Evaluation of non-additive genetic variation in feed-related traits of broiler chickens

Y. Li,*,[1] R. Hawken,† R. Sapp,† A. George,‡ S. A. Lehnert,* J. M. Henshall,† and A. Reverter*

*CSIRO Agriculture and Food, St. Lucia, QLD 4067, Australia; †Cobb-Vantress Inc., Siloam Springs, Arkansas 72761-1030; and ‡CSIRO Data61, Dutton Park, QLD 4102, Australia

ABSTRACT Genome-wide association mapping and genomic predictions of phenotype of individuals in livestock are predominately based on the detection and estimation of additive genetic effects. Non-additive genetic effects are largely ignored. Studies in animals, plants, and humans to assess the impact of non-additive genetic effects in genetic analyses have led to differing conclusions. In this paper, we examined the consequences of including non-additive genetic effects in genome-wide association mapping and genomic prediction of total genetic values in a commercial population of 5,658 broiler chickens genotyped for 45,176 single nucleotide polymorphism (SNP) markers. We employed mixed-model equations and restricted maximum likelihood to analyze 7 feed related traits (TRT1 – TRT7). Dominance variance accounted for a significant proportion of the total genetic variance in all 7 traits, ranging from 29.5% for TRT1 to 58.4% for TRT7. Using a 5-fold cross-validation schema, we found that in spite of the large dominance component, including the estimated dominance effects in the prediction of total genetic values did not improve the accuracy of the predictions for any of the phenotypes. We offer some possible explanations for this counter-intuitive result including the possible confounding of dominance deviations with common environmental effects such as hatch, different directional effects of SNP additive and dominance variations, and the gene-gene interactions' failure to contribute to the level of variance.

Key words: non-additive genetic variation, dominance, accuracy of genomic prediction, feed-related trait, broiler

2017 Poultry Science 96:754–763
http://dx.doi.org/10.3382/ps/pew333

INTRODUCTION

The total genetic variance of a trait can be partitioned into 3 different variance components — an additive component, a dominance component, and an epistatic component (Fisher, 1918). The additive component measures the variance due to the additive effect of genes on a trait. The dominance component measures the variance due to the interaction of alleles of the same gene locus. The epistatic component measures the variance due to genes interacting across loci. The last 2 components are non-additive genetic variances. However, traditionally it is challenging to obtain accurate estimates for non-additive genetic variances using pedigree information, especially if family sizes are small as in sheep and cattle in which populations consist mainly of half-sib families, due to large sampling errors (Falconer and Mackey, 1996; Lynch and Walsh, 1998; Hill et al., 2008). Even with large full-sib family sizes as in pigs and poultry, the precision of estimation for non-additive variances is still problematic as the resemblance between relatives has a high degree of confounding with common environmental effects (Hill et al., 2008). Today, with the advent of high-throughput genotyping technologies and the availability of whole-genome single nucleotide polymorphism (**SNP**) marker data, it is feasible to calculate these non-additive variance components from marker information (Toro and Varona, 2010; Wittenburg et al., 2011).

Since Meuwissen et al. (2001) pioneered the genome-wide selection method using high-density SNP markers in early phenotype prediction, there have been large efforts in examining the impacts of parametric and non-parametric methods on the performance of predicting phenotypes that include non-additive effects. For example, parametric approaches have been used to estimate both additive and non-additive variances via SNP marker information as opposed to pedigree information (Su et al., 2012; Vitezica et al., 2013; Munoz et al., 2014; Nishio and Satoh 2014; Bolormaa et al., 2015; Wittenburg et al., 2015; Zhu et al., 2015; Aliloo et al., 2016). Non-parametric and machine learning methods such as reproducing kernel Hilbert spaces (**RKHS**) regression (González-Recio et al., 2008; Gianola et al., 2014), neural networks (Gianola et al., 2011), and radial basis function (**RBF**) regression (Long et al., 2011), Random Forests (**RF**), Bayesian LASSO (**BLASSO**), and

Received April 22, 2016.
Accepted August 11, 2016.
[1]Corresponding author: yutao.li@csiro.au

Bayesian additive regression trees (**BART**) (Azvedo et al., 2015; Waldmann 2016) have been evaluated for their predictability of phenotypes. Findings from these studies have been mixed. Munoz et al. (2014) found that in trees, marker-based estimation of the covariance structures led to more precise separation of the additive and non-additive genetic variance components. Vitezica et al. (2013), in their study of mice, showed that the individual breeding values were more accurately estimated with marker-based genomic relationship information than pedigree-based estimation. Su et al. (2012) demonstrated that in pigs, genomic prediction was improved by implementing analyses that include non-additive effects and marker-based calculation of the covariance matrices. However, Bolormaa et al. (2015) found that for growth, carcass, and fertility traits in beef cattle, there was very little improvement in prediction accuracy of phenotypic values using best linear unbiased prediction (**BLUP**) values by accounting for non-additive variation. Furthermore, Zhu et al. (2015) conducted a genome-wide association study (**GWAS**) involving 79 complex traits in humans. They also found little benefit in accounting for dominance variation in their association mapping analyses. They concluded that the missing heritability problem in humans is most likely not due to unexplained dominance variation (Zhu et al., 2015). Recently, Aliloo et al. (2016) found that including dominance effects in a genomic prediction model improved slightly the accuracy for fat yield in Holstein cows, but not for fertility traits.

To date, other than the study by González-Recio et al. (2008) that applied non-parametric methods to include non-additive dominance and epistatic effects in the genomic prediction of mortality rates in broiler chicken, there has been no other study undertaken to evaluate the impact of including non-additive genetic variation on the genomic prediction of phenotypic values in a poultry population. In this paper, we aimed a) to characterize non-additive genetic variation in genome-wide association mapping of 7 feed-related traits; and b) to examine the consequence of including non-additive genetic effects in prediction of total genetic values for a commercial poultry population. More specifically, our objective was to investigate whether there was merit in including non-additive genetic variation in a genomic prediction program of phenotypes.

MATERIALS AND METHODS

Animals, Phenotypes, and SNP Data

A total of 5,658 broiler chickens from 3 contemporary groups of a selection line in Cobb-Vantress Inc. were genotyped with a Cobb-Vantress custom-designed chip containing 52,232 SNPs. They were the progeny from 79 sires and 496 dams that formed 575 full-sib families. The population consisted of 3,979 females and 1,679 males with 7 feed-related phenotypic traits measured at less than 7 wk of age. Summary statistics for the 7 traits are given in Table 1. Due to commercial confidentiality restrictions, the 7 traits were coded as TRT1, TRT2, TRT3, TRT4, TRT5, TRT6, and TRT7, respectively. The values of mean, minimum, and maximum of each trait in Table 1 are presented in standard deviation units. All 7 traits are continuous traits that follow a normal distribution.

