, with 2DS, 4DS and 4DL involved in 10 or more significant pairs. The 4D chromosome in particular was involved in almost half (21) of the interacting arm pairs for PH. 7DL was involved in all but three of the interacting pairs detected for the TW trait. Arm interactions for HD did not cluster to one or a few arms in the same way as PH and TW, but 6AS and 7BL were each involved in five interacting pairs for this trait.

■ **Table 1** Table of significant homeologous chromosome arm interactions. The proportion of genetic variance attributed to each arm and their corresponding interaction are shown with statistical significance from a nested likelihood ratio test

Trait	arm _i	arm _j	(h ² _{arm_i} , h ² _{arm_j}) ^a	h ² _{arm_i × arm_j}	ρ ^b
GY	5BS	5DS	(0.038, 0.000)	0.028** ^c	0.27***
GY	7AL	7BL	(0.018, 0.000)	0.041*	0.10***
PH	2AS	2DS	(0.021, 0.079)***	0.033***	-0.04
PH	4AS	4DS	(0.000, 0.039)***	0.017*	0.19***
PH	4AL	4BL	(0.013, 0.034)*	0.029*	0.10***
PH	4AL	4DL	(0.015, 0.004)	0.027***	0.07**
PH	4BS	4DS	(0.002, 0.031)***	0.049***	0.18***
PH	4BL	4DL	(0.048, 0.003)*	0.058***	-0.65***
PH	6AL	6DL	(0.110, 0.005)**	0.024*	0.06*
PH	7AL	7BL	(0.000, 0.070)	0.029**	0.45***
TW	1BS	1DS	(0.000, 0.000)	0.073***	0.00
TW	4BL	4DL	(0.096, 0.049)***	0.013*	0.14***
TW	6AL	6BL	(0.031, 0.000)	0.047*	0.12***
TW	7AL	7DL	(0.019, 0.030)	0.140***	-0.04
TW	7BL	7DL	(0.043, 0.061)	0.092***	0.16***
HD	1BS	1DS	(0.000, 0.000)	0.018*	0.00
HD	4BS	4DS	(0.000, 0.002)	0.014**	0.15***
HD	6AS	6BS	(0.008, 0.041)*	0.049***	0.02
HD	6AS	6DS	(0.014, 0.000)	0.046***	-0.03
HD	6AL	6BL	(0.009, 0.110)*	0.013*	-0.21***
HD	7AS	7DS	(0.013, 0.045)**	0.032*	-0.05*
HD	7AL	7BL	(0.000, 0.045)***	0.025*	0.14***
HD	7BS	7DS	(0.013, 0.054)***	0.012*	0.29***

^ah² represents the proportion of the chromosome arm additive or interaction variance component estimates to the total genetic variance.
^bρ indicates the correlation between the product of the additive arm effects and their interaction effect with correlation coefficients significantly different from zero indicated by asterisks. If only one additive effect had a non-zero variance, the correlation coefficient shown is the correlation between the additive effect with the non-zero variance and the interaction effect.
^c*, **, and *** correspond to p-values < 0.05, 0.01, and a Bonferroni correction of 0.05/42 = 0.0012, respectively.

Most correlations between the additive products and the epistatic effect were low in magnitude (*i.e.*, < 0.3), particularly for the TW and HD traits. Notable exceptions include the 4BL/4DL pair for PH, which had a highly negative correlation, as previously noted. Pairs with moderate magnitude tended to also include the 4DL chromosome arm, but other pairs with moderate correlations between the product of their additive and interaction effects included the 1AL/2AL, 1AL/7AL, and 3AS/6AS arm pairs.

DISCUSSION

Centromere positions

While our assigned position for the 3B centromere position is an estimate, most of the other chromosome estimates were close to the centromere position provided by the IWGSC (Supplemental Figure S1). The centromere position estimate reported here should be sufficient to assign most of the 3B markers to the correct chromosome arm for the subsequent analyses.

Model fit and p-value distribution

The distribution of p-values from the likelihood ratio test should be uniform if no true interactions exist. If interactions are important, then we would expect to see a skewed distribution with many small p-values. However, the p-values were often calculated to be 1 because the variance components were estimated on the parameter boundary (*i.e.*, zero), resulting in the U-shaped distribution. When variance parameters are estimated on the parameter boundary, the p-value becomes 1 simply due to the fact that the variance component is zero. This is likely due to a lack of sufficient population size to distinguish and resolve multiple small variance components. Perhaps another explanation may be provided by the use of the the average information algorithm to fit the mixed model, which may lose a small portion of information by avoiding the calculation of the second derivative of the likelihood function. While other algorithms exist for solving REML problems, the computational burden of resolving multiple variance components with dense covariance structures may be restrictive. Further investigation is necessary to determine how large a population need be to resolve

