Published in final edited form as:

Ann Hum Genet. 2015 January ; 79(1): 46–56. doi:10.1111/ahg.12089.

A family-based joint test for mean and variance heterogeneity for quantitative traits

YING CAO^{1,2}, TAYLOR J. MAXWELL³, and PENG WEI^{1,2,*}

¹ Division of Biostatistics, The University of Texas School of Public Health, Houston, Texas, USA

² Human Genetics Center, The University of Texas School of Public Health, Houston, Texas, USA

³ Computational Biology Institute, The George Washington University, Ashburn, Virginia, USA

SUMMARY

Traditional quantitative trait locus (QTL) analysis focuses on identifying loci associated with mean heterogeneity. Recent research has discovered loci associated with phenotype variance heterogeneity (vQTL), which is important in studying genetic association with complex traits, especially for identifying gene-gene and gene-environment interactions. While several tests have been proposed to detect vQTL for unrelated individuals, there are no tests for related individuals, commonly seen in family-based genetic studies. Here we introduce a likelihood ratio test (LRT) for identifying mean and variance heterogeneity simultaneously or either effect alone, adjusting for covariates and family relatedness using a linear mixed effect model approach. The LRT test statistic for normally distributed quantitative traits approximately follows χ^2 -distributions. To correct for inflated Type I error for non-normally distributed quantitative traits, we propose a parametric bootstrap based LRT that removes the best linear unbiased prediction (BLUP) of family random effect. Simulation studies show that our family-based test controls Type I error and has good power, while Type I error inflation is observed when family relatedness is ignored. We demonstrate the utility and efficiency gains of the proposed method using data from the Framingham Heart Study to detect loci associated with body mass index (BMI) variability.

Keywords

variance heterogeneity; QTL; linear mixed model; family data; BLUP

INTRODUCTION

Genetic studies of quantitative traits have focused on detecting loci associated with mean difference across genotypes (Liu *et al.*, 2013, Sun, 2012). However, there has been increasing evidence for genetic loci influencing variability of quantitative traits (Jimenez-Gomez *et al.*, 2011, Shen *et al.*, 2012). The genetic locus associated with phenotype

*Corresponding Author: Peng Wei, PhD, Division of Biostatistics and Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, 1200 Pressler Street, RAS E817, Houston, TX 77030, USA, Telephone: 1-713-500-9565, Fax: 1-713-500-9530, Peng.Wei@uth.tmc.edu.

variability has been referred to as a variance-heterogeneity quantitative trait locus (vQTL) (Rönnegård & Valdar, 2011). In addition, phenotype variance heterogeneity across genotypes can also be the result of an interaction effect with another genetic locus or environmental factor, known as gene-gene or gene-environment interactions (Paré *et al.*, 2010, Struchalin *et al.*, 2010). If a locus interacts with another factor, the phenotype distribution of this genotype group will be a mixture of distributions with different means, leading to phenotype variance heterogeneity across genotype groups at this locus. Thus, vQTL can serve as candidates for gene-gene and gene-environment interaction analyses, which will significantly reduce the dimension of interaction tests compared with pair-wise interaction screens. Moreover, vQTL can also be induced by linkage disequilibrium (LD) with a functional mean effect locus with a disparate minor allele frequency (MAF) (Cao *et al.*, 2014, Balding, 2009).

In light of the great promise of vQTL, different models and tests have been suggested for vQTL detection, for example, double generalized linear model (DGLM) (Rönnegård & Valdar, 2011), squared residual value linear modeling (SVLM) of Struchalin *et al.* (2012) and Levene's test (Paré *et al.*, 2010, Shen *et al.*, 2012, Struchalin *et al.*, 2010). DGLM can adjust for covariates, but is not robust to non-normality of quantitative traits. In contrast, Levene's test for variance heterogeneity is robust to non-normality, but cannot directly accommodate covariates. We recently proposed a likelihood ratio test (LRT) for vQTL detection in a linear model framework, which can adjust for covariates and control Type I error satisfactorily for non-normally distributed quantitative traits using parametric bootstrap (Cao *et al.*, 2014). The linear model based framework is versatile for a joint test for both mean and variance heterogeneity or a variance effect only test. Variance heterogeneity is known for being more difficult to detect than mean heterogeneity. A joint test can reveal more genetic loci associated with variance heterogeneity, which are also associated with mean heterogeneity, but may be missed by a variance effect only test (Cao *et al.*, 2014, Shen *et al.*, 2012).

Family-based designs, such as the Framingham Heart Study (FHS) (Splansky *et al.*, 2007), have been widely used in genetic studies of complex diseases. When compared with population-based designs, family-based designs have advantages that they are robust against population substructure if certain analyses, such as the transmission disequilibrium test (TDT), are performed, and allow tests for both linkage and association (Laird & Lange, 2006). For quantitative traits, linear mixed effect models (LMMs) have been commonly used to account for familial correlations. Specifically, LMM models family relatedness as a random effect following a multivariate normal distribution with a zero mean vector and a covariance matrix proportional to the kinship matrix (Aulchenko *et al.*, 2007, Chen *et al.*, 2013, Fisher, 1918, Kang *et al.*, 2010, Ouakacha *et al.*, 2013, Yu *et al.*, 2006). The kinship matrix represents the expected correlation for each pair of subjects that can be calculated based on pedigree information.

To the best of our knowledge, no vQTL test is yet developed specifically for family-based designs, although traits from related samples could be transformed to be independent using GRAMMAR (Aulchenko *et al.*, 2007) or GRAMMAR+ transformation (Belonogova *et al.*, 2013) and existing vQTL tests for independent samples could be subsequently employed. To

fill in this gap, here we propose family-based LRTs (famLRTs) in a LMM framework for vQTL detection using family-based data. The tests are denoted as famLRT_{MV}, famLRT_M, and famLRT_V for joint effects, mean effect only, and variance effect only, respectively. We demonstrate that the proposed famLRT test statistic approximately follows a χ^2 -distribution when the quantitative trait of interest is normally distributed. However, p-value calculation for non-normally distributed traits requires parametric bootstrapping after removing the best linear unbiased prediction (BLUP) of the family random effect. Simulation studies show that existing vQTL tests for unrelated individuals have inflated Type I errors for family data, while famLRTs control Type I error well and have good statistical power under various scenarios. We also use simulations to compare the proposed famLRTs with the two-step approach based on the GRAMMAR+ transformation. Finally, we demonstrate the utility of the proposed tests using the FHS data to detect single nucleotide polymorphisms (SNPs) associated with body mass index (BMI) variability and confirmed one previously identified vQTL of BMI.