Table 1. Summary of basic statistics for 7 feed-related traits in the study population. N – number of animals; Mean – mean value in a standard deviation unit, Min – minimum value in a standard deviation unit, Max – maximum value in a standard deviation unit. Heterozygosity levels were raw values without standardization.

	Female (N = 3979)			Male (N = 1679)		
Trait	Mean	Min	Max	Mean	Min	Max
TRT1	10.93	7.94	13.90	10.33	7.26	13.44
TRT2	18.92	15.45	22.92	14.52	11.20	18.52
TRT3	17.67	13.68	21.28	15.22	11.42	18.25
TRT4	8.80	3.75	11.92	9.00	4.65	12.04
TRT5	6.31	2.92	9.16	6.36	2.68	9.33
TRT6	11.79	9.21	17.14	11	8.6	15.80
TRT7	−0.0034	−5.27	5.70	0.00057	−3.41	3.97
Heterozygosity	0.34	0.25	0.38	0.36	0.31	0.40

For consistency and ease of implementing in an industry practice, a non-conventional method of coding SNP genotypes was applied. Instead of treating the minor allele of each SNP as the first allele, we coded each SNP genotype based on the alphabetical order of the 4 nucleotides of genome sequence (e.g., for an "AT" SNP, "AA" was always coded as 0, "AT" as 1, and "TT" as 2).

Prior to any analyses, quality control of the SNP data was carried out. A total of 7,056 SNPs with MAF < 0.05 or calling rate $< 95\%$ were removed. Our final marker data contained genotypes from 45,176 SNPs.

Statistical Analysis

Genome-wide Association Study (GWAS) with Additive and Dominance Effects. Here, we departed from typical genome-wide association mapping in which only additive effects of SNP are included in the linear mixed model. Instead, our model included both fixed additive and dominance SNP effects. The strength of a marker-trait association was assessed for each SNP separately, using the same analysis procedure as presented in Bolormaa et al. (2015). A linear mixed model was constructed that consisted of fixed and random effects. Our linear mixed model, given observed animal phenotypes $\mathbf{y}_{(n\times1)}$, is as follows:

$$\mathbf{y} = \mathbf{1}_n\,\mu + \mathbf{Xb} + \mathbf{w1}_i\alpha_i + \mathbf{w2}_i\beta_i + \mathbf{Za} + \mathbf{e} \qquad (1)$$

where $\boldsymbol{\mu}$ is the population mean, $\mathbf{X}_{(n\times b)}$ is a design matrix, $\mathbf{b}_{(b\times1)}$ is a vector of fixed effect consisting of contemporary group and sex, $\mathbf{w1}_{i(n\times1)}$ and $\mathbf{w2}_{i(n\times1)}$ are vectors containing the additive and dominance genotype codes at the i^{th} SNP, respectively, $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ are the

scalar regression coefficients for the additive and dominance effects, respectively, $\mathbf{Z}_{(n\times n)}$ is an identity matrix, $\mathbf{a}_{(n\times 1)}$ is a vector of random animal additive (polygenic) effects, and $\mathbf{e}_{(n\times 1)}$ is a vector of errors. It is assumed that **a** and **e** follow a normal distribution with mean zero and variance $\mathbf{A}_{(n\times n)}\ \boldsymbol{\sigma}_{\mathbf{a}}^{\mathbf{2}}$ and $\mathbf{I}_{(n\times n)}\ \sigma_{\mathbf{e}}^{\mathbf{2}}$, respectively. Here, $\boldsymbol{\sigma}_{\mathbf{a}}^{\mathbf{2}}$ and $\sigma_{\mathbf{e}}^{\mathbf{2}}$ are additive genetic and error variances and **A** is the numerator relationship matrix based on pedigree information. The linear mixed models were fitted and variance components estimated using residual maximum likelihood as implemented in the stand-alone package Qxpak v5.02 (Pérez-Enciso and Misztal, 2011).

For each of the 45,176 SNPs, we conducted 3 separate GWAS as follows: (1) First, we fitted an additive model where the fixed dominance effect $\mathbf{w2}_{\mathbf{i}}\beta_{\mathbf{i}}$was dropped from the model; (2) second, we fitted a dominance model where the fixed additive effect $\mathbf{w1}_{\mathbf{i}}\alpha_{\mathbf{i}}$was dropped from the model; and (3) third, we fitted a joint additive dominance model where both $\mathbf{w1}_{\mathbf{i}}\alpha_{\mathbf{i}}$ and $\mathbf{w2}_{\mathbf{i}}\beta_{\mathbf{i}}$ were retained in the model. For our analyses, we assumed a threshold value of $P < 10^{-4}$ at which the association between a SNP and a trait was declared significant. For each model, the false discovery rate (**FDR**) of SNP effects was calculated as in Bolormaa et al. (2011). That is, FDR = $P^*(1\text{-}Q)/[Q^*(1\text{-}P)]$, where P is the significance threshold value (e.g., $P = 10^{-4}$) and Q is the fraction of SNPs detected to the total number of SNPs genotyped. The R function Chi.test() in R program (version 3.1.2) was applied to test the imbalance of the number of SNPs with positive effects versus the number of SNPs with negative effects in each GWAS model.

Joint Estimation of Additive and Dominance Variance. We also performed analyses to examine the benefits of partitioning the genetic variance into its additive and dominance components. We used the same linear mixed model as described by Bolormaa et al. (2015). Our linear mixed model is as follows:

$$\mathbf{y} = \mathbf{1}_{\mathbf{n}}\,\mu + \mathbf{Xb} + \mathbf{h}\beta + \mathbf{Za} + \mathbf{Zd} + \mathbf{e} \qquad (2)$$

where the term **Xb** is the same as the model for GWAS (Eq. 1), $\mathbf{h}_{(n\times 1)}$ is a vector containing the average heterozygosity across all SNPs for each animal, $\boldsymbol{\beta}$ is the heterozygosity regression coefficient for the trait, the term **Za** is as defined for the model for GWAS (**Z** is an identity matrix), $\mathbf{d}_{(n\times 1)}$ is a vector of random dominance effects, and **e** is a vector of random errors. The distributions of the random effects **a** and **e** have been defined above, except that here, **a** has variance matrix $\sigma_{\mathbf{a}}^2 GRM$ where GRM is a genomic relationship matrix with its values calculated from the SNP information. For the random dominance effect **d**, we assume it follows a normal distribution with mean zero and variance matrix $\sigma_d^2 DRM$, where σ_d^2 is the dominance variance parameter and DRM is a dominance relationship matrix with its values estimated from the SNP information.

