Joint analysis of multiple traits in rare variant association studies

Zhenchuan Wang[1], **Xuexia Wang**[2], **Qiuying Sha**[1], and **Shuanglin Zhang**[1,*]

[1]Department of Mathematical Sciences, Michigan Technological University, Houghton, Michigan

[2]Joseph J. Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin

SUMMARY

The joint analysis of multiple traits has recently become popular since it can increase statistical power to detect genetic variants and there is increasing evidence showing that pleiotropy is a widespread phenomenon in complex diseases. Currently, the majority of existing methods for the joint analysis of multiple traits test association between one common variant and multiple traits. However, the variant-by-variant methods for common variant association studies may not be optimal for rare variant association studies due to the allelic heterogeneity as well as the extreme rarity of individual variants. Current statistical methods for rare variant association studies are for one single trait only. In this paper, we propose an Adaptive Weighting Reverse Regression (AWRR) method to test association between multiple traits and rare variants in a genomic region. AWRR is robust to the directions of effects of causal variants and is also robust to the directions of association of traits. Using extensive simulation studies, we compare the performance of AWRR with canonical correlation analysis (CCA), Single-TOW, and the Weighted Sum Reverse Regression (WSRR). Our results show that, in all of the simulation scenarios, AWRR is consistently more powerful than CCA. In most scenarios, AWRR is more powerful than Single-TOW and WSRR.

Keywords

rare variants; multiple traits; association studies; multiple correlated phenotypes; pleiotropy

Introduction

There is increasing evidence showing that pleiotropy, the effect of one variant on multiple traits, is a widespread phenomenon in complex diseases (Sivakumaran et al., 2011). Furthermore, in genetic association studies of complex diseases, multiple related traits are usually measured. For example, hypertension is evaluated using systolic and diastolic blood pressures, the Metabolic Syndrome is based on observing three of five criteria (Sattar et al.,

*Corresponding author: Shuanglin Zhang, Department of Mathematical Sciences, Michigan Technological University, 1400 Townsend Drive, Houghton, Michigan, 49931, USA, Tel: 1 (906) 487-2095, Fax: 1 (906) 487-3133, shuzhang@mtu.edu.

Conflict of Interest

The authors declare no conflict of interest.

2008), and neuropsychiatric disorders depend on a range of overlapping clinical characteristics (O'Reilly et al., 2012). Although most published genome-wide association studies (GWAS) analyze each of the related traits separately, the joint analysis of multiple traits not only can increase statistical power to detect genetic variants (Solovieff et al., 2013; Stephens, 2013; Yang & Wang, 2012; Zhou & Stephens, 2014), but can also be crucial to understand the genetic architecture of the disease of interest (Aschard et al., 2014). Thus, the joint analysis of multiple traits has recently become popular. Several statistical methods for the joint analysis of multiple traits have been developed. These methods can be roughly divided into three groups: regression methods, combining test statistics from univariate analysis, and dimension reduction methods. Regression methods include mixed effects models (Korte et al., 2012; Zhou & Stephens, 2014) and reverse regression models (O'Reilly et al., 2012; Yan et al., 2013). By modeling the covariance structure of correlated traits and dependence structure between individuals, mixed effects models not only can incorporate multiple correlated traits, but also can be robust to population stratification. Reverse regression models consider genotypes as the response variable and all the traits as independent variables, therefore, reverse regression models do not need to know the complex distributions of the traits and can be applied to a large number of mixed types of traits. For combining the test statistics from univariate analysis, one first obtains univariate test statistics by performing association tests for each trait individually and then combines the univariate test statistics by linear combinations (O'Brien, 1984; van der Sluis et al., 2013; Yang et al., 2010). The dimension reduction methods include canonical correlation analysis (CCA) (Tang & Ferreira, 2012), principal components of traits (PCT) (Aschard et al., 2014), and principal components of heritability (PCH) (Klei et al., 2008; Lange et al., 2004; Ott & Rabinowitz, 1999). CCA is to find a linear combination of traits and a linear combination of genotypes at multiple variants such that the correlation between the two linear combinations reaches its maximum. PCT is the principal component analysis to the traits. The PCT methods are usually based on the first PC or first few PCs of the traits (Feng et al., 2007; Klei et al., 2008). Aschard et al. (2014) showed that contrary to the widespread practice, tests based on only the first few PCs often have low power, whereas combining signals across all PCs can have greater power. PCH is to find a linear combination of multiple traits such that this linear combination has the maximum heritability.

Almost all of the aforementioned methods are to test association between one common variant and multiple traits. However, the variant-by-variant methods for common variant association studies may not be optimal for rare variant association studies due to the allelic heterogeneity as well as the extreme rarity of individual variants (Li & Leal, 2008). Recent studies show that complex diseases are caused by both common and rare variants (Bodmer & Bonilla, 2008; Kang et al., 2010; Pritchard, 2001; Pritchard & Cox, 2002; Stratton & Rahman, 2008; Teer & Mullikin, 2010; Walsh & King, 2007). Next-generation sequencing technology allows sequencing of the whole genome of large groups of individuals, and thus makes rare variant association studies feasible (Andres et al., 2007; Metzker, 2010). Recently, statistical methods for rare variant association studies with a single trait have been developed by summarizing genotype information from multiple variants. These methods include burden tests (Li & Leal, 2008; Madsen & Browning, 2009; Morgenthaler & Thilly, 2007; Price et al., 2010; Zawistowski et al., 2010), quadratic tests (Neale et al., 2011; Sha et

al., 2012; Wu et al., 2011), and combined tests (Derkach et al., 2013; Lee et al., 2012; Sha & Zhang, 2014). Burden tests collapse rare variants in a genomic region into a single burden variable and then regress the trait on the burden variable to test for the cumulative effects of rare variants in the region. These tests implicitly assume that all rare variants are causal and that the directions of the effects are all the same. Quadratic tests include tests with statistics of quadratic form of score vector, as well as adaptive weighting methods. These tests are robust to the directions of the effects of causal variants and are less affected by neutral variants than burden tests. Burden tests can only outperform quadratic tests when most of the rare variants are causal and the directions of the effects of causal variants are all the same. Combined tests combine information from burden tests, quadratic tests, and possibly other tests aiming to have advantages of multiple tests.

