## ANIMAL GENETICS AND GENOMICS

# Association analysis of loci implied in "buffering" epistasis

Antonio Reverter,[†,1] Zulma G. Vitezica,[‡] Marina Naval-Sánchez,[†]
John Henshall,[∥] Fernanda S. S. Raidan,[$] Yutao Li,[†] Karin Meyer,[¶]
Nicholas J. Hudson,[**] Laercio R. Porto-Neto,[†] and Andrés Legarra[‡]

[†]CSIRO Agriculture & Food, St. Lucia, Brisbane, QLD 4067, Australia, [‡]INRA/INPT, UMR 1388 GenPhySE, F-31326 Castanet-Tolosan, France, [∥]Cobb-Vantress Inc., Siloam Springs, AR 72761-1030, [$]CSIRO Agriculture & Food, Hobart, TAS 7004, Australia, [¶]Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351, Australia, [**]School of Agriculture and Food Sciences, The University of Queensland, Gatton, QLD 4343, Australia

[1]Corresponding author: tony.reverter-gomez@csiro.au

## Abstract

The existence of buffering mechanisms is an emerging property of biological networks, and this results in the buildup of robustness through evolution. So far, there are no explicit methods to find loci implied in buffering mechanisms. However, buffering can be seen as interaction with genetic background. Here we develop this idea into a tractable model for quantitative genetics, in which the buffering effect of one locus with many other loci is condensed into a single statistical effect, multiplicative on the total additive genetic effect. This allows easier interpretation of the results and simplifies the problem of detecting epistasis from quadratic to linear in the number of loci. Using this formulation, we construct a linear model for genome-wide association studies that estimates and declares the significance of multiplicative epistatic effects at single loci. The model has the form of a variance components, norm reaction model and likelihood ratio tests are used for significance. This model is a generalization and explanation of previous ones. We test our model using bovine data: Brahman and Tropical Composite animals, phenotyped for body weight at yearling and genotyped at high density. After association analysis, we find a number of loci with buffering action in one, the other, or both breeds; these loci do not have a significant statistical additive effect. Most of these loci have been reported in previous studies, either with an additive effect or as footprints of selection. We identify buffering epistatic SNPs present in or near genes reported in the context of signatures of selection in multi-breed cattle population studies. Prominent among these genes are those associated with fertility (*INHBA*, *TSHR*, *ESRRG*, *PRLR*, and *PPARG*), growth (*MSTN*, *GHR*), coat characteristics (*KIT*, *MITF*, *PRLR*), and heat resistance (*HSPA6* and *HSPA1A*). In these populations, we found loci that have a nonsignificant statistical additive effect but a significant epistatic effect. We argue that the discovery and study of loci associated with buffering effects allow attacking the difficult problems, among others, of the release of maintenance variance in artificial and natural selection, of quick adaptation to the environment, and of opposite signs of marker effects in different backgrounds. We conclude that our method and our results generate promising new perspectives for research in evolutionary and quantitative genetics based on the study of loci that buffer effect of other loci.

**Key words:** beef cattle, epistasis, GWAS, selection

**Abbreviations**

| | |
|---|---|
| BB | Brahman |
| FDR | false discovery rate |
| GEBV | genomic estimated breeding value |
| GWAS | genome-wide association study |
| PLAG1 | pleomorphic adenoma gene 1 |
| PRLR | prolactin receptor |
| PRPF4B | pre-mRNA processing factor 4B |
| QTL | quantitative trait loci |
| REML | restricted maximum likelihood |
| SP | single nucleotide polymorphism |
| TBX5 | T-box 5 transcription factor |
| TC | Tropical Composite |
| YWT | yearling weight |

## Introduction

The availability of high-density single nucleotide polymorphism (**SP**) genotypes in livestock species allows for the exploration of nonadditive effects to a degree not often captured by pedigree relationships alone. In particular, epistasis—the interaction between loci—is thought to play a key role in defining the genetic architecture of complex traits (Mackay, 2014). However, exploring all possible SNP to SNP combinations is computationally prohibitively and statistically underpowered. Hence, alternative compromises are being proposed such as the identification of the interaction between one SNP and the polygenic background (Crawfor et al., 2017).

Epistasis against the polygenic background is one of the biological phenomena in which epistasis is likely implied in "buffering" (Visser et al., 2003; Flatt, 2005), a mechanism that would allow the buildup of robustness through evolution (see Flatt, 2005, for examples). A known example is chaperones (Visser et al., 2003; Kitano, 2004). Loci implied in buffering would mitigate heritable perturbations. For instance, for a trait with intermediate optima, too high total genotypic values would not be expressed.

Inspired by these models, here we present a unique computational approach for the rapid identification of buffering epistatic SNPs based on those with a significant effect on the phenotype, however with an opposed effect depending on the genetic background of the sampled population. Preliminary derivations were presented by Reverter et al. (2018) with a mechanical heuristic and a fast approximate numerical method applied to one population of Brahman (**BB**) cattle. Here, we expand that work by deriving an exact maximum likelihood method and using a second larger population of Tropical Composite (**TC**) cattle.

## Materials and Methods

Animal Care and Use Committee approval was not obtained for this study because historical data were used and no animals were handled as part of the study. Analyses were performed on phenotypic data and DNA samples that had been collected previously as part of the Australian Cooperative Research Centre for Beef Genetic Technologies (Beef CRC; http://www.beefcrc.com/).

### Animals, phenotypes, and genotypes

Animals, phenotypes, and genotypes used in this study were a subset of those recently described in Raidan et al. (2018). In brief, we used body yearling weight (**YWT**) data of 2,111 BB and 2,550 TC cows and bulls genotyped using either the BovineSNP50 (Matukumalli et al., 2009) or the BovineHD BeadChip array (Illumina Inc., San Diego, CA). Animals that were genotyped with the lower density array had their genotypes imputed to higher density as described previously by Bolormaa et al. (2014).

The imputation was performed on 30 iterations of BEAGLE 3.3 (Browning and Browning, 2011), within breed using as reference 519 BB and 351 TC genotyped using the BovineHD chip. SNPs were mapped to the UMD3.1 bovine genome assembly (Zimin et al., 2009). The SNP positions were based on the *Bos taurus* genome assembly ARS-UCD1.2 (https://www.animalgenome.org/repository/cattle/UMC_bovine_coordinates/). After selecting autosomal SNP with minor allele frequency > 1%, we retained 651,253 SNPs for BB and 689,818 SNPs for TC.

The average, minimum, and maximum of YWT (kg) were 227.7, 115, and 353 kg for BB, and 247.07, 120.5, and 394.5 kg for TC. Moreover, the average, minimum, and maximum of age at YWT was 360, 302, and 416 days for BB, and 361, 319, and 403 days for TC.

### Genome-wide association study (GWAS) for epistasis: one locus against the polygenic background

A typical model is $y = \ldots + Wu + \ldots$ where $u$ is estimated as a vector with additive polygenic effects. Let $Var(u) = G\sigma_u^2$ and assume that there is an epistatic deviation quantitative trait loci (**QTL**) at position $i$ with statistical (not biologically functional) effect $(\alpha\alpha)_i$ and that the epistasis is against the polygenic background. A model for total genotypic value is: $g = u + i = u + (\alpha\alpha)_i z_i \odot u$ (Jannink, 2007), where $z_i$ is a centered vector with $\{2 - 2p, \ 1 - 2p, \ 0 - 2p\}$ for genotypes $\{aa, Aa, AA\}$, and the symbol $\odot$ indicates de Hadamard product.

Equivalently, $g = u + (\alpha\alpha)_i Z_i^* u$, where $Z_i^*$ is a matrix whose diagonal contains the coding of the different genotypes at locus i. Thus, $(\alpha\alpha)_i$ can be seen as the regression of the remaining genetic value once the polygenic additive effect $u$ has been removed from $g$.

Imagine for instance the epistatic effect is $(\alpha\alpha)_i = -0.2$ and $p = freq(a) = 0.6$. For an individual with $u = 20$ and carrier of $aa$ genotype, the epistatic effect is negative: $(\alpha\alpha)_i z_i u = -0.2 \ \times \ 0.8 \ \times \ 20 = -3.2$, $2 - 2p = 0.8$, and the total genetic value is $g = 20 \ - \ 3.2 = 16.8$. Similarly, for an individual with $u = 0$, the epistatic QTL has no effect; for an individual with $u = -20$, the epistatic effect is positive.