Several different marker-based approaches to calculating the elements of the covariance matrices GRM and DRM have been developed (e.g., Vitezica et al., 2013; Munoz et al., 2014; and Zhu et al., 2015). In this paper, we adopted the approach of Bolormaa et al. (2015). In brief, GRM and DRM are calculated as $GRM = T\,T^t/m$ and $DRM = H\,H^t/m$, where $\mathrm{T}_{(n\times m)}$ and $\mathrm{H}_{(n\times m)}$ are matrices containing the additive and dominance contributions for the n animals and m SNP loci with $T_{ij} \in T$ and $H_{ij} \in H$. For j^{th} animal at i^{th} SNP locus, the additive contribution $\mathrm{T_{ij}}$ is calculated as $-2p_i/\sqrt{2p_iq_i}$, $(q_i - p_i)/\sqrt{2p_iq_i}$, and $2q_i/\sqrt{2p_iq_i}$ for when the marker genotype is AA, AB, and BB, respectively. The dominance contribution $\mathrm{H_{ij}}$ is calculated *as* $-p_i/q_i$, -1, and $-q_i/p_i$, for when the marker genotype is AA, AB, and BB, respectively. Here, q_i and p_i are the allele frequencies for alleles A and B, respectively, at the i^{th} SNP locus. Variance component estimates were obtained from Eq. 2 using the package Qxpak v5.02 (Pérez-Enciso and Misztal, 2011). The independence of matrices GRM and DRM was examined using the correlation of their off-diagonal elements.

The significance of the dominance variance was examined by comparing the difference between -2Log-Likelihood values of additive genomic model (**AM**) and additive and dominance genomic model (**ADM**) for each trait to a χ^2 distribution with one degree of freedom (Bolormaa et al., 2015).

Cross-validation for Determining Accuracy of Prediction under Different Models. A 5-fold cross-validation approach was applied to predict the total genetic values for individual animals in a validation population using the additive and dominance genetic values. The genomic prediction accuracy was calculated as the correlation between the predicted total genetic values (additive in the presence of dominance variations) and the corrected phenotypic values (Bolormaa et al., 2015). The corrected phenotypic values were derived after adjusting the original phenotypes for all fixed effects (i.e., = phenotype – fixed effects).

For validation purposes, the study population was divided into 5 parts of similar size using 2 different methods. The first method (denoted as "Across-Family Group") was to initially apply the Cytoscape 3.0 program (©2001-2013 Cytoscape Consortium) to visualize the family relationships among 79 sires, 496 dams, and 5,658 progeny, then split the sires into 5 groups and place all offspring of the same sire into their corresponding sire group (see Fig. 1 for illustration). As a result, no animals in a validation dataset would have related full-sibs in the other 4 reference datasets. The second method (denoted as "Random Group") was to randomly allocate the offspring of each sire to one of 5 groups. For both methods, one of 5 groups took turns as the validation population and the other 4 as the reference population.

For comparison purposes, the prediction accuracies were obtained for 3 genomic models, namely an additive model (AM, i.e., fitting only GRM in the model),

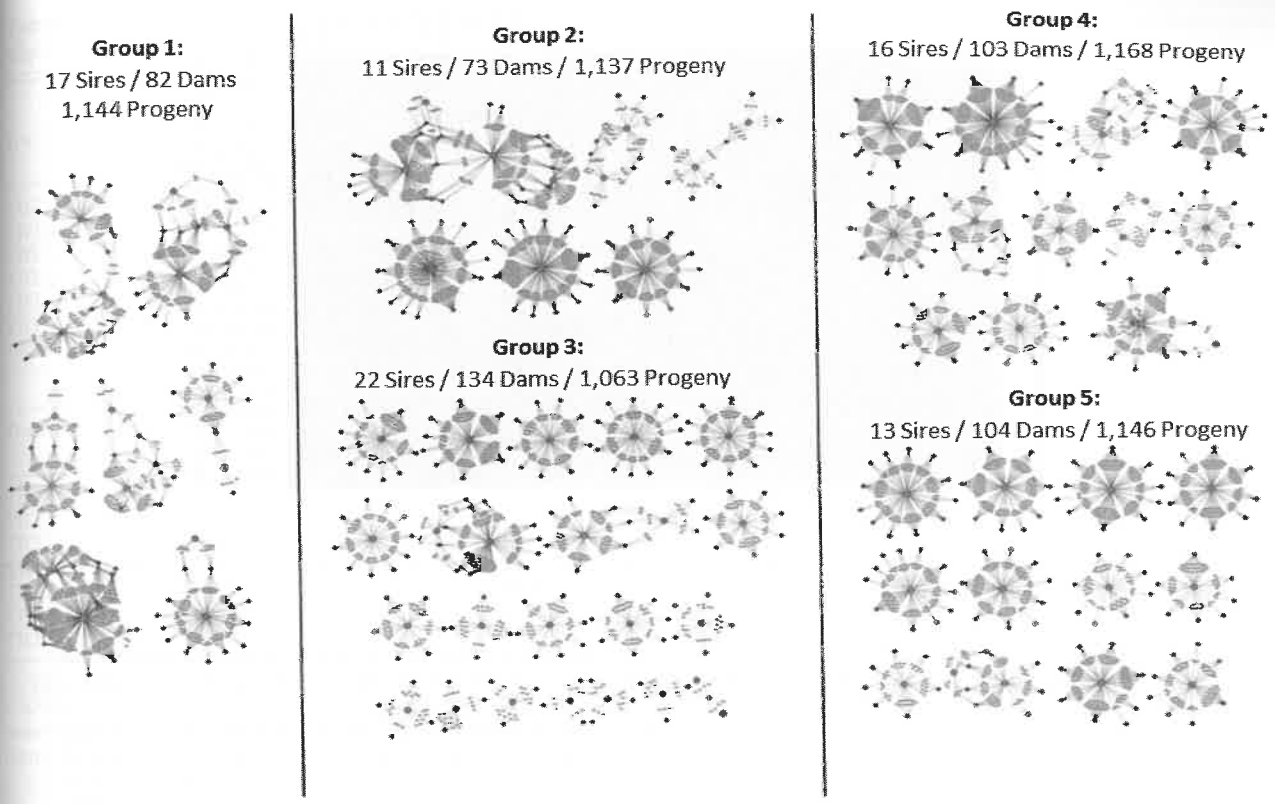

Figure 1. Cytoscape visualisation of groups used in the Across-Family cross-validation illustrating the large number of full-sib families in the ulation and a balanced separation of group animals. Red dots represent sires, blue dots refer to dams, and turquoise dots represent progeny h genotypes and phenotypes.

dominance model (DM, fitting DRM only) and the ditive and dominance model (ADM, fitting both RM and DRM), and 2 5-fold cross validation schemes cross-Family Group and Random Group).