In this article, we propose an adaptive weighting reverse regression (AWRR) method to test association between multiple traits and rare variants in a genomic region. In AWRR, we first propose adaptive weights to collapse genotypes. Then, we use the score test to test association based on the reverse regression, in which collapsed genotypes is treated as the response variable and multiple traits are treated as independent variables. Using extensive simulation studies, we compare the performance of AWRR with CCA, Single-TOW, and the Weighted Sum Reverse Regression (WSRR). In Single-TOW, we first calculate the TOW statistic (Sha et al., 2012) to test the association between each trait and variants in a genomic region and then the statistic of Single-TOW is the largest of TOW statistics. In WSRR, we first calculate the weighted sum (Madsen & Browning, 2009) of genotypes at variants in a genomic region and then the statistic of WSRR is the score test statistic under reverse regression model, in which the weighted sum of genotypes is the response variable and traits are predictor variables. Our results show that, in all of the simulation scenarios, AWRR is consistently more powerful than CCA. In most scenarios, AWRR is more powerful than Single-TOW and WSRR.

Method

We consider a sample with n unrelated individuals. Each individual has K (potentially correlated) traits and has been genotyped at M variants in a genomic region. Let y_{ik} denote the kth trait value of the ith individual. Let x_{im} denote the genotype score of the mth variant of the ith individual, where x_{im} is the number of minor alleles of the ith individual carried at the mth variant. We denote $Y_i = (y_{i1}, \ldots, y_{iK})^T$ as the K traits for the ith individual. We propose an adaptive weighting reverse regression (AWRR) method to test the null hypothesis H_0: none of the K traits are associated with the M variants in the genomic region. For constructing the test statistic of AWRR, we first collapse the M dimensional genotype (x_{i1}, $\ldots, x_{iM}$) into a one dimensional number $x_i = \sum_{m=1}^{M} w_m x_{im}$, where w_m is the adaptive weight for the mth variant. The adaptive weight w_m should satisfy the properties that w_m should be large if the mth variant has strong association with the K traits and w_m should have different signs for risk and protective variants. Then, the statistic of AWRR is the score test statistic

under the reverse regression model $x_i = \beta_0 + \sum_{k=1}^{K} \beta_k y_{ik} + \varepsilon_i$. In details, the AWRR method has the following steps.

1. We define a weight W_m for the mth variant such that W_m will be large if the mth variant has strong association with the K traits and W_m will be also large if the mth variant is a rare variant. For these purposes, we propose $W_m = \dfrac{1}{\sqrt{p_m(1-p_m)}} T_m$, where p_m is the minor allele frequency of the mth variant and T_m is the score statistic to test the null hypothesis $H_0: \beta_1 = \cdots = \beta_K = 0$ under the reverse regression model $\log\dfrac{p_{im}}{1-p_{im}} = \beta_0 + \beta_1 y_{i1} + \cdots + \beta_K y_{iK}$, where we assume a dominant model $p_{im} = \Pr(x_{im}=1) = \Pr(x_{im}=2)$. In fact, for rare variants, x_{im} essentially is 0 or 1. The score statistic is given by $T_m = U_m^T V_m^{-1} U_m$, where $U_m = \sum_{i=1}^{n} Y_i(x_{im} - \overline{x}_m)$ and $V_m = \dfrac{1}{n}\sum_{i=1}^{n}(x_{im} - \overline{x}_m)^2 \sum_{i=1}^{n}(Y_i - \overline{Y})(Y_i - \overline{Y})^T$, where $\overline{x}_m = \dfrac{1}{n}\sum_{i=1}^{n} x_{im}$, and $\overline{Y} = \dfrac{1}{n}\sum_{i=1}^{n} Y_i$. Under the null hypothesis, T_m follows a χ^2 distribution with degrees of freedom K. However, W_m does not consider the direction of the effects of causal variants.

2. In this step, we will define a direction of W_m. We first select a trait (the selected trait denoted as the $k^{\tilde{t}h}$ trait). We use $\text{sign}(\rho(y_k, \tilde{x}_m))$ to denote the direction of the associations of the mth variant, where $y_k \tilde{=} (y_{1k}, \ldots, y_{nk})^T$ and $x_m = (x_{1m}, \ldots, x_{nm})^T$. If the $k^{\tilde{t}h}$ trait has no association with the M variants, the directions of the association will be random. In order to try to avoid random directions, we propose to choose the trait that has the strongest association with the M variants. Let T_{TOW}^k denote the statistic of TOW (Sha et al., 2012) to test association between the kth trait and the M variants. T_{TOW}^k is defined as $T_{TOW}^k = \sum_{i=1}^{n}(y_{ik} - \overline{y}_k)(x_i^o - \overline{x}^o)$, where $x_i^o = \sum_{m=1}^{M} w_m^o x_{im}$ and $w_m^O = \sum_{i=1}^{n}(y_{ik} - \overline{y}_k)(x_{im} - \overline{x}_m) / \sum_{i=1}^{n}(x_{im} - \overline{x}_m)^2$. We choose the $k^{\tilde{t}h}$ trait such that $T_{TOW}^{\tilde{k}} = \max_{1 \le k \le K} T_{TOW}^k$.

