ORIGINAL ARTICLE

Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model

J Jiang[1,5], Q Zhang[1], L Ma[2], J Li[3], Z Wang[4] and J-F Liu[1]

Predicting organismal phenotypes from genotype data is important for preventive and personalized medicine as well as plant and animal breeding. Although genome-wide association studies (GWAS) for complex traits have discovered a large number of trait- and disease-associated variants, phenotype prediction based on associated variants is usually in low accuracy even for a high-heritability trait because these variants can typically account for a limited fraction of total genetic variance. In comparison with GWAS, the whole-genome prediction (WGP) methods can increase prediction accuracy by making use of a huge number of variants simultaneously. Among various statistical methods for WGP, multiple-trait model and antedependence model show their respective advantages. To take advantage of both strategies within a unified framework, we proposed a novel multivariate antedependence-based method for joint prediction of multiple quantitative traits using a Bayesian algorithm via modeling a linear relationship of effect vector between each pair of adjacent markers. Through both simulation and real-data analyses, our studies demonstrated that the proposed antedependence-based multiple-trait WGP method is more accurate and robust than corresponding traditional counterparts (Bayes A and multi-trait Bayes A) under various scenarios. Our method can be readily extended to deal with missing phenotypes and resequence data with rare variants, offering a feasible way to jointly predict phenotypes for multiple complex traits in human genetic epidemiology as well as plant and livestock breeding.
Heredity (2015) 115, 29–36; doi:10.1038/hdy.2015.9; published online 15 April 2015

INTRODUCTION

In human genetic epidemiology, accurate prediction of disease risk is vital for disease prevention and personalized medicine. Although thousands of genome-wide association studies (GWAS) have discovered a large number of variants that are significantly associated with complex human traits and diseases (Hindorff et al., 2014), it is still a challenge to directly implement these findings to predict yet-to-be observed phenotypes and thus to advance preventive and personalized medicine. Partly, this is owing to the so-called 'missing heritability' that these associated variants can typically account for only a small fraction of the total genetic variance (Manolio et al., 2009). The issue of 'missing heritability' probably results from the limitations of statistical techniques commonly used in GWAS, as pointed out by previous studies (Yang et al., 2010; de los Campos et al., 2010; Makowsky et al., 2011); for example, in single-locus association analyses for either SNP markers or gene expression levels, we can merely derive a small number of statistically significant loci even for a complex trait affected by a large number of small-effect variants.

To address the weakness of conventional GWAS on phenotype prediction, some of recent studies have turned to take advantage of the whole-genome prediction (WGP) method initially proposed by Meuwissen et al. (2001). WGP seeks to model genome-wide SNPs simultaneously to predict yet-to-be observed phenotypes, for example, human height (Makowsky et al., 2011; de Los Campos et al., 2013),

human lifespan (de los Campos et al., 2012) and skin cancer risk (Vazquez et al., 2012). These studies have demonstrated the strength of WGP methods for predicting human complex traits of which the underlying genetic architecture likely consists of numerous variants of small effect. In addition, a number of studies suggested that many human diseases, like schizophrenia and bipolar disorder (Purcell et al., 2009), multiple sclerosis (Bush et al., 2010) and rheumatoid arthritis (Stahl et al, 2012), also have such a polygenic architecture, implicating a broad utility for WGP methods to predict human disease risk. Besides phenotype prediction in human genetics, WGP has been receiving much attention and widely employed to genetically improve economically important traits in domestic animals, including milk production performance and fitness traits of dairy cows (VanRaden et al., 2009).

Since the seminal work of Meuwissen et al. (2001) for predicting genomic breeding values in animal and plant breeding, a number of WGP methods have been developed and extensively investigated based on different algorithms (Gianola et al., 2006; VanRaden, 2008; Gianola et al., 2010; Habier et al., 2011; Legarra et al., 2011). Among these methods, the most popular are the parametric approaches, for example, GBLUP (VanRaden, 2008) and the Bayesian alphabet (Gianola, 2013), which are basically based on the framework of linear regression. Generally, these parametric methods assume that the effect of each marker is independently distributed with a specific prior

[1]Department of Animal Genetics, Breeding and Reproduction, China Agricultural University, Beijing, China; [2]Department of Animal and Avian Sciences, University of Maryland, College Park, MD, USA; [3]Institute of Animal Science, Chinese Academy of Agricultural Science, Beijing, China and [4]Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Alberta, Canada
[5]Current address: Department of Animal and Avian Sciences, University of Maryland, College Park, MD, USA.
Correspondence: Dr J-F Liu, Department of Animal Genetics, Breeding and Reproduction, China Agricultural University, 2 Yuanmingyuan West Road, Beijing 100193, China.
E-mail: liujf@cau.edu.cn

distribution given by corresponding statistical methods. Clearly, such an assumption of independent distribution for each SNP effect is statistically inappropriate, especially when the adjacent markers are in high linkage disequilibrium (LD) with the same causal gene. This unrealistic assumption potentially sacrifices the prediction accuracy to some extent. To address this issue, Yang and Tempelman (2012) proposed a first-order antedependence model to account for the nonstationary correlations between SNP markers through assuming a linear relationship between the effects of adjacent markers. As expected, the proposed antedependence-based WGP models outperformed their conventional counterparts in the prediction accuracy of genomic merit in the context of single-trait analyses.

Compared with single-trait analyses, multiple-trait joint analyses have been widely confirmed as having obvious advantages in statistical power and parameter estimation accuracy in earlier studies of quantitative trait locus (QTL) mapping and GWAS (Liu et al., 2007; Liu et al., 2009; Jiang and Zeng, 1995; Korte et al., 2012). Following this evidence, it can be speculated that in WGP fields, joint prediction of multiple traits should benefit from genetic correlation between traits and thus have the potential to obtain higher prediction accuracy than single-trait methods, especially for a low-heritability trait that is genetically correlated with a high-heritability trait. Statistical methods for joint prediction of multiple traits have received close attention in recent years. Particularly, Bayes A, Bayes C and Bayes Cπ have been extended to be applicable for multiple trait analyses (Calus and Veerkamp, 2011; Jia and Jannink, 2012), demonstrating clear advantages over single-trait methods.