### Mechanical heuristics

In layman terms, our proposed approach proceeds in five main steps as follows (Reverter et al., 2018):

(1) Rank individuals from lowest to highest genomic estimated breeding value (**GEBV**);

(2) Divide the ranked list into five equally sized bins with BIN1 containing the 20% of individuals with the lowest GEBVs, BIN2 the next 20% of individuals based on GEBVs, and so on until BIN5 containing the 20% individuals with the highest GEBVs;

(3) Perform a GWAS of SNPs on phenotypes, within bin, and with the whole population. Note that because relationships are considered when estimating GEBVs, bins should contain similar individuals;

(4) Collect SNPs with significant yet opposed effect in BIN1 and BIN5 and a monotonic pattern of effect from BIN1 to BIN5 (e.g., strong positive, mild positive, zero, mild negative, and strong negative);

(5) Confirm the SNPs collected are not significant in the GWAS with the whole population.

The interpretation of the heuristic is that we try to find SNPs with significant and opposing effect in the extreme populations,

but that are not significant in the current population when analyzed as a whole.

## Fast approximate numerical method

The quantity of interest is the regression of $y$ on $Z_i^* u$, which can be approximated as follows:

(1) Run a GBLUP model with additive effect;
(2) Extract residuals $\hat{e}$ and GEBVs $\hat{u}$ from the output;
(3) For each SNP marker i:

    a) Multiply $\hat{u}$ by centered gene contents to obtain $Z_i^* \hat{u}$
    b) Run a single marker regression $\hat{e} = 1\mu + (\alpha\alpha)_i Z_i^* \hat{u} + \epsilon$ to estimate $(\alpha\alpha)_i$
    c) Obtain a $t$-test and associated P-value from the output.

This approximate method is very fast, but ignores the uncertainty in the estimation of $\hat{e}$ and $\hat{u}$, and therefore it will be accurate for informative datasets—that is, for moderate to highly heritable traits or large datasets. It may be used for a fast screening followed by a restricted maximum likelihood (**REML**) analysis (Jannink, 2007; Crawford et al., 2017) for a subset.

## Exact maximum likelihood method

A GWAS model would be:

$$y = Xb + u + k_i Z_i^* u + e \tag{1}$$

Where $y$ are quantitative phenotypes of interest, $b$ is the fixed effects vector (e.g., herd-sex-year contemporary group), $X$ is a design matrix relating records to fixed effects, $Z_i^*$ is a diagonal matrix whose diagonal contains $z_i$, the coding of the different genotypes at locus i, and $k_i$ is the buffering effect of locus i, expressed as a regression of phenotype on the product gene content at i multiplied the additive polygenic effect. Note that the term $k_i Z_i^* u = k_i \sum_{j=1,n} z_i \odot z_j \alpha_j$ is equivalent to equation 1 in Crawford et al. (2017) and their terms $\alpha_j$ are equivalent to our terms $k_i \alpha_j$. However, they do not present their model in terms of buffering, and their matrices $x$ ($z$ in our notation) are not centered, which leads to a lack of orthogonality of their model (Alvarez-Castro and Carlborg, 2007; Vitezica et al., 2017).

Our model in equation 1 is not usable because both the terms $k_i$ and $u$ implied in the regression are unknown. However, the epistatic component $\phi = k_i Z_i^* u$ defines a covariance matrix for $\phi$ in the i-th locus:

$$Cov(\phi) = Z_i^* G \sigma_u^2 k_i^2 Z_i^{*'}$$

which suggests a linear model with the form $y = Xb + u + \phi + e$, with covariance as above (Jannink, 2007; Crawford et al., 2017). Unfortunately, GWAS tests with this formulation imply computing and inverting $Cov(\phi)$ matrix at each locus (which is computing intensive) and can result in lack of convergence (Crawford et al., 2017). We instead propose an equivalent formulation that uses:

$$g = u + \phi = u + Z_i^* u^i$$

where $u^i = k_i u$. We have, therefore, defined a random effect, $u^i$, which multiplies real values given by the covariable $Z_i^*$. This is known as a reaction norm or random regression model (Laird and Ware, 1982; Schaeffer, 2004). Using this formulation, there

are two additive genetic traits in this model: a general additive trait $u$ with variance $G\sigma_u^2$, where $G$ is a relationship matrix (Wright, 1922; VanRaden, 2008) and a transformation of the buffering action of locus i into another additive trait: $u^i = uk_i$, $Var(u^i) = G\sigma_u^2 k_i^2$. The joint covariance matrix is:

$$Var\begin{pmatrix} u \\ u^i \end{pmatrix} = \begin{pmatrix} \sigma_u^2 & k_i \sigma_u^2 \\ k_i \sigma_u^2 & k_i^2 \sigma_u^2 \end{pmatrix} \otimes G = G_0 \otimes G$$

where $\begin{pmatrix} \sigma_u^2 & k_i \sigma_u^2 \\ k_i \sigma_u^2 & k_i^2 \sigma_u^2 \end{pmatrix} = G_0$ is a non-full rank matrix because $u^i = uk_i$, and $\otimes$ indicates the Kronecker product. The final linear model, considering the epistatic interaction of locus i with all other loci, is:

$$y = Xb + u + Z_i^* u^i + e$$

with $Var\begin{pmatrix} u \\ u^i \end{pmatrix} = \begin{pmatrix} \sigma_u^2 & k_i \sigma_u^2 \\ k_i \sigma_u^2 & k_i^2 \sigma_u^2 \end{pmatrix} \otimes G = G_0 \otimes G$. This model can be used in an exact method as described below.

The exact method proceeds by likelihood ratio test of the two alternative hypothesis, using random regression with the Model $H_1$:

$$y = Xb + z_i \alpha_i + u + Z_i^* u^i + e \tag{2}$$

and a simpler model excluding random regression with the Model $H_0$:

$$y = Xb + z_i \alpha_i + u + e \tag{3}$$

Where the actual parameter being tested is $k_i^2 \sigma_u^2 \neq 0$. The regression on gene content $z_i \alpha_i$ corrects for eventual statistical additive (not epistatic) effects of locus i. Parameters are estimated by REML (Patterson and Thompson, 1971).

After fitting the two models, the likelihood ratio test of the competing models is distributed as a mixture of 0 and 1 degrees of freedom chi-square, from which P-values can be obtained. In addition, from the estimated covariance matrix $G_0$ estimated under $H_1$, the estimated buffering epistatic effect can be obtained as:

$$\hat{k}_i = \hat{G}_0[2,2]/\hat{G}_0[1,2] \tag{4}$$

Contrary to Crawford et al. (2017), the matrix $G$ has to be computed and inverted only once, because the inclusion of the i-th locus has zero influence on the result (Gianola et al., 2016) and the matrix $Z_i^* G \sigma_u^2 Z_i^{*'}$ is never explicitly computed. This results in great savings of computing time. Matrix $G_0$ is a non-full rank matrix, which slows down convergence. An easy solution is to use a reduced rank model fitting one principal component (Meyer and Kirkpatrick, 2005) as implemented in Wombat (Meyer, 2007). Convergence takes a few iterations in this case compared with hundreds using a standard REML algorithm.

## False discovery rate

Following Bolormaa et al. (2014) and with equivalent original derivations from Storey (2002), false discovery rate (**FDR**) was calculated as:

$$FDR = \frac{P\left(1 - \frac{A}{T}\right)}{\left(\frac{A}{T}\right)(1 - P)}$$

Where $P$ is the $P$-value tested, $A$ is the number of SNPs that were significant at the $P$-value tested, and $T$ is the total number of SNP tested ($T$ = 651,253 and 689,818 for BB and TC, respectively).

### Analytical implementation

We implemented the analyses using shell scripts to manipulate the data, Wombat (Meyer, 2007) for the REML analyses (see Supplementary Appendix), and our own programs for the fast approximate analyses. Scripts and source code are available from the corresponding author upon request. The GWAS was performed in a high-performance cluster computer (Genotoul Toulouse bioinfo platform), in which the epistatic effect of each marker ($k_i$) was estimated separately from each other. Wall clock computing time was approximately 4 d for the exact analysis run in parallel and a few minutes for the fast approximate one.

## Results and Discussions

Figure 1 shows the steps of the mechanical heuristics proposed to identify SNPs with buffering epistatic effect. For the BB population, the distribution of GEBV averaged $4.26 \times 10^{-7}$ kg and ranged from –43.08 to 41.12 kg (Figure 1A). Similarly, for the TC population, the distribution of GEBV averaged $–6.27 \times 10^{-7}$ kg and ranged from –58.50 to 57.53 kg (Figure 1A). We split the animals into five equally sized bins according to their GEBV (Figure 1B). Using a nominal $P < 0.05$ from the GWAS within the extreme bins (BIN1 and BIN5) with opposite effect sign, plus a monotonic pattern of effect across bins as well as a $P > 0.10$ in the GWAS using the whole data, we found 243 and 143 epistatic SNPs in the BB and TC population, respectively (Figure 1C).