3. The final weight for the mth variant is given by $w_m = \text{sign}(\rho(y_k, \tilde{x}_m)) W_m$. Let $x_i = \sum_{m=1}^{M} w_m x_{im}$. Then, we consider the reverse regression model $x_i = \beta_0 + \sum_{k=1}^{K} \beta_k y_{ik} + \varepsilon_i$. We apply a score test to test the null hypothesis $H_0: \beta_1 = \cdots = \beta_K = 0$. The score statistic is given by $T_{AWRR} = u^T V^{-1} u$, where

$$u = \sum_{i=1}^{n}(x_i - \overline{x})Y_i, \quad V = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})(Y_i - \overline{Y})^T, \quad \overline{x} = \frac{1}{n}\sum_{i=1}^{n}x_i, \text{ and}$$

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n}Y_i.$$

4. In this step, we evaluate the p-value of T_{AWRR}. Since w_m depends on the trait values and the genotype scores, T_{AWRR} does not follow a χ^2 distribution with degrees of freedom K. We use a permutation procedure to evaluate the p-value of T_{AWRR}. In each permutation, we randomly shuffle $Y_1,\ldots,Y_n$ and keep the genotypes of each individual unchanged. We repeat step 1 to step 3 based on each permutated data. Let T_{AWRR}^{0} denote the test statistic of T_{AWRR} based on the original data and T_{AWRR}^{per} denote the test statistic based on the permuted data. Then the p-value of the test T_{AWRR} is the proportion of the number of permutations with $T_{AWRR}^{\text{per}} \geq T_{AWRR}^{0}$.

Comparison of Tests

We compare the performance of the proposed test AWRR with those of the canonical correlation analysis (CCA) (Tang & Ferreira, 2012), the Single-TOW method (Sha et al., 2012), and the Weighted Sum Reverse Regression (WSRR) method (Madsen & Browning, 2009).

CCA method

although the asymptotical distribution of the CCA statistic works well for common variants, it is very conservative for rare variants. Thus, we propose to use a permutation procedure to evaluate the p-value of the CCA statistic.

Single-TOW method

let T_{TOW}^{k} denote the statistic of TOW (Sha et al., 2012) to test association between the kth trait and the M variants. The statistic of Single-TOW is given by $T_{\text{Single}-TOW} = \max_{1 \leq k \leq K} T_{TOW}^{k}$. The p-value of $T_{Single\text{-}TOW}$ is evaluated by a permutation procedure.

WSRR method

let $X_i = \sum_{m=1}^{M} w_m x_{im}$, where $w_m = 1/\sqrt{p_m(1-p_m)}$ and p_m is the minor allele frequency of the mth variant. We consider the reverse regression model $X_i = \beta_0 + \sum_{k=1}^{K}\beta_k y_{ik} + \varepsilon_i$. The statistic of WSRR, T_{WSRR}, is the score test statistic to test the null hypothesis $H_0: \beta_1 = \cdots = \beta_K = 0$. Under the null hypothesis, T_{WSRR} follows a χ^2 distribution with degrees of freedom K.

Simulation

Our simulations follow that of Sha et al. (2012). In detail, the empirical Mini-Exome genotype data provided by the genetic analysis workshop 17 (GAW17) are used for simulation studies. This dataset contains genotypes of 697 unrelated individuals on 3205 genes. We conduct two sets of simulations. In the first set of simulations, we choose four genes: *ELAVL4, MSH4, PDE4B*, and *ADAMTS4* with 10, 20, 30, and 40 variants, respectively. We merge the four genes to form a super gene (Sgene1) with 100 variants (Sha et al., 2012). In the second set of simulations, we choose ten genes: *ELAVL4, FAM73A, PSMB4, FSHR, GMCL1, HNMT, GALNT13, NEUROD1, MYEOV2*, and *TWF2* with 10 variants in each of them. We merge the ten genes to form a super gene (Sgene2) with 100 variants. In our simulation studies, we generate genotypes based on the genotypes of 697 individuals in the Sgene1 and Sgene2. To generate a qualitative disease affection status, we use a liability threshold model based on a quantitative trait. For a qualitative trait, an individual is defined as affected if the individual's corresponding quantitative trait is at least one standard deviation larger than the phenotypic mean. This yields a prevalence of 16% for the simulated disease in the general population. In the following, we describe how to generate a quantitative trait.

To evaluate the type I error, we generate K traits of an individual independent of the genotypes by using

$$Y = (y_1, \ldots, y_K)^T = \sqrt{\rho}\mathrm{B}u + \sqrt{1-\rho}\varepsilon,$$

where $u = (u_1, \ldots, u_{n_u})^T \sim MVN(0, I)$ is a vector of n_u independent standard normal latent variables, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_K)^T \sim MVN(0, I)$ is a vector of errors, B is a $K \times n_u$ loading matrix, the values of n_u and B are based on two variance models: (1) $n_u = 1$, $\mathrm{B} = (1, \ldots, 1)^T$ and (2) $n_u = 2$, $\mathrm{B} = \begin{bmatrix} e_{\lceil K/2 \rceil} & 0 \\ 0 & e_{K - \lceil K/2 \rceil} \end{bmatrix}$. Thus, $Y \sim MVN(0, \Sigma)$, where $\Sigma = \rho\mathrm{BB}^T + (1-\rho)I$.

To evaluate power, we consider that all causal variants are rare (MAF<0.01). We randomly choose n_c rare variants as causal variants, where n_c is determined by the percentage of causal variants among rare variants. Denote n_r and n_p as the number of risk rare variants and protective rare variants, respectively, where $n_r + n_p = n_c$. Let x_{qi}^r and x_{ji}^p denote the genotypic scores of the qth risk rare variant and the jth protective rare variant for the ith individual, respectively. Suppose that causal variants have impact on L traits among the K traits. Among the L traits, there are L_p traits positively correlated with risk variants and there are L_n traits negatively correlated with risk variants. Let h denote the heritability of all the n_c rare causal variants on each of the L traits. Generate n_c random numbers $r_1, \ldots, r_{n_c}$ from a uniform distribution between 0 and 1. The heritability of the ith causal variant is given by

$h_i = hr_i \Big/ \sum\limits_{j=1}^{n_c} r_j$. Under this assumption, we simulated K traits by

$$y_{ik} = \begin{cases} \sum\limits_{q=1}^{n_r} \beta_{kq}^r x_{qi}^r - \sum\limits_{j=1}^{n_p} \beta_{kj}^p x_{ji}^p + \varepsilon_{ik}, & 1 \le k \le L_p \\ -\left(\sum\limits_{q=1}^{n_r} \beta_{kq}^r x_{qi}^r - \sum\limits_{j=1}^{n_p} \beta_{kj}^p x_{ji}^p \right) + \varepsilon_{ik}, & L_p + 1 \le k \le L, \\ \varepsilon_{ik}, & L < k \le K \end{cases}$$

where $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iK})^T$ can be generated in the same way as generating traits of evaluating type I error, β_{kq}^r and β_{kj}^p are constants and their values depend on the heritability.