■ **Table 2** Table of significant three-way homeologous chromosome arm interactions. The proportion of genetic variance attributed to each arm and their corresponding interaction are shown with statistical significance from a nested likelihood ratio test indicated by asterisks

Trait	arm _i	arm _j	arm _k	(h ² _{arm_i} , h ² _{arm_j} , h ² _{arm_k}) ^a	(h ² _{arm_i × arm_j} , h ² _{arm_i × arm_k} , h ² _{arm_j × arm_k})	h ² _{arm_i × arm_j × arm_k}
PH	3AS	3BS	3DS	(0.017, 0.017, 0.054)** ^b	(0.000, 0.000, 0.007)**	0.010*
TW	7AL	7BL	7DL	(0.017, 0.044, 0.035)	(0.005, 0.057, 0.000)	0.051**
HD	6AS	6BS	6DS	(0.005, 0.031, 0.000)*	(0.035, 0.021, 0.000)*	0.019*

^ah² represents the proportion of the chromosome arm additive or interaction variance component estimates to the total genetic variance.
^b* and ** correspond to p-values < 0.05 and 0.01, respectively.

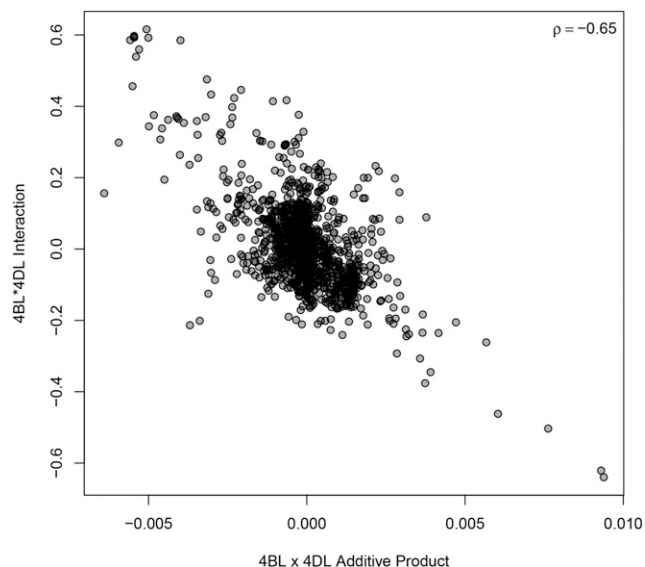


Figure 1 Interaction effect of chromosome 4BL by 4DL plotted against the product of the additive effects for 4BL and 4DL for PH. ρ indicates the Pearson correlation coefficient.

multiple genetic variance parameters with magnitudes of 1% or less of the total variance.

Homeologous arm tests

Most of the homeologous chromosome arm interactions detected across all traits involved homeologs 4 and 7. The less-than-additive trend observed for the 4BL/4DL pair for PH may suggest a significant degree of gene functional redundancy between these two arms. Functional gene redundancy between homeologous alleles should result in a less-than-additive effect similar to partial dominance at alleles. Despite having a weak positive additive genetic trait correlation between PH and TW (0.3, Santantonio *et al.* 2018a), the 4BL/4DL pair had a weak, yet positive relationship for TW. This provides evidence that the observed pattern is not simply a genetic artifact and may indicate differential gene function for these two traits.

A negative correlation for PH was not observed for the 4BS/4DS chromosome arm pair, as would be expected from previous results for the *Rht* genes that reside on those chromosome arms (Santantonio *et al.* 2018b). In that study, a significant less-than-additive effect of the wild-type *Rht* homeologs suggested that the two genes have partial functional redundancy. The mutant alleles are insensitive to the plant hormone Gibberellic acid (GA), and result in a semi-dwarf plant stature which greatly reduces lodging in high nitrogen environments. Plants with both mutant alleles are far shorter than expected based on the additive effect of the two mutant alleles. From the perspective of functionality, one functional homeolog is able to recover most of the gene pathway function, with relatively little gained from the addition of a second functional homeolog. Therefore, the loss of one gene does not result in complete loss of the gene pathway, but merely a non-additive reduction in its total activity.

The lack of a negative correlation between the additive and interaction effects for the 4BS/4DS pair cast some doubt on the usefulness of these correlations to infer the direction of the epistatic effect. The relationship between the product of the additive effects and the interaction was thought to mirror the $\{-1, 1\}$ Additive \times Additive epistatic model using a multi-locus approach (Hill *et al.* 2008; Santantonio *et al.* 2018b), but it is unclear what is driving these trends. Further

investigation into the relationships between regional additive and epistatic effects is warranted.