RESULTS

Genome-wide Association Studies (GWAS) Fitting Either Additive or Dominance or both Additive and Dominance Variations

Figure 2 shows the distribution of the number of sig- ificant SNPs ($P < 10^{-4}$) identified on each chromo- ome for each of 7 traits and 3 GWAS models (i.e., addi- ve, dominance, or both additive and dominance mod- s). Several chromosomes contained a large number of NPs associated with 2 or more traits. These included e chromosome 4 for TRT4 and TRT5, chromosome 6 r TRT2, TRT3, TRAIT4, and TRT7, chromosome 12 r TRT1 and TRT3, chromosome 14 for TRAIT4 and RT7, and chromosome Z for TRT1, TRT3, TRT6, and RT7, respectively.

It is clear that the additive model (illustrated by the ostfix "_A"), and the combined additive and domi- ance models (see the postfix "_AD") produced the ajority of significant SNPs. This indicates that the ditive effects of these SNP markers were the major rce contributing to the significant results.

When only the dominance effect of a SNP was fitted in the GWAS model (Fig. 2, illustrated by the postfix "_D"), 3 chromosomes with a large number of significant dominance effects were identified: chromosome 6 for TRT2, TRT3, TRT4, and TRT7, chromosome 1 for TRT5, and chromosome Z for TRT2 and TRT3.

When comparing the 3 models within traits, for the majority of traits, including the dominance effects into the GWAS model contributed little change in the significant SNP profiles of the additive model (e.g., TRT6_A vs. TRT6_D vs. TRT6_AD).

Table 2 presents detailed information on the number of significant SNPs ($P < 10^{-4}$) for each trait that belongs to unique groups (individual model or intersection of 2 models or 3 models, i.e., the components of a Venn diagram). A total of 1,327 unique markers contributed to these significant results. A close examination of the Hardy-Weinberg equilibrium test using Pearson's chi-square method found that the genotypic frequencies of 8,141 SNP (18%) did not follow the expected proportions ($P < 0.001$). Among the 1,327 significant SNP, 318 (24%) failed the Hardy-Weinberg equilibrium test. Of the 318 SNP, 95 (30%) were from the sex chromosome Z.

Table 3 shows the composition of direction of significant SNP effects for individual traits. It is clear that for all 7 traits more SNPs were detected under the additive model (ADD) than under dominance (DOM) or the combined additive and dominance model (AD).

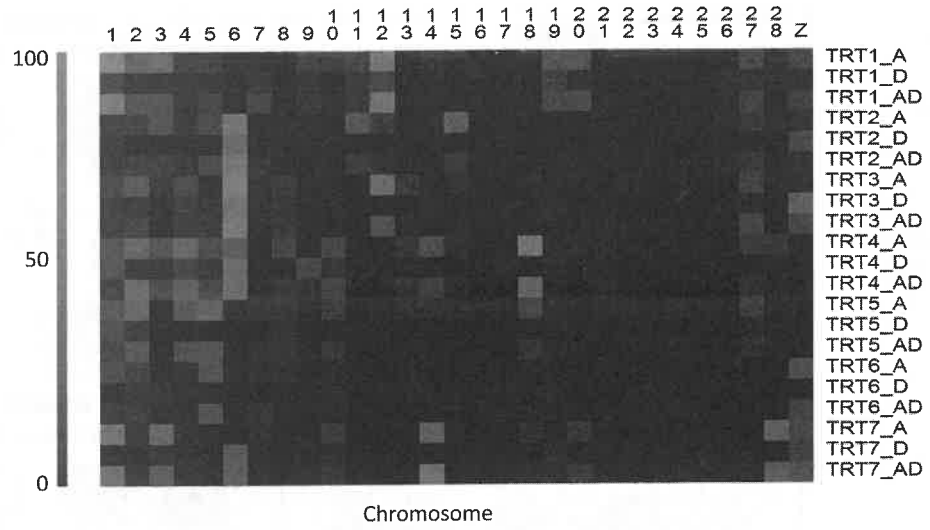

Figure 2. Heatmap visualisation of chromosome distribution of significant SNP ($P < 0.0001$) for each of the 7 traits (TRT1 – TRT7) and models (only Additive: "_A"; only Dominance: "_D"; additive and dominance: "_AD").

Table 2. The number of significant SNPs ($P < 10^{-4}$) identified from GWAS for each trait in each model that belong to various exclusive groups (i.e., the components of a Venn diagram). Add – additive model, Dom – dominance model and AD – additive and dominance model.

Trait	Only Add	Only Dom	Only AD	Add + Dom	Dom + AD	Add + AD	Add + Dom + AD	Total[A]
TRT1	127	19	17	1	12	203	21	400
TRT2	85	16	6	0	2	61	44	216
TRT3	77	35	3	1	19	60	57	252
TRT4	105	19	21	7	24	142	40	358
TRT5	78	8	14	5	0	66	9	180
TRT6	33	3	4	0	0	31	2	73
TRT7	65	7	17	1	33	145	0	268

[A]Across the 7 traits, the total number of unique SNPs identified to be significant was 1,327.

The false discovery rate (FDR) was smaller under the ADD (ranged from 1.27 to 6.84%) than under the DOM or AD. As expected, under the ADD, the number of SNPs with positive effects and the number with negative effects were balanced (see Chi-squared P-value in Table 3). Under the DOM, the percentage of FDR ranged from 4.0 to 90.4%. Except for TRT1, TRT5, and TRT6, the number of SNPs with negative effects was significantly larger than the number with positive effects for all other traits (χ^2 $P < 0.05$). Under the AD (Table 3), except for TRT5, the class of SNP with different directional effects followed a skewed distribution (Chi-squared $P < 0.05$) for 6 other traits. Regardless of the direction of additive effects, the number of SNPs with the negative dominance effects far exceeded the number of SNPs with positive dominance effects for TRT2, TRT3, TRT4, and TRT5. For TRT6 and TRT7, the opposite was true, i.e., the majority of identified significant SNPs had positive dominant effects (e.g., 73 and 66 for TRT6, 147 and 157 for TRT7).

Estimates of Dominance Variance

The REML estimates of genetic variances using additive and both additive and dominance models are shown in Table 4. In comparison to the additive model, fitting both additive and dominance genetic effects in the genomic model resulted in the reduction of the proportion of genetic variance explained by the additive effect in all traits (see Table 4 h_a^2 values, narrow-sense heritability) except TRT5. The decrease in this value was substantial for TRT3 (from 0.31 to 0.087) and TRT4 (from 0.33 to 0.092). The total phenotypic variance explained by the dominance variance (V_d/V_{pADM}) was significant for all traits (Table 4, $V_d/V_{pADM} > 0.10$, $P < 0.01$). Among all, TRT1 had the highest V_d/V_{pADM} value (0.135) and TRT7 had the lowest value (0.105). The broad sense heritability (the proportion of phenotypic variance explained by the total genetic variance that consisted of additive and dominance variances, V_g/V_{pADM}) was much higher than the narrow sense of heritability for TRT1, TRT2, TRT5, TRT6, and TRT7, while for TRT3 and TRT4, the opposite was true. In all traits, the proportion of genetic variance explained by dominance variance (calculated by $V_d/V_{g,}$) was high, ranging from 29.5% for TRT1 to 57% for TRT2.