Logically, a more natural strategy for enhancing WGP is to take into consideration relations between multiple traits, as well as those between SNP effects simultaneously for achieving high prediction accuracy from the viewpoint of statistics. So far, there is still a gap between multiple-trait joint prediction approaches and models considering correlated SNP effects, for example, antedependence-based WGP models. Therefore, aiming at developing an improved prediction methodology, we are in an attempt to construct a multiple-trait antedependence-based WGP model to bridge such a gap aforementioned under the framework of Bayesian algorithm. Specifically, we proposed a novel multivariate method with two different types of prediction models via setting the antedependence parameter as either a matrix or a scalar, respectively. Theoretically, our proposed models can relax the conventional assumption of independence of marker effects while simultaneously taking advantage of the correlation between traits by modeling a linear relationship of effect vector between each pair of adjacent markers. Using simulations as well as the publicly available data sets including the 16th QTL-MAS workshop data and the heterogeneous stock mice real data, we compared our proposed method with the classical approaches including the Bayes A version by Habier et al. (2011) and multi-trait Bayes A by Jia and Jannink (2012) to further validate the performance of our proposed method. Our study clearly demonstrated that the proposed multiple-trait antedependence-based WGP method is more accurate and robust than corresponding traditional counterparts for genomic prediction, offering a feasible way to jointly predict complex traits in human genetic epidemiology, as well as plant and livestock breeding.

MATERIALS AND METHODS

Bayesian multivariate antedependence model

We developed a Bayesian version of first-order multivariate antedependence model (Zimmerman and Nunez-Anton, 2010). The performance of proposed method herein was evaluated by comparison with traditional single-trait Bayes A (Meuwissen et al., 2001; Habier et al., 2011), as well as the multi-trait Bayes A

method initially developed by Jia and Jannink (2012) (see Supplementary Method, Section 1 for details). Considering different types of correlation between SNP effects, we developed two forms of first-order multivariate antedependence models considering the antedependence parameter (Zimmerman and Nunez-Anton, 2010) as a matrix as well as a scalar, named as 'matrix model' and 'scalar model', respectively.

Matrix model. Assuming n individuals were genotyped at p SNP markers, the matrix model for joint prediction of m traits is expressed as,

$$y_i | \mu, \alpha_\bullet, \Sigma \sim N(\mu + \sum_{j=1}^{p} Z_{ij}\alpha_j, \Sigma)\ i = 1, ..., n$$

$$\alpha_j = \begin{cases} \delta_j\ j = 1 \\ T_{j,j-1}\alpha_{j-1} + \delta_j\ j = 2, ..., p \end{cases}$$

$$T_{j,j-1} | M, V_{\delta_j}, V_T \sim MatrixNormal(M, V_{\delta_j}, V_T)\ j = 2, ..., p$$

$$\delta_j | V_{\delta_j} \sim N(\mathbf{0}, V_{\delta_j})\ j = 1, ..., p$$

$$V_T \sim W^{-1}(scale = I, d.f. = m)$$

$$\Sigma | \Psi_\Sigma, v_\Sigma \sim W^{-1}(scale = \Psi_\Sigma, d.f. = v_\Sigma)$$

$$V_{\delta_j} | \Psi_V, v_V \sim W^{-1}(scale = \Psi_V, d.f. = v_V)j = 1, ..., p$$

$$\Psi_V \sim W_m(scale = I, d.f. = m)$$

$$\Psi_\Sigma \sim W_m(scale = I, d.f. = m) \quad (1)$$

where y_i is an m-element phenotypic vector for individual i ($i = 1, ..., n$); μ is the vector of overall population mean of m traits; α_j is an m-element vector for the effects of the jth SNP marker on all m traits ($j = 1, ..., p$); Z_{ij} is the SNP genotype code for individual i at marker j; and Σ is an $m \times m$ covariance matrix of the residual effects. In this study, all markers are considered to be biallelic, and marker genotypes were coded as 0, 1 or 2 corresponding to the number of copies of an allele at a locus.

As shown in the model (1), each marker effect is assumed to have a linear relationship with that of the preceding adjacent marker based on the physical position or the genetic position along a chromosome, that is, $\alpha_j = T_{j,j-1}\alpha_{j-1} + \delta_j$. It should be pointed out here that in the traditional univariate model, the prior of antedependence parameter t generally follows a normal distribution (Yang and Tempelman, 2012). For the case of multivariate model, we accordingly adopted a random matrix $T_{j,j-1}$ herein with the prior of normal distribution. It has been demonstrated elsewhere (Minka, 2001) how to use such a matrix normal prior in similar analyses. Following this initial work, we constructed corresponding prior normal distribution for the matrix antedependence parameter $T_{j,j-1}$ in our proposed methodology. Here $T_{j,j-1}$ is a $m \times m$ matrix corresponding to the jth maker, following a matrix normal distribution with mean matrix M ($m \times m$), among-row covariance V_{δ_j} and among-column covariance V_T. δ_j is an m-element vector for the residual part of the jth marker's effects, following a multivariate normal distribution with mean zero and covariance V_{δ_j}. Note that V_{δ_j} is the among-row covariance of the prior distribution of $T_{j,j-1}$ as well as the covariance of the prior distribution of δ_j. Introducing V_{δ_j} to the prior of T herein can clearly facilitate the implementation of Gibbs sampling. The reason is that we can readily get a close form of the conditional posterior of T under such case (see Page 7 in Supplementary Method). As we know, an explicit expression of posterior probability density function (pdf) is a prerequisite for performing Gibbs sampling. Further, we specified for V_T an inverse Wishart prior distribution with an identity scale matrix and m degrees of freedom. Accordingly, we assigned an inverse Wishart prior distribution with scale matrix Ψ_Σ and v_Σ degrees of freedom for matrix Σ and an inverse Wishart prior distribution with scale matrix Ψ_V and v_V degrees of freedom for V_{δ_j} ($j = 1, ..., p$). Both Ψ_Σ and Ψ_V were further assumed to follow a Wishart distribution with an identity scale matrix and m degrees of freedom.

In our application, we specified the hyperparameters in model (1) as $\{v_\Sigma = 4.2, v_V = 4.2, M = \mathbf{0}_{m,m}\}$. The matrix model was solved through Gibbs sampling. We give the details on the Gibbs sampling procedures in Supplementary Method, Section 2. In total, 200 000 MCMC iterations were conducted for the matrix model and the first 100 000 iterations were discarded as burn in. Every 10th sample was kept for the follow-up inference.