For inbred allopolyploids, multi-subunit protein complexes can be comprised of genes from a single subgenome, or from multiple subgenomes. If functional copies of subunits exist on both genomes, the formation of subgenome hetero-complexes may occur. Protein complexes comprised of evolutionarily divergent subunits may have increased or, more likely, decreased functionality. If hetero-complexes display decreased functionality, then we would expect the relationship between the additive and epistatic effects to be negative.

It is unlikely that all homeologous interactions are so large in effect that they are quickly fixed after the hybridization event. The distribution of epistatic effects is likely similar in shape to the distribution of additive effects. These distributions will change based on the complexity of the trait. If a trait is governed by relatively few loci, the relatively few epistatic interactions may have larger effects, and may be easier to detect. In contrast, a large number of small effect additive loci may also result in a large number of small effect interactions that are too small to detect in populations of moderate size.

All pairwise arm tests

PH appears to exhibit a higher degree of epistasis than either TW or HD. However, the number of interacting loci or chromosome arms detected was not directly related to the observed increase in genomic prediction accuracy by inclusion of epistatic predictors. Santantonio *et al.* (2018a) found HD to have the largest percent increase in accuracy from the additive model by including all pairwise additive by additive interactions, yet had the fewest detectable interacting chromosome arms, other than GY.

GY showed no evidence of important epistatic interactions in this study, as has been previously shown (Santantonio *et al.* 2018a,b). This may be due to one of two explanations. The first and most obvious is that grain yield is not subject to epistatic gene action. This would mean that all genes contribute additively to the collection and allocation of resources to vegetative tissue, and then reallocation to the ear during flowering and grain fill. The second and more likely explanation is that GY is the culmination of essentially all the genes working in concert to produce the final outcome, and interactions with such small effects may simply be too small to detect (Xu and Jia 2007; Wu *et al.* 2012). Differential response to environmental stress across years and locations may further reduce the ability to detect interactions if they are only important in certain environments.

While we corrected for population structure on both the additive and epistatic levels (*i.e.*, using additive and additive by additive genetic covariance terms), it is possible that residual structure is causing the observed additive and epistasis relationships. The drastically different patterns in the arm pair test results for each trait suggests otherwise. If these interactions were due to population structure, we would expect to see similar patterns of significance across all traits. When we omitted the background epistatic effect, most of the 861 interactions were significant (results not shown). We deemed this to be due to chromosome arm epistatic relationship matrices modeling close relationships in the population regardless of which unit of chromatin was used to determine those relationships. However, it is possible that these interactions are far more prevalent than suggested here, and that correction for background epistatic effects is diluting true genetic signal.

The prevalence of a few chromosome arms interacting with many other arms is of particular interest, due to the potential for one site to influence the expression of so many other sites. These sites appear to act as hubs in interaction networks, and have been shown to be prevalent in