We also observed that the level of population heterozygosity had a significant impact ($P < 0.05$) on TRT1, TRT2, TRT3, TRT4, and TRT5 (see Table 4 t_{Het} values), but not on TRT6 and TRT7. The

Table 3. Composition of significant SNPs ($P < 10^{-4}$) with additive or dominance effects or both additive and dominance effects.

	Additive model					Dominance model				
Trait	Total No.	FDR (%)	+ve Nb	−ve Nb	Chi-squared *P*-value	Total No.	FDR (%)	+ve Nb	−ve Nb	Chi-squared *P*-value
TRT1	352	1.27	170	182	0.522	53	8.51	34	19	0.039
TRT2	192	2.34	91	101	0.470	62	7.28	9	53	0.000
TRT3	195	2.31	94	101	0.616	112	4.02	16	96	0.000
TRT4	294	1.53	157	137	0.243	90	5.01	24	66	0.000
TRT5	158	2.85	75	83	0.524	22	20.53	14	8	0.201
TRT6	66	6.84	30	36	0.460	5	90.35	1	4	0.180
TRT7	211	2.13	107	104	0.836	41	11.01	4	37	0.000

	Additive and dominance model						
Trait	Total No.	FDR (%)	+ve Add +ve Dom	−ve Add −ve Dom	+ve Add −ve Dom	−ve Add +ve Dom	Chi-squared *P*-value
TRT1	253	1.78	45	81	82	43	0.0001
TRT2	115	3.92	5	59	45	6	0.0000
TRT3	139	3.24	7	64	62	6	0.0000
TRT4	227	1.98	54	72	67	34	0.0017
TRT5	89	5.07	26	23	27	13	0.1376
TRT6	37	12.20	12	2	4	19	0.0002
TRT7	195	2.31	58	36	33	68	0.0005

Total No. - Total number of SNPs with $P < 10^{-4}$; FDR – False discovery rate; +ve Nb – number of SNPs with positive effect; −ve Nb – SNPs with negative effect; +ve Add +ve Dom – SNPs with positive additive and dominance effects; −ve Add −ve Dom – SNPs with negative additive and dominance effects; +ve Add −ve Dom – SNPs with positive additive and negative dominance effects; −ve Add +ve Dom – SNPs with negative additive and positive dominance effects; Chi-squared *P*-value – *P*-value from χ^2 test.

Table 4. Proportion of genetic variance estimated from the additive genomic model (AM) and the additive and dominance genomic model (ADM) for each trait.

	AM	ADM				
Trait	h_a^2	t_{Het}	h_a^2	V_d/V_{pADM}	V_g/V_{pADM}	V_d/V_g (%)
TRT1	0.35 (0.12)	1.88*	0.323 (0.11)	0.135 (0.040)***	0.458	29.5
TRT2	0.12 (0.018)	6.65***	0.081 (0.009)	0.108 (0.028)***	0.189	57.3
TRT3	0.30 (0.10)	4.41***	0.088 (0.011)	0.108 (0.027)***	0.195	55.1
TRT4	0.33 (0.11)	3.24***	0.093 (0.011)	0.106 (0.027)***	0.199	53.3
TRT5	0.13 (0.021)	2.54*	0.282 (0.098)	0.133 (0.041)**	0.415	32.0
TRT6	0.19 (0.047)	−0.38	0.214 (0.065)	0.130 (0.042)**	0.344	37.8
TRT7	0.12 (0.018)	1.12	0.0749 (0.0084)	0.105 (0.027)**	0.180	58.4

h_a^2 – heritability of additive effect; V_{pADM} – total phenotypic variance from the additive and dominance model; V_d – dominance variance; V_g – genotypic variance; t_{Het} – t-value of heterozygosity effect.
*significantly different at $P < 0.05$.
**significantly different at $P < 0.01$.
***significantly different at $P < 0.001$.

traits on which heterozygosity had a significant impact correspond to the traits for which the majority of SNPs had the negative dominance effects, while in TRT6 and TRT7 (no significant impact of population heterozygosity found) the majority of SNPs had positive dominance effects (Table 4).

Accuracy of Prediction of Phenotypic Values

Tables 5 and 6 summarize the genomic prediction accuracy of phenotypic values (calculated as the correlation between the predicted total genetic values and the corrected phenotypic values) for each trait using 3 genomic prediction models (additive, dominance or both additive and dominance) and 2 5-fold cross-validation schemes. When the Across-Family group approach was used, in comparison to the additive genomic model AM (fitting GRM only), the average prediction accuracy for each trait in the dominance model DM (fitting DRM only) was close to zero (Table 5, the values ranged from −0.002 (0.046) to 0.056 (0.029)). When using the Random Group approach (Table 6), the average prediction accuracy in the dominance model ranged from 0.072 (0.017) for TRT6 to 0.13 (0.029) for TRT1.

When comparing the genomic prediction accuracies of phenotypic values from both additive and dominance models (ADM) to those of AM, surprisingly, the overall prediction accuracy declined in all traits (Tables 5 and 6), although the magnitude of decrease was not significant ($P > 0.05$). For example, for TRT1, the average accuracy under the AM was 0.394 (0.0146), while under the ADM, the prediction accuracy was reduced to 0.387 (0.0134) (Table 6). Therefore, including SNP dominance effects in a genomic prediction model had a slight negative effect on the prediction accuracy.

Table 5. Accuracies of predicted phenotypic values for different models and the 5-fold cross-validation scheme – Across family groups.