Results

For type I error evaluation, we only consider the first set of simulations, but consider different sample sizes, different significance levels, different variance models and different types of traits. In each simulation scenario, the p-values of AWRR, Single-TOW and CCA are estimated by 10,000 permutations (the p-values of WSRR are estimated by a χ^2 distribution) and the type I error rates of all of the four tests are evaluated using 10,000 replicated samples. For 10,000 replicated samples, the 95% confidence intervals (CIs) for the type I error rates at the nominal levels 0.05 and 0.01 are (0.046, 0.054) and (0.008, 0.012), respectively. The estimated type I error rates of the four tests are summarized in Tables 1, 2, S1 and S2. From these tables, we can see that only two estimated type I error rates of CCA are not within the CIs and these two type I error rates (one is 0.0126 for nominal level 0.01 in Table 1, and the other one is 0.05505 for nominal level 0.05 in Table S1) are very close to the upper bounds of the corresponding CIs, which indicates that the four tests are all valid.

For power comparisons, we consider 10 traits and we assume that all causal variants are rare. For each type of traits and each variance model, we consider different values of heritability, different percentages of protective variants, and different percentages of causal variants. In each of the simulation scenarios, the p-values of AWRR, Single-TOW and CCA are estimated using 1,000 permutations (the p-values of WSRR are estimated by a χ^2 distribution) and the power of all of the four tests is evaluated using 500 replicated samples at a significance level of 0.05.

We first consider the first set of simulations for quantitative traits under variance model 1. Figure 1 provides the power comparisons of the four tests (WSRR, AWRR, CCA and Single-TOW) for the power as a function of heritability. This figure shows that WSRR is the least powerful one and AWRR is the most powerful one. It is little complicated to compare the power of Single-TOW with the power of CCA. When genotypes impact on only one trait, Single-TOW is more powerful than CCA; otherwise, CCA is more powerful than Single-TOW. Since Single-TOW only depends on the trait that has the strongest association with genotypes, it is more favorable for Single-TOW when genotypes impact on fewer traits. Power comparisons of the four tests for the power as a function of percentage of protective variants are given by Figure 2. This figure shows that, with an increasing percentage of protective variants, the power of WSRR decreases while the power of the other three methods does not change. Other patterns of the power comparisons are similar to those

shown in Figure 1. The power comparisons of the four tests for the power as a function of the percentage of causal variants are given by Figure 3. As shown in this figure, with an increasing percentage of causal variants, the power of WSRR increases while the power of the other three methods does not change. WSRR is the least powerful method when the percentage of causal variants is small (0.15), while WSRR is the most powerful test when the percentage of causal variants is large (0.3). The patterns of the power comparisons of CCA, AWRR and Single-TOW are similar to those shown in Figure 1.

Under the first set of simulations, we also compare the powers of the four methods for quantitative traits under variance model 2 and for qualitative traits under variance models 1 and 2. These results are given in Figures S1–S9. For each type of trait, the patterns of the power comparisons are similar under variance models 1 and 2. For qualitative traits, CCA is consistently less powerful than Single-TOW and AWRR because CCA is designed for quantitative traits. For qualitative traits, the powers of AWRR, Single-TOW, and CCA decrease with the increase of the percentage of protective variants, although decrease not as fast as that of WSRR. As pointed out by Wu et al. (2011) and Sha et al. (2012), the decrease in the powers of AWRR, Single-TOW, and CCA in the presence of both risk and protective variants is due to the fact that protective variants lower MAFs in cases and thus make observing rare variants in the cases more difficult. The larger decrease in power of WSRR is additionally driven by the sensitivity to the direction of the effect due to aggregation of genotypes.

Under the second set of simulations, we compare the powers of the four methods for quantitative traits under variance model 1. Results are given in Figures S10–S12. Comparing Figures S10–S12 with Figures 1–3, the patterns of the power comparisons under the second set of simulations are very similar to those under the first set of simulations. Under the second set of simulations, we also compare the powers of the four methods for quantitative traits under variance model 2 and for qualitative traits under variance models 1 and 2 (results are not shown). Results also show that the patterns of the power comparisons under the second set of simulations are very similar to those under the first set of simulations.

In summary, for all simulation scenarios, AWRR is consistently more powerful than CCA and the power of WSRR increases with an increasing percentage of causal variants or with a decreasing o percentage of protective variants. For quantitative traits, the powers of AWRR, CCA and Single-TOW are robust to the percentage of protective variants and to the percentage of causal variants, while for qualitative traits, the powers of AWRR, CCA and Single-TOW decrease with the increasing of the percentage of protective variants and are relatively robust to the percentage of causal variants.

Discussion

In this article, we proposed the AWRR method to perform joint analysis of multiple traits in rare variant association studies based on the following reasons: (1) the development of next-generation sequencing technology has made directly testing all rare variants feasible and (2) there is increasing evidence showing that pleiotropy is a widespread phenomenon in complex diseases and multiple related traits are usually measured in genetic association

studies of complex diseases. We used extensive simulation studies to compare the performance of AWRR with CCA, WSRR and Single-TOW. Our results showed that AWRR has correct type I error rates, is robust to the directions of the association of causal variants for quantitative traits, and is robust to the percentage of causal variants. AWRR is consistently more powerful than CCA. AWRR is more powerful than Single-TOW and WSRR in most simulation scenarios.

Our simulation studies showed that the performance of each of AWRR, WSRR and Single-TOW depends strongly upon the number of traits impacted by genetic variants, the percentage of protective variants, and the percentage of causal variants. And no method demonstrates consistently good power. To increase the robustness of the test, we can combine AWRR, WSRR and Single-TOW aiming to have advantages of the three methods. Let p_{AWRR}, p_{WSRR} and $p_{Single-TOW}$ denote the p-values of AWRR, WSRR, and Single-TOW, respectively. The combined test statistic can be defined as $T_{combined} = \min\{p_{AWRR}, p_{WSRR}, p_{Single-TOW}\}$. However, the performance of the combined test needs further investigation.