METHODS

A Likelihood Ratio Test for Detecting Mean and Variance Heterogeneity in Family-based Samples

Let y_i , $i = 1, \dots, n$, denote the quantitative trait value of subject i , and G_{0i} , G_{1i} and G_{2i} be indicator variables indicating subject i is major allele homozygous, heterozygous and minor allele homozygous, respectively. We model y_i via a LMM:

$$y_i = \beta_0 + \mathbf{Z}_i \boldsymbol{\beta} + G_{1i} \gamma_1 + G_{2i} \gamma_2 + \delta_i + \varepsilon_i \quad (1)$$

where $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{im})$ are the m covariates of subject i , $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)'$ is the regression coefficients of m covariates, γ_1 and γ_2 are genotypes' fixed effects on the phenotypic mean, δ_i is the random effect capturing familial correlation, and ε_i is the residual. To model

variance heterogeneity of different genotype groups, we let $\varepsilon_i \sim N(0, G_{0i}\sigma_0^2 + G_{1i}\sigma_1^2 + G_{2i}\sigma_2^2)$, where σ_0^2 , σ_1^2 and σ_2^2 are the residual variances of major allele homozygote, heterozygote and minor allele homozygote, respectively. The random effect vector $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)'$ and the residual vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)'$ are assumed to follow multivariate normal distributions and are independent of each other: $\boldsymbol{\delta} \sim N(\mathbf{0}, \sigma_F^2 \boldsymbol{\Phi})$ and $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{V}_E)$, where $\boldsymbol{\Phi}$ is two times the kinship matrix of size $n \times n$ calculated using pedigree information representing the expected correlation for each pair of subjects, $\mathbf{V}_E$ is a diagonal matrix of size $n \times n$ with the i th element on the diagonal being $G_{0i}\sigma_0^2 + G_{1i}\sigma_1^2 + G_{2i}\sigma_2^2$.

The joint test for mean and variance heterogeneity in family samples amounts to testing $H_0: \gamma_1 = \gamma_2 = 0$ and $\sigma_0^2 = \sigma_1^2 = \sigma_2^2$ vs. H_1 : at least one “=” does not hold in model (1). We propose to employ a likelihood ratio test to test the full model (1) versus the reduced model (2):

$$y_i = \beta_0 + \mathbf{Z}_i \boldsymbol{\beta} + \delta_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \quad (2)$$

Under the assumptions of model (1), $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\alpha}, \boldsymbol{\Sigma})$, where $\mathbf{y} = (y_1, \dots, y_n)'$, the i th row of $\mathbf{X}$ is

$$(1, Z_{i1}, \dots, Z_{im}, G_{1i}, G_{2i}), \boldsymbol{\alpha} = (\beta_0, \beta_1, \dots, \beta_m, \gamma_1, \gamma_2)' \text{ and } \boldsymbol{\Sigma} = \sigma_F^2 \boldsymbol{\Phi} + \mathbf{V}_E = \sigma_F^2 \left(\boldsymbol{\Phi} + \frac{1}{\sigma_F^2} \mathbf{V}_E \right).$$

Let $\mathbf{V} = \boldsymbol{\Phi} + \frac{1}{\sigma_F^2} \mathbf{V}_E$. The log-likelihood of the LMM (1) is:

$$\begin{aligned} l &= C - \frac{1}{2} \left\{ \log |\boldsymbol{\Sigma}| + (\mathbf{y} - \mathbf{X}\boldsymbol{\alpha})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}) \right\} \\ &= C - \frac{1}{2} \left\{ n \log (\sigma_F^2) + \log |\mathbf{V}| + \frac{1}{\sigma_F^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\alpha})' \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}) \right\} \end{aligned}$$

For a given $\mathbf{V}$, the values of $\boldsymbol{\alpha}$ and σ_F^2 that maximize the log-likelihood function can be written as:

$$\hat{\boldsymbol{\alpha}}(\mathbf{V}) = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1} \mathbf{y}$$

$$\hat{\sigma}_F^2(\mathbf{V}) = \frac{1}{n} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\alpha}}(\mathbf{V}))' \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\alpha}}(\mathbf{V}))$$

Plugging the expressions of $\hat{\boldsymbol{\alpha}}(\mathbf{V})$ and $\hat{\sigma}_F^2(\mathbf{V})$, the profile log-likelihood of the LMM (1) can be written as:

$$l_P(\mathbf{V}) = C' - \frac{1}{2} \left\{ n \log (\hat{\sigma}_F^2(\mathbf{V})) + \log |\mathbf{V}| \right\}$$

Direct maximization of the profile log-likelihood $l_P(\mathbf{V})$ is slow and numerically unstable given that $\mathbf{V}$ is a sparse block diagonal matrix, especially when the number of pedigrees is large. To overcome this issue, we propose to employ generalized Cholesky decomposition to decompose $\mathbf{V}$ as $\mathbf{L}\mathbf{D}\mathbf{L}'$, where $\mathbf{L}$ is a lower triangular matrix and $\mathbf{D}$ is a diagonal matrix. We use the “gchol” function in the R package “kinship” to perform the generalized Cholesky decomposition in each iteration of the likelihood function maximization. Then $\hat{\boldsymbol{\alpha}}(\mathbf{V})$ and $\hat{\sigma}_F^2(\mathbf{V})$ can be solved as an ordinary least square problem with new outcome vector $\mathbf{y}_{new} = \mathbf{D}^{-1/2} \mathbf{L}^{-1} \mathbf{y}$ and new design matrix $\mathbf{X}_{new} = \mathbf{D}^{-1/2} \mathbf{L}^{-1} \mathbf{X}$. Finally, the profile likelihood function $l_P(\mathbf{V})$ is maximized with respect to $\mathbf{V}$ using the “BFGS” method in the R function “optim”.

The joint likelihood ratio test statistic for mean and variance heterogeneity in family samples is $-2(l_2 - l_1)$, where l_1 and l_2 are the maximum log-likelihood of full model (1) and reduced model (2), respectively. When the total sample size n is large, the test statistic approximately follows a χ^2 -distribution with 4 degrees of freedom. Similarly, to test for mean heterogeneity only while adjusting for variance heterogeneity in family samples, we test the full model (1) against the reduced model (3):

$$y_i = \beta_0 + \mathbf{Z}_i \beta + \delta_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, G_{0i}\sigma_0^2 + G_{1i}\sigma_1^2 + G_{2i}\sigma_2^2) \quad (3)$$

The likelihood ratio test statistic $-2(l_3 - l_1)$ approximately follows a χ^2 -distribution with 2 degrees of freedom when n is large, where l_3 is the maximum log-likelihood of model (3). To test for variance heterogeneity only while adjusting for mean heterogeneity in family samples, we test the full model (1) against the reduced model (4):

$$y_i = \beta_0 + \mathbf{Z}_i \beta + G_{1i}\gamma_1 + G_{2i}\gamma_2 + \delta_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \quad (4)$$

The likelihood ratio test statistic $-2(l_4 - l_1)$ approximately follows a χ^2 -distribution with 2 degrees of freedom when n is large, where l_4 is the maximum log-likelihood of model (4).

Parametric Bootstrap for Non-normally Distributed Quantitative Traits

Variance tests are known to be more sensitive to deviation from normality; see, for example, page 189 of (Lehmann, 1999). We previously observed that the variance test (LRT_V) and the joint test for mean and variance heterogeneity (LRT_{MV}) had inflated Type I error in unrelated samples when the residual normality assumption was violated, and we proposed a parametric bootstrap procedure to control the Type I error rate (Cao *et al.*, 2014). In simulation studies for family samples, we also observed Type I error inflation of famLRT_{MV} and famLRT_V when residuals were not normally distributed, while famLRT_M could control Type I error well. We intended to find a resampling method to accommodate adjustment for both covariates and family relatedness simultaneously. However, due to the complication introduced by the random effect accounting for familial correlation, none of the resampling methods we tried that simultaneously adjust for covariates and family relatedness controlled Type I error rate satisfactorily, for example, permutation of residuals and wild bootstrap for models with clustered errors (Cameron *et al.*, 2008). We therefore propose to perform parametric bootstrap as in (Cao *et al.*, 2014) after removing the family random effect, as follows:

1. Fit the reduced model (2) and obtain BLUP of random effect δ , denoted as $\hat{\delta} = \Phi \hat{\mathbf{V}}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\alpha}})$, where $\hat{\mathbf{V}}$ and $\hat{\boldsymbol{\alpha}}$ are the maximum likelihood estimate (MLE) of $\mathbf{V}$ and $\boldsymbol{\alpha}$ in model (2), respectively.
2. Calculate the new outcome vector with BLUP removed as $\tilde{\mathbf{y}} = \mathbf{y} - \hat{\delta}$.
3. With the new outcome vector $\tilde{\mathbf{y}} = (\tilde{y}_1, \dots, \tilde{y}_n)^T$, the n samples can be treated as independent and the famLRT_{MV} test is reduced to LRT_{MV} test for independent samples for which parametric bootstrap can be carried out as:
 - 3a. Fit the null model $\tilde{y}_i = \beta_0 + \mathbf{Z}_i \beta + \varepsilon_i$ where $\varepsilon_i \sim N(0, \sigma^2)$, and the full model $\tilde{y}_i = \beta_0 + \mathbf{Z}_i \beta + G_{1i}\gamma_1 + G_{2i}\gamma_2 + \varepsilon_i$, where $\varepsilon_i \sim N(0, G_{0i}\sigma_0^2 + G_{1i}\sigma_1^2 + G_{2i}\sigma_2^2)$. Then calculate LRT_{MV} test statistic as in (Cao *et al.*, 2014).
 - 3b. Obtain MLE $\hat{\beta}_0$ and $\hat{\beta}$ of the null model in 3a and calculate residuals $e_i = \tilde{y}_i - \hat{\beta}_0 - \mathbf{Z}_i \hat{\beta}$, for $i = 1, \dots, n$.

- 3c. Permute e_i 's to generate e_i^* 's and make $\tilde{y}_i^* = \hat{\beta}_0 + \mathbf{Z}_i \hat{\beta}$, for $i = 1, \dots, n$.
- 3d. Calculate test statistic LRT_{MV}^* by replacing $\tilde{y}_i$ with $\tilde{y}_i^*$ (for $i = 1, \dots, n$) as in 3a.
- 3e. Repeat 3c and 3d for B times and the parametric bootstrap famLRT_{MV} p-value is $\sum_{b=1}^B I(LRT_{MV}^{*(b)} > LRT_{MV}) / B$.

The parametric bootstrap famLRT_V p-value can be calculated similarly by fitting the null model $\tilde{y}_i = \beta_0 + \mathbf{Z}_i \beta + G_{1i} \gamma_1 + G_{2i} \gamma_2 + \varepsilon_i$ where $\varepsilon_i \sim N(0, \sigma^2)$, to obtain LRT_V test statistic in 3a and calculating LRT_V^* in 3d. As to be shown in the simulation studies, the parametric bootstrap based famLRT_{MV} test and famLRT_V test after removing BLUP of family random effect can control the Type I error rate satisfactorily for non-normally distributed residuals. Noticeably, an attractive property of the BLUP in step 1 is that it does not require the random effect δ to be normally distributed (Robinson, 1991).

Simulation Studies

To make the simulation studies representative of real family studies, we used pedigree information of 150 randomly selected families from the FHS and two real genome wide association study (GWAS) SNPs, one with MAF of 0.44 and the other with MAF of 0.14. We also used sex and age at the first clinical visit. The data we used for simulation studies includes 1019 individuals. The largest pedigree has 20 individuals and the smallest pedigree has 3 individuals.

We first evaluated the performance of famLRTs for normally distributed quantitative traits in four different scenarios: 1. There is no association between the SNP and the simulated quantitative trait; 2. The SNP is only associated with the mean heterogeneity of the trait; 3. The SNP is only associated with the variance heterogeneity of the trait; 4. The SNP is associated with both mean and variance heterogeneity of the trait. The model $y_i = 0.5 \times \text{sex}_i + 0.05 \times \text{age}_i + G_{1i} \gamma_1 + G_{2i} \gamma_2 + \delta_i + \varepsilon_i$, for $i = 1, \dots, 1019$, was used to simulate quantitative trait y_i 's, where G_{1i} and G_{2i} are indicator variables indicating heterozygote and minor allele homozygote, respectively; δ_i is the family random effect; and ε_i is the residual. $\delta = (\delta_1, \dots, \delta_{1019})'$ was simulated from the multivariate normal distribution $N(\mathbf{0}, \Phi)$, where Φ is two times the kinship matrix. In scenario 1 and 3 without mean effect, we let $\gamma_1 = \gamma_2 = 0$. In scenario 2 and 4 with mean effect, we let $\gamma_1 = 0.15$ and $\gamma_2 = 0.35$. Similarly, in scenario 1 and 2 without variance effect, ε_i was generated randomly from $N(0, 1)$. In scenario 3 and 4 with variance effect, ε_i was generated randomly from $N(0, 1)$ for the common allele homozygote, $N(0, 1.15^2)$ for the heterozygote, $N(0, 1.4^2)$ for the minor allele homozygote. For both of the GWAS SNPs, 1000 simulated datasets were generated for each of the four scenarios. Empirical Type I error/power of famLRT_{MV}, famLRT_M, famLRT_V, parametric bootstrap based famLRT_{MV} and famLRT_V were calculated at significance levels of 0.05 and 0.01. For parametric bootstrap based tests, 1000 resamplings were conducted for each simulated dataset. LRTs ignoring familial correlation, including LRT_{MV}, LRT_M, and LRT_V, were included in scenario 1 to evaluate the impact on Type I error when family relatedness was ignored. Of note, the recently proposed GRAMMAR+ method transforms traits from related individuals to be independent and preserves high power in the subsequent

independent sample test (Belonogova *et al.*, 2013). We therefore also included LRT_{MV} , LRT_M , and LRT_V tests using GRAMMAR+ residuals for comparison.

For non-normally distributed quantitative traits, we conducted two sets of simulation studies, one with residuals generated from t -distributions representing quantitative traits with heavy tails, and the other with residuals generated from χ^2 -distributions representing skewed quantitative traits. For each set of simulation studies, we also considered two SNPs with different MAFs in four scenarios and used the same model for quantitative trait generation as we did for normally distributed traits. The only difference was residual generation. For heavy tail quantitative traits, residuals were generated randomly from a t -distribution ($df=10$) in scenario 1 and 2, and from t -distributions with $df=10$, 5, and 3 for the common allele homozygote, heterozygote, and minor allele homozygote, respectively, in scenario 3 and 4. For skewed quantitative traits, residuals were generated randomly from a χ^2 -distribution ($df=3$) in scenario 1 and 2, and from χ^2 -distributions with $df=3$, 4, and 6 for the common allele homozygote, heterozygote, and minor allele homozygote, respectively, in scenario 3 and 4. Residuals from χ^2 -distributions were scaled to have mean 0 in each genotype group and then divided by 3 to make variances comparable to variances in other simulation scenarios. We simulated 1000 replicates for each scenario. Empirical Type I error/power was calculated at significance levels of 0.05 and 0.01. For parametric bootstrap based $famLRT_{MV}$ and $famLRT_V$, 1000 resamplings were conducted for each replicate to obtain parametric bootstrap p-values.

Application to the Framingham Heart Study

The Framingham Heart Study (FHS) is a family-based study conducted for three generations: the Original Cohort, the Offspring Cohort, and the Third Generation Cohort (Splansky *et al.*, 2007). We downloaded the FHS genotype and phenotype data from NCBI dbGaP (<http://www.ncbi.nlm.nih.gov/gap>). We examined the association between mean or variance heterogeneity of BMI and SNPs near previously identified SNPs associated with BMI. The phenotype information was extracted from the fourth clinic exam for the Original Cohort, the third clinic exam for the Offspring Cohort, and the first clinic exam for the Third Generation Cohort to make the average age when phenotype information was recorded comparable across the three cohorts. BMI was calculated as weight in kilograms divided by the square of height in meters. Large-scale meta-analyses have identified SNPs associated with mean heterogeneity of BMI (Berndt *et al.*, 2013, Speliotes *et al.*, 2010, Willer *et al.*, 2009) and variance heterogeneity of BMI (Yang *et al.*, 2012). We extracted SNPs within ± 500 kilobase pairs (kb) of previously identified BMI associated SNPs from the genotype data of FHS SNP Health Association Resource (SHARe) GWAS. 6484 individuals from three generations with both phenotype and genotype information available were included in the analysis. We restricted our analysis to SNPs with at least 20 observations in each genotype group. 4640 SNPs were tested using $famLRT_M$, and parametric bootstrap based $famLRT_{MV}$ and $famLRT_V$, adjusting for sex, age, and age squared. 1,000,000 resamplings were performed for parametric bootstrap based tests.

RESULTS

Simulation Studies

Table 1 shows the simulation study results for normally distributed quantitative traits. For family data, tests without adjustment for family relatedness (LRT_{MV} , LRT_M and LRT_V) have inflated Type I errors for both SNPs with different MAFs, while tests adjusting for family relatedness, including parametric tests ($famLRT_{MV}$, $famLRT_M$ and $famLRT_V$), parametric bootstrap based tests ($famLRT_{MV}(PB)$ and $famLRT_V(PB)$) and LRTs using GRAMMAR+ residuals, control Type I errors satisfactorily at significance levels of 0.05 and 0.01 based on 1000 replications (Table 1A). To assess powers of family-based tests, we conducted simulation studies in three scenarios (Table 1B). When a quantitative trait only shows mean difference across genotype groups, $famLRT_M$ is the most powerful test; $famLRT_{MV}$ loses power due to larger degrees of freedom; $famLRT_V$ controls Type I error well. Similarly, when there is only variance difference of a quantitative trait across different genotype groups, $famLRT_V$ is the most powerful test; $famLRT_{MV}$ also loses power in this scenario; $famLRT_M$ controls Type I error well. In contrast, when a SNP is associated with both mean and variance heterogeneity of a quantitative trait, the joint test is more powerful than either of the tests for a single effect. We noticed that powers of parametric bootstrap based $famLRT_{MV}$ and $famLRT_V$ are slightly lower compared with their parametric versions, but the power loss is moderate. The two-step tests using GRAMMAR+ residuals have power comparable to the one-step $famLRT$ s in all three scenarios. In addition, with the same magnitude of simulated effects, tests for the SNP with lower MAF have lower power than tests for the SNP with higher MAF. This is as expected because sample sizes of heterozygotes and minor allele homozygotes decrease as MAF decreases.

