RESEARCH

Open Access

A comprehensive comparison of multilocus association methods with summary statistics in genome-wide association studies

Zhonghe Shao[1†], Ting Wang[1†], Jiahao Qiao[1†], Yuchen Zhang[1], Shuiping Huang[1,2,3,4,5] and Ping Zeng[1,2,3,4,5*]

†Zhonghe Shao, Ting Wang and Jiahao Qiao are co-first authors

*Correspondence: zpstat@xzhmu.edu.cn

[1] Department of Biostatistics, School of Public Health, Xuzhou Medical University, Xuzhou 221004, Jiangsu, China
[2] Center for Medical Statistics and Data Analysis, Xuzhou Medical University, Xuzhou 221004, Jiangsu, China
[3] Key Laboratory of Human Genetics and Environmental Medicine, Xuzhou Medical University, Xuzhou 221004, Jiangsu, China
[4] Key Laboratory of Environment and Health, Xuzhou Medical University, Xuzhou 221004, Jiangsu, China
[5] Engineering Research Innovation Center of Biological Data Mining and Healthcare Transformation, Xuzhou Medical University, Xuzhou 221004, Jiangsu, China

Abstract

Background: Multilocus analysis on a set of single nucleotide polymorphisms (SNPs) pre-assigned within a gene constitutes a valuable complement to single-marker analysis by aggregating data on complex traits in a biologically meaningful way. However, despite the existence of a wide variety of SNP-set methods, few comprehensive comparison studies have been previously performed to evaluate the effectiveness of these methods.

Results: We herein sought to fill this knowledge gap by conducting a comprehensive empirical comparison for 22 commonly-used summary-statistics based SNP-set methods. We showed that only seven methods could effectively control the type I error, and that these well-calibrated approaches had varying power performance under the simulation scenarios. Overall, we confirmed that the burden test was generally underpowered and score-based variance component tests (e.g., sequence kernel association test) were much powerful under the polygenic genetic architecture in both common and rare variant association analyses. We further revealed that two linkage-disequilibrium-free *P* value combination methods (e.g., harmonic mean *P* value method and aggregated Cauchy association test) behaved very well under the sparse genetic architecture in simulations and real-data applications to common and rare variant association analyses as well as in expression quantitative trait loci weighted integrative analysis. We also assessed the scalability of these approaches by recording computational time and found that all these methods can be scalable to biobank-scale data although some might be relatively slow.

Conclusion: In conclusion, we hope that our findings can offer an important guidance on how to choose appropriate multilocus association analysis methods in post-GWAS era. All the SNP-set methods are implemented in the R package called MCA, which is freely available at https://github.com/biostatpzeng/.

Keywords: Genome-wide association study, Multilocus method, SNP-set analysis, Summary statistics, *P* value combination method, Common and rare variant association study, Integrative analysis, Expression quantitative trait loci

Background

Over the past two decades, genome-wide association studies (GWASs) have successfully identified a large number of genetic loci associated with many complex traits/diseases by examining millions of single nucleotide polymorphisms (SNPs) across the whole genome [1–4]. However, the contribution of associated SNPs to disease susceptibility and phenotypic variation is far less than expected, leading to the so-called problem of "missing heritability" [5–8]. One plausible interpretation for such an issue is that the single-marker analysis commonly used in GWAS is underpowered [9]; many potential genetic variants that exhibit significant but weak effects on traits/diseases have yet been discovered. As an effective supplementary strategy of single-marker analysis, multilocus methods have been increasingly applied [10]. Multilocus analysis often jointly examines a set of SNPs that are pre-defined within a functional unit such as gene to evaluate the overall association evidence at the gene level; it is thus also referred to as SNP-set or gene-based approach.

Compared to the conventional single-marker analysis, SNP-set analysis has several statistical and biological advantages. First, susceptibility genes may contain multiple independent pathogenic variants; SNP-set inference can hence substantially increase power by gathering different signals within the gene. The potential of improving power also results from the reduced burden of multiple comparisons. Second, SNP-set analysis can solve the problem of allelic heterogeneity [11], producing more consistent results across distinct studies [12]. Third, many biological processes are driven by complicated mechanisms involving more than one genetic variant; gene (or SNP-set) based inference can thus offer more biologically meaningful interpretation as genes are important functional units in living organisms [13]. Fourth, SNP-set analysis can be easily extended to pathway or network analysis [14–20]. Fifth, SNP-set analysis has already become the standard operation for rare variant association in whole genome sequencing studies [21–27]. Sixth, SNP-set analysis can easily take functional information into account [21, 28–33], which hence improves power and facilitates interpretation of GWAS discoveries. Finally, besides its own importance, SNP-set analysis is a critical step toward many other post-GWAS functional explorations, including gene-centric pleiotropy identification [34, 35], TWAS with bulk-cell sequencing RNA data [36, 37] and integrative gene analysis of GWAS with single-cell RNA sequencing data [38, 39].

