Title

A general statistic to test an optimally weighted combination of common and/or rare variants.

Permalink

https://escholarship.org/uc/item/3rq6r6mb

Journal

Genetic Epidemiology, 43(8)

Authors

Zhang, Jianjun
Wu, Baolin
Sha, Qiuying
et al.

Publication Date

2019-12-01

DOI

10.1002/gepi.22255

Peer reviewed

A General Statistic to Test an Optimally Weighted Combination of Common and/or Rare Variants

Jianjun Zhang[1], **Baolin Wu**[2], **Qiuying Sha**[3], **Shuanglin Zhang**[3], **Xuexia Wang**[1,*]

[1]Department of Mathematics, University of North Texas;

[2]Division of Biostatistics, School of Public Health, University of Minnesota;

[3]Department of Mathematical Sciences, Michigan Technological University

Abstract

Both genome-wide association study and next generation sequencing data analyses are widely employed in order to identify disease susceptible common and/or rare genetic variants. Rare variants generally have large effects though they are hard to detect due to their low frequencies. Currently, many existing statistical methods for rare variants association studies employ a weighted combination scheme, which usually puts subjective weights or suboptimal weights based on some adhoc assumptions (e.g. ignoring dependence between rare variants). In this study, we analytically derived optimal weights for both common and rare variants and proposed a General and novel approach to Test association between an Optimally Weighted combination of variants (G-TOW) in a gene or pathway for a continuous or dichotomous trait while easily adjusting for covariates. Results of the simulation studies show that G-TOW has properly controlled type I error rates and it is the most powerful test among the methods we compared when testing effects of either both rare and common variants or rare variants only. We also illustrate the effectiveness of G-TOW using the Genetic Analysis Workshop 17 (GAW17) data. Additionally, we applied G-TOW and other competitive methods to test disease associated genes in a real data of schizophrenia. The G-TOW has successfully verified genes FYN and VPS39 which are associated with schizophrenia reported in existing publications. Both of these genes are missed by the weighted sum statistic (WSS) and the sequence kernel association test (SKAT). Simulation study and real data analysis indicate that G-TOW is a powerful test.

Keywords

rare variants; common variants; optimal weights; association studies

Introduction

Genome-wide association studies (GWAS) have identified numerous common variants that are associated with complex human traits (Heid et al., 2010; Allen et al. 2010; Plenge et al.,

---

*Corresponding author: Xuexia Wang, Associate Professor in Statistics, Department of Mathematics, University of North Texas, 1155 Union Circle Denton, TX 76201. Xuexia.Wang@unt.edu.

2007; Saxena et al., 2007; Thomson et al., 2007; Zeggini et al., 2007). However, common variants that have been identified through GWAS can only explain small fractions of estimated trait heritabilities (Bansal et al., 2010; McCarthy et al., 2008; Schork et al., 2009). Rare variants may play an important role in studying the etiology of complex human diseases. Next-generation sequencing technology (Cirulli and Goldstein, 2010) provides potential opportunities to find the missing heritability due to rare genetic variants (Cohen et al., 2006; Ji et al., 2008; Manolio et al., 2009; Marini et al., 2008; Nejentsev et al., 2009; Zhu et al., 2010) since the new sequencing technology makes directly testing for rare variants possible (Andrés et al., 2007).

Due to allelic heterogeneity and the extreme rarity of individual variants, statistical methods used to detect common variants may potentially be unsuccessful when applied to detect rare variants unless sample sizes or effect sizes are very large (Li and Leal, 2008). Recently, several statistical methods for detecting associations of rare variants have been developed. Most of these methods can be divided into two classes: burden tests and quadratic tests. Burden tests collapse rare variants in a genomic region into a single burden variable by using weighted combination and then testing the association between a phenotype and the single burden variable. Choosing appropriate weights is critical to the performance of the burden tests. Different weights reflect different burden test methods, such as the cohort allelic sums test (CAST) (Morgenthaler and Thilly, 2007), the combined multivariate and collapsing (CMC) method (Li and Leal, 2008), the weighted sum statistic (WSS) (Madsen and Browning, 2009), and the variable threshold (VT) method (Price et al., 2010). However, burden tests could suffer a loss of statistical power if different directions of effects exist at causal variants. Quadratic tests include tests with statistics of quadratic forms of the score vector such as the sequence kernel association test (SKAT) (Wu et al., 2011), data-adaptive sum (aSUM) (Han and Pan, 2010), adaptive weighting (AW) methods (Sha, Wang, and Zhang, 2013), and weighted sum of squared score (SSU) method (Pan 2009). Even though quadratic tests are robust to the directions of the effects of causal variants, they could be less powerful due to the large number of degrees of freedom of the test statistics when we consider large amounts of rare variants, because the degree of freedom of the quadratic tests depend on the number of rare variants where the quadratic tests usually follows a chi-square distribution with degree freedom related to the number of variants in the considered region.

Sha et al., (2012) proposes a method to detect disease associated rare variants by testing an optimally weighted combination of variants (TOW), in which they derives optimal weights by maximizing the score test statistic. There are similarities between the TOW method and the burden tests regarding weights. The TOW method chooses optimal weights under a certain criterion, whereas the traditional burden test methods give a subjective weight directly. The TOW can also be viewed as a quadratic test from the form of the test statistic. Sha et al., (2012) points out that TOW is related to the sequence kernel association test (SKAT) proposed by Wu et al., (2011) and the weighted sum of squared score (SSU) proposed by Pan (2009), which are both quadratic tests. This property means that the TOW method can combine the advantages of both burden tests and quadratic tests, making it robust to the directions of effects of causal variants and the large number of rare variants.

Due to low minor allele counts of rare causal variants, rare variants are often assumed to be independent, either implicitly or explicitly in the development of association test methods. The TOW method is derived under the independent assumption among rare variants. However, simulations based on coalescent models have shown that two variants with low MAFs are likely to be negatively correlated with a negative linkage disequilibrium (LD) value (Hudson, 1985), and ignoring such correlations can reduce the power of aggregation based rare variant association tests (Kinnamon, Hershberger, and Martin, 2012). Talluri and Shete (2013) propose a method for detecting rare variants that utilizes the information of LD between variants (measured by R^2) to select the best subset of rare variants, subsequently improving the power of detecting association. Turkmen and Lin (2017) show that two commonly used LD measures are not capable of detecting LD when rare variants are involved and proposed a new LD measure method.

In this paper, we develop a General and novel approach to Test association between an Optimally Weighted combination of variants (G-TOW) between common and/or rare variants in a gene or pathway and a complex trait of interest. The optimal weights are analytically derived and can be estimated from sampled genotypes and phenotypes under a certain criterion while we consider the correlations among variants in the considered region simultaneously. G-TOW can be applied to both quantitative and qualitative traits, allows covariates, and is robust to directions of effects of causal variants. Extensive simulation studies and applications to the Genetic Analysis Workshop 17 (GAW17) data were used to compare the performance of G-TOW with that of three existing methods (TOW, WSS, and SKAT). Results show that G-TOW is a valid method and consistently more powerful than the other three methods. Additionally, we applied all the four methods to a verification study for schizophrenia. Results show that G-TOW verified more important schizophrenia associated genes than other methods.

Dependence between genetic variants

The correlation between two SNPs are often denoted as LD, which plays a fundamental role in genetic association studies that map genes associated with specific traits of interests. GWAS often relies on high LDs to identify markers that are associated with traits where identified common variants are often in LD with causal variants. However, due to low minor allele frequencies, rare variants are often assumed to be independent. Understanding precisely the LD structure in the human genome including both rare and common variants may lead to better-formed statistical tests that could improve the power of association studies.