For each population, six of these SNPs, including three of each pattern (positive to negative, and negative to positive), are listed in Table 1. Table 1 also lists the effect of an SNP in the coding regions of *PLAG1* (*pleomorphic adenoma gene 1*), a well-known locus affecting stature in humans (Gudbjartsson et al., 2008), as well as growth and fertility traits in cattle (Karim et al., 2011; Fortes et al., 2013). The SNP on *PLAG1* was found to be significant only in the GWAS of the middle bins (BIN3 and BIN2 in BB and TC, respectively) as well as the GWAS of the whole dataset in both the populations.

Among the genes listed in Table 1, we highlight *LRIG3* (Leucine-rich repeats and immunoglobulin-line domains protein 3), a body size-related gene found to be under positive
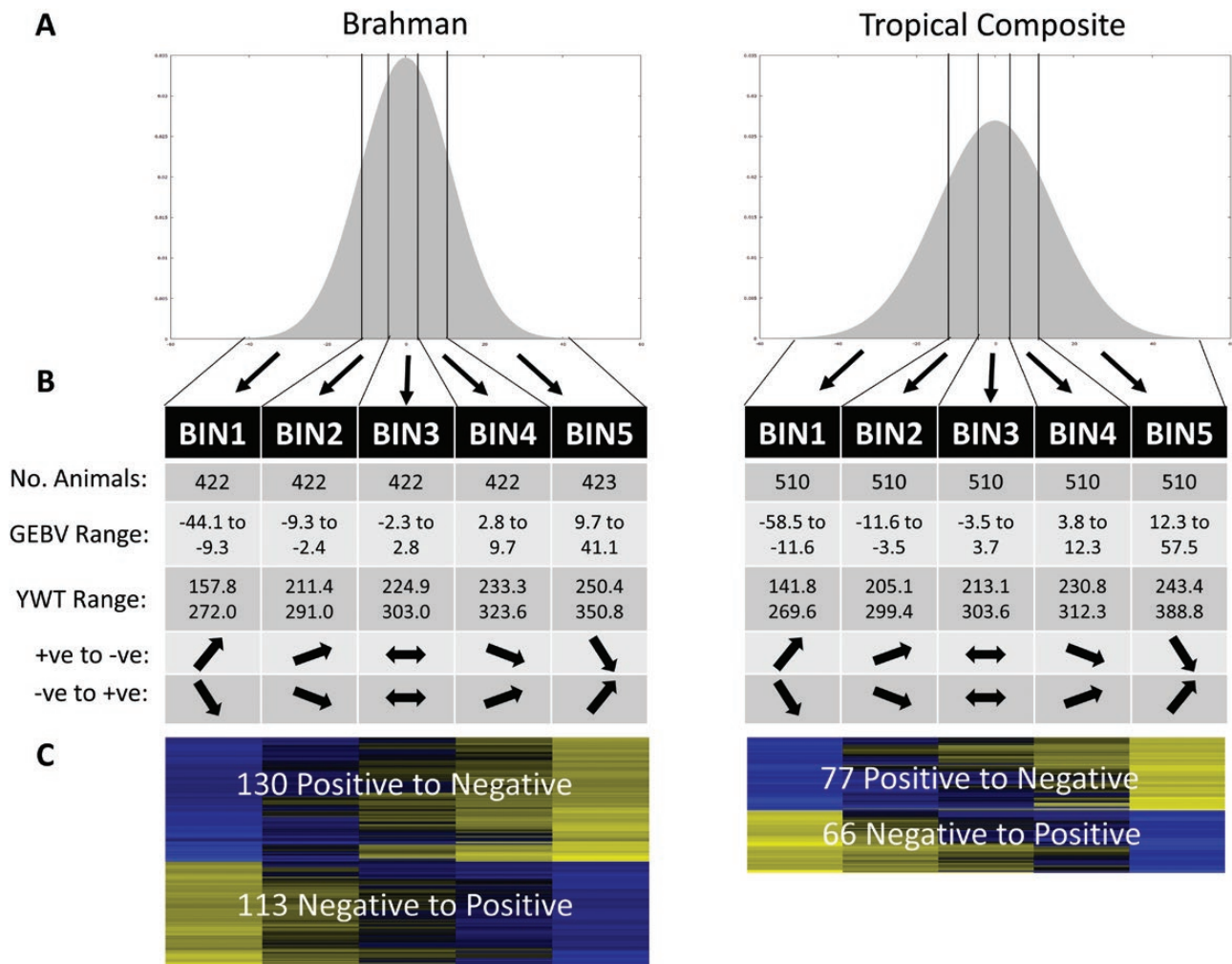


**Figure 1.** Schematic flowchart of the mechanical heuristic to identify buffering epistatic SNP: (A) Distribution of yearling weight GEBV for 2,111 BB and 2,550 TC cattle with five equally sized bins clearly demarked, BIN1 to BIN5; (B) Across bins the range of GEBV are by construct nonoverlapping, but the range of phenotypes overlap across bins. A GWAS of SNP genotype on phenotype is performed with the intention to capture SNPs with significant yet opposed additive effects in BIN1 and BIN5, and with a monotonic pattern of effects across bins; (C) We find 243 and 143 buffering epistatic SNPs in the BB and TC population, respectively.

**Table 1.** Mechanical heuristics: estimated SNP effects in *t*-statistic units (estimated effect divided by standard error) in the GWAS within BINs and across the whole population: three examples each of "Negative to Positive" and "Positive to Negative" pattern as well as for a SNP in the PLAG1 coding region for BB and TC populations.

| SNP Chr:Mb (Gene) | BIN1 | BIN2 | BIN3 | BIN4 | BIN5 | Whole |
|---|---|---|---|---|---|---|
| | | | Brahman | | | |
| 18:56.5 (CPT1C) | −7.58* | −1.38 | −0.86 | 2.67 | 4.84* | 0.60 |
| 23:50.0 (PRPF4B) | −4.50* | 0.91 | 1.39 | 2.14 | 3.58* | 1.28 |
| 28:23.3 (CTNNA3) | −8.00* | −3.05 | −1.93 | 1.55 | 5.06* | −0.05 |
| 4:71.4 (OSBPL3) | 5.95* | 1.45 | −0.50 | −1.89 | −6.01* | 0.29 |
| 5:54.9 (LRIG3) | 4.86* | −0.30 | −0.79 | −2.05 | −3.94* | 1.18 |
| 27:1.1 (CSMD1) | 4.53* | 1.18 | −0.66 | −0.69 | −2.37* | 0.58 |
| 14:25.0 (PLAG1) | 0.74 | 2.07 | 3.46* | 2.15 | 2.02 | 4.71* |
| | | | Tropical Composite | | | |
| 14:84.3 (SNTB1) | −4.21* | 0.25 | 1.69 | 1.93 | 3.42* | 0.64 |
| 16:28.5 (CNIH3) | −3.39* | −2.00 | 0.65 | 1.10 | 2.74* | 0.52 |
| 23:15.3 (FOXP4) | −6.20* | −1.94 | −0.76 | 1.87 | 5.14* | −0.26 |
| 1:5.39 (GRIK1) | 2.90* | 0.74 | 0.16 | −0.83 | −2.98* | 1.58 |
| 10:75.6 (KCNH5) | 3.56* | 2.07 | 1.00 | −0.12 | −3.63* | −0.54 |
| 22:57.3 (PPARG) | 3.49* | 1.12 | −0.09 | −2.66* | −2.81* | 1.01 |
| 14:25.0 (PLAG1) | −0.08 | 2.83* | 2.47 | 1.02 | 0.44 | 6.16* |

*indicates significance at $P < 0.001$.

**Table 2.** Number of significant buffering epistatic SNP (*N*) and false discovery rate (FDR) at decreasing levels of *P*-value for the BB and TC populations

| P-value | Brahman | | Tropical Composite | |
|---|---|---|---|---|
| | N | FDR, % | N | FDR, % |
| < 0.05 | 66,789 | 46.06 | 87,545 | 36.21 |
| < 0.01 | 20,542 | 31.01 | 30,803 | 21.61 |
| < 0.005 | 12,503 | 25.67 | 19,831 | 16.98 |
| < 0.001 | 3,679 | 17.62 | 6,972 | 9.80 |
| <0.0005 | 2,186 | 14.85 | 4,484 | 7.64 |
| <0.0001 | 662 | 9.83 | 1,571 | 4.38 |
| <0.00005 | 384 | 8.47 | 1,028 | 3.35 |
| <0.00001 | 91 | 7.15 | 342 | 2.02 |

selection in a study of five bovine breeds including BB (Xu et al., 2015). This finding is of most relevance because genes found under selection in a breed comparison study are bound to have little variation in their coding region and/or no additive effect in any given breed, and are only identified as relevant, such as harboring signatures of selection, in a multi-breed comparison.