In association studies based on unrelated individuals, it has been long recognized that population stratification can seriously confound association results (Knowler et al., 1988; Lander & Schork, 1994). Several methods have been developed to control for population stratification for association studies based on unrelated individuals. These methods include the GC approach (Devlin & Roeder, 1999; Devlin et al., 2001; Reich & Goldstein, 2001), the PC approach (Bauchet et al., 2007; Chen et al., 2003; Price et al., 2006; Zhang et al., 2003; Zhu et al., 2002), and the MLM approach (Kang et al., 2010; Zhang et al., 2010). Like most association tests based on unrelated individuals, AWRR is subject to bias due to population stratification. To make AWRR robust to population stratification, we can use the PC approach. Let $P_i = (p_{i1},\ldots,p_{iL})^T$ denote the first L PCs of the genotypes at a set of genomic markers for the ith individual. In step 3 of AWRR, we can use the residuals of the regression $x_i = \alpha + \beta^T P_i + \varepsilon_i$ to replace x_i and use the residuals of the regression $y_{ik} = \alpha_k + \beta_k^T P_i + \varepsilon_{ik}$ to replace y_{ik}. The performance of using the PC approach to control for population stratification in AWRR also needs further investigation.

The computation time required for running AWRR depends on the sample size, the number of variants in the genomic region, the number of traits, and the number of permutations. The running time of AWRR with 1000 permutations on the data set with 1000 individuals, 10 traits, and 100 variants in the genomic region on a laptop with four Intel Cores @ 2.00GHz and 4 GB memory is no more than 0.5s. To perform genome-wide studies, we can first select genomic regions that show evidence of association based on a small number of permutations (e.g. 1000), and then a large number of permutations can be used to test the selected regions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Andres AM, Clark AG, Shimmin L, Boerwinkle E, Sing CF, Hixson JE. Understanding the accuracy of statistical haplotype inference with sequence data of known phase. Genet Epidemiol. 2007; 31:659–671. [PubMed: 17922479]

Aschard H, Vilhjalmsson BJ, Greliche N, Morange PE, Tregouet DA, Kraft P. Maximizing the power of principal-component analysis of correlated phenotypes in genome-wide association studies. Am J Hum Genet. 2014; 94:662–676. [PubMed: 24746957]

Bauchet M, Mcevoy B, Pearson LN, Quillen EE, Sarkisian T, Hovhannesyan K, Deka R, Bradley DG, Shriver MD. Measuring European population stratification with microarray genotype data. Am J Hum Genet. 2007; 80:948–956. [PubMed: 17436249]

Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet. 2008; 40:695–701. [PubMed: 18509313]

Chen HS, Zhu X, Zhao H, Zhang S. Qualitative semi-parametric test for genetic associations in case-control designs under structured populations. Ann Hum Genet. 2003; 67:250–264. [PubMed: 12914577]

Derkach A, Lawless JF, Sun L. Robust and powerful tests for rare variants using Fisher's method to combine evidence of association from two or more complementary tests. Genet Epidemiol. 2013; 37:110–121. [PubMed: 23032573]

Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]

Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. Theor Popul Biol. 2001; 60:155–166. [PubMed: 11855950]

Feng T, Zhang S, Sha Q. A method dealing with a large number of correlated traits in a linkage genome scan. BMC Proc. 2007; 1(Suppl 1):S84. [PubMed: 18466587]

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. Variance component model to account for sample structure in genome-wide association studies. Nat Genet. 2010; 42:348–354. [PubMed: 20208533]

Klei L, Luca D, Devlin B, Roeder K. Pleiotropy and principal components of heritability combine to increase power for association analysis. Genet Epidemiol. 2008; 32:9–19. [PubMed: 17922480]

Knowler WC, Williams RC, Pettitt DJ, Steinberg AG. Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. Am J Hum Genet. 1988; 43:520–526. [PubMed: 3177389]

Korte A, Vilhjalmsson BJ, Segura V, Platt A, Long Q, Nordborg M. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nat Genet. 2012; 44:1066–1071. [PubMed: 22902788]

Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994; 265:2037–2048. [PubMed: 8091226]

Lange C, Van Steen K, Andrew T, Lyon H, Demeo DL, Raby B, Murphy A, Silverman EK, Macgregor A, Weiss ST, Laird NM. A family-based association test for repeatedly measured quantitative traits adjusting for unknown environmental and/or polygenic effects. Stat Appl Genet Mol Biol. 2004; 3 Article17.

Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, Team, N.G.E.S.P.-E.L.P. Christiani DC, Wurfel MM, Lin X. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. Am J Hum Genet. 2012; 91:224–237. [PubMed: 22863193]

Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet. 2008; 83:311–321. [PubMed: 18691683]

Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. PLoS Genet. 2009; 5:e1000384. [PubMed: 19214210]

Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet. 2010; 11:31–46. [PubMed: 19997069]

Morgenthaler S, Thilly WG. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). Mutat Res. 2007; 615:28–56. [PubMed: 17101154]

Neale BM, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, Kathiresan S, Purcell SM, Roeder K, Daly MJ. Testing for an unusual distribution of rare variants. PLoS Genet. 2011; 7:e1001322. [PubMed: 21408211]

O'Brien PC. Procedures for comparing samples with multiple endpoints. Biometrics. 1984; 40:1079–1087. [PubMed: 6534410]

O'Reilly PF, Hoggart CJ, Pomyen Y, Calboli FC, Elliott P, Jarvelin MR, Coin LJ. MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. PLoS One. 2012; 7:e34861. [PubMed: 22567092]

Ott J, Rabinowitz D. A principal-components approach based on heritability for combining phenotype information. Hum Hered. 1999; 49:106–111. [PubMed: 10077732]