Scalar model. Basically, the scalar model is a simplified version of the matrix model. The antedependence parameter in the scalar model was specified as a scalar rather than a matrix, that is,

$$\boldsymbol{\alpha}_j = \begin{cases} \boldsymbol{\delta}_j \; j = 1 \\ t_{j,j-1}\boldsymbol{\alpha}_{j-1} + \boldsymbol{\delta}_j \; j = 2, ..., p, \end{cases} \quad (2)$$

where $t_{j,j-1}$ is the scalar antedependence parameter of $\boldsymbol{\alpha}_j$ on $\boldsymbol{\alpha}_{j-1}$. Using the scalar means that the antedependence parameter at a marker is identical for all traits. The simplified model can decrease the sampling complexity in MCMC cycles and still reflect the correlation between each pair of adjacent markers. Further, we set $t_{j,j-1}$ a normal prior as

$$t_{j,j-1}|\mu_t, \sigma_{t_j}^2 \sim N(\mu_t, \sigma_{t_j}^2) \, j = 2, ..., p,$$

where $\sigma_{t_j}^2$ was further given a scaled inverse χ^2 prior distribution as

$$\sigma_{t_j}^2|v_{\sigma_t}, \tau_{\sigma_t}^2 \sim \chi^{-2}(d.f. = v_{\sigma_t}, scale = \tau_{\sigma_t}^2) \, j = 2, ..., p.$$

The hyperparameter, $\tau_{\sigma_t}^2$, is given a Gamma (shape = 1, rate = 1) prior and need to be estimated in the model. All of other underlying parameters, including $\boldsymbol{\delta}_j$ $(j = 1, ..., p)$, $\boldsymbol{V}_{\delta_j}$ $(j = 1, ..., p)$, $\boldsymbol{\Psi}_V$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}_\Sigma$, were separately given a prior as the same as in model (1).

We herein specified the hyperparameters in the scalar model as $\{v_\Sigma = 4.2, v_V = 4.2, v_{\sigma_t} = 4.2, \mu_t = 0\}$. The scalar model was solved through Gibbs sampling similar to the matrix model, and the details on the Gibbs sampling procedures were given in Supplementary Method, Section 3.

Prediction of multiple traits. We used a training population to predict the effect of each marker, $\boldsymbol{\alpha}_j$ $(j = 1, ..., p)$, and then estimated the total genetic values of m traits for an individual as $g_i = \sum_{j=1}^p Z_{ij}\boldsymbol{\alpha}_j$, where g_i is an m-element vector of additive genetic values of m traits for the ith individual. Prediction accuracy is defined as the correlation between the estimated total genetic values and known true total genetic values in the validation population. If the true total genetic values are unknown in the validation population, we used predictive ability to compare the various methods. Predictive ability is defined as the correlation between the estimated total genetic values and their corresponding phenotypes in the validation population.

To validate the performance of the proposed methods, we conducted extensive simulation and the analyses on the 16th QTL-MAS workshop data set and the heterogeneous stock mice real data to compare it with the single-trait and multi-trait Bayes A models with respect to prediction accuracy or predictive ability.

Simulation analyses
We conducted extensive simulation by considering different levels of genetic and phenotypic correlations between traits, SNP densities and the numbers of underlying QTLs to validate the prediction accuracy and robustness of our proposed models.

Data simulation. A whole-genome simulation program, GPOPSIM (Zhang et al., 2012), which is based on the mutation-drift equilibrium model, was used to generate biallelic markers and QTLs. The simulation started with a base population of 100 individuals (50 unrelated males and 50 unrelated females), followed by 5000 generations of random mating with the same population size, to generate data with realistic LD patterns caused by mutation and drift. The entire genome was composed of one chromosome of length 1 Morgan. All of 20 001 potential SNP markers were equally spaced on the chromosome with a QTL placed directly in the middle of each interval of adjacent markers. The mutation rate for both SNP markers and QTLs was specified to be 1.0E-4 per locus per generation. In generation 5001, the population size was expanded to 500 by randomly mating 50 males with 50 females from generation 5000. Generation 5002 was generated by randomly mating 50 males with 50 females from generation 5001 and also had a population size of 500. Similar to the study by Yang and Tempelman (2012), we randomly selected 30 QTLs with a minor allele frequency (MAF) > 0.05 to generate true breeding values for individuals in generations 5001 and 5002. Following Jia and Jannink (2012), QTL effects on two phenotypic traits were sampled from a standard bivariate normal distribution with correlation 0.5, which assumes a certain level of pleiotropy effects at all loci. The true breeding value for each individual was the

genotype-based sum of the QTL effects for each trait. The total genetic covariance matrix for the two traits was determined as $\sum_{k=1}^{30} 2(1 - p_k)p_k\boldsymbol{g}_k\boldsymbol{g}_k'$, where p_k and $\boldsymbol{g}_k$ is the MAF and the sampled vector of effect at QTL k, respectively. Normal error deviates were added to achieve heritabilities of 0.5 for trait 1 and 0.1 for trait 2. All individuals have phenotypes on both traits. The covariance of errors between traits was set to zero.

We considered the above simulation scenario as the default validation data set. Accordingly, we perturbed a single-simulation parameter at a time from the default scenario to generate simulated data under other scenarios for comparison. Perturbed parameters for simulations included genetic correlation between traits (0.2, 0.5 and 0.8), error correlation (−0.2, 0 and 0.2), and number of selected QTLs (30 vs 300). Each of these simulation scenarios was repeated 30 times for producing convincing results. For all simulation scenarios, generation 5001 was used as the training population and generation 5002 as the validation population.

Estimation of prediction accuracy in simulated data. For each simulated data set, we filtered the 20 001 SNP markers using a criterion of MAF > 0.05, leaving ~ 4100 markers. The filtered SNPs were sorted merely according to the physical position in base pair and then used to predict genomic breeding values for the validation population. The prediction accuracy for each data set was calculated as the correlation between the estimated genomic breeding values and the known true genomic breeding values in the validation population. In each scenario, we used a paired t-test to test if there was a significant difference of prediction accuracy between our proposed models and the conventional models based on the 30 replicates.

Furthermore, based on the default scenario, we also employed the subsets of SNPs through sampling every 10th and 25th SNPs from the full set of SNPs respectively to predict genomic breeding values for testing the effect of LD between adjacent markers on the prediction accuracy.

Analysis of the 16th QTL-MAS workshop data set
In addition to our simulated data sets, we analyzed the 16th QTL-MAS workshop data set (http://qtl-mas-2012.kassiopeagroup.com/en/index.php) to further validate the prediction accuracy of our proposed methods.