Due to the usefulness, distinct SNP-set methods have been recently developed [17, 21, 25, 29, 40–51], many of which can be implemented with only GWAS summary statistics [17, 45, 52–54], greatly generalizing their applicability due to the widespread availability of summary-level data [55]. With distinct SNP-set approaches for multilocus association studies, one naturally wonders which one should be chosen in practice. Moreover, existing SNP-set methods are not used without deficiencies, potential limitations include insufficient power [56], inability to provide statistically valid tests under certain parameter settings [57], and reliance on permutation sampling [58]. Unfortunately, despite the importance of multilocus analysis in GWAS and the vast number of SNP-set methods, few comprehensive comparison studies have been performed to evaluate their effectiveness. Subsequently, due to the lack of consensus on the most suitable SNP-set method, the realization of the above advantages and benefits is to some extent currently hindered.

In the present work, we sought to fill this knowledge gap by conducting a comprehensive comparison for 22 commonly-used summary-statistics based SNP-set methods in the hope that our results could serve as an important guidance for practitioners on how to choose appropriate SNP-set analysis methods in post-GWAS era. In the following, we first evaluated the performance of these various methods in type I error control. We further assessed the power of these SNP-set methods which could maintain well-calibrated control of type I error under various simulation scenarios including common variant association analysis, rare variant association analysis and expression quantitative trait loci (eQTL) weighted integrative association analysis. We also assessed the scalability of these SNP-set approaches by recording computational time in simulation studies. Finally, corresponding to the three main simulation scenarios above, we applied these well-calibrated SNP-set methods to common variant summary statistics of six psychiatric disorders available from the Psychiatric Genomics consortium (PGC) [59, 60], rare variant summary statistics of four plasma lipid traits yielded from the Global Lipids Genetics consortium (GLGC) [61], and two-stage transcriptome-wide association study (TWAS) [31–33, 62–64] by integrating eQTL weights obtained from the Geuvadis project [65] and common variant summary statistics of nine immune-related diseases [63].

Materials and methods

Overview of SNP-set analysis methods

As a flexible and powerful strategy alternative to single-marker analysis in association studies, many SNP-set methods have been developed over the past few years [17, 21, 40–45, 51, 66–74], where a group of pre-assigned genetic variants are analyzed collectively to examine their joint influence on diseases/traits. We here have retrieved and compiled a list of 22 widely-used SNP-set methods (Table 1), which can be grouped into distinct categories in terms of input, requirement of external linkage disequilibrium (LD) and computational manner for *P* value of the aggregated test statistic. Particularly, these approaches include LD-dependent linear or quadratic combination of *Z*-scores with an additional correlation matrix accounting for dependence among SNPs (e.g., SKAT and optimal SKAT (SKATO)) [17, 45, 51], and LD-free *P* value combination methods which might be robust against correlation between SNPs (e.g., HMP and aggregated Cauchy association test (ACAT)) [54, 75].

On the other hand, some methods efficiently obtain their *P* value in an analytical way (e.g., SKAT and HMP) [17, 45, 51, 54, 75], whereas other yield *P* value via a simulation-based algorithm (e.g., GATES) [79], which would be time-consuming. Moreover, besides the general input of *Z*-scores (or *P* values) and LD matrix, some methods additionally require tuning parameters to first remove potentially null SNPs which have large *P* values [78]. From a modeling perspective, some methods (e.g., MLP and FLM) were built under the framework of fixed-effects model [45, 76, 77], whereas other (e.g., SKAT and SKATO) were established within the context of random-effect model [21, 24].

An overview of the 22 SNP-set methods with their corresponding modeling characteristic is summarized in Table 1, with technical details given in the Additional files 1 and 2. Three important applications of these SNP-set based association approaches to various genomic research fields would be discussed below (Fig. 1). The R code for implementing each method is freely available at https://github.com/biostatpzeng/MCA. It needs to

Table 1 An overview of 22 SNP-set methods and their corresponding modeling characteristics

No	Year	Method	Input					Calculate *P* value		References
			P	*Z*	*W*	R	Other	Analytical	Simulation	
1	1960	MLR		√	√	√	*N*	√		[76]
2	2008	FLM		√	√	√	*N*	√		[45, 77]
3	2004	HC		√	√			√		[66]
4	2017	GHC		√	√	√		√		[49]
5	2019	BJ		√	√			√		[51]
6	2019	GBJ		√	√	√		√		[51]
7	2020	DOT		√	√	√		√		[67]
8	2017	BT		√	√	√		√		[45]
9	2013	SKATO		√	√	√		√		[45]
10	2018	SKAT		√	√	√		√		[45]
11	1986	Simes	√					√		[68]
12	1992	FCP	√					√		[69]
13	2002	TPM	√				τ	√		[70]
14	2003	RTP	√				*k*	√		[71]
15	2007	minP	√			√		√		[72]
16	2019	ART	√				*k*	√		[78]
17	2019	ART-A	√		√		*k*	√	√	[78]
18	2007	GM	√				*a*	√		[73]
19	2008	SimpleM	√			√		√		[74]
20	2011	GATES	√			√			√	[79]
21	2019	HMP	√		√			√		[75]
22	2020	ACAT	√		√			√		[54]