Model	Trait	Group 1	Group 2	Group 3	Group 4	Group 5	Average	SD
Additive	TRT1	0.297	0.351	0.269	0.250	0.313	0.296	0.0392
	TRT2	0.199	0.193	0.122	0.176	0.195	0.177	0.0320
	TRT3	0.181	0.229	0.202	0.226	0.154	0.198	0.0316
	TRT4	0.183	0.252	0.269	0.276	0.202	0.236	0.0416
	TRT5	0.190	0.213	0.255	0.277	0.161	0.219	0.0472
	TRT6	0.168	0.174	0.193	0.206	0.169	0.182	0.0168
	TRT7	0.152	0.166	0.213	0.137	0.192	0.172	0.0306
	Average	0.196	0.225	0.217	0.221	0.198	0.212	0.0135
Dominance	TRT1	0.013	0.047	0.067	0.091	0.064	0.056	0.0289
	TRT2	−0.004	0.030	−0.011	0.045	0.028	0.018	0.0240
	TRT3	−0.044	−0.009	0.014	0.054	0.008	0.004	0.0357
	TRT4	−0.051	0.010	0.022	0.041	0.033	0.011	0.0366
	TRT5	−0.060	−0.027	0.020	0.037	0.044	0.003	0.0447
	TRT6	−0.055	−0.043	0.036	0.004	0.047	−0.002	0.0457
	TRT7	−0.004	−0.006	0.019	0.014	0.033	0.011	0.0164
	Average	−0.030	0.000	0.024	0.041	0.037	0.014	0.0295
Add. + Dom.	TRT1	0.275	0.337	0.280	0.270	0.297	0.292	0.0272
	TRT2	0.150	0.185	0.095	0.182	0.151	0.152	0.0362
	TRT3	0.119	0.192	0.187	0.232	0.116	0.169	0.0503
	TRT4	0.105	0.217	0.238	0.261	0.143	0.193	0.0661
	TRT5	0.106	0.169	0.216	0.246	0.126	0.173	0.0589
	TRT6	0.125	0.146	0.187	0.178	0.165	0.160	0.0250
	TRT7	0.103	0.117	0.177	0.109	0.146	0.130	0.0308
	Average	0.140	0.195	0.197	0.211	0.163	0.181	0.0290

Table 6. Accuracies of predicted phenotypic values for different models and the 5-fold cross-validation scheme – Random groups.

Model	Trait	Group 1	Group 2	Group 3	Group 4	Group 5	Average	SD
Additive	TRT1	0.382	0.402	0.400	0.376	0.411	0.394	0.0146
	TRT2	0.264	0.288	0.267	0.238	0.361	0.284	0.0468
	TRT3	0.275	0.298	0.291	0.290	0.353	0.301	0.0300
	TRT4	0.290	0.331	0.336	0.363	0.369	0.338	0.0314
	TRT5	0.264	0.274	0.291	0.337	0.354	0.304	0.0396
	TRT6	0.265	0.238	0.252	0.275	0.296	0.265	0.0221
	TRT7	0.279	0.267	0.276	0.263	0.276	0.272	0.0068
	Average	0.288	0.300	0.302	0.306	0.346	0.308	0.0221
Dominance	TRT1	0.133	0.102	0.173	0.125	0.104	0.127	0.0288
	TRT2	0.085	0.073	0.130	0.112	0.084	0.097	0.0235
	TRT3	0.094	0.106	0.127	0.151	0.059	0.107	0.0347
	TRT4	0.087	0.128	0.131	0.161	0.053	0.112	0.0422
	TRT5	0.068	0.157	0.120	0.133	0.067	0.109	0.0401
	TRT6	0.042	0.075	0.075	0.078	0.087	0.072	0.0172
	TRT7	0.120	0.078	0.139	0.144	0.089	0.114	0.0295
	Average	0.090	0.103	0.128	0.129	0.078	0.105	0.0227
Add. + Dom.	TRT1	0.370	0.390	0.403	0.377	0.395	0.387	0.0134
	TRT2	0.243	0.261	0.265	0.238	0.354	0.272	0.0471
	TRT3	0.262	0.290	0.292	0.297	0.332	0.294	0.0250
	TRT4	0.262	0.323	0.324	0.349	0.343	0.320	0.0345
	TRT5	0.226	0.290	0.286	0.312	0.328	0.288	0.0388
	TRT6	0.236	0.233	0.248	0.263	0.279	0.252	0.0193
	TRT7	0.257	0.237	0.269	0.264	0.273	0.260	0.0142
	Average	0.265	0.289	0.298	0.300	0.329	0.296	0.0230

DISCUSSION

A better understanding of the non-additive genetic effects of a large number of SNPs on complex phenotypes can provide important knowledge on the improvement of genomic prediction of individual animals. To date several studies have examined the impact of non-additive SNP genetic variation on genomic prediction accuracy of complex traits in different species. However, the results were conflicting. This raised the question whether the differences from these studies were due to different population structures of these species, i.e., large full-sib families (e.g., pig and tree species) vs. large half-sib families (e.g., cattle). In this study, using data from large full-sib families of broilers (575 full-sib families consisting of 79 sires, 496 dams, and 5,658 progeny) and the same methods as Bolormaa et al. (2015), we evaluated the SNP dominance effects on 7 feed-related traits with different heritabilities (0.12 to 0.35). Our results clearly indicate that a significant

amount of SNP dominance variance existed for all traits and that this explained 10.5 to 13.5% of the phenotypic variance and 29.5 to 58.4% of the total genetic variance, respectively. The magnitude of the dominance variances identified in this study was much larger than those of growth, carcass, and fertility traits in beef cattle (Bolormaa et al., 2015). However, despite an ideal large full-sib family structure and large dominance variations in all traits, including both additive and dominance SNP effects in a genomic model for the prediction of phenotypic values had little impact on improving the prediction accuracy of an additive model. Our results confirmed the findings of Bolormaa et al. (2015). The theoretical explanations by Hill et al. (2008) may shed some light for our findings: a) In cases of full-sib populations, epistatic or dominance variance components are confounded with the common environment shared by full-sibs (in this study, it would be hatch effect); and b) the gene-gene interactions fail to contribute to the level of variance.

The intriguing results found in our study were that including additive and dominance SNP effects in the prediction model resulted in a slight decline of prediction accuracy in all 7 traits. This was not seen in the other studies (e.g., Su et al., 2012; Bolormaa et al., 2015). This may be due to the fact that the genomic prediction model, e.g., Su et al. (2012) used, included not only additive and dominance effects of SNP markers, but also the additive × additive interactions. As a result, their model produced a slightly higher accuracy (29.5%) than a simple additive model (28.5%) (Su et al., 2012). In our study, none of the models considered an epistatic interaction. In addition, a close examination of all SNPs identified with significant additive and dominance effects on several traits in our study revealed that the number of SNPs with negative dominance effects was far greater than the number with positive effects in 5 out of 7 traits. We also observed that the total genetic values predicted from the model including additive and dominance effects were smaller (ranging from 2.7 to 44%) than the genetic values from the additive effects only model. The results indicate that the additive and dominance effects may have cancelled each other out when both were included in the same genomic prediction model.

In addition, we also found that when including additive and dominant effects into the REML analysis, the 5 traits (TRT1, TRT2, TRT3, TRT4, and TRT7) showed reduced additive variances. The recent study in apples (Kumar et al., 2015) also found that the genomic prediction models excluding non-additive components produced upward bias in estimates of additive variance.