The workshop data were generated to investigate performance of various WGP approaches when dealing with multiple correlated traits. Specifically, a base population of 1020 unrelated individuals (20 males and 1000 females) was generated with a 499.750 Mb long genome consisting of five chromosomes. Each chromosome had a size of 99.95 Mb and carried 2000 equally distributed SNP. A total of 50 QTLs were generated, and each QTL had an effect on at least two traits. The detailed information of the QTLs can be found at http://qtl-mas-2012.kassiopeagroup.com/en/dataset.php. Each of four generations (G1–G4) consisted of 20 males and 1000 females and was generated from the previous generation by randomly mating each male with 51 females. The female prolificacy was set to 1 except for the 20 dams of male that generated two offspring (1 male and 1 female). No Generation overlapped in the process of data simulation. The mimic phenotypes consisted of three genetically correlated milk production traits.

In the analyses, we removed all SNP markers with MAF = 0, leaving 9969 markers for further analyses. The SNP markers had been sorted by physical position. The 3000 females from G1–G3 had observations for all three traits, which were used as the training population. The remaining 1020 individuals from G4 with known total genetic values were used as the validation population. The prediction accuracies obtained by the different multi-trait methods were compared with officially reported results (http://qtl-mas-2012.kassiopeagroup.com/en/index.php). Briefly, four different groups independently submitted their results, and all the 14 analytical methods were merely focused on single-trait prediction, including GBLUP, the Bayesian alphabet and various regularized regression methods. The highest prediction accuracies obtained by the methods were 0.794, 0.853 and 0.828 for traits 1, 2 and 3, respectively, which were used as the results of best single-trait methods and as a benchmark for comparison with the results herein.

Real data analyses on heterogeneous stock mice data set

The heterogeneous stock mice data set (http://mus.well.ox.ac.uk/GSCAN) consists of 1940 genotyped mice generated based on intercross mating among eight inbred strains (Valdar *et al.*, 2006). We randomly selected two pairs of immunological traits, %CD4+/CD3+ and %CD8+, and %CD4+ and %CD8+, to compare prediction accuracy between the methods used in our study. Phenotypes for these traits had been preadjusted for marginally significant covariates.

Following the study by Speed and Balding (2014), SNPs were removed if they had MAF <0.01, HWE $P<10^{-4}$, or call rate <0.99. Missing SNP genotypes were imputed on the basis of their corresponding allelic frequencies in the data set following Legarra *et al.* (2008). Individuals were removed unless they had phenotypes of all the selected traits. After quality control, 1404 individuals having both phenotypes and genotypes were kept, and each individual had genotypes of 9159 SNPs mapped to autosomes. In addition, we ordered SNP genotypes on the basis of their genetic positions along the chromosome.

We randomly divided individuals into two nearly equal-sized subsets as training and validation sets, respectively. This was replicated 20 times for performance comparison among different prediction methods.

RESULTS

Results from the simulations

Prediction accuracy under different LD levels of adjacent markers. We investigated the performance of the proposed methods with default simulation parameters while dealing with SNPs with different densities to explore the effect of LD extent on prediction accuracy. Table 1 presented the three average LD levels ($r^2 = 0.33$, 0.22 and 0.14) of adjacent markers corresponding to the full set of markers and two subsets of markers sampled from the full set of markers with intervals of 10 and 25, respectively.

As clearly shown in Figure 1, our proposed methods always exhibited the best performance among all methods across different LD levels. Especially, with the increase of average LD, our methods were becoming more advantageous over other analytical approaches for the low-heritability trait 2.

Prediction accuracy under scenarios with a varied number of QTLs. Besides the default 30-QTL scenario, we considered a 300-QTL scenario for further comparison among different analytical methods. Despite the clear superiority of the proposed methods under the 30-QTL scenario, there were no significant differences of prediction accuracy between our proposed models and the traditional multi-trait Bayes A model for the 300-QTL scenario (Figure 2).

Prediction accuracy under scenarios with varied genetic correlations. It can be clearly shown from Figure 3 that our proposed methods were always among the top rank on the prediction accuracy compared with other analytical models under different levels of genetic correlations.

**Table 1 Number of used markers and LD level across 30 replicates under varied scenarios**

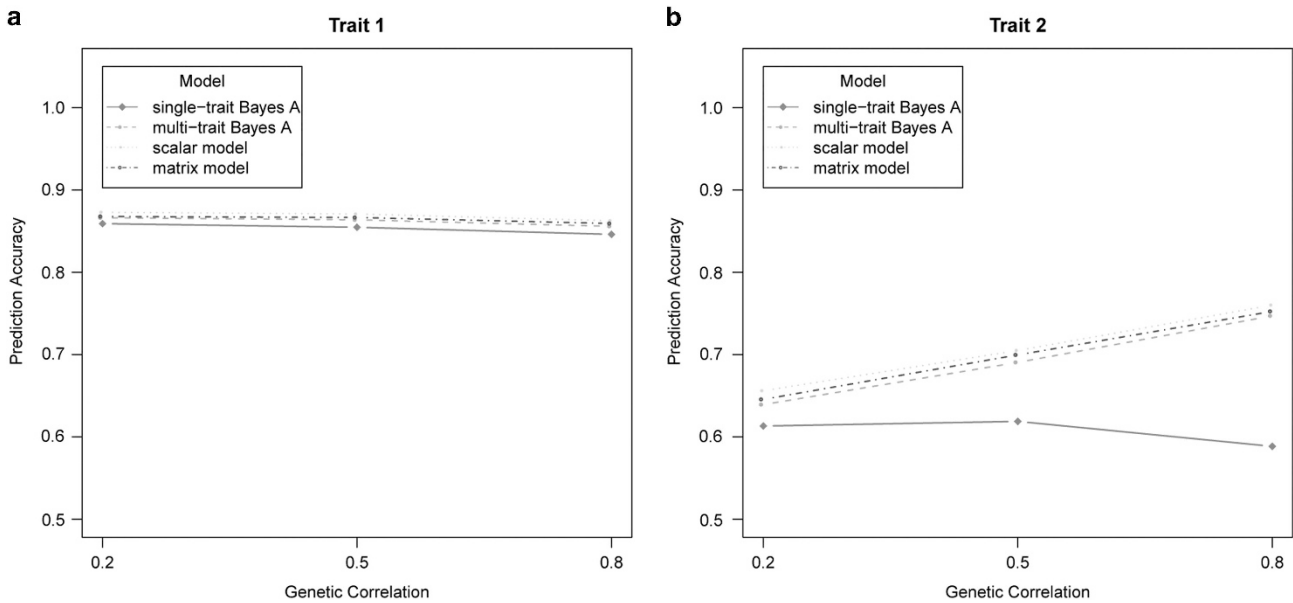| Scenario no.	Genetic correlation	Error correlation	No. of QTLs	Way of using markers	Average no. of used markers (std.)	Average LD level of adjacent markers (std.)[a]
1 (default)	0.5	0	30	All	4119 (159)	0.333 (0.016)
				Every 10th	411 (16)	0.220 (0.018)
				Every 25th	164 (6)	0.136 (0.015)
2	0.5	0	300	All	4140 (126)	0.333 (0.015)
3	0.2	0	30	All	4044 (162)	0.336 (0.013)
4	0.5	−0.2	30	All	4030 (164)	0.338 (0.015)
5	0.5	0.2	30	All	4071 (141)	0.334 (0.013)