Price AL, Kryukov GV, De Bakker PI, Purcell SM, Staples J, Wei LJ, Sunyaev SR. Pooled association tests for rare variants in exon-resequencing studies. Am J Hum Genet. 2010; 86:832–838. [PubMed: 20471002]

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–909. [PubMed: 16862161]

Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet. 2001; 69:124–137. [PubMed: 11404818]

Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant…or not? Hum Mol Genet. 2002; 11:2417–2423. [PubMed: 12351577]

Reich DE, Goldstein DB. Detecting association in a case-control study while correcting for population stratification. Genet Epidemiol. 2001; 20:4–16. [PubMed: 11119293]

Sattar N, Mcconnachie A, Shaper AG, Blauw GJ, Buckley BM, De Craen AJ, Ford I, Forouhi NG, Freeman DJ, Jukema JW, Lennon L, Macfarlane PW, Murphy MB, Packard CJ, Stott DJ, Westendorp RG, Whincup PH, Shepherd J, Wannamethee SG. Can metabolic syndrome usefully predict cardiovascular disease and diabetes? Outcome data from two prospective studies. Lancet. 2008; 371:1927–1935. [PubMed: 18501419]

Sha Q, Wang X, Wang X, Zhang S. Detecting association of rare and common variants by testing an optimally weighted combination of variants. Genet Epidemiol. 2012; 36:561–571. [PubMed: 22714994]

Sha Q, Zhang S. A rare variant association test based on combinations of single-variant tests. Genet Epidemiol. 2014; 38:494–501. [PubMed: 25065727]

Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, Rudan I, Mckeigue P, Wilson JF, Campbell H. Abundant pleiotropy in human complex diseases and traits. Am J Hum Genet. 2011; 89:607–618. [PubMed: 22077970]

Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet. 2013; 14:483–495. [PubMed: 23752797]

Stephens M. A unified framework for association analysis with multiple related phenotypes. PLoS One. 2013; 8:e65245. [PubMed: 23861737]

Stratton MR, Rahman N. The emerging landscape of breast cancer susceptibility. Nat Genet. 2008; 40:17–22. [PubMed: 18163131]

Tang CS, Ferreira MA. A gene-based test of association using canonical correlation analysis. Bioinformatics. 2012; 28:845–850. [PubMed: 22296789]

Teer JK, Mullikin JC. Exome sequencing: the sweet spot before whole genomes. Hum Mol Genet. 2010; 19:R145–R151. [PubMed: 20705737]

Van Der Sluis S, Posthuma D, Dolan CV. TATES: efficient multivariate genotype-phenotype analysis for genome-wide association studies. PLoS Genet. 2013; 9:e1003235. [PubMed: 23359524]

Walsh T, King MC. Ten genes for inherited breast cancer. Cancer Cell. 2007; 11:103–105. [PubMed: 17292821]

Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011; 89:82–93. [PubMed: 21737059]

Yan T, Li Q, Li Y, Li Z, Zheng G. Genetic association with multiple traits in the presence of population stratification. Genet Epidemiol. 2013; 37:571–580. [PubMed: 23740720]

Yang Q, Wang Y. Methods for Analyzing Multivariate Phenotypes in Genetic Association Studies. J Probab Stat. 2012; 2012:652569. [PubMed: 24748889]

Yang Q, Wu H, Guo CY, Fox CS. Analyze multivariate phenotypes in genetic association studies by combining univariate association tests. Genet Epidemiol. 2010; 34:444–454. [PubMed: 20583287]

Zawistowski M, Gopalakrishnan S, Ding J, Li Y, Grimm S, Zollner S. Extending rare-variant testing strategies: analysis of noncoding sequence and imputed genotypes. Am J Hum Genet. 2010; 87:604–617. [PubMed: 21070896]

Zhang S, Zhu X, Zhao H. On a semiparametric test to detect associations between quantitative traits and candidate genes using unrelated individuals. Genet Epidemiol. 2003; 24:44–56. [PubMed: 12508255]

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES. Mixed linear model approach adapted for genome-wide association studies. Nat Genet. 2010; 42:355–360. [PubMed: 20208535]

Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nat Methods. 2014; 11:407–409. [PubMed: 24531419]