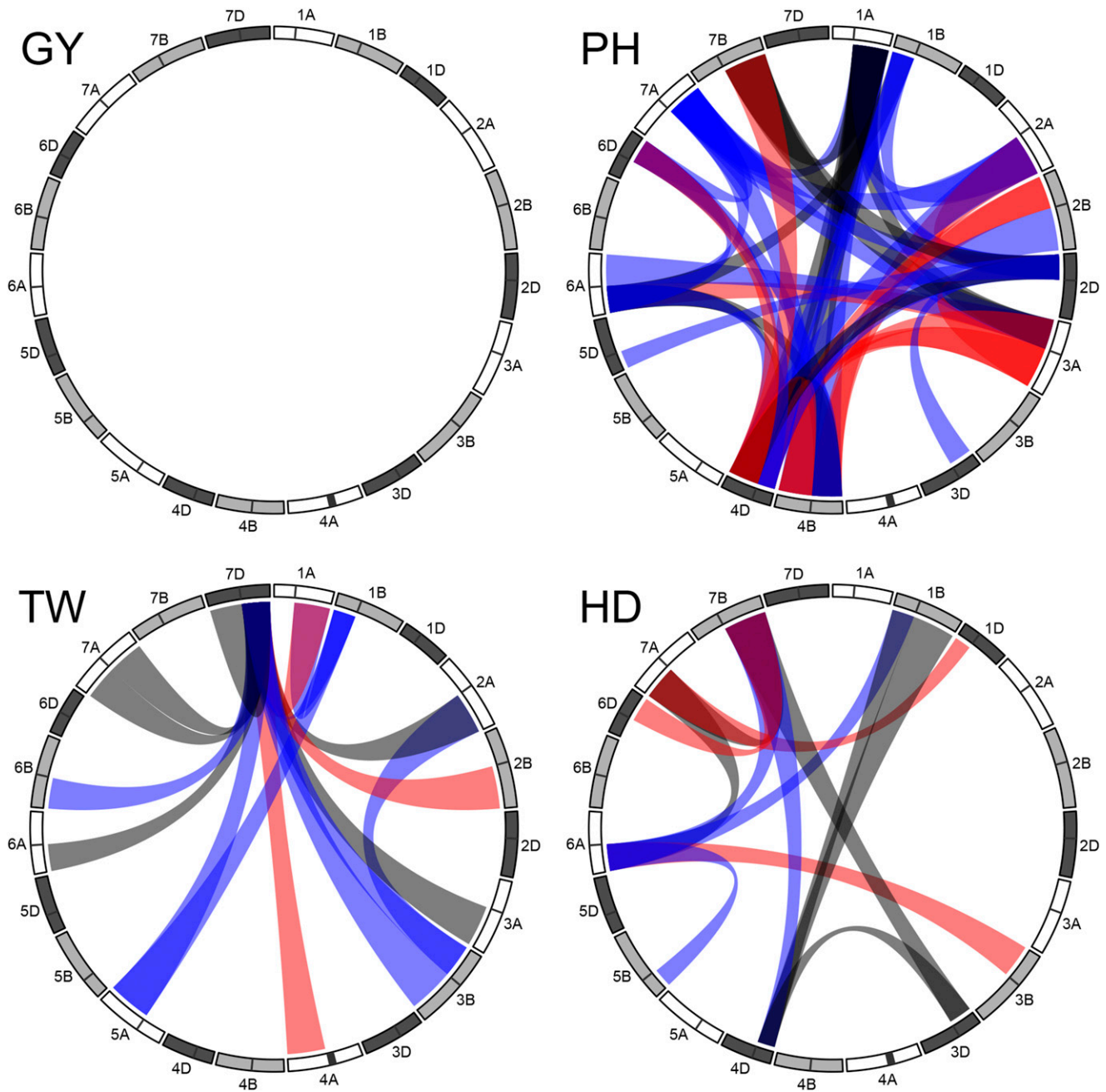


Figure 2 Chromosome arm interactions significant at a Bonferroni correction of $0.05/861 = 5.8 \times 10^{-5}$. Blue and red bridges indicate interactions with a significant positive or negative correlation between the product of the additive effects and their interaction effect, respectively. Black bridges indicate significant interactions that did not have a significant correlation between additive products and the interaction effect.

yeast (Forsberg *et al.* 2017). Jiang *et al.* (2017) observed a large proportion of the epistatic interactions affecting GY involved chromosomes 4A and 7D in a large population of hybrid wheat. While we did not detect a large number of interactions involving 4A, 7D was particularly important for TW. However, the interactions that they detected appear to be on the short arm of chromosome 7D, instead of the long arm as we observed. It appears that the hub loci detected in this study are not the same as those of Jiang *et al.* (2017), although they used a different genome assembly than used in this report. The signal detected for hub arms may be due to the presence of functional and

non-functional alleles at important upstream regulators, such as transcription factors. In this case, a non-functional transcription factor could cause the suppression of differential additive alleles.

The detection of chromosome arm interactions not identified in the homeologous marker sets of Santantonio *et al.* (2018b) suggests that single marker sets may miss important interactions. It is unclear if these interactions would have been detected if they had tested all pairwise epistatic interactions between markers. While all possible tests can be conducted, this increases the multiple testing problem drastically and may result in the loss of ability to detect any interactions. It is unclear

how large the effect sizes of a single pair of interacting loci would need to be to show up in a variance component estimated from multiple loci. While this method may not work well for a single large effect interaction, it may work well for many small effect interactions as might be expected for homeologous regions.

It should be noted that epistatic relationships formed from the Hadamard product of covariance matrices have the property of shrinking distant relationships while emphasizing close ones. For example, two lines with an additive covariance of 0.1 will have an epistatic covariance of 0.01, whereas two lines with an additive covariance of 0.9 will have an epistatic covariance of 0.81. It may be that there are several levels of relatedness that must be considered to properly account for genetic relatedness. The pedigree is an example of a covariance estimation procedure that emphasizes close relationships and deemphasizes more distant ones. Considering both pedigree and marker based covariance matrices has been shown to be more predictive than using either alone (de los Campos *et al.* 2009b; Crossa *et al.* 2010). Other methods, including Reproducing Kernel Hilbert Spaces (RKHS), can be used to model these various degrees of genetic relatedness (de los Campos *et al.* 2009a; Crossa *et al.* 2010), but may have less genetic interpretability than the method presented here.