Abbreviation: QTLs, quantitative trait loci.
[a]Average $=$ sum(A_i)/30 and std. $=$ sqrt(sum((A_i-Average)^2)/29), where A_i is the average LD level in the *i*th replicate.
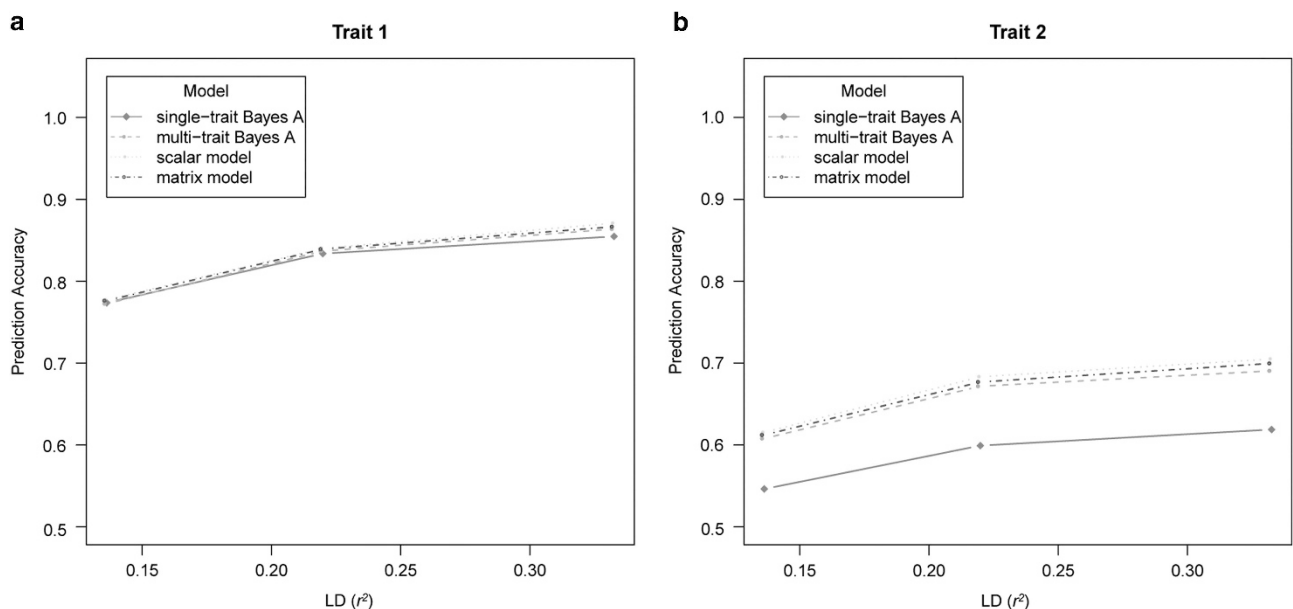
Figure 1 Prediction accuracies of various models for the high-heritability trait (**a**) and the low-heritability trait (**b**) under different LD levels of adjacent markers.
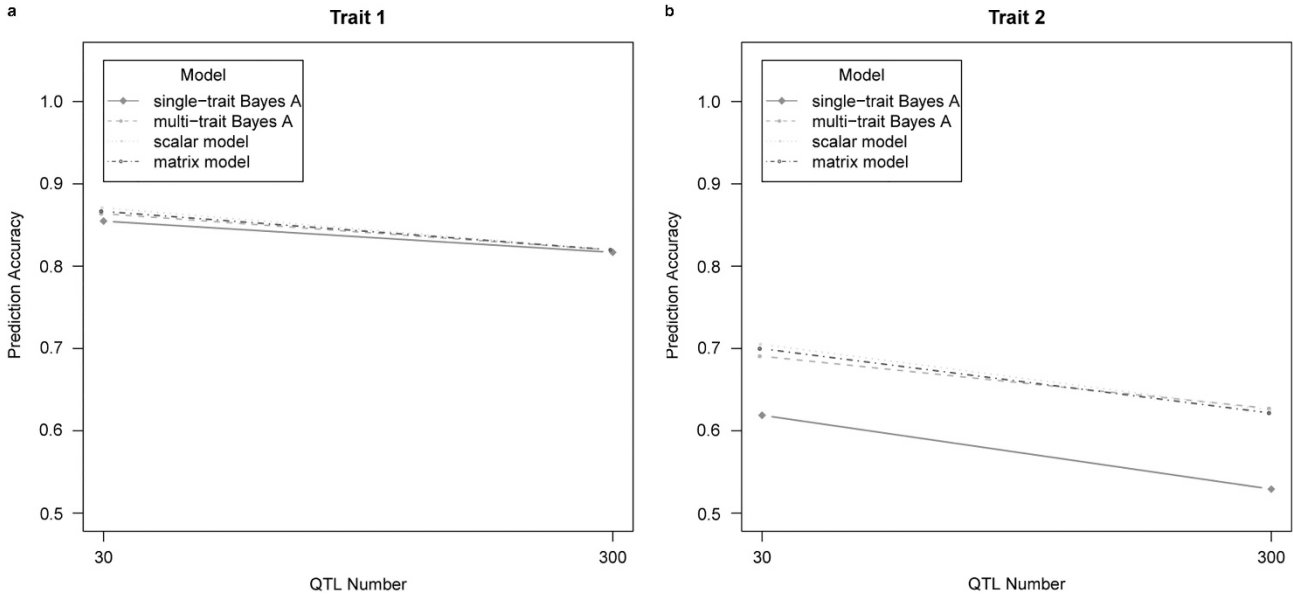
Figure 2 Prediction accuracies of various models for the high-heritability trait (**a**) and the low-heritability trait (**b**) under scenarios with a varied number of underlying QTLs.
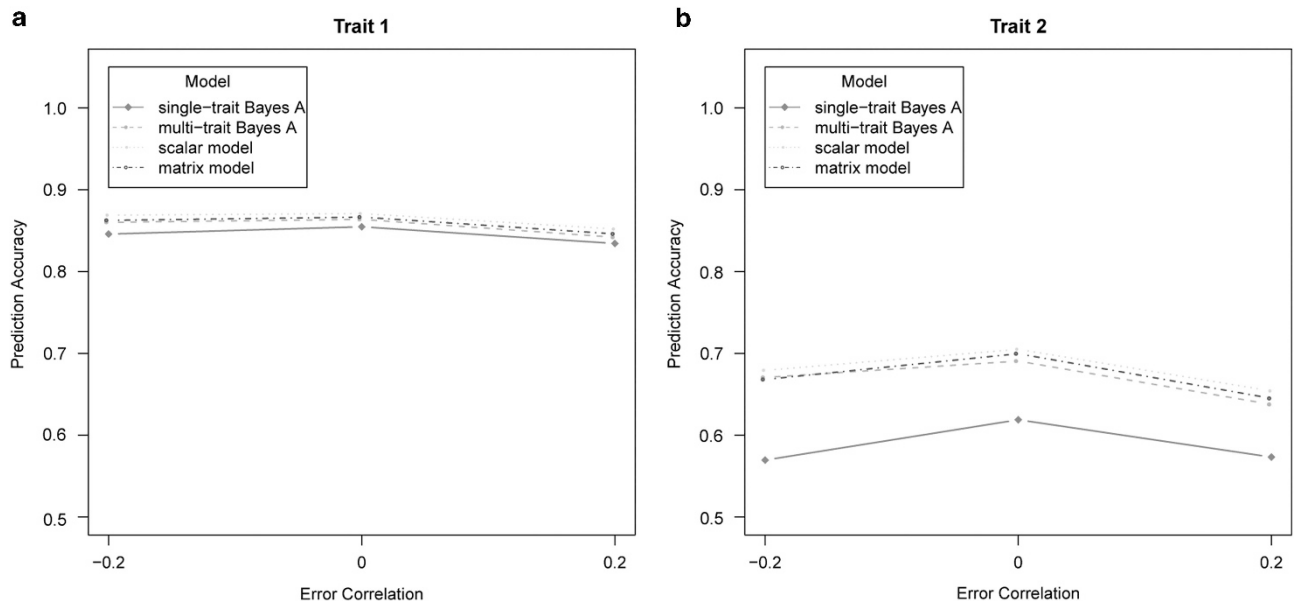

