Journal of the Royal Statistical Society
SERIES C Applied Statistics
Appl. Statist. (2014)
63, *Part* 3, *pp.* 483–498

A generalized quasi-likelihood scoring approach for simultaneously testing the genetic association of multiple traits

Zeny Feng

University of Guelph, Canada

[Received December 2011. Final revision June 2013]

Summary. In the genetic association analysis of Holstein cattle data, researchers are interested in testing the association between a genetic marker with more than one estimated breeding value phenotype. It is well known that testing each trait individually may lead to problems of controlling the overall type I error rate and simultaneous testing of the association between a marker and multiple traits is desired. The analysis of Holstein cattle data has additional challenges due to complicated relationships between subjects. Furthermore, phenotypic data in many other genetic studies can be quantitative, binary, ordinal, count data or a combination of different types of data. Motivated by these problems, we propose a novel statistical method that allows simultaneous testing of multiple phenotypes and the flexibility to accommodate data from a broad range of study designs. The empirical results indicate that this new method effectively controls the overall type I error rate at the desired level; it is also generally more powerful than testing each trait individually at a given overall type I error rate. The method is applied to the analysis of Holstein cattle data as well as to data from the Collaborative Study on the Genetics of Alcoholism to demonstrate the flexibility of the approach with different phenotypic data types.

Keywords: Genetic association study; Multiple testing; Quasi-likelihood

1. Introduction

In genetic studies of complex traits, data for multiple phenotypic traits are typically collected for each individual. For example, the Holstein cattle data set that we shall analyse later includes more than 20 estimated breeding value (EBV) phenotypic traits of proven bulls—those that have shown to have high economic value in producing progenies and/or producing progenies that have high profile in economical important traits. The EBVs of a bull predict its genetic merit on economic traits. For example, the milk yield, protein yield, fat yield, protein percentage and fat percentage EBVs of a bull predict the quality and quantity of milk products of its female descendants. Animal breeding researchers are interested in finding single-nucleotide polymorphisms (SNPs) or genome regions that are associated with any of these EBV traits. The use of genomic information in genetic evaluation of animals enhances the animal selection process for breeding. Testing each trait individually is subject to problems of multiple testing. Procedures such as familywise error rate control or Bonferroni control will generally lead to a stringent and thus less powerful test. This issue becomes more problematic as the number of phenotypes to be tested increases.

Address for Correspondence: Zeny Feng, Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada.
E-mail: zfeng@uoguelph.ca

 0035–9254/14/63483

In genetics, the effect of a single variant influencing more than one trait, or pleiotropy, was first introduced by Ludwig Plate in 1910. In complex traits, pleiotropic effects are considered to play indispensable roles. In general, numerous genes are involved in a complex trait. With a massive number of complex traits, the majority of genes play roles in more than one trait. Many examples of pleiotropy have been published. For example, about 40% of cats with white fur and blue eyes are found to be deaf owing to a single gene that is responsible for both the pigmentation and hearing (Hartl and Jones, 2005). Following this line, pleiotropic genetic effects influencing pigment phenotypes and hearing are also found in other animals such as chickens and mice (Sunquist, 2007). In human diseases, phenotypic traits such as insulin secretion and insulin action are reported to share common genetic variants (Mitchell *et al.*, 1996). The gene TP53 is found to play a prominent role in longevity and cancer suppression (Rodier *et al.*, 2007). Multiple research groups have confirmed pleiotropic genetic effects contributing to hypertension-related traits (Mahaney *et al.*, 1995; Edwards *et al.*, 1999; Kullo *et al.*, 2005). With respect to mental illness, several genes have been found to influence both schizophrenia and bipolar disorder (Lin and Mitchell, 2008), with more still to be determined. Thus, research to discover genes with pleiotropic effects is an important direction in genetics, and statistical methods for identifying genetic pleiotropic effects by simultaneously testing for multiple traits become appealing.

In many studies, phenotypic traits can be of different types of data (e.g. quantitative, binary, ordinal, count data or a combination thereof). For example, in the Collaborative Study on the Genetics of Alcoholism data, the phenotypic data comprise different types of data, including

(a) ordinal traits, such as the affection status (ALADX1), which is determined on the basis of both the DSM-III-R criteria of the American Psychiatric Association (1987) and the criteria for definite alcoholism specified by Feighner *et al.* (1972),
(b) quantitative traits, such as the maximum number of drinks in a 24-h period and
(c) binary traits, such as if an individual has physical health problems from drinking or not.

In this case, a method that can accommodate traits with different types of data is attractive.

Recently, problems with respect to testing the genetic association of multiple traits have received much attention. Several methods have been proposed for the simultaneous testing of multiple traits. Lange *et al.* (2003) extended the family-based association test (FBAT) to multivariate traits using generalized estimating equations. The FBAT–generalized estimating equation approach does not require any distributional assumption for the phenotypes and hence can be applied to association studies including different phenotypic data types. The FBAT is also well known to be robust to population stratifications (Horvath *et al.*, 2001; Laird *et al.*, 2000). However, as an extension of the FBAT, the FBAT–generalized estimating equation approach has inherited some limitations. The FBAT statistic depends on the known allele transmission pattern from parents to offspring, and subjects with unknown parental genotypes or homozygous parents are not useful in the analysis. This problem is quite common in the study of late onset disease or when the SNP being tested has a small minor allele frequency. In addition, for a binary trait, such as the affection status of a disease, the population prevalence of a disease must be known to make use of unaffected offspring in the test statistic. Thus, accurate knowledge of population prevalence is crucial for the analysis. The information that is provided by unaffected offspring also becomes extraneous in the case of a rare disease. Moreover, the FBAT ignores relationships across related families because it typically breaks down a large pedigree into small nuclear families and treats them independently. For these reasons, the FBAT approach is generally less powerful (Bourgain *et al.*, 2003; Feng *et al.*, 2011; Risch and Teng, 1998; Thornton

and McPeek, 2007). Zhang *et al.* (2011) proposed a method based on the generalized Kendall τ, which allows testing for association with a set of traits of different types. The method is based on dissimilarity measures of the traits of individuals and their parental genotypes of the marker. Empirical results show that the simultaneous testing of multiple traits is more powerful than testing each trait individually. However, this method is limited to family-based study designs. Parental information is crucial as, otherwise, the conditional expectation of parental genotype given the offspring's observed genotype will be used. In summary, these methods are useful for data from family-based study designs.

Generally, a flexible method that can accommodate data from different study designs is desired. This paper is motivated by the challenges of analysing data from Holstein cattle. Performing genetic association studies on these data is difficult owing to the complicated relationship between subjects in the data. First, dams are not genotyped, and sires are genotyped only if they appear as proven bulls in the data set. Second, most of the bulls in the data are inbred. For these reasons, both the FBAT and the generalized Kendall's τ-method are unsuitable for the analysis of this data set. The aim of this paper is to propose a method to test simultaneously genetic association with multiple traits and to accommodate data from a broad range of study designs, including conventional population-based studies with unrelated subjects, founder or isolated population-based studies with related subjects, extended large pedigree studies, family-based studies or hybrids of these designs. Our method is also flexible for simultaneously testing associations involving different phenotypic data types. To demonstrate this flexibility, we also apply the method to the Collaborative Study on the Genetics of Alcoholism data.

Our method is based on the generalized quasi-likelihood scoring (GQLS) approach that was proposed by Feng *et al.* (2011). The GQLS method focuses on testing the association between a marker and a trait. The observed allele frequency is treated as the response variable and the phenotypic value of the trait is treated as a covariate. For a biallelic marker such as an SNP, a logistic link is used to model the association between the distribution of the marker allele frequency and the phenotypic values. Given the relationship between subjects, an exact covariance matrix of the response vector can be explicitly computed. Provided that the mean and the covariance of the response vector are known, a quasi-likelihood scoring function and the quasi-likelihood statistic (which is denoted by W_G) can be expressed explicitly. Extending to the test involving multiallelic markers, the W_G-statistic follows an asymptotic χ^2_a-distribution with a determined by the number of alleles of the marker. Because the phenotype is treated as a covariate, the distribution of the phenotype does not need to be specified. The GQLS method can then be extended to test multiple traits simultaneously. We name this new method the generalized quasi-likelihood scoring method (GQLSM) for multiple traits. In this paper, we shall derive the proposed GQLSM in Section 2. Section 3 describes the simulation studies to assess the validity and the power of the method proposed. In Section 4, we shall apply our method to the Holstein cattle data and the Collaborative Study on the Genetics of Alcoholism data. Discussions and suggestions for future work are provided in Section 5.

Some example data and the programs that were used to analyse them can be obtained from

`http://wileyonlinelibrary.com/journal/rss-datasets`

2. Methods

We first describe the model for subjects sampled from a single extended family or from an isolated or founder population in Section 2.1. Then, we extend the model to the case when the subjects are from multiple independent families in Section 2.2.

2.1. Single large family study design

Suppose that in a genetic study we have phenotypic information on k traits for a sample of n individuals from a single pedigree or an isolated or founder population. The relationships between individuals are assumed to be known. To accommodate an inbred population, the inbreeding configuration within each individual is also assumed to be known. This requirement can be relaxed if genomewide genetic data are available from which the relationships can be inferred. Let $\mathbb{X}$ be an $n \times (k+1)$ design matrix such that

$$\mathbb{X} = \begin{pmatrix} 1 & X_{11} & \cdots & X_{1k} \\ 1 & X_{21} & \cdots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{nk} \end{pmatrix},$$

where the $(j+1)$th column is for the jth trait, $j = 1, \ldots, k$. Let $\mathbf{X}_i' = (1, X_{i1}, X_{i2}, \ldots, X_{ik})$ be the ith row of $\mathbb{X}$ for individual i. Note that the phenotypes are now treated as covariates. Thus, they can be of any type of data (e.g. binary, ordinal, count or continuous). For a biallelic marker to be tested, without loss of generality, label the two alleles '0' and '1'. Let $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ represent the proportion of allele 1 in the observed genotype for each individual, so that Y_i equals 0, $\frac{1}{2}$ or 1. Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)' = E(\mathbf{Y}|\mathbb{X})$, $0 < \mu_i < 1$, for all i. Then, $2Y_i$ follows a binomial $(2, \mu_i)$ distribution. With a logistic link,

$$\mu_i = E(Y_i|\mathbf{X}_i) = \frac{\exp(\mathbf{X}_i'\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i'\boldsymbol{\beta})}, \tag{1}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)'$. If the marker under investigation is not associated with any trait, all coefficients that are associated with the traits should be 0. Thus, we can perform a simultaneous association test for all traits in the form of

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0 \quad \text{against} \quad H_a : \text{at least one } \beta_j \neq 0.$$

Under the null hypothesis, $\mu_i = \mu = \exp(\beta_0)/\{1+\exp(\beta_0)\}$ for all i and the mean response vector becomes $\boldsymbol{\mu} = E(\mathbf{Y}|\mathbb{X}) = E(\mathbf{Y}) = \mu\mathbf{1}$ with $\mathbf{1}$ being an n-vector of 1s. The covariance matrix of $\mathbf{Y}$ under H_0 is given by $\boldsymbol{\Sigma}_0 = \frac{1}{2}\mu(1-\mu)\boldsymbol{\rho}$, where

$$\boldsymbol{\rho} = \begin{pmatrix} 1+\phi_1 & 2\phi_{12} & \cdots & 2\phi_{1n} \\ 2\phi_{12} & 1+\phi_2 & \cdots & 2\phi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 2\phi_{1n} & 2\phi_{2n} & \cdots & 1+\phi_n \end{pmatrix} \tag{2}$$

is the correlation matrix of $\mathbf{Y}$. In $\boldsymbol{\rho}$, ϕ_i is the inbreeding coefficient of individual i and ϕ_{ij} is the kinship coefficient between individual i and individual j. When two individuals are not related, $\phi_{ij} = 0$. When a population is outbred, $\phi_i = 0$. See Feng *et al.* (2011) for justification.

The quasi-likelihood score functions are in a $(k+1)$-vector that has the form

$$U(\boldsymbol{\beta}) = (U_{\beta_0}(\boldsymbol{\beta}), U_{\beta_1}(\boldsymbol{\beta}), \ldots, U_{\beta_k}(\boldsymbol{\beta}))' = D'\boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\boldsymbol{\mu}), \tag{3}$$

where D is an $n \times (k+1)$ derivative matrix of the form

$$D = \frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\beta}} = \left(\frac{\partial\boldsymbol{\mu}}{\partial\beta_0}, \frac{\partial\boldsymbol{\mu}}{\partial\beta_1}, \ldots, \frac{\partial\boldsymbol{\mu}}{\partial\beta_k}\right) \tag{4}$$

and $\boldsymbol{\Sigma}$ is the covariance matrix of $\mathbf{Y}$. Under the null hypothesis, $\boldsymbol{\mu} = \mu\mathbf{1}$, the covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, $D = \mu(1-\mu)\mathbb{X}$, and

$$U(\boldsymbol{\beta}) = 2\mathbb{X}'\boldsymbol{\rho}^{-1}(\mathbf{Y} - \mu\mathbf{1}). \tag{5}$$

Under the null hypothesis H_0 that $(\beta_1, \ldots, \beta_k) = \mathbf{0}$, setting the quasi-likelihood score of $U_{\beta_0}(\beta_0, \mathbf{0}) = 0$ yields the estimate of μ given by

$$\hat{\mu} = (\mathbf{1}'\boldsymbol{\rho}^{-1}\mathbf{1})^{-1}\mathbf{1}'\boldsymbol{\rho}^{-1}\mathbf{Y}. \tag{6}$$

According to Cox and Hinkley (1974) and Heyde (1997), the quasi-score statistic is given by

$$W = U_{\boldsymbol{\beta}_{-0}}(\hat{\beta}_0, \mathbf{0})' \operatorname{cov}_0^{-1}\{U_{\boldsymbol{\beta}_{-0}}(\hat{\beta}_0, \mathbf{0})\} U_{\boldsymbol{\beta}_{-0}}(\hat{\beta}_0, \mathbf{0}), \tag{7}$$

where $U_{\boldsymbol{\beta}_{-0}}(\hat{\beta}_0, \mathbf{0})$ is a vector of score functions given by equation (3) under the null hypothesis and omitting the score function for β_0 and $\operatorname{cov}_0^{-1}\{U_{\boldsymbol{\beta}_{-0}}(\hat{\beta}_0, \mathbf{0})\}$ is a $k \times k$ matrix that omits the first row and the first column of the inverse of the information matrix $\mathbf{I}(\beta)$. Both are computed under the null hypothesis H_0 that $\beta_1 = \beta_2 = \ldots = \beta_k = 0$. We derive an explicit form for the W-statistic and denote it as W_{M}, where the subscript M stands for multiple traits. The W_{M}-statistic is given by

$$W_{\mathrm{M}} = \frac{2}{\hat{\mu}(1-\hat{\mu})}(\mathbf{Y} - \hat{\mu}\mathbf{1})'\boldsymbol{\rho}^{-1}\mathbb{X}_{-1}[(\mathbb{X}'\boldsymbol{\rho}^{-1}\mathbb{X})^{-1}]_{-1,-1}\mathbb{X}'_{-1}\boldsymbol{\rho}^{-1}(\mathbf{Y} - \hat{\mu}\mathbf{1}). \tag{8}$$

Here, $\mathbb{X}_{-1}$ indicates the removal of the first column (the intercept column) of the design matrix $\mathbb{X}$, and $[(\mathbb{X}'\boldsymbol{\rho}^{-1}\mathbb{X})^{-1}]_{-1,-1}$ indicates the removal of the first row and the first column of the matrix $(\mathbb{X}'\boldsymbol{\rho}^{-1}\mathbb{X})^{-1}$. Under the null hypothesis, W_{M} follows a χ^2-distribution with the degrees of freedom determined by the rank of $\mathbb{X}'\boldsymbol{\rho}^{-1}\mathbb{X}$. If the k phenotypes being tested are linearly independent, $W_{\mathrm{M}} \sim \chi^2_k$ asymptotically. When $k = 1$, we test only a single trait, and the W_{G}-statistic that was described in Feng *et al.* (2011) for a test on a single trait becomes a special case of our W_{M}-statistic.

2.2. Multiple-families study design

Now we derive the W_{M}-statistic when a sample of n individuals is from F independent families. Suppose that, among n individuals, n_f are from the fth family and $n = \Sigma_{f=1}^{F} n_f$. Let $\mathbf{Y}_f = (Y_{1f}, \ldots, Y_{n_f f})'$ for the observed genotypes of individuals from family f. Similarly, the correlation matrix $\boldsymbol{\rho}_f$ is defined for family f. We arrange the response vector $\mathbf{Y} = (\mathbf{Y}'_1, \ldots, \mathbf{Y}'_F)'$ and, correspondingly, the design matrix is of the form

$$\mathbb{X} = (\mathbb{X}'_1, \mathbb{X}'_2, \cdots, \mathbb{X}'_f, \cdots, \mathbb{X}'_F)',$$

where

$$\mathbb{X}_f = \begin{pmatrix} 1 & X_{1f1} & \cdots & X_{1fk} \\ 1 & X_{2f1} & \cdots & X_{2fk} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n_f f1} & \cdots & X_{n_f fk} \end{pmatrix}.$$

Again, under the null hypothesis, the mean response vector becomes a constant vector in the form of $\boldsymbol{\mu} = E(\mathbf{Y}|\mathbb{X}) = E(\mathbf{Y}) = \mu\mathbf{1}$. The overall covariance matrix of $\mathbf{Y}$ for the whole sample has the form

$$\boldsymbol{\Sigma}_0 = \tfrac{1}{2}\mu(1-\mu)\boldsymbol{\rho}, \tag{9}$$

where

$$\boldsymbol{\rho} = \begin{pmatrix} \boldsymbol{\rho}_1 & 0 & \cdots & 0 \\ 0 & \boldsymbol{\rho}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \boldsymbol{\rho}_F \end{pmatrix}.$$

Under the null hypothesis that $\beta_1 = \ldots = \beta_k = 0$, the quasi-likelihood estimate of μ is given by equation (6) or can be rewritten as

$$\hat{\mu} = \left(\sum_{f=1}^{F} \mathbf{1}'_f \boldsymbol{\rho}_f^{-1} \mathbf{1}_f \right)^{-1} \left(\sum_{f=1}^{F} \mathbf{1}'_f \boldsymbol{\rho}_f^{-1} \mathbf{Y}_f \right),$$

where $\mathbf{1}_f$ is the n_f-vector of 1s. We derive the explicit form for the W_{M}-statistic as

$$W_{\mathrm{M}} = \frac{2}{\hat{\mu}(1-\hat{\mu})} \left(\sum_{f=1}^{F} \mathbb{X}'_{f,-1} \boldsymbol{\rho}_f (\mathbf{Y}_f - \hat{\mu}\mathbf{1}_f) \right)' \left[\left(\sum_{f=1}^{F} \mathbb{X}'_f \boldsymbol{\rho}_f \mathbb{X}_f \right)^{-1} \right]_{-1,-1} \left(\sum_{f=1}^{F} \mathbb{X}'_{f,-1} \boldsymbol{\rho}_f (\mathbf{Y}_f - \hat{\mu}\mathbf{1}_f) \right). \tag{10}$$

Similarly, $\mathbb{X}'_{f,-1}$ indicates the removal of the first column of design matrix $\mathbb{X}'_f$ for the fth family and $[(\Sigma_{f=1}^{F} \mathbb{X}'_f \boldsymbol{\rho}_f \mathbb{X}_f)^{-1}]_{-1,-1}$ indicates the removal of the first row and the first column of the matrix $(\Sigma_{f=1}^{F} \mathbb{X}'_f \boldsymbol{\rho}_f \mathbb{X}_f)^{-1}$. If the k phenotypes being tested are not linearly dependent, the W_{M}-statistic follows the χ^2_k-distribution asymptotically. Here, equation (10) is the alternative form of equation (8). When the sample size is large, equation (10) breaks down the whole sample to F independent families, which avoids the manipulation of high dimensional matrices and makes the computation more feasible.

3. Simulation studies

We conduct simulation studies to validate the approximation of the distribution of the W_{M}-statistic under the null hypothesis by the χ^2-distribution. We also compare the power that is achieved by testing multiple traits simultaneously with the power achieved by testing each trait individually. SNP data are simulated. We consider three study designs: a single large pedigree study design, a multiple small family study design and a multiple large family study design. In the first study, all phenotypes are quantitative. In the second study, we consider a combination of three different types of trait: quantitative, binary and ordinal. In the third study, we consider combinations of different genetic models (e.g. additive, epistasis and recessive) in the disease model as well as different numbers of traits influenced by a single gene. The software `KinInbCoef` (Bourgain and Zhang, 2009) is used to compute the kinship coefficients for calculating the correlation matrix $\boldsymbol{\rho}$.

3.1. Study 1: single large pedigree study design

We grow a family from a single individual, where each single individual has a probability of 0.8 of being assigned a spouse or a probability of 0.2 of remaining single. For each couple, a number of offspring are generated according to a Poisson distribution with a mean of 3. We grow families for six generations. Family members of the top three generations are removed to mimic the practical situations in which phenotypic information and deoxyribonucleic acid samples are less likely to be available for more than three generations back. However, the genealogy information of the entire pedigree remains for calculating the correlation matrix $\boldsymbol{\rho}$. In our simulation procedure, any family that stops growing before the completion of six generations by natural degeneration, or any family with the number of family members in the last three generations less than the

Table 1. Empirical null rejection rate comparison for the single large pedigree study design

α	*Quantitative trait*	*Rejection rates for the following sample sizes:*		
		100	*200*	*500*
0.05	1	0.055	0.047	0.045
	2	0.056	0.055	0.053
	3	0.042	0.053	0.043
	4	0.056	0.046	0.058
	5	0.066	0.056	0.047
	Union	0.243	0.213	0.222
	Simultaneous	0.056	0.054	0.055
0.01	1	0.01	0.012	0.013
	2	0.012	0.013	0.013
	3	0.011	0.010	0.009
	4	0.011	0.010	0.013
	5	0.013	0.011	0.013
	Union	0.056	0.054	0.057
	Simultaneous	0.009	0.011	0.013

desired size, is disregarded. Here, we generate large outbred pedigrees that have sizes of at least 100, 200 and 500 in the last three generations. For each founder (an individual with parental information unknown), the marker genotype is generated by random mating. The genotypes of descendants are generated according to the Mendelian law of segregation.

In the first study, we consider an association test with five quantitative phenotypes. To assess the type I error rate, five quantitative traits are generated for each individual. To do this, we first generate an SNP with the minor allele frequency set to 0.3. Denote the genotype of the SNP by $G = 0, 1, 2$ for individuals having no, one or two copies of the minor allele respectively. We denote the five traits for individual i by $X_{i1}, X_{i2}, X_{i3}, X_{i4}$ and X_{i5}. Given the genotype G_i of the SNP, each quantitative trait X_{ij} is generated from $N\{b_j(-1+G_i), \sigma^2\}$, $j = 1, \ldots, 5$, where $b = (0.6, 0.4, 0.2, -0.1, 0.3)$ are the effects of the SNP on each of the five traits respectively, and $\sigma^2 = 1$. By doing this, an individual's traits are genetically correlated. Then, another SNP that is not linked to the causal SNP is generated. The minor allele frequency of this unlinked SNP is set at 0.3. We generate 1000 data sets. Here, the analysis of the causal SNP is used to access the power and the analysis of the unlinked SNP is to evaluate the type I error control.

For each simulated data set, we perform the association test on five traits simultaneously. We compute the W_{M}-statistic and take the $(1-\alpha)$th quantile of the χ^2_5-distribution to be the rejection threshold. We also perform the association test on each individual trait. For each trait, we compute the W_{G}-statistic that was given by Feng *et al.* (2011) and take the rejection threshold to be the $(1-\alpha)$th quantile of the χ^2_1-distribution.

The empirical rejection rates are summarized in Table 1. For our simultaneous test, the empirical null rejection rates are very close to each nominal level of significance. For each individual test, the empirical null rejection rates are again very close to each nominal level of significance; however, the overall type I error rate, calculated as the union of the null rejections from the five individual tests, is elevated from 0.213 to 0.243 for $\alpha = 0.05$ and from 0.054 to 0.057 for $\alpha = 0.01$. These overall type I error rates are very close to the theoretical familywise error rates. Denote the familywise error rate by α_{F} and $\alpha_{\mathrm{F}} = 1-(1-\alpha)^k$ if there are k individual tests. When $k = 5$,

Table 2. Power comparison with the single large pedigree study design

	Quantitative trait	*Results for the following sample sizes:*		
		100	*200*	*500*
$\alpha_F = 0.05$				
0.01†	1	0.461	0.898	0.992
0.01	2	0.183	0.525	0.834
0.01	3	0.038	0.119	0.209
0.01	4	0.018	0.030	0.38
0.01	5	0.108	0.227	0.537
	Union	0.599	0.952	1
0.05	Simultaneous	0.720	0.985	1
$\alpha_F = 0.01$				
0.002	1	0.290	0.780	0.967
0.002	2	0.075	0.344	0.674
0.002	3	0.011	0.054	0.087
0.002	4	0.004	0.008	0.013
0.002	5	0.035	0.139	0.333
	Union	0.366	0.870	0.990
0.01	Simultaneous	0.511	0.943	0.999
$\alpha_F = 0.001$				
0.0002	1	0.118	0.548	0.917
0.0002	2	0.024	0.156	0.398
0.0002	3	0.001	0.015	0.023
0.0002	4	0.000	0.001	0.000
0.0002	5	0.008	0.046	0.146
	Union	0.141	0.625	0.950
0.001	Simultaneous	0.248	0.844	0.973

†For an individual test, $\alpha = \alpha_F/5$.

an individual test of $\alpha = 0.05$ corresponds to an α_F of 0.226 and an individual test of $\alpha = 0.01$ corresponds to an α_F of 0.049. These empirical results show that the union of the individual null rejection rate is very close to the α_F. Our method effectively controls the overall type I error rate α_F at nominal levels. In addition, the Q–Q-plots in Fig. S1 in the on-line supplementary materials show that the χ^2_5-distribution very closely approximates the distribution of the W_M-statistic under the null hypothesis for three different sample sizes.

In assessing the power of the GQLSM test and comparing the power with multiple individual tests, we control the overall type I error rate α_F at 0.05, 0.01 and 0.001 to obtain a fair comparison. The α-value for each individual test can be obtained by using $\alpha_F = 1 - (1-\alpha)^k$ or using the Bonferroni control $\alpha = \alpha_F/k$, which give almost identical values.

The results that are summarized in Table 2 show that our simultaneous testing of five traits is consistently more powerful than the union of five individual tests over different α_F levels of significance and sample sizes. The gain in power is particularly obvious when the sample size is relatively small.

3.2. Study 2: multiple small pedigrees study design

For the multiple small pedigree study design, families are grown for a maximum of three genera-

tions. We allow the simulated sample to consist of families and independent individuals. Family sizes range from 1 to 22 with an average of about 6. We let the total sample sizes be about 100, 200 and 500. To show the flexibility of our method for the analysis of phenotypic traits of different types of data, we consider three different traits: quantitative, binary and ordinal. They are denoted for individual i by X_{i1}, X_{i2} and X_{i3} respectively. Similarly, we first generate an SNP G with the minor allele frequency set to 0.3. Given the genotype G_i of the SNP, the quantitative trait X_{i1} is generated from $N\{b(-1+G_i), \sigma^2\}$, where $b=0.3$ and $\sigma^2=1$. The binary trait X_{i2} is generated from a Bernoulli(π) distribution where $\pi=0.5-0.2G_i$. The ordinal trait X_{i3} will take a value of 1, 2 or 3 according to the probability of $\pi_1=1-\pi$, $\pi_2=\pi-\pi^2$, $\pi_3=\pi^2$ and $\Sigma_{i=1}^{3}\pi_i=1$. By doing this, phenotypes of individuals within a family are genetically connected. Then, another SNP that is not linked to the causal SNP is generated. The minor allele frequency of this unlinked SNP is set at 0.3. We generate 1000 data sets. For each simulated data set, we perform the association test on the three traits simultaneously with both the causal SNP and the unlinked SNP. We compute the W_{M}-statistic and take the $(1-\alpha)$th quantile of the χ_3^2-distribution to be the rejection threshold. We also perform the association test on each individual trait. For each trait, we compute the W_{G}-statistic and take the rejection threshold to be the $(1-\alpha)$th quantile of the χ_1^2-distribution.

The empirical rejection rates are summarized in Table S1 in the on-line supplementary materials. The results show a similar pattern to that of the results based on the single large pedigree study design. The empirical null rejection rates based on the simultaneous test are very close to each nominal level of significance. For each individual test, the empirical null rejection rates are again very close to each nominal level of significance; however, the overall type I error rates α_{F} are elevated from 0.132 to 0.16 for $\alpha=0.05$ and from 0.024 to 0.036 for $\alpha=0.01$. These overall type I error rates are very close to the theoretical familywise error rates of an individual test with $\alpha=0.05$ (an α_{F} of 0.142) and an individual test with $\alpha=0.01$ (an α_{F} of 0.0297) for the union of three individual tests. In addition, the Q–Q-plots in Fig. S2 in the on-line supplementary materials show that the χ_3^2-distribution well approximates the distribution of the W_{M}-statistic under the null hypothesis for the three different sample sizes.

The power assessment results are summarized in Table S2 in the supplementary materials. The results show that our simultaneous testing of three traits is consistently more powerful than the union of three individual tests over different α_{F} levels of significance and sample sizes.

3.3. Study 3: multiple large pedigrees study design

With the multiple large pedigree study design, families are grown for six generations. In contrast with study 1, the number of family members in the last three generations could be relatively small and, thus, each sample of size of at least 100, 200 and 500 in the last three generations would consist of more than one family.

In this study, we consider three traits, X_1, X_2 and X_3, with five causal SNPs G_1, G_2, G_3, G_4 and G_5 and their minor allele frequencies set at 0.2, 0.3, 0.15, 0.25 and 0.25 respectively. As shown in Table 3, SNPs G_1, G_2 and G_3 influence all three traits. SNP G_4 influences trait 1 and 2. SNP G_5 influences only trait 3. These five SNPs are on different chromosomes. So, genotypes of each SNP are generated independently.

Given the genotypes G_{i1}–G_{i5} for individual i, we generate the three traits according to the following models:

$$X_{i1}=\mu_1+\gamma_1(G_{i1},G_{i2})+aG_{i3}+bG_{i4}+\varepsilon_{i1},$$
$$X_{i2}=\mu_2+aG_{i1}+bG_{i2}+\gamma_2(G_{i3},G_{i4})+\varepsilon_{i2},$$

Table 3. Involvement of five SNPs to the three traits in study 3

Trait	*SNPs*				
	G_1	G_2	G_3	G_4	G_5
X_1	Yes	Yes	Yes	Yes	No
X_2	Yes	Yes	Yes	Yes	No
X_3	Yes	Yes	Yes	No	Yes

Table 4. Interaction effect between G_1 and G_2 on X_1 in study 3

G_1	γ_1 *for the following values of* G_2*:*		
	0	*1*	*2*
0	−1.25	−1.25	0.25
1	−1.25	−1.25	0.25
2	0.25	0.25	1.25

Table 5. Interaction effect between G_3 and G_4 on X_2 in study 3

G_3	γ_2 *for the following values of* G_4*:*		
	0	*1*	*2*
0	−1	−1	−1
1	−1	2	2
2	−1	2	2

$$X_{i3} = \mu_3 + aG_{i1} + bG_{i2} + cG_{i3} + \gamma_3(G_{i5}) + \varepsilon_{i3}.$$

Without loss of generality, we set $\mu_j = 0$ for $j = 1, 2, 3$. We set $a = 1$, $b = 0.5$ and $c = 1.2$, with ε_{ij} representing random environmental error generated from $N(0, 1)$ for all traits. Genotypes of all G_{ij} are in the form of 0, 1 and 2. As shown in Tables 4 and 5, γ_1 quantifies the interaction effect between G_1 and G_2 on X_1, and γ_2 quantifies the interaction effect between G_3 and G_4 on X_2. We set $\gamma_3 = 2$ if $G_{i5} = 2$, and $\gamma_3 = 0$ otherwise, for the recessive effect of G_5 on the third trait X_3.

We simulate 1000 data sets. For each simulated data set, we perform the simultaneous test on the three traits and individual tests on each trait. We set $\alpha_F = 0.05$, 0.01, 0.001. The results are summarized in Table S3 in the on-line supplementary materials. Note that rejection rates in parentheses are empirical type I error rates as, for example, G_4 is not responsible for the third trait X_3. We see that the empirical null rejection rates are close to their nominal levels

of significance. Overall, our GQLSM method is consistently more powerful than the union of individual tests when the SNP under investigation influences more than one trait. When the SNP influences only one single trait, the power that is achieved by our GQLSM method is comparable with the power that is achieved by individual tests at the same α_F-level.

4. Real data analysis

4.1. Application to Holstein cattle study

There are 821 progeny-tested proven bulls in the data. The relationship between bulls and their sires and dams is complicated; some bulls are also the sires of other bulls. Most of the bulls in the sample have non-zero inbreeding coefficients. A genealogy of the population, tracing back 25 generations to the earliest animal born in 1909, was used to compute the kinship–inbreeding coefficient with the software CFC (Sargolzaei *et al.*, 2006). Each bull was genotyped by using the Affymetrix MegAllele GeneChip bovine mapping 10K SNP array (Affymetrix, 2005). SNPs on sex chromosomes or with more than 20% values missing were excluded from the study. A total of 6418 SNPs are analysed. The experimental design is mainly a granddaughter design, in that the milk productivities of daughter and granddaughter cows of a bull are used to estimate the EBVs of the bull. The phenotypes that were used in the analysis were trait EBVs released in November 2008 and provided by the Canadian Dairy Network (Guelph, Canada). In this study, we aim to identify SNPs or genome regions that are associated with the five milk-product-related EBVs: milk yield MY, protein yield PY, fat yield FY, protein percentage PP and fat percentage FP.

First, we simultaneously test the association between each SNP and the five EBV traits by using our GQLSM method. Second, we individually test the association between each SNP and each of the five traits. We also record the union of significant SNPs from the five individual tests. In Table 6, we report the number of significant SNPs found by the individual tests and by the simultaneous test. The Venn diagrams in Fig. 1 show the overlap of significant SNPs between the simultaneous test and the union of the five individual tests for each EBV trait while controlling the overall type I error rate at $\alpha_F = 0.05, 0.01, 0.001$. Over three levels of significance, the simultaneous test covers the majority of the significant SNPs found by the union of the individual tests. However, the union of the individual tests misses more than half of the SNPs that were found to be significant by the simultaneous test. After Bonferroni adjustment for multiple testing, 149 of the 6418 SNPs tested at $\alpha_F = 0.05$ for the simultaneous test (or, at the 7.78×10^{-6} level of significance for each SNP) remain significant; however, only 47 SNPs in the union of individual tests remain significant at the $\alpha_F = 0.05$ level (or at the 1.56×10^{-6} level of significance for each SNP). Of these 47 significant SNPs, 45 are also found to be significant by the simultaneous test at the same α_F-level.

Many of the 104 significant SNPs that were found by the simultaneous test but not by the

Table 6. Number of significant SNPs found by individual tests and the simultaneous test in the Holstein cattle data analysis

α_F	*Individual test results*						*Simultaneous test result*
	MY	*PY*	*FY*	*PP*	*FP*	*Union*	
0.05	312	264	318	266	290	880	1619
0.01	156	124	140	129	155	436	844
0.001	73	31	54	61	64	168	397

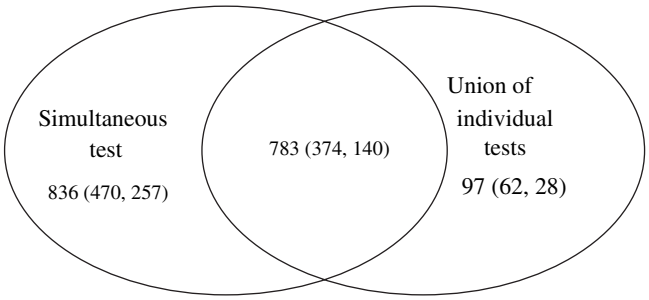

Fig. 1. Venn diagram to show the overlap between the simultaneous test and the union of the five individual tests for significance level α_F at 0.05: results of α_F at 0.01 and 0.001 are in parentheses

union of individual tests have been confirmed by others in the literature. For example, on the first Bos taurus autosome chromosome BTA 1, an SNP at 47.9 cM is found to have a strong level of significance in the simultaneous test (p-value 5.44×10^{-8}) but moderate levels of significance with milk yield (p-value 2.81×10^{-5}), protein yield (p-value 1.61×10^{-5}) and fat yield (p-value 0.00841). This result is consistent with an SNP at 46 cM that was reported to be associated with protein yield by Rodriguez-Zas *et al.* (2002) and an SNP at 47 cM that was reported to be associated with milk yield and protein yield by Daetwyler *et al.* (2008). Another nearby SNP at 48.7 cM is also found to be strongly significant in the simultaneous test (p-value 2.23×10^{-12}). However, it is not significantly associated with any of the five traits (p-values are all greater than 0.05) in the individual tests. On the same chromosome, an SNP at 110.6 cM has a strong association signal (p-value 1×10^{-16} or less) according to the simultaneous test and a moderate significant level of association with protein yield (p-value 0.00458). This confirms the association with protein yield that was reported by Daetwyler *et al.* (2008), Heyen *et al.* (1999) and Rodriguez-Zas *et al.* (2002). On BTA 3, an SNP at 73.34 cM (p-value 1.14×10^{-6}) is surrounded by SNPs at 72 and 74 cM reported to be associated with fat yield by Rodriguez-Zas *et al.* (2002) and an SNP at 77.6 cM reported to be associated with protein percentage by Bagnato *et al.* (2008). On BTA 6, an SNP with a strong significant signal (p-value 1×10^{-16} or less) at 92.79 cM is flanked by an SNP at 91.5 cM that has been reported to be associated with milk yield (Bagnato *et al.*, 2008) and an SNP at 95 cM that has been reported to be associated with fat percentage (Viitala *et al.*, 2003). On BTA 8, an SNP with a p-value of 3.93×10^{-8} at 35.69 cM is flanked by two SNPs at 31.4 and 38 cM that are associated with protein percentage and protein yield respectively (Daetwyler *et al.*, 2008; Rodriguez-Zas *et al.*, 2002). On BTA 26, a significant SNP at 58 cM (p-value 7.37×10^{-7}) is very close to SNPs at 55 and 56 cM that are associated with milk yield (Boichard *et al.*, 2003; Daetwyler *et al.*, 2008) and an SNP at 57 cM that is associated with fat yield and protein yield (Boichard *et al.*, 2003).

There are also many literature confirmations of those SNPs found to be significant in both the simultaneous test and the union of individual tests. For example, on BTA 1, an SNP at 142.64 cM is strongly significant in the simultaneous test (p-value 2.61×10^{-14}) and also significantly associated with milk yield (p-value 2.73×10^{-8}), protein yield (p-value 3.53×10^{-6}) and fat percentage (p-value 5.4×10^{-11}). This is consistent with Daetwyler *et al.* (2008) who report an SNP at 142 cM associated with milk yield and protein yield. On BTA 2, an SNP at 40.47 cM is significant in the simultaneous test (p-value 4.44×10^{-15}) and is also significantly associated with fat yield (p-value 3.54×10^{-7}) and fat percentage (p-value 4.62×10^{-12}) and moderately significant with protein percentage (p-value 0.00018). Heyen *et al.* (1999) also reported that a microsatellite marker TGLA377 at 40.6 cM has a significant effect on the protein percentage trait.

Table 7. Number of significant SNPs by individual tests and the simultaneous test

α_F	*Individual test results*				*Simultaneous test result*
	ALDX1	*MaxDrink*	*Health*	*Union*	
0.05	205	203	206	555	617
0.01	46	35	51	128	139
0.001	9	7	9	24	26

4.2. Genetics of alcoholism data

To illustrate the flexibility of our method in the simultaneous analysis of different phenotypic data types, we apply our method to the Collaborative Study on the Genetics of Alcoholism data provided by the Genetic Analysis Workshop 14. This data set contains 1074 'white, non-Hispanic' individuals from 119 independent families. They were genotyped for a panel of 11 555 SNPs by using Affymetrix. A set of alcoholism phenotypes is also provided. The alcohol DX-DSM3R + Feighner (ALDX1) phenotype is considered as an ordinal trait with four levels: 1, 'pure unaffected', 2, 'never drink', 3, 'unaffected with some symptoms', and 4, 'affected'. The maximum number of drinks in a 24-h period (MaxDrink) phenotype is treated as a quantitative trait. The physical health problems for drinking (Health) phenotype is treated as a binary trait: 1 for yes and 0 for no. We analyse only SNPs that are on autosomes. SNPs that show only one allele type in the data set have been excluded from the analysis. In total, 10 684 SNPs are tested for association with the three traits.

The number of significant SNPs for different levels of significance in terms of α_F are summarized in Table 7. More SNPs are found to be significant in the simultaneous test than in the union of individual tests when α_F is controlled at the same level. Note that, after adjusting for Bonferroni's corrections on 10 684 SNPs tested at $\alpha_F = 0.05$, five SNPs remain significant in the simultaneous test (at the 4.68×10^{-6} level of significance for each SNP) and four SNPs remain significant in the union of individual tests (at the 1.56×10^{-6} level of significance for each SNP).

In chromosome 2, SNP tsc0057308 is strongly significant in both the simultaneous test (p-value 4.22×10^{-6}) and in the union of individual tests (p-value 4.6×10^{-7}) and is less than 2 M bases away from the KIAA1912 gene. In addition, two neighbouring SNPs, tsc0521398 and tsc1707844, which are less than 1 M base from genes KIAA1912 and EFEMP1 respectively, are found to be significant (p-values 0.0083 and 0.0097 respectively) in the simultaneous test. Heath *et al.* (2011) also identified two SNPs close to these two genes that are significantly associated with alcohol dependence factor score and weekly alcohol consumption. In chromosome 3, SNP tsc0050826 is strongly significant in both tests (p-values 9.96×10^{-5} and 0.0001 respectively) and is less than 500 k bases from gene MGLL. In chromosome 4, SNP tsc0280570 is less than 1 M base from genes DKK2 and PAPSS1 and is strongly significant in the simultaneous test (p-value 0.0003) but less significant in the union of individual tests (p-values 0.0279 and 0.0374 with ALDX1 and MaxDrink). Kalsi *et al.* (2010) also reported the association of genes DKK2 and PAPSS1 with a quantitative trait of alcohol dependence symptom counts.

5. Discussion

In many genetic studies of complex traits, multiple traits are often recorded. In most cases,

genetic association tests are performed for each individual trait and p-values must be adjusted for multiple tests. When more traits are considered, the test will be more stringent and, as a result, the test will have a lower power to detect true associations. Reducing the number of hypotheses tested becomes an appealing way to control the overall type I error rate. Simulation studies show that the simultaneous testing of multiple traits effectively controls the overall type I error rate. When the level of significance is controlled at the same level, simultaneous tests are found to be more powerful than the union of individual tests.

In a genomewide association study, when multiple traits are considered, the GQLSM method proposed would be applied in the first step to screen for SNPs that are associated with any one of the traits. Then, subsequent individual tests would be performed on the subset of significant SNPs to look at which particular traits the SNPs associated with. Results of simulation studies suggest that the GQLSM method proposed is appropriate for identifying SNPs that influence only one single trait, as the simultaneous test has a comparable power compared with individual tests. In contrast, the GQLSM method outperforms the union of individual tests for identifying SNPs that influence more than one trait. In other words, the GQLSM method would be a more powerful approach to screen SNPs for genetic pleiotropic effects.

The GQLSM method proposed is an extension of the GQLS method and thus inherits some of its features. For example, the GQLS method is a general method that enables the analysis of data from a broad range of study designs and with different phenotypic data types. The GQLSM method can be further generalized to test for multiple traits that are of different types of data or hybrids of different types of data. When computing the test statistic W_{M}, all trait information is used. Unless traits are linearly dependent, the number of degrees of freedom of the χ^2-distribution equals the number of traits being tested. When testing a large number of correlated quantitative traits, techniques such as principal components analysis might be considered to reduce the dimensions. Instead of testing all traits simultaneously, one may test only the most important principal components. By doing this, a smaller degree of freedom of the χ^2-distribution is used to approximate the distribution of the W_{M}-statistic and thus a smaller threshold value will be used. However, selection of the principal components for the association test must be done carefully. The first m most important principal components among all p principal components could contribute nothing to explaining the variation in the response variable. This problem has been reported by Hadi and Lin (1998) and Jolliffe (1982) via various examples in real analysis. Therefore, the possibility of incorporating dimension reduction techniques in the GQLSM method to increase the power of the test further by lowering the threshold value is worth investigating.

In genetic association studies, confounding factors such as other environmental factors that contribute to the variation of the phenotypic traits should also be considered. However, in the current GQLSM model, the phenotypic traits are treated as covariates and confounding factors that affect these traits cannot be directly incorporated in the model. One possible way to adjust for the confounding effects is to perform a so-called 'two-stage residual outcome' analysis. In the two-stage residual outcome analysis, regression analysis of the phenotype on the confounding factors is performed in stage 1 to obtain residuals. Then, the residual is treated as a phenotype for the genetic association test in the second stage. For application of our GQLSM method in the two-stage residual analysis, multiple residual outcomes can be obtained through multiple individual regression analyses on each trait. Then, in the second stage, we test the association between an SNP and multiple residual outcomes simultaneously by using our GQLSM method. However, the current two-stage residual analysis method mainly focuses on quantitative traits and a normality assumption on the residual is often required. Further investigation is warranted with respect to the application of the GQLSM method in the two-stage residual outcome strategy

and extension of the current method to accommodate other types of data, such as binary, ordinal or count phenotypic traits, as well as other study designs.

6. Supplementary materials

Please see the on-line supporting information for the link to the R code (R Development Core Team, 2011) implementing the GQLSM method and additional simulation study results.

Acknowledgements

This study makes use of Holstein cattle data and Collaborative Study on the Genetics of Alcoholism data. The Holstein cattle data were provided by Dr Flavio Schenkel, Department of Animal and Poultry Science, University of Guelph. The Collaborative Study on the Genetics of Alcoholism data were provided by the Collaborative Study on the Genetics of Alcoholism (U10 AA008401); they were assembled for the Genetic Analysis Workshop 14 and are described in detail by Edenberg *et al.* (2005).

This work was supported by a Natural Sciences and Engineering Research Council of Canada individual discovery grant.

References

Affymetrix (2005) Affymetrix MeAllele GeneChip Bovine 10K SNP array. Affymetrix, South San Francisco. (Available from `http://www.affymetrix.com/support/technical/datasheets/bovine10k_snp_datasheet.pdf`.)

American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R)*. Washington DC: American Psychiatric Association.

Bagnato, A., Schiavini, F., Rossoni, A., Maltecca, C., Dolezal, M., Medugorac, I., Solkner, J., Russo, V., Fontanesi, L., Friedmann, A., Soller, M. and Lipkin, E. (2008) Quantitative trait loci affecting milk yield and protein percentage in a three-country Brown Swiss population. *J. Dairy Sci.*, **91**, 767–783.

Boichard, D., Grohs, C., Bourgeois, F., Cerqueira, F., Faugeras, R., Neau, A., Rupp, R., Amigues, Y., Boscher, M. Y. and Levziel, H. (2003) Detection of genes influencing economic traits in three French dairy cattle breeds. *Genet. Selectn Evoln*, **35**, 77–101.

Bourgain, C., Hoffjan, S., Nicolae, R., Newman, D., Steiner, L., Walker, K., Reynolds, R., Ober, C. and McPeek, M. S. (2003) Novel case-control test in a founder population identifies P-seletin as an Antopy-susceptibility locus. *Am. J. Hum. Genet.*, **73**, 612–626.

Bourgain, C. and Zhang, Q. (2009) KinInbCoef: calculation of kinship and inbreeding coefficients based on pedigree information. University of Chicago, Chicago. (Available from `http://galton.uchicago.edu/~mcpeek/software/KinInbCoef/index.html`.)

Cox, D. R. and Hinkley, D. V. (1974) *Theoretical Statistics*. London: Chapman and Hall.

Daetwyler, H. D., Schenkel, F. S., Sargolzaei, M. and Robinson, J. A. B. (2008) A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. *J. Dairy Sci.*, **91**, 3225–3236.

Edenberg, H. J., Bierut, L. J., Boyce, P., Cao, M., Cawley, S., Chiles, R., Doheny, K. F., Hansen, M., Hinrichs, T., Jones, K., Kennedy, G. C., Liu, G., Marcus, G., McBride, C., Murray, S. S., Oliphant, O., Pettengill, J., Porjesc, B., Pugh, E. W., Rice, J. P., Rubano, T., Shannon, S., Steeke, R., Tischfield, J. A., Tsai, Y. Y., Zhang, C. and Begleiter, H. (2005) Description of the data from the Collaborative Study on the Genetics of Alcoholism (COGA) and single-nucleotide polymorphism genotyping for Genetic Analysis Workshop 14. *BMC Genet.*, **6**, suppl. I, article S2.

Edwards, K. L., Mahaney, M. C., Motulsky, A. G. and Austin, M. A. (1999) Pleiotropic genetic effects on LDL size, plasma triglyceride, and HDL cholesterol in families. *Artscler. Thromb. Vasc. Biol.*, **19**, 2456–2464.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, Jr, R. A., Winokur, G. and Munoz, R. (1972) Diagnostic criteria for use in psychiatric research. *Arch. Gen. Psychiatr.*, **26**, 57–63.

Feng, Z., Wong, W., Gao, X. and Schenkel, F. (2011) Generalized genetic association study with samples of related individuals. *Ann. Appl. Statist.*, **5**, 2109–2130.

Hadi, A. S. and Ling, R. F. (1998) Some cautionary notes on the use of principal components regression. *Am. Statistn*, **52**, 15–19.

Hartl, D. L. and Jones, E. W. (2005) *Genetics: Analysis of Genes and Genomes*, 6th edn. Boston: Jones and Bartlett.

Heath, A. C., Whitfield, J. B., Martin, N. G., Pergadia, M. L., Goate, A. M., Lind, P. A., McEvoy, B. P., Schrage, A. J., Grant, J. D., Chou, Y., Zhu, R., Henders, A. K., Medland, S. E., Gordon, S. D., Nelson, E. C., Agrawal, A., Nyholt, D. R., Bucholz, K. K., Madden, P. A. F. and Montgomery, G. W. (2011) A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol. Psychiatr.*, **70**, 513–518.

Heyde, C. (1997) *Quasi-likelihood and Its Application: a General Approach to Optimal Parameter Estimation*. New York: Springer.

Heyen, D. W., Weller, J. I., Ron, M., Band, M., Beever, J. E., Feldmesser, E., Da, Y., Wiggans, G. R., VanRaden, P. M. and Lewin, H. A. (1999) A genome scan for QTL influencing milk production and health traits in dairy cattle. *Physiol. Genom.*, **1**, 165–175.

Horvath, S., Xu, X. and Laird, N. M. (2001) The family based association test method: strategies for studying general genotype–phenotype associations. *Eur. J. Hum. Genet.*, **9**, 301–306.

Jolliffe, I. T. (1982) A note on the use of principal components in regression. *Appl. Statist.*, **31**, 300–303.

Kalsi, G., Kuo, P., Aliev, F., Alexander, J., McMichael, O., Patterson, D. G., Walsh, D., Zhao, Z., Schuckit, M., Nurnberger, Jr, J., Edenberg, H., Kramer, J., Hesselbrock, V., Tischfield, J. A., Vladimirov, V., Prescott, C. A., Dick, D. M., Kendler, K. S. and Riley, B. P. (2010) A systematic gene-based screen of chr4q22-q32 identifies association of a novel susceptibility gene, DKK2, with the quantitative trait of alcohol dependence symptom counts. *Hum. Molec. Genet.*, **19**, 2497–2506.

Kullo, I. J., de Andrade, M., Boerwinkle, E., McConnell, J. P., Kardia, S. L. and Turner, S. T. (2005) Pleiotropic genetic effects contribute to the correlation between HDL cholesterol, triglycerides, and LDL particle size in hypertensive sibships. *Am. J. Hyptens.*, **18**, 99–103.

Laird, N. M., Horvath, S. and Xu, X. (2000) Implementing a unified approach to family-based tests of association. *Genet. Epidem.*, **19**, suppl. 1, S36–S42.

Lange, C., Silverman, E. K., Xu, X., Weiss, S. T. and Laird, N. M. (2003) A multivariate family-based association test using generalized estimating equations: FBAT-GEE. *Biostatistics*, **4**, 195–206.

Lin, P. I. and Mitchell, B. D. (2008) Approaches for unraveling the joint genetic determinants of schizophrenia and bipolar disorder. *Schizphren. Bull.*, **34**, 791–797.

Mahaney, M. C., Blangero, J., Comuzzie, A. G., VandeBerg, J. L., Stern, M. P. and MacCluer, J. W. (1995) Plasma HDL cholesterol, triglycerides, and adiposity: a quantitative genetic test of the conjoint trait hypothesis in the San Antonio Family Heart Study. *Circulation*, **92**, 3240–3248.

Mitchell, B. D., Kammerer, C. M., Mahaney, M. C., Blangero, J., Comuzzie, A. G., Atwood, L. D., Haffner, S. M., Stern, M. P. and MacCluer, J. W. (1996) Pleiotropic effects of genes influencing insulin levels on lipoprotein and obesity measures. *Artscler. Thromb. Vasc. Biol.*, **16**, 281–288.

R Development Core Team (2011) *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Risch, N. and Teng, J. (1998) The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases: I, DNA pooling. *Genome Res.*, **8**, 1273–1288.

Rodier, F., Campisi, J. and Bhaumik, D. (2007) Two faces of p53: aging and tumor suppression. *Nucleic Acids Res.*, **35**, 7475–7484.

Rodriguez-Zas, S. L., Southey, B. R., Heyen, D. W. and Lewin, H. A. (2008) Interval and composite interval mapping of somatic cell score, yield, and components of milk in dairy cattle. *J. Dairy Sci.*, **85**, 3081–3091.

Sargolzaei, M., Iwaisaki, H. and Colleau, J. J. (2006) CFC: a tool for monitoring genetic diversity. *8th Wrld Congr. Genetics Applied to Livestock Production, Belo Horizonte, Aug. 13th–18th*.

Sunquist, F. (2007) Malaysian mystery leopards. *Natn. Wldlfe*, **45**.

Thornton, T. and McPeek, M. S. (2007) Case-control association testing with related individuals: a more powerful quasi-likelihood score test. *Am. J. Hum. Genet.*, **81**, 321–337.

Viitala, S. M., Schulman, N. F., Koning, D. J., Elo, K., Kinos, R., Virta, A., Virta, J., Mäki-Tanila, A. and Vilkki, J. H. (2003) Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle. *J. Dairy Sci.*, **86**, 1828–1836.

Zhang, H., Liu, C. and Wang, X. (2011) An association test for multiple traits based on the generalized Kendall's tau. *J. Am. Statist. Ass.*, **105**, 473–481.

Supporting information

Additional 'supporting information' may be found in the on-line version of this article:

'Supplementary Materials for "A generalized quasi-likelihood scoring approach for simultaneously testing the genetic association of multiple traits"'.