The utility of the low resolution approach will depend on the magnitude and direction of effects, as well as the number and distribution of causal interacting loci. This is further complicated by the relative LD between these loci, the markers tagging them, and the loci with which they interact. While single markers may not be in high LD with causal loci, the combination of some markers can form haplotypes that are in high LD with one or more causal loci. Haplotype combinations across regions may then flag true interacting loci that would be otherwise undetectable using single marker interactions. While our method does not explicitly define these haplotypes, it does capture these relationships through the additive genetic covariance of the region. We suspect that this method will be the most appropriate for traits with moderate genetic complexity, as demonstrated by the lack of interacting regions detected for perhaps the most complex trait, grain yield.

CONCLUSION

The interacting pairs presented here do not have the precision to make claims of interacting genes. Nor are these interactions necessarily targets for selection. They do, however, demonstrate that there appears to be global patterns of epistasis across the genome. Seemingly additive only traits have often been shown to be under a high degree of epistasis when careful investigation is used to elucidate the trait (Carlborg *et al.* 2006; Forsberg *et al.* 2017). Some have argued that much, if not most, of genetic variation is subject to epistasis (Carlborg and Haley 2004; Carlborg *et al.* 2006; Huang *et al.* 2012; Forsberg *et al.* 2017), where the rest of the genome must be functional to express additive differences in alleles.

This is evident when we consider the the complexity of the cell, where no genes truly work independently of one another. In order to create the complex structure of the cell, proteins may interact with other proteins, both alike and dislike to them, to form multi-subunit complexes. Therefore allelic variation alone should be sufficient to produce epistatic variation. It is merely our inability to separate this variation from “additive” variation under classic parameterizations that leads many to conclude that epistasis is not important (Carlborg *et al.* 2006; Hill *et al.* 2008; Huang *et al.* 2012; Huang and Mackay 2016; Forsberg *et al.* 2017).

Further research into this methodology might be used to identify meaningful haplotypes. Once interacting segments are identified, they can each be split into multiple pieces for further refinement of the

method, while nominally increasing the number of tests performed. The low resolution epistasis mapping approach presented here emphasizes the power of using multiple genetic markers to test for interacting genomic regions, albeit at the cost of low precision.

ACKNOWLEDGMENTS

Funding of this research was provided by the USDA National Needs Fellowship for N. Santantonio, in partial fulfillment of the requirements for a Ph.D in Plant Breeding and Genetics at Cornell University. Additionally, the field trials comprising the phenotypic data for the CNLM population were funded in part by the Hatch Project # 149-447. Genotyping was funded by the Wheat Coordinated Agricultural Project (WheatCAP). The authors thank the International Wheat Genome Sequencing Consortium for prepublication access to IWGSC RefSeq v1.0 and Martin Mascher at the Leibniz-Institute of Plant Genetics and Crop Plant Research IPK (collaborating with IWGSC) for providing chromosome centromere positions.

Note added in proof: See Santantonio *et al.* 2019 (pp. 285–298) in this issue and (pp. 1105–1122) in the Genetics March 2019 issue, for related works.

LITERATURE CITED

- Adams, K. L., and J. F. Wendel, 2005 Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* 8: 135–141. <https://doi.org/10.1016/j.pbi.2005.01.001>
- Akdemir, D., and J.-L. Jannink, 2015 Locally epistatic genomic relationship matrices for genomic association and prediction. *Genetics* 199: 857–871. <https://doi.org/10.1534/genetics.114.173658>
- Akdemir, D., J.-L. Jannink, and J. Isidro-Sánchez, 2017 Locally epistatic models for genome-wide prediction and association by importance sampling. *Genet. Sel. Evol.* 49: 74. <https://doi.org/10.1186/s12711-017-0348-8>
- Akdemir, D. and U. G. Okeke, 2015 EMMREML: *Fitting mixed models with known covariance structures*. R package version 3.1.
- Bateson, W., 1909 *Mendel's principles of heredity*, Cambridge Univ. Press, New York.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57: 289–300.
- Birchler, J. A., H. Yao, S. Chudalayandi, D. Vaiman, and R. A. Veitia, 2010 Heterosis. *Plant Cell* 22: 2105–2112. <https://doi.org/10.1105/tpc.110.076133>
- Butler, D., 2009 *asreml: asreml() fits the linear mixed model*. R package version 3.0.
- Carlborg, Ö., and C. S. Haley, 2004 Epistasis: too often neglected in complex trait studies? *Nat. Rev. Genet.* 5: 618–625. <https://doi.org/10.1038/nrg1407>
- Carlborg, Ö., L. Jacobsson, P. Åhlgren, P. Siegel, and L. Andersson, 2006 Epistasis and the release of genetic variation during long-term selection. *Nat. Genet.* 38: 418–420. <https://doi.org/10.1038/ng1761>
- Carlson, C. S., M. A. Eberle, L. Kruglyak, and D. A. Nickerson, 2004 Mapping complex disease loci in whole-genome association studies. *Nature* 429: 446–452. <https://doi.org/10.1038/nature02623>
- Chen, Z. J., 2010 Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci.* 15: 57–71. <https://doi.org/10.1016/j.tplants.2009.12.003>
- Chen, Z. J., 2013 Genomic and epigenetic insights into the molecular bases of heterosis. *Nat. Rev. Genet.* 14: 471–482. <https://doi.org/10.1038/nrg3503>
- Choulet, F., A. Alberti, S. Theil, N. Glover, V. Barbe *et al.*, 2014 Structural and functional partitioning of bread wheat chromosome 3b. *Science* 345: 1249721. <https://doi.org/10.1126/science.1249721>
- Cowman, T., and M. Koyutürk, 2017 Prioritizing tests of epistasis through hierarchical representation of genomic redundancies. *Nucleic Acids Res.* 45: e131. <https://doi.org/10.1093/nar/gkx505>
- Crawford, L., P. Zeng, S. Mukherjee, and X. Zhou, 2017 Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative

- traits. *PLoS Genet.* 13: e1006869. <https://doi.org/10.1371/journal.pgen.1006869>
- Crossa, J., G. de Los Campos, P. Pérez, D. Gianola, J. Burgueño *et al.*, 2010 Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* 186: 713–724. <https://doi.org/10.1534/genetics.110.118521>
- Dahl, D. B., 2016 *xtable: Export Tables to LaTeX or HTML*. R package version 1.8–2.
- de los Campos, G., D. Gianola, and G. J. Rosa, 2009a Reproducing kernel hillbert spaces regression: a general framework for genetic evaluation. *J. Anim. Sci.* 87: 1883–1887. <https://doi.org/10.2527/jas.2008-1259>
- de los Campos, G., H. Naya, D. Gianola, J. Crossa, A. Legarra *et al.*, 2009b Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics* 182: 375–385. <https://doi.org/10.1534/genetics.109.101501>
- Doebley, J., A. Stec, and C. Gustus, 1995 *teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* 141: 333–346.
- Doust, A. N., L. Lukens, K. M. Olsen, M. Mauro-Herrera, A. Meyer *et al.*, 2014 Beyond the single gene: How epistasis and gene-by-environment effects influence crop domestication. *Proc. Natl. Acad. Sci. USA* 111: 6178–6183. <https://doi.org/10.1073/pnas.1308940110>
- Fisher, R. A., 1919 Xv.–the correlation between relatives on the supposition of mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 52: 399–433. <https://doi.org/10.1017/S0080456800012163>
- Forsberg, S. K., J. S. Bloom, M. J. Sadhu, L. Kruglyak, and Ö. Carlborg, 2017 Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. *Nat. Genet.* 49: 497–503. <https://doi.org/10.1038/ng.3800>
- Gilmour, A., 1997 *Asreml for testing fixed effects and estimating multiple trait variance components*. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* 12: 386–390.
- Gu, Z., L. Gu, R. Eils, M. Schlesner, and B. Brors, 2014 *circize* implements and enhances circular visualization in *r*. *Bioinformatics* 30: 2811–2812. <https://doi.org/10.1093/bioinformatics/btu393>
- Hansen, T. F., 2013 Why epistasis is important for selection and adaptation. *Evolution* 67: 3501–3511. <https://doi.org/10.1111/evo.12214>
- Henderson, C., 1985 Best linear unbiased prediction of nonadditive genetic merits in noninbred populations. *J. Anim. Sci.* 60: 111–117. <https://doi.org/10.2527/jas1985.6011111x>
- Hill, W. G., M. E. Goddard, and P. M. Visscher, 2008 Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* 4: e1000008. <https://doi.org/10.1371/journal.pgen.1000008>
- Huang, W., and T. F. Mackay, 2016 The genetic architecture of quantitative traits cannot be inferred from variance component analysis. *PLoS Genet.* 12: e1006421. <https://doi.org/10.1371/journal.pgen.1006421>
- Huang, W., S. Richards, M. A. Carbone, D. Zhu, R. R. Anholt *et al.*, 2012 Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. *Proc. Natl. Acad. Sci. USA* 109: 15553–15559. <https://doi.org/10.1073/pnas.1213423109>
- International Wheat Genome Sequencing Consortium, 2014 A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* 345: 1251788. <https://doi.org/10.1126/science.1251788>
- IWGSC, I. W. G. S. C., 2018 Shifting the limits in wheat research and breeding using a fully annotated reference genome by the international wheat genome sequencing consortium (iwgsc). *Science* 361: eaar7191.
- Jiang, Y., and J. C. Reif, 2015 Modeling epistasis in genomic selection. *Genetics* 201: 759–768. <https://doi.org/10.1534/genetics.115.177907>
- Jiang, Y., R. H. Schmidt, and J. C. Reif, 2018 Haplotype-based genome-wide prediction models exploit local epistatic interactions among markers. *G3: Genes, Genomes, Genetics* 8: 1687–1699. <https://doi.org/10.1534/g3.117.300548>
- Jiang, Y., R. H. Schmidt, Y. Zhao, and J. C. Reif, 2017 A quantitative genetic framework highlights the role of epistatic effects for grain-yield heterosis in bread wheat. *Nat. Genet.* 49: 1741–1746. <https://doi.org/10.1038/ng.3974>