Unlike growth, carcass, and fertility traits in beef cattle (Bolormaa et al., 2015), in which the direction of the overall heterozygosity effect (reflected by the sign of t_{Het} value) for each trait largely agreed with the direction of the majority of SNPs identified to have significant dominance effects, in our study, we found that the opposite was true. That is, the traits with an overall significant positive effect of heterozygosity (e.g., TRT2, TRT3, TRT4, and TRT5 in Table 5, t_{Het}), except TRT1, linked to a large number of SNPs with negative dominance effects. This suggests that the underlying dominance effects favored the opposite of directional selection on the trait of interest. Another possible explanation is that the direction of SNP dominance effects could be the consequence of the allele coding scheme we applied to all SNP markers.

Two different 5-fold cross-validation schemes were used in this study to compare the genomic prediction accuracy. In the case of Across-Family groups, in which no progeny of the reference population families was included in the validation population of the other 4 groups, in theory the off-diagonal elements of the additive GRM and those of DRM should be orthogonal (i.e., zero correlation). However, we did find a positive correlation between them (Fig. 3, $r = 0.27$, $P < 2.2^{*}10^{-16}$). The departure from the Hardy-Weinberg equilibrium in 18% of 45,176 SNPs that were used to build both GRM and DRM could be one of the contributing factors for the un-orthogonal correlation. In theory the orthogonality of additive and dominance components happens only when there is no genetic drift, no selection, and no mutation in a population, and all genotypic frequencies follow a linkage equilibrium pattern (Falconer and Mackay, 1996). When investigating the relationship between individual genome-wide heterozygosity and individual inbreeding coefficients, we found a significant negative correlation between the 2 parameters (Fig. 4, $r = -0.2194$, $P < 2.2^{*}10^{-16}$)".

In this study, the only method evaluated for estimating the SNP dominance effects on different feed-related traits was that of Bolormaa et al. (2015). Using simulation data, Long et al. (2011) investigated the non-parametric method – RBF regression for predicting quantitative traits under different additive, dominance, and epistasis gene effects. Their results suggest that RBF had better predicting merit of individuals than the Bayesian LASSO when non-additive gene effects exist, but was the opposite under additive gene action (Long et al., 2011). Azvedo et al. (2015) evaluated the performance of 10 different additive-dominance predictive models, including Ridge, Lasso, and Bayesian methods, in predicting genomic breeding and total genotypic values from additive and dominance models. Using the simulated datasets, they demonstrated that a modified Bayesian/Lasso method performed the best in comparison to other methods. This suggests that the different methods may yield different outcomes. Recently, using both simulation data and real pig data, Waldmann (2016) also found that in the presence of dominance and epistasis, the non-parametric method – BART, developed by Chipman et al., 2010, gave a smaller genomic prediction error and increased prediction accuracy of phenotypic values than RF, BLASSO, GBLUP, and RKHS regression methods. Therefore, further examination with different models needs to be conducted

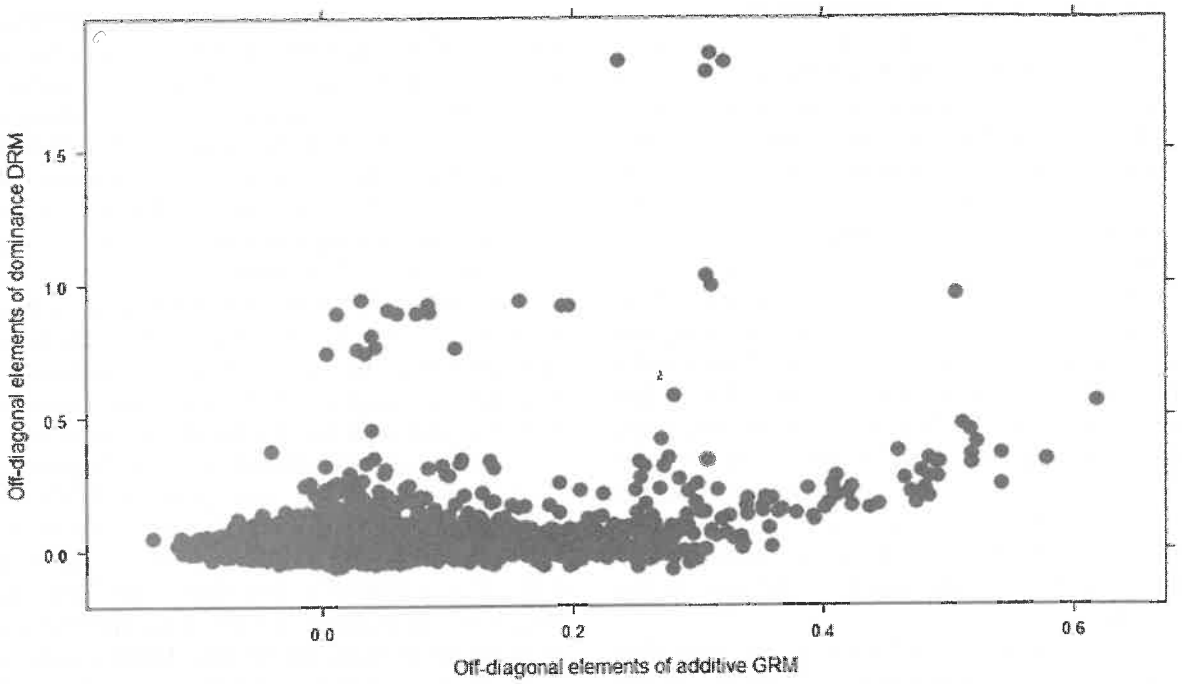

Figure 3. The relationship between off-diagonal elements of the additive genomic relationship (GRM) and dominance genomic relationship (DRM). The correlation is 0.2722, which is significantly different from zero ($P < 2.2e\text{-}16$).

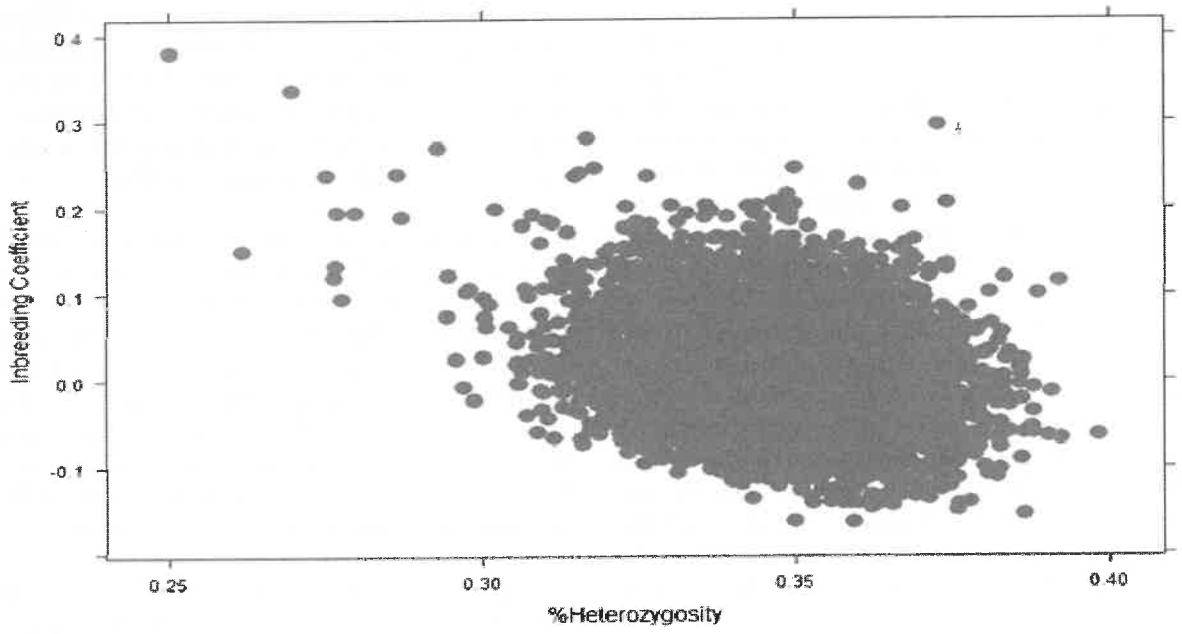

Figure 4. Relationship between individual % heterozygosity and the inbreeding coefficient. The correlation is −0.2194, which is highly and significantly different from zero ($P < 2.2e\text{-}16$).

to evaluate the effect of non-additive genetic variance (including non-additive and dominance epistasis interactions).

CONCLUSIONS

This study quantified the non-additive genetic variance - dominance in a commercial broiler chicken population for feed-related traits using a large number of SNP markers and large full-sib families. The results demonstrate that 1): SNP dominance variance accounted for a significant proportion of the total genetic variance in all 7 traits; 2) including dominance effects in a GWAS model contributed little in the SNP profiles of an additive model; and 3) there is little merit in including non-additive SNP genetic effects into a genomic prediction of phenotypic values of animals.

ACKNOWLEDGMENTS

The work has been financed by CSIRO and Cobb-Vantress Inc. We would like to thank the 2 CSIRO internal referees Sonja Dominik and Laercio Porto-Neto for their valuable comments and suggestions.

REFERENCES

Aliloo, H., J. E. Pryce, O. Gonzalez-Recio, B. G. Cocks, and B. J. Hayes. 2016. Accounting for dominance to improve genomic evaluations of dairy cows for fertility and milk production traits. Genet. Sel. Evol. 48:8.

Azvedo, C. F., M. D. de Resende, F. F. de Silva, J. M. S. Viana, M. S. F. Valente, M. F. R. Resende, Jr., and P. Munoz. 2015. Ridge, LASSO and Bayesian additive – dominance genomic models. BMC Genetics. 16:105.

Bolormaa, S., B. J. Hayes, K. Savin, R. Hawken, W. Barendse, P. Arthur, R. M. Herd, and M. E. Goddard. 2011. Genome wide association studies for feedlot and growth traits in cattle. J. Anim. Sci. 89:1684–97.

Bolormaa, S., J. E. Pryce, Y. Zhang, A. Reverter, W. Barendse, B. Hayes, and M. E. Goddard. 2015. Non-additive genetic variation in growth, carcass and fertility traits of beef cattle. Genet. Sel. Evol. 47:26.

Chipman, H. A., E. I. George, and R. E. McCulloch. 2010. BART: Bayesian additive regression trees. Ann Appl Stat. 4:266–298.

Falconer, D. D., and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. England: Longman.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Trans. Roy. Soc. Edinburgh. 52:399–433.

Gianola, D., H. Okut, K. A. Weigel,, and G. J. M. Rosa. 2011. Predicting complex quantitative traits with Bayesian neural networks: A case study with Jersey cows and wheat. BMC genetics 12:87.

Gianola, D., G. Morota, and J. Crossa. 2014. Genome-enabled prediction of complex traits with kernel methods: What have we learned? Proceedings of 10th World Congress on Genetics Applied to Livestock Production. 212.

González-Recio, D. Gianola, N. Long, K. A. Weigel, G. J. M. Rosa, and S. Avendaño 2008. Nonparametric methods for incorporating genomic information into genetic evaluations: an application to mortality in broilers. Genetics 178:2305–2313.

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.

Kumar, S., C. Molloy, P. Muñoz, H. Daetwyler, D. Chagné, and R. Volz. 2015. Genome-enabled estimates of additive and nonadditive genetic variances and prediction of apple phenotypes across enviroments. G3 Genes— Genomes— Genetics. 5:2711–2718.

Long, N., D. Gianola, G. J. M. Rosa, and K. A. Weigel. 2011. Marker-assisted prediction of non-additive genetic values. Genetica 139:843–854.

Lynch, M., and B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associates, Inc.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Munoz, P. R., M. F. R. Resende, Jr., S. A. Gezan, M. D. V. Resende, G. de los Campos, M. Kirst, D. Huber, and G. Peter. 2014. Unravelling additive from nonadditive effects using genomic relationship matrices. Genetics, 198:1759–1768.

Nishio, M., and M. Satoh. 2014. Including dominance effects in the genomic BLUP methods for genomic evaluation. PLoS One. 9:e85792. doi:10.1371/journal.pone.0085792.

Pérez-Enciso, M., and I. Misztal. 2011. Qxpak.5: Old mixed model solutions for new genomics problems. BMC Bioinformatics. 12:202.

Su, G., O. F. Christensen, T. Ostersen, M. Henryon, and M. S. Lund. 2012. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense dingle nucleotide polymorphism markers. PLoS One. 7:e45293.

Toro, M. A., and L. Varona. 2010. A note on mate allocation for dominance handling in genomic selection. Genet. Sel. Evol. 42:33.

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.

Waldmann, P. 2016. Genome-wide prediction using Bayesian additive regression trees. Genet Sel Evol., 48:42. DOI 10.1186/s12711-016-0219-8.

Wittenburg, D., N. Melzer, and N. Reinsch. 2011. Including non-additive genetic effects in Bayesian methods for the prediction of genetic values based on genome-wide markers. BMC Genetics 12:74.

Wittenburg, D., N. Melzer, and N. Reinsch. 2015. Genomic additive and dominance variance of milk performance traits. J. Anim. Breed. Genet. 132:3–8.

Zhu, Z., A. Bakshi, A. A. E. Vinkhuyzen, G. Hemani, S. H. Lee, I. M. Nolte, J. V. van Vliet-Ostaptchouk, H. Snieder, T. Esko, L. Milani, R. Magi, A. Metspalu, W. G. Hill, B. S. Weir, M. E. Goddard, P. M. Visscher, and J. Yang, and The lifelines Cohort Study. 2015. Dominance genetic variation contributes little to the missing heritability for human complex traits. Am. J. Hum. Genet. 96:377–385.