



A Novel Approach Integrating Hierarchical Clustering and Weighted Combination for Association Study of Multiple Phenotypes and a Genetic Variant

Liwan Fu^{1,2}, Yuquan Wang¹, Tingting Li¹ and Yue-Qing Hu^{1,3*}

¹ State Key Laboratory of Genetic Engineering, Human Phenome Institute, Institute of Biostatistics, School of Life Sciences, Fudan University, Shanghai, China, ² Center for Non-communicable Disease Management, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China, ³ Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China

OPEN ACCESS

Edited by:

Sebastian Zöllner,
University of Michigan, United States

Reviewed by:

Qiuying Sha,
Michigan Technological University,
United States
Peter Holmans,
Cardiff University, United Kingdom

*Correspondence:

Yue-Qing Hu
yuehu@fudan.edu.cn

Specialty section:

This article was submitted to
Statistical Genetics and Methodology,
a section of the journal
Frontiers in Genetics

Received: 16 February 2021

Accepted: 20 April 2021

Published: 17 June 2021

Citation:

Fu L, Wang Y, Li T and Hu Y-Q
(2021) A Novel Approach Integrating
Hierarchical Clustering and Weighted
Combination for Association Study of
Multiple Phenotypes and a Genetic
Variant. *Front. Genet.* 12:654804.
doi: 10.3389/fgene.2021.654804

As a pivotal research tool, genome-wide association study has successfully identified numerous genetic variants underlying distinct diseases. However, these identified genetic variants only explain a small proportion of the phenotypic variation for certain diseases, suggesting that there are still more genetic signals to be detected. One of the reasons may be that one-phenotype one-variant association study is not so efficient in detecting variants of weak effects. Nowadays, it is increasingly worth noting that joint analysis of multiple phenotypes may boost the statistical power to detect pathogenic variants with weak genetic effects on complex diseases, providing more clues for their underlying biology mechanisms. So a Weighted Combination of multiple phenotypes following Hierarchical Clustering method (WCHC) is proposed for simultaneously analyzing multiple phenotypes in association studies. A series of simulations are conducted, and the results show that WCHC is either the most powerful method or comparable with the most powerful competitor in most of the simulation scenarios. Additionally, we evaluated the performance of WCHC in its application to the obesity-related phenotypes from Atherosclerosis Risk in Communities, and several associated variants are reported.

Keywords: GWAS, hierarchical cluster, multiple phenotypes, score test, obesity

INTRODUCTION

Traditionally, Genome-Wide Association Studies (GWASs) aim to identify genetic variants associated with certain phenotypes for explaining complex diseases (O'Reilly et al., 2012; Yang and Wang, 2012). In GWASs, multiple related phenotypes of diseases are typically collected for getting better understand complex diseases (Yang Q. et al., 2010). For example, hypertension is directly dependent on the magnitudes of Systolic Blood Pressures (SBP) and Diastolic Blood Pressures (DBP) (Yang and Wang, 2012). The correlation coefficient between SBP and DBP is greater than 0.5 in 95% of patients (Gavish et al., 2008), and researchers could acquire SBP and DBP together for studying hypertension. Similarly, Type 2 Diabetes (T2D) study often gathers relevant risk factors and a number of diabetes-related quantitative phenotypes. Moreover, metabolic syndrome encompasses insulin resistance, obesity, atherosclerotic dyslipidemia,

and hypertension; and these factors are interrelated to share potential genetic mediators, pathways, and mechanisms (Huang, 2009). In statistical genetics, jointly analyzing multiple phenotypes can enhance the power of association tests to identify genetic markers associated with one or more phenotypes (Aschard et al., 2014). One of the common approaches for analyzing multiple related phenotypes is to conduct single-phenotype separately and report the results for each phenotype (O'Reilly et al., 2012). However, analysis for one phenotype at a time will be inevitably subject to multiple testing corrections, which leads to a power loss in GWAS (Yang Q. et al., 2010). In recent years, joint analysis of multiple phenotypes has become catching on because of its enhanced statistical power in the detection of genetic variants compared to analysis for each phenotype separately (Yang Q. et al., 2010; Aschard et al., 2014).

Nowadays, jointly analyzing multiple phenotypes has been developed rapidly, which is of two categories: univariate analysis and multivariate analysis. Univariate analysis, as the name suggests, combines various test statistics or p -values of univariate association analysis by means of different strategies. Recently, some approaches of univariate analysis have been proposed for exploring the association between multiple phenotypes and a genetic variant (van der Sluis et al., 2013; Liang et al., 2016; Yang et al., 2016). For example, Kwak et al. (Kwak and Pan, 2016) established an adaptive testing approach, which employs summary statistics from GWASs to evaluate the relationship between multiple phenotypes and a genetic variant. TATES mainly conducts p -values from the association between phenotypes and Single Nucleotide Polymorphisms (SNPs) and concurrently adjusting the correlations among various phenotypes (van der Sluis et al., 2013). Adaptive Fisher's combination (AFC) (Liang et al., 2016) combines a number of optimal p -values from the traditional GWASs. Compared to multivariate analysis, univariate analysis is generally in a unified framework and tends to ignore the crucial information among multiple phenotypes, which may result in reducing statistical power. In recent years, a series of multivariate analysis approaches including mixed-effects models (Korte et al., 2012; Zhou and Stephens, 2014; Casale et al., 2015), Generalized Estimating Equation (GEE) (Zeger and Liang, 1986; Zhang et al., 2014), and reverse regression methods (O'Reilly et al., 2012; Yan et al., 2013; Wang et al., 2016) have been developed. Mixed-effects models comprise Linear Mixed Effects model (LME) model and Generalized Linear Mixed effects Model (GLMM), where the genetic variants are regarded as the fixed effects and the correlation among phenotypes is considered as random effects. The GEE method collapses the random effects and random residual errors in marginal regression models, which makes it different from LME. The reverse regression methods regard genotypes as the response variable and multiple phenotypes as predictors, such as the proportional odds logistic regression for joint model of multiple phenotypes (MultiPhen) (O'Reilly et al., 2012). Multivariate analysis methods are complicated, and few available software has been developed to implement these methods (Yang and Wang, 2012).

In this article, we establish a novel allele-based approach aiming at detecting association between multiple phenotypes

and a genetic variant for improving the power in association studies. We first employ the Hierarchical Clustering based on Different methods for calculating Correlation coefficients (HCDC) (Fu, 2020) to cluster the enrolled phenotypes into several groups. Then, inspired by Weighted Combination of multiple Phenotypes (WCmulP) (Zhu et al., 2018), which provides optimal weights in linear combination, we perform WCmulP in each cluster to generate a novel phenotype by virtual of combining the multiple phenotypes. Subsequently, for every cluster, score test derived from the logistic regression model is constructed to test the association between the genetic variant and the novel phenotype. The permutation procedure is employed to evaluate the p -values of the score test statistics, and their minimum is taken as the test statistic for detecting association between the genetic variant and all phenotypes. Consequently, the Weighted Combination of multiple phenotypes following Hierarchical Clustering method (WCHC) is proposed. Using extensive simulation scenarios, we compare the performance of WCHC with that of six existing methods: O'Brien's method (O'Brien, 1984), MultiPhen (O'Reilly et al., 2012), MANOVA (Cole et al., 1994), SHet (Zhu et al., 2015), TATES (van der Sluis et al., 2013), and WCmulP (Zhu et al., 2018). The results reveal that WCHC is either the most powerful test or comparable with the most powerful tests among the methods we compared in most of the simulation scenarios. Finally, we evaluate the performance of WCHC approach by utilizing the obesity-related phenotypes from a real dataset, Atherosclerosis Risk in Communities (ARIC) Study from dbGaP, and 11 obesity-associated SNPs are detected.

MATERIALS AND METHODS

Proposed WCHC

Suppose a sample of N individuals each have M quantitative phenotypes Y_1, Y_2, \dots, Y_M and genotype G at a genetic variant. It is straightforward to calculate the correlation coefficient between two sets of phenotypes. Based on our previous work (Fu, 2020), the hierarchical clustering is conducted, and finally we have K clusters C_1, C_2, \dots, C_K . Let M_k denote the number of phenotypes in the k th cluster C_k , $k = 1, 2, \dots, K$. We take the first cluster C_1 as an example to show the subsequent procedure. Without loss of generality, assume Y_1, Y_2, \dots, Y_{M_1} are the M_1 phenotypes in the first cluster. Borrowing the allele-based regression idea (Zhu et al., 2018), we introduce $x_{2i-1} = x_{2i} = 1$, $x_{2i-1} = x_{2i} = 0$, and $x_{2i-1} = 1$ and $x_{2i} = 0$, if the genotype of the i th individual is AA , aa , and Aa , respectively, $i = 1, 2, \dots, N$. By analogy, let $y_{2i,j} = y_{2i-1,j}$ be the value of the j th phenotype of individual i , $i = 1, 2, \dots, N$, $j = 1, 2, \dots, M_1$. Based on $\{x_l, y_{l,1}, y_{l,2}, \dots, y_{l,M_1}\}_{l=1}^{2N}$, we establish the following model:

$$\begin{aligned} \text{logit } P(x_l = 1 | y_{l,1}, y_{l,2}, \dots, y_{l,M_1}) &= \beta_0 + \beta_1 y_{l,1} + \beta_2 y_{l,2} \\ &+ \dots + \beta_{M_1} y_{l,M_1}, l = 1, 2, \dots, 2N \end{aligned}$$

to test the association between multiple phenotypes Y_1, Y_2, \dots, Y_{M_1} and a genetic variant.

Instead of the conventional score test that is vulnerable in the case of big M_1 , we adopt the following test statistic (Zhu et al., 2018):

$$T_1 = \sum_{l=1}^{2N} (x_l - \bar{x})(y_l - \bar{y}),$$

where $\bar{x} = \frac{1}{2N} \sum_{l=1}^{2N} x_l$, $\bar{y} = \frac{1}{2N} \sum_{l=1}^{2N} y_l$, $y_l = \sum_{j=1}^{M_1} w_j y_{l,j}$, $w_j = \frac{\sum_{l=1}^{2N} (x_l - \bar{x})(y_{l,j} - \bar{y}_j)}{\sum_{l=1}^{2N} (y_{l,j} - \bar{y}_j)^2}$, $\bar{y}_j = \frac{1}{2N} \sum_{l=1}^{2N} y_{l,j}$, $j = 1, 2, \dots, M_1$. Similarly, we have the corresponding test statistics T_2, \dots, T_K when we study the association of the genetic variant and the multiple phenotypes in the clusters C_2, \dots, C_K , respectively. Further, let p_1, p_2, \dots, p_K be the p -values of T_1, T_2, \dots, T_K , respectively, and we propose our test statistic:

$$T_{WCHC} = \min\{p_1, p_2, \dots, p_K\}.$$

As it is not easy to derive the distribution of the test statistics T_1, T_2, \dots, T_K under the null hypothesis of no association, the permutation procedure described below is employed to calculate the p -value of T_{WCHC} .