- Keshet, I., J. Lieman-Hurwitz, and H. Cedar, 1986 Dna methylation affects the formation of active chromatin. *Cell* 44: 535–543. [https://doi.org/10.1016/0092-8674\(86\)90263-1](https://doi.org/10.1016/0092-8674(86)90263-1)
- Kusterer, B., J. Muminovic, H. F. Utz, H.-P. Piepho, S. Barth *et al.*, 2007 Analysis of a triple testcross design with recombinant inbred lines reveals a significant role of epistasis in heterosis for biomass-related traits in *Arabidopsis*. *Genetics* 175: 2009–2017. <https://doi.org/10.1534/genetics.106.069005>
- Lamkey, K. R., B. J. Schnicker, and A. E. Melchinger, 1995 Epistasis in an elite maize hybrid and choice of generation for inbred line development. *Crop Sci.* 35: 1272–1281. <https://doi.org/10.2135/cropsci1995.0011183X003500050004x>
- Lee, J. A., C. C. Cockerham, and F. Smith, 1968 The inheritance of gossypol level in *Gossypium* i. additive, dominance, epistatic, and maternal effects associated with seed gossypol in two varieties of *Gossypium hirsutum* l. *Genetics* 59: 285–298.
- Li, L., K. Lu, Z. Chen, T. Mu, Z. Hu *et al.*, 2008 Dominance, over-dominance and epistasis condition the heterosis in two heterotic rice hybrids. *Genetics* 180: 1725–1742. <https://doi.org/10.1534/genetics.108.091942>
- Lin, D., and D. Zeng, 2006 Likelihood-based inference on haplotype effects in genetic association studies. *J. Am. Stat. Assoc.* 101: 89–104. <https://doi.org/10.1198/016214505000000808>
- Lukens, L. N., and J. Doebley, 1999 Epistatic and environmental interactions for quantitative trait loci involved in maize evolution. *Genet. Res.* 74: 291–302. <https://doi.org/10.1017/S0016672399004073>
- Malmberg, R. L., S. Held, A. Waits, and R. Mauricio, 2005 Epistasis for fitness-related quantitative traits in *Arabidopsis thaliana* grown in the field and in the greenhouse. *Genetics* 171: 2013–2027. <https://doi.org/10.1534/genetics.105.046078>
- Martini, J. W., V. Wimmer, M. Erbe, and H. Simianer, 2016 Epistasis and covariance: how gene interaction translates into genomic relationship. *Theor. Appl. Genet.* 129: 963–976. <https://doi.org/10.1007/s00122-016-2675-5>
- McClelland, M., M. Nelson, and E. Raschke, 1994 Effect of site-specific modification on restriction endonucleases and dna modification methyltransferases. *Nucleic Acids Res.* 22: 3640–3659. <https://doi.org/10.1093/nar/22.17.3640>
- Melchinger, A., H. Geiger, and F. Schnell, 1986 Epistasis in maize (*Zea mays* l.). *Theor. Appl. Genet.* 72: 231–239. <https://doi.org/10.1007/BF00266997>
- Microsoft and R Core Team, 2017 *Microsoft R Open*. Microsoft Corporation, Redmond, Washington. <https://mran.microsoft.com>
- Nagamine, Y., R. Pong-Wong, P. Navarro, V. Vitart, C. Hayward *et al.*, 2012 Localising loci underlying complex trait variation using regional genomic relationship mapping. *PLoS One* 7: e46501. <https://doi.org/10.1371/journal.pone.0046501>
- Neyman, J., and E. S. Pearson, 1933 Ix. on the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. Lond. A* 231: 289–337. <https://doi.org/10.1098/rsta.1933.0009>
- Okeke, U. G., D. Akdemir, I. Rabbi, P. Kulakow, and J.-L. Jannink, 2018 Regional heritability mapping provides insights into dry matter content in african white and yellow cassava populations. *Plant Genome* 11: 1–18. <https://doi.org/10.3835/plantgenome2017.06.0050>
- Paux, E., P. Sourdille, J. Salse, C. Saintenac, F. Choulet, *et al.*, 2008 A physical map of the 1-gigabase bread wheat chromosome 3b. *Science* 322: 101–104.
- Phillips, P. C., 2008 Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* 9: 855–867. <https://doi.org/10.1038/nrg2452>
- R Core Team, 2015 *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
- Resende, R. T., M. D. V. Resende, F. F. Silva, C. F. Azevedo, E. K. Takahashi *et al.*, 2017 Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in *Eucalyptus*. *New Phytol.* 213: 1287–1300. <https://doi.org/10.1111/nph.14266>

