Received: 19 May 2016 | Revised: 1 September 2016 | Accepted: 19 September 2016 | Published on: 5 December 2016

DOI: 10.1002/gepi.22014

RESEARCH ARTICLE

WILEY Genetic Epidemiology

OFFICIAL JOURNAL INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY www.geneticepi.org

A comparison study of multivariate fixed models and Gene Association with Multiple Traits (GAMuT) for next-generation sequencing

Chi-Yang Chiu[1] | Jeesun Jung[2] | Yifan Wang[3] | Daniel E. Weeks[4] | Alexander F. Wilson[5] | Joan E. Bailey-Wilson[5] | Christopher I. Amos[7] | James L. Mills[6] | Michael Boehnke[8] | Momiao Xiong[9] | Ruzong Fan[1]

[1]Biostatistics and Bioinformatics Branch, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health (NIH), Bethesda, MD, USA

[2]Laboratory of Epidemiology and Biometry, National Institute on Alcohol, Abuse and Alcoholism, NIH, Bethesda, MD, USA

[3]Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

[4]Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA, USA

[5]Computational and Statistical Genomics Branch, National Human Genome Research Institute, NIH, Bethesda, MD, USA

[6]Epidemiology Branch, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health (NIH), Bethesda, MD, USA

[7]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Lebanon, NH, USA

[8]Department of Biostatistics, School of Public Health, The University of Michigan, Ann Arbor, MI, USA

[9]Human Genetics Center, University of Texas—Houston, Houston, TX, USA

Correspondence
Ruzong Fan, Department of Biostatistics, Bioinformatics, and Biomathematics, 4000 Reservoir Road NW, Building D-180, Georgetown University Medical Center, Washington, DC 20057.
Email: rf740@georgetown.edu

ABSTRACT

In this paper, extensive simulations are performed to compare two statistical methods to analyze multiple correlated quantitative phenotypes: (1) approximate F-distributed tests of multivariate functional linear models (MFLM) and additive models of multivariate analysis of variance (MANOVA), and (2) Gene Association with Multiple Traits (GAMuT) for association testing of high-dimensional genotype data. It is shown that approximate F-distributed tests of MFLM and MANOVA have higher power and are more appropriate for major gene association analysis (i.e., scenarios in which some genetic variants have relatively large effects on the phenotypes); GAMuT has higher power and is more appropriate for analyzing polygenic effects (i.e., effects from a large number of genetic variants each of which contributes a small amount to the phenotypes). MFLM and MANOVA are very flexible and can be used to perform association analysis for (i) rare variants, (ii) common variants, and (iii) a combination of rare and common variants. Although GAMuT was designed to analyze rare variants, it can be applied to analyze a combination of rare and common variants and it performs well when (1) the number of genetic variants is large and (2) each variant contributes a small amount to the phenotypes (i.e., polygenes). MFLM and MANOVA are fixed effect models that perform well for major gene association analysis. GAMuT can be viewed as an extension of sequence kernel association tests (SKAT). Both GAMuT and SKAT are more appropriate for analyzing polygenic effects and they perform well not only in the rare variant case, but also in the case of a combination of rare and common variants. Data analyses of European cohorts and the Trinity Students Study are presented to compare the performance of the two methods.

KEYWORDS

association mapping, common variants, complex traits, functional data analysis, multivariate analysis of variance (MANOVA), multivariate functional linear models (MFLM), quantitative trait loci, rare variants

1 | INTRODUCTION

Since multi-phenotype analysis can increase power to dissect complex disorders, analysis of pleiotropic traits has become a very important topic. One method to analyze pleiotropic traits is to analyze a single polymorphism at a time to evaluate the effect of common variants as is routinely done in genome-wide association studies (GWAS) or exome studies (Allison et al., 1998; Chavali et al., 2010; Ferreira & Purcell, 2009; Galesloot et al., 2014; Huang et al., 2011; O'Reilly et al., 2012; Ried et al., 2012; Sivakumaran et al., 2011; Solovieff et al., 2013). In recent years, next-generation sequencing technologies have provided rich resources to search for causal genetic variants. Researchers are facing

ever-increasing amounts of data and the need to analyze such data efficiently to enable novel discoveries (Ansorge, 2009; Mardis, 2008; Metzker, 2010; Rusk & Kiermer, 2008; Shendure & Ji, 2008). There are increasing interest in developing gene-based methods to analyze next-generation sequencing data of pleiotropic traits (Broadaway et al., 2016; Maity et al., 2012; Vsevolozhskaya et al., 2016; Wang et al., 2015). The gene-based methods have several advantages such as combining multiple variants for a unified analysis, thereby increasing power, and reducing the number of multiple comparisons. In practice, the advantages of different methods are not always clear. In this article, we aim at evaluating the performance of two gene-based procedures described below to understand the pros and cons of each procedure.

In Wang et al. (2015), multivariate functional linear models (MFLM) were proposed to perform gene-based analysis of pleiotropic traits. The MFLM are very flexible and can be used to analyze rare variants or common variants or a combination of the two. Here the rare variants' minor allele frequencies (MAF) are less than $0.01 \sim 0.05$. Broadaway et al. (2016) proposed a method of Gene Association with Multiple Traits (GAMuT) for association testing of phenotypes with high-dimensional rare variant data. By using simulated data of 30 kb regions using COSI (Schaffner et al., 2005), the authors compared power levels of GAMuT and approximate F-distributed tests of MFLM, and found that GAMuT had higher power than the approximate F-distributed tests of MFLM for six and ten correlated quantitative phenotypes. In addition, Broadaway et al. (2016) analyzed four phenotypic measures of cardiovascular health using data from the Genetic Epidemiology Network of Arteriopathy (GENOA) (Daniels et al., 2004), and found that MFLM inflates P-values. An interesting question is: why and how this happens?

The data analyzed in Broadaway et al. (2016) included 48,712 rare genetic variants (MAF $< 3\%$) that fell within 3,277 genes. Hence, each gene region has about 15 rare variants in the data analysis. Note that MFLM are designed to analyze high-dimensional next-generation sequencing data of multiple quantitative traits (Wang et al., 2015). For a gene region with about 15 rare variants, the number of parameters of MFLM is about 60 for four phenotypes if one uses B-spline basis functions suggested by Wang et al. (2015). Therefore, the number of parameters is much larger than the number of rare variants in the data analysis making it almost impossible for MFLM to perform well. If there is only a small number of variants in a gene region, it would be possible to use linear regressions to perform model selection to pick up the important variants, and then one may be able to get a final optimal model to analyze the data. In that case, neither MFLM nor GAMuT is necessary because they are mainly for large number variant analysis.

In the simulation studies of Wang et al. (2015), genetic variants located in 3 kb regions were simulated using the package COSI (Schaffner et al., 2005). In the simulations of rare variants (defined as MAF $< 3\%$), the 3 kb regions contain a mean of 53 variants. In the case that some variants are common and the rest are rare, the 3 kb regions contain a mean of 59 variants and about 10% are common. If the simulated data used in Broadaway et al. (2016) are similar, the 30 kb regions would contain more than 500 rare variants (and each causal variant contributes a small amount to the traits). Hence, the simulation studies of Broadaway et al. (2016) were based on high-dimensional genotype data. In the Supplementary Information, Broadaway et al. (2016) presented a power comparison using genetic variants located in 3 kb regions for three phenotypes and found that GAMuT performed similarly to MFLM when genetic effect sizes are relatively large.

Some interesting questions and issues stand out: how do the two methods of GAMuT and MFLM perform for more simulation scenarios? When does the GAMuT perform better and when do the fixed models including MFLM perform better and why? MFLM are very flexible and can be used to perform association analysis for (i) rare variants, (ii) common variants, and (iii) a combination of rare and common variants. Can GAMuT be used to analyze a combination of rare and common variants (or just common variants), although it was designed to analyze rare variants only? Here we perform extensive simulations to evaluate the performance of the approximate F-distributed tests of fixed effect models and GAMuT for quantitative traits by using genetic variants located in 3–30 kb regions of simulated COSI data. Data analyses of European cohorts and Trinity Students Study (TSS) are presented to compare the performance of the two methods.

2 | MODELS

In gene-based association analysis, the research goal is to model the association between multiple genetic variants and phenotypic traits. In this section, we briefly introduce the two procedures (i.e., GAMuT and MFLM) for gene-based analysis of pleiotropic traits.

2.1 | Gene Association with Multiple Traits

GAMuT utilizes a kernel distance covariance to build a nonparametric test of independence between multiple phenotypes and multiple genetic variants, and can be viewed as an extension of sequence kernel association tests (SKAT) (Ionita-Laza et al., 2013; Lee et al., 2012; Wu et al., 2011). GAMuT can analyze both quantitative and categorical phenotypes adjusting for covariates. The kernel distance covariance framework used by GAMuT assesses if pairwise phenotypic similarity is independent of pairwise rare-variant genotypic similarity. The phenotypic similarity and genotypic similarity can be formulated as matrices using a projection or a weighted linear kernel function. An MAF-weighted linear kernel is recommended for the genotypic similarity (Broadaway et al., 2016).

2.2 | Multivariate fixed effect models

Consider n individuals who are sequenced in a genomic region that has m variants. We assume that the m variants are located in a region with ordered physical positions $0 \le t_1 < \cdots < t_m = T$. To make the notation simpler, we normalize the region $[t_1, T]$ to be [0, 1]. For the ith individual, let $X_i = (x_i(t_1), \ldots, x_i(t_m))'$ denote her/his genotypes at the m variants and $Z_i = (z_{i1}, \ldots, z_{ic})'$ denote her/his covariates. Hereafter, $'$ denotes the transpose of a vector or matrix. For genotypes, we assume that $x_i(t_j)(= 0, 1, 2)$ is the number of minor alleles of the individual at the jth variant located at the position t_j. For each individual, we assume that there are L quantitative traits, $L \ge 1$. We assume that the quantitative traits are normally distributed. For the ith individual, let $y_{i\ell}$ ($\ell = 1, 2, \ldots, L$) denote her/his quantitative traits, respectively.

2.2.1 | Traditional additive effect models of MANOVA

To model the relationship between the quantitative traits and the m variants, one may use the following additive effect models of multivariate analysis of variance (MANOVA)

$$y_{i\ell} = \alpha_{\ell 0} + Z_i'\alpha_\ell + \sum_{j=1}^{m} x_i(t_j)\beta_{\ell j} + \varepsilon_{i\ell}, \ell = 1, 2, \ldots, L, \quad (1)$$

where $\alpha_{\ell 0}$ is the overall mean, $\alpha_\ell = (\alpha_{\ell 1}, \ldots, \alpha_{\ell c})'$ is a $c \times 1$ column vector of regression coefficients of covariates, $\beta_{\ell j}$ is the effect of genetic variant $x_i(t_j)$, and $\varepsilon_{i\ell}$ is an error term. For each i, the error vector $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iL})'$ is normally distributed with a mean vector of zeros and a $L \times L$ variance-covariance matrix Σ. Moreover, $\varepsilon_1, \ldots, \varepsilon_n$ are assumed to be independent. When the number of genetic variants is large, the number of parameters in the model (1) can be large, which may lead to low power. Before fitting the model (1), the QR decomposition can be applied to the genotype data to remove the redundancy, i.e., to decompose the genotype matrix into the product of an orthogonal matrix Q and a triangular matrix R via Gram–Schmidt process. Since dense variants in a region can be highly correlated to each other, the QR decomposition could significantly reduce the dimensionality and could be useful in data analysis.