Table 2 presents the number of significant SNP and FDR at various *P*-value thresholds. At any given *P*-value, the number of significant SNP was lower in BB than in TC. As a result, the FDR was lower in TC than in BB for a given *P*-value. For instance, at *P*-value < 0.0001, the FDR was 9.83% and 4.38% for BB and TC, respectively. The higher number of epistatic SNPs identified in the TC compared with the BB population was attributed to the distinct allele frequencies observed in the two populations. Across all SNPs, the first (reference) allele was found to be either mostly absent (reference allele frequency near 0) or nearly fixated (reference allele frequency near 1) in the BB population, while intermediate allele frequencies (highly polymorphic and hence more informative SNP) were predominant in the TC population (Figure 2 in (Reverter et al. 2017)). In other words, the distribution of allele frequencies in BB are U-shaped.

Figure 2 illustrates the two-way relationship between the mechanical heuristics and the exact test based on bivariate random regression REML analyses (Figure 2A and C), and between the numerical approximation based on the regression of residuals on GEBV and the exact text (Figure 2B and D). In the BB population, we observed a correlation of $r = 0.728$ (Figure 2A) between SNP epistatic effects according to the mechanical heuristics and the exact test, while a correlation of $r = 0.979$ (Figure 2B) was found between the approximation and the exact test. In the TC population, the same pair of values were $r = 0.805$ (Figure 2C) and $r = 0.990$ (Figure 2D). These results corroborate the superiority of the numerical approximation compared with the mechanical heuristics, with the latter being affected by arbitrary parameters, including the number of bins and the criteria for the identification of epistatic SNPs (e.g., monotonic pattern of effects from the first to the last bin).

The performance of the mechanical heuristics could be further affected if only GEBV with an accuracy beyond a nominal threshold were to be employed. Also, issues related to the sensitivity of the mechanical heuristics to GEBV accuracies, number of bins and relatedness of individuals within bin have not been considered with any degree of detail. Obviously, the minimum number of bins would be two (for negative and positive GEBVs), but that would likely result in many false positives. On the other extreme, a large number of bins can only be considered with large sample sizes so that a within-bin GWAS can be undertaken with sufficient statistical power. At this stage, we acknowledge that the mechanical heuristics lacks theoretical formality and its value is limited to provide a graphical understanding of buffering epistasis.

Figure 3 shows the Manhattan plots for the GWAS for additive and epistatic effects in the BB and TC populations. The most likely candidate genes in the most significantly associated regions are also given in Figure 3. It can already be noted that additive and epistatic gene effects are mutually orthogonal.

In the BB population, the strongest significance for epistatic effect corresponded to SNP BovineHD1700017822 mapped to 60,216,894 bp of BTA17 at 11,627 bp of the coding region of **TBX5** (T-box 5 transcription factor) and with an estimated epistatic
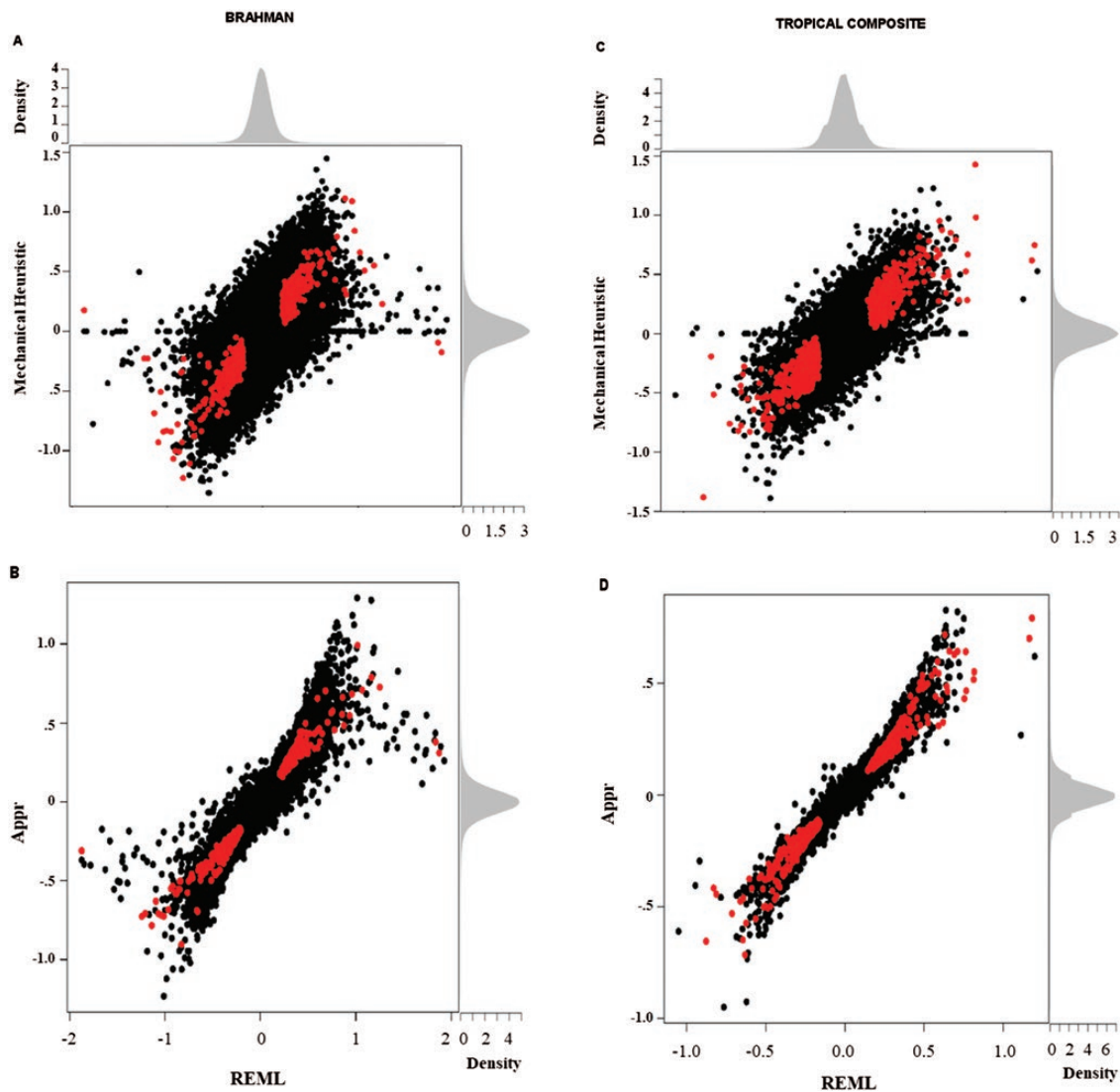
**Figure 2.** Methods comparison: relationship between the SNP buffering epistatic effects estimated via the bivariate REML analyses and the mechanical heuristic (A and C), and between the bivariate REML analyses and the approximation based on the regression of residuals on GEBV (B and D) for the Brahman (A and B) and Tropical Composite (C and D) populations. Highlighted red are the significant SNP (P < 0.0001) according to REML.

effect of −0.533 ($-\mathrm{Log}_{10}$(P-value) = 6.849). The corresponding human chromosome segment is involved in ulnar–mammary syndrome (Klopocki et al., 2006), and a recent large meta-GWAS study reveals *TBX5* as a candidate gene for mammary gland morphology in Fleckvieh cattle (Pausch et al., 2016).

Following *TBX5*, we found the second strongest signal for buffering epistasis in BB to SNP BovineHD2000011094 (estimated epistatic effect of −0.282 and $-\mathrm{Log}_{10}$(P-value) = 6.669) mapped to 38.97 Mb of BTA20 and 2.6 kb downstream of prolactin receptor (**PRLR**). *PRLR* is in a region captured by selection signatures for adaptation in beef cattle (Boitard et al., 2016) and mutations on this gene have been found to have a major genetic effect on hair length and coat structure characteristics of cattle (Littlejohn et al., 2014; Porto-Neto et al., 2018).

The third strongest signal corresponded to SNP BovineHD2300014569 mapped to 50.10 Mb of BTA23 in the coding region of **PRPF4B** (*pre-mRNA processing factor 4B*) with an estimated effect of 0.324 ($-\mathrm{Log}_{10}$(P-value) = 6.309). With no reported function in the context of bovine breeding and genetics, *PRPF4B* is an essential kinase induced by estrogen (Lahsaee et al.,

2016) and its loss promotes sustained growth factor signaling (Corkery et al., 2018). Quite strikingly, loci on the coding region of *SPEN* (SNP BovineHD1600014616, epistatic effect = 0.213, $-\mathrm{Log}_{10}$(P-value) = 2.137) and *GHR* (SNP BovineHD2000009203, epistatic effect = 0.782, $-\mathrm{Log}_{10}$(P-value) = 2.234) were found to be significantly epistatic in our study. *SPEN* is an estrogen receptor cofactor and a key regulator of fat deposition and energy balance (Hazegh et al., 2017). Furthermore, an SNP-based co-association gene network by our group previously identified *ESRRG* and *PPARG* as key regulators of age at puberty in BB cows (Fortes et al., 2013).