- Riggio, V., and R. Pong-Wong, 2014 Regional heritability mapping to identify loci underlying genetic variation of complex traits. *BMC Proc.* 8: S3. <https://doi.org/10.1186/1753-6561-8-S5-S3>
- Ritchie, M. D., 2011 Using biological knowledge to uncover the mystery in the search for epistasis in genome-wide association studies. *Ann. Hum. Genet.* 75: 172–182. <https://doi.org/10.1111/j.1469-1809.2010.00630.x>
- Santantonio, N., J.-L. Jannink, and M. E. Sorrells, 2018a Prediction of subgenome additive and interaction effects in allohexaploid wheat. *G3 (Bethesda)* 9: 685–698. <https://doi.org/10.1534/g3.118.200613>
- Santantonio, N., J.-L. Jannink, and M. E. Sorrells, 2018b A subfunctionalization epistasis model to evaluate homeologous gene interactions in allopolyploid wheat. *bioRxiv* 376731. <https://doi.org/10.1101/376731>
- Shen, G., W. Zhan, H. Chen, and Y. Xing, 2014 Dominance and epistasis are the main contributors to heterosis for plant height in rice. *Plant Sci.* 215–216: 11–18. <https://doi.org/10.1016/j.plantsci.2013.10.004>
- Shirali, M., R. Pong-Wong, P. Navarro, S. Knott, C. Hayward *et al.*, 2016 Regional heritability mapping method helps explain missing heritability of blood lipid traits in isolated populations. *Heredity* 116: 333–338. <https://doi.org/10.1038/hdy.2015.107>
- Stuber, C., and R. Moll, 1971 Epistasis in maize (*zea mays* L.). ii: Comparison of selected with unselected populations. *Genetics* 67: 137–149.
- VanRaden, P., 2008 Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91: 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Vitezica, Z. G., L. Varona, and A. Legarra, 2013 On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics* 195: 1223–1230. <https://doi.org/10.1534/genetics.113.155176>
- Wendel, J. F., 2000 Genome evolution in polyploids, pp. 225–249 in *Plant molecular evolution*, Springer, New York. https://doi.org/10.1007/978-94-011-4221-2_12
- Wolf, D. P., and R. Hallauer, 1997 Triple testcross analysis to detect epistasis in maize. *Crop Sci.* 37: 763–770. <https://doi.org/10.2135/cropsci1997.0011183X003700030012x>
- Wolfe, M. D., P. Kulakow, I. Y. Rabbi, and J.-L. Jannink, 2016 Marker-based estimates reveal significant non-additive effects in clonally propagated cassava (*manihot esculenta*): implications for the prediction of total genetic value and the selection of varieties. *G3 (Bethesda)* 6: 3497–3506. <https://doi.org/10.1534/g3.116.033332>
- Wu, X., X. Chang, and R. Jing, 2012 Genetic insight into yield-associated traits of wheat grown in multiple rain-fed environments. *PLoS One* 7: e31249. <https://doi.org/10.1371/journal.pone.0031249>
- Xu, S., and Z. Jia, 2007 Genomewide analysis of epistatic effects for quantitative traits in barley. *Genetics* 175: 1955–1963. <https://doi.org/10.1534/genetics.106.066571>
- Yu, S., J. Li, C. Xu, Y. Tan, Y. Gao *et al.*, 1997 Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. USA* 94: 9226–9231. <https://doi.org/10.1073/pnas.94.17.9226>
- Zhang, L., R. Liu, Z. Wang, D. A. Culver, and R. Wu, 2012 Modeling haplotype-haplotype interactions in case-control genetic association studies. *Front. Genet.* 3: 2. <https://doi.org/10.3389/fgene.2012.00002>

Communicating editor: J. Wendel