2.2.2 | General MFLM

In this subsection, we introduce general MFLM to connect genetic variants to the traits (Fan et al., 2013, 2014, 2015, 2016a,b,c; Ramsay & Silverman, 2005; Wang et al., 2015). We view the ith individual's genotype data as a genetic variant function (GVF) as $X_i(t), t \in [0, 1]$. We assume that the GVF $X_i(t)$ is continuous, but this assumption can be removed as in the beta-smooth models (6).

Note that the sample includes n discrete realizations or observations $X_i = (x_i(t_1), \ldots, x_i(t_m))'$ of the human genome. By using the genetic variant information X_i, we may estimate the related GVF $X_i(t)$. To relate the GVF to the quantitative traits adjusting for covariates, we consider the following MFLM

$$y_{i\ell} = \alpha_{\ell 0} + Z_i'\alpha_\ell + \int_0^1 X_i(t)\beta_\ell(t)dt + \varepsilon_{i\ell}, \quad \ell = 1, 2, \ldots, L, \quad (2)$$

where $\beta_\ell(t)$ is the genetic effect of GVF $X_i(t)$ at the position t, and the other terms are similar to those in the MANOVA model (1).

Estimation of genetic variant functions. To estimate the GVF $X_i(t)$ from the genotypes X_i, we use an ordinary linear square smoother (Fan et al., 2013, 2014; Wang et al., 2015). The ordinary linear square smoother method assumes that the GVF is smooth. Let $\phi_k(t), k = 1, \ldots, K$, be a series of K basis functions, such as the B-spline basis and Fourier basis functions. Denote $\phi(t) = (\phi_1(t), \ldots, \phi_K(t))'$. Let Φ denote the m by K matrix containing the values $\phi_k(t_j)$, where $j \in 1, \ldots, m$. Using the discrete realizations $X_i = (x_i(t_1), \ldots, x_i(t_m))'$, we may estimate the GVF $X_i(t)$ using an ordinary linear square smoother as follows (Ramsay & Silverman, 2005)

$$\hat{X}_i(t) = (x_i(t_1), \ldots, x_i(t_m))\Phi[\Phi'\Phi]^{-1}\phi(t). \quad (3)$$

We consider two types of basis functions: (1) the B-spline basis: $\phi_k(t) = B_k(t), k = 1, \ldots, K$; and (2) the Fourier basis: $\phi_1(t) = 1, \phi_{2r+1}(t) = \sin(2\pi rt)$, and $\phi_{2r}(t) = \cos(2\pi rt), r = 1, \ldots, (K-1)/2$. Here for the Fourier basis, K is taken as a positive odd integer (de Boor, 2001; Ferraty & Romain, 2010; Horváth & Kokoszka, 2012; Ramsay, Hooker, & Graves, 2009; Ramsay & Silverman, 2005).

Revised functional regression models. The genetic effect functions $\beta_\ell(t)$ are assumed to be continuous/smooth. One may expand them by B-spline or Fourier basis functions. Formally, let $\psi_k(t), k = 1, \ldots, K_\beta$, be a series of K_β basis functions. We expand the genetic effect function $\beta_\ell(t)$ by $\psi(t) = (\psi_1(t), \ldots, \psi_{K_\beta}(t))'$ as

$$\beta_\ell(t) = (\psi_1(t), \ldots, \psi_{K_\beta}(t))(\beta_{\ell 1}, \ldots, \beta_{\ell K_\beta})' = \psi(t)'\beta_\ell, \quad (4)$$

where $\beta_\ell = (\beta_{\ell 1}, \ldots, \beta_{\ell K_\beta})'$ is a vector of coefficients $\beta_{\ell 1}, \ldots, \beta_{\ell K_\beta}$. Replacing $X_i(t)$ in MFLM (2) by $\hat{X}_i(t)$ in (3) and $\beta_\ell(t)$ by the expansion (4), we have the following revised MFLM

$$\begin{aligned} y_{i\ell} &= \alpha_{\ell 0} + Z_i'\alpha_\ell \\ &+ \left[(x_i(t_1), \ldots, x_i(t_m))\Phi[\Phi'\Phi]^{-1}\int_0^1 \phi(t)\,\psi'(t)dt\right]\beta_\ell + \varepsilon_{i\ell} \\ &= \alpha_{\ell 0} + Z_i'\alpha_\ell + W_i'\beta_\ell + \varepsilon_{i\ell}, \end{aligned} \quad (5)$$

where $W_i' = (x_i(t_1), \ldots, x_i(t_m))\Phi[\Phi'\Phi]^{-1}\int_0^1 \phi(t)\psi'(t)dt$.

2.2.3 | MFLM: beta-smooth only approach

We now introduce a simplified version of our MFLM, i.e., beta-smooth only model (de Boor, 2001; Fan et al., 2013, 2014; Ferraty & Romain, 2010; Horváth & Kokoszka, 2012; Ramsay et al., 2009; Ramsay & Silverman, 2005; Wang et al., 2015). The beta-smooth only MFLM were developed to define the relationship between the ℓth quantitative trait and the m variants (Wang et al., 2015)

$$y_{i\ell} = \alpha_{\ell 0} + Z_i'\alpha_\ell + \sum_{j=1}^{m} x_i(t_j)\beta_\ell(t_j) + \varepsilon_{i\ell}, \quad \ell = 1, 2, \ldots, L, \tag{6}$$

where $\beta_\ell(t_j)$ is the genetic effect at the physical position t_j, and the other terms are similar to those in the model (1). As for the general MFLM (2), the genetic effect $\beta_\ell(t)$ are expanded by a series of basis functions by relations (4). Replacing $\beta_\ell(t_j)$ by the expansion, the models (6) can be revised as

$$\begin{aligned} y_{i\ell} &= \alpha_{\ell 0} + Z_i'\alpha_\ell + \left[\sum_{j=1}^{m} x_i(t_j)(\psi_1(t_j), \ldots, \psi_{K_\beta}(t_j))\right]\beta_\ell + \varepsilon_{i\ell} \\ &= \alpha_{\ell 0} + Z_i'\alpha_\ell + W_i'\beta_\ell + \varepsilon_{i\ell}, \end{aligned} \tag{7}$$

where $W_i' = \sum_{j=1}^{m} x_i(t_j)(\psi_1(t_j), \ldots, \psi_{K_\beta}(t_j))$. In the model (6) and its revised version (7), we use the raw genotype data $X_i = (x_i(t_1), \ldots, x_i(t_m))'$. The genetic effect functions $\beta_\ell(t)$ are assumed to be smooth. Thus, the models are called beta-smooth only. In our previous work, we showed that beta-smooth only models perform similarly to the general MFLM in real data analysis and simulation studies (Fan et al., 2013, 2014, 2015, 2016a,b,c; Wang et al., 2015).

2.2.4 | Null hypotheses and test statistics

Consider the additive effect model of MANOVA (1) and the revised MFLM (5) and (7). To test for association between the m genetic variants and the quantitative traits as a group, the null hypothesis is $H_0 : \beta_\ell = (\beta_{\ell 1}, \ldots, \beta_{\ell m})' = 0, \ell = 1, \ldots, L$, for model (1) and $H_0 : \beta_\ell = (\beta_{\ell 1}, \ldots, \beta_{\ell K_\beta})' = 0, \ell = 1, \ldots, L$, for models (5) and (7). We may test the null $H_0 : \beta_1 = \cdots = \beta_L = 0$ by approximate F-distributed tests based on Pillai-Bartlett trace, Hotelling-Lawley trace, and Wilks's Lambda using standard statistical approaches (Anderson, 1984; Rao, 1973).

2.2.5 | Functional data analysis parameters

In the data analysis and simulations, we used the functional data analysis procedure in the statistical package R. We used two functions from the functional data analysis (fda) R package as follows to create the bases:

```
basis = create.bspline.basis(norder =
order, nbasis = bbasis)
basis = create.fourier.basis(c(0,1), nbasis =
fbasis)
```

The three parameters were taken as *order = 4, bbasis = 15, fbasis = 21* for quantitative traits in all simulations. Specifically, the order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 15$, the number of Fourier basis functions was $K = K_\beta = 21$. To make sure that the results are valid and stable, we tried a wide range of parameters that $10 \le K = K_\beta \le 21$ and the results are very close to each other (data not shown).

3 | SIMULATION STUDIES

We utilize two fixed models: (a) MFLM and (b) additive models (1) of MANOVA. Simulations were performed to evaluate the performance of the fixed models and GAMuT with sample sizes 500, 1,000, and 1,500. We used the European ancestry simulated sequence data (Lee et al., 2012; Wu et al., 2011). The sequence data are from 10,000 simulated chromosomes covering a 1 Mb region simulated using the calibrated coalescent model programmed in COSI (Schaffner et al., 2005). The generated European haplotypes mimic CEPH Utah individuals with ancestry from northern and western Europe in terms of site frequency spectrum and linkage disequilibrium (LD) pattern.

Type I error simulations. To evaluate whether the approximate F-distributed tests control false-positive rates accurately, we consider either three or six correlated phenotypes for each individual. For the three phenotype case, we generated three correlated quantitative traits using the model

$$\begin{aligned} y_{i1} &= 0.5z_{i1} + 0.5z_{i2} + \varepsilon_{i1}, \\ y_{i2} &= 0.3z_{i1} + 0.7z_{i2} + \varepsilon_{i2}, \\ y_{i3} &= 0.6z_{i1} + 0.4z_{i2} + \varepsilon_{i3}, \end{aligned} \tag{8}$$

where z_{i1} is a continuous covariate from a standard normal distribution $N(0, 1)$, z_{i2} is a dichotomous covariate taking values 0 and 1 with a probability of 0.5, and $(\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3})'$ follows a normal distribution with a mean vector of 0 and a 3×3 variance-covariance matrix $\Sigma = \begin{pmatrix} 1.00 & 0.60 & -0.35 \\ 0.60 & 1.00 & -0.45 \\ -0.35 & -0.45 & 1.00 \end{pmatrix}$. The 3×3 variance-covariance matrix Σ is taken from an empirical analysis of three traits from The TSS (Wang et al., 2015).

For the six phenotype case, we use the same strategy of Broadaway et al. (2016) to generate the correlation matrix Σ. That is, we consider scenarios of low residual correlation among phenotypes [pairwise correlation among phenotypes selected from a uniform (0, 0.3) distribution], moderate residual correlation [pairwise correlation selected from a uniform (0.3, 0.5) distribution], and high residual correlation [pairwise correlation selected from a uniform (0.5, 0.7) distribution].

TABLE 1 Empirical type I error rates of the approximate F-distribution tests based on Pillai-Bartlett trace of six traits and moderate correlation, when the variants are either rare or common

Region size	Sample size	Nominal Level α	Basis of both GVF and $\beta_\ell(t)$		Basis of beta-smooth only		MANOVA Model (1)
			B-sp basis	Fourier basis	B-sp basis	Fourier basis	
6 kb	500	0.001	0.000896	0.000986	0.000894	0.000987	0.000942
		0.0001	0.000082	0.000090	0.000082	0.000090	0.000087
	1,000	0.001	0.000994	0.001006	0.000994	0.001006	0.000957
		0.0001	0.000103	0.000100	0.000103	0.000100	0.000094
	1,500	0.001	0.001035	0.000974	0.001034	0.000974	0.000974
		0.0001	0.000093	0.000097	0.000093	0.000097	0.000098
9 kb	500	0.001	0.000910	0.000897	0.000910	0.000897	0.000887
		0.0001	0.000089	0.000081	0.000089	0.000081	0.000105
	1,000	0.001	0.000995	0.000976	0.000995	0.000976	0.000934
		0.0001	0.000094	0.000113	0.000094	0.000113	0.000091
	1,500	0.001	0.000969	0.000996	0.000969	0.000996	0.000947
		0.0001	0.000098	0.000085	0.000098	0.000085	0.000088
12 kb	500	0.001	0.000907	0.000944	0.000907	0.000944	0.000881
		0.0001	0.000095	0.000096	0.000095	0.000096	0.000090
	1,000	0.001	0.000930	0.000954	0.000930	0.000954	0.000928
		0.0001	0.000083	0.000088	0.000083	0.000088	0.000101
	1,500	0.001	0.001012	0.000948	0.001012	0.000948	0.000989
		0.0001	0.000088	0.000092	0.000088	0.000092	0.000115
15 kb	500	0.001	0.000931	0.000953	0.000931	0.000953	0.000997
		0.0001	0.000086	0.000102	0.000086	0.000102	0.000094
	1,000	0.001	0.000976	0.000958	0.000976	0.000958	0.000955
		0.0001	0.000115	0.000088	0.000115	0.000088	0.000102
	1,500	0.001	0.000955	0.000889	0.000955	0.000889	0.001003
		0.0001	0.000111	0.000100	0.000111	0.000100	0.000106
18 kb	500	0.001	0.000870	0.000943	0.000870	0.000943	0.000938
		0.0001	0.000076	0.000081	0.000076	0.000081	0.000098
	1,000	0.001	0.000958	0.001013	0.000958	0.001013	0.000966
		0.0001	0.000099	0.000113	0.000099	0.000113	0.000099
	1,500	0.001	0.000937	0.000956	0.000937	0.000956	0.000925
		0.0001	0.000077	0.000089	0.000077	0.000089	0.000083
21 kb	500	0.001	0.000923	0.000917	0.000923	0.000917	0.000893
		0.0001	0.000089	0.000065	0.000089	0.000065	0.000088
	1,000	0.001	0.000945	0.000961	0.000945	0.000961	0.000969
		0.0001	0.000073	0.000089	0.000073	0.000089	0.000102
	1,500	0.001	0.000947	0.000986	0.000947	0.000986	0.000980
		0.0001	0.000093	0.000100	0.000093	0.000100	0.000101
24 kb	500	0.001	0.000939	0.000919	0.000939	0.000919	0.000914
		0.0001	0.000093	0.000095	0.000093	0.000095	0.000093
	1,000	0.001	0.000984	0.000959	0.000984	0.000959	0.000961
		0.0001	0.000104	0.000091	0.000104	0.000091	0.000100
	1,500	0.001	0.001003	0.000931	0.001003	0.000931	0.001003
		0.0001	0.000086	0.000089	0.000086	0.000089	0.000104
27 kb	500	0.001	0.000979	0.001003	0.000979	0.001003	0.000925
		0.0001	0.000091	0.000081	0.000091	0.000081	0.000091
	1,000	0.001	0.000919	0.000966	0.000919	0.000966	0.000956

(continues)

TABLE 1 *(Continued)*

Region size	Sample size	Nominal Level α	Basis of both GVF and $\beta_\ell(t)$		Basis of beta-smooth only		MANOVA Model (1)
			B-sp basis	Fourier basis	B-sp basis	Fourier basis	
		0.0001	0.000088	0.000101	0.000088	0.000101	0.000114
	1,500	0.001	0.000933	0.000922	0.000933	0.000922	0.000976
		0.0001	0.000085	0.000089	0.000085	0.000089	0.000079
30 kb	500	0.001	0.000981	0.001031	0.000981	0.001031	0.000895
		0.0001	0.000102	0.000104	0.000102	0.000104	0.000093
	1,000	0.001	0.000979	0.000972	0.000979	0.000972	0.001001
		0.0001	0.000092	0.000093	0.000092	0.000093	0.000087
	1,500	0.001	0.000966	0.000969	0.000966	0.000969	0.000971
		0.0001	0.000096	0.000097	0.000096	0.000097	0.000102

The results of "Basis of both GVF and $\beta_\ell(t)$" were based on smoothing both GVF and genetic effect functions $\beta_\ell(t)$ of model (5), the results of "Basis of beta-smooth only" were based on the smoothing $\beta_\ell(t)$ only approach of model (7). The order of B-spline basis was 4, and the number of basis functions of B-spline was $K = K_\beta = 15$; the number of Fourier basis functions was $K = K_\beta = 21$.

The six correlated quantitative traits were generated using the model

$$\begin{aligned} y_{i1} &= 0.2z_{i1} + 0.8z_{i2} + \varepsilon_{i1}, \\ y_{i2} &= 0.3z_{i1} + 0.7z_{i2} + \varepsilon_{i2}, \\ y_{i3} &= 0.4z_{i1} + 0.6z_{i2} + \varepsilon_{i3}, \\ y_{i4} &= 0.5z_{i1} + 0.5z_{i2} + \varepsilon_{i4}, \\ y_{i5} &= 0.6z_{i1} + 0.4z_{i2} + \varepsilon_{i5}, \\ y_{i6} &= 0.7z_{i1} + 0.3z_{i2} + \varepsilon_{i6}, \end{aligned} \tag{9}$$

where z_{i1} and z_{i1} are the same as those of (8).

To be sure that the false positives are properly controlled, empirical type I errors are calculated for the approximate F-distributed tests. For the three trait case, the type I error rates were reported in Tables 3 and 4 of Wang et al. (2015). For six traits, the type I errors of the approximate F-distributed tests are reported in Tables 1 and 2, and they are around the nominal levels and so the false-positive rates are accurately controlled.

Empirical power simulations. For empirical power simulations of quantitative traits, we assumed that 5% of the variants were causal. We considered two scenarios: (1) all variants are rare (MAF < 0.03), and (2) some variants are common and the rest are rare. Once a subregion of size 3–30 kb was selected from the 1 Mb region, a subset of p causal variants located in the subregion was then randomly selected to obtain ordered genotypes $(x_i(t_1), \ldots, x_i(t_p))$. Then, we generated the quantitative traits by adding genetic contributions to models (8) and (9). For instance, the three quantitative traits were generated by

$$\begin{aligned} y_{i1} &= 0.5z_{i1} + 0.5z_{i2} + \beta_{11}x_i(t_1) + \cdots + \beta_{1p}x_i(t_p) + \varepsilon_{i1}, \\ y_{i2} &= 0.3z_{i1} + 0.7z_{i2} + \beta_{21}x_i(t_1) + \cdots + \beta_{2p}x_i(t_p) + \varepsilon_{i2}, \\ y_{i3} &= 0.6z_{i1} + 0.4z_{i2} + \beta_{31}x_i(t_1) + \cdots + \beta_{3p}x_i(t_p) + \varepsilon_{i3}, \end{aligned} \tag{10}$$

where z_{i1}, z_{i2}, and $(\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3})'$ are the same as in the model (8), and the βs are additive effects for the causal variants defined as follows. We used $|\beta_{ij}| = c_i |\log_{10}(MAF_j)|$, where MAF_j was the MAF of the jth variant. For the three trait model (10), we assume that 5% of the variants were causal and the constants c_i are defined by

$$\begin{aligned} c_1 &= \log(10)/(2k), \quad c_2 = \log(8.5)/(2k), \\ c_3 &= \log(7)/(2k); \end{aligned} \tag{11}$$

for the six trait case, we also assume that 5% of the variants were causal and the constants $c_i = 4.0/k$ for all six traits, where k depends on region size. The constants k and genetic effect sizes decrease as region sizes increase:

$$k = \begin{cases} 1.0 & \text{if region size = 3 kb,} \\ 2.0 & \text{if region size = 6 kb,} \\ \cdots & \cdots \\ 9.0 & \text{if region size = 27 kb,} \\ 10.0 & \text{if region size = 30 kb.} \end{cases} \tag{12}$$

It can be seen that the effect sizes $|\beta_{ij}|$ are smaller and smaller when the region sizes in (12) increase. In particular, the number of causal variants is large and each causal variant contributes a small amount to the traits if the region sizes are larger than 12 kb for the three trait case (i.e., $c_i \le \log(10)/(2 \times 4) \approx 0.29$). For the six trait case, the constant $c_i = 0.4$ when region size is 30 kb and this is the same as that in the simulations of Figure 3, Broadaway et al. (2016), except for an additional random contribution $N(0, 1)|\log_{10}(MAF_j)|$. For the three trait case, we also consider a second type of constants: $k = 3.0$, i.e., effect sizes $|\beta_{ij}|$ do not depend on region sizes and are relatively large.

For each setting of empirical power calculations, 1,000 datasets were simulated to calculate the empirical power levels as the proportion of P-values that are smaller than a given $\alpha = 0.01$ level for three traits and $\alpha = 2.5 \times 10^{-6}$

TABLE 2 Empirical type I error rates of the approximate F-distribution tests based on Pillai-Bartlett trace of six traits and moderate correlation, when the variants are only rare

Region size	Sample size	Nominal Level α	Basis of both GVF and $\beta_\ell(t)$		Basis of beta-smooth only		MANOVA Model (1)
			B-sp basis	Fourier basis	B-sp Basis	Fourier basis	
6 kb	500	0.001	0.000906	0.000919	0.000906	0.000916	0.000921
		0.0001	0.000091	0.000088	0.000093	0.000090	0.000085
	1,000	0.001	0.000996	0.000930	0.000998	0.000930	0.000918
		0.0001	0.000096	0.000091	0.000096	0.000091	0.000089
	1,500	0.001	0.000985	0.000991	0.000985	0.000991	0.000984
		0.0001	0.000094	0.000099	0.000094	0.000099	0.000095
9 kb	500	0.001	0.000940	0.000925	0.000940	0.000923	0.000912
		0.0001	0.000090	0.000095	0.000090	0.000095	0.000100
	1,000	0.001	0.000906	0.000969	0.000906	0.000969	0.000900
		0.0001	0.000092	0.000086	0.000092	0.000086	0.000092
	1,500	0.001	0.000981	0.000980	0.000981	0.000980	0.000952
		0.0001	0.000111	0.000091	0.000111	0.000091	0.000076
12 kb	500	0.001	0.000930	0.000901	0.000930	0.000901	0.000909
		0.0001	0.000086	0.000089	0.000086	0.000089	0.000078
	1,000	0.001	0.000905	0.000930	0.000905	0.000930	0.000946
		0.0001	0.000094	0.000085	0.000094	0.000085	0.000094
	1,500	0.001	0.000965	0.000983	0.000965	0.000983	0.000984
		0.0001	0.000099	0.000099	0.000099	0.000099	0.000097
15 kb	500	0.001	0.000950	0.000947	0.000950	0.000947	0.000940
		0.0001	0.000093	0.000099	0.000093	0.000099	0.000093
	1,000	0.001	0.000951	0.000946	0.000951	0.000946	0.000965
		0.0001	0.000103	0.000094	0.000103	0.000094	0.000098
	1,500	0.001	0.000925	0.000966	0.000925	0.000966	0.000987
		0.0001	0.000098	0.000089	0.000098	0.000089	0.000104
18 kb	500	0.001	0.000896	0.000957	0.000896	0.000957	0.000913
		0.0001	0.000077	0.000088	0.000077	0.000088	0.000105
	1,000	0.001	0.000979	0.000955	0.000979	0.000955	0.000946
		0.0001	0.000093	0.000078	0.000093	0.000078	0.000105
	1,500	0.001	0.000969	0.000985	0.000969	0.000985	0.000962
		0.0001	0.000083	0.000114	0.000083	0.000114	0.000105
21 kb	500	0.001	0.000888	0.000929	0.000888	0.000929	0.000936
		0.0001	0.000086	0.000085	0.000086	0.000085	0.000077
	1,000	0.001	0.000879	0.000940	0.000879	0.000940	0.001018
		0.0001	0.000092	0.000095	0.000092	0.000095	0.000093
	1,500	0.001	0.000919	0.000932	0.000919	0.000932	0.000989
		0.0001	0.000086	0.000079	0.000086	0.000079	0.000086
24 kb	500	0.001	0.000943	0.000846	0.000943	0.000846	0.000931
		0.0001	0.000087	0.000091	0.000087	0.000091	0.000076
	1,000	0.001	0.000968	0.000986	0.000968	0.000986	0.000975
		0.0001	0.000085	0.000084	0.000085	0.000084	0.000085
	1,500	0.001	0.000989	0.000990	0.000989	0.000990	0.001014
		0.0001	0.000110	0.000096	0.000110	0.000096	0.000090
27 kb	500	0.001	0.000935	0.000960	0.000935	0.000960	0.000946
		0.0001	0.000105	0.000107	0.000105	0.000107	0.000092
	1,000	0.001	0.000988	0.000974	0.000988	0.000974	0.000984

(continues)

TABLE 2 *(Continued)*

Region size	Sample size	Nominal Level α	Basis of both GVF and $\beta_\ell(t)$		Basis of beta-smooth only		MANOVA Model (1)
			B-sp basis	Fourier basis	B-sp Basis	Fourier basis	
		0.0001	0.000105	0.000106	0.000105	0.000106	0.000098
	1,500	0.001	0.000999	0.000993	0.000999	0.000993	0.000966
		0.0001	0.000097	0.000113	0.000097	0.000113	0.000097
30 kb	500	0.001	0.000900	0.000916	0.000900	0.000916	0.000942
		0.0001	0.000069	0.000082	0.000069	0.000082	0.000083
	1,000	0.001	0.000953	0.000940	0.000953	0.000940	0.000938
		0.0001	0.000109	0.000083	0.000109	0.000083	0.000104
	1,500	0.001	0.000997	0.000940	0.000997	0.000940	0.000980
		0.0001	0.000095	0.000098	0.000095	0.000098	0.000097

The results of 'Basis of both GVF and $\beta_\ell(t)$" were based on smoothing both GVF and genetic effect functions $\beta_\ell(t)$ of model (5), the results of 'Basis of beta-smooth only" were based on the smoothing $\beta_\ell(t)$ only approach of model (7). The order of B-spline basis was 4, and the number of basis functions of B-spline was $K = K_\beta = 15$; the number of Fourier basis functions was $K = K_\beta = 21$.

level for six traits. The results of two combinations of traits are reported: one trivariate combination (y_1, y_2, y_3) and one bivariate combination (y_1, y_2) for three trait case. We calculated the empirical power levels for the approximate F-distributed tests based on Pillai-Bartlett trace, Hotelling-Lawley trace, and Wilks's Lambda. The results of approximate F-distributed tests based on the Pillai-Bartlett trace are reported, which are similar to the results of approximate F-distributed tests based on Hotelling-Lawley trace and Wilks's Lambda. An MAF-weighted linear kernel is used for the genotypic similarity.

Three traits: Power comparison when the constants k are given by relations (12). In this case, genetic effect sizes $|\beta_{ij}|$ decrease as region sizes increase. When some variants are common and the rest are rare, we report in Figure 1 the empirical power of the approximate F-distributed tests of additive models of MANOVA (1) and MFLM (7) and GAMuT at $\alpha = 0.01$. When the region sizes are between 3 kb and 12 kb, both the additive models of MANOVA and MFLM perform better than GAMuT, and the additive models of MANOVA perform better than MFLM. When the region sizes are 15 kb and 18 kb, both the additive models of MANOVA and MFLM perform similarly to GAMuT based on projection matrix, and the additive models of MANOVA start to perform worse than MFLM. When the region sizes are between 21 kb and 27 kb, both the additive models of MANOVA and MFLM perform worse than GAMuT based on projection matrix, and the additive models of MANOVA perform worse than MFLM.

When all variants are rare, we report empirical power levels in Figure 2. When the region sizes are between 3 kb and 9 kb, the additive models of MANOVA perform the best (i.e., better than GAMuT and MFLM), and MFLM performs better than or similar to GAMuT. When the region sizes are 12 kb and 15 kb, the additive models of MANOVA perform similarly to GAMuT based on projection matrix. When the region sizes are between 18 kb and 27 kb, the GAMuT based on projection matrix performs the best.

In Figures 1 and 2, GAMuT based on projection matrix performs similarly to GAMuT based on linear kernel when the region sizes are between 3 kb and 9 kb; When the region sizes are between 12 kb and 27 kb, GAMuT based on projection matrix perform better than GAMuT based on linear kernel.

Three traits: Power comparison when the constant $k = 3.0$. In these cases, genetic effect sizes $|\beta_{ij}|$ do not depend on the region sizes and are relatively large. When some variants are common and the rest are rare, the power levels are presented in Figure 3. When all variants are rare, the power levels are presented in Figure 4. In these figures, the results of 9 kb region sizes are not plotted because they are the same as those in plots (a3) of Figures 1 and 2. The obvious features of Figures 3 and 4 are that the additive models of MANOVA perform the best (i.e., better than GAMuT and MFLM). When some variants are common and the rest are rare, MFLM perform better than GAMuT. When all variants are rare, MFLM perform worse than GAMuT.

Six Traits: Power comparison when the constants k are given by relations (12). If the residual correlations are moderate, the empirical power levels are plotted in Figures 5 and 6. When some variants are common and the rest are rare, the power levels are presented in Figure 5. When all variants are rare, the power levels are presented in Figure 6. It can be seen that the additive models of MANOVA perform the best (i.e., better than GAMuT and MFLM) in Figures 5 and 6. When some variants are common and the rest are rare, MFLM perform better than GAMuT. When all variants are rare, MFLM perform better than GAMuT when the region sizes are between 6 kb and 15 kb, MFLM perform similarly

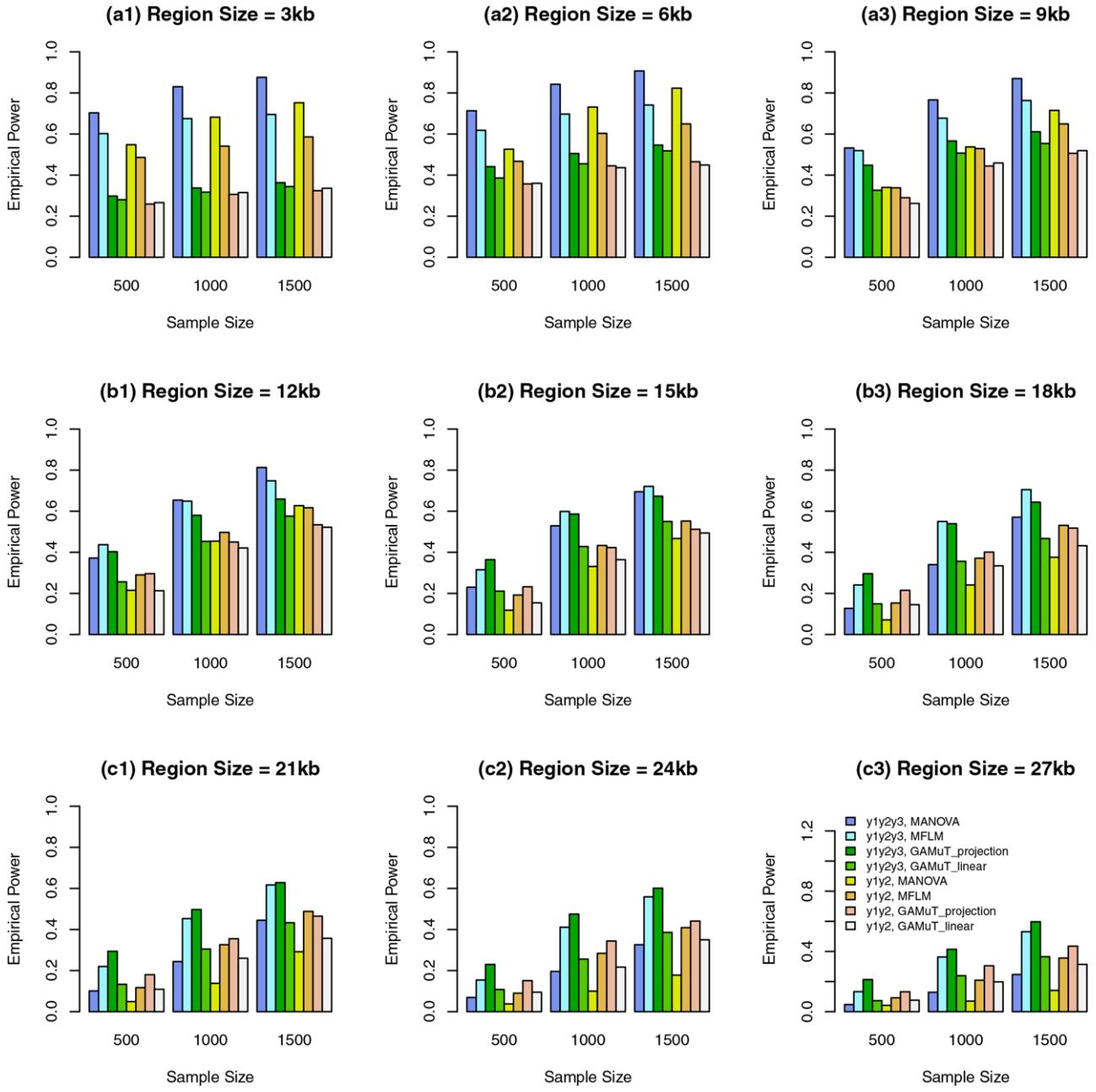

FIGURE 1 The empirical power of the approximate F-distributed tests of the additive models of MANOVA (1) and MFLM (7) using B-spline basis based on Pillai-Bartlett trace and GAMuT at $\alpha = 0.01$, when some variants are common and the rest are rare, the constants k are given by relations (12), 20%/80% causal variants have negative/positive effects for each of three traits, and 5% variants are causal. The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 15$

to GAMuT when the region sizes are between 18 kb and 24 kb, and MFLM perform similarly to or worse than GAMuT when the region sizes are 27 kb and 30 kb.

In supplementary Figures S1 and S2, the power levels are plotted when the residual correlations are low. In supplementary Figures S3 and S4, the power levels are plotted when the residual correlations are high. The features of supplementary Figures S1 and S3 are similar to those of the Figure 1 when some variants are common and the rest are rare, and the features of supplementary Figures S2 and S4 are similar to those of the Figure 2 when all variants are rare.

4 | APPLICATION TO REAL DATA

In Wang et al. (2015), we analyzed data from the TSS and European lipid studies by fixed models. In this report, we analyzed the data by GAMuT. Table 3 reports results of MFLM, additive models of MANOVA, and GAMuT. In the European lipid studies, four lipid quantitative traits were analyzed in 22 gene regions: high-density lipoprotein (HDL) levels, low-density lipoprotein (LDL) levels, triglycerides (TG), and total cholesterol (CHOL). Three quantitative traits (i.e., A, B, and C) from the TSS were analyzed in the region of an enzyme

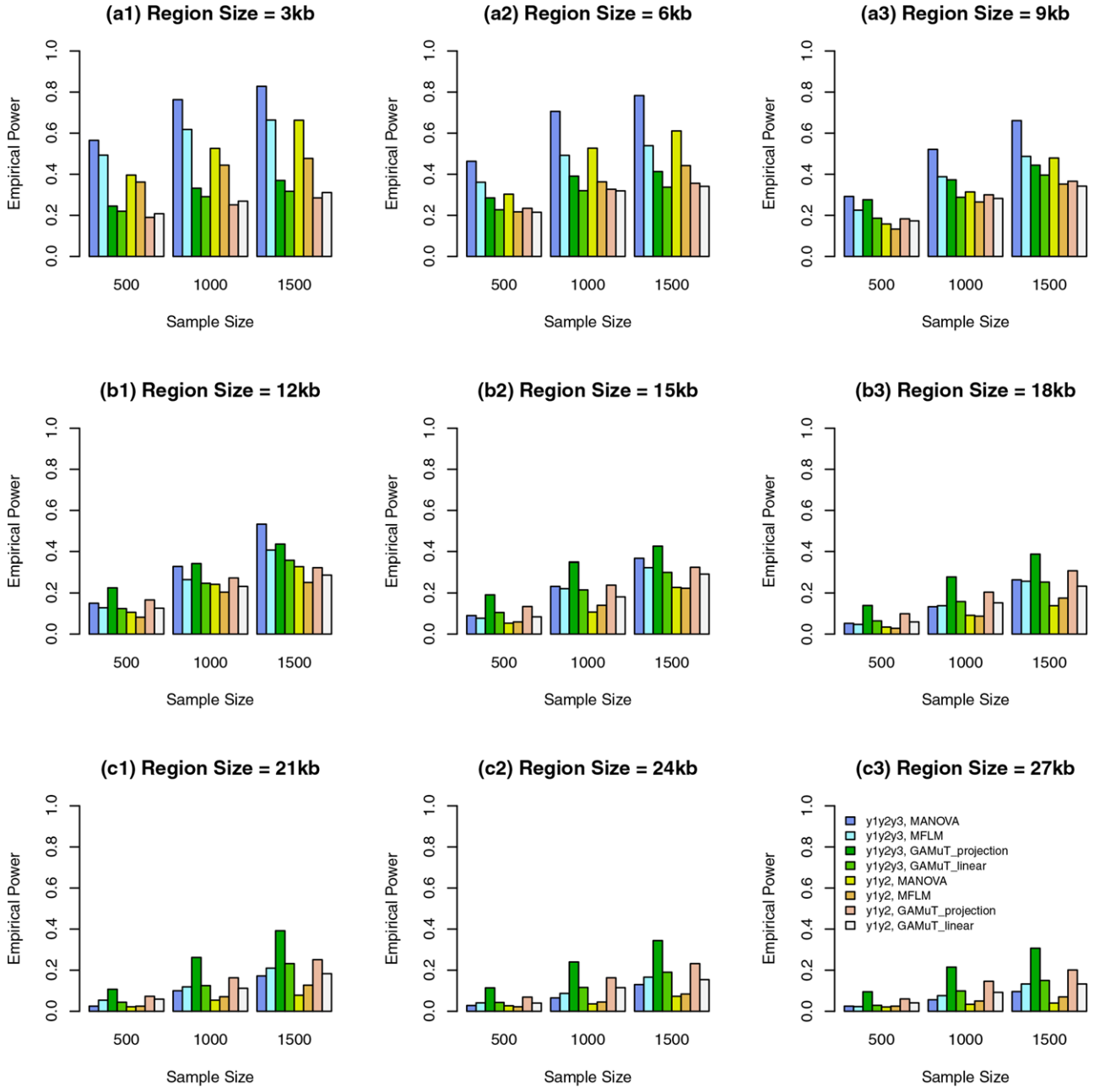

FIGURE 2 The empirical power of the approximate F-distributed tests of the additive models of MANOVA (1) and MFLM (7) using B-spline basis based on Pillai-Bartlett trace and GAMuT at $\alpha = 0.01$, when all variants are rare, the constants k are given by relations (12), 20%/80% causal variants have negative/positive effects for each of three traits, and 5% variants are causal. The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 15$

gene. The associations that attain a threshold significance of $P < 3.1 \times 10^{-6}$ are highlighted in red (Liu et al., 2014). If the P-values are around 10^{-5} but larger than 3.1×10^{-6}, we claim the association as tentative.

In Table 3, the results of GAMuT are new but the other results are mainly from Wang et al. (2015). GAMuT detected only one association signal at gene *LPL* in the FUSION study based on projection matrix for a combination of (LDL, TG, CHOL) ($P = 2.29 \times 10^{-6}$), and this is one of the two cases that MFLM and MANOVA failed to detect an association (the other instance is from a combination of (LDL, TG) at gene *LPL* in study of D2d-2007). In addition, GAMuT based on projection matrix detected seven tentative association signals and GAMuT based on linear kernel detected five. By MFLM and additive models of MANOVA, however, quite a few combinations of lipid traits from the five European studies showed associations or tentative association signals in the regions of the *APOE* and *LDLR* genes, and all combinations of three traits (i.e., A, B, and C) in the TSS showed association with the enzyme gene (Table 3). Moreover, the P-values of

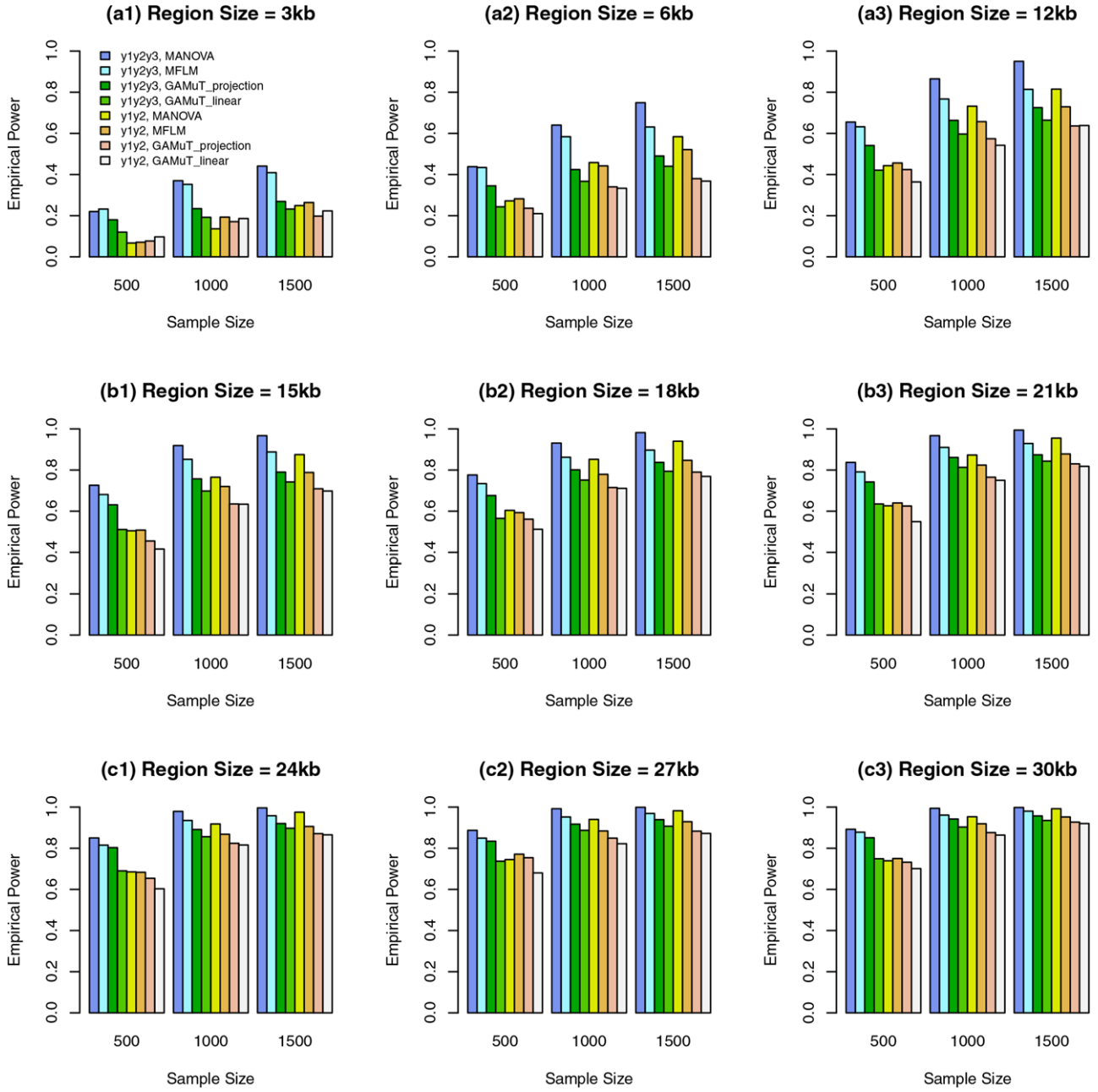

FIGURE 3 The empirical power of the approximate F-distributed tests of the additive models of MANOVA (1) and MFLM (7) using B-spline basis based on Pillai-Bartlett trace and GAMuT at $\alpha = 0.01$, when some variants are common and the rest are rare, the constant $k = 3.0$, 20%/80% causal variants have negative/positive effects for each of three traits, and 5% variants are causal. The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 15$

the approximate F-distributed tests of fixed models are generally much smaller than those of GAMuT. Therefore, the fixed effect MFLM and MANOVA perform better than GAMuT.

In supplementary Tables S3 and S4, we report the results of data analysis of the European lipid studies by dividing the data into rare and common variants based on a cutoff of 0.03. It is worth noting that the gene regions contain both rare and common variants and the associations are mainly from common variants. GAMuT detected a tentative association at gene *LPL* in the FUSION study in supplementary Table S4 based on common variants for the combination (LDL, TG, CHOL) ($P = 2.99 \times 10^{-5}$), but no association signal was detected in supplementary Table S3 based on rare variants ($P = 3.02 \times 10^{-1}$). After combining rare and common variants into one group, GAMuT detected an association signal at gene *LPL* in the FUSION study based on projection matrix in Table 3 ($P = 2.29 \times 10^{-6}$). Interestingly, GATuT was designed to analyze rare variants while the only association was detected in a combination of rare and common variants at gene *LPL*.

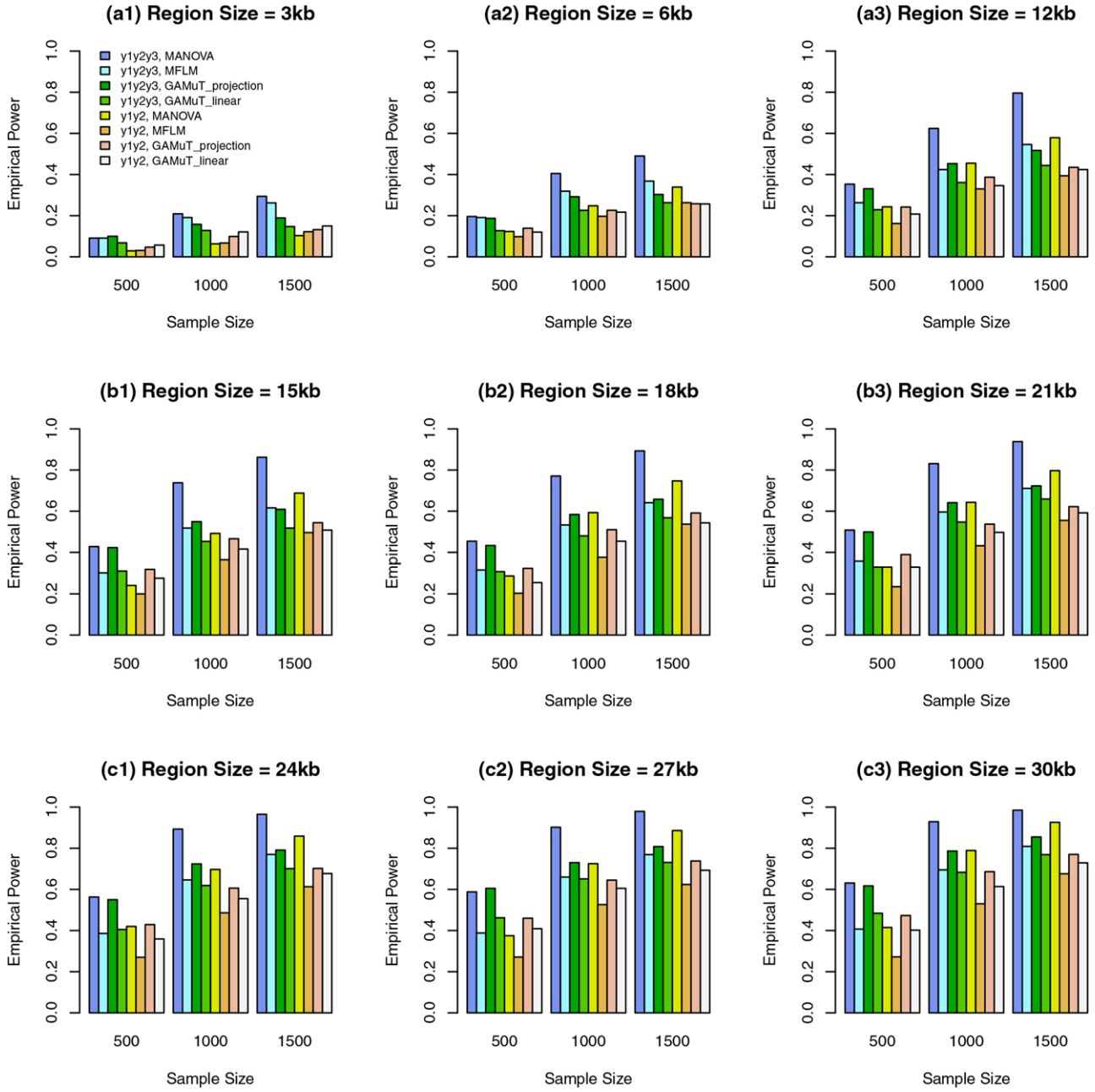

FIGURE 4 The empirical power of the approximate F-distributed tests of the additive models of MANOVA (1) and MFLM (7) using B-spline basis based on Pillai-Bartlett trace and GAMuT at $\alpha = 0.01$, when all variants are rare, the constant $k = 3.0$, 20%/80% causal variants have negative/positive effects for each of three traits, and 5% variants are causal. The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 15$

5 | DISCUSSION

In this study, extensive simulations were performed to evaluate the performance of tests of fixed effect models and GAMuT, by using simulated genetic variants located in 3–30 kb regions. We carried out simulation analyses for two scenarios: (1) all variants are rare; (2) some variants are common and the rest are rare. No matter which scenario, fixed effect MFLM and MANOVA perform better than GAMuT when the genetic effect sizes are relatively large, and GAMuT performs better when the region sizes are large and the genetic effect sizes are small. When the region size grows, MFLM and MANOVA gradually perform worse and GAMuT performs better if the genetic effect sizes are smaller and smaller. In short, MFLM and MANOVA perform well if the effective sizes are relatively large and GAMuT performs well when the effective sizes are small, which was also pointed out in Broadaway et al. (2016).

In prior studies, fixed effect functional regression models were found to outperform SKAT, its optimal unified test (SKAT-O), and a combined sum test of rare and common variant effect (SKAT-C) in most cases (Fan et al., 2013, 2014,

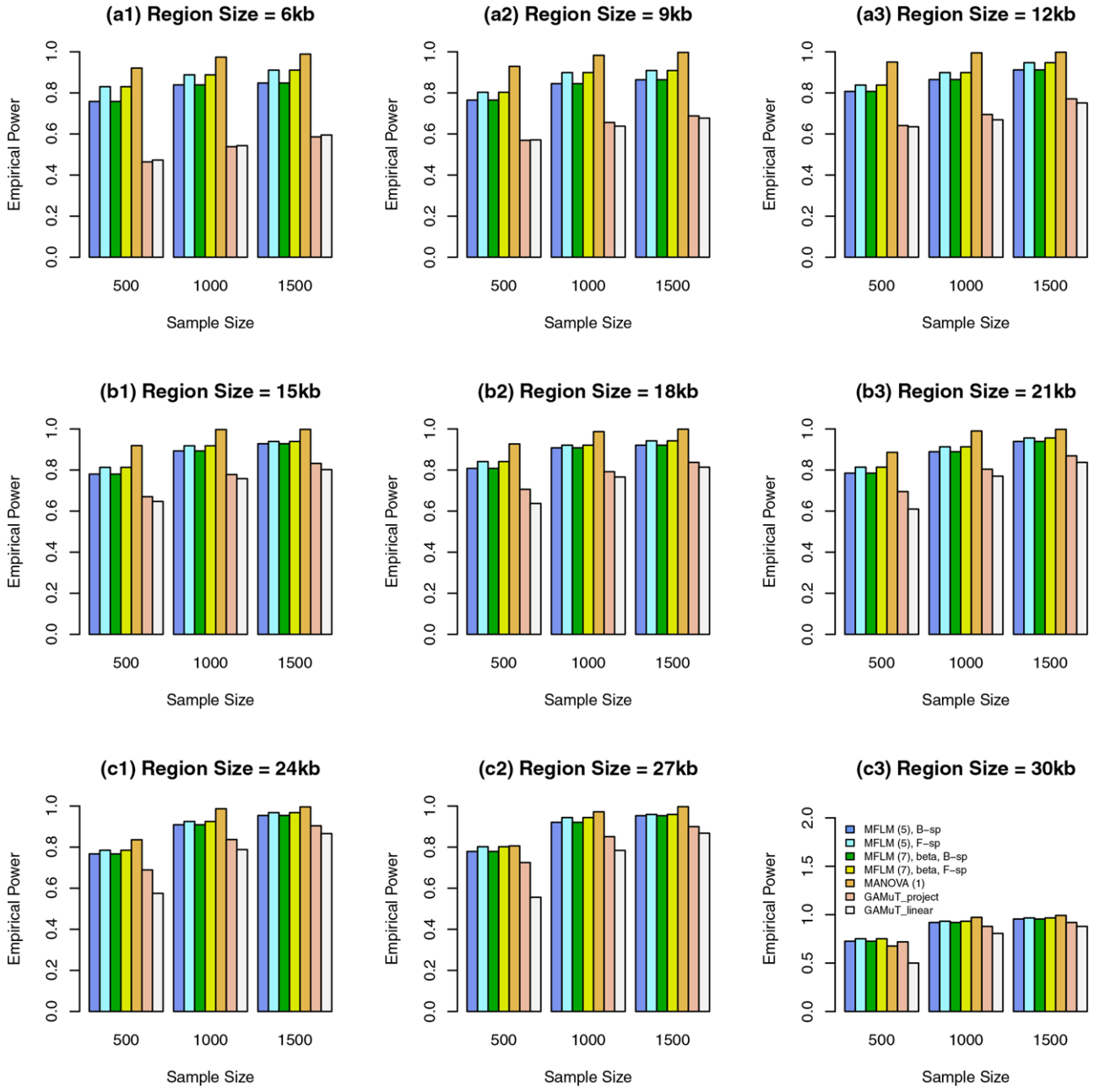

FIGURE 5 The empirical power of the approximate F-distributed tests of the additive models of MANOVA (1) and MFLM (5) and (7) based on Pillai-Bartlett trace and GAMuT at $\alpha = 2.5 \times 10^{-6}$ for six traits and moderate correlation, when some variants are common and the rest are rare, 20%/80% causal variants have negative/positive effects for each of six traits, and 5% variants are causal. The order of B-spline basis was 4, the number of B-spline basis functions was $K = K_\beta = 15$, and the number of Fourier basis functions was $K = K_\beta = 21$

2015, 2016a,b,c; Luo et al., 2011, 2012, 2013; Svishcheva et al., 2015; Vsevolozhskaya et al., 2014, 2016). In Fan et al. (2016c), we compared the performance of MFLM and MANOVA, and the performance of SKAT/SKAT-O/SKAT-C and the univariate fixed models (Fan et al., 2013). For multivariate analysis, no comparison was made because there was no multivariate version of SKAT/SKAT-O/SKAT-C to compare with in Fan et al. (2016c). In this paper, we fill the gap by comparing the performance of MFLM and MANOVA with GAMuT.

Geneticists have long known of the existence of polygenes, which have small effects on phenotypes (Fisher, 1918). If the number of causal genetic variants at a gene locus is very large and each variant contributes a small amount to the traits, SKAT/SKAT-O/SKAT-C and GAMuT perform better than the tests of fixed models. Thus, SKAT/SKAT-O/SKAT-C as well as GAMuT are more appropriate for analyzing polygenic effects. In major gene association analysis, we look for genes that have relatively large effects (otherwise, they are not major genes). When the number of causal genetic variants at

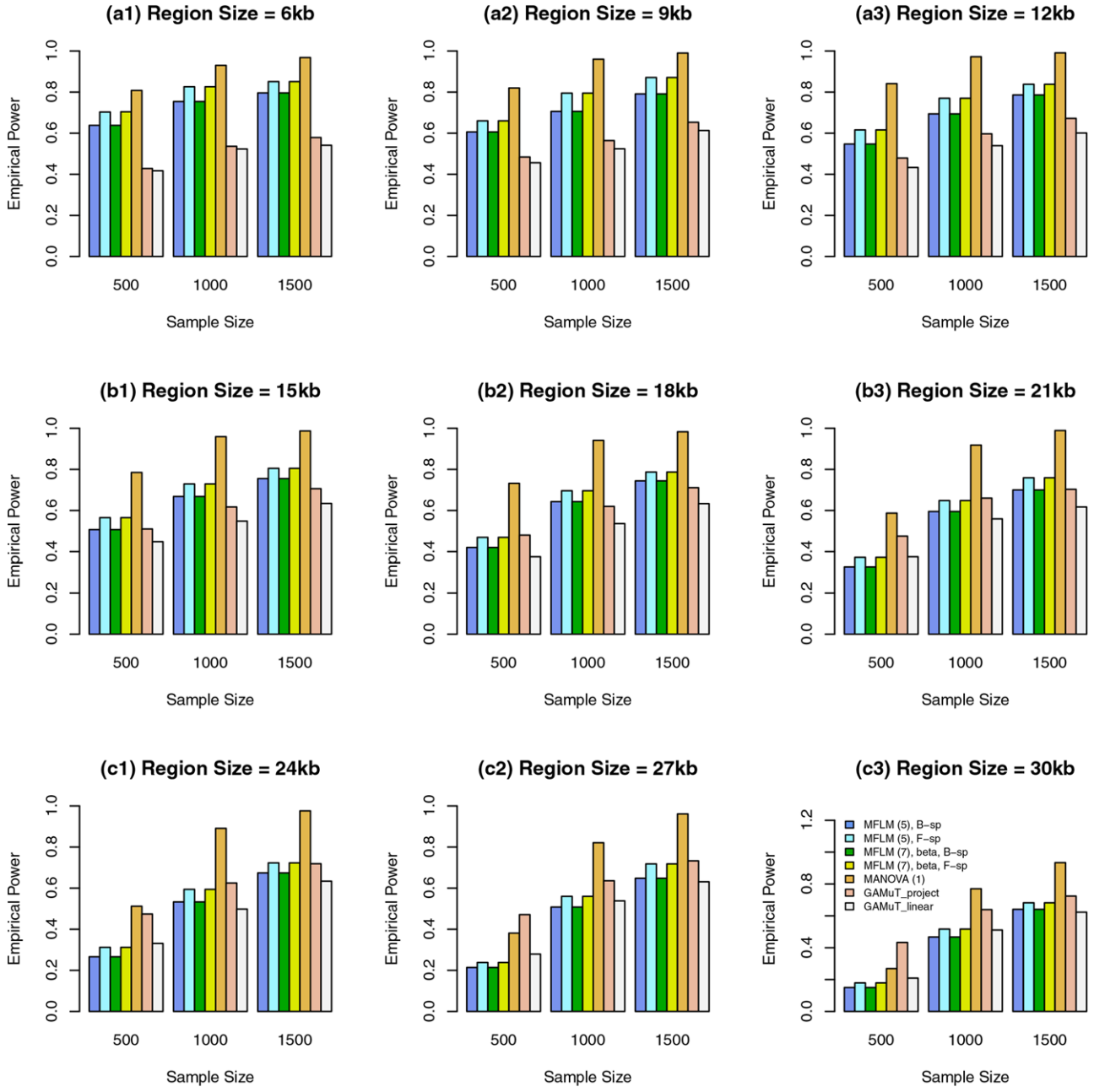

FIGURE 6 The empirical power of the approximate F-distributed tests of the additive models of MANOVA (1) and MFLM (5) and (7) based on Pillai-Bartlett trace and GAMuT at $\alpha = 2.5 \times 10^{-6}$ for six traits and moderate correlation, when all variants are rare, 20%/80% causal variants have negative/positive effects for each of six traits, and 5% variants are causal. The order of B-spline basis was 4, the number of B-spline basis functions was $K = K_\beta = 15$, and the number of Fourier basis functions was $K = K_\beta = 21$

a major gene locus is not very large and the contribution of a few causal variants to the traits is reasonably large, the fixed models should work well, which should be the case for most complex disorders.

The GAMuT procedure was designed for the analysis of rare variants but we use GAMuT to analyze a combination of common and rare variants. As noted in Ionita-Laza et al. (2013), this would be suboptimal and would lead to the common variants drowning out the effects of rare variants. It is very likely that GAMuT can be revised to improve power to analyze a combination rare and common variants by implementing a strategy similar to the combined sum test outlined in Ionita-Laza et al. (2013). In terms of MFLM, it does not need to be weighted by MAF. The genetic effect functions $\beta_\ell(t)$ is actually the effect of the GVFs at the location t, which can be thought of as a weighted effect. In Fan et al. (2014), we explored the issues using weighted GVFs defined by the MAF, and found that the power is very similar to the power without weights. Hence, it is not necessary to add weights in functional regression models. One benefit of treating genotype data functionally is that the genetic effect function naturally serves as a weighting function; this function is

TABLE 3 Results of association analysis of four lipid traits in five European studies in the regions of *APOE, LPL*, and *LDLR* genes and three traits of the Trinity Students Study in the region of an enzyme gene using the F-approximation based on Pillai-Bartlett trace

			P-values of the F-approximation based on Pillai–Bartlett trace					P-values of GAMuT	
			Basis of both GVF and $\beta_\ell(t)$		Basis of beta-smooth only				
Study	Gene	Combinations of traits	B-sp basis	Fourier basis	B-sp basis	Fourier basis	MANOVA Model (1)	Projection matrix	Linear kernel
D2d-2007	APOE	LDL,TG	4.33×10^{-23}	8.96×10^{-23}	4.33×10^{-23}	8.96×10^{-23}	4.92×10^{-22}	2.01×10^{-4}	2.01×10^{-4}
		LDL,CHOL	1.21×10^{-20}	2.08×10^{-19}	1.21×10^{-20}	2.08×10^{-19}	7.91×10^{-19}	4.62×10^{-4}	4.62×10^{-4}
		TG,CHOL	2.98×10^{-18}	2.69×10^{-18}	2.98×10^{-18}	2.69×10^{-18}	1.20×10^{-17}	1.61×10^{-3}	1.61×10^{-3}
		LDL,TG,CHOL	9.10×10^{-20}	3.45×10^{-19}	9.10×10^{-20}	3.45×10^{-19}	1.84×10^{-18}	7.31×10^{-5}	3.51×10^{-4}
	LPL	LDL,TG	5.15×10^{-2}	2.85×10^{-2}	5.15×10^{-2}	2.85×10^{-2}	4.32×10^{-1}	4.36×10^{-5}	4.36×10^{-5}
FUSION	APOE	LDL,TG	3.05×10^{-7}	2.02×10^{-8}	3.05×10^{-7}	2.02×10^{-8}	3.83×10^{-8}	1.90×10^{-1}	1.90×10^{-1}
		LDL,CHOL	1.20×10^{-7}	1.29×10^{-8}	1.20×10^{-7}	1.29×10^{-8}	1.75×10^{-8}	4.88×10^{-2}	4.88×10^{-2}
		TG,CHOL	4.25×10^{-4}	1.06×10^{-5}	4.25×10^{-4}	1.06×10^{-5}	1.93×10^{-5}	4.95×10^{-1}	4.95×10^{-1}
		LDL,TG,CHOL	8.02×10^{-6}	6.44×10^{-7}	8.02×10^{-6}	6.44×10^{-7}	1.11×10^{-6}	1.33×10^{-1}	9.41×10^{-2}
	LPL	LDL,TG	7.11×10^{-5}	2.82×10^{-3}	7.11×10^{-5}	2.82×10^{-3}	2.73×10^{-2}	2.71×10^{-5}	2.71×10^{-5}
		LDL,TG,CHOL	8.51×10^{-4}	1.79×10^{-2}	8.51×10^{-4}	1.79×10^{-2}	6.32×10^{-2}	2.29×10^{-6}	8.61×10^{-4}
Norway	APOE	LDL,TG	1.42×10^{-25}	8.16×10^{-25}	1.42×10^{-25}	8.16×10^{-25}	4.72×10^{-24}	2.43×10^{-4}	2.43×10^{-4}
		LDL,CHOL	8.12×10^{-29}	1.64×10^{-27}	8.12×10^{-29}	1.64×10^{-27}	6.70×10^{-27}	1.13×10^{-4}	1.13×10^{-4}
		TG,CHOL	5.32×10^{-20}	1.46×10^{-19}	5.32×10^{-20}	1.46×10^{-19}	6.08×10^{-19}	1.66×10^{-3}	1.66×10^{-3}
		LDL,TG,CHOL	1.18×10^{-24}	3.06×10^{-23}	1.18×10^{-24}	3.06×10^{-23}	1.68×10^{-22}	8.33×10^{-5}	2.20×10^{-4}
DIAGEN	APOE	LDL,TG	1.78×10^{-8}	1.76×10^{-7}	1.78×10^{-8}	1.76×10^{-7}	4.47×10^{-7}	3.73×10^{-3}	3.73×10^{-3}
		LDL,CHOL	1.24×10^{-9}	1.44×10^{-8}	1.24×10^{-9}	1.44×10^{-8}	3.24×10^{-8}	1.60×10^{-1}	1.60×10^{-1}
		TG,CHOL	2.99×10^{-6}	2.49×10^{-5}	2.99×10^{-6}	2.49×10^{-5}	4.51×10^{-5}	1.71×10^{-1}	1.71×10^{-1}
		LDL,TG,CHOL	1.81×10^{-10}	4.43×10^{-9}	1.81×10^{-10}	4.43×10^{-9}	1.19×10^{-8}	1.25×10^{-3}	1.83×10^{-2}
METSIM	APOE	LDL,TG	2.70×10^{-7}	3.45×10^{-7}	2.70×10^{-7}	3.45×10^{-7}	7.77×10^{-7}	6.29×10^{-4}	6.29×10^{-4}
		LDL,CHOL	3.87×10^{-5}	5.63×10^{-5}	3.87×10^{-5}	5.63×10^{-5}	9.45×10^{-5}	3.08×10^{-3}	3.08×10^{-3}
		LDL,TG,CHOL	1.09×10^{-6}	2.08×10^{-6}	1.09×10^{-6}	2.08×10^{-7}	3.91×10^{-6}	9.51×10^{-4}	1.06×10^{-3}
	LDLR	LDL,TG	1.20×10^{-4}	2.59×10^{-5}	1.20×10^{-4}	2.59×10^{-5}	2.51×10^{-5}	2.85×10^{-4}	2.85×10^{-4}
		LDL,CHOL	3.24×10^{-5}	2.99×10^{-7}	3.24×10^{-5}	2.99×10^{-7}	7.83×10^{-7}	1.12×10^{-5}	1.12×10^{-5}
		TG,CHOL	5.49×10^{-4}	2.03×10^{-5}	5.49×10^{-4}	2.03×10^{-5}	2.09×10^{-5}	1.76×10^{-5}	1.76×10^{-5}
		LDL,TG,CHOL	4.26×10^{-5}	1.19×10^{-6}	4.26×10^{-5}	1.19×10^{-6}	1.72×10^{-6}	6.16×10^{-5}	3.24×10^{-5}
Trinity	An	A,B	2.14×10^{-20}	3.14×10^{-18}	2.14×10^{-20}	3.14×10^{-18}	7.67×10^{-17}	4.21×10^{-3}	2.44×10^{-3}
Students	enzyme	A,C	1.08×10^{-17}	9.53×10^{-16}	1.08×10^{-17}	9.53×10^{-16}	4.46×10^{-15}	2.36×10^{-3}	2.53×10^{-3}
Study	gene	B,C	6.54×10^{-15}	9.51×10^{-12}	6.54×10^{-15}	9.51×10^{-12}	1.05×10^{-10}	8.96×10^{-2}	5.83×10^{-2}
		A,B,C	2.30×10^{-21}	5.87×10^{-18}	2.30×10^{-21}	5.87×10^{-18}	1.56×10^{-16}	7.42×10^{-3}	3.91×10^{-3}

The associations that attain a threshold significance of $P < 3.1 \times 10^{-6}$ are highlighted in red (Liu et al. 2014). The results of "Basis of both GVF and $\beta_\ell(t)$" were based on smoothing both GVF and genetic effect functions $\beta_\ell(t)$ of model (5), and the results of "Basis of β-smooth only" were based on smoothing $\beta_\ell(t)$ only approach of model (7). GVF, genetic variant function.

determined by the data, and takes marker spacing and LD and similarity among individuals into account. In short, the functional regression models are data-driven approaches.

By using gene-based tests, one may discover associations with a variant set. Gene-based tests do not reveal precisely which variants are associated with the disease, but the findings can suggest targeted follow-up and laboratory investigation (Zuk et al., 2014). If all variants had small effects on the phenotypes, it would be hard to locate them. If the contribution of some causal variants to the traits is reasonably large, it would be possible to locate them. We argue that MFLM and MANOVA perform better in most major gene association studies.

In our real data analysis, we found that multivariate fixed models perform better than GAMuT in most gene regions. Note that the European lipid data contain both rare and common variants. As argued by Ionita-Laza et al. (2013), it is reasonable to assume that a combination of rare and common variants affects the risk of many complex disorders. GAMuT detected only one association signal at gene *LPL*, while multivariate fixed models failed to confirm it. Hence, the two methods can be complementary instead of competing with

each other. It is our hope that our work may shed more light in gene-based association analysis to facilitate dissection of complex disorders.

ACKNOWLEDGMENT

Two anonymous reviewers and the editors, Dr. Shete and Dr. Cordell, provided very good and insightful comments for us to improve the manuscript. We greatly thank the European cohorts groups for letting us analyze the data and using them as examples. Dr. Heather M. Stringham and Dr. Tanya M. Teslovich kindly sent us the data of the European cohorts and patiently answered many questions about the cohorts, and we greatly appreciate their help. This study was supported by the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (R.F., C.-Y.C., and J.L.M.), by the Intramural Research Program of the National Human Genome Research Institute (A.F.W. and J.E.B.-W.), National Institutes of Health, Bethesda, MD. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov).

Computer Program. The methods proposed in this paper are implemented by using procedures from functional data analysis (fda) R package. The R codes for multivariate fixed models are available from the web site http://www.nichd.nih.gov/about/org/diphr/bbb/software/fan/Pages/default.aspx

REFERENCES

Allison, D. B., Thiel, B., St Jean, P., Elston, R. C., Infante, M. C., & Schork, N. J. (1998). Multiple phenotype modeling in gene-mapping studies of quantitative traits: Power advantages. *American Journal of Human Genetics 63*, 1190–1201.

Anderson, T. W. (1984). *An introduction to multivariate statistical analysis* (2nd ed.). New York: John Wiley & Sons.

Ansorge, W. J. (2009). Next-generation DNA sequencing techniques. *New Biotechnology 25*, 195–203.

Broadaway, K. A., Cutler, D. J., Duncan, R., Moore, J. L., Ware, E. B., Jhun, M. A., … Epstein, M. P. (2016). A statistical approach for testing cross-phenotype effects of rare variants. *American Journal of Human Genetics 98*, 525–540.

Chavali, S., Barrenas, F., Kanduri, K., & Benson, M. (2010). Network properties of human disease genes with pleiotropic effects. *BMC Systems Biology 4*, 78.

Daniels, P. R., Kardia, S. L., Hanis, C. L., Brown, C. A., Hutchinson, R., Boerwinkle, E., … Genetic Epidemiology Network of Arteriopathy Study. (2004). Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *American Journal of Medicine 116*, 676–681.

de Boor, C. (2001). *Applied mathematical sciences 27. A practical guide to splines*, revised version. New York: Springer.

Fan, R., Wang, Y., Mills, J. L., Wilson, A. F., Bailey-Wilson, J. E., & Xiong, M. M. (2013). Functional linear models for association analysis of quantitative traits. *Genetic Epidemiology 37*, 726–742.

Fan, R., Wang, Y., Mills, J. L., Carter, T. C., Lobach, I., Wilson, A. F., … Xiong, M. M. (2014). Generalized functional linear models for case-control association studies. *Genetic Epidemiology 38*, 622–637.

Fan, R. Z., Wang, Y. F., Boehnke, M., Chen, W., Li, Y., Ren, H. B., … Xiong, M. M. (2015). Gene level meta-analysis of quantitative traits by functional linear models. *Genetics 200*, 1089–1104.

Fan, R. Z., Wang, Y. F., Chiu, C. Y., Chen, W., Ren, H. B., Li, Y., … Xiong, M. M. (2016a). Meta-analysis of complex diseases at gene level with generalized functional linear models. *Genetics 202*, 457–470.

Fan, R. Z., Wang, Y. F., Qi, Y., Ding, Y., Weeks, D. E., Lu, Z. H., … Chen, W. (2016b). Gene-based association analysis for censored traits via functional regressions. *Genetic Epidemiology 40*, 133–143.

Fan, R. Z., Chiu, C. Y., Jung, J. S., Weeks, D. E., Wilson, A. F., Bailey-Wilson, J. E., … Xiong, M. M. (2016c). A comparison study of fixed and mixed effect models for gene level association studies of complex traits. *Genetic Epidemiology 40*, 702–721.

Ferraty, F., & Romain, Y. (2010). *The Oxford handbook of functional data analysis*. New York: Oxford University Press.

Ferreira, M. A., & Purcell, S. M. (2009). A multivariate test of association. *Bioinformatics 25*, 132–133.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Philosophical Transactions of the Royal Society of Edinburg 52*, 399–433.

Galesloot, T. E., van Steen, K., Kiemeney, L. A., Janss, L. L., & Vermeulen, S. H. (2014). A comparison of multivariate genome-wide association methods. *PLoS One 9*, e95923.

Horváth, L., & Kokoszka, P. (2012). *Inference for functional data with applications* New York: Springer.

Huang, J., Johnson, A. D., & ODonnell, C. J. (2011). PRIMe: A method for characterization and evaluation of pleiotropic regions from multiple genome-wide association studies. *Bioinformatics 27*, 1201–1206.

Ionita-Laza, I., Lee, S., Makarov, V., Buxbaum, J. D., & Lin, X. (2013). Sequence kernel association tests for the combined effect of rare and common variants. *American Journal of Human Genetics 92*, 841–853.

Lee, S., Emond, M. J., Bamshad, M. J., Barnes, K. C., Rieder, M. J., Nickerson, D. A., … Lin, X. (2012). Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *American Journal of Human Genetics 91*, 224–237.

Liu, D. J., Peloso, G. M., Zhan, X., Holmen, O. L., Zawistowski, M., Feng, S., Nikpay, M., Auer, P. L., Goel, A., Zhang, H., Peters, U., Farrall, M., Orho-Melander, M., Kooperberg, C., McPherson, R., Watkins, H., Willer, C. J., Hveem, K., Melander, O., Kathiresan, S., Abecasis, G. R. (2014). Meta-analysis of gene-level tests for rare variant association. *Nature Genetics 46*, 200–204.

Luo, L., Boerwinkle, E., & Xiong, M. M. (2011). Association studies for next-generation sequencing. *Genome Research 21*, 1099–1108.

Luo, L., Zhu, Y., & Xiong, M. M. (2012). Quantitative trait locus analysis for next-generation sequencing with the functional linear models. *Journal of Medical Genetics 49*, 513–524.

Luo, L., Zhu, Y., & Xiong, M. M. (2013). Smoothed functional principal component analysis for testing association of the entire allelic spectrum of genetic variation. *European Journal of Human Genetics 21*, 217–224.

Maity, A., Sullivan, P. F., & Tzeng, J. Y. (2012). Multivariate phenotype association analysis by marker-set kernel machine regression. *Genetic Epidemiology 36*, 686–695.

Mardis, E. R. (2008). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics 9*, 387–402.

Metzker, M. L. (2010). Sequencing technologies the next generation. *Nature Reviews Genetics 11*, 31–46.

O'Reilly, P. F., Hoggart, C. J., Pomyen, Y., Calboli, F. C., Elliott, P., Jarvelin, M. R., & Coin, L. J. (2012). MultiPhen: Joint model of multiple phenotypes can increase discovery in GWAS. *PLoS One 7*, e34861.

Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis* (2nd ed.). New York: Springer.

Ramsay, J. O., Hooker, G., & Graves, S. (2009). *Functional data analysis with R and Matlab*. New York: Springer.

Rao, C. R. (1973). *Linear statistical inference and its applications* (2nd ed.). New York: John Wiley & Sons.

Ried, J. S., Döring, A., Oexle, K., Meisinger, C., Winkelmann, J., Klopp, N., … Gieger, C. (2012). PSEA: Phenotype Set Enrichment Analysis—A new method for analysis of multiple phenotypes. *Genetic Epidemiology 36*, 244–252.

Rusk, N., & Kiermer, V. (2008). Primer: Sequencing the next generation. *Nature Methods 5*, 15.

Schaffner, S. F., Foo, C., Gabriel, S., Reich, D., Daly, M. J., & Altshuler, D. (2005). Calibrating a coalescent simulation of human genome sequence variation. *Genome Research 15*, 1576–1583.

Shendure, J., & Ji, H. (2008). Next-generation DNA sequencing. *Nature Biotechnology 26*, 1135–1145.

Sivakumaran, S., Agakov, F., Theodoratou, E., Prendergast, J. G., Zgaga, L., Manolio, T., … Campbell, H. (2011). Abundant pleiotropy in human complex diseases and traits. *American Journal of Human Genetics 89*, 607–618.

Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in complex traits: Challenges and strategies. *Nature Reviews Genetics 14*, 483–495.

Svishcheva, G. R., Belonogova, N. M., & Axenovich, T. I. (2015). Region-based association test for familial data under functional linear models. *PLoS One 10*, e0128999.

Vsevolozhskaya, O. A., Zaykin, D. V., Greenwood, M. C., Wei, C., & Lu, Q. (2014). Functional analysis of variance for association studies. *PLoS One 9*(9), e105074.

Vsevolozhskaya, O. A., Zaykin, D. V., Barondess, D. A., Tong, X., Jadhav, S., & Lu, Q. (2016). Uncovering local trends in genetic effects of multiple phenotypes via functional linear models. *Genetic Epidemiology 40*, 210–221.

Wang, Y. F., Liu, A. Y., Mills, J. L., Boehnke, M., Wilson, A. F., Bailey-Wilson, J. E., … Fan, R. Z. (2015). Pleiotropy analysis of quantitative traits at gene level by multivariate functional linear models. *Genetic Epidemiology 39*, 259–275.

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., & Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *American Journal of Human Genetics 89*, 82–93.

Zuk, O., Schaffner, S. F., Samocha, K., Do, R., Hechter, E., Kathiresan, S., … Lander, E. S. (2014). Searching for missing heritability: Designing rare variant association studies. *Proceedings of the National Academy of Sciences 111*(4), E455–E464.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.