P denotes a vector of *P* values, Z denotes a vector of *Z* scores, *W* is a vector of weights, R denotes the SNP-by-SNP correlation matrix, τ indicates a fixed value that *P* is less than in TPM, with the default being 0.2; *k* is the number of *P* values to be combined in RTP, ARTP, ART, ART-A, with the default value being 2/*M*, where *M* is the number of SNPs for a given gene; *a* is a shape parameter in GM, with the default being 0.0383; *N* is the sample size

MLR Multiple linear regression, *FLM* Functional multiple linear regression model, *HC* Higher criticism test, *GHC* Generalized higher criticism test, *BJ* Berk–Jones test, *GBJ* Generalized Berk–Jones test, *DOT* Decorrelation by orthogonal transformation, *BT* Burden test, *SKATO* Optimal sequence kernel association test, *SKAT* Sequence kernel association test, *Simes* Simes's test, *FCP* Fisher combined probability, *TPM* Truncated product method, *RTP* Rank truncated product, *ART* Augmented rank truncation, *ART-A* Adaptive augmented rank truncation, *GM* Gamma method, *GATES* Gene-based association test that uses extended Simes procedure, *HMP* The harmonic mean *P* value test, *ACAT* Aggregated Cauchy association test

first point out that we here did not consider some other SNP-set methods as they enjoy the similar principle of approaches described in the present work. For instance, fastBAT [80] and MAGMA [17] were constructed based on the same rationale of SKAT.

LD matrix estimation

In general, the LD matrix required in some of the these SNP-set methods (e.g., SKAT) is computed with genotypes of ancestry-matched individuals from an external reference panel such as the 1000 Genomes Project [81]. Denote **G** the standardized genotypes matrix of a given gene, and n the sample size of the reference panel. Intuitively, the empirical LD, $\hat{\mathbf{R}} = \mathbf{G}^T\mathbf{G}/(n-1)$, can be used, which however is in general not well-conditioned in the sense that the smaller eigenvalues of $\hat{\mathbf{R}}$ are underestimated because n is often not sufficiently large [82]. As a result, it would lead to inflated false discoveries. To handle this issue, many sophisticated approaches have been proposed to calculate large-dimensional covariance and correlation matrices [83]. We here estimate LD using

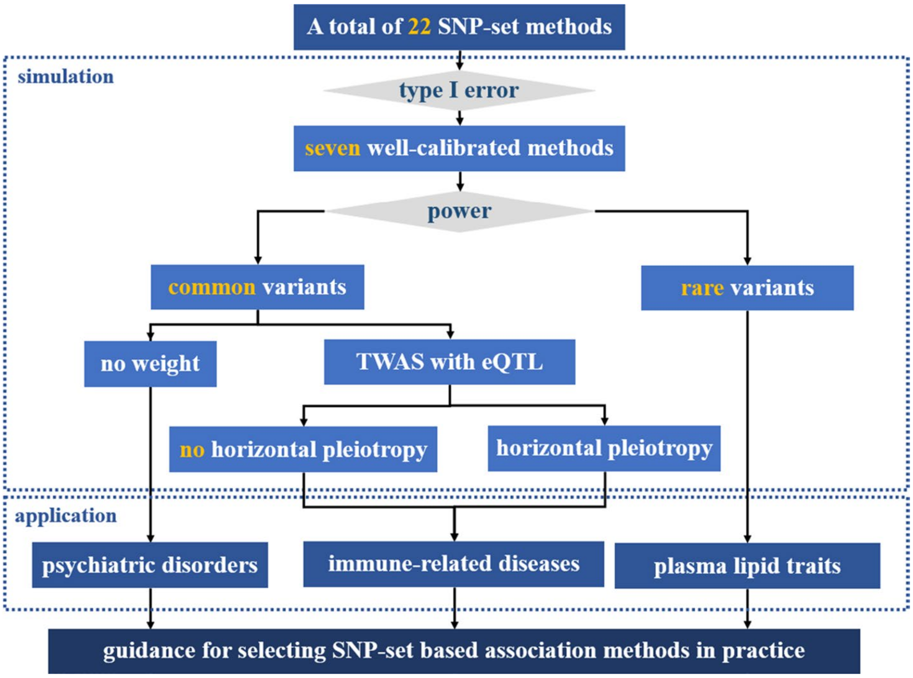

Fig. 1 Statistical analysis framework for the theoretical and application comparison of SNP-set based association methods with summary statistics

a simple, shrinkage fashion relying on the empirical one: $\mathbf{R} = \delta \times \hat{\mathbf{R}} + (1 - \delta) \times \mathbf{I}$, where δ is the shrinkage parameter and $\mathbf{I}$ is the identify matrix. We set δ to 0.95 throughout our analyses following prior studies [63, 84, 85].

Numerical studies for evaluating type I error control and power

Simulