


# Application of Whole-Genome Prediction Methods for Genome-Wide Association Studies: A Bayesian Approach

Rohan FERNANDO , Ali TOOSI, Anna WOLC, Dorian GARRICK, and Jack DEKKERS

Data that are collected for whole-genome prediction can also be used for genome-wide association studies (GWAS). This paper discusses how Bayesian multiple-regression methods that are used for whole-genome prediction can be adapted for GWAS. It is argued here that controlling the posterior type I error rate (PER) is more suitable than controlling the genomewide error rate (GER) for controlling false positives in GWAS. It is shown here that under ideal conditions, i.e., when the model is correctly specified, PER can be controlled by using Bayesian posterior probabilities that are easy to obtain. Computer simulation was used to examine the properties of this Bayesian approach when the ideal conditions were not met. Results indicate that even then useful inferences can be made.

**Key Words:** Bayesian multiple regression; Genome-wide association studies; Genomewide error rate; Posterior type I error rate; Whole-genome prediction.

## 1. INTRODUCTION

High-density SNP genotypes are currently being used in livestock for whole-genome prediction (VanRaden et al. 2009; Hayes et al. 2009; Habier et al. 2010; Wolc et al. 2011). This requires obtaining genotypes and phenotypes on several thousand animals in a training population to estimate effects of the SNP genotypes on the traits of interest. These estimated SNP effects are then used to predict the breeding values of selection candidates that may not have any phenotypes recorded but have been genotyped (Meuwissen et al. 2001). The genotype and phenotype data obtained for whole-genome prediction can also be used for genome-wide association studies (GWAS) to locate causal variants (QTL) for traits of economic importance.

---

Rohan Fernando (✉), Ali Toosi, Anna Wolc, Dorian Garrick, and Jack Dekkers are in Department of Animal Science, Iowa State University, Ames, IA 50011-3150, USA. (E-mail: [rohan@iastate.edu](mailto:rohan@iastate.edu)).

© 2017 The Author(s). This article is an open access publication  
*Journal of Agricultural, Biological, and Environmental Statistics*, Volume 22, Number 2, Pages 172–193  
DOI: 10.1007/s13253-017-0277-6

Many GWAS for quantitative traits are based on testing one SNP at a time using simple regression models or using mixed models with a fixed substitution effect of the SNP genotype along with a polygenic effect correlated according to a pedigree-based or a genomic relationship matrix to capture the effects of all the other genes. Such GWAS have been successful in detecting many associations, but the established associations typically explain only a small fraction of the genetic variability of quantitative traits (Maher 2008; Manolio et al. 2009; Visscher et al. 2010).

Although it can be shown that a mixed model that uses a genomic relationship matrix is equivalent to a multiple-regression model that simultaneously fits all SNPs as random effects and jointly explains a large proportion of the genetic variance (Hayes et al. 2010; Fan et al. 2011), any given SNP in these analyses may have only a weak association even with a close QTL. The reason for this is that in these multiple-regression models, the association of a SNP with the phenotype is at best a partial association conditional on all the other SNPs. Further, in a high-density SNP panel, many SNP genotypes within a narrow genomic segment are expected to be highly correlated with each other and with any QTL that are close to them. So, any one SNP may contribute only a little more to explain the variability of the QTL in addition to the other SNPs in the neighborhood. These problems are exacerbated by the number of SNP covariates being much larger than the number of observations. On the other hand, even if each SNP in a neighborhood is only weakly associated with a QTL, the SNPs in the neighborhood may jointly explain much more of the variability of a QTL than any SNP by itself. Therefore, in multiple-regression models, SNPs in a genomic window should be used to locate QTL (Sahana et al. 2010; Hayes et al. 2010; Fan et al. 2011).

Inferences on genomic windows by frequentist methods, however, are computationally very challenging because they require repeated analyses of the data with bootstrap or permuted samples to obtain significance levels for tests (Hayes et al. 2010; Fan et al. 2011). It can be shown that Bayesian posterior probabilities obtained from a single analysis can be used to make inferences on genomic windows (Sahana et al. 2010). This approach to inference is related to the frequentist approach of controlling the posterior type I error rate (PER), which is the conditional probability of a false positive (type I error) given a positive (significant) result (Morton 1955). It is PER that traditionally has been used in human genetics to control false positives in linkage analyses of monogenic traits (Elston 1997) rather than the usual type I error rate, which is the conditional probability of a false positive result given that the null hypothesis is true.

A requirement for controlling PER is knowledge of the distribution of the test statistic under the null hypothesis of no association, which is also required to control the usual type I error rate. In addition to this requirement, controlling PER requires knowing the proportion  $\pi$  of SNPs for which the null hypothesis is true and the average power of the test, which is the average probability of rejecting the null hypothesis when it is not true. These quantities are almost never known in a GWAS of a quantitative trait, and thus PER cannot be controlled in the sense that the usual type I error rate can be controlled (Elston 1997).

In a QTL mapping study, Soller and colleagues were the first to show that  $\pi$  can be estimated from a histogram of  $p$ -values (Mosig et al. 2001). In this multiple-test setting, they showed how the estimate of  $\pi$  can be used to obtain an adjusted false discovery rate (FDR) (Benjamini and Hochberg 1995) that results in increased power to detect QTL relative

to the usual FDR approach. Their seminal paper also showed how to estimate the average power of tests (Mosig et al. 2001). Such estimates of  $\pi$  and average power can also be used to estimate PER for a test with a specified significance level (type I error rate). However, when the resulting estimates of PER are used for inference, the estimates of  $\pi$  and average power are treated as known values because it is not straightforward to incorporate the errors in these estimates into the calculations. In contrast to frequentist methods, when the Bayesian multiple-regression models (Meuwissen et al. 2001) that are used for whole-genome prediction are applied to GWAS, posterior probabilities that are similar to PER can be obtained even without the requirement of knowing the null distribution of any test statistic. Further, in Bayesian analyses,  $\pi$  and the partial regression coefficients of markers, which determine average power, can be formally treated as unknowns such that their uncertainties are incorporated in the inference.

It can be shown that using PER or related posterior probabilities for inference has the following advantage. Suppose PER is controlled, for example, at 0.05 for each of many tests, which may be dependent. Then, among the positive results that accumulate over many independent experiments, the proportion of false positives (PFP) will converge to 0.05. As a result, regardless of the number of tests in an experiment and their interdependence, the proportion of false positives that accumulate in the literature can be controlled by controlling the PER for individual tests. It has been shown that controlling PER for individual tests in an experiment is related to controlling FDR or its close relatives pFDR (Storey 2002) or PFP (Southey and Fernando 1998; Fernando et al. 2004) for the collection of tests from the experiment.

Posterior probabilities that are related to PER are still not widely used for inference in GWAS by plant and animal geneticists. One reason for this is that to manage false positives in a GWAS, controlling the genomewide error rate (GER), which is the probability of one or more type I errors in the complete collection of tests, is better accepted than controlling the PER for individual tests or FDR for the experiment. A second reason could be that the relationship between PER and the Bayesian posterior probabilities for inference in GWAS has not been studied in a realistic setting. A third reason for preferring GER over PER is given by Chen and Storey (2006). Their reasoning can be interpreted as follows. In linkage analysis, a QTL is in principle linked to every marker on the same chromosome, and this makes the null hypothesis that a marker is not linked to a QTL false for every marker on a chromosome that has even a single QTL. Then, a marker with a significant linkage signal on a chromosome that has a QTL will be a true positive even if the marker is not close to the QTL. This makes controlling the proportion of false positives among significant results meaningless (Chen and Storey 2006). They refer to this problem as signal dependence. Thus, the objectives of this paper are to: (1) address the problem of signal dependence and review the advantages of managing false positives in GWAS by controlling PER or related measures such as FDR or PFP rather than by controlling the GER, and (2) use computer simulation to examine the relationship between Bayesian posterior probabilities of association and the frequency of a true association.

In Sect. 2.1, the problem of signal dependence is addressed and the advantages of managing false positives by controlling PER rather than GER are reviewed. Section 2.2 describes how multiple-regression models that were developed for whole-genome prediction can be

used for GWAS. Section 2.3 shows how the PER is related to the Bayesian posterior probabilities used for inference in GWAS. A computer simulation that was used to study the properties of the Bayesian approach is described in Sect. 2.4. An application to real data is described in Sect. 2.5. Results from the simulation are discussed and summarized in Sect. 3.

## 2. METHODS

### 2.1. CONTROLLING FALSE POSITIVES

Many papers have discussed the advantages of controlling FDR, and related measures rather than controlling GER (Benjamini and Hochberg 1995; Storey 2002; Fernando et al. 2004; Stephens and Balding 2009). Chen and Storey (2006), however, have argued that measures such as FDR that control the proportion of false positives among significant results are not suitable for controlling false positives in QTL studies. The essence of their argument stems from the implicit assumption that the linkage signal from a QTL spans the whole chromosome. Thus, the null hypothesis that a marker is not linked to a QTL is equivalent to that of no QTL on the same chromosome as the marker. In the approach that is proposed by Chen and Storey (2006), the significance threshold for GER was derived under the null hypothesis of no linkage between markers and QTL. Under this null hypothesis, PER would be 1.0 regardless of the threshold for significance. Further, regardless of the approach used to control false positives, rejection of the null hypothesis of no linkage between markers and QTL does not imply that the significant marker is close to a QTL. To address this problem, others have employed multiple-regression models where the signal from a QTL is localized to a relatively narrow segment of the chromosome (Zeng 1993, 1994; Southey and Fernando 1998).

Suppose data are available from a backcross population derived from two completely homozygous lines. In such a population, the following multiple-regression model can be used to test the null hypothesis that a chromosomal segment with a central marker  $C$  bracketed by two markers  $L$  on the left and  $R$  on the right does not contain any QTL:

$$Y_i = \beta_0 + X_{Li}\beta_L + X_{Ci}\beta_C + X_{Ri}\beta_R + \epsilon_i, \quad (1)$$

where  $Y_i$  is the trait value,  $\beta_0$  is the intercept,  $\beta_L$ ,  $\beta_C$ , and  $\beta_R$  are regression coefficients for three markers  $L$ ,  $C$  and  $R$ , respectively, and  $X_{Li}$ ,  $X_{Ci}$  and  $X_{Ri}$  are 0 or 1 depending on whether the marker genotype is heterozygous or homozygous. It can be shown that  $\beta_C$  is non-null if and only if the chromosomal segment bracketed by  $L$  and  $R$  contains a QTL (Zeng 1993). Thus, with this multiple-regression model, the signal from the trait gene is localized to a chromosomal segment of, for example, 10 cM in length, and testing the null hypothesis that such a segment does not contain a QTL is meaningful in a QTL study. In this setting, we will consider the difference between controlling GER and PER.

A hypothetical example is used next to show why controlling PER is more meaningful than controlling GER in a QTL study. We will assume the trait is controlled by a single gene, the genome consists of 30 chromosomes each of length 1 M, and the tests are to detect the presence of a QTL in a chromosomal segment of length 10 cM.

First, suppose only one chromosomal segment, randomly selected from all 300 segments, is tested. Then, GER can be calculated as:

$$\begin{aligned} \text{GER} &= \alpha\pi \\ &= 0.05 \times \frac{299}{300} \\ &= 0.0498, \end{aligned}$$

where  $\pi$  is the probability that the null hypothesis is true for a randomly selected segment, and  $\alpha = 0.05$  is the type I error rate. The PER for this situation can be much higher. To calculate PER, we assume that if the null hypothesis is not true, the power of the test is  $p = 90\%$  when a type I error rate of  $\alpha = 0.05$  is used. Then, PER is

$$\text{PER} = \frac{\alpha\pi}{\alpha\pi + p(1 - \pi)} \quad (2)$$

$$= \frac{0.05 \times \frac{299}{300}}{0.05 \times \frac{299}{300} + 0.9 \times \frac{1}{300}} \quad (3)$$

$$= 0.94, \quad (4)$$

where the numerator of (4) is the joint probability that the null hypothesis is true and it is rejected, and the denominator is the marginal probability of rejecting the null hypothesis. Thus, using  $\alpha = 0.05$  results in a GER of 0.0498, but a significant result will be a false positive with probability of over 0.9! In other words, among significant results over 90% will be false positives. If we want PER to be 0.05,  $\alpha$  must be reduced to 0.0001584 while maintaining the same power. Thus, controlling GER can be very different from controlling PER.

Suppose now a chromosomal segment of 10 cM is chosen at random from each of the 30 chromosomes. For each of these segments, the probability  $\pi$  that the null hypothesis is true remains  $\frac{299}{300}$ . So, if the significance level is kept at  $\alpha = 0.0001584$ , the PER will be 0.05 for the results from each chromosome. In other words, if results are accumulated over repetitions of the experiment, the proportion of false positives among the significant results would remain 0.05. Thus, the proportion of false positives in the collection of significant results from all chromosomes will also be 0.05.

Given that the 30 tests are independent, the GER can be calculated as follows (Sidak 1967). The probability of not getting a false positive result is  $(1 - \alpha)$  for a randomly sampled segment from any of the 29 chromosomes that do not contain the gene. Thus, the probability of not getting any false positive results from these 29 chromosomes is  $(1 - \alpha)^{29}$ . The probability of not getting a false positive result from the chromosome with the gene is:  $\frac{1}{10} \times 1 + \frac{9}{10}(1 - \alpha)$ , where the first term is the probability that the sampled segment contains the gene and the test does not result in a false positive, and the second term is the probability that the sampled segment does not contain the gene and the test does not result in a false positive. Thus, GER for this situation is:

$$\text{GER} = 1 - (1 - \alpha)^{29} \left[ \frac{1}{10} + \frac{9}{10}(1 - \alpha) \right] = 0.004726. \quad (5)$$

Note that as the number of tests increased, GER also increased from 0.0001584 for one test to 0.004726 for 30 tests, while PER stayed constant at 0.05. So, if GER is used, as the number of tests increases, the significance level for individual tests has to be decreased to keep GER at 0.05, and this will result in lower power. This is not the case with PER.

## 2.2. BAYESIAN MULTIPLE REGRESSION

Here, we will explain how multiple-regression models that were developed for whole-genome prediction can be used for GWAS. In these models, inference is based on posterior probabilities that, as will be shown in Sect. 2.3, are similar to PER.

Following [Meuwissen et al. \(2001\)](#), consider the mixed linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \sum_{j=1}^K \mathbf{z}_j \alpha_j + \mathbf{e}, \quad (6)$$

where  $\mathbf{y}$  is the vector of trait phenotypes,  $\mathbf{X}$  is a matrix relating the vector of non-genetic, fixed effects  $\boldsymbol{\beta}$  to  $\mathbf{y}$ ,  $\mathbf{z}_j$  is a vector of genotype covariates (coded as 0, 1 or 2) for SNP  $j$ ,  $\alpha_j$  is the random, partial regression coefficient for SNP  $j$ , and  $\mathbf{e}$  is a vector of residuals. In this model, the fixed effects are assumed to have a constant prior, the  $\alpha_j$  are a priori assumed independently distributed as

$$\alpha_j | \pi, \sigma_{\alpha_j}^2 = \begin{cases} 0 & \text{with probability } \pi \\ \sim \text{N}(0, \sigma_{\alpha_j}^2) & \text{with probability } (1 - \pi), \end{cases} \quad (7)$$

where the  $\sigma_{\alpha_j}^2$  are a priori assumed independently and identically distributed (iid) scaled inverse chi-square variables with scale  $S_{\alpha}^2$  and degrees of freedom  $\nu_{\alpha}$ . The residuals are assumed iid normal with null mean and variance  $\sigma_e^2$ , with a scaled inverse chi-square prior for  $\sigma_e^2$  with scale  $S_e^2$  and degrees of freedom  $\nu_e$ . Inferences on the unknowns in the model are made from their marginal posterior distributions, using Markov chain Monte Carlo (MCMC) methods ([Meuwissen et al. 2001](#); [Habier et al. 2010](#)).

Although this model was first proposed for whole-genome prediction ([Meuwissen et al. 2001](#)), it can also be used to locate genomic segments that contain QTL ([Yi et al. 2003](#); [Sun et al. 2011](#); [Fan et al. 2011](#)). Consider a model where  $\pi$  is close to one, i.e., a model where most segments of the genome do not have markers that are associated with the trait. This is the model called BayesB by [Meuwissen et al. \(2001\)](#). Given such a model, the posterior probability that  $\alpha_j$  is nonzero for at least one SNP  $j$  in a window or segment can be used to make inferences on the presence of QTL in that segment. We will refer to this probability as the window posterior probability of association (WPPA). The underlying assumption here is that if a genomic window contains a QTL, one or more SNPs in that window will have nonzero  $\alpha_j$ . Thus, WPPA, which is estimated by counting the number of MCMC samples in which  $\alpha_j$  is nonzero for at least one SNP  $j$  in the window, can be used as a proxy for the posterior probability that the genomic window contains a QTL.

It is possible, however, that SNPs in a window that does not contain any QTL are in association with a QTL outside the window, which is what has been called signal dependence

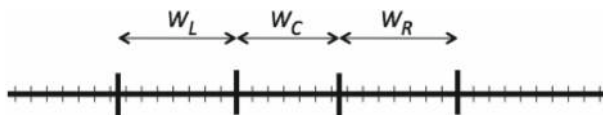


Figure 1. Illustration of composite genomic window  $W$  consisting of central window  $W_C$  and flanking windows  $W_L$  and  $W_R$ . To test the null hypothesis of no QTL in  $W$ , window PPA (WPPA) is computed by counting the number of MCMC samples in which  $\alpha_j$  is nonzero for at least one SNP in the central window  $W_C$ .

(Chen and Storey 2006). Depending on factors such as LD and marker density, the linkage signal from LD extends over shorter distances compared to that from cosegregation, and as described below, when all SNPs are fitted simultaneously, signal dependence is further reduced. Let  $W_C$  denote the window for which WPPA is estimated. Let  $W_L$  and  $W_R$  be windows of length  $k$  cM to the left and right of  $W_C$  as illustrated in Fig. 1.

A high-window WPPA for  $W_C$  is taken as evidence of a QTL in the “composite” window  $W$  comprised of  $W_L$ ,  $W_C$ , and  $W_R$ . Because WPPA for  $W_C$  is a partial association conditional on all other SNPs in the model, including those in the flanking windows  $W_L$  and  $W_R$ , the influence of QTL from outside the composite window on the WPPA signal for  $W_C$  will be inversely related to the length  $k$  of the flanking windows. In other words, as the number of markers between a QTL and  $W_C$  increases, the influence of the QTL on the WPPA signal for  $W_C$  will decrease. Thus, as shown in more detail in the next section, WPPA computed for  $W_C$  can be used to locate QTL.

### 2.3. RELATIONSHIP OF POSTERIOR PROBABILITIES TO PER

Here, we show that Bayesian posterior probabilities such as WPPA are related to PER, and therefore they can be used to control false positives in a multiple-test setting. The relationship of WPPA to PER is most straightforward when the priors used in the Bayesian analysis have a frequentist interpretation. This would be the case in a simulation study where data are generated and analyzed as follows. Suppose SNP genotypes are available on  $n$  individuals at  $K$  markers. Then, phenotypes for these  $n$  individuals can be generated according to model (6) by taking the matrix  $X$  to be a vector of ones and the vector  $\beta$  to be any constant (for example 0), sampling the partial regression coefficients,  $\alpha_j$ , according to (7), and sampling the residuals from a normal distribution with null mean and variance  $\sigma_e^2$ . Now, in the Bayesian analysis of the simulated data, suppose the distributions used to simulate the partial regression coefficients and the residuals are used as priors. Note that in this setting, the QTL are included in the marker panel, and thus, the QTL signal is expected to stay within its window. So,  $W_L$  and  $W_R$  can be set to have a length of zero.

Then, the WPPA calculated in the analysis for a genomic window,  $W_C$ , is the conditional probability of a true association (i.e., a QTL) within that window given the data. The frequentist interpretation of this probability is as follows. Suppose the simulation and analysis described above are repeated many times. Then, among all genomic windows with WPPA equal to  $q$ , a proportion  $q$  is expected to be truly associated with the trait, i.e., contain a QTL within that window.

Recall that PER is associated with the test of a hypothesis. The null hypothesis  $H_0$  in this case is that the genomic window  $W$  does not contain any SNPs associated with the trait. Using this notation, WPPA is the conditional probability that  $H_0$  is false given the observed data, while PER is the conditional probability that  $H_0$  is true given that  $H_0$  has been rejected based on some statistical test. Suppose the test is based on WPPA and  $H_0$  is rejected whenever WPPA is larger than some value  $t$ . Then, PER is the probability that  $H_0$  is true given WPPA is larger than  $t$ , and it can be written as:

$$\begin{aligned} \text{PER} &= \Pr(H_0 \text{ is true} | \text{WPPA} > t) \\ &= \frac{\int_t^1 (1 - q) f(q) dq}{\int_t^1 f(q) dq} \\ &= \text{E}[(1 - \text{WPPA}) | \text{WPPA} > t] \\ &\leq 1 - t, \end{aligned} \quad (8)$$

where  $f(q)$  is the density function of WPPA. Thus, in hypothetical repetitions of the analysis, for any interval with  $\text{WPPA} > t$  the proportion of false positives among significant results will be  $\leq (1 - t)$ .

One of the advantages of inference based on posterior probabilities such as WPPA is that derivation of the distribution of the test statistic is not required, but when sufficient data are not available, posterior probabilities may be sensitive to the priors used. Further, the posterior probabilities of interest are easily obtained from the MCMC samples from the Bayesian analysis. Consider, for example, the situation where interest is in detecting segments that explain more than some proportion  $v$  of the total genetic variance ( $\sigma_g^2$ ) (Hayes et al. 2010; Fan et al. 2011). The variance that is attributed to a genomic segment or window is defined as follows. First, the component of the genotypic value that corresponds to genomic window  $W_c$  is defined as

$$\mathbf{g}_w = \mathbf{Z}_w \boldsymbol{\alpha}_w, \quad (9)$$

where  $\mathbf{Z}_w$  is the matrix of genotype covariates and  $\boldsymbol{\alpha}_w$  is the vector of partial regression coefficients for SNPs in the window  $W_c$ . In other words,  $\mathbf{g}_w$  is the sum of the genotypic values for SNPs in window  $W_c$ . Then, the window variance is defined as the variance of a random element sampled from  $\mathbf{g}_w$ , and this variance is:

$$\sigma_{g_w}^2 = \frac{\sum_j g_{w_j}^2}{n} - \left( \frac{\sum_j g_{w_j}}{n} \right)^2. \quad (10)$$

In most Bayesian multiple-regression analyses, samples of the partial regression coefficients are drawn from their posterior distribution, and these samples can be used to obtain samples of window variances. Samples of the total genetic variance can also be similarly obtained from samples of  $\mathbf{g}$ . Then, the posterior probability that the window variance is larger than  $v\sigma_g^2$  can be estimated by counting the number of MCMC samples where the window variance is larger than  $v\sigma_g^2$ .

In this section, the relationship between WPPA and PER was discussed in the context of a Bayesian analysis of simulated data, where the distributions used in the generation of the



data were used as priors in the Bayesian analysis. In analysis on real data, the distribution of unobservable quantities such as the partial regression coefficients and residuals is not known. Thus, based on available knowledge of the problem, priors that lead to computationally efficient algorithms are used. Computer simulation will be used here to study the impact of such priors on the relationship between WPPA and PER in a realistic GWAS setting. As the amount of data that are combined with the priors increases, the impact of the priors on the posterior distributions is expected to decrease (Karaman et al. 2016).

#### 2.4. COMPUTER SIMULATION

The simulation described here was used to test if WPPA can be employed to control false positives in GWAS, where the tests are dependent. Actual SNP genotypes of purebred Angus bulls were used to simulate QTL and phenotypes as described in Kizilkaya et al. (2010). Exactly 100 data sets with 1000 observations and another 100 with 3570 observations were simulated, using genotypes at 52,910 SNP loci on 3570 purebred Angus bulls. The 1000 bulls were randomly sampled without replacement for inclusion in the data sets with 1000 observations, whereas all bulls were used in the data sets with 3570 observations.

In each of the 200 data sets, SNP effects of markers were sampled according to the prior of the BayesC model of (Habier et al. 2011) with  $\pi = 0.995$ , where a proportion  $\pi$  of the loci have null effects and the remaining loci have normally distributed effects with null mean and common variance  $\sigma_\alpha^2$  of SNP effects. The value of the common variance of SNP effects was calculated as  $\sigma_\alpha^2 = \frac{\sigma_g^2}{(1-\pi)K^2pq}$  (Fernando et al. 2007; Fernando and Garrick 2013), where  $\overline{2pq}$  is the average heterozygosity, and using  $\sigma_g^2 = 0.9$  for the additive genetic variance of the trait. The average number of QTL in the data sets was about 260. The residual variance for the trait was set at 0.1 to give a heritability of 0.9. This high value for the heritability was used to ensure a sufficient number of windows with high values of WPPA in order to quantify the relationship between WPPA and the frequency of true associations from 100 replications of the simulation. Further, although a heritability of 0.9 may be very high for most traits, it is not that high when estimated breeding values of sires are used as phenotypes, which is often the case in livestock, or when plot means are used as phenotypes, which is often the case in plants.

The data sets with 1000 observations were analyzed with and without including the SNPs that represent the QTL in the marker panel. The data sets with 3570 observations were only analyzed without including the QTL in the marker panel. Posterior inferences were based on 10,000 MCMC samples after a burn-in of 1000 samples. WPPA was also computed from 50,000 MCMC samples after a burn-in of 1000 samples for one of the data sets with 1000 observations. The correlation between WPPA values computed from 10,000 and 50,000 samples was 0.99.

In all analyses, the genome was divided into 2676 one Mb intervals according to the bovine map. The WPPA and window variance were computed for each such window,  $W_C$ , as explained previously. In order to study the relationship between WPPA and the true probability of a QTL, each genomic window,  $W_C$ , was classified into one of 10 WPPA classes of length 0.1 between 0 and 1. For example, all windows with WPPA between 0

and 0.1 were classified into the first class, and those with WPA between 0.1 and 0.2 to the second class. Note that the argument used in (8) implies that the posterior probability of a QTL in the window with WPPA between  $t_1$  and  $t_2$  will be at least  $t_1$ . The true probability of a QTL for a WPPA class was estimated by the observed relative frequency as:

$$f_j = \frac{\sum_i^{100} Q_{ij}}{\sum_i^{100} R_{ij}}, \quad (11)$$

where  $R_{ij}$  is the total number of genomic windows ( $W_C$ ) that belong to WPPA class  $j$  from data set  $i$ , and  $Q_{ij}$  is the subset of these windows that contain a QTL either in  $W_C$  or in windows  $W_L$  or  $W_R$  flanking  $W_C$ .

BayesC with  $\pi = 0.995$  or  $0.999$  was used for the analyses where the QTL were included in the marker panel. Recall that the prior of BayesC with  $\pi = 0.995$  was used in the simulation of SNP effects. Thus, with  $\pi = 0.995$ , WPPA is expected to agree well with the actual frequency of the QTL. The relationship between the lower bound of WPPA class interval and the actual frequency of QTL for these analyses are given for flanking windows of length  $k = 0, 1, \text{ or } 2$  Mb in Fig. 2.

Further, for these analyses with QTL included in the marker panel, each individual marker (not windows) was classified into ten posterior probability of association (PPA) classes as described above for WPPA. Figure 3 shows the relationship between the lower bound of the PPA class interval and the actual frequency of QTL among markers that belong to each class.

BayesC and BayesB with  $\pi = 0.995$ , and BayesC $\pi$ , where  $\pi$  is treated as an unknown, were used to analyze the data sets with 1000 observations without including the QTL in the marker panel. The relationship between the lower bound of WPPA class interval and the actual frequency of QTL for these analyses are given in Fig. 4.

BayesC $\pi$  was used to analyze the data sets with 3570 observations without including the QTL in the marker panel (Fig. 5).

In addition to computing WPPA for each genomic window,  $W_C$ , the posterior probability that the window variance ( $\sigma_C^2$ ) exceeds 1/1000 of the posterior mean of the total variance ( $\sigma_g^2$ ) was computed and grouped into 10 equally spaced posterior probability classes. For each such class, the actual frequency of a QTL in  $W_C$  or in the flanking windows having a variance ( $v_C$ ) that exceeds 1/1000 of the total QTL variance ( $V_g$ ) was also computed. The relationships between these probabilities and corresponding actual frequencies are given in Fig. 6 using BayesC $\pi$ .

## 2.5. AN APPLICATION TO REAL DATA

To demonstrate the use of the method described here, it was used to identify genomic windows that account for more than 0.1% of the genetic variance with a posterior probability of 90% or greater for weight of first three eggs laid in a White Leghorn line. Results are compared to those of a previous study of the same trait. Marker genotypes from 630,954 SNPs were available. Only markers that passed the following quality checks were used in the analysis. The quality checks were the proportion of missing genotypes  $< 0.05$ , minor allele

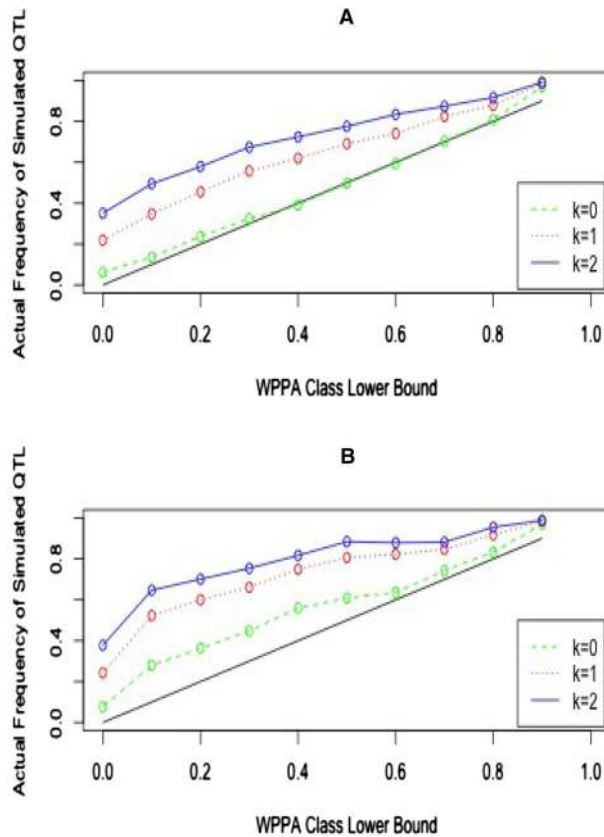


Figure 2. Relationship between window posterior probability of association (WPPA) and the actual frequency of simulated QTL in analyses where the QTL were included in the marker panel. WPPA was computed for each 1 Mb window ( $W_C$ ) of the genome and grouped into 10 WPPA classes. For each WPPA class, the actual frequency of simulated QTL in the composite window consisting of  $W_C$  and the flanking windows of  $k$  Mb ( $k = 0, 1, \text{ or } 2$ ) in length is given in the  $y$ -axis against the lower bound of the WPPA class interval on the  $x$ -axis. Results are for BayesC with  $\pi = 0.995$  (plot A) and BayesC with  $\pi = 0.999$  (plot B) from 100 data sets of 1000 observations.

frequency  $> 0.001$ , and detected parent-offspring mismatches  $< 5\%$ . Following quality checks, 171,941 SNP genotypes were used in the analysis with  $n = 4866$  observations. Of these observations, 2132 were individual observations with genotypes and phenotypes from the same individual; the rest were “family-mean” observations, where the genotype was the mean genotype of the parents and the phenotype was the mean phenotype of the offspring from this pair of parents. Residuals of family-mean observations were weighted by  $w_p = \frac{1-h^2}{1-0.5h^2/p}$  to account for the additional residual variance and differences in family size  $p$  (Garrick et al. 2009), where  $h^2$  is the heritability for the trait, which was estimated as 0.66 in a previous study (Wolc et al. 2012). The genome was divided into 951 non-overlapping windows of 1 Mb based on Build galGal4 ([http://uswest.ensembl.org/Gallus\\_gallus/Info/Index](http://uswest.ensembl.org/Gallus_gallus/Info/Index)) of the chicken genome. The variance that can be attributed to each such segment was computed using Eq. (10). The GenSel package (Garrick and Fernando 2013) was used for the analysis using the BayesB method with  $\pi = 0.99$ .

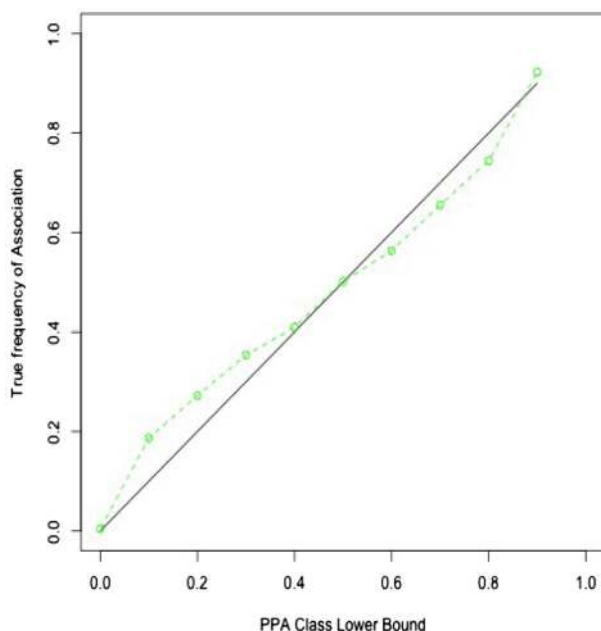


Figure 3. Relationship between posterior probability of association (PPA) of individual markers and the actual frequency of simulated QTL in analyses where the QTL were included in the marker panel. PPA was computed for each marker in the panel and grouped into 10 PPA classes. For each PPA class, the actual frequency of markers in that class being a QTL is given in the *y-axis* against the lower bound of the PPA class interval on the *x-axis*. Results are for BayesC with  $\pi = 0.995$  from 100 data sets each with 1000 observations.

Nine genomic windows were identified (Table 1) that each accounts for more than 0.1% of the genetic variance with a posterior probability of 90% or greater. In a previous study of the same trait in a purebred brown egg layer line, three genomic windows were identified on chromosomes 4, 5 and 8 (Wolc et al. 2012) that each accounts for more than 0.1% of the genetic variance with a posterior probability of 90% or greater. In that study, 1-Mb windows were defined according to Build galGal3. In the description of the results given here, the window positions have been translated to correspond to Build galGal4.

Two of the three windows in the previous study by Wolc et al. (2012) that explain >0.1% of the genetic variance were also identified in the present study with posterior probability  $\geq 0.9$ . The window that explains the largest proportion (>23%) of the genetic variance in the present analysis starts at position 75000211 on chromosome 4 and contains SNP rs14491030, which was the most significant SNP in the window explaining the largest proportion (>30%) of the genetic variance in Wolc et al. (2012). This region on chromosome 4 was identified as carrying a QTL in multiple other studies ([www.animalgenome.org](http://www.animalgenome.org)). Further, in the present study, six more windows that explain >0.1% of the genetic variance were identified, but these windows were not found to explain >0.1% of the variance with probability 0.9 in the brown line. These differences are not surprising as white and brown egg layers are two very distinct breeds (Moiseyeva et al. 2003).

The simulation of QTL genotypes and trait phenotypes, and the Bayesian analyses presented here were based on version 4.0 of the GenSel program (Fernando and Garrick 2013).

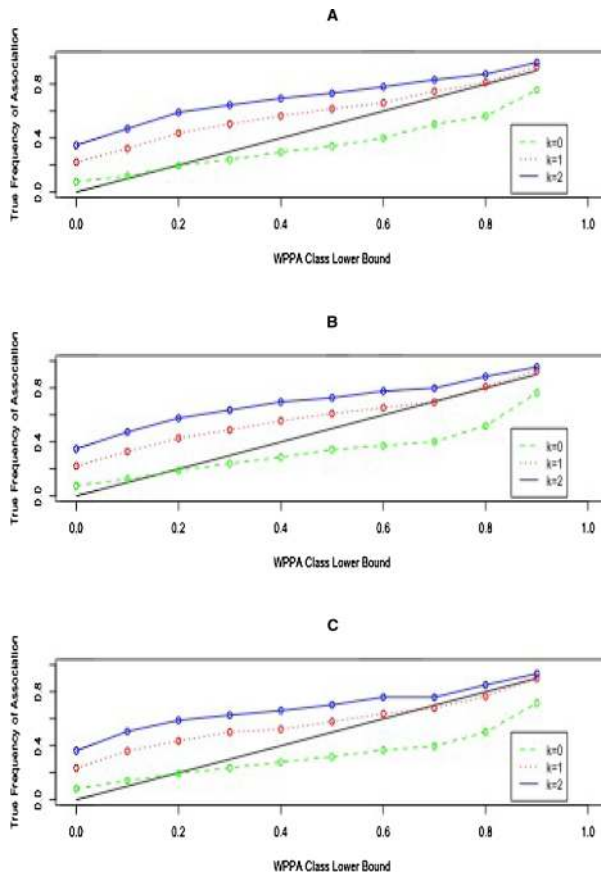


Figure 4. Relationship between window posterior probability of association (WPPA) and the actual frequency of simulated QTL. WPPA was computed for each 1 Mb window ( $W_C$ ) of the genome and grouped into 10 WPPA classes. For each WPPA class, the actual frequency of simulated QTL in the composite window consisting of  $W_C$  and the flanking windows of  $k$  Mb ( $k = 0, 1, \text{ or } 2$ ) in length is given in the  $y$ -axis against the lower bound of the WPPA class interval on the  $x$ -axis. Results are for BayesB with  $\pi = 0.995$  (plot A), BayesC with  $\pi = 0.995$  (plot B), and BayesC $\pi$  (plot C) from 100 data sets each with 1000 observations. The QTL were not included in the marker panel.

### 3. DISCUSSION

There are two main approaches to control false positives in genome-wide association studies. The most widely used approach is based on controlling the genomewide type I error rate (GER). The other is controlling the proportion of false positives among significant results, which, as we have shown here, is equivalent to controlling the posterior type I error rate (PER) (Morton 1955) for each test.

The longstanding practice of using a lod score of three for declaring linkage between a monogenic disease locus and a random marker is based on control of PER to about 0.05 (Elston 1997). The examples given in Sect. 2.1 show that the GER and PER values of a test can be very different. If only one marker is tested, controlling GER to about 0.05 will result in a much larger value for PER, i.e., among significant results a large proportion would

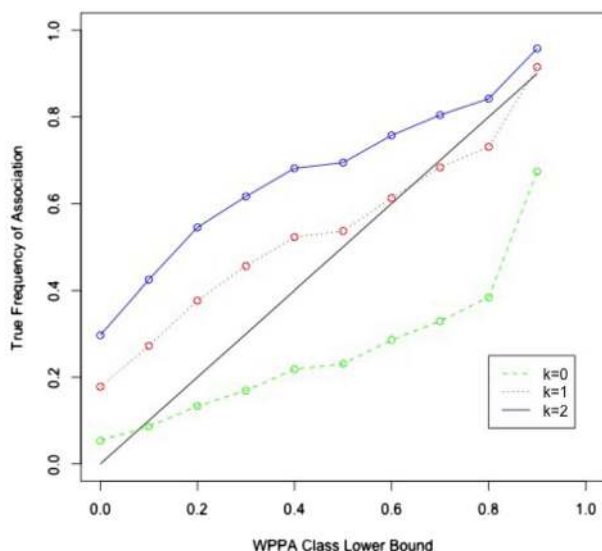


Figure 5. Relationship between window posterior probability of association (WPPA) and the actual frequency of simulated QTL. WPPA was computed for each 1 Mb window ( $W_C$ ) of the genome and grouped into 10 WPPA classes. For each WPPA class, the actual frequency of simulated QTL in the composite window consisting of  $W_C$  and the flanking windows of  $k$  Mb ( $k = 0, 1, \text{ or } 2$ ) in length is given in the  $y$ -axis against the lower bound of the WPPA class interval on the  $x$ -axis. Results are for BayesC $\pi$ , which treats  $\pi$  as unknown, from 100 data sets each with 3570 observations. The QTL were not included in the marker panel.

be false positives (94% in the example). Thus, to control PER to 0.05, a more stringent significance threshold has to be used (0.0001584 in the example). Now, suppose several markers are tested with the same stringent significance threshold each with the same prior probability of linkage to a QTL and power of test. Then, for each test, PER would be 0.05, even when the tests are not independent. Thus, provided the prior probability of linkage and power are constant, the same significance threshold can be used to control PER regardless of the number of tests. Thus, when PER is used for inference there is no multiple-test penalty (Stephens and Balding 2009). This is not true for GER, which for a given significance threshold increases with the number of tests.

To control PER, however, it is required to know the value of  $\pi$ , which is the probability that the null hypothesis is true for a test, and the power of the test. Mosig et al. (2001) have shown that these quantities can be estimated from data, and these estimates can be used to estimate PER. On the other hand, in Bayesian analyses, these quantities can be treated as unknowns and an upper bound for PER can be obtained from Bayesian posterior probabilities, as described in Sect. 2.3. For example, the PER for the test of a QTL in a genomic window,  $W$ , is obtained as  $1 - \text{WPPA}$ , where WPPA is estimated by counting the number of MCMC samples in which  $\alpha_j$  is nonzero for any SNP  $j$  in  $W_C$ , the central window of  $W$  (Fig. 1).

We have argued here that by using multiple-regression models the linkage signal can be localized to relatively short segments of the genome and thereby avoid the problem of signal dependence raised by Chen and Storey (2006). In model (1), assuming no interference, segregation of alleles at locus  $C$  is conditionally independent of the segregation of alleles at

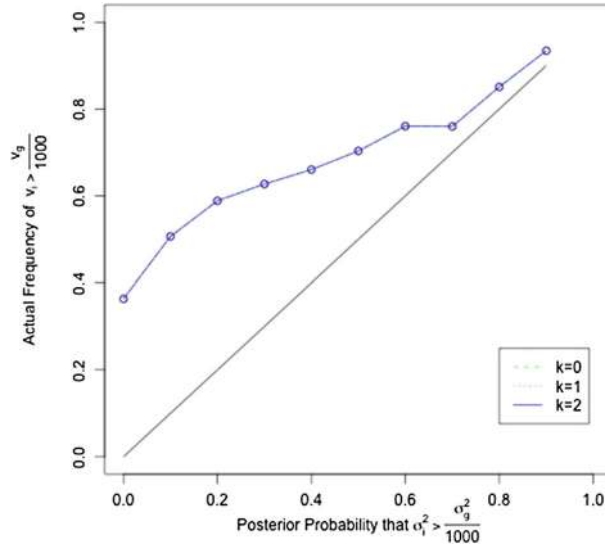


Figure 6. Relationship between posterior probability that variance of the central window  $W_C$  exceeds 1/1000 of total variance and corresponding actual frequency of the simulated QTL variance for the composite window  $W$ . The posterior probability that the window variance ( $\sigma_C^2$ ) exceeded 1/1000 of the posterior mean of the total variance ( $\sigma_g^2$ ) was computed for each 1 Mb window ( $W_C$ ) of the genome and grouped into 10 posterior probability classes. For each such class, the actual frequency of simulated QTL in the composite window consisting of  $W_C$  and the flanking windows of  $k$  Mb ( $k = 0, 1, \text{ or } 2$ ) in length having a variance ( $v_C$ ) that exceeds 1/1000 of the total QTL variance ( $V_g$ ) is given in the  $y$ -axis against the lower bound of the posterior probability class interval on the  $x$ -axis. Results are for BayesC $\pi$ , which treats  $\pi$  as unknown, from 100 data sets each with 3570 observations. The QTL were not included in the marker panel.

Table 1. 1-Mb windows that explain >0.1% of genetic variance with posterior probability  $\geq 0.9$  for weight of first three eggs laid from a White Leghorn line.

Chromosome	1 Mb window number	Number of SNPs	% genetic variance	$P > 0.1$
4	76	146	23.5	1.0
5	13	166	1.8	1.0
6	11	209	3.0	1.0
8	22	211	0.8	0.92
9	15	348	1.0	1.0
14	13	251	1.0	0.98
15	2	410	1.0	0.91
17	2	237	0.6	0.95
23	2	363	1.6	1.0

Each chromosome was divided into one mega base non-overlapping windows, containing a variable number of SNPs. Results were obtained using the GenSel package with the BayesB method for  $\pi = 0.99$ .

loci to the left of locus L and to the right of locus R given the segregation events at L and R. Thus, this model can be used to test the null hypothesis of no QTL between loci L and R in a linkage analysis, where the signal for QTL detection comes from the cosegregation of alleles at markers and QTL. In an association analysis, the signal for QTL detection comes from LD between markers and QTL, which is the non-independence between the allele states at

these loci. Unfortunately, even under the assumption of no interference, allele states at locus C may not be conditionally independent of allele states at loci to the left of locus L and to the right of locus R given the allele states at loci L and R. However, the LD signal, depending on the effective population size, decays much faster with distance between loci than that due to cosegregation. Further, in the Bayesian multiple-regression models that we have used here, all SNPs are used in the analysis. Thus, in association analyses signal dependence is less of a problem than in linkage analysis, and when multiple-regression models are used with high-density SNP markers, the signal from a QTL is expected to be almost completely explained by the markers within a narrow genomic window containing it. If the QTL is at the edge of a window, however, its signal may bleed into the next window. Thus, we considered using a composite window (Fig. 1). The signal is measured only in the central window to test the null hypothesis of no QTL in the composite window.

In this study, posterior probabilities of association were computed for genomic windows of 1 Mb (WPPA). When the QTL were included in the marker panel used in the analysis and the distribution used to simulate the QTL effects was used as the prior for marker effects in the Bayesian analysis, Fig. 2 (plot A) shows good agreement between WPPA for a 1 Mb window and the frequency of QTL in that window ( $k = 0$ ). For example, among genomic windows with WPPA between 0 and 0.1, about 5% contained one or more QTL, and among windows with WPPA between 0.9 and 1.0, about 95% contained QTL.

Figure 2a also shows the frequency of QTL in composite windows consisting of a 1 Mb central window and left and right flanking windows of length  $k = 0, 1, \text{ or } 2$  Mb. In all three curves, WPPA was computed for the central window, but the actual QTL frequency was computed for either a 1, 3, or 5 Mb window centered at the window for which WPPA was computed. Windows with WPPA between 0 and 0.1 had actual QTL frequencies of 0.06, 0.22, or 0.35 for windows of length 1, 3 or 5 Mb. Given that about 260 QTL were simulated and uniformly placed in 2676 genomic windows, the prior probability that a composite window would contain one or more QTL is 0.09, 0.25 or 0.38 for  $k = 0, 1 \text{ or } 2$ . Thus, the actual frequencies were reduced from the prior values toward a WPPA of 0.05, which is the midpoint of the interval. This is most evident when  $k = 0$ , where the actual frequency was 0.06, indicating that in this case, the WPPA is mainly influenced by the presence or absence of QTL in the central window. In windows where WPPA was between 0.40 and 0.70, the actual frequency of QTL for  $k = 0$  was slightly lower than the lower bound for the class. This indicates that the WPPA is slightly inflated by the presence of QTL outside  $W_C$  whose signal bleeds into the signal observed for  $W_C$ .

Figure 2 (plot B) shows the same relationships when the Bayesian analysis used  $\pi = 0.999$  although  $\pi = 0.995$  was used in the simulation (QTL were included in the marker panel). The high value for  $\pi$  makes it difficult for a locus to have a nonzero effect, and this can explain why the actual QTL frequencies for  $k = 0$  were higher than the upper bound for the WPPA class when WPPA was  $< 0.60$ . This was not the case for when WPPA was larger than 0.60, indicating that when the QTL had a large effect, the information from the phenotype was able to overwhelm the incorrect information from prior.

Figure 3 shows the relationships between PPA for each marker and the actual frequency that the corresponding marker is a QTL (QTL were included in the marker panel). Here,



the actual frequencies of QTL were even lower than in plot A of Fig. 2 for WPPA between 0.50 and 0.90.

In plot A of Fig. 2, the overestimation of the QTL frequency for  $k = 0$  was thought to be due to the presence of QTL outside  $W_C$  whose signal bleeds into the signal observed for  $W_C$ . In Fig. 3, this can happen with markers in the window that contains the QTL. Thus, the observation that QTL frequencies are lower in Fig. 3 than in plot A of Fig. 2 is consistent with the expectation that markers within the QTL window would have higher LD with the QTL than markers from adjacent windows. This, however, indicates a mixing problem given the 10,000 MCMC samples used for inference in the simulation. If that is the case, use of a longer chain would give better results. It also indicates that alternative sampling strategies to improve mixing should be investigated.

In the remaining analyses, the QTL were not included in the marker panel. Figure 4 presents results from three analyses of the 100 data sets with 1000 observations, and Figs. 5 and 6 give results for the 100 data sets with 3570 observations. In these analyses that did not have the QTL included in the marker panels, in genomic windows of 1 Mb and  $k = 0$  WPPA for  $W_C$  substantially overestimated the frequency of QTL in the window  $W_C$  when WPPA was greater than about 0.2. For example, in plot B of Fig. 4, which shows the relationship between WPPA and the frequency of QTL for BayesC with  $\pi = 0.995$ , in genomic windows of 1 Mb with WPPA between 0.9 and 1.0 the frequency of QTL was about 0.72 and in genomic windows of 1 Mb with WPPA between 0.8 and 0.9 the frequency of QTL was only about 0.5. When the QTL were included in the analysis, the comparable QTL frequencies were 0.97 and 0.81 (plot A of Fig. 2). Thus, when the QTL were not in the panel, WPPA overestimated the frequency of QTL in  $W_C$ . Following are two possible reasons for this. The first is that the prior used for marker effects does not agree with the actual distribution of effects. When the QTL are not included in the marker panel, only markers that are in complete LD with the QTL will have effects that are distributed as the QTL. In Angus, the average LD between adjacent markers for the 50k SNP panel is about 0.2 (Goddard and Hayes 2009). Thus, the distribution of marker effects may be quite different from that of the QTL and this may have an impact on the relationship between WPPA for a genomic interval and the frequency of QTL in that interval even when the distribution used to generate the QTL effects is used as the prior for marker effects as in the BayesC analysis with  $\pi = 0.995$ . The second reason is violation of the assumption that WPPA is equivalent to the posterior probability that  $W_C$  contains a QTL (WPPQ). Recall that WPPA is the posterior probability that a marker in window  $W_C$  has a nonzero regression coefficient. When the QTL are included in the panel, WPPA is also the posterior probability of a QTL in  $W_C$  because QTL by definition have nonzero effects on the trait. However, when the QTL are not included in the panel, WPPA is not equivalent to probability of a QTL in  $W_C$ . A marker in  $W_C$  may have a nonzero effect even when  $W_C$  does not contain any QTL due to it being in LD with a QTL in an adjacent window. Thus, some intervals without QTL may have high values of WPPA, which is consistent with our results where the frequency of QTL was often lower than WPPA when the panel did not contain QTL.

It can be argued that both of the reasons given above played a role in the observed overestimation of QTL frequencies in  $W_C$  by WPPA. Violation of the assumption that WPPA is equivalent to WPPQ, however, seems to have played a greater role. The three plots in

Fig. 4 were obtained using three different priors. Plot A is from a BayesB analysis with  $\pi = 0.995$ , where a central  $t$  distribution with four degrees of freedom was used as the prior for marker effects. Plot B is from BayesC with  $\pi = 0.995$ , where a normal distribution is used for marker effects, and plot C is from BayesC $\pi$ , where  $\pi$  is treated as unknown with a uniform prior between 0 and 1 and a normal prior for marker effects. The results from these three analyses being very similar indicates that with 1000 observations these differences in priors had a negligible effect on the relationships between WPPA and QTL frequencies. Further, if the overestimation of QTL frequencies by WPPA was due to the prior for marker effects not being appropriate, then better results would be expected in the data sets with 3570 observations. However, this was not the case. Overestimation was even greater with the bigger data sets (Fig. 5). On the other hand, if the observed overestimation of QTL frequencies was due to markers in  $W_C$  being in LD with QTL in adjacent windows, it is possible that with more data associations with even more distant QTL could further inflate WPPA. Comparison of QTL frequencies in plot C of Fig. 4 with those in Fig. 5 for genomic windows with WPPA between 0.8 and 0.9 and  $k = 0, 1, \text{ and } 2$  suggests that with the bigger data sets more distant QTL contributed to the WPPA value calculated for  $W_C$ .

In these analyses that did not include the QTL in the marker panel, there was good agreement between WPPA and the actual frequency of the QTL in the composite window  $W$  with  $k = 2$  when WPPA was larger than 0.8. At lower values of WPPA, WPPA underestimated the QTL frequency for  $k = 2$ . In genomic windows with WPPA between 0 and 0.1, the QTL frequency with  $k = 0$  was almost 0.05, agreeing very well with WPPA. This is because the QTL in these windows have only very small effects. Thus, only the QTL from  $W_C$  contribute to WPPA. As mentioned previously, with  $k = 2$ , the prior probability of a QTL in  $W$  is 0.38. The observed frequency of QTL in  $W$  with WPPA between 0 and 0.1 was 0.3 for  $k = 2$ , which is lower than the prior value of 0.38. Genomic windows with higher values of WPPA, for example, between 0.2 and 0.3, consist of a mixture of windows containing QTL of moderate size in the flanking windows which affect WPPA computed for  $W_C$  and smaller QTL that do not affect WPPA computed in the central window  $W_C$ . As WPPA gets higher, most windows contain large QTL that contribute to the high value of WPPA.

Figure 6 shows the relationship between the posterior probability that the window variance (PPWV) exceeds 1/1000 of the total variance and the corresponding actual frequency for the QTL variance. PPWV are especially useful for GWAS using models such as BayesA (Meuwissen et al. 2001) and Bayesian Lasso (de los Campos et al. 2009), where all markers are assumed to have non-null effects and, thus, WPPA is always 1.

Here, we will use window variances to examine the signal from the flanking windows. In the BayesC $\pi$  analyses (Figs. 4, 5), the actual frequencies of QTL were in good agreement with WPPA for WPPA values larger than 0.85 and  $k = 2$ . However, when  $k = 0$ , the actual frequencies were much lower than WPPA. This indicates that the high values of WPPA are due partly to strong signals from the flanking windows. This can be tested by examining the QTL signal in the central and flanking windows for segments with PPWV = 0.95 in comparison with the corresponding signal in segments with PPWV = 0.05. In segments with PPWV = 0.95, the QTL in the central window ( $k = 0$ ) had a mean that was 1.1% of the total variance, and those in the flanking windows ( $k = 2$ ) had a mean that was 1.3% of the total variance. In segments with PPWV = 0.05, the QTL in the central window had a

mean that was 0.1% of the total variance, and those in the flanking windows had a mean that was 0.3% of the total variance. Thus, in segments with  $PPWV = 0.95$ , there was a QTL with a strong signal in the central window or one with even a stronger signal in the flanking windows.

These simulation results show that there is good agreement between posterior probabilities and the actual frequencies for the corresponding events when the priors used for the analysis represent the actual distribution of the marker effects. When the QTL are not included in the marker panel, the distribution of marker effects is not known even for simulated data. In analysis of real data, this will be even more of a problem. Even when the “correct” prior was not used in the analysis, there was good agreement between the posterior probabilities and the actual frequencies for low values of the posterior probabilities and  $k = 0$  and at high values of the posterior probabilities and  $k = 2$ . The width of the genomic interval that gives good agreement between the posterior probabilities and the actual frequencies may depend on the distribution of the QTL effects, the LD structure between the markers and the QTL, and the amount of data. However, based on these simulation results it is expected that a genomic window with a high value for WPPA or PPWV would have equally high frequencies of large QTL in  $W_C$  or close to it. These results are based on a simulated a trait with a heritability of 0.9. In a subsequent study, a heritability of 0.5 was used, and results from that study also showed good agreement between WPPA and true frequency of QTL (Zeng 2015).

We have argued here that the primary reason WPPA over estimates the true frequency of QTL in  $W_C$  (WPPQ) when the QTL are not included on the marker panel is because, in this situation, the assumption that WPPA is equivalent to WPPQ is violated. In a subsequent study that had a marker density that was over ten times higher than in this study, the agreement between WPPA and WPPQ for genomic windows of one megabase was much better (Zeng 2015). This could be due to the larger number of markers in a window capturing most of the variability of the QTL in that window and thus making WPPA equivalent to WPPQ. Also, preliminary results indicate that when the marker density is low, using a single five-megabase window gives good agreement between WPPA and WPPQ (<https://www.slideshare.net/RohanFernando14/discovery-of-qtll-using-a-qtll-effects-model>). Thus, when the marker density is low, larger windows seem to be needed to make WPPA equivalent to WPPQ. Further study is needed to verify this hypothesis and establish guidelines for appropriate window sizes that would depend on marker density and effective population size.

Models that fit haplotype effects may give results similar to those obtained here through the use of genomic windows. However, fitting haplotype effects requires the additional step of determining haplotypes, which can introduce another source of error, especially when pedigree information is not available.

Application of this approach to a egg-weight trait in a White Leghorn line resulted in the identification of nine genomic windows that each accounted for more than 0.1% of the genetic variance with a posterior probability of 0.9 or greater. In a previous study of the same trait in brown egg layers (Wolc et al. 2012), three genomic windows with  $WPPA \geq 0.9$  were identified. Two of these windows were also found in the present study. This demonstrates that the method can detect previously identified regions with large effects.

In summary, we have argued here that PER is more suitable for controlling false positives in GWAS than GER. Controlling PER at individual tests results in control of false positives for the collection of tests that may be dependent. Further, we have shown the relationship between PER and WPPA under the ideal situation where the “correct” prior is known. Computer simulation was used to examine the impact of not knowing the “correct” prior. If a high value of WPPA or PPWV is used to detect QTL, among the positive results that accumulate over many experiments, the proportion of false positives can be expected to be low. Further, use of multiple regression models allows inference on the presence or absence of QTL to be specific to relatively narrow segments of the genome. Thus, the problem of signal dependence (Chen and Storey 2006) is to a large degree avoided. Further research is needed to determine the optimum size of the central and flanking window sizes, which may depend on the nature of LD, the number of observations, heritability of the trait and possibly other factors such as the number of QTL which may be unknown.

## ACKNOWLEDGEMENTS

Authors are grateful to Dr. Soller for his instruction and research that has been a source of inspiration for much of our work. We are also grateful to Dr. Dan Nettleton and two anonymous reviewers who provided very useful suggestions. This work was supported in part by the US Department of Agriculture, Agriculture and Food Research Initiative National Institute of Food and Agriculture Competitive Grant No. 2015-67015-22947.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

[Received July 2016. Accepted March 2017. Published Online April 2017.]