Figure 3 Prediction accuracies of various models for the high-heritability trait (**a**) and the low-heritability trait (**b**) under scenarios with varied genetic correlations between traits.

Furthermore, all multi-trait prediction methods significantly outperformed the single-trait analyses on the traits investigated under all scenarios.

It is worth to note that the gain in prediction accuracies of multi-trait over single-trait models were becoming more significant for the low-heritability trait 2 as genetic correlation increased between traits. In contrast, for the high-heritability trait 1, no obvious change in prediction accuracy was observed as the genetic correlation increased from 0.2 to 0.8. The similar trend was also observed in previous studies (Calus and Veerkamp, 2011; Jia and Jannink, 2012). This implied that the superiority of our proposed models over the multi-trait Bayes A model resulted from the feature that the model accounted for the dependence on the effects of adjacent markers.

Prediction accuracy under scenarios with varied error correlations between traits. Besides the default scenario with error correlation of 0.0, we also considered another two scenarios with varied error correlations of − 0.2 and 0.2, respectively. From Figure 4, for both low- and high-heritability traits, the prediction accuracies were largely consistent for each individual analytical across the three scenarios. Similar profiles as those aforementioned, our proposed methods still achieved highest prediction accuracies across various scenarios, and all the multi-trait prediction methods rendered obvious advantages over the single-trait approach.

Results from the analysis of the 16th QTL-MAS workshop data set
We also used the 16th QTL-MAS workshop data set to validate our proposed methods. As shown in Figure 5, multi-trait methods,

Figure 4 Prediction accuracies of various models for the high-heritability trait (**a**) and the low-heritability trait (**b**) under scenarios with varied error correlations between traits.
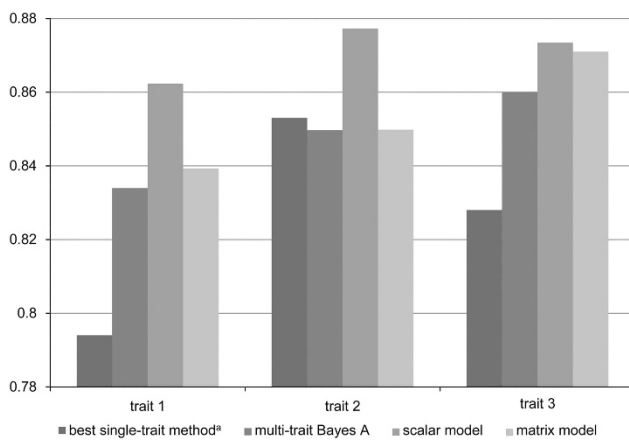
Figure 5 Prediction accuracies of various models for the 16th QTL-MAS workshop data set. [a]Means the best single-trait method officially reported in the 16th QTL-MAS workshop.

including the multi-trait Bayes A model, the scalar model and the matrix model, showed considerable advantage over single-trait methods.

It can be found clearly that our proposed scalar model obtained the best results with regard to the 16th QTL-MAS workshop data set. The prediction accuracies of the scalar model for the three traits increased by as high as 8.6%, 2.8% and 5.5%, respectively, compared with the reported best single-trait methods.

Results from the analysis of real heterogeneous stock mice data set
As shown in Table 2, for both of trait pairs, our proposed models always ranked top two in prediction abilities among all four methods for each of traits, and all three joint prediction methods outperformed the single-trait model significantly ($P < 0.01$).

Similar as simulation analyses, our antedependence models obtained predictive abilities significantly higher than the multi-trait Bayes A for both of trait pairs. It is worth noting herein that the proposed matrix model had a better performance than the scalar

model for each of trait ($P < 0.01$) in the first trait pair, exhibiting its potential superiority in practice in contrast to its simplified version, the scalar model.

DISCUSSION
In WGP analyses, the inferred adjacent SNP effects should be spatially correlated due to chromosomally proximal effects of potential QTL (Yang and Tempelman, 2012), which is most likely in LD with the SNPs surrounding it, even when no biological mechanism exists among these SNPs. The antedependence model can model such nonstationary correlated effects via incorporating antedependence parameters between adjacent SNPs into the model. This is the theoretical basis that the antedependence model can work better than the traditional independence models given no additional information about markers but genotypes.

In the current study, we firstly developed two types of multi-trait antedepedence models for joint phenotype prediction. We hypothesized that our proposed WGP methods could enhance the prediction performance owing to considering information of correlation among traits and adjacent SNPs simultaneously. This theoretical superiority was further validated herein through extensive simulation and real-data analyses, and the resulting prediction patterns under various scenarios consistently revealed the significant performance gain of our proposed methods compared with the corresponding counterpart methods. Our proposed WGP methods offer a feasible alternative in joint prediction of multiple complex traits in human genetic epidemiology as well as livestock breeding. We have developed C programs to implement the corresponding prediction models, which are freely available upon request.

In our methods, we employed a matrix and a scalar antedependence parameters, respectively, for our proposed multivariate prediction method. The introduced antedepedence parameter matrix T considered respective effect correlation between pair of adjacent markers for each of correlated traits, which should outperform the scalar model in theory. However, our simulation data analyses and the analysis of the 16th QTL-MAS data set have not demonstrated obvious advantage of the matrix model over the scalar model. This may be owing to the

Table 2 Predictive ability of various methods for the heterogeneous stock mice data across 20 replicates

Trait pair	Trait	Predictive ability of various methods			
		Single-trait Bayes A	Multitrait Bayes A	Scalar model	Matrix model
1	%CD4+/CD3+	0.609	0.616	0.620	0.625
	%CD8+	0.663	0.664	0.668	0.671
2	%CD4+	0.358	0.372	0.375	0.375
	%CD8+	0.665	0.667	0.670	0.670

extreme case that was mimicked in our simulation scenarios as well as the 16th QTL-MAS data set. Specifically, in our simulation, all simulated QTL effects were assumed being sampled from the same distribution and contributing to both traits simultaneously. In the 16th QTL-MAS data set, each simulated QTL contributed to at least two traits out of all three. In such case, the antedependence parameters for each of correlated traits tend to be equal. Accordingly, the scalar model suits the data sets with the such genetic structures better than the matrix model and was prone to obtaining a higher prediction accuracy.

In the simulation scenario when considering all 300 QTLs with very small effects, the antedependence parameters are inefficient to model the dependence between SNPs, lowering its advantage over the traditional methods. Another aspect in our simulation is that a small base population of 100 followed by a large number of generations was simulated. Although this scheme was usually adopted in previous studies (Zhang *et al.*, 2010; Calus and Veerkamp, 2011; Yang and Tempelman, 2012), it would generate data with extremely small effective population size such that the simulated data could not represent the population structure of real data.

Owing to limited simulation scale in the studies, it is not feasible to simulate various scenarios with genetic structures consistent of all possible built-in mechanisms for performance validation. Hence we turned to further perform real-data analyses on two sets of complex traits from a publicly available heterogeneous stock mice data to explore the feasibility of our methods. The results clearly demonstrated that the proposed multivariate antedependence model outperformed corresponding traditional counterparts, and the matrix model has also shown the advantage over its simplified version, the scalar model, as we expected.

Yang and Tempelman (2012) have reported that the antedependence methods outperformed their corresponding classical counterparts in single-trait prediction, and the increase in prediction accuracy contributed by the single-trait antedependence model was generally <0.05. This increase in prediction accuracy may be not comparable to that for low-heritability traits benefiting from joint prediction of multiple highly correlated traits which can be usually >0.10 (Jia and Jannink, 2012). This has been clearly reflected in our analyses. As shown in Figures 1,2,3,4, the low-heritability trait generally benefited more from joint prediction of multiple traits than the high-heritability trait in various scenarios.

A different aspect in our prediction model from that by Yang and Tempelman (2012) is regarding the expectation of the prior of the antedependence parameter. In the study by Yang and Tempelman (2012), a marker-specific antedependence parameter was assumed a normal distributed prior with unknown expectation which needs to be estimated in the model. In contrast, we directly set the prior expectation as zero. Our consideration for this is that the sign of the additive effect at any marker is merely determined by the way we code its genotypes (for example, genotypes *BB* can be coded as 0 or 2)

and the antedependence parameter should have a sign determined by the signs of SNPs effects therein disregarding the residual term of Equation (2) aforementioned. Thus, as the genotypes of biallelic markers are coded as the number of copies of one arbitrary allele in our study, it is equally probable for the antedependence parameter to be positive or negative, and then it is reasonable to assume its expectation as zero.

It should be pointed out herein that although we focused on the use of common variants in current study, rare variants can also be readily incorporated into our proposed method by drawing idea from the collapsing approach (Morris and Zeggini, 2010). It will be imperative to incorporate rare variants into prediction in future for the situation where resequence data instead of traditional SNP chips are widely used in WGP.

Furthermore, our developed antedependence version of the multi-trait Bayes A models can be modified to other types of Bayes model, for example, Bayes B and Bayes $C\pi$ (Habier *et al.*, 2011), which have been two popular Bayesian variable selection methods for single-trait WGP. For example, a multi-trait antedependence version of Bayes B can be developed based on model (2) aforementioned with a mixture prior distribution of $\boldsymbol{\delta}_j$ ($j=1,..,p$) as

$$(\boldsymbol{\delta}_j | \pi, \boldsymbol{V}_{\delta_j}) = \begin{cases} \sim N(\boldsymbol{0}, \boldsymbol{V}_{\delta_j}) & \text{probability } (1-\pi) \\ 0 & \text{probability } \pi \end{cases}$$

. Then the conditional posterior distribution of each unknown parameter can be derived drawing idea from the single-trait antedependence-based Bayes B (Yang and Tempelman, 2012).

Aside from predicting yet-to-be observed phenotypes, our proposed models are also useful in QTL-mapping studies according to previous reports on applying WGP methodology to GWAS (Peters *et al.*, 2012; Garrick and Fernando, 2013). By adding up genetic variance contributed by SNPs in a SNP window, we can calculate the genetic variance of consecutive SNP windows along the genome and then consider windows contributing relatively large genetic variance to be QTLs (Peters *et al.*, 2012). Bootstrap analysis can be further used to determine the significance of detected QTLs as described by Peters *et al.* (2012).

We believed that it is necessary to find a way to integrate multiple advantageous strategies for developing an optimal WGP model, for example, joint prediction of multiple traits, the antedependence model, incorporating dominant and epistatic effects (for example, Wang *et al.*, 2012; Nishio and Satoh, 2014), modeling genotype × environment interaction (for example, Crossa, 2012). As these strategies are utilizing different principles, it is anticipated that the strength contributed by each of the strategies can be accumulated if these various strategies are integrated in a proper way. Our study can be considered as an example of such integration, which takes advantage of both antedependence and multivariate models for joint prediction.

DATA ARCHIVING

All simulated data analyzed in the present study have been deposited in Dryad (http://doi.org/10.5061/dryad.dd60v).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Bush WS, Sawcer SJ, de Jager PL, Oksenberg JR, McCauley JL, Pericak-Vance MA *et al.* (2010). Evidence for polygenic susceptibility to multiple sclerosis–the shape of things to come. *Am J Hum Genet* **86**: 621–625.

Calus MP, Veerkamp RF (2011). Accuracy of multi-trait genomic selection using different methods. *Genet Sel Evol* **43**: 26.

Crossa J (2012). From genotype x environment interaction to gene x environment interaction. *Curr Genomics* **13**: 225–244.

de los Campos G, Gianola D, Allison DB (2010). Predicting genetic predisposition in humans: the promise of whole-genome markers. *Nat Rev Genet* **11**: 880–886.

de los Campos G, Klimentidis YC, Vazquez AI, Allison DB (2012). Prediction of expected years of life using whole-genome markers. *PloS One* **7**: e40964.

de Los Campos G, Vazquez AI, Fernando R, Klimentidis YC, Sorensen D (2013). Prediction of complex human traits using the genomic best linear unbiased predictor. *PLoS Genet* **9**: e1003608.

Garrick DJ, Fernando RL (2013). Implementing a QTL detection study (GWAS) using genomic prediction methodology. *Methods Mol Biol* **1019**: 275–298.

Gianola D (2013). Priors in whole-genome regression: the bayesian alphabet returns. *Genetics* **194**: 573–596.

Gianola D, Fernando RL, Stella A (2006). Genomic-assisted prediction of genetic value with semiparametric procedures. *Genetics* **173**: 1761–1776.

Gianola D, Wu XL, Manfredi E, Simianer H (2010). A non-parametric mixture model for genome-enabled prediction of genetic value for a quantitative trait. *Genetica* **138**: 959–977.

Habier D, Fernando RL, Kizilkaya K, Garrick DJ (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics* **12**: 186.

Hindorff L, MacArthur J, Morales J, Junkins H, Hall P, Klemm A *et al.* (2014). A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies.

Jia Y, Jannink JL (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics* **192**: 1513–1522.

Jiang C, Zeng ZB (1995). Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* **140**: 1111–1127.

Korte A, Vilhjalmsson BJ, Segura V, Platt A, Long Q, Nordborg M (2012). A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat Genet* **44**: 1066–1071.

Legarra A, Robert-Granie C, Croiseau P, Guillaume F, Fritz S (2011). Improved Lasso for genomic selection. *Genet Res (Camb)* **93**: 77–87.

Legarra A, Robert-Granie C, Manfredi E, Elsen JM (2008). Performance of genomic selection in mice. *Genetics* **180**: 611–618.

Liu J, Liu Y, Liu X, Deng HW (2007). Bayesian mapping of quantitative trait loci for multiple complex traits with the use of variance components. *Am J Hum Genet* **81**: 304–320.

Liu J, Pei Y, Papasian CJ, Deng HW (2009). Bivariate association analyses for the mixture of continuous and binary traits with the use of extended generalized estimating equations. *Genet Epidemiol* **33**: 217–227.

Makowsky R, Pajewski NM, Klimentidis YC, Vazquez AI, Duarte CW, Allison DB *et al.* (2011). Beyond missing heritability: prediction of complex traits. *PLoS Genet* **7**: e1002051.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ *et al.* (2009). Finding the missing heritability of complex diseases. *Nature* **461**: 747–753.

Meuwissen TH, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**: 1819–1829.

Minka TP (2001). *Bayesian Linear Regression*. Technical report, MIT Media Lab: Cambridge, MA, USA.

Morris AP, Zeggini E (2010). An evaluation of statistical approaches to rare variant analysis in genetic association studies. *Genet Epidemiol* **34**: 188–193.

Nishio M, Satoh M (2014). Including Dominance Effects in the Genomic BLUP Method for Genomic Evaluation. *PloS One* **9**: e85792.

Peters SO, Kizilkaya K, Garrick DJ, Fernando RL, Reecy JM, Weaber RL *et al.* (2012). Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. *J Anim Sci* **90**: 3398–3409.

Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al.* (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**: 748–752.

Speed D, Balding DJ (2014). MultiBLUP: improved SNP-based prediction for complex traits. *Genome Res* **24**: 1550–1557.

Stahl EA, Wegmann D, Trynka G, Gutierrez-Achury J, Do R, Voight BF *et al.* (2012). Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet* **44**: 483–489.

Valdar W, Solberg LC, Gauguier D, Burnett S, Klenerman P, Cookson WO *et al.* (2006). Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat Genet* **38**: 879–887.

VanRaden PM (2008). Efficient methods to compute genomic predictions. *J Dairy Sci* **91**: 4414–4423.

VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF *et al.* (2009). Invited review: reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci* **92**: 16–24.

Vazquez AI, de los Campos G, Klimentidis YC, Rosa GJ, Gianola D, Yi N *et al.* (2012). A comprehensive genetic approach for improving prediction of skin cancer risk in humans. *Genetics* **192**: 1493–1502.

Wang D, Salah El-Basyoni I, Stephen Baenziger P, Crossa J, Eskridge KM, Dweikat I (2012). Prediction of genetic values of quantitative traits with epistatic effects in plant breeding populations. *Heredity (Edinb)* **109**: 313–319.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR *et al.* (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**: 565–569.

Yang W, Tempelman RJ (2012). A Bayesian antedependence model for whole genome prediction. *Genetics* **190**: 1491–1501.

Zhang Z, Ding X, Liu J, Ni G, Li J, Zhang Q (2012). *4th International Conference on Computer Modeling and Simulation (ICCMS 2012), Vol. 22*. IACSIT Press: Hong Kong, China. pp 87–93.

Zhang Z, Liu J, Ding X, Bijma P, de Koning DJ, Zhang Q (2010). Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. *PloS One* **5**: e12648.

Zimmerman DL, Nunez-Anton VA (2010). *Antedependence Models for Longitudinal Data*. Chapman & Hall: London/New York.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)