Zhu X, Zhang S, Zhao H, Cooper RS. Association mapping, using a mixture model for complex traits. Genet Epidemiol. 2002; 23:181–196. [PubMed: 12214310]
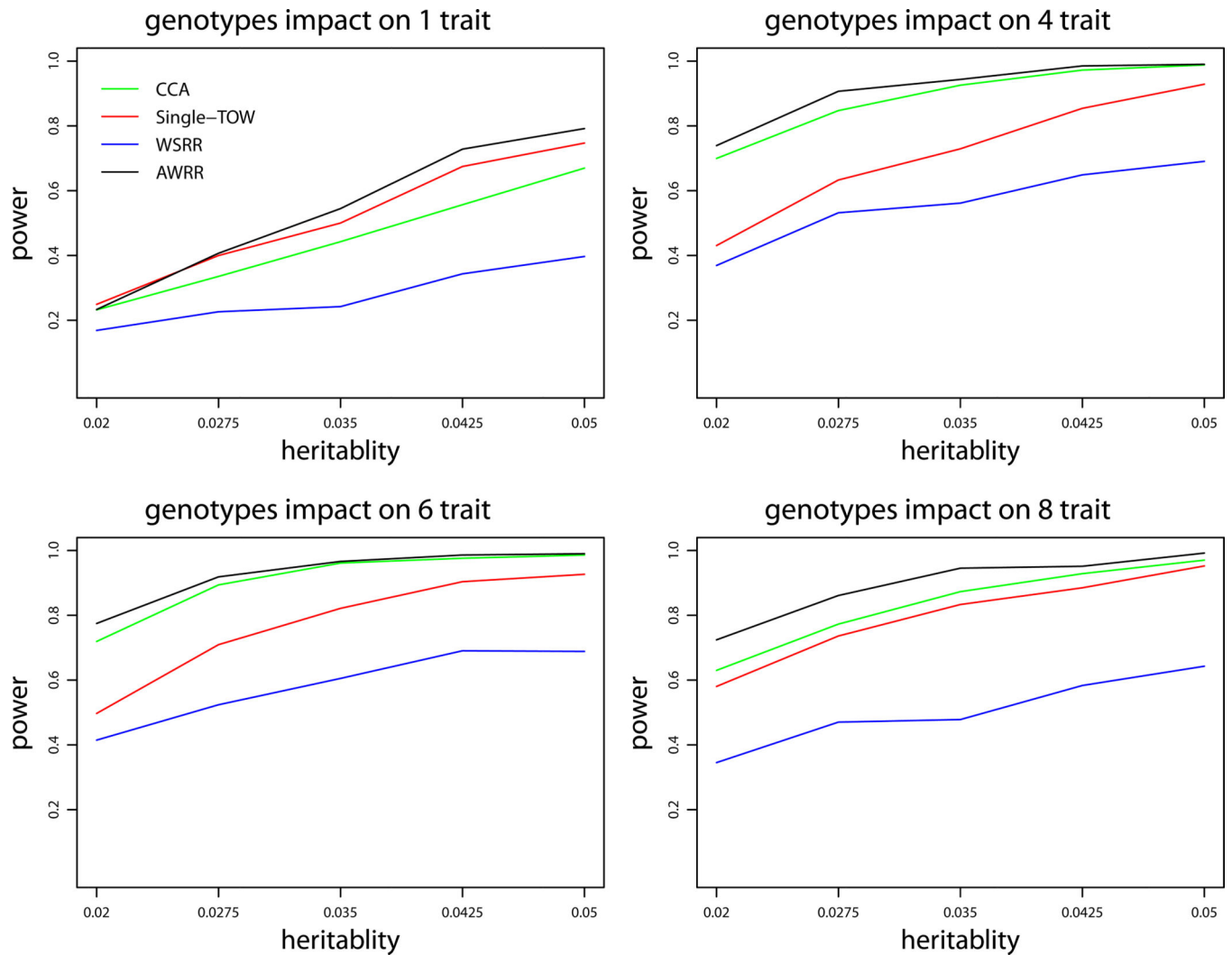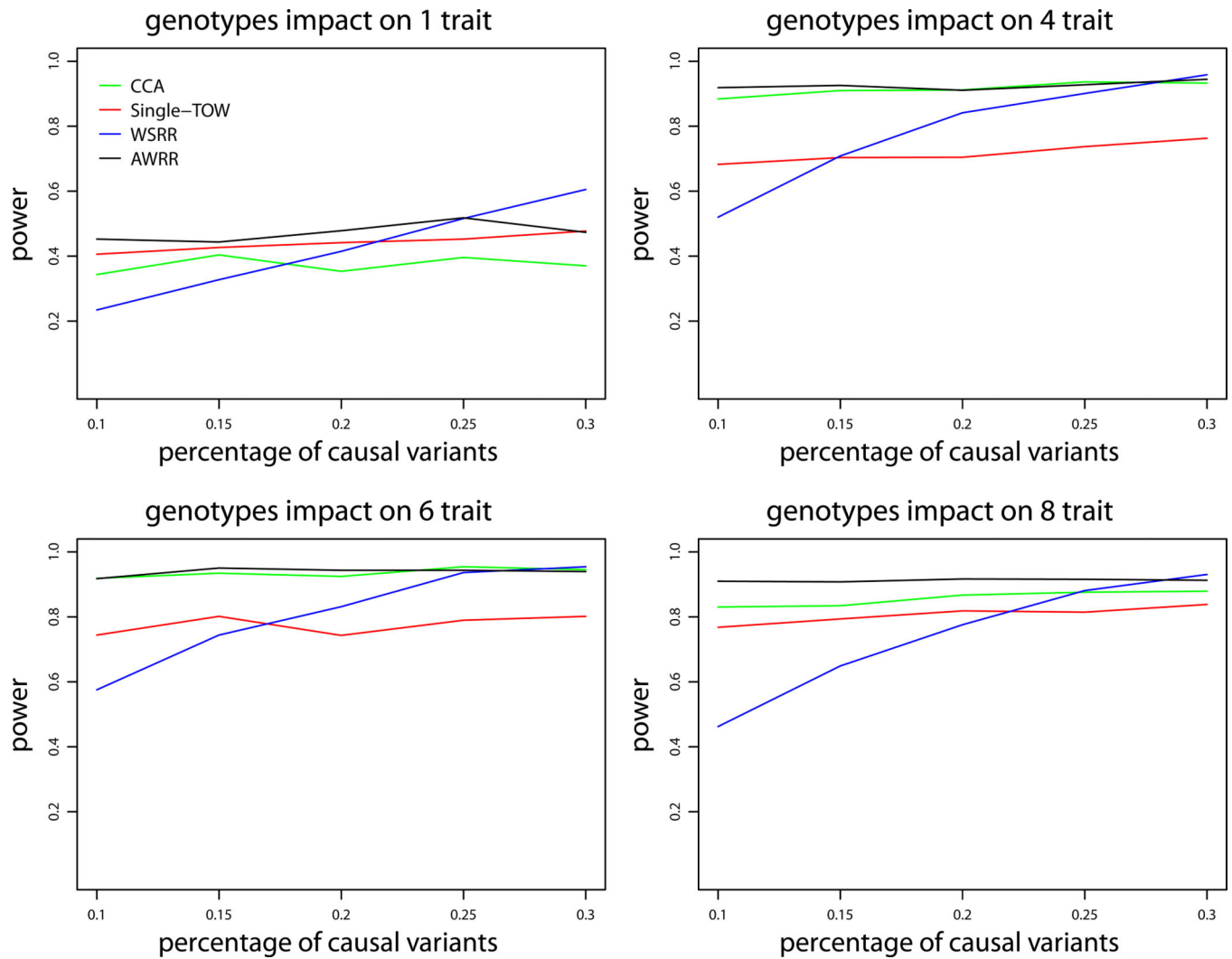
Figure 1.
Power comparisons of the four tests (WSRR, AWRR, CCA and Single-TOW) for the power as a function of heritability for quantitative traits under variance model 1. The sample size is 1000 and $\rho = 0.5$. The percentage of the causal variants is 0.1. All causal variants are risk variants. The total number of traits is 10. This set of simulations is based on Sgene1.
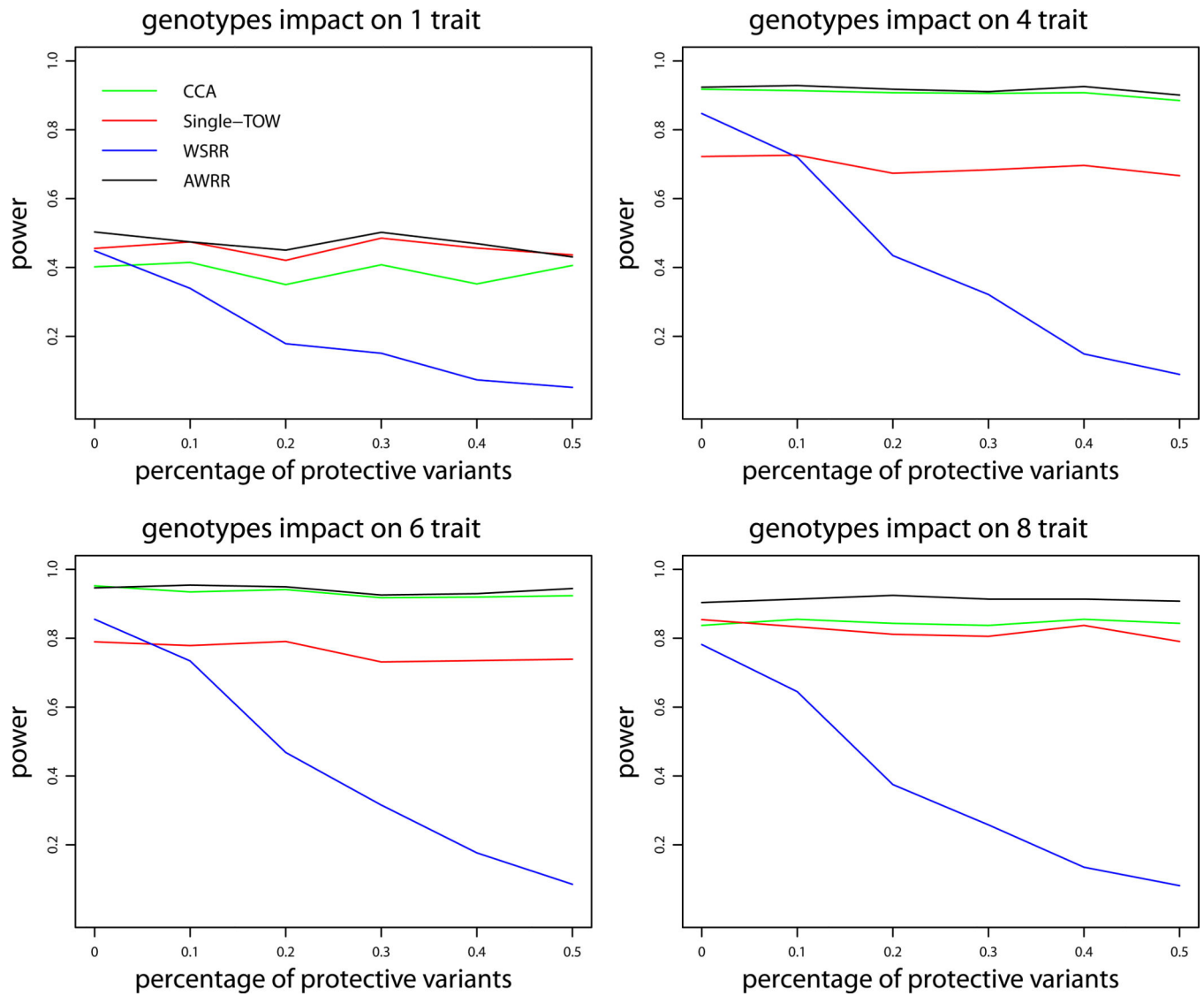
Figure 2.
Power comparisons of the four tests (WSRR, AWRR, CCA and Single-TOW) for the power as a function of percentage of protective variants for quantitative traits under variance model 1. The sample size is 1000, the percentage of causal variants is 0.2, the total heritability is 0.03, and $\rho = 0.5$. The total number of traits is 10. This set of simulations is based on Sgene1.