Two commonly used statistical measures have been proposed to quantify the degree of LD between a pair of variants: D' and R^2, which have different properties and may be applied for different application purposes (Hill and Robertson, 1968; Lewontin, 1964). Specifically, consider two loci, locus A with minor allele A and locus B with minor allele B. Then D' and R^2 are defined as:

$$R^2 = \frac{D^2}{p_A p_B (1 - p_A)(1 - p_B)} \qquad (1)$$

and

$$D' = \begin{cases} \dfrac{D}{\max\{-p_A p_B,\ -(1 - p_A)(1 - p_B)\}} & if \quad D < 0 \\[4mm] \dfrac{D}{\min\{p_A(1 - p_B),\ p_B(1 - p_A)\}} & if \quad D > 0 \end{cases} \qquad (2)$$

where $D = p_{AB} - p_A p_B$ and p_A, p_B, p_{AB} denote the frequencies of allele A, allele B, and haplotype AB, respectively. Hudson (1985) showed that two low frequency variants are likely to be negatively correlated by simulations based on coalescent models, that is $D < 0$ for two rare variants. If locus A and locus B are both rare variants, we can see that the numerator and denominator are approximated by $(p_A p_B)^2$ and $p_A p_B$ in equation (1), respectively, and the numerator and denominator are all approximated by $-p_A p_B$ in equation (2). This means that regardless of whether these two rare variants are related to each other, the value of R^2 always tends to be zero and the value of D' tends to be 1. Turkmen and Lin (2017) showed that the results from these two measurements are especially sensitive to the selection of tolerance cutoffs for pairs with small MAFs, where the tolerance cutoffs are set in the optimization function to obtain the maximum likelihood estimate of frequencies with the genetic package (Warnes 2013). They also showed that $D^* = \text{sign}(D)D'$ will likely attenuate the boundary value to -1, while $R = \text{sign}(D)\sqrt{R^2}$ will likely be close to 0 for rare variant pairs even if they come from similar (small) MAFs.

From the aforementioned, we can see that two commonly used LD measures are not capable of quantifying LD when rare variants are involved. The limitations of these two traditional LD measurements lead us to assume that rare variants are independent. Turkmen and Lin (2017) proposed an alternative measure of LD method, the polychoric correlation (PCor), and showed that rare variants in close physical proximity can show strong patterns of local correlation/LD when such variants are located on just a few haplotypes. Dickson et al, (2010) showed that some of the signals that have been detected from common variants could come from the effect of rare variants. Correctly utilizing existing LD among rare variants or between common and rare variants can help researchers to understand the genetic architecture of the human genome and improve power for detecting causal genetic variants. In this study, we propose a novel genetic association test in which we fully consider the correlation situation as either dependent or independent between variants.

Method

Consider a sample of n individuals and each individual has been genotyped at M variants in a genomic region. Denote y_i as the trait value of the ith individual for either a quantitative trait or a qualitative trait (1 for cases and 0 for controls for a qualitative trait) and denote $X_i = (x_{i1}, \cdots, x_{iM})'$ as the genotypic score of the ith individual, where $x_{im} \in \{0, 1, 2\}$ is the number of minor alleles that the ith individual has at the mth variant.

Without loss of generality, we first describe our methods without considering covariates and then extend it to incorporate covariates naturally, following Sha et al., (2011). We use the generalized linear model to model the relationship between trait values and genotypes (Nelder and Wedderburn, 1972):

$$g\big(\mathbb{E}[y_i|X_i]\big) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_M x_{iM} \tag{3}$$

where $g(\cdot)$ is a monotone link function and $\boldsymbol{\beta} = (\beta_0, \cdots, (\beta_M)$ are parameters. Two commonly used models under the generalized linear model framework are the linear model with the identity link for continuous or quantitative traits and the logistic regression model with the Logit link for a binary trait. Under the generalized linear model, the score test statistic to test the null hypothesis $H_0 : \boldsymbol{\beta} = 0$ is given by Sha et al., (2011):

$$S = U'V^{-1}U$$

where $U = \sum_{i=1}^{n}(y_i - \bar{y})(X_i - \bar{X})$ and $V = \frac{1}{n}\sum_{i-1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})'$. The statistic S asymptotically follows a chi-square distribution with $k = \text{rank}(V)$ degrees of freedom (df). However, for rare variants, the score test may suffer power loss due to the sparse data and a large df. To solve this issue, we can develop methods to test the genetic association in the considered region by testing the effect of a weighted combination of variants, $\sum_{m=1}^{M} w_m x_{im}$, where w_i are the weights that will be given later. That is, the model is written as:

$$g\big(\mathbb{E}[y_i|X_i]\big) = \beta_0 + \beta\left(\sum_{m=1}^{M} w_m x_{im}\right) \tag{4}$$

Under this model, to test the genetic association in the considered region is equivalent to test the null hypothesis $H_0 : \beta = 0$.

Let $\boldsymbol{X} = (X_1, \cdots, X_n)'$ and $Y = (y_1, \cdots, y_n)'$. To test the effect of the weighted combination of variants $X_i'\boldsymbol{W}$ where $\boldsymbol{W} = (w_1, \cdots, w_M)'$, let $P = I_n - \frac{1}{n}1_n 1_n'$ where 1_n represents a column vector containing all ones. Then, the score test statistic S as a function of W can be written as:

$$\begin{aligned}
S(w_1, \cdots, w_M) &= n\frac{(PY)'PXW(PY)'PXW}{(PY)'PY(PXW)'PXW} \\
&= n\frac{W'X'PYY'PXW}{Y'PYW'X'PXW} \\
&= n\frac{W'X'PYY'PXW}{W'X'PXY'PYW}
\end{aligned}$$

where the second equation holds because we have $P = P'$ and $PP' = P$ and the third equation holds because $Y'PY$ is a constant.

When $E = X'PX$ is positive definite, maximizing $S(w_1, \cdots, w_M)$ is equivalent to maximizing

$$S\left(w_1, \cdots, w_M\right) = \frac{W'X'PYY'PXW}{W'X'PXW}$$
$$= \frac{W'LL^{-1}X'PYY'PX\left(L^{-1}\right)'L'W}{W'LL'W}$$

where L is the lower triangular matrix obtained from the Cholesky decomposition of $E = LL'$. However, when the matrix of E is not full rank, it may cause the matrix E to be not positive definite (semi-positive definite). If the matrix is not positive definite, we consider two options. First, we can perform correlation pruning so that all pairwise correlation between SNPs is less than 1. Second, we can introduce a ridge parameter λ_0, for which we suggest $\lambda_0 = \sqrt{1/n}$ (referred to Baker et al. 2018), where n is the number of individuals in the data. This modification, $E = X'PX + \lambda_0 I$, allows us to mitigate the flaw of the non-positive matrix E in order to avoid the instability. We adopt the first option in the following simulation study and real data analysis if the matrix is not positive definite.

To maximize $S(w_1, \cdots, w_M)$, let us denote $C = L^{-1}X'PYY'PX(L^{-1})'$, and c be the eigenvector corresponding to the largest eigenvalue of the matrix C, then $S\left(w_1, \cdots, w_M\right) = \frac{W'LCL'W}{W'LL'W}$ reaches its maximum when $L'W = c$. Therefore, we obtain the optimal weight $W = (L')^{-1}c$. Specially, when we assume that the genetic variants are all independent, the weight of the mth variant can be expressed as

$$w_m = \frac{(PY)'\left(PX_{m\cdot}\right)}{\left(PX_{m\cdot}\right)'\left(PX_{m\cdot}\right)} = \frac{Y'PX_{m\cdot}}{X'_{m\cdot}PX_{m\cdot}} \tag{5}$$

This is the same as the optimal weight in the TOW method (Sha et al., 2012), where $X_{m\cdot} = (x_{m1}, \cdots, x_{mn})$. The equation (5) is equivalent to $w_m = \frac{Corr\left(Y, X_{m\cdot}\right)}{\sqrt{X'_{m\cdot}PX_{m\cdot}}}$ where the numerator is the correlation coefficient between phenotype Y and the mth variant genotypic score $X_{m\cdot}$ and the denominator can be viewed as the variance of the $X_{m\cdot}$. This means that w_m has the same direction as the correlation between the phenotype Y and the mth variant $X_{m\cdot}$, and puts big weight to the variant that has strong association with the trait of interest and to the variant with low variance which is usually a low frequency variant.

We define the General statistic to Test the effect of the Optimally Weighted combination (G-TOW) of variants, XW, as:

$$T_{G-TOW} = \frac{W'X'PYY'PXW}{W'X'PXW} \tag{6}$$

We use permutation methods to evaluate the P-value of T_{G-TOW}. Based on an assumption that rare variants are independent, TOW derives the optimal weights in order to maximize $S(w_1, \cdots, w_M)$. As we mentioned above, correlations between variants play an important role for association analysis and ignoring such correlations can reduce the power of association tests, especially those low frequent variants based association tests.

The G-TOW method can be extended to incorporate covariates. Suppose that there are p covariates. Let z_{il} denote the lth covariate of the ith individual. We adjust both trait value y_i and genotypic score x_{im} for the covariates by using linear regression models. That is,

$$y_i = \alpha_0 + \alpha_1 z_{i1} + \cdots + \alpha_p z_{ip} + \epsilon_i \text{ and } x_{im} = \alpha_{0m} + \alpha_{1m} z_{i1} + \cdots + \alpha_{pm} z_{ip} + \tau_{im}$$

Let $\tilde{y}_i$ and $\tilde{x}_{im}$ denote the residuals of y_i and x_{im}, respectively. We incorporate the covariate effects in G-TOW by replacing y_i and x_{im} in equation (6) by $\tilde{y}_i$ and $\tilde{x}_{im}$. With covariates, the statistic of G-TOW is defined as:

$$T_{G-TOW} = T_{G-TOW}\big|_{y_i = \tilde{y}_i, \, x_{im} = \tilde{x}_{im}}$$

Comparison of tests

We compared our method, (G-TOW) with three other approaches: 1) Weighted Sum Statistic method, (WSS), (Madsen and Browning 2009); 2) Testing the effect of the Optimally Weighted combination of variants, (TOW), (Sha et al., 2012); 3) Sequence Kernel Association Test, (SKAT), (Wu et al., 2011).

Simulation

Gene TP53 is crucial in multicellular organisms as it prevents cancer formation, thus functioning as a tumor suppressor. In simulation studies, we generate genotype data using the haplotypes of gene TP53 obtained from the 1000 Genome Project http://www.sph.umich.edu/csg/abecasis/MACH/download/1000G-2010-08.html. This dataset contains 566 hyplotypes for 74 variants (27 common variants and 47 rare variants). Here, we divide variants into rare (MAF<the rare variant threshold (RVT)) and common (MAF>RVT), and consider RVT as 0.05. In our simulation studies, we generate genotypes for individuals based on this haplotype pool. To generate the genotype of an individual, we select two haplotypes according to the haplotype frequencies.

Correlation analysis of the gene TP53

To be able to empirically illustrate the effects of varying values of MAF on correlation measures, we classified SNPs into five index groups using the following cutoffs: Index 1 (MAF 0.05), Index 2 (0.03 MAF<0.05), Index 3 (0.0 MAF<0.03), Index 4 (0.01 MAF<0.02) and Index 5 (MAF<0.01). Here, Index 1 represents common variants, whereas the remaining indices represent rare variants. Index 5 depicts the most extreme rare scenario. Detailed information for this classification and the number of SNPs in each index set for the dataset is summarized in Table 1.

For illustration purpose, we have chosen several SNP pairs from the TP53 Data to demonstrate the specific effects of varying values of MAF on R^2, D', R, D^* and PCor (Turkmen and Lin 2017). Table 2 summarizes the results reported with different LD measures and indicates the limitations of existing LD measures for rare variants. As shown

in Table 2, if a SNP is rare, it is very likely that the frequency of the haplotype carrying the minor allele can not be estimated accurately, whereas the two known popular LD measures R^2 and D' are calculated under the assumption that the population distribution of haplotypes is known. This will lead to D' attaining 1 or R^2 tending to equal 0 regardless of the level of real LD. From the values of PCor, we can see that these correlations/LD exist between rare variants or even rare and common variants in the gene TP53 data.

Simulation setting

To assess the performance of the proposed method G-TOW, we conduct extensive simulation studies based on the gene TP53 data. We compare the performance of G-TOW with the existing methods TOW, WSS, and SKAT.

To evaluate type I error rates of the four methods, we generate 1000 trait values which are independent of genotypes by using the following model (7) :

$$y = 0.5z_1 + 0.5z_2 + \epsilon \tag{7}$$

where z_1 is a continuous covariate generated from a standard normal distribution, z_2 is a binary covariate taking values 0 and 1 with a probability of 0.5, and ϵ is the error term following a standard normal distribution.

To evaluate power, we randomly choose one common variant and 20% or 40% rare variants (n_c) as causal variants where the common variant and half of rare causal variants ($n_r = n_c/2$) are risk variants and the other half of the rare causal variants ($n_p = n_c/2$) are protective variants. Denote x_i^r, x_j^p, and x^c as genotypes of the ith risk rare variant, the jth protective rare variant, and the common variant, respectively. Then, we generate a quantitative trait using the following model (8):

$$y = 0.5z_1 + 0.5z_2 + \beta_c x^c + \sum_{i=1}^{n_r} \beta_i^r x_i^r - \sum_{j=1}^{n_p} \beta_j^p x_j^p + \epsilon \tag{8}$$

where z_1, z_2 and ϵ are the same as those in equation (7). In model (8), β_c, β_i^r and β_j^p are coefficients. The values of these three coefficients depend on the total heritability (h_{total}) and the ratio (R^0) of heritability of the rare causal variant to heritability of the common causal variant. Given h_{total} and R^0, we can calculate the heritability of the rare casual variants and the common causal variants, respectively. We assume that all the rare causal variants have the same heritability such that rarer variants have larger effects. When we consider common variant and rare variants simultaneously, the values of β^c, β_i^r, and β_j^p can be determined by:

$$\beta_i^r = \sqrt{\frac{h_{total}R^0}{\text{var}\left(x_i^r\right)n_c(1 - h_{total})\left(1 + R^0\right)}}$$

$$\beta_j^p = \sqrt{\frac{h_{total}R^0}{\text{var}\left(x_j^p\right)n_c(1 - h_{total})\left(1 + R^0\right)}} \quad (9)$$

$$\beta^c = \sqrt{\frac{h_{total}}{\text{var}\left(x_k\right)(1 - h_{total})\left(1 + R^0\right)}}$$

When we only consider rare variants, the values of β_i^r and β_j^p are given as:

$$\beta_i^r = \sqrt{\frac{h_{total}}{\text{var}\left(x_i^r\right)n_c(1 - h_{total})}} \quad \beta_j^p = \sqrt{\frac{h_{total}}{\text{var}\left(x_j^p\right)n_c(1 - h_{total})}}$$

To generate a qualitative trait, we use a liability threshold model based on the quantitative trait (continuous phenotype) described above. An individual is defined to be affected (case) if the individual's phenotype is at least one standard deviation larger than the phenotypic mean, which yields a approximate prevalence of 16% for the simulated disease in a general population.

Simulation results

For type I error evaluation, we consider different disease models for different types of traits (qualitative trait or quantitative trait). The type I error rates are evaluated using 100,000 replicated samples and the P-values are estimated using 100,000 permutations for G-TOW, TOW and SKAT test methods. We use asymptotic distribution to evaluate the type I error rates for WSS test method. For the 100,000 replicated samples, the 95% confidence intervals (CIs) for the estimated type I error rates at nominal levels 0.05, 0.01, 0.001, and 0.0001 are (0.0486, 0.0514), (0.0094, 0.0106), (0.00080, 0.00120), and (0.000038, 0.000162), respectively. The estimated type I error rates of these four tests (G-TOW, TOW, WSS and SKAT) are summarized in Table 3. From this table, we can see that all of the estimated type I error rates are either within 95% CIs or close to the boundary of the corresponding 95% CIs, which indicate that the existing methods and our proposed method are all valid.

In power comparisons, the P-values of G-TOW, TOW and SKAT are estimated using 1,000 permutations, while the P-values of WSS are estimated by asymptotic distributions. The power of each of the four tests is evaluated using 1,000 replicated samples at the significance level of 0.05 (Figure 1–4). For power comparisons, we consider two different cases: (1) rare causal variants in which all causal variants are rare (MAF<RVT) and (2) both rare and common causal variants in which there is one common variant and the heritability of the common variant is twice as big as the heritability of all the rare causal variants.

Power comparisons of the four tests (G-TOW, TOW, WSS and SKAT) as a function of heritability for a quantitative trait are given in Figure 1 (20% causal rare variants) and Figure 3 (40% causal rare variants). As shown in Figure 1 or Figure 3, G-TOW and TOW are

consistently more powerful than WSS and SKAT. We can also see that G-TOW is the most powerful test when the causal variants contain both common and rare variants. Additionally, G-TOW is the most powerful test when the heritability is less than 0.05 and the casual variants are all rare variants. The power of both G-TOW and TOW is close to 1 when the heritability is greater than 0.05. The method SKAT performs worse than G-TOW and TOW under all of cases in Figure 1 and Figure 3, because the MAFs of a large portion of rare causal variants are under 0.01 and SKAT will lose power when the MAFs of causal variants are not in the range (0.01,0.035) (Yang et al. 2017). G-TOW is more powerful than TOW, especially when the causal variants contain both common and rare variants because TOW does not consider the correlation between variants. The method WSS is the least powerful test, because WSS will suffer substantial losses of power when both risk and protective variants are present.

Power comparisons of the four tests for different values of heritability based on a qualitative trait are given in Figure 2 (20% causal rare variants) and Figure 4 (40% causal rare variants). Comparing Figure 1 with Figure 2 (or Figure 3 with Figure 4), we can see that the patterns of power comparisons based on a qualitative trait are very similar to those based on a quantitative trait. However, it seems that the power of all four tests in most cases is smaller based on a qualitative trait than based on a quantitative trait. This may be because that converting to qualitative variables from quantitative variables results in some information loss. Comparing the case of 20% causal rare variants in Figure 1–2 with that of 40% causal rare variants in Figure 3–4, we can see that patterns of power comparisons based on 20% causal rare variants are very similar to those based on 40% causal rare variants. The power of G-TOW, TOW and SKAT is relatively robust to the increase of neutral variants while the power of WSS still suffers a substantial loss of power in the presence of both risk and protective variants.

Because the method TOW can be viewed as a special case of G-TOW, G-TOW is consistently more powerful than TOW when correlation between variants exist and has similar power when the variants are all independent. In summary, G-TOW is consistently more powerful or has similar power to the method TOW, and G-TOW is more powerful than WSS and SKAT. The power of G-TOW is also relatively robust to an increase in the percentage of rare variants.

Analysis of the GAW17 dataset

The GAW17 dataset consists of real genotypes and 200 replicates of the simulated phenotypes for a collection of 697 unrelated individuals, which is provided by the 1000 Genomes Project for their pilot3 study. A total of three quantitative traits (Q1, Q2, and Q4) are simulated and covariates include age, sex, and smoking status. Because the quantitative trait Q4 is not influenced by any of the genotyped exonic SNPs, we did not consider Q4 for the purpose of power comparisons. The P-values of G-TOW, TOW, and SKAT are evaluated by 10,000 permutations and the P-values of WSS are evaluated by asymptotic distributions. For a better comparison of these four test methods, we merge two replicates into one replicate to increase the sample size and all powers are estimated at a significance level of 0.001. In all of the cases, all causal variants are risk variants. There are nine causal genes for

Q1 and thirteen causal genes for Q2. We omit causal genes that have one variant, causal genes in which there is only one copy of haplotype among the individuals for each causal variant, causal genes in which all of the four tests have 100% power, and causal genes in which all of the four tests have a power less than 10%. There are four genes associated with Q1 and seven causal genes with Q2. The powers of WSS, TOW and SKAT are not consistent with those in Table 2 of Sha et al., (2012), where they did not adjust trait values and genotypes for covariates.

The powers of the four tests to detect association between each of the 11 causal genes and Q1 or Q2 are given in Table 4. As shown in Table 2, G-TOW is the most powerful test in four out of 11 genes, WSS is the most powerful test in three out of 11 genes, TOW is the most powerful test in one out of 11 genes and SKAT is the most powerful test in three out of 11 genes. Comparing G-TOW to TOW, G-TOW is either most powerful test or has similar power to the TOW method in ten of 11 casual genes.

Analysis of the Schizophrenia dataset

Schizophrenia is a severe and disabling mental illness with onset typically in early adult life. It is associated with low fecundity but nevertheless remains fairly common with a lifetime prevalence of around 1% (Power et al., 2013). Many genes including common and rare variants have been reported association with schizophrenia (Curtis et al. 2017, Genovese et al. 2016, Purcell et al. 2014, Fromer et al. 2014, Takata et al. 2014, Rujescu et al. 2009, Xu et al. 2012, Millar et al. 2000). In this study, we seek to verify the 14 associated genes (see Table 5) for schizophrenia by analyzing exome-sequencing of 2,536 schizophrenia cases and 2,543 controls (total 5079 individuals) of a Swedish sample. SNPs on the 14 genes with MAF that equals to zero and missing rate greater than 20% are excluded in the analysis.

The P-values of these four methods to test the association between each of the 14 genes and the case-control trait of schizophrenia are given in Table 5. As shown in Table 5, the number of genes verified by these four methods G-TOW, TOW, WSS and SKAT are 3,3,2, and 1, respectively, when we use the significant level as 0.05. When we consider multiple testing issue by using a significance level as 0.0036, among these four methods, only gene FYN is identified by the G-TOW (P-value = 0.0037). At this gene, the P-value of G-TOW is one order of magnitude smaller than the P-value of TOW method. The FMR1 gene can be verified by the other three tests other than TOW. Compared to the other three methods, we can see that our method G-TOW performs best overall from Table 5.

For the three verified genes by G-TOW (using significance level as 0.05), gene FYN, VPS39 and FMR1 have been reported significantly associated with schizophrenia by Tsavou et al. (2019), Curtis et al. (2017), Curtis et al. (2018), Xu et al. (2012) and Purcell et al. (2014), respectively. Gene FYN was identified as a 'prioritized candidate gene', and an intronic marker, rs7757969, on this gene was significant at $p = 4.8 \times 10^{-8}$ (Li et al., 2017). Gene FYN is involved in neuronal apoptosis, brain development and synaptic transmission, and lower expression has been observed in the platelets of schizophrenic patients compared with controls. Xu et al. (2012) showed that gene VPS39 was affected by recurrent novo events in a sequenced 795 exomes from 231 parent-proband trios enriched for sporadic schizophrenia

cases, as well as 34 unaffected trios. Purcell et al. (2014) proposed that targets of the fragile × mental retardation protein (FMRP, product of FMR1) were enriched in cases with schizophrenia.

Discussion

Most of the recently developed methods for rare variants association studies essentially assume that rare variants are independent. The two popular LD measures D' and R^2 are not capable of correctly measuring the correlation involving rare variants. Ignoring such correlation between variants will reduce the power of rare variants association tests. Motivated by TOW method, we analytically derived the optimal weights where we fully considered the correlations between variants whether they are common or rare. Based on the optimal weights, we proposed G-TOW to test the effect of the optimally weighted combination of variants. We used extensive simulation studies and application to the GAW17 dataset and the real schizophrenia data to compare the performance of our G-TOW method with existing methods (TOW, WSS and SKAT). The results show that our method is consistently more powerful than the other methods we compared with.

When we use burden tests to test rare variants, most of them put subjective weights on variants (usually put large weights on rare variants and small weights on common variants). By putting incorrect weights on variants, these methods lose power when testing the effects of both rare and common variants. To test the effects of both rare and common variants, WSS puts weights as the inverse square root of the expected variance based on allele frequencies on rare and common variants. By putting positive weights on variants, WSS loses power when causal variants contain both risk and protective variants. The SKAT method puts the beta distribution density function with pre-specified parameters determined by MAFs of variants in the data. However, SKAT has better performance only for those variants with MAF in a certain range (Sha et al., 2012). Our proposed method G-TOW which chooses weights adaptively by combining correlation between variants under a certain criterion, has relatively better performance than TOW in testing the effects of both rare and common variants or rare variants only, since TOW ignores the correlation between variants.

Acknowledgments

References

Allen HL, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F,… & Ferreira T (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467(7317): 832. [PubMed: 20881960]

Andrés AM, Clark AG, Shimmin L, Boerwinkle E, Sing CF, &, Hixson JE (2007). Understanding the accuracy of statistical haplotype inference with sequence data of known phase. Genetic epidemiology 31(7): 659–671. [PubMed: 17922479]

Baker E, Schmidt KM, Sims R, O'Donovan MC, Williams J, Holmans P,… & Consortium WTG (2018). POLARIS: Polygenic LD-adjusted risk score approach for set-based analysis of GWAS data. Genetic epidemiology 42(4): 366–377. [PubMed: 29532500]

Bansal V, Libiger O, Torkamani A, & Schork NJ (2010). Statistical analysis strategies for association studies involving rare variants. Nature Reviews Genetics 11(11): 773.

Cirulli ET, & Goldstein DB (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nature Reviews Genetics 11(6): 415.

Cohen JC, Pertsemlidis A, Fahmi S, Esmail S, Vega GL, Grundy SM, & Hobbs HH (2006). Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. Proceedings of the National Academy of Sciences of the United States of America 103(6): 1810–1815. [PubMed: 16449388]

Curtis D, Coelewij L, Liu SH, Humphrey J, &, Mott R (2018). Weighted burden analysis of exome-sequenced case-control sample implicates synaptic genes in schizophrenia aetiology. Behavior genetics 48(3): 198–208. [PubMed: 29564678]

Dickson SP, Wang K, Krantz I, Hakonarson H, & Goldstein DB (2010). Rare variants create synthetic genome-wide associations. PLoS biology 8(1): e1000294. [PubMed: 20126254]

Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P,… & Carrera N (2014). De novo mutations in schizophrenia implicate synaptic networks. Nature 506(7487): 179. [PubMed: 24463507]

Genovese G, Fromer M, Stahl EA, Ruderfer DM, Chambert K, Landn M.,… & Hultman CM (2016). Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. Nature neuroscience 19(11): 1433. [PubMed: 27694994]

Han F, & Pan W (2010). A data-adaptive sum test for disease association with multiple common or rare variants. Human heredity 70(1): 42–54. [PubMed: 20413981]

Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V,… & Workalemahu T (2010). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nature genetics 42(11): 949. [PubMed: 20935629]

Hill WG, & Robertson A (1968). Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38(6): 226–231. [PubMed: 24442307]

Hudson RR (1985). The sampling distribution of linkage disequilibrium under an infinite allele model without selection. Genetics 109(3): 611–631. [PubMed: 3979817]

Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, … & Lifton RP (2008). Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nature genetics 40(5): 592. [PubMed: 18391953]

Kinnamon DD, Hershberger RE, & Martin ER (2012). Reconsidering association testing methods using single-variant test statistics as alternatives to pooling tests for sequence data with rare variants. PLoS One 7(2): e30238. [PubMed: 22363423]

Lewontin RC (1964). The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 49(1): 49–67. [PubMed: 17248194]

Li B, & Leal SM (2008). Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. The American Journal of Human Genetics 83(3): 311–321. [PubMed: 18691683]

Li Z, Chen J, Yu H, He L, Xu Y, Zhang D, … & Song Z (2017). Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. Nature genetics 49(11): 1576. [PubMed: 28991256]

Madsen BE, & Browning SR (2009). A groupwise association test for rare mutations using a weighted sum statistic. PLoS genetics 5(2): e1000384. [PubMed: 19214210]

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, … & Cho JH (2009). Finding the missing heritability of complex diseases. Nature 461(7265): 747. [PubMed: 19812666]

Marini NJ, Gin J, Ziegle J, Keho KH, Ginzinger D, Gilbert DA, & Rine J (2008). The prevalence of folate-remedial MTHFR enzyme variants in humans. Proceedings of the National Academy of Sciences 105(23): 8055–8060.

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, & Hirschhorn JN (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nature Reviews Genetics 9(5): 356.

Millar JK, Wilson-Annan JC, Anderson S, Christie S, Taylor MS, Semple CA,… & Porteous DJ (2000). Disruption of two novel genes by a translocation co-segregating with schizophrenia. Human molecular genetics 9(9): 1415–1423. [PubMed: 10814723]

Morgenthaler S, & Thilly WG (2007). A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 615(1): 28–56. [PubMed: 17101154]

Nejentsev S, Walker N, Riches D, Egholm M, & Todd JA (2009). Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science 324(5925): 387–389. [PubMed: 19264985]

Nelder JA, & Baker RJ (1972). Generalized linear models. John Wiley & Sons, Inc..

Pan W (2009). Asymptotic tests of association with multiple SNPs in linkage disequilibrium. Genetic epidemiology 33(6): 497–507. [PubMed: 19170135]

Plenge RM, Cotsapas C, Davies L, Price AL, De Bakker PI, Mailer J, … & Parkin M (2007). Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nature genetics 39(12): 1477. [PubMed: 17982456]

Power RA, Kyaga S, Uher R, MacCabe JH, Lngstrm N, Landen M.,… & Svensson AC (2013). Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. JAMA psychiatry 70(1): 22–30. [PubMed: 23147713]

Price AL, Kryukov GV, de Bakker PI, Purcell SM, Staples J, Wei LJ, &, Sunyaev SR (2010). Pooled association tests for rare variants in exon-resequencing studies. The American Journal of Human Genetics 86(6): 832–838. [PubMed: 20471002]

Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, … & Duncan L (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506(7486): 185. [PubMed: 24463508]

Rujescu D, Ingason A, Cichon S, Pietilinen OP, Barnes MR, Toulopoulou T, … & Murray R (2008). Disruption of the neurexin 1 gene is associated with schizophrenia. Human molecular genetics 18(5): 988–996. [PubMed: 18945720]

Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, … & Hughes TE (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 316(5829): 1331–1336. [PubMed: 17463246]

Schork NJ, Murray SS, Frazer KA, & Topol EJ (2009). Common vs. rare allele hypotheses for complex diseases. Current opinion in genetics & development 19(3): 212–219. [PubMed: 19481926]

Sha Q, Wang S, &, Zhang S (2013). Adaptive clustering and adaptive weighting methods to detect disease associated rare variants. European Journal of Human Genetics 21(3): 332. [PubMed: 22781093]

Sha Q, Wang X, Wang X, & Zhang S (2012). Detecting association of rare and common variants by testing an optimally weighted combination of variants. Genetic epidemiology 36(6): 561–571. [PubMed: 22714994]

Sha Q, Zhang Z, & Zhang S (2011). An improved score test for genetic association studies. Genetic epidemiology 35(5): 350–359. [PubMed: 21484862]

Takata A, Xu B, Ionita-Laza I, Roos JL, Gogos JA, & Karayiorgou M (2014). Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. Neuron 82(4): 773–780. [PubMed: 24853937]

Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, … & Wilson AG (2007). Rheumatoid arthritis association at 6q23. Nature genetics 39(12): 1431. [PubMed: 17982455]

Tsavou A, & Curtis D (2019). In-silico investigation of coding variants potentially affecting the functioning of the glutamatergic N-methyl-D-aspartate receptor in schizophrenia. Psychiatric genetics 29(2): 44–50. [PubMed: 30664045]

Turkmen A, & Lin S (2017). Are rare variants really independent?. Genetic epidemiology 41(4): 363–371. [PubMed: 28300291]

Warnes G with contributions from Gorjanc G, Leisch F, & Man M (2013). Genetics: Population genetics. R package version 1.3.8.1. Retrieved from http://CRAN.R-project.org/package=genetics.

Wu MC, Lee S, Cai T, Li Y, Boehnke M, & Lin X (2011). Rare-variant, association testing for sequencing data with the sequence kernel association test. The American Journal of Human Genetics 89(1): 82–93. [PubMed: 21737059]

Xu B, Ionita-Laza I, Roos JL, Boone B, Woodrick S, Sun Y.,… & Karayiorgou M (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. Nature genetics 44(12): 1365. [PubMed: 23042115]

Yang X, Wang S, Zhang S, & Sha Q (2017). Detecting association of rare and common variants based on cross-validation prediction error. Genetic epidemiology 41(3): 233–243. [PubMed: 28176359]

Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, … &, Barrett JC (2007). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 316(5829): 1336–1341. [PubMed: 17463249]

Zhu X, Feng T, Li Y, Lu Q, & Elston RC (2010). Detecting rare variants for complex traits using family and unrelated data. Genetic epidemiology 34(2): 171–187. [PubMed: 19847924]
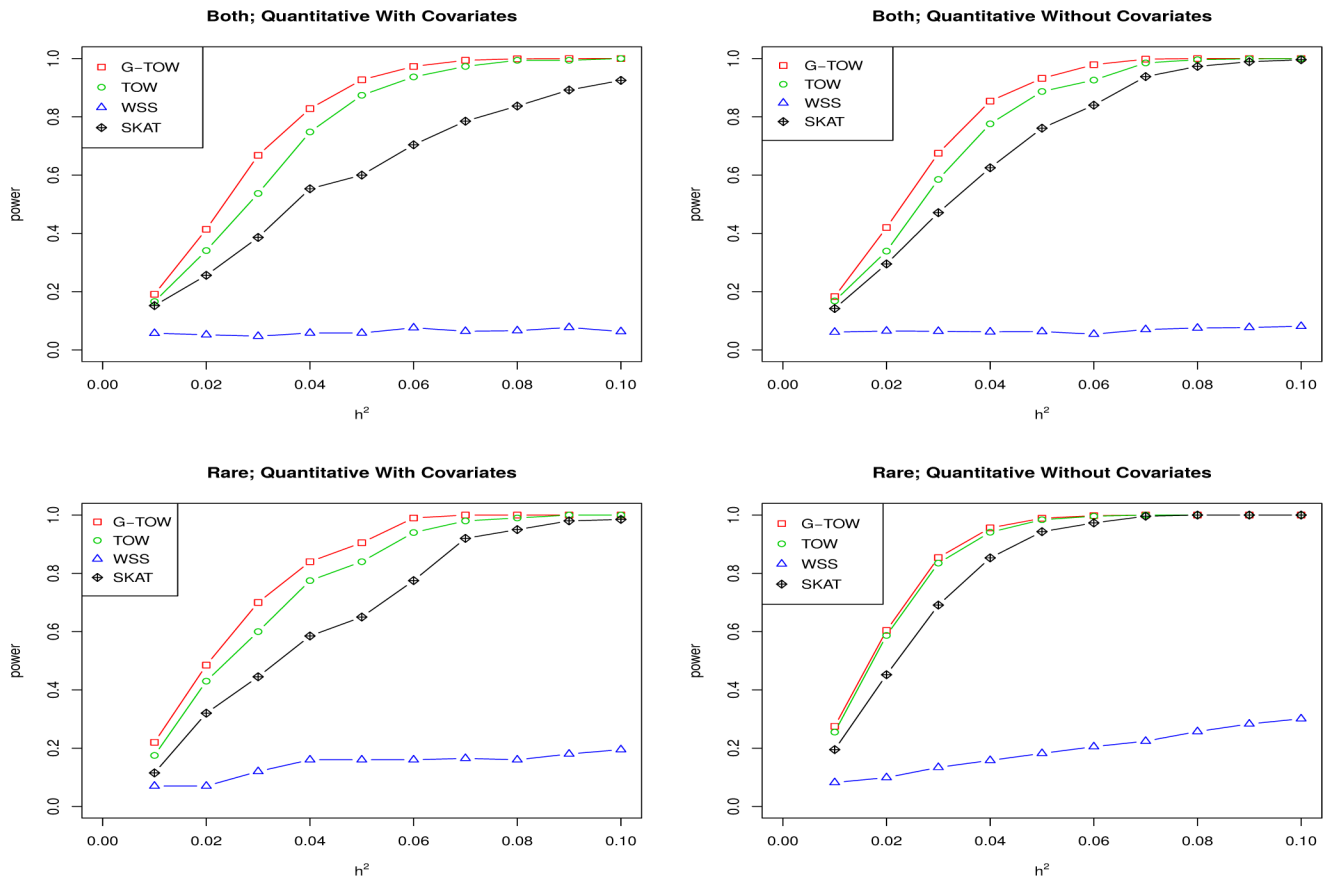
Figure 1.
Power comparisons of the four tests (G-TOW, TOW, WSS and SKAT) for the power as a function of heritability for a quantitative trait. Rare means that all causal variants are rare. Both means that causal variants contain both rare and common (one common variant) and the heritability of the common variant is as twice as the heritability of all the rare causal variants. The x-axis represents the total heritability of all causal variants. Sample size is 1,000 and 20% of rare variants are causal, half of rare causal variants are risk and the other half of causal variants are positive. The powers (y-axis) are evaluated at a significance level of 0.05.
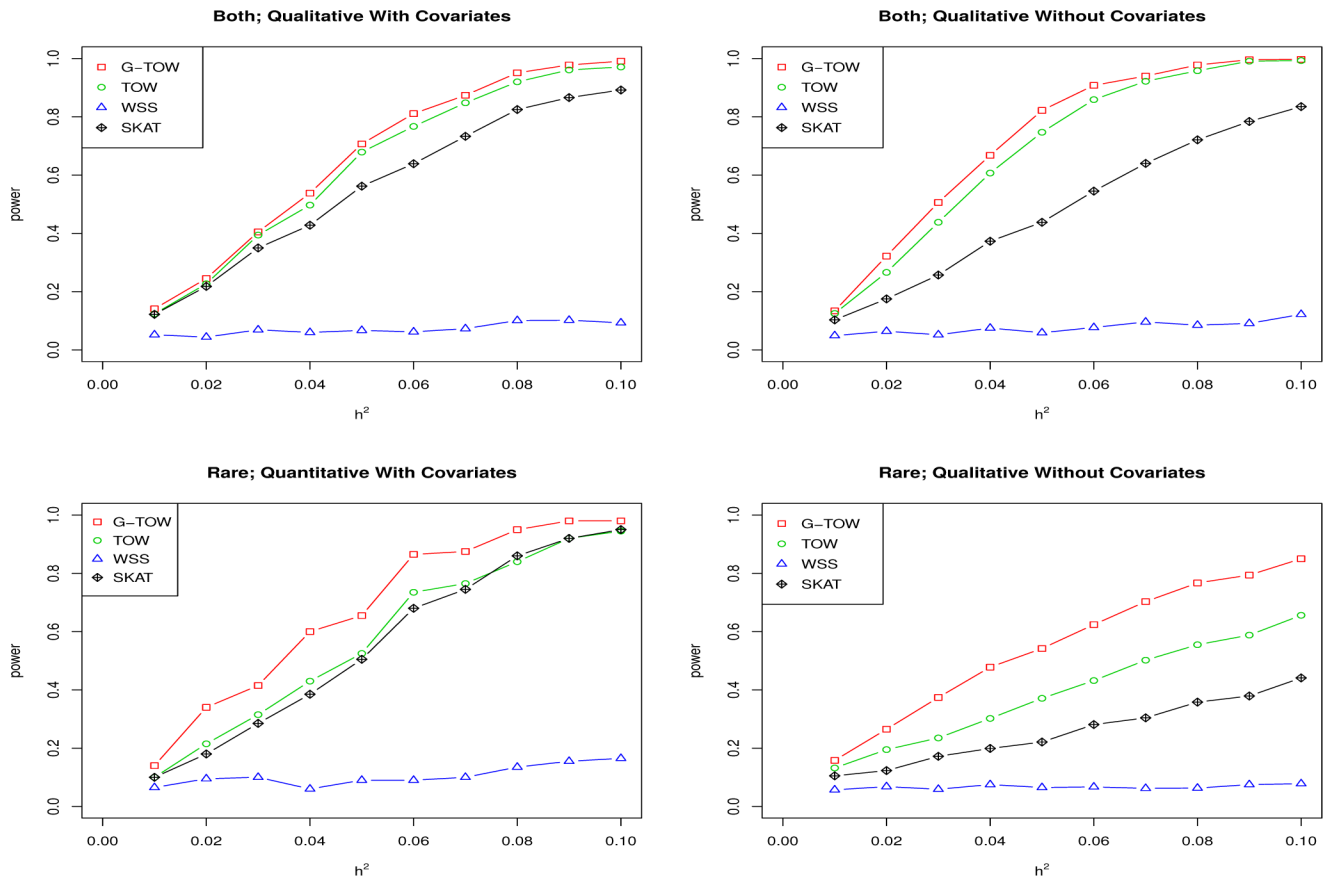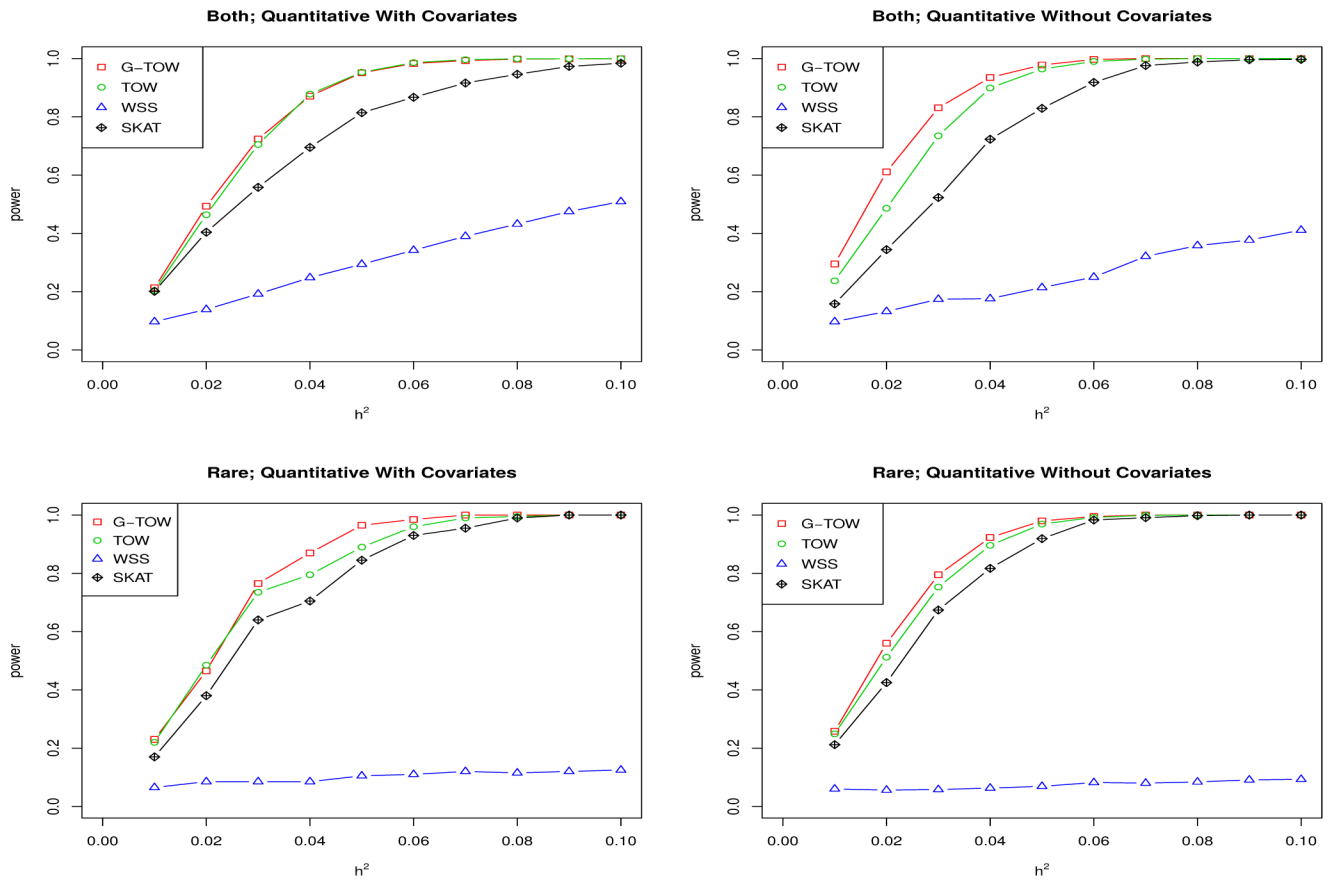
Figure 2.
Power comparisons of the four tests (G-TOW, TOW, WSS and SKAT) for the power as a function of heritability for a qualitative trait. Rare means that all causal variants are rare. Both means that causal variants contain both rare and common (one common variant) and the heritability of the common variant is as twice as the heritability of all the rare causal variants. The x-axis represents the total heritability of all causal variants. Sample size is 1,000 and 20% of rare variants are causal, half of rare causal variants are risk and the other half of causal variants are positive. The powers (y-axis) are evaluated at a significance level of 0.05.

Figure 3.
Power comparisons of the four tests (G-TOW, TOW, WSS and SKAT) for the power as a function of heritability for a quantitative trait. Rare means that all causal variants are rare. Both means that causal variants contain both rare and common (one common variant) and the heritability of the common variant is as twice as the heritability of all the rare causal variants. The x-axis represents the total heritability of all causal variants. Sample size is 1,000 and 40% of rare variants are causal, half of rare causal variants are risk and the other half of causal variants are positive. The powers (y-axis) are evaluated at a significance level of 0.05.
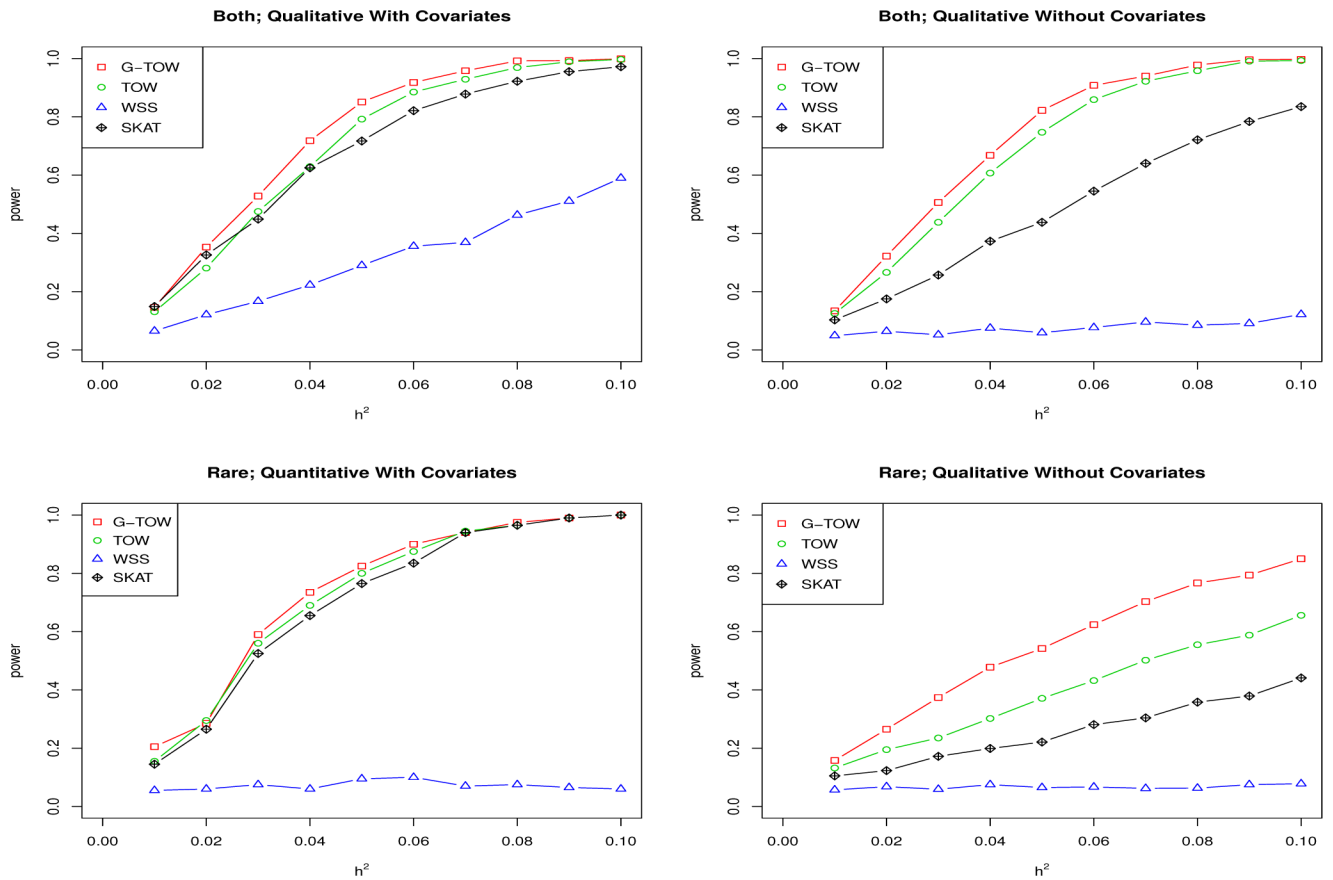
Figure 4.
Power comparisons of the four tests (G-TOW, TOW, WSS and SKAT) for the power as a function of heritability for a qualitative trait. Rare means that all causal variants are rare. Both means that causal variants contain both rare and common (one common variant) variants. The heritability of the common variant is as twice as the heritability of all the rare causal variants. The x-axis represents the total heritability of all causal variants. Sample size is 1,000 and 40% of rare variants are causal, half of rare causal variants are risk and the other half of causal variants are positive. The powers (y-axis) are evaluated at a significance level of 0.05.

Table 1:

Index Definitions for TP53 Data based on MAF Cutoffs

Group	MAF Range	#SNPs
Index 1	MAF 0.05	27
Index 2	0.03 MAF<0.05	8
Index 3	0.02 MAF<0.03	4
Index 4	0.01 MAF<0.02	6
Index 5	MAF<0.01	29

Table 2:

R^2, D', R, D^* and PCor values for the Selected SNP-Pairs on TP53 Data

Index		MAF1	MAF2	R^2	D'	R	D^*	PCor
1	1	0.297	0.290	0.844	0.936	0.919	0.936	0.993
1	2	0.297	0.040	0.099	1.000	0.313	1.000	0.989
1	3	0.297	0.025	0.060	1.000	0.246	1.000	0.983
1	4	0.297	0.019	0.008	1.000	−0.089	−1.000	−0.951
1	5	0.297	0.003	0.001	1.000	−0.034	−1.000	−0.942
2	2	0.040	0.042	0.950	1.000	0.974	1.000	0.999
2	3	0.040	0.025	0.001	1.000	−0.032	−1.000	−0.990
2	4	0.040	0.019	0.001	1.000	−0.028	−1.000	−0.898
2	5	0.040	0.003	0.000	1.000	−0.010	−1.000	−0.894
3	3	0.025	0.028	0.907	1.000	0.952	1.000	0.999
3	4	0.025	0.019	0.001	1.000	−0.021	−1.000	−0.991
3	5	0.025	0.003	0.000	1.000	−0.008	−1.000	−0.890
4	4	0.019	0.015	0.000	1.000	−0.017	−1.000	−0.899
4	5	0.019	0.003	0.000	1.000	−0.007	−1.000	−0.821
5	5	0.003	0.001	0.000	1.000	−0.001	−1.000	−0.820

Table 3:

The estimated type I error rates of the four test methods

	α-level	Quantitative trait			
		G-TOW	**TOW**	**WSS**	**SKAT**
$n = 1000$	0.05	0.0497	0.0502	0.0500	0.0497
	0.01	0.0097	0.0104	0.0100	0.0099
	0.001	0.00093	0.00110	0.00100	0.00105
	0.0001	0.00013	0.00010	0.00012	0.00008
		Qualitative trait			
$n = 1000$	0.05	0.0496	0.0498	0.0500	0.0497
	0.01	0.0102	0.0098	0.0100	0.0096
	0.001	0.00104	0.00097	0.00100	0.00087
	0.0001	0.00011	0.00009	0.00008	0.00011

Table 4:

Power of the four methods to test the association between each of the four causal genes and quantitative trait Q1 and between each of the seven causal genes and quantitative trait Q2

Traits	Gene name	No. of variants, no. of causal variants	Min, max, mean MAF	G-TOW	TOW	WSS	SKAT
Q1	ARNT	18,5	0.07,1.15,0.33	0.58	0.59	0.08	**0.94**
	FLT4	10,2	0.07,0.14,0.11	0.66	0.69	**0.85**	0.66
	HIF1A	8,4	0.07,1.22,0.39	0.57	0.59	0.62	**0.90**
	VEGFA	6,1	0.22,0.22,0.22	0.17	0.16	**0.35**	0.11
Q2	BCHE	29,13	0.07,0.29,0.10	**0.38**	0.37	0.17	0.13
	LPL	20,3	0.07,1.58,0.60	0.25	0.23	0.00	**0.36**
	PDGFD	11,4	0.07,0.86,0.29	**0.18**	0.15	0.01	0.14
	SIRT1	23,9	0.07,0.22,0.12	0.31	**0.61**	0.34	0.56
	SREBF1	24,10	0.07,0.43,0.22	0.22	0.17	**0.24**	0.04
	VNN1	7,2	0.57,17.1,8.82	**0.83**	0.71	0.08	0.03
	VNN3	15,7	0.07,9.83,2.06	**0.72**	0.53	0.31	0.37

Note: Min, max, mean MAF: the minimum, maximum, and mean MAF (in percentage) at causal variants. In each row, the boldfaced number represents the highest power in the row

Table 5:

P-values of these four methods to test association between genes and case-control trait

Gene name	Min, max, mean MAF	G-TOW	TOW	WSS	SKAT
ARC	0.0001, 0.4618, 0.0339	0.2583	0.2817	0.4645	0.7751
DISCI	0.0001, 0.3111, 0.0135	0.7645	0.6939	0.7736	0.4124
DISC2	0.0001, 0.1579, 0.0309	0.5306	0.5056	0.6880	0.5407
DPYD	0.0001, 0.2481, 0.0102	0.1475	0.1959	0.4782	0.1119
FMR1	0.0001, 0.0847, 0.0078	**0.0339**	0.0598	**0.0313**	**0.0252**
FYN	0.0001, 0.1377, 0.0064	*0.0037*	**0.0143**	0.9254	0.1628
KL	0.0001, 0.4038, 0.0213	0.5505	0.5829	0.6533	0.8737
LAMA2	0.0001, 0.4669, 0.0183	0.1154	**0.0138**	0.5967	0.1166
NMDAR	0.0001, 0.4118, 0.0192	0.3431	0.3272	0.3804	0.4312
NRXN1	0.0001, 0.0090, 0.0009	0.5764	0.4227	**0.0303**	0.4997
SETD1A	0.0001, 0.3788, 0.0037	0.6878	0.5235	0.6386	0.7821
TAF13	0.0001, 0.1276, 0.0109	0.7455	0.7414	0.3661	0.7041
TRRAP	0.0001, 0.2383, 0.0034	0.1046	0.4087	0.1040	0.6638
VPS39	0.001, 0.1830, 0.0115	**0.0442**	**0.0245**	0.4767	0.0752

Note: Min, max, mean MAF: the minimum, maximum, and mean MAF (in percentage) at causal variants. In each row, the boldfaced number represents the significant P-value in the row (significant level = 0.05); the italic number in each row represents the nearly significant P-value after correcting for multiple testing issue (significant level = 0.0036).