In the TC population, we found the strongest signal in SNP BovineHD0100028404 (epistatic effect = 0.294, $-\mathrm{Log}_{10}$(P-value) = 9.261) mapped to 98.71 Mb of BTA1 in the coding region of *LOC100139843* (*mCG140927-like*) with limited information known about its function, but quite strikingly, recently reported to be associated with age at puberty in Angus bulls (Fernández et al., 2016). We found the second and third strongest signal in the coding region of *ZNF521* (SNP BovineHD2400008618 mapped to BTA24:31,439,030 with an estimated epistatic effect = −0.260,
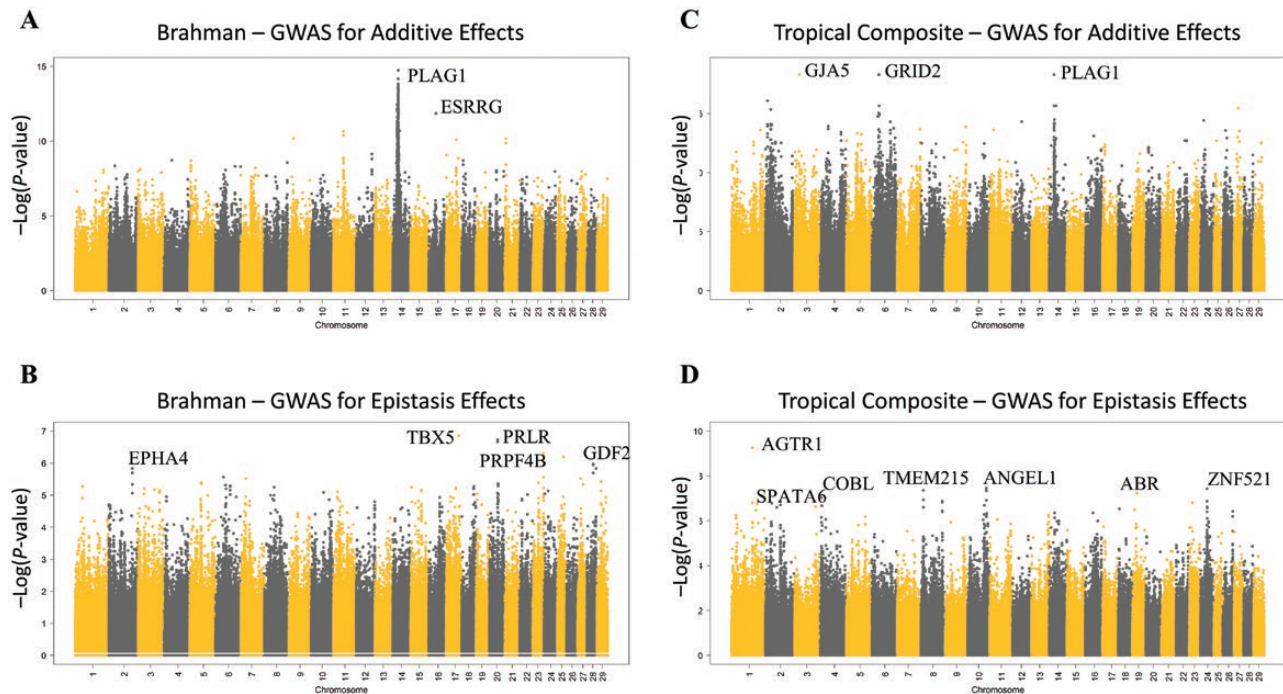
**Figure 3.** Genome-wide additive and epistasis association: Manhattan plots of the additive (A and C) and buffering epistatic (B and D) association of SNPs across the 29 bovine autosomal chromosomes for the Brahman (A and B) and Tropical Composite (C and D) populations. The most likely candidate genes in the most significantly associated regions are annotated where an obvious candidate could be identified according to the bovine reference genome assembly ARS-UCS1.2. SNPs on odd-numbered chromosomes are in black and those on even-numbered chromosomes are in yellow.

$-\text{Log}_{10}$(P-value) = 7.432) and *AGTR1* (SNP BovineHD0100034098 mapped to BTA1: 119,483,491 with an estimated epistatic effect = −0.194, $-\text{Log}_{10}$(P-value) = 3.204), respectively. The loci on *ZNF521* has been found to associate with female fertility in Nordic Red cattle, consisting of three different populations from Finland, Sweden, and Denmark (Höglund et al., 2015). Whereas the role in bovine fertility of *AGTR1* (*angiotensin II receptor type 1*) has long been documented (Portela et al., 2008; Marey et al., 2016) including its differential expression at the level of the oviduct between *Bos taurus* and *Bos indicus* cattle (Fontes et al., 2018).

The relationship between epistatic and additive effect in each population is illustrated in Figure 4. It can be seen that effects are empirically orthogonal as expected. Note that unlike additive effects, epistatic effects have no units as they are defined as a multiplier on additive effects. Figure 5 shows the relationship between the epistatic effects in both populations, BB and TC. Significant simultaneously in both populations were 42 SNPs of which 24 where located within 50 kb of the coding region of known genes and these are listed in Table 3. Porto-Neto et al. (2014) showed that Linkage disequilibrium dropped below 0.2 at distances of 50 kb.

Among those listed in Table 3, prominent genes for their reported role in mammalian fertility including bovine are: *ALDH1A1* (*aldehyde dehydrogenase 1 family member A1*), *PCSK5* (*proprotein convertase subtilisin/kexin type 5*), and *TSHR* (*thyroid-stimulating hormone receptor*), and *IL21* (*Interleukin-21*).

The role of *ALDH1A1* during bovine ovarian development has recently been established (Hatzirodos et al., 2019; Hummitzsch et al., 2019). Antenos et al. (2011) reported the role of PCSK5 in mouse ovarian follicle development. Similarly, *TSHR* is a well-known regulator of growth, fat metabolism, and

fertility. Dias et al. (2017) identified a candidate QTL in *TSHR* affecting puberty in five cattle breeds across the taurine and Indicine lineages: Brangus, BB, Nellore, Angus, and Holstein. Also, one of the most prominent selective sweeps found in all domestic chickens occurred at the locus for *TSHR* (Rubin et al., 2010). Finally, the immune system response gene *IL21* has been shown to harbor selection signatures among divergently selected subpopulations of Polish Red cattle (Gurgul et al., 2019), and among goats and sheep indigenous to a hot arid environment (Kim et al., 2016).

The existence of buffering mechanisms is an emerging propriety of networks (Mackay, 2014), and, therefore, because biochemical and gene networks are pervasive in nature, buffering loci must exist. Moreover, the existence of segregating (not fixed) buffering epistatic loci would explain several phenomena that are not well understood: environmental robustness, release of additive variance after disturbing events, (Visser et al., 2003; Flatt, 2005), maintenance of genetic variance in selected populations previously under stabilizing selection (Gimelfarb, 1989), and opposite signs of GWAS associations in different populations (Huang et al., 2012).

A simulated dataset would assist in understanding the implications of the analytical approach presented here, but we argue about the difficulty of simulating a dataset in a realistic fashion. Jannink (2007) simulated a trait with a genetic variance of 60 and a phenotypic variance of 100. An epistatic variance of 20 was generated with the marked QTL interacted with nine other unlinked loci based on a compound epistatic network proposed in the context of plant breeding (Cooper et al., 2002). Further research is warranted to understand how and if this approach could work to simulate buffering epistasis by which a set of QTLs have opposing effects on the extremes of the additive
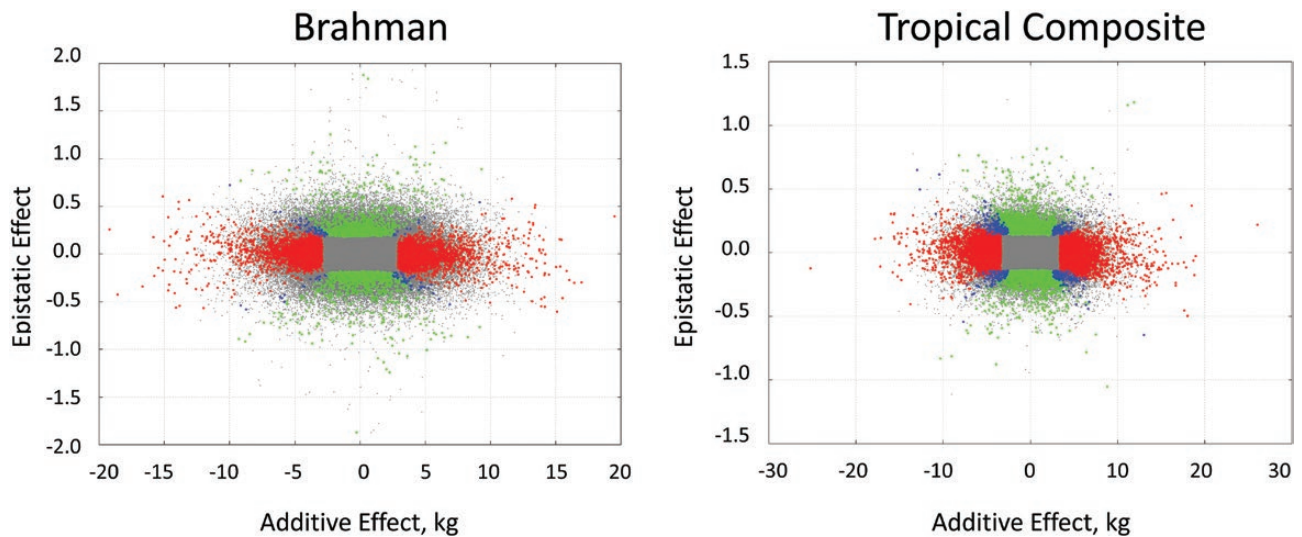
**Figure 4.** Scatter plot of the relationship between SNP additive (x-axis) and buffering epistatic effects (y-axis) for the Brahman (left panel) and Tropical Composite (right) populations. Red, green, and blue indicate significance (P-value < 0.001) for additivity, epistatic and both, respectively.
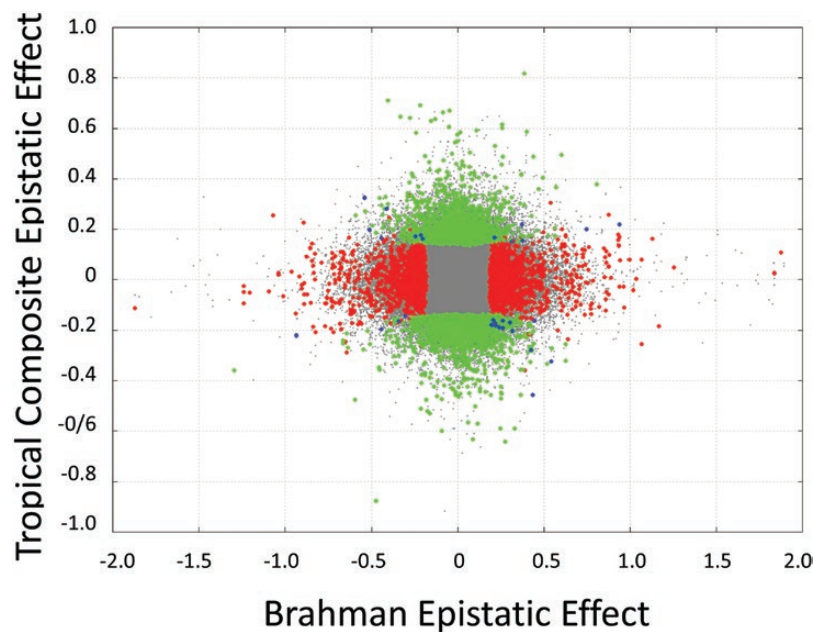


**Figure 5.** Scatter plot of the relationship between SNP buffering epistatic effect in the Brahman (x-axis) and Tropical Composite (y-axis) populations. Red, green, and blue indicate significance (P-value < 0.001) in the Brahman, Tropical Composite, and both populations, respectively.

polygenic value and in the presence of selection. Importantly, in his simulations, Jannink (2007) observed the conversion of additive-by-additive epistasis to additive variance which is what we hypothesize based on our mechanical heuristics and the empirical results.

In multi-generation selection programs, the continued response has been seen, classically for over 100 generations in the Illinois maize kernel content lines [38], and there have been large and still continuing genetic improvements in livestock populations, notably in broiler chickens [39]. Whilst the obvious source of continued response is de novo mutation, some of the additive variation being utilized may have derived from existing mutations whose behavior changes from epistatic to

additive in response to changes in the remainder of the genome. Paixão and Barton (2016) argued that "epistasis sustains additive genetic variance for longer: Alleles that were initially deleterious or near-neutral may acquire favorable effects as the genetic background changes, 'converting' epistatic variance into additive, and so prolonging the response to selection." Similarly, Hill (2017) concluded that "It seems better to concentrate on utilizing additive variance, and hope for a bonus from converting epistatic variance." These assessments beg the questions: 1) how would one go about strategically increasing this conversion? and 2) are some functional groups and/or genomic regions more likely to convert than others? We hypothesize the epistatic SNPs found here are in effect "dormant" with an additive effect

**Table 3.** Annotation of SNP with significant (*P* < 0.001) buffering epistatic effect in both populations (BB and TC) including genome position, distance to nearest gene, gene (according to the bovine reference genome assembly ARS-UCS1.2), reference allele frequency (RAF), estimated effect, and significance (–Log(*P*))

| SNP | Chr | Bp | Distance | Gene | Brahman | | | Tropical Composite | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | RAF | Effect | –Log(P) | RAF | Effect | –Log(P) |
| BovineHD0200001651 | 2 | 5,747,611 | 38,487 | NEMP2 | 0.0263 | –0.5406 | 3.0404 | 0.0729 | 0.3258 | 3.5760 |
| BovineHD0200001652 | 2 | 5,748,515 | 37,583 | NEMP2 | 0.0263 | –0.5406 | 3.0404 | 0.0727 | 0.3237 | 3.5065 |
| BovineHD0200001654 | 2 | 5,757,112 | 28,986 | NEMP2 | 0.9737 | 0.5433 | 3.0404 | 0.9272 | –0.3234 | 3.5297 |
| ARS-BFGL-NGS-100083 | 8 | 11,133,497 | 0 | SCARA3 | 0.9528 | –0.5113 | 3.4369 | 0.6920 | 0.1975 | 4.1983 |
| BovineHD0800014800 | 8 | 49,143,207 | 36,501 | ALDH1A1 | 0.9820 | 0.7474 | 3.5992 | 0.7831 | 0.2002 | 3.9915 |
| BovineHD0800015819 | 8 | 52,427,816 | 8,480 | PCSK5 | 0.0452 | –0.4416 | 3.3439 | 0.2349 | –0.1963 | 3.6686 |
| Hapmap48089-BTA-81187 | 8 | 52,524,895 | 0 | GCNT1 | 0.9403 | 0.3760 | 3.0638 | 0.6829 | 0.1562 | 3.2041 |
| BovineHD1000026980 | 10 | 92,312,649 | 0 | TSHR | 0.6054 | 0.2160 | 3.5761 | 0.4686 | –0.1712 | 4.0605 |
| BovineHD1000026985 | 10 | 92,331,573 | 0 | TSHR | 0.3946 | –0.2140 | 3.5761 | 0.5317 | 0.1779 | 4.1064 |
| BovineHD1000030828 | 10 | 92,335,071 | 0 | TSHR | 0.4216 | –0.2017 | 3.2507 | 0.7002 | 0.1623 | 3.2973 |
| BovineHD1200025989 | 12 | 84,830,432 | 15,785 | 5S_rRNA | 0.7141 | 0.2066 | 3.2041 | 0.6998 | –0.1603 | 3.0872 |
| BovineHD1400011237 | 14 | 37,055,691 | 0 | ENSBTAG00000002076 | 0.1033 | –0.3042 | 3.0404 | 0.5288 | –0.1442 | 3.0872 |
| BovineHD1400011280 | 14 | 37,211,716 | 0 | ENSBTAG00000050514 | 0.9095 | 0.3733 | 3.9915 | 0.6543 | 0.2204 | 5.9710 |
| BovineHD1400013234 | 14 | 44,638,868 | 5,792 | PMP2 | 0.6080 | 0.1951 | 3.0169 | 0.6512 | –0.1791 | 3.9225 |
| BovineHD1600016214 | 16 | 56,665,443 | 0 | TNR | 0.1409 | 0.2618 | 3.0872 | 0.6231 | –0.1635 | 3.6917 |
| BovineHD1700009805 | 17 | 35,070,459 | 0 | IL21 | 0.9377 | 0.4452 | 3.7841 | 0.7129 | –0.1633 | 3.2973 |
| BovineHD1700009807 | 17 | 35,079,284 | 8,519 | IL21 | 0.0623 | –0.4429 | 3.7610 | 0.2859 | 0.1659 | 3.2274 |
| BovineHD1700009808 | 17 | 35,083,771 | 13,006 | IL21 | 0.0623 | –0.4429 | 3.7610 | 0.2859 | 0.1659 | 3.2274 |
| BovineHD1700009811 | 17 | 35,094,499 | 23,734 | IL21 | 0.9377 | 0.4452 | 3.7841 | 0.7131 | –0.1626 | 3.2740 |
| BovineHD1900012991 | 19 | 45,742,542 | 18,774 | ENSBTAG00000053203 | 0.2399 | 0.2629 | 3.5297 | 0.3921 | –0.1908 | 4.5416 |
| BovineHD4100015365 | 21 | 65,988,493 | 33,080 | bta-mir-656 | 0.0559 | 0.4343 | 3.1340 | 0.0502 | –0.4564 | 5.9259 |
| BovineHD2500009619 | 25 | 34,418,988 | 0 | SSC4D | 0.1843 | 0.3028 | 3.6455 | 0.4378 | –0.1702 | 3.8764 |
| BovineHD2500009621 | 25 | 34,427,964 | 0 | SSC4D | 0.1831 | 0.3168 | 3.8764 | 0.3959 | –0.2030 | 4.9974 |
| BovineHD2800005959 | 28 | 22,571,976 | 0 | CTNNA3 | 0.1338 | 0.3147 | 3.1807 | 0.3500 | 0.1516 | 3.2973 |