Figure 3.
Power comparisons of the four tests (WSRR, AWRR, CCA and Single-TOW) for the power as a function of the percentage of causal variants for quantitative traits under variance model 1. The sample size is 1000 and $\rho = 0.5$, and the total heritability is 0.03. All causal variants are risk variants. The total number of traits is 10. This set of simulations is based on Sgene1.

Table 1

The estimated type I error rates of four methods for quantitative traits under variance model 1. 10,000 replicates are used. This set of simulations is based on Sgene1.

		Sample size		
		500	**1000**	**2000**
α = 0.05	CCA	0.0518	0.0519	0.04645
	Single-TOW	0.04995	0.05255	0.0506
	WSRR	0.0464	0.0506	0.0496
	AWRR	0.0519	0.0527	0.0531
α = 0.01	CCA	0.012	0.00845	0.0126
	Single-TOW	0.0117	0.00965	0.012
	WSRR	0.0081	0.009	0.0097
	AWRR	0.01075	0.0097	0.01135

Table 2

The estimated type I error rates of four methods for qualitative traits under variance model 1. 10,000 replicates are used. This set of simulations is based on Sgene1.

		Sample size		
		500	**1000**	**2000**
α = 0.05	CCA	0.052	0.0527	0.04985
	Single-TOW	0.0519	0.0534	0.05
	WSRR	0.0502	0.0493	0.0487
	AWRR	0.054	0.0505	0.05265
α = 0.01	CCA	0.0101	0.01115	0.00955
	Single-TOW	0.0109	0.01165	0.01115
	WSRR	0.0106	0.0092	0.0106
	AWRR	0.00955	0.00965	0.012