Table 2 shows the simulation study results for non-normally distributed quantitative traits, one set simulated from t -distributions representing heavy tail traits and the other set simulated from χ^2 -distributions representing skewed traits. Parametric tests for variance, including both $famLRT_{MV}$ and $famLRT_V$, cannot control Type I errors when quantitative traits are not normally distributed, while the mean only parametric test $famLRT_M$ still performs satisfactorily. In contrast, parametric bootstrap based tests for variance can fix the problem and control Type I errors at the nominal levels for both heavy tail and skewed traits (Table 2A). We then evaluated the power of the tests that can control Type I errors in three different scenarios (Table 2B). The same as the simulation study results for normally distributed traits, the joint test is the most powerful when a SNP is associated with both trait mean and variance heterogeneity, while it sacrifices power in a mean effect or variance effect only scenario because of the extra degrees of freedom of the test. For the mean effect only scenario, parametric bootstrap $famLRT_V$ controls Type I errors well, while for the variance effect only scenario, $famLRT_M$ controls Type I errors well. The powers of all tests decrease as the MAF of the associated SNP decreases because sample sizes of heterozygotes and minor allele homozygotes decrease. We also applied the LRTs for unrelated individuals to the GRAMMAR+ residuals for simulated non-normally distributed traits. As expected, although the GRAMMAR+ transformation can effectively remove family relatedness, it does not overcome the issue of inflated Type I error for vQTL test (results not shown).

Analysis of the FHS Data

We investigated the association between BMI and BMI candidate SNPs using the family-based FHS data. BMI candidate SNPs are SNPs within $\pm 500\text{kb}$ of previously identified 50 SNPs significantly associated with either mean or variance heterogeneity of BMI by large-scale meta-analysis (Berndt *et al.*, 2013, Speliotes *et al.*, 2010, Willer *et al.*, 2009). As shown in Supplemental Figure 1, the distributions of both BMI and BMI residuals are quite positively skewed and we would expect inflated Type I errors for famLRT_{MV} and famLRT_V. As a result, we applied the proposed parametric bootstrap-based famLRT tests. The effective number of tests of the 4640 BMI candidate SNPs is 2731, corresponding to explanation of 99.5% of the total variation of 4640 SNPs, which was calculated following the method proposed by (Gao *et al.*, 2008). However, none of the tested SNPs passed the significance level of 1.8×10^{-5} after multiple test correction. The top 10 SNPs with the smallest p-values of parametric bootstrap based famLRT_{MV} are listed in Table 3. Four SNPs are in the same region near gene *CADM2*. Out of the ten SNPs, nine SNPs have p-values less than 0.05 for both the mean effect only and the variance effect only tests, showing evidence of heterogeneity in both mean and variance. For seven SNPs, the mean and variance joint test p-value is smaller than the p-values of both individual effect tests, confirming that the test for both mean and variance heterogeneity is more powerful for detecting vQTLs when a SNP is also associated with mean heterogeneity. BMI residual plots after removing covariate effects and family relatedness show observable variance heterogeneity across genotypes for the top two SNPs rs3120625 and rs7987059 (Figure 1). Furthermore, rs12328474 was identified as a top SNP associated with BMI variability by a large-scale genetic association study, with p-value of 1.2×10^{-6} in a discovery set of 104,640 samples and p-value of 0.02 in a replication set of 32,403 samples (Yang *et al.*, 2012). Our analysis confirmed the vQTL of BMI with parametric bootstrap based famLRT_V p-value of 4.2×10^{-3} using the 6484 FHS samples (Table 3). In contrast, if we performed vQTL test in 2380 unrelated FHS samples using the LRT_V test, the p-value for rs12328474 is 0.22, which is much less significant than the famLRT_V p-value based on all FHS samples. In addition, for the top ten SNPs listed in Table 3, all the tests using independent samples have less significant p-values than their corresponding family-based tests using all samples. This demonstrates the power gain enabled by our proposed family-based tests on all related individuals in a family design.

DISCUSSION

In this paper, we have proposed a joint test of mean and variance heterogeneity for family data in the LMM framework. The model is flexible in that it can accommodate a joint test and a mean or variance effect only test with adjustment for both covariates and family relatedness. To our knowledge, famLRT_{MV} is the first test developed for jointly testing mean and variance effects for family data.

Family design is common in genetic studies. We have demonstrated that the joint test for mean and variance heterogeneity has inflated Type I error when familial correlation is ignored. An alternative is to extract unrelated individuals for analysis, which will significantly reduce the total sample size and lead to power loss. For example, 6484

individuals were included in our case study of the FHS data, while only 2380 individuals out of the 6484 individuals are unrelated. We would discard approximately two thirds of the entire sample if only unrelated samples were considered, which is a big loss of valuable phenotype and genotype information that has been collected.

We showed that the $\text{famLRT}_{\text{MV}}$ test statistic of a normally distributed trait approximately follows a χ^2 -distribution, while Type I error inflation was observed for a non-normally distributed trait. Transformation of a non-normally distributed trait to be normal is often used in linear models, including Box-Cox transformation and inverse normal transformation. However, transformation is not preferred for a variance test because it may distort the variance related information in the untransformed trait, which may lead to bias and loss of power (Cao *et al.*, 2014). Resampling based methods are more appropriate for p-value calculation in this scenario. Commonly used resampling methods for LMMs, including wild bootstrap (Cameron *et al.*, 2008), case bootstrap and bootstrap for both random effects and residuals (Cameron *et al.*, 2008), are not applicable to the model we proposed. Wild bootstrap only takes into account of mean heterogeneity, but not variance heterogeneity. Case bootstrap is not appropriate because different families may have very different sizes so that the total sample size of each resampled sample using case bootstrap may vary dramatically. Due to the fact that each family has a unique correlation structure and the correlation within each family is not exchangeable, bootstrap for random effects is not applicable either. Therefore, we suggested a two-step approach for parametric bootstrap by removing the BLUP of the family random effect at the first step and performing parametric bootstrap for the uncorrelated samples in the second step. Decreased power was observed for the two-step parametric bootstrap based $\text{famLRT}_{\text{MV}}$ when compared with one-step parametric $\text{famLRT}_{\text{MV}}$ (Table 1B). The power decrease is subtle for the variance effect only and joint effect scenarios, while it is a little more for the mean effect only scenario. However, the two-step approach is necessary for Type I error control when the quantitative trait of interest is not normally distributed. In addition, removing the BLUP of the family random effect avoids decomposing the sparse block diagonal kinship matrix when performing parametric bootstrap, which is much more computationally efficient. For genome-wide pedigree-based QTL analysis, Aulchenko *et al.* (2007) also suggested a two-step approach to shorten computation time by removing familial correlation in the first step. Furthermore, for a given quantitative trait of interest, the empirical distribution of the $\text{famLRT}_{\text{MV}}$ test statistic constructed based on parametric bootstrap is invariant to the SNP being tested after removal of the BLUP of the family random effect (Supplemental Figure 2). In other words, the parametric bootstrap based $\text{famLRT}_{\text{MV}}$ only needs to be conducted for one SNP given a trait and the empirical distribution of the test statistic can then be used to calculate p-values for all the SNPs. Although the resampling based method is necessary for $\text{famLRT}_{\text{MV}}$ p-value calculation when the quantitative trait is not normally distributed, the fact that only one set of resampling suffices for all the SNPs makes the test much more practical and useful.

To conduct $\text{famLRT}_{\text{MV}}$ for a non-normally distributed quantitative trait in a large scale association analysis, e.g., GWAS, we suggest an initial screening by calculating parametric p-values for all the SNPs based on χ^2 -distribution. If there are any SNPs passing the

significance level, one set of parametric bootstrap can then be conducted to construct the empirical null distribution of the test statistic, which can be used to determine the significance of all the SNPs (Supplemental Figure 2).

In addition to famLRTs, independent sample tests using GRAMMAR+ residuals can be an option for vQTL detection using family data as shown in Table 1. GRAMMAR+ transformation is an approximation method (Belonogova *et al.*, 2013) and any tests using GRAMMAR+ residuals are two-step tests where the covariate effects are removed in the first step. Although tests using GRAMMAR+ residuals have power comparable to famLRTs shown by our simulation study (Table 1), we would expect power loss in the mean test part if the tested SNP is correlated with any covariates (Cao *et al.*, 2014). The advantage of GRAMMAR+ transformation is that GRAMMAR+ residuals can be tested using Levene's test or Fligner-Killeen test for variance heterogeneity. These two tests are robust to non-normality and are computationally efficient. As for computational time, GRAMMAR+ transformation followed by LRT_{MV} took approximately 90 minutes, while famLRT_{MV} took approximately 5 minutes for 1000 replications of the simulated data with 1019 subjects on the same computer.

As demonstrated by simulation studies and application to the FHS data, the joint test for detecting mean and variance heterogeneity is more powerful than a single effect test when a SNP is associated with both mean and variance heterogeneity of a quantitative trait. Mean heterogeneity is a more common phenomenon than variance heterogeneity in QTL analysis (Yang *et al.*, 2012). In addition, vQTLs can occur due to LD with a causal functional SNP with mean heterogeneity. In this case the vQTL will show mean heterogeneity as well (Cao *et al.*, 2014). The real data analysis showed that nine out of the ten top SNPs are associated with both mean and variance heterogeneity of BMI using nominal significance level of 0.05 (Table 3). By both simulation (Table 1B and 2B) and empirical studies, we found that the power loss of the joint test is subtle compared with the variance only test when a SNP has variance only effect. For example, rs2889756 has a joint test p-value of 4.0×10^{-4} , which is even smaller than the variance test p-value of 4.3×10^{-4} given a mean test p-value of 0.12; rs12328474 has a joint test p-value of 7.3×10^{-3} , which is only slightly larger than the variance test p-value of 4.2×10^{-3} given a mean test p-value of 0.26. Therefore, we recommend vQTL screening using joint effect test famLRT_{MV}, followed by variance effect only test famLRT_V for the significant SNPs identified by famLRT_{MV} test to increase the power of vQTL detection. The significance level of the subsequent famLRT_V test only needs to be adjusted for the number of SNPs that are globally significant based on famLRT_{MV} test. Our previous simulation study has demonstrated that a variance test after rejecting the null hypothesis of the joint test can control family-wise Type I error rate at the nominal level (Cao *et al.*, 2014). The famLRT_{MV}, famLRT_M and famLRT_V tests are implemented as R functions, which will be posted on our website at <https://sites.google.com/site/utpengwei/>

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

This work was supported by the National Institutes of Health grants R01HL105502 (to TJM), R01HL116720 and R01CA169122 (to PW). The authors declare that there are no conflicts of interest. The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University (Contract No. N01-HC-25195). This manuscript was not prepared in collaboration with investigators of the Framingham Heart Study and does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or NHLBI. Funding for SHARe Affymetrix genotyping was provided by NHLBI Contract N02-HL-64278. SHARe Illumina genotyping was provided under an agreement between Illumina and Boston University. We thank the anonymous reviewers for their constructive comments that improved the presentation of the paper.

References

- Aulchenko YS, De Koning D-J, Haley C. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics*. 2007; 177:577–585. [PubMed: 17660554]
- Balding, DJ. The advance of Bayesian methods for genetic association analysis; Presented at the 59th Annual Meeting of The American Society of Human Genetics; Honolulu, HI. Nov 22. 2009 2009
- Belonogova NM, Svishcheva GR, van Duijn CM, Aulchenko YS, Axenovich TI. Region-based association analysis of human quantitative traits in related individuals. *PLoS one*. 2013; 8(6):e65395. [PubMed: 23799013]
- Berndt SI, Gustafsson S, Maegi R, Ganna A, Wheeler E, Feitosa MF, Justice AE, Monda KL, Croteau-Chonka DC, Day FR. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nature genetics*. 2013; 45:501–512. [PubMed: 23563607]
- Cameron AC, Gelbach JB, Miller DL. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*. 2008; 90:414–427.
- Cao Y, Wei P, Bailey M, Kauwe JS, Maxwell TJ. A versatile omnibus test for detecting mean and variance heterogeneity. *Genetic Epidemiology*. 2014; 38:51–59. [PubMed: 24482837]
- Chen H, Meigs JB, Dupuis J. Sequence kernel association test for quantitative traits in family samples. *Genetic Epidemiology*. 2013; 37:196–204. [PubMed: 23280576]
- Fisher RA. The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the royal society of Edinburgh*. 1918; 52(2):399–433.
- Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genetic Epidemiology*. 2008; 32:361–9. [PubMed: 18271029]
- Jimenez-Gomez JM, Corwin JA, Joseph B, Maloof JN, Kliebenstein DJ. Genomic analysis of QTLs and genes altering natural variation in stochastic noise. *PLoS genetics*. 2011; 7:e1002295. [PubMed: 21980300]
- Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-Y, Freimer NB, Sabatti C, Eskin E. Variance component model to account for sample structure in genome-wide association studies. *Nature genetics*. 2010; 42:348–354. [PubMed: 20208533]
- Laird NM, Lange C. Family-based designs in the age of large-scale gene-association studies. *Nat Rev Genet*. 2006; 7:385–94. [PubMed: 16619052]
- Lehmann, EL. Elements of large-sample theory. Springer; New York: 1999.
- Liu CT, Monda KL, Taylor KC, Lange L, Demerath EW, Palmas W, Wojczynski MK, Ellis JC, Vitols MZ, Liu S, Papanicolaou GJ, Irvin MR, Xue L, Griffin PJ, Nalls MA, Adeyemo A, Liu J, Li G, Ruiz-Narvaez EA, Chen WM, Chen F, Henderson BE, Millikan RC, Ambrosone CB, Strom SS, Guo X, Andrews JS, Sun YV, Mosley TH, Yanek LR, Shriner D, Haritunians T, Rotter JJ, Speliotes EK, Smith M, Rosenberg L, Mychaleckyj J, Nayak U, Spruiell I, Garvey WT, Pettaway C, Nyante S, Bandera EV, Britton AF, Zonderman AB, Rasmussen-Torvik LJ, Chen YD, Ding J, Lohman K, Kritchevsky SB, Zhao W, Peyser PA, Kardina SL, Kabagambe E, Broeckel U, Chen G, Zhou J, Wassertheil-Smoller S, Neuhauser ML, Rasmussen E, Psaty B, Kooperberg C, Manson JE, Kuller LH, Ochs-Balcom HM, Johnson KC, Sucheston L, Ordovas JM, Palmer JR, Haiman CA, McKnight B, Howard BV, Becker DM, Bielak LF, Liu Y, Allison MA, Grant SF, Burke GL,

- Patel SR, Schreiner PJ, Borecki IB, Evans MK, Taylor H, Sale MM, Howard V, Carlson CS, Rotimi CN, Cushman M, Harris TB, Reiner AP, Cupples LA, North KE, Fox CS. Genome-wide association of body fat distribution in African ancestry populations suggests new loci. *PLoS genetics*. 2013; 9:e1003681. [PubMed: 23966867]
- Ouallacha K, Dastani Z, Li R, Cingolani PE, Spector TD, Hammond CJ, Richards JB, Ciampi A, Greenwood CM. Adjusted sequence kernel association test for rare variants controlling for cryptic and family relatedness. *Genetic Epidemiology*. 2013; 37:366–376. [PubMed: 23529756]
- Paré G, Cook NR, Ridker PM, Chasman DI. On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the Women's Genome Health Study. *PLoS genetics*. 2010; 6:e1000981. [PubMed: 20585554]
- Robinson GK. That BLUP is a good thing: The estimation of random effects (with discussion). *Statist. Sci.* 1991; 6
- Rönnegård L, Valdar W. Detecting major genetic loci controlling phenotypic variability in experimental crosses. *Genetics*. 2011; 188:435–447. [PubMed: 21467569]
- Shen X, Pettersson M, Rönnegård L, Carlborg Ö. Inheritance beyond plain heritability: variance-controlling genes in *Arabidopsis thaliana*. *PLoS genetics*. 2012; 8:e1002839. [PubMed: 22876191]
- Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Allen HL, Lindgren CM, Luan JA, Mägi R. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature genetics*. 2010; 42:937–948. [PubMed: 20935630]
- Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'agostino RB, Fox CS, Larson MG, Murabito JM. The third generation cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *American journal of epidemiology*. 2007; 165:1328–1335. [PubMed: 17372189]
- Struchalin MV, Dehghan A, Witteman JCM, Van Duijn C, Aulchenko YS. Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations. *BMC genetics*. 2010; 11:92. [PubMed: 20942902]
- Struchalin MV, Amin N, Eilers PH, van Duijn CM, Aulchenko YS. An R package "VariABEL" for genome-wide searching for potentially interacting loci by testing genotypic variance heterogeneity. *BMC genetics*. 2012; 13:4. [PubMed: 22272569]
- Sun W. A statistical framework for eQTL mapping using RNA-seq data. *Biometrics*. 2012; 68:1–11. [PubMed: 21838806]
- Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, Berndt SI, Elliott AL, Jackson AU, Lamina C. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nature genetics*. 2009; 41:25–34. [PubMed: 19079261]
- Yang J, Loos RJJ, Powell JE, Medland SE, Speliotes EK, Chasman DI, Rose LM, Thorleifsson G, Steinthorsdottir V, Mägi R. FTO genotype is associated with phenotypic variability of body mass index. *Nature*. 2012; 490:267–272. [PubMed: 22982992]
- Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature genetics*. 2006; 38:203–208. [PubMed: 16380716]

Figure 1.

BMI residual box plots and density plots by genotype for rs3120625 and rs7987059. AA, Aa, and aa represent major allele homozygote, heterozygote, and minor allele homozygote, respectively. Residuals were obtained after removing covariate effects and family random effect. rs3120625 is the most significant SNP for mean and variance joint effect. rs7987059 is the most significant SNP for variance only effect.

Table 1A

Empirical Type I error/power of association tests for normally distributed traits at significance levels of 0.05 and 0.01. PB; parametric bootstrap. A. Type I error comparison of family-based tests and tests ignoring family relatedness.

SNP MAF	α level	famLRT						GRAMMAR+						LRT		
		famLRT _{MV}			famLRT _{MV} (PB)			famLRT _V			famLRT _V (PB)			LRT _{MV}		
		famLRT _{MV}	famLRT _{MV} (PB)	famLRT _M	famLRT _M	famLRT _V	famLRT _V (PB)	LRT _{MV}	LRT _M	LRT _V	LRT _M	LRT _M	LRT _V	LRT _{MV}	LRT _M	LRT _V
0.44	0.05	0.054	0.049	0.052	0.053	0.046	0.046	0.060	0.061	0.054	0.078	0.079	0.062	0.078	0.079	0.062
	0.01	0.012	0.010	0.009	0.012	0.008	0.008	0.012	0.014	0.011	0.023	0.022	0.018	0.023	0.022	0.018
0.14	0.05	0.048	0.044	0.056	0.055	0.046	0.046	0.057	0.056	0.061	0.074	0.075	0.064	0.074	0.075	0.064
	0.01	0.012	0.009	0.012	0.011	0.008	0.008	0.014	0.011	0.012	0.015	0.018	0.016	0.015	0.018	0.016

Table 1B

Power comparison of famLRTs and LRTs after GRAMMAR+ transformation in three different scenarios: mean effect, variance effect, mean and variance joint effect.

Effect	SNP MAF	α level	famLRT				GRAMMAR+			
			famLRT _{MV}	famLRT _{MV(PB)}	famLRT _M	famLRT _V	famLRT _{V(PB)}	LRT _{MV}	LRT _M	LRT _V
Mean	0.44	0.05	0.573	0.521	0.692	0.048	0.042	0.559	0.687	0.041
		0.01	0.327	0.285	0.446	0.012	0.008	0.297	0.424	0.010
	0.14	0.05	0.266	0.236	0.334	0.053	0.050	0.263	0.340	0.062
		0.01	0.098	0.077	0.155	0.009	0.011	0.103	0.150	0.012
Variance	0.44	0.05	0.759	0.735	0.039	0.846	0.835	0.711	0.062	0.816
		0.01	0.555	0.508	0.007	0.663	0.634	0.515	0.010	0.618
	0.14	0.05	0.336	0.315	0.056	0.373	0.365	0.298	0.055	0.358
		0.01	0.147	0.144	0.012	0.188	0.175	0.129	0.009	0.161
Mean & Variance	0.44	0.05	0.919	0.915	0.580	0.837	0.819	0.918	0.578	0.822
		0.01	0.816	0.781	0.339	0.647	0.621	0.791	0.343	0.626
	0.14	0.05	0.533	0.512	0.282	0.376	0.362	0.496	0.289	0.360
		0.01	0.317	0.292	0.118	0.177	0.160	0.292	0.126	0.160

Table 2A

Empirical Type I error/power of association tests for t-distributed and χ^2 -distributed quantitative traits at significance levels of 0.05 and 0.01. PB: parametric bootstrap. A. Type I error comparison.

Quantitative Traits	SNP MAF	α level	famLRT _{MV}	famLRT _{MV} (PB)	famLRT _M	famLRT _V	famLRT _V (PB)
t-distributed	0.44	0.05	0.083	0.045	0.055	0.090	0.044
		0.01	0.022	0.008	0.012	0.031	0.010
	0.14	0.05	0.079	0.045	0.056	0.093	0.048
		0.01	0.035	0.012	0.012	0.037	0.012
	0.04	0.05	0.079	0.045	0.056	0.093	0.048
		0.01	0.035	0.012	0.012	0.037	0.012
χ^2 -distributed	0.44	0.05	0.111	0.049	0.055	0.125	0.055
		0.01	0.036	0.010	0.011	0.044	0.009
	0.14	0.05	0.104	0.056	0.056	0.105	0.053
		0.01	0.030	0.011	0.013	0.040	0.013
	0.04	0.05	0.104	0.056	0.056	0.105	0.053
		0.01	0.030	0.011	0.013	0.040	0.013

Table 2B

Power comparison in three different scenarios: mean effect, variance effect, mean and variance joint effect.

Effect	SNP MAF	α level	<i>t</i> -distributed			χ^2 -distributed		
			famLRT _{MV} (PB)	famLRT _M	famLRT _V (PB)	famLRT _{MV} (PB)	famLRT _M	famLRT _V (PB)
Mean	0.44	0.05	0.460	0.633	0.055	0.537	0.770	0.056
		0.01	0.222	0.379	0.009	0.253	0.563	0.013
	0.14	0.05	0.190	0.311	0.048	0.236	0.412	0.047
		0.01	0.070	0.142	0.013	0.078	0.197	0.012
	0.04	0.05	0.100	0.150	0.005	0.080	0.120	0.005
		0.01	0.020	0.030	0.001	0.010	0.020	0.001
Variance	0.44	0.05	0.619	0.051	0.657	0.449	0.043	0.584
		0.01	0.408	0.010	0.433	0.218	0.014	0.363
	0.14	0.05	0.380	0.057	0.398	0.173	0.046	0.256
		0.01	0.184	0.013	0.202	0.059	0.011	0.103
	0.04	0.05	0.100	0.010	0.100	0.050	0.005	0.050
		0.01	0.020	0.001	0.020	0.010	0.001	0.010
Mean & Variance	0.44	0.05	0.799	0.513	0.675	0.823	0.696	0.577
		0.01	0.603	0.281	0.459	0.557	0.373	0.335
	0.14	0.05	0.504	0.275	0.384	0.356	0.273	0.244
		0.01	0.289	0.117	0.220	0.138	0.059	0.106
	0.04	0.05	0.100	0.010	0.100	0.050	0.005	0.050
		0.01	0.020	0.001	0.020	0.010	0.001	0.010

Table 3

Summary of Framingham Heart Study BMI association analysis results.

SNP	Chromosome	Position	Nearest Gene	MAF	p-value using all samples			p-value using independent samples		
					famLRT _{MV} (PB)	famLRT _M	famLRT _V (PB)	LRT _{MV} (PB)	LRT _M	LRT _V (PB)
rs3120625	1	109768889	<i>SARS</i>	0.344	7.7×10^{-5}	3.8×10^{-5}	3.7×10^{-4}	0.37	0.12	0.55
rs7987059	13	27541625	<i>USP12</i>	0.181	9.0×10^{-5}	1.9×10^{-2}	8.4×10^{-5}	0.14	0.24	0.14
rs9310004	3	86302812	<i>CADM2</i>	0.121	2.2×10^{-4}	5.3×10^{-3}	3.7×10^{-4}	0.12	0.33	0.11
rs7592497	2	659958	<i>TMEM18</i>	0.284	2.7×10^{-4}	2.2×10^{-3}	4.7×10^{-4}	2.1×10^{-3}	1.1×10^{-2}	3.6×10^{-3}
rs7621130	3	86344362	<i>CADM2</i>	0.120	3.0×10^{-4}	3.7×10^{-3}	6.1×10^{-4}	9.9×10^{-2}	0.32	8.4×10^{-2}
rs352973	2	142582263	<i>LRP1B</i>	0.241	3.2×10^{-4}	9.2×10^{-3}	2.1×10^{-4}	3.6×10^{-2}	9.5×10^{-2}	3.1×10^{-2}
rs7653885	3	86344304	<i>CADM2</i>	0.120	3.4×10^{-4}	4.1×10^{-3}	6.8×10^{-4}	0.13	0.37	0.11
rs2889756	3	186252771	<i>CRYGS</i>	0.076	4.0×10^{-4}	0.12	4.3×10^{-4}	2.1×10^{-2}	0.18	1.8×10^{-2}
rs9675886	18	57969322	<i>MC4R</i>	0.265	4.2×10^{-4}	1.3×10^{-3}	1.1×10^{-3}	8.8×10^{-2}	0.26	7.6×10^{-2}
rs6781294	3	86331094	<i>CADM2</i>	0.120	5.1×10^{-4}	5.7×10^{-3}	9.0×10^{-4}	0.13	0.37	0.11
rs12328474*	2	140922100	<i>LRP1B</i>	0.249	7.3×10^{-3}	0.26	4.2×10^{-3}	0.26	0.45	0.22

This table lists the 10 SNPs with smallest p-values of parametric bootstrap based famLRT_{MV}, adjusting for sex, age, and age squared.

* denotes the SNP that was identified as a top SNP associated with BMI variability [Yang et al., 2012]