## REFERENCES

- Benjamini, Y., and Hochberg, Y. (1995), “Controlling the false discovery rate: a practical and powerful approach to multiple testing.” *J. R. Statist. Soc. B*, 57, 289–300.
- Chen, L., and Storey, J. D. (2006), “Relaxed significance criteria for linkage analysis,” *Genetics*, 173(4), 2371–2381.
- de los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., Weigel, K., and Cotes, J. M. (2009), “Predicting quantitative traits with regression models for dense molecular markers and pedigree,” *Genetics*, 182(1), 375–385. <http://www.hubmed.org/display.cgi?uids=19293140>
- Elston, R. C. (1997), “1996 William Allan Award Address: Algorithms and inferences: The challenges of multifactorial diseases,” *American Journal of Human Genetics*, 60, 225–262.
- Fan, B., Onteru, S. K., Du, Z.-Q., Garrick, D. J., Stalder, K. J., and Rothschild, M. F. (2011), “Genome-Wide Association Study Identifies Loci for Body Composition and Structural Soundness Traits in Pigs,” *PLoS ONE*, 6(2), e14726.
- Fernando, R. L., and Garrick, D. (2013), “Bayesian methods applied to GWAS.” *Methods in molecular biology (Clifton, N.J.)*, 1019, 237–274.
- Fernando, R. L., Habier, D., Stricker, C., Dekkers, J. C. M., and Totir, L. R. (2007), “Genomic selection,” *Acta Agriculturae Scandinavica, Section A - Animal Science*, 57(4), 192–195.
- Fernando, R. L., Nettleton, D., Southey, B., Dekkers, J., Rothschild, M., and Soller, M. (2004), “Controlling the proportion of false positives in multiple dependent tests,” *Genetics*, 166(611-619).

- Garrick, D. J., and Fernando, R. L. (2013), "Implementing a QTL detection study (GWAS) using genomic prediction methodology," *Methods in molecular biology (Clifton, N.J.)*, .
- Garrick, D. J., Taylor, J. F., and Fernando, R. L. (2009), "Deregressing estimated breeding values and weighting information for genomic regression analyses," *Genet Sel Evol*, 41(1), 55–55. <http://www.hubmed.org/display.cgi?uids=20043827>
- Goddard, M. E., and Hayes, B. J. (2009), "Mapping genes for complex traits in domestic animals and their use in breeding programmes," *Nat Rev Genet*, 10(6), 381–391. doi:10.1038/nrg2575
- Habier, D., Fernando, R., Kizilkaya, K., and Garrick, D. J. (2010), Extension of the Bayesian alphabet for genomic selection., in *Proc. 9th World Congress on Genet. Appl. Livest. Prod.*, Vol. 9, p. 468.
- Habier, D., Fernando, R. L., Kizilkaya, K., and Garrick, D. (2011), "Extension of the bayesian alphabet for genomic selection," *BMC Bioinformatics*, 12, 186.
- Habier, D., Tetens, J., Seefried, F.-R., Lichtner, P., and Thaller, G. (2010), "The impact of genetic relationship information on genomic breeding values in German Holstein cattle," *Genetics Selection Evolution*, 42(1), 5.
- Hayes, B., Bowman, P., Chamberlain, A., Verbyla, K., and Goddard, M. (2009), "Accuracy of genomic breeding values in multi-breed dairy cattle populations," *Genetics Selection Evolution*, 41(1), 51.
- Hayes, B. J., Pryce, J., Chamberlain, A. J., Bowman, P. J., and Goddard, M. E. (2010), "Genetic Architecture of Complex Traits and Accuracy of Genomic Prediction: Coat Colour, Milk-Fat Percentage, and Type in Holstein Cattle as Contrasting Model Traits," *PLoS Genet*, 6(9), e1001139.
- Karaman, E., Cheng, H., Firat, M. Z., Garrick, D. J., and Fernando, R. L. (2016), "An Upper Bound for Accuracy of Prediction Using GBLUP," *PLoS ONE*, 11(8), e0161054–18.
- Kizilkaya, K., Fernando, R. L., and Garrick, D. J. (2010), "Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes," *J Anim Sci*, 88(2), 544–551.
- Maher, B. (2008), "The case of the missing heritability.," *Nature*, 456, 18–21.
- Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., McCarthy, M. I., Ramos, E. M., Cardon, L. R., Chakravarti, A., Cho, J. H., Guttmacher, A. E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C. N., Slatkin, M., Valle, D., Whittemore, A. S., Boehnke, M., Clark, A. G., Eichler, E. E., Gibson, G., Haines, J. L., Mackay, T. F., McCarroll, S. A., and Visscher, P. M. (2009), "Finding the missing heritability of complex diseases," *Nature*, 461(7265), 747–753.
- Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001), "Prediction of total genetic value using genome-wide dense marker maps," *Genetics*, 157, 1819–1829.
- Moiseyeva, I. G., Romanov, M. N., Nikiforov, A. A., Sevastyanova, A. A., and Semyenova, S. K. (2003), "Evolutionary relationships of Red Jungle Fowl and chicken breeds," *Genetics Selection Evolution*, 35(5), 403–423. doi:10.1186/1297-9686-35-5-403
- Morton, N. (1955), "Sequential tests for the detection of linkage," *American Journal of Human Genetics*, 7, 277–318.
- Mosig, M., Lipkin, E., Khutoreskaya, G., Tchourzyna, E., Solter, M., and Friedmann, A. (2001), "A whole genome scan for QTL affecting milk protein percent in Israel-Holstein cattle by means of selective milk pooling in a daughter design, using an adjusted false discovery rate criterion," *Genetics*, 157, 1683–1698.
- Sahana, G., Guldbbrandtsen, B., Janss, L., and Lund, M. S. (2010), "Comparison of association mapping methods in a complex pedigreed population," *Genetic Epidemiology*, 34, 455–462.
- Sidak, Z. (1967), "Rectangular confidence regions for the means of multivariate normal distributions.," *J. Am. Stat. Assoc.*, 62, 626–633.
- Southey, B. R., and Fernando, R. L. (1998), Controlling the proportion of false positives among significant results in QTL detection., in *Proc. 6th Wld. Cong. Genet. App. Liv. Prod.*, Vol. 26, Armidale, Australia, pp. 221–224.
- Stephens, M., and Balding, D. J. (2009), "Bayesian statistical methods for genetic association studies," *Nat Rev Genet*, 10(10), 681–690.
- Storey, J. D. (2002), "A direct approach to false discovery rates," *Journal of the Royal Statistical Society, Series B*, 64, 479–498.

- Sun, X., D. H., R.L. F., Garrick, D., and J.C.M., D. (2011), "Genomic breeding value prediction and QTL mapping of QTLMAS-2010 data using Bayesian methods.," *BMC proceedings*, 5(Suppl 3), S13.
- VanRaden, P. M., Van Tassell, C. P., Wiggans, G. R., Sonstegard, T. S., Schnabel, R. D., Taylor, J. F., and Schenkel, F. S. (2009), "Invited review: reliability of genomic predictions for North American Holstein bulls," *J Dairy Sci*, 92(1), 16–24.
- Visscher, P. M., Yang, J., and Goddard, M. E. (2010), "A commentary on 'common SNPs explain a large proportion of the heritability for human height' by Yang et al. (2010)," *Twin Res Hum Genet*, 13(6), 517–524. <http://www.ncbi.nlm.nih.gov/pubmed/21142928>
- Wolc, A., Arango, J., Settar, P., Fulton, J., Sullivan, N. P., Preisinger, R., Habier, D., Fernando, R., Garrick, D. J., Hill, W. G., and Dekkers, J. C. M. (2012), "Genome-wide association analysis and genetic architecture of egg weight and egg uniformity in layer chickens.," *Animal Genetics*, 43 (Suppl 1), 87–96.
- Wolc, A., Stricker, C., Arango, J., Settar, P., Fulton, J., O'Sullivan, N., Preisinger, R., Habier, D., Fernando, R., Garrick, D., Lamont, S., and Dekkers, J. (2011), "Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model," *Genetics Selection Evolution*, 43(1), 5.
- Yi, N., George, V., and Allison, D. B. (2003), "Stochastic Search Variable Selection for Identifying Multiple Quantitative Trait Loci," *Genetics*, 164(3), 1129–1138.
- Zeng, J. (2015), PhD thesis, Whole genome analyses accounting for structures in genotype data.
- Zeng, Z.-B. (1993), "Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci," *Proc. Natl. Acad. Sci. USA*, 90, 10972–10976.
- Zeng Z-B (1994), "Precision mapping of quantitative trait loci," *Genetics*, 136, 1457–1468.