(1) In each of the B permutations, we random shuffle the genotypes and then get the statistics $T_1^{(b)}, T_2^{(b)}, \dots, T_K^{(b)}$, $b = 0, 1, 2, \dots, B$. Note that $b = 0$ is corresponding to the original data (no permutation).

(2) Calculate $p_k^{(b)}$ by:

$$p_k^{(b)} = \frac{\#\{d: T_k^{(d)} > T_k^{(b)} \text{ for } d=0,1,\dots,B\}}{B}, \text{ for } k = 1, 2, \dots, K.$$

and then $T_{WCHC}^{(b)} = \min\{p_1^{(b)}, p_2^{(b)}, \dots, p_K^{(b)}\}$ for $b = 0, 1, 2, \dots, B$;

(3) Then, the p -value of T_{WCHC} is given by:

$$\frac{\#\{b : T_{WCHC}^{(b)} < T_{WCHC}^{(0)} \text{ for } b = 1, 2, \dots, B\}}{B}.$$

The hierarchical clustering based on our previous work (Fu, 2020) is as follows: In summary, we can find a partition ψ that partitioned M phenotypes into K disjoint clusters C_1, C_2, \dots, C_K , where $\psi = \{C_1, C_2, \dots, C_K\}$ with $\bigcup_{k=1}^K C_k = \{1, 2, \dots, M\}$ and $C_k \cap C_l = \emptyset (k \neq l)$. Specifically, applying the bottom-up hierarchical clustering approach, we begin with each phenotype as a singleton cluster and then subsequently merge pairs of clusters with the largest similarity until all clusters have been merged into a single cluster that contains all phenotypes. The largest similarity in each iteration is referred as the height of the merged cluster in the dendrogram. A stopping criterion determines the number of clusters, which is similar to an established principle (Bühlmann et al., 2013). Suppose h_b is the largest similarity between two clusters in iteration b ($b \geq 1$) or the height of iteration b . We define:

$$\hat{b} = \arg \min_{b \geq 1} (h_{b+1} - h_b).$$

Then, the number of clusters identified at the iteration \hat{b} is chosen to determine the K clusters C_1, C_2, \dots, C_K . On the calculation of correlation coefficient, the Pearson correlation coefficient, multiple correlation coefficient, and canonical correlation

coefficient are respectively employed according to the number of phenotypes in the merged two clusters.

The source code for WCHC method can be found in <https://github.com/YQHFD/WCHC>.

Comparison of Methods

For convenience, let $\mathbf{1}_n$ be the all ones vector of length n and $\mathbf{0}_n$ be the all zeroes vector of length n , where n is a positive integer. We first list the following existing methods for power comparison with the proposed WCHC.

OB (O'Brien's method) (O'Brien, 1984): Using a linear combination of univariate statistics, the OB statistic, $\mathbf{1}_M^T \Sigma^{-1} T_{uni}$, is developed. It is the most powerful statistic when a class of statistics is a linear combination of T_{uni} , where T_{uni} is the vector of univariate statistics and Σ is the variance-covariance matrix of T_{uni} .

MultiPhen (Joint model of Multiple Phenotypes) (O'Reilly et al., 2012): Modeling the genotype data as ordinal response and phenotypes as predictors, MultiPhen employs likelihood ratio test to evaluate the null hypothesis in the proportional odds logistic regression.

MANOVA (Multivariate Analysis Of Variance) (Cole et al., 1994): In the standard MANOVA, there are a total of M phenotypes, and the $M \times M$ symmetrical background variance-covariance matrix Σ is unconstrained. It has $((M+1) \times M)/2$ freely estimated elements in covariances and variances. Standard MANOVA tests the null hypothesis that the M regression coefficients are all zeroes, which is asymptotically equal to the F -test.

SHet (Test for Heterogeneous genetic effects) (Zhu et al., 2015): The test statistic of SHet, S_{Het} , is based on S_{Hom} , which is the most powerful statistic when the genetic effects are homogeneous. $S_{Hom} = \frac{\mathbf{1}_M^T (CorrW)^{-1} T_{uni} (\mathbf{1}_M^T (CorrW)^{-1} T_{uni})^T}{\mathbf{1}_M^T (WCorrW)^{-1} \mathbf{1}_M}$, where $Corr$ is the correlation matrix of T_{uni} , W is a diagonal matrix of weights for the univariate statistic. S_{Het} is the maximum of S_{Hom} 's satisfying various thresholds. Specifically, only the statistics with absolute values greater than the given threshold are employed; $Corr$ and W are partially used corresponding to the selected statistics. The p -value of S_{Het} could be estimated by simulation.

TATES (Trait-based Association Test that uses Extended Simes procedure) (van der Sluis et al., 2013): TATES combines the p -values of univariate analysis for getting a comprehensive p -value, while correcting the correlation between phenotypes. The TATES p -value is denoted as $\min \left(\frac{M_e p_{(j)}}{M_{e(j)}} \right)$, where $p_{(j)}$ is the j^{th} ($j = 1, \dots, M$) sorted p -value in ascending order; M_e and $M_{e(j)}$ denote the effective number of independent p -values among all M phenotypes and m specific phenotypes, respectively. The effective numbers can be obtained from the correlation matrix of p -values.

WCmulP (Weighted Combination of multiple Phenotypes) (Zhu et al., 2018): WCmulP can be taken as a component of WCHC. The original phenotypes are not used clustering and directly applied the logistic regression. Then the T statistic is proposed to test the association between the phenotypes and genetic variants. Lastly, the permutation procedure is used to derive the distribution of the test statistic T .

Simulation Studies

Assume that the population is in Hardy–Weinberg equilibrium (HWE), and the genotypes of the genetic variants follow the binomial distribution with parameter 2 and the minor allele frequency (MAF). We set MAF = 0.3 in this simulation study for all scenarios. The multiple phenotypes are generated via the following factor model (van der Sluis et al., 2013):

$$y = \lambda x + c\gamma f + \sqrt{1 - c^2} \times \varepsilon,$$

where $y = (y_1, \dots, y_M)^T$ is the M phenotypes; x is the genotype; $\lambda = (\lambda_1, \dots, \lambda_M)^T$ is the vector of effect sizes of the variant on the M phenotypes; f is the vector of factors; $f = (f_1, \dots, f_R)^T \sim MVN(0, \Sigma)$, $\Sigma = (1 - \rho)I + \rho A$; I is the identity matrix; A is a matrix with elements of 1; R is the number of factors; and ρ is the correlation between factors; γ is an $M \times R$ matrix; c is a constant; $\varepsilon = (\varepsilon_1, \dots, \varepsilon_M)^T$ is a vector of random errors; and $\varepsilon_1, \dots, \varepsilon_M$ are mutually independent and follow the standard normal distributions. Consider the following six models with varied numbers of factors.

Model 1: There is only one factor, and the genotype has an influence on all phenotypes with the same effect size. Namely, $R = 1$, $\lambda = \beta \mathbf{1}_M$, and $\gamma = \mathbf{1}_M$.

Model 2: There are two factors and a genotype has an effect on one factor with the same effect. That is, $R = 2$, $\lambda = (\mathbf{0}_{M/2}^T, \beta \mathbf{1}_{M/2}^T)^T$, and $\gamma = bdiag(\mathbf{1}_{M/2}, \mathbf{1}_{M/2})$, which represents the block diagonal matrix of $\mathbf{1}_{M/2}$ and $\mathbf{1}_{M/2}$.

Model 3: There are two factors, and a genotype has an effect on the second factor with different sizes. That is, $R = 2$, $\lambda = (\mathbf{0}_{M/2}^T, \frac{\beta}{M+1} [1 : M/2]^T + \beta \mathbf{1}_{M/2}^T)^T$ and $\gamma = bdiag(\mathbf{1}_{M/2}, \mathbf{1}_{M/2})$, where $[1 : M/2]$ represents the vector of components 1, 2, ..., $M/2$.

Model 4: There are four factors, and a genotype has an impact on the last factor with the same size. That is, $R = 4$, $\lambda = (\mathbf{0}_{3M/4}^T, \beta \mathbf{1}_{M/4}^T)^T$, and $\gamma = bdiag(\mathbf{1}_{M/4}, \mathbf{1}_{M/4}, \mathbf{1}_{M/4}, \mathbf{1}_{M/4})$.

Model 5: There are four factors, and a genotype has an effect on the last factor with different sizes. Namely, $R = 4$, $\lambda = (\mathbf{0}_{3M/4}^T, \frac{\beta}{M+1} [1 : M/4]^T + \beta \mathbf{1}_{M/4}^T)^T$, $\gamma = bdiag(\mathbf{1}_{M/4}, \mathbf{1}_{M/4}, \mathbf{1}_{M/4}, \mathbf{1}_{M/4})$.

Model 6: There are four factors, and a genotype has an impact on the last two factors with different effect directions. That is, $R = 4$, $\lambda = (\mathbf{0}_{M/2}^T, -\frac{\beta}{M+1} [1 : M/4]^T - \beta \mathbf{1}_{M/4}^T, \beta \mathbf{1}_{M/4}^T)^T$, $\gamma = bdiag(\mathbf{1}_{M/4}, \mathbf{1}_{M/4}, \mathbf{1}_{M/4}, \mathbf{1}_{M/4})$.

For these six models, the within-factor correlation is c^2 and the between-factor correlation is ρc^2 . For estimating type I error rates and powers, we fix $N = 1,000$ unrelated subjects, the number of phenotypes $M = 16, 32$. By means of setting $\beta = 0$, we generate all phenotypes that is independent of genotypes to evaluate the type I error rates of all methods, including OB, MultiPhen, MANOVA, SHet, TATES, WCMulP, and WCHC. The corresponding Q–Q plot of type I error rates is shown in **Supplementary Figures 1–6**. Importantly, to evaluate powers, we not only vary the values of β (while within-factor correlation $c^2 = 0.5$ and between-factor correlation $\rho c^2 = 0.1$) but also change the values of within-factor correlation $c^2 = 0.3, 0.5, 0.7$, and 0.9 (while between-factor correlation $\rho c^2 = 0.1$).

The calculation of heritability is as follows: the heritability of genotypes to the j -th phenotype is given by $h^2(y_j) = \frac{var(x)\lambda_j^2}{var(x)\lambda_j^2 + 1} \approx var(x)\lambda_j^2$. The heritability of genotypes to the total M phenotypes is given by $h^2 = \sum_{j=1}^M h^2(y_j) \approx var(x) \sum_{j=1}^M \lambda_j^2$. Then given the parameters λ , M , and MAF, we can calculate h^2 for the different models.

Simulation Results

We set different nominal significance levels, various numbers of phenotypes, and distinct number of factors to estimate the type I error rates of WCHC and other six methods. For each simulation scenario, the p -values of WCHC, WCMulP, and SHet are evaluated by 2,000 permutations; and the p -values of MANOVA, MultiPhen, TATES, and OB are evaluated by their asymptotic distributions. The type I error rates of the seven methods are estimated using 2,000 replicated samples. For 2,000 replicated samples, the 95% confidence intervals (CIs) for type I error rates of nominal levels 0.01 and 0.05 are about (0.0056, 0.0144) and (0.0404, 0.0596), respectively. The evaluated type I error rates of WCHC and other six methods are presented in **Table 1** ($M = 16$) and **Table 2** ($M = 32$). It is observed from these two tables that most of the type I error rates of WCHC are within 95% CIs, which shows the validity of the developed WCHC. Meanwhile, the type I error rates of WCMulP, SHet, MANOVA, MultiPhen, TATES, and OB are not obviously deviated from the nominal levels. See more information in Q–Q plots (**Supplementary Figures 1–6**).

In order to compare powers of these seven methods, we plot power against the genetic effect β (in **Figures 1, 2**) and the within-factor correlation c^2 (in **Figures 3, 4**). Note in the calculation of power, the p -values of WCHC, WCMulP, and SHet are evaluated by 1,000 permutations; the powers of the seven methods are estimated based on 1,000 replicated samples at a significance level of 0.05. The following observations can be drawn from the simulation.

(1) As expected, in each model, the powers of all seven methods increase as the genetic effect β increases (see **Figures 1, 2**). (2) Except in models 1 and 6, WCHC is the most powerful test in all the methods under most of the simulation scenarios (see **Figures 1–4**). (3) As the number of phenotypes increases from $M = 16$ to $M = 32$, WCHC exhibits more obvious advantages over other methods except in Model 1 and 6 (see **Figures 1, 2**). (4) No matter changes of genetic effects β or variations of correlation coefficients between different phenotypes, MANOVA and MultiPhen have the similar performance in all six models. (5) Generally, in each model, the power of all methods decreases with the increase of correlation coefficients of within factors between phenotypes. (6) OB is the most powerful test when the genetic effects are homogeneous (Model 1). However, OB's power decreases when there exist opposite directions (Model 6) or when the genetic variant has an influence on a small proportion of phenotypes (Model 5). (7) In general, WCHC, WCMulP, and TATES are more powerful than SHet, OB, MANOVA, and MultiPhen when the genetic variant affects a portion of phenotypes (Models 2–6). (8) WCHC shows

TABLE 1 | Type I error rates of the seven methods in three simulation settings.

Methods	Type I error rates					
	<i>R</i> = 1		<i>R</i> = 2		<i>R</i> = 4	
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$
WCHC	0.009	0.0505	0.011	0.065	0.007	0.0455
WCmulP	0.0095	0.0495	0.012	0.0595	0.01	0.047
MANOVA	0.013	0.0495	0.0105	0.054	0.0075	0.0555
MultiPhen	0.014	0.0495	0.011	0.055	0.0095	0.0505
TATES	0.0105	0.0465	0.011	0.049	0.007	0.0445
SHet	0.008	0.0515	0.009	0.0535	0.0115	0.0425
OB	0.007	0.045	0.0095	0.055	0.009	0.0475

Sample size $N = 1,000$, the number of phenotypes $M = 16$, $c^2 = 0.5$, $\rho c^2 = 0.1$, and minor allele frequency (MAF) = 0.3. The *p*-values of WCHC, WCmulP, and SHet are estimated by 2,000 permutations. The type I error rates of all seven method are estimated via 2,000 replicated samples at the significance of α . *R* is the number of factors. The number in bold indicates that the type I error rate is out of 95% CIs of the nominal significance level.

TABLE 2 | Type I error rates of the seven methods in three simulation settings.

Methods	Type I error rates					
	<i>R</i> = 1		<i>R</i> = 2		<i>R</i> = 4	
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$
WCHC	0.0095	0.0525	0.0105	0.0525	0.0105	0.0565
WCmulP	0.013	0.0545	0.0135	0.051	0.01	0.05
MANOVA	0.0175	0.072	0.0115	0.059	0.009	0.0535
MultiPhen	0.0155	0.072	0.0115	0.054	0.01	0.055
TATES	0.0105	0.048	0.0115	0.0475	0.011	0.049
SHet	0.0125	0.061	0.0095	0.0475	0.005	0.0415
OB	0.011	0.0555	0.0135	0.0555	0.01	0.044

Sample size $N = 1,000$, the number of phenotypes $M = 32$, $c^2 = 0.5$, $\rho c^2 = 0.1$, and minor allele frequency (MAF) = 0.3. The *p*-values of WCHC, WCmulP, and SHet are estimated by 2,000 permutations. The type I error rates of all seven method are estimated via 2,000 replicated samples at the significance of α . *R* is the number of factors. The number in bold indicates that the type I error rate is out of 95% CIs of the nominal significance level.

obvious advantages over other methods when the genetic variant only affects part of the phenotypes with the same direction. One possible reason is that in the models of generating data, the genetic variant has effects of the same directions on some phenotypes and has no effect on the remaining ones. The hierarchical clustering is capable of grouping similar phenotypes together, so as to reduce the dimensions of association test for improving the power to detect the associated phenotypes.

Overall, from all the power simulation results, we could draw that our proposed WCHC has advantages over other methods in most scenarios, and especially in some scenes, the ascendancy is obvious. In other scenarios, WCHC is comparable with the most powerful test.

REAL DATA ANALYSIS

We applied our proposed method WCHC to the real data analysis from ARIC study (see more details in The ARIC Investigators, 1989). In brief, sponsored by the National Heart, Lung, and Blood Institute (NHLBI), ARIC is a prospective cohort study of atherosclerosis risk in community. It records the

changes of the incidence of atherosclerosis-related diseases and cardiovascular risk factors in distinct races, regions, genders, and time, aiming at investigating the etiology and natural process of atherosclerosis (Morrison et al., 2013). We obtain the genotyped and clinical phenotypic data in ARIC from dbGaP server of the United States National Center for Biotechnology Information (accession number: phs000090.v4.p1).

To evaluate the performance of WCHC in real data, we use the seven methods to analyze obesity-related phenotypes in ARIC. We selected nine continuous traits with regard to obesity including weight, body mass index (BMI), average skinfold thickness of the triceps brachii, mean subscapular skinfold thickness, waist, hip girth, waist-to-hip ratio, calf girth, and wrist breadth and three covariates including age, gender, and race. The specific description of these variables is listed in **Table 3**, and the correlation structure of obesity-related phenotypes is given in **Supplementary Figure 7**. A set of 12,701 subjects across 272,027 SNPs were left for subsequent analysis after excluding subjects with missing data in any of the 12 variables as well as the genetic variants with missing rate greater than 0.2 or HWE $< 10^{-4}$. Every phenotype is adjusted for those three covariates using linear regression model.

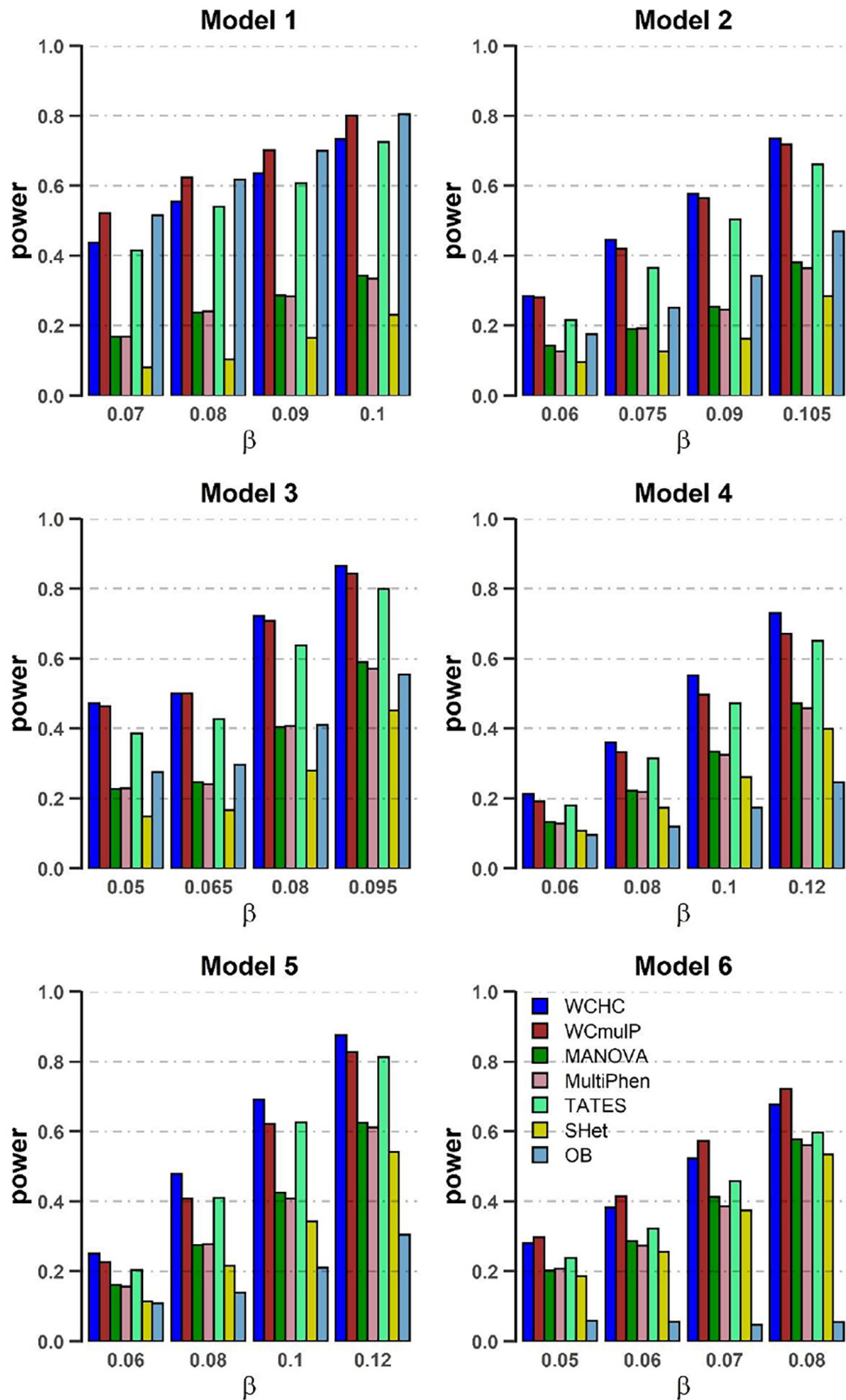


FIGURE 1 | Power comparisons of the seven methods as a function of β in the six models. Sample size is $N = 1,000$, the number of phenotypes is $M = 16$, $c^2 = 0.5$, $\rho c^2 = 0.1$, and $MAF = 0.3$. The power of all the seven methods is estimated by 1000 replicated samples at a significance level of 0.05.

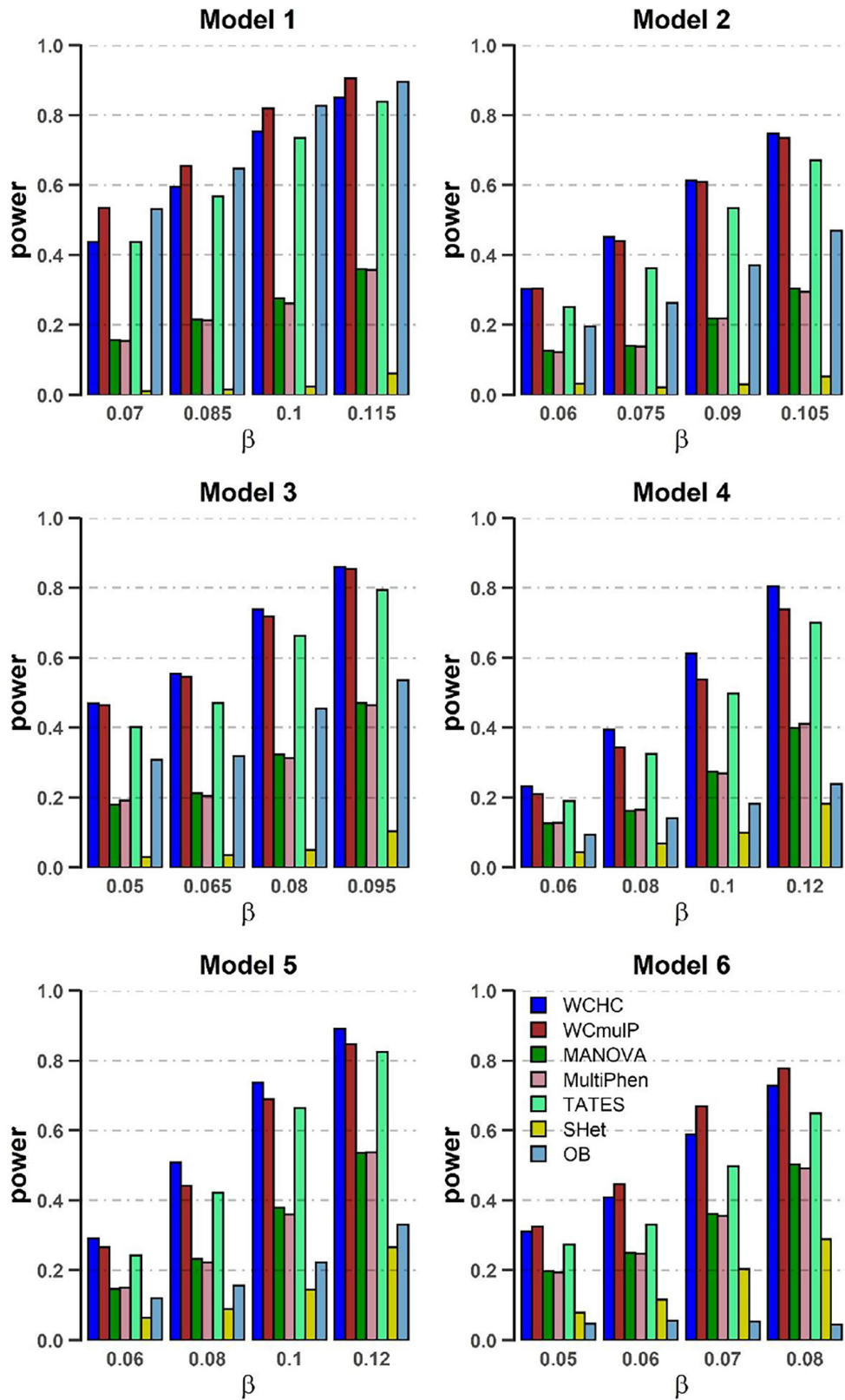


FIGURE 2 | Power comparisons of the seven methods as a function of β in the six models. Sample size is $N = 1,000$, the number of phenotypes is $M = 32$, $c^2 = 0.5$, $\rho c^2 = 0.1$, and MAF = 0.3. The power of all the seven methods is estimated by 1000 replicated samples at a significance level of 0.05.

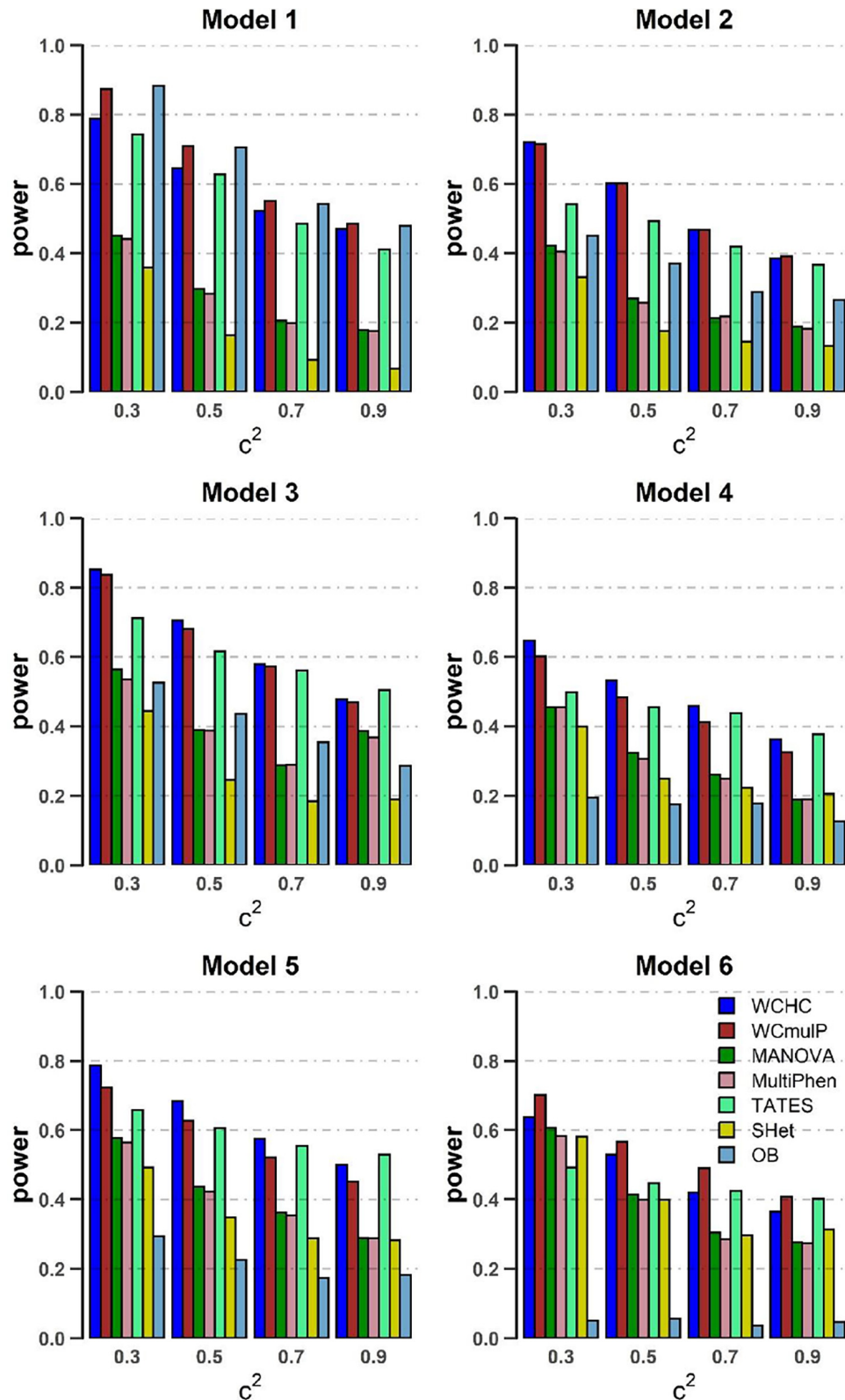


FIGURE 3 | Power comparisons of the seven methods as a function of c^2 in the six models. Sample size is $N = 1,000$, the number of phenotypes is $M = 16$, $\rho c^2 = 0.1$ and $MAF = 0.3$. $\beta = 0.09$ for model 1 and 2; $\beta = 0.08$ for model 3; $\beta = 0.1$ for model 4 and 5; $\beta = 0.07$ for model 6. The power of all the seven methods is estimated by 1000 replicated samples at a significance level of 0.05.

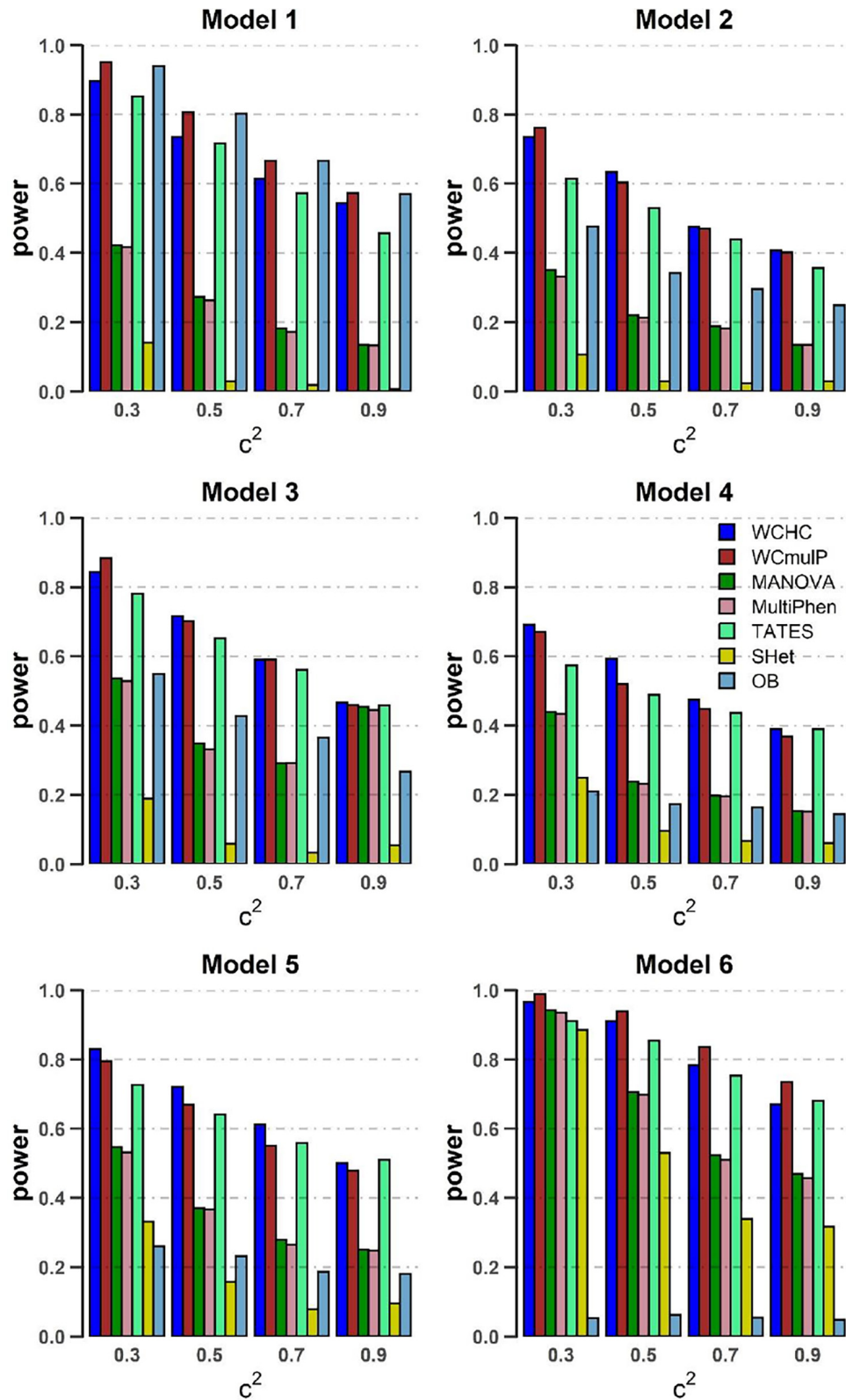


FIGURE 4 | Power comparisons of the seven methods as a function of c^2 in the six models. Sample size is $N = 1,000$, the number of phenotypes is $M = 32$, $\rho c^2 = 0.1$, and MAF = 0.3. $\beta = 0.1$ for model 1 and 4–6; $\beta = 0.09$ for model 2; $\beta = 0.08$ for model 3. The power of all the seven methods is estimated by 1000 replicated samples at a significance level of 0.05.

Based on these adjusted phenotypes related to obesity, we employ WCHC and other six methods to detect associated SNPs. Two groups are obtained after clustering the nine phenotypes in the real data analyses by the hierarchical clustering in WCHC. One of the clusters only includes wrist breadth, while

the other encompasses the remaining phenotypes. Because of multiple testing correction, we adopt the significance threshold of 1×10^{-7} , not the traditional genome-wide significance threshold of 5×10^{-8} . There are totally 11 SNPs that are significant for at least one method (Table 4). Previous studies (Frayling et al., 2007;

TABLE 3 | Summary statistics of obesity-related indexes and covariates in ARIC.

Index	All	Gender			Race		
		Male	Female	p-value	White	Black	p-value
N	12,771	5,704	7,067	–	9,633	3,138	–
Male, %	44.66	–	–	–	47.02	37.44	9.11×10^{-21}
Age, years	54.09 ± 5.73	54.450 ± 5.75	53.76 ± 5.69	6.76×10^{-13}	54.34 ± 5.68	53.34 ± 5.80	5.51×10^{-17}
Weight, lb	173.13 ± 36.85	188.27 ± 31.46	160.92 ± 36.36	$<2.2 \times 10^{-16}$	169.61 ± 35.69	183.99 ± 38.25	1.90×10^{-74}
Weight missing, %	0.149	0.158	0.142	0.995	0.083	0.351	0.002
BMI, kg/m ²	27.66 ± 5.30	27.54 ± 4.18	27.75 ± 6.05	0.020	27.01 ± 4.86	29.65 ± 6.05	9.98×10^{-104}
BMI missing, %	0.149	0.158	0.142	0.995	0.083	0.351	0.002
Triceps, mm	25.26 ± 10.02	19.34 ± 7.87	30.04 ± 8.97	$<2.2 \times 10^{-16}$	24.54 ± 9.08	27.48 ± 12.23	1.73×10^{-34}
Triceps missing, %	0.157	0.175	0.142	0.798	0.093	0.351	0.004
Scapular, mm	24.48 ± 11.59	22.22 ± 9.19	26.31 ± 12.92	1.13×10^{-94}	21.85 ± 9.33	32.59 ± 13.89	1.60×10^{-299}
Scapular missing, %	0.446	0.561	0.354	0.107	0.353	0.733	0.009
WC, cm	96.94 ± 13.83	99.23 ± 10.93	95.09 ± 15.54	1.25×10^{-68}	96.19 ± 13.33	99.25 ± 15.02	5.34×10^{-24}
WC missing, %	0.141	0.123	0.156	0.798	0.104	0.255	0.092
HC, cm	104.55 ± 10.31	102.85 ± 8.09	105.93 ± 11.63	2.81×10^{-68}	103.50 ± 9.478	107.79 ± 11.98	7.52×10^{-72}
HC missing, %	0.141	0.140	0.142	0.999	0.104	0.255	0.092
WHtR	0.926 ± 0.078	0.963 ± 0.054	0.895 ± 0.081	$<2.2 \times 10^{-16}$	0.928 ± 0.079	0.920 ± 0.076	4.66×10^{-8}
WHtR missing, %	0.149	0.140	0.156	0.999	0.114	0.255	0.131
Calf, cm	37.44 ± 3.67	38.06 ± 3.17	36.95 ± 3.95	1.48×10^{-68}	37.39 ± 3.58	37.60 ± 3.93	0.006
Calf missing, %	0.157	0.210	0.113	0.248	0.114	0.287	0.062
Wrist, mm	53.62 ± 5.18	57.78 ± 3.66	50.27 ± 3.53	$<2.2 \times 10^{-16}$	53.59 ± 5.26	53.74 ± 4.91	0.137
Wrist missing, %	0.117	0.123	0.113	0.999	0.073	0.255	0.022

N is the number of subjects; BMI is body mass index; Triceps is average skinfold thickness of the triceps brachii; Scapular is mean subscapular skinfold thickness; WC is waist; HC is hip girth; WHtR is waist-to-hip ratio; Calf is calf girth; and Wrist is wrist breadth. The distributions of normal index is described by mean ± standard deviation; the distributions of non-normal indicators are described by means (25% quantile, 75% quantile). For normal distribution indicators, the differences between groups are evaluated using the t-test (the variances of two groups are homogeneous) or the approximate t-test (the variances of two groups are heterogeneous). For non-normally indicators, Wilcoxon signed rank test is used to test the differences between indicators to get the p-values of differences. For discrete indicators, the chi-square test is used for hypothesis testing and then deriving p-values. Bold number indicates $p < 0.05$. ARIC, Atherosclerosis Risk in Communities.

TABLE 4 | Significant SNPs and the corresponding p-values in the analysis of ARIC.

Chr	SNP	OB	MultiPhen	MANOVA	SHet	TATES	WCMuIP	WCHC
3	rs17017947	1.57×10^{-12}	NA	1.02×10^{-11}	0	0.314	0.513	0.672
10	rs41470552	0.062	NA	6.25×10^{-9}	1.15×10^{-4}	0.035	0.078	0.141
11	rs7927943	0.099	3.33×10^{-6}	5.57×10^{-6}	8.00×10^{-7}	1.16×10^{-8}	1×10^{-7}	1.00×10^{-7}
11	rs1945647	0.038	6.27×10^{-6}	1.2×10^{-5}	7.00×10^{-7}	1.77×10^{-8}	0	1.00×10^{-7}
11	rs7114045	3.73×10^{-5}	5.47×10^{-6}	5.66×10^{-8}	0.003	0.018	0.051	0.108
12	rs7968682	0.414	5.36×10^{-8}	8.34×10^{-8}	0	0.018	0.079	0.006
16	rs9939609	0.082	3.39×10^{-8}	1.85×10^{-8}	0	2.97×10^{-10}	0	1.00×10^{-7}
16	rs8050136	0.186	8.66×10^{-8}	4.29×10^{-8}	0	2.86×10^{-9}	0	1.00×10^{-7}
20	rs201561	0.138	2.91×10^{-6}	2.48×10^{-6}	6.30×10^{-6}	7.99×10^{-7}	0.035	1.00×10^{-7}
20	rs1570004	0.184	7.77×10^{-5}	5.28×10^{-5}	1.90×10^{-6}	6.12×10^{-8}	0.001	1.00×10^{-7}
20	rs1014883	0.457	3.00×10^{-5}	1.83×10^{-5}	1.60×10^{-6}	3.19×10^{-4}	0.011	1.07×10^{-4}

The p-values of WCHC, SHet, and WCMuIP are estimated using 10^7 permutations; the p-values of OB, MultiPhen, MANOVA, and TATES are estimated using asymptotic distributions; and the bold number indicates p-value $\leq 1 \times 10^{-7}$. HWE $\geq 10^{-4}$ for the identified SNPs; "NA" indicates MultiPhen is not available, as the genotype at the specified SNP does not take all three values of 0, 1, and 2 in these data. **0** indicates p-value $< 10^{-7}$. SNP, single-nucleotide polymorphism; ARIC, Atherosclerosis Risk in Communities; HWE, Hardy-Weinberg equilibrium.

Heard-Costa et al., 2009; Lindgren et al., 2009; Meyre et al., 2009; Thorleifsson et al., 2009; Willer et al., 2009; Heid et al., 2010; Speliotes et al., 2010; Bradfield et al., 2012; Wen et al., 2012; Berndt et al., 2013; Monda et al., 2013; Locke et al., 2015; Shungin et al., 2015) have reported that *FTO* leads to obesity through population studies and experimental researches elaborating relevant mechanisms. Among the 11 identified SNPs, rs9939609 and rs8050136 are involved in *FTO*. Additionally, rs7968682 is reported to be associated with height (Yang T. L. et al., 2010; Takeshita et al., 2011). Few other SNPs have been assessed to explore the association with obesity or obesity-related phenotypes. From **Table 4**, we can see that both WCHC and MANOVA identified six SNPs; TATES identified five SNPs; both WCMulP and SHet identified four SNPs; MultiPhen identified three SNPs; and OB only identified one SNP, which may be due to that the true genetic effects of most of SNPs are heterogeneous for all phenotypes. In summary, the number of SNPs identified by WCHC is comparable with the largest number of SNPs identified by other tests. These real data analysis results are consistent with our simulation results.

CHARACTERISTICS OF THE SIGNIFICANT VARIANTS

Table 5 shows the annotations of the identified SNPs based on the Ensemble website¹ and SCAN website². From **Table 5**, we can see that the significant SNPs are located in intergenic or intron region, and most of them have been reported to be associated with BMI, height, or T2D. Generally, they have been reported in GWAS. We could also explore the expression of genes associated with the significant SNPs, although they are located in intergenic or intron region. Therefore, we make full use of Qtlizer³, eQTLGen⁴, and PsychENCODE⁵, which are the largest integrating various tissues, blood, and brain expression Quantitative Trait Locus (eQTL) samples, respectively. Instead of restricting analysis to the SNPs in **Table 5**, we considered using a larger list of SNPs with proxy variants, which are in Linkage Disequilibrium (LD) with the SNPs in **Table 5** ($r^2 \geq 0.8$) via Qtlizer website. We restricted the eQTL association criteria with False Discovery rate (FDR) < 0.05. The results of eQTLs in Qtlizer, eQTLGen, and PsychENCODE are displayed in **Supplementary Data Sheets**.

In order to further study the biology function of the genetic variants, we performed enrichment analysis on genes associated with these 11 SNPs in **Table 5** and the proxy variants (see qtlizer.results in Supplementary Data Sheet) in the three websites/consortiums (Qtlizer, eQTLGen, and PsychENCODE). After summarizing all genes in the three tables (see qtlizer.results, eQTLGen.results, and PsychENCODE.results

in **Supplementary Data Sheet**), we got all the genes associated with the eQTLs (see Gene sheet in Supplementary Data Sheet). A total of 76 genes were obtained to do the gene set analysis by virtue of different biological databases for investigating biological processes, cell components, molecular functions, metabolic pathways, phenotypes with relevant diseases, and protein interactions. The results of enrichment analysis and protein-protein interaction (PPI) are given in **Figures 5, 6**. According to the Gene Ontology (GO) enrichment analysis chart in **Figure 5**, GO items mainly focus on the cellular response to hydrogen and regulation of lipid kinase activity, which may be parts of the metabolic process. Moreover, the Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathway in **Figure 6** presents that the enriched genes possess taurine and hypotaurine metabolism and endocrine resistance, which indicates that the obesity-related variants detected by WCHC and other methods in ARIC might be involved in the metabolic pathways, which releases the signal that our results are in a certain degree of credibility. Subsequently, we draw a PPI network diagram through the STRING⁶ to reveal that most of the proteins (nodes) encoded by the genes have certain interactions (edges), which suggests the proteins related the expression of genes might interact with each other for controlling a variety of biological phenomena including endocrine development, cellular response to hydrogen, and metabolic processes.

Overall, our results showed that WCHC and other six methods could identify significant genetic variants for obesity phenotypes in real data analysis from ARIC. More importantly, functional annotations of genetic variants and enrichment analysis support that the variants are closely related to biological functions and metabolic pathways of obesity.

DISCUSSION

In this article, we proposed WCHC to perform multivariate analysis of multiple phenotypes in association studies due to the following reasons. (1) Multiple correlated phenotypes are usually measured in complex disease for genetic association studies. Compared to univariate analysis, multivariate analysis considers multidimensional structure information. It indicates certain variance-covariance is included in multiple phenotypes. (2) Association analysis of multiple phenotypes separately cannot present genetic interactions between phenotypes. More and more evidence reveals that joint analysis of multiple related phenotypes, which considers the interactions between phenotypes comprehensively, can boost the power of detecting genetic variants associated with complex diseases. No matter whether the effects of genetic variants on phenotypes are consistent or not, WCHC provides a relatively simple way to incorporate the correlations between phenotypes into analysis. (3) Actually, we are not sure which phenotype or linear combination of phenotypes is more likely to elucidate the genetic structure of complex diseases. WCHC adopts clustering

¹<https://asia.ensembl.org>

²<http://scandb.org>

³<http://genehopper.de/qtlizer>

⁴<https://www.eqtlgen.org/>

⁵<http://resource.psychencode.org/>

⁶<https://www.string-db.org/>

TABLE 5 | Characteristics of the significant SNPs.

SNPs	Chr.	Position (GRCh38)	Alleles (Alt/Ref)	Gene (nearest)	Feature	Expression genes	Reported (yes/no)	Reported phenotypes	GWAS references
rs17017947	3	276,171	A/C	<i>CHL1</i>	Intron	–	No	–	–
rs41470552	10	102,222,133	T/G	<i>PITX3</i>	Intergenic	–	No	–	–
rs7114045	11	101,479,689	C/T	<i>TRPC6</i>	Intron	–	No	–	–
rs1945647	11	81,602,715	C/T	<i>MTND6P25</i>	Intergenic	<i>GNAI2, STK40, LIMK1, LIG4, HLTf, ZNF511, CBL1, NUDT17, POLR3C, DAGLB, KDELR2, NUP93, PRCC, C16orf80, RAB33B, LRP8</i>	No	–	–
rs7927943	11	81,637,194	C/T	<i>MTND6P25</i>	Intergenic	<i>WSCD2, GNAI2, ZFH3, NUP93, FAM60A, LIMK1, MAP4, FLJ31958, LIG4, HLTf</i>	No	–	–
rs7968682	12	65,978,100	G/T	<i>HMG2</i>	Intergenic	–	Yes	Height, birth weight	–
rs8050136	16	53,782,363	C/A	<i>FTO</i>	Intron	<i>HES7, LATS2</i>	Yes	BMI, T2D, adiposity	–
rs9939609	16	53,786,615	T/A	<i>FTO</i>	Intron	<i>CR1, CR1L, ZNRF1, ANKRD50, LATS2, TSPYL4, HES7</i>	Yes	BMI, T2D	–
rs1014883	20	21,863,992	A/G	<i>RPL41P1</i>	Intergenic	<i>ANTXR2</i>	No	–	–
rs1570004	20	35,370,450	A/T	<i>UQCC</i>	Intron	–	Yes	Height	–
rs201561	20	22,018,575	G/C	<i>RPL41P1</i>	Intergenic	<i>P2RX3, EHD4</i>	Yes	Balding type 1	–

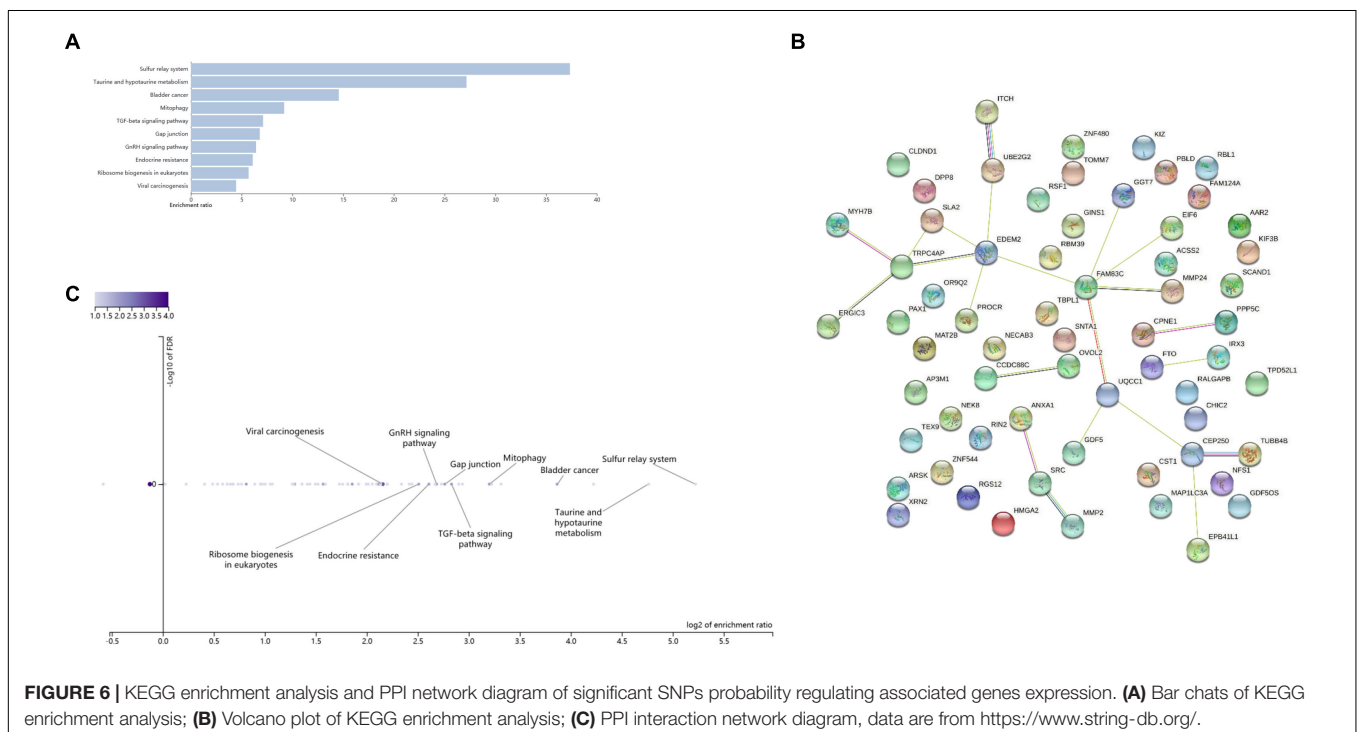
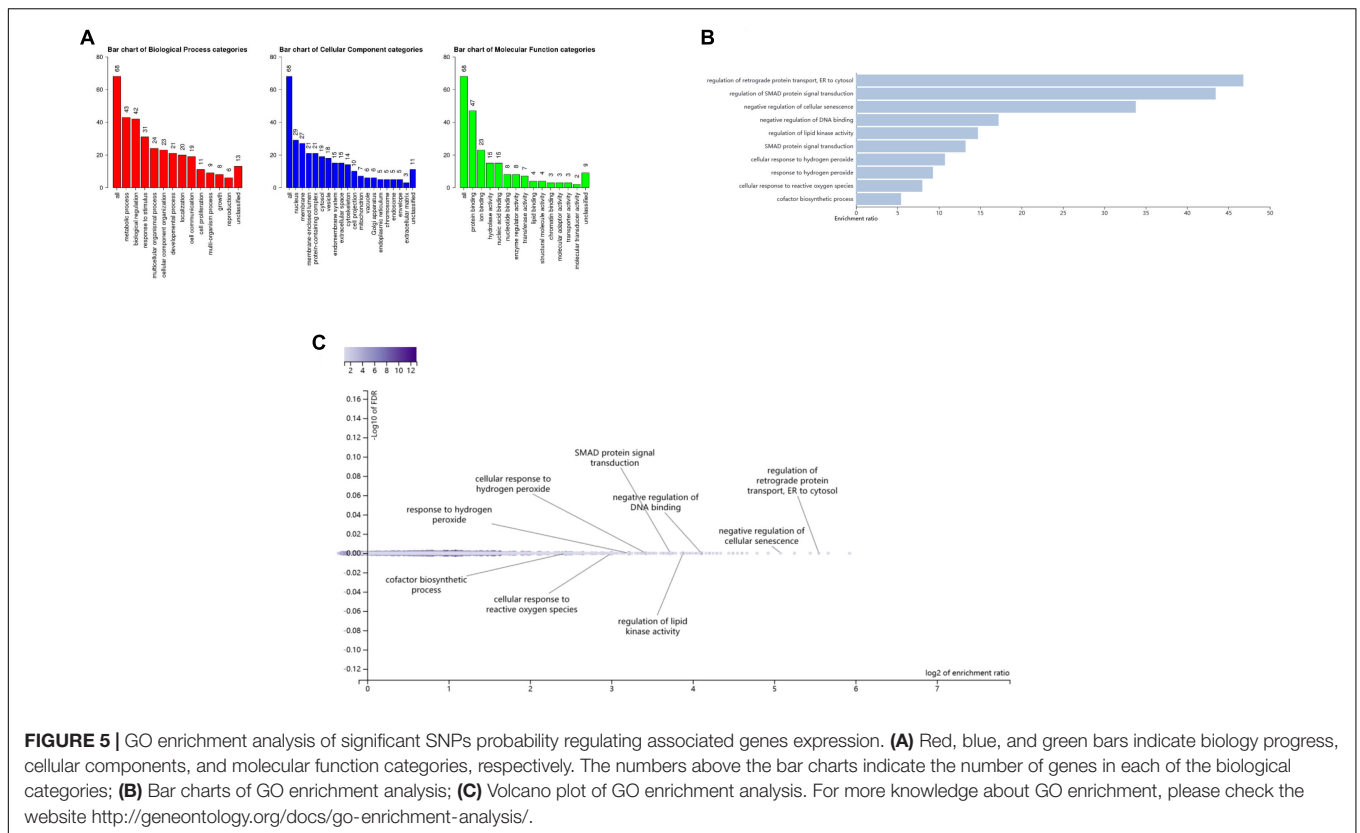
Annotations are from Ensemble website (<https://asia.ensembl.org>) and SCAN website (<http://scandb.org>); intron denotes the SNP is located between exons; intergenic denotes the SNP is located between genes. Expression genes denotes annotations added after analysis of transcriptional levels of eQTL in cell lines from HapMap CEU and YRI samples using Affymetrix human exon 1.0 ST array. GWAS references refer to the identifications of PubMed. SNP, single-nucleotide polymorphism; GWAS, genome-wide association study; eQTL, expression quantitative trait locus.

approach and linear combination of multiple phenotypes to account for the complex genetic information, which not only takes the similarity between phenotypes into consideration but also considers the heterogeneity, so it is helpful to explore the genetic mechanism of diseases.

Our results manifested that WCHC has correct type I error rates and is either the most powerful test or comparable with the most powerful tests among the seven methods we adopted. None of the other methods observes consistently good performances under the simulation scenarios. OB is the most powerful test when the genetic effects are homogeneous, while it loses power dramatically when genetic effects are heterogeneous, especially if there exists opposite directions of genetic effects. In most simulation scenarios, SHet, MANOVA, MultiPhen, and TATES have similar powers, and they are less powerful than WCHC, and WCHC is more powerful when the genetic variant influences a part of phenotypes. However, WcmulP is less powerful in this scenario. Furthermore, in real data analysis, WCHC and MANOVA identified the largest number of significant SNPs (six SNPs). Therefore, the real data analysis results demonstrate that WCHC has excellent performance in detecting SNPs associated with complex disease with multiple related phenotypes such as obesity. As for the

methods giving such different results when applied to the real ARIC data, we think that the parametric information of real data is unknown for us. Therefore, we may try various methods to analyze the real data for getting reliable results as much as possible. In our opinion, no method can guarantee 100% accuracy. We can only be cautious to say that the significant loci are more likely to be true signals, but further verification is still needed.

In the context of association studies, population stratification (PS) refers to allele frequency difference between populations uncorrelated to the outcome of interest, but due to systematic ancestry differences. PS may cause confounding effects seriously if not adjusted properly (Knowler et al., 1988; Lander and Schork, 1994). Methods such as principal component analysis (PCA) (Zhu et al., 2002; Chen et al., 2003; Zhang et al., 2003; Price et al., 2006; Bauchet et al., 2007), linear mixed model (LMM) (Kang et al., 2010; Zhang et al., 2010; Hoffman, 2013), multidimensional scaling (MDS) (Li and Yu, 2008), robust PCA based on resampling by half means (RPCA-RHM) (Liu et al., 2013), and robust PCA based on the projection pursuit (RPCA-PP) (Liu et al., 2013) can be used to adjust for PS. We propose to apply PCA to control for PS when samples from different populations are involved.



In real data analysis, as the number of phenotypes elevates, the chance of missing at least one subject increases exponentially, especially in epidemiological and clinical

research (Ali et al., 2011; Dahl et al., 2016). We removed 412 subjects with missing either phenotypes or covariates from 13,113 observations. It is worth noting that the sample

mean substitution (Ali et al., 2011; van der Sluis et al., 2013) is a simple, unconditional method that does not depend on other variables, which is a common strategy replacing the missing values with plausible values for the variable with missing values. However, it may contribute to biased estimates where data are not missing completely at random (Ali et al., 2011). Additionally, imputation is a more complicated approach that fills in missing values with estimated values via model-based methods or conditional imputation, comprising multiple imputation (MI), multivariate normal imputation (MVNI), and fully conditional specification (FCS) (Raghunathan et al., 2001; Buuren et al., 2006; De Silva et al., 2017).

One weakness of WCHC is that the test statistic does not have an asymptotic distribution and its p -value needs to be calculated by permutation procedure, which is time-consuming as compared with approaches whose test statistics have asymptotic distributions. To conduct GWAS, a small number of permutations (e.g., 1,000) can be used to select genetic variants that reveal evidence of association, and then a large number of permutations are employed to estimate the selected significant genetic variants. We adopted this strategy to analyze the real dataset. Consequently, it seems to be efficient, and the bioinformatics analysis of significant variants supports our results. In conclusion, in the field of genotype–phenotype association studies, WCHC is an effective method for association analysis of multiple phenotypes, which considers both the correlations and differences among the multiple phenotypes. WCHC provides a convenient approach of association analysis for researchers to discover potential genes causing complex diseases, which does not need to assume the genetic model, and there is no limit to the number of phenotypes. Because the genetic structure of phenotypes is usually unknown, WCHC provides a convenient statistical method for the application of massive multi-phenotypic data in the future.

REFERENCES

- Ali, A. M., Dawson, S. J., Blows, F. M., Provenzano, E., Ellis, I. O., Baglietto, L., et al. (2011). Comparison of methods for handling missing data on immunohistochemical markers in survival analysis of breast cancer. *Br. J. Cancer* 104, 693–699. doi: 10.1038/sj.bjc.6606078
- Aschard, H., Vilhjálmsson, B. J., Greliche, N., Morange, P.-E., Tréguoët, D.-A., and Kraft, P. (2014). Maximizing the power of principal-component analysis of correlated phenotypes in genome-wide association studies. *Am. J. Hum. Genet.* 94, 662–676. doi: 10.1016/j.ajhg.2014.03.016
- Bauchet, M., McEvoy, B., Pearson, L. N., Quillen, E. E., Sarkisian, T., Hovhannesian, K., et al. (2007). Measuring European population stratification with microarray genotype data. *Am. J. Hum. Genet.* 80, 948–956. doi: 10.1086/513477
- Berndt, S. I., Gustafsson, S., Mägi, R., Ganna, A., Wheeler, E., Feitosa, M. F., et al. (2013). Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat. Genet.* 45, 501–512. doi: 10.1038/ng.2606
- Bradfield, J. P., Taal, H. R., Timpson, N. J., Scherag, A., Lecoeur, C., Warrington, N. M., et al. (2012). A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat. Genet.* 44, 526–531. doi: 10.1038/ng.2247

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: phs000090.v4.p1. The source code for WCHC method can be found in <https://github.com/YQHuFD/WCHC>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ARIC. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LF and Y-QH: study concept and design and drafting of the manuscript. LF, Y-QH, and YW: acquisition of data. LF: methodology and interpretation of data. LF, YW, TL, and Y-QH: critical revision of the manuscript for important intellectual content. All authors have read and approved the final version of manuscript. We thank all reviewers and editors for their valuable suggestions on revision.

FUNDING

This study was supported by grants to Y-QH from the National Natural Science Foundation of China (grant nos. 11971117 and 11571082).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.654804/full#supplementary-material>

- Bühlmann, P., Rutimann, P., van de Geer, S., and Zhang, C. H. (2013). Correlated variables in regression: clustering and sparse estimation. *J. Stat. Plan Infer.* 143, 1835–1858. doi: 10.1016/j.jspi.2013.05.019
- Buuren, S., Brand, J., Groothuis-Oudshoorn, C., and Rubin, D. (2006). Fully conditional specification in multivariate imputation. *J. Stat. Comput. Simul.* 76, 1049–1064. doi: 10.1080/10629360600810434
- Casale, F. P., Rakitsch, B., Lippert, C., and Stegle, O. (2015). Efficient set tests for the genetic analysis of correlated traits. *Nat. Methods* 12, 755–758. doi: 10.1038/nmeth.3439
- Chen, H. S., Zhu, X., Zhao, H., and Zhang, S. (2003). Qualitative semi-parametric test for genetic associations in case-control designs under structured populations. *Ann. Hum. Genet.* 67, 250–264. doi: 10.1046/j.1469-1809.2003.00036.x
- Cole, D. A., Maxwell, S. E., Arvey, R., and Salas, E. (1994). How the power of MANOVA can both increase and decrease as a function of the intercorrelations among the dependent variables. *Psychol. Bull.* 115, 465–474.
- Dahl, A., Iotchkova, V., Baud, A., Johansson, Å, Gyllensten, U., Soranzo, N., et al. (2016). A multiple-phenotype imputation method for genetic studies. *Nat. Genet.* 48, 466–472. doi: 10.1038/ng.3513
- De Silva, A. P., Moreno-Betancur, M., De Livera, A. M., Lee, K. J., and Simpson, J. A. (2017). A comparison of multiple imputation methods for handling missing

- values in longitudinal data in the presence of a time-varying covariate with a non-linear association with time: a simulation study. *BMC Med. Res. Methodol* 17:114. doi: 10.1186/s12874-017-0372-y
- Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316, 889–894. doi: 10.1126/science.1141634
- Fu, L. (2020). *Multi-Phenotype Association studies and Mendelian Randomization for Pleiotropic Genetic Variants and their Applications*. Ph.D. thesis. Shanghai: Fudan University.
- Gavish, B., Ben-Dov, I. Z., and Bursztyn, M. (2008). Linear relationship between systolic and diastolic blood pressure monitored over 24 h: assessment and correlates. *J. Hypertens.* 26, 199–209. doi: 10.1097/HJH.0b013e3282f25b5a
- Heard-Costa, N. L., Zillikens, M. C., Monda, K. L., Johansson, A., Harris, T. B., Fu, M., et al. (2009). NRXN3 is a novel locus for waist circumference: a genome-wide association study from the charge consortium. *PLoS Genet.* 5:e1000539. doi: 10.1371/journal.pgen.1000539
- Heid, I. M., Jackson, A. U., Randall, J. C., Winkler, T. W., Qi, L., Steinthorsdottir, V., et al. (2010). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* 42, 949–960. doi: 10.1038/ng.685
- Hoffman, G. E. (2013). Correcting for population structure and kinship using the linear mixed model: theory and extensions. *PLoS One* 8:e75707. doi: 10.1371/journal.pone.0075707
- Huang, P. L. (2009). A comprehensive definition for metabolic syndrome. *Dis. Model. Mech.* 2, 231–237. doi: 10.1242/dmm.001180
- Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-Y., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548
- Knowler, W. C., Williams, R. C., Pettitt, D. J., and Steinberg, A. G. (1988). Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. *Am. J. Hum. Genet.* 43, 520–526.
- Korte, A., Vilhjálmsson, B. J., Segura, V., Platt, A., Long, Q., and Nordborg, M. (2012). A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat. Genet.* 44, 1066–1071. doi: 10.1038/ng.2376
- Kwak, I. Y., and Pan, W. (2016). Adaptive gene- and pathway-trait association testing with GWAS summary statistics. *Bioinformatics* 32, 1178–1184. doi: 10.1093/bioinformatics/btv719
- Lander, E. S., and Schork, N. J. (1994). Genetic dissection of complex traits. *Science* 265, 2037–2048.
- Li, Q., and Yu, K. (2008). Improved correction for population stratification in genome-wide association studies by identifying hidden population structures. *Genet. Epidemiol.* 32, 215–226. doi: 10.1002/gepi.20296
- Liang, X., Wang, Z., Sha, Q., and Zhang, S. (2016). An adaptive Fisher's combination method for joint analysis of multiple phenotypes in association studies. *Sci. Rep.* 6:34323. doi: 10.1038/srep34323
- Lindgren, C. M., Heid, I. M., Randall, J. C., Lamina, C., Steinthorsdottir, V., Qi, L., et al. (2009). Genome-wide association scan meta-analysis identifies three loci influencing adiposity and fat distribution. *PLoS Genet.* 5:e1000508. doi: 10.1371/journal.pgen.1000508
- Liu, L., Zhang, D., Liu, H., and Arendt, C. (2013). Robust methods for population stratification in genome wide association studies. *BMC Bioinformatics* 14:132. doi: 10.1186/1471-2105-14-132
- Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518, 197–206. doi: 10.1038/nature14177
- Meyre, D., Delplanque, J., Chèvre, J.-C., Lecoq, C., Lobbens, S., Gallina, S., et al. (2009). Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat. Genet.* 41, 157–159. doi: 10.1038/ng.301
- Monda, K. L., Chen, G. K., Taylor, K. C., Palmer, C., Edwards, T. L., Lange, L. A., et al. (2013). A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. *Nat. Genet.* 45, 690–696. doi: 10.1038/ng.2608
- Morrison, A. C., Voorman, A., Johnson, A. D., Liu, X., Yu, J., Li, A., et al. (2013). Whole-genome sequence-based analysis of high-density lipoprotein cholesterol. *Nat. Genet.* 45, 899–901. doi: 10.1038/ng.2671
- O'Brien, P. C. (1984). Procedures for comparing samples with multiple endpoints. *Biometrics* 40, 1079–1087.
- O'Reilly, P. F., Hoggart, C. J., Pomyen, Y., Calboli, F. C., Elliott, P., Jarvelin, M. R., et al. (2012). MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. *PLoS One* 7:e34861. doi: 10.1371/journal.pone.0034861
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909. doi: 10.1038/ng1847
- Raghunathan, T., Lepkowski, J., Hoewyk, J., and Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Surv. Methodol* 27, 85–95.
- Shungin, D., Winkler, T. W., Croteau-Chonka, D. C., Ferreira, T., Locke, A. E., Mägi, R., et al. (2015). New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 518, 187–196. doi: 10.1038/nature14132
- Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A. U., et al. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* 42, 937–948. doi: 10.1038/ng.686
- Takeshita, H., Fujihara, J., Soejima, M., Koda, Y., Kimura-Kataoka, K., Ono, R., et al. (2011). Confirmation that SNPs in the high mobility group-a2 gene (HMGA2) are associated with adult height in the Japanese population; wide-ranging population survey of height-related SNPs in HMGA2. *Electrophoresis* 32, 1844–1851. doi: 10.1002/elps.201100128
- The ARIC Investigators (1989). The atherosclerosis risk in communities (ARIC) study: design and objectives. *Am. J. Epidemiol.* 129, 687–702.
- Thorleifsson, G., Walters, G. B., Gudbjartsson, D. F., Steinthorsdottir, V., Sulem, P., Helgadóttir, A., et al. (2009). Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat. Genet.* 41, 18–24. doi: 10.1038/ng.274
- van der Sluis, S., Posthuma, D., and Dolan, C. V. (2013). TATES: efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLoS Genet.* 9:e1003235. doi: 10.1371/journal.pgen.1003235
- Wang, Z., Sha, Q., and Zhang, S. (2016). Joint analysis of multiple traits using “optimal” maximum heritability test. *PLoS One* 11:e0150975. doi: 10.1371/journal.pone.0150975
- Wen, W., Cho, Y.-S., Zheng, W., Dorajoo, R., Kato, N., Qi, L., et al. (2012). Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat. Genet.* 44, 307–311. doi: 10.1038/ng.1087
- Willer, C. J., Speliotes, E. K., Loos, R. J. F., Li, S., Lindgren, C. M., Heid, I. M., et al. (2009). Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.* 41, 25–34. doi: 10.1038/ng.287
- Yan, T., Li, Q., Li, Y., Li, Z., and Zheng, G. (2013). Genetic association with multiple traits in the presence of population stratification. *Genet. Epidemiol.* 37, 571–580. doi: 10.1002/gepi.21738
- Yang, J. J., Li, J., Williams, L. K., and Buu, A. (2016). An efficient genome-wide association test for multivariate phenotypes based on the Fisher combination function. *BMC Bioinformatics* 17:19. doi: 10.1186/s12859-015-0868-6
- Yang, Q., and Wang, Y. (2012). Methods for analyzing multivariate phenotypes in genetic association studies. *J. Probab. Stat.* 2012:652569. doi: 10.1155/2012/652569
- Yang, Q., Wu, H., Guo, C. Y., and Fox, C. S. (2010). Analyze multivariate phenotypes in genetic association studies by combining univariate association tests. *Genet. Epidemiol.* 34, 444–454. doi: 10.1002/gepi.20497
- Yang, T. L., Guo, Y., Zhang, L. S., Tian, Q., Yan, H., Guo, Y. F., et al. (2010). HMGA2 is confirmed to be associated with human adult height. *Ann. Hum. Genet.* 74, 11–16. doi: 10.1111/j.1469-1809.2009.00555.x
- Zeger, S. L., and Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42, 121–130.
- Zhang, S., Zhu, X., and Zhao, H. (2003). On a semiparametric test to detect associations between quantitative traits and candidate genes using unrelated individuals. *Genet. Epidemiol.* 24, 44–56. doi: 10.1002/gepi.10196
- Zhang, Y., Xu, Z., Shen, X., and Pan, W. (2014). Testing for association with multiple traits in generalized estimation equations, with application to

- neuroimaging data. *Neuroimage* 96, 309–325. doi: 10.1016/j.neuroimage.2014.03.061
- Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546
- Zhou, X., and Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* 11, 407–409. doi: 10.1038/nmeth.2848
- Zhu, H., Zhang, S., and Sha, Q. (2018). A novel method to test associations between a weighted combination of phenotypes and genetic variants. *PLoS One* 13:e0190788. doi: 10.1371/journal.pone.0190788
- Zhu, X., Feng, T., Tayo, B. O., Liang, J., Young, J. H., Franceschini, N., et al. (2015). Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *Am. J. Hum. Genet.* 96, 21–36. doi: 10.1016/j.ajhg.2014.11.011
- Zhu, X., Zhang, S., Zhao, H., and Cooper, R. S. (2002). Association mapping, using a mixture model for complex traits. *Genet. Epidemiol.* 23, 181–196. doi: 10.1002/gepi.210

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Fu, Wang, Li and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.