waiting to be "released" when selection moves the population to either tail of the genetic value distribution. Consistent with the argument of Carlborg et al. (2006), we further argue that these SNPs provide an answer to the long-standing paradox by which genetic variation does not diminish with selection as fast as theory would anticipate, and instead epistasis is responsible for the release of genetic variation during long-term selection.

## Conclusions

We have developed three novel analytical methodologies to identify and estimate the effect of an SNP with an epistatic effect against the polygenic background. Two of the approaches require the use of previously computed GEBV, which makes them particularly suited to domestic species of agricultural importance that have adopted genomic selection for their breeding purposes.

## Supplementary Data

Supplementary data are available at *Journal of Animal Science* online.

## Acknowledgments

## Conflict of interest statement

The authors declare no real or perceived conflicts of interest.

## Literature Cited

Alvarez-Castro, J. M., and O. Carlborg. 2007. A unified model for functional and statistical epistasis and its application in quantitative trait loci analysis. *Genetics* 176:1151–1167. doi:10.1534/genetics.106.067348

Antenos, M., L. Lei, M. Xu, A. Malipatil, S. Kiesewetter, and T. K. Woodruff. 2011. Role of PCSK5 expression in mouse ovarian follicle development: identification of the inhibin α- and β-subunits as candidate substrates. *Plos One* 6:e17348. doi:10.1371/journal.pone.0017348

Boitard, S., M. Boussaha, A. Capitan, D. Rocha, and B. Servin. 2016. Uncovering adaptation from sequence data: lessons from genome resequencing of four cattle breeds. *Genetics* 203:433–450. doi:10.1534/genetics.115.181594

Bolormaa, S., J. E. Pryce, A. Reverter, Y. Zhang, W. Barendse, K. Kemper, B. Tier, K. Savin, B. J. Hayes, and M. E. Goddard. 2014. A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. *PLoS Genet.* 10:e1004198. doi:10.1371/journal.pgen.1004198

Browning, B. L., and S. R. Browning. 2011. A fast, powerful method for detecting identity by descent. *Am. J. Hum. Genet.* 88:173–182. doi:10.1016/j.ajhg.2011.01.010

Carlborg, O., L. Jacobsson, P. Ahgren, P. Siegel, and L. Andersson. 2006. Epistasis and the release of genetic variation during

long-term selection. *Nat. Genet.* **38**:418–420. doi:10.1038/ng1761

Cooper, M., K. P. Podlich, K. P. Micallef, O. S. Smith, N. M. Jensen, S. C. Chapman, and N. L. Kruger. 2002. Complexity, quantitative traits and plant breeding: a role for simulation modelling in the genetic improvement of crops. In: Kang, M. S., editor. *Quantitative genetics, genomics and plant breeding*. Wallingford, UK: CAB International; p. 143–166.

Corkery, D. P., L. E. Clarke, S. Gebremeskel, J. Salsman, J. Pinder, C. Le Page, L. Meunier, Z. Xu, A. M. Mes-Masson, J. N. Berman, *et al.* 2018. Loss of PRP4K drives anoikis resistance in part by dysregulation of epidermal growth factor receptor endosomal trafficking. *Oncogene* **37**:174–184. doi:10.1038/onc.2017.318

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. doi:10.1371/journal.pgen.1006869

Dias, M. M., A. Cánovas, C. Mantilla-Rojas, D. G. Riley, P. Luna-Nevarez, S. J. Coleman, S. E. Speidel, R. M. Enns, A. Islas-Trejo, J. F. Medrano, *et al.* 2017. SNP detection using RNA-sequences of candidate genes associated with puberty in cattle. *Genet. Mol. Res.* **16**(1):gmr16019522. doi:10.4238/gmr16019522

Fernández, M. E., A. Prando, A. Rogberg-Muñoz, P. Peral-García, A. Baldo, G. Giovambattista, and J. P. Lirón. 2016. Association of a region of bovine chromosome 1 (BTA1) with age at puberty in Angus bulls. *Reprod. Fertil. Dev.* **28**:1618–1621. doi:10.1071/RD14511

Flatt, T. 2005. The evolutionary genetics of canalization. *Q. Rev. Biol.* **80**:287–316. doi:10.1086/432265

Fontes, P. K., R. L. Ereno, A. R. Peixoto, R. F. Carvalho, W. R. Scarano, L. A. Trinca, C. M. Barros, and A. C. S. Castilho. 2018. Can the antral follicular count modulate the gene expression of bovine oviducts in Aberdeen Angus and Nelore heifers? *PLoS One* **13**:e0202017. doi:10.1371/journal.pone.0202017

Fortes, M. R., K. Kemper, S. Sasazaki, A. Reverter, J. E. Pryce, W. Barendse, R. Bunch, R. McCulloch, B. Harrison, S. Bolormaa, *et al.* 2013. Evidence for pleiotropism and recent selection in the PLAG1 region in Australian Beef cattle. *Anim. Genet.* **44**:636–647. doi:10.1111/age.12075

Gianola, D., M. I. Fariello, H. Naya, and C. C. Schön. 2016. Genome-wide association studies with a genomic relationship matrix: a Case Study with wheat and *Arabidopsis*. *G3 (Bethesda)*. **6**:3241–3256. doi:10.1534/g3.116.034256

Gimelfarb, A. 1989. Genotypic variation for a quantitative character maintained under stabilizing selection without mutations: epistasis. *Genetics* **123**:217–227.

Gudbjartsson, D. F., G. B. Walters, G. Thorleifsson, H. Stefansson, B. V. Halldorsson, P. Zusmanovich, P. Sulem, S. Thorlacius, A. Gylfason, S. Steinberg, *et al.* 2008. Many sequence variants affecting diversity of adult human height. *Nat. Genet.* **40**:609–615. doi:10.1038/ng.122

Gurgul, A., I. Jasielczuk, E. Semik-Gurgul, T. Szmatoła, A. Majewska, E. Sosin-Bzducha, and M. Bugno-Poniewierska. 2019. Diversifying selection signatures among divergently selected subpopulations of Polish Red cattle. *J. Appl. Genet.* **60**:87–95. doi:10.1007/s13353-019-00484-0

Hatzirodos, N., K. Hummitzsch, H. F. Irving-Rodgers, J. Breen, V. E. A. Perry, R. A. Anderson, and R. J. Rodgers. 2019. Transcript abundance of stromal and thecal cell related genes during bovine ovarian development. *PLoS One* **14**:e0213575. doi:10.1371/journal.pone.0213575

Hazegh, K. E., T. Nemkov, A. D'Alessandro, J. D. Diller, J. Monks, J. L. McManaman, K. L. Jones, K. C. Hansen, and T. Reis. 2017. An autonomous metabolic role for Spen. *PLoS Genet.* **13**:e1006859. doi:10.1371/journal.pgen.1006859

Hill, W. G. 2017. "Conversion" of epistatic into additive genetic variance in finite populations and possible impact on long-term selection response. *J. Anim. Breed. Genet.* **134**:196–201. doi:10.1111/jbg.12270

Höglund, J. K., B. Buitenhuis, B. Guldbrandtsen, M. S. Lund, and G. Sahana. 2015. Genome-wide association study for female fertility in Nordic Red cattle. *BMC Genet.* **16**:110. doi:10.1186/s12863-015-0269-x

Huang, W., S. Richards, M. A. Carbone, D. Zhu, R. R. Anholt, J. F. Ayroles, L. Duncan, K. W. Jordan, F. Lawrence, M. M. Magwire, *et al.* 2012. Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. *Proc. Natl. Acad. Sci. U. S. A.* **109**:15553–15559. doi:10.1073/pnas.1213423109

Hummitzsch, K., N. Hatzirodos, H. F. Irving-Rodgers, M. D. Hartanti, V. E. A. Perry, R. A. Anderson, and R. J. Rodgers. 2019. Morphometric analyses and gene expression related to germ cells, gonadal ridge epithelial-like cells and granulosa cells during development of the bovine fetal ovary. *PLoS One* **14**:e0214130. doi:10.1371/journal.pone.0214130

Jannink, J. L. 2007. Identifying quantitative trait locus by genetic background interactions in association studies. *Genetics* **176**:553–561. doi:10.1534/genetics.106.062992

Karim, L., H. Takeda, L. Lin, T. Druet, J. A. Arias, D. Baurain, N. Cambisano, S. R. Davis, F. Farnir, B. Grisart, *et al.* 2011. Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. *Nat. Genet.* **43**:405–413. doi:10.1038/ng.814

Kim, E. S., A. R. Elbeltagy, A. M. Aboul-Naga, B. Rischkowsky, B. Sayre, J. M. Mwacharo, and M. F. Rothschild. 2016. Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity (Edinb)*. **116**:255–264. doi:10.1038/hdy.2015.94

Kitano, H. 2004. Biological robustness. *Nat. Rev. Genet.* **5**:826–837. doi:10.1038/nrg1471

Klopocki, E., L. M. Neumann, H. Tönnies, H. H. Ropers, S. Mundlos, and R. Ullmann. 2006. Ulnar-mammary syndrome with dysmorphic facies and mental retardation caused by a novel 1.28 Mb deletion encompassing the TBX3 gene. *Eur. J. Hum. Genet.* **14**:1274–1279. doi:10.1038/sj.ejhg.5201696

Lahsaee, S., D. P. Corkery, L. E. Anthes, A. Holly, and G. Dellaire. 2016. Estrogen receptor alpha (ESR1)-signaling regulates the expression of the taxane-response biomarker PRP4K. *Exp. Cell Res.* **340**:125–131. doi:10.1016/j.yexcr.2015.12.013

Laird, N. M., and J. H. Ware. 1982. Random-effects models for longitudinal data. *Biometrics* **38**:963–974. doi:10.2307/2529876

Littlejohn, M. D., K. M. Henty, K. Tiplady, T. Johnson, C. Harland, T. Lopdell, R. G. Sherlock, W. Li, S. D. Lukefahr, B. C. Shanks, *et al.* 2014. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. *Nat. Commun.* **5**:5861. doi:10.1038/ncomms6861

Mackay, T. F. C. 2014. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat. Rev. Genet* **15**: 22–33. doi:10.1038/nrg3627

Marey, M. A., M. S. Yousef, J. Liu, K. Morita, M. Sasaki, H. Hayakawa, T. Shimizu, and A. Miyamoto. 2016. Angiotensin II increases sperm phagocytosis by neutrophils in vitro: a possible physiological role in the bovine oviduct. *Mol. Reprod. Dev.* **83**:630–639. doi:10.1002/mrd.22672

Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. F. Allan, M. P. Heaton, J. O'Connell, S. S. Moore, T. P. Smith, T. S. Sonstegard, *et al.* 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One* **4**:e5350. doi:10.1371/journal.pone.0005350

Meyer, K., and M. Kirkpatrick. 2005. Random regression analyses using B-splines to model growth of Australian Angus cattle. *Genet. Sel. Evol.* **37**:473–500. doi:10.1051/gse:2005012

Meyer, K. 2007. WOMBAT: a tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *J. Zhejiang Univ. Sci. B* **8**:815–821. doi:10.1631/jzus.2007.B0815

Paixão, T., and N. H. Barton. 2016. The effect of gene interactions on the long-term response to selection. *Proc. Natl. Acad. Sci. U.S.A.* **113**: 4422–4427. doi:10.1073/pnas.1518830113

Patterson, H., and R. Thompson. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**:545.

Pausch, H., R. Emmerling, H. Schwarzenbacher, and R. Fries. 2016. A multi-trait meta-analysis with imputed sequence variants reveals twelve QTL for mammary gland morphology in Fleckvieh cattle. *Genet. Sel. Evol.* **48**:14. doi:10.1186/s12711-016-0190-4

Portela, V. M., P. B. Gonçalves, A. M. Veiga, E. Nicola, J. Buratini Jr, and C. A. Price. 2008. Regulation of angiotensin type 2 receptor in bovine granulosa cells. *Endocrinology* **149**:5004–5011. doi:10.1210/en.2007-1767

Porto-Neto, L. R., D. M. Bickhart, A. J. Landaeta-Hernandez, Y. T. Utsunomiya, M. Pagan, E. Jimenez, P. J. Hansen, S. Dikmen, S. G. Schroeder, E. S. Kim, *et al*. 2018. Convergent evolution of slick coat in cattle through truncation mutations in the prolactin receptor. *Front. Genet.* **9**:57. doi:10.3389/fgene.2018.00057

Porto-Neto, L. R., J. W. Kijas, and A. Reverter. 2014. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genet. Sel. Evol.* **46**:22. doi:10.1186/1297-9686-46-22

Raidan, F. S. S., L. R. Porto-Neto, Y. Li, S. A. Lehnert, Z. G. Vitezica, and A. Reverter. 2018. Evaluation of nonadditive effects in yearling weight of tropical beef cattle. *J. Anim. Sci.* **96**:4028–4034. doi:10.1093/jas/sky275

Reverter, A., J. Henshall, L. R. Porto-Neto, F. Raidan, Y. Li, M. Naval-Sánchez, S. A. Lehnert, K. Meyer, Z. Vitezica, *et al*. 2018. A rapid method for the identification of epistatic 'dormant' SNPs. Proceedings of the 11th World Congress on Genetics Applied to Livestock Production Methods and Tools-GWAS: 260; Auckland, New Zealand, February 10 to 16, 2018. www.wcgalp.org/proceedings/2018/rapid-method-identification-epistatic-%E2%80%98dormant%E2%80%99-snps

Reverter, A., L. R. Porto-Neto, M. R. S. Fortes, P. Kasarapu, M. A. R. de Cara, H. M. Burrow, and S. A. Lehnert. 2017. Genomic inbreeding depression for climatic adaptation of tropical beef cattle. *J. Anim. Sci.* **95**:3809–3821. doi:10.2527/jas2017.1643

Rubin, C. J., M. C. Zody, J. Eriksson, J. R. Meadows, E. Sherwood, M. T. Webster, L. Jiang, M. Ingman, T. Sharpe, S. Ka, *et al*. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**:587–591. doi:10.1038/nature08832

Schaeffer, L. R. 2004. Application of random regression models in animal breeding. *Livest. Prod. Sci.* **86**:35–45. doi:10.1016/S0301-6226(03)00151-9

Storey, J. D. 2002. A direct approach to false discovery rates. *J. Royal Stat. Soc.* **64**: 479–498. doi: 10.1111/1467–9868.00346

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**:4414–4423. doi:10.3168/jds.2007-0980

Visser, J. A. G. M. de, J. Hermisson, G. P. Wagner, L. A. Meyers, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao, J. M. Cheverud, S. F. Elena, W. Fontana, *et al*. 2003. Perspective: evolution and detection of genetic robustness. *Evolution* **57**:1959–1972. doi:10.1111/j.0014-3820.2003.tb00377.x

Vitezica, Z. G., A. Legarra, M. A. Toro, and L. Varona. 2017. Orthogonal estimates of variances for additive, dominance and epistatic effects in populations. *Genetics* **116**:199406. doi:10.1534/genetics.116.199406

Wright, S. 1922. Coefficients of inbreeding and relationship. *Am. Nat.* **56**:330–338.

Xu, L., D. M. Bickhart, J. B. Cole, S. G. Schroeder, J. Song, C. P. Tassell, T. S. Sonstegard, and G. E. Liu. 2015. Genomic signatures reveal new evidences for selection of important traits in domestic cattle. *Mol. Biol. Evol.* **32**:711–725. doi:10.1093/molbev/msu333

Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. P. Van Tassell, T. S. Sonstegard, *et al*. 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* **10**:R42. doi:10.1186/gb-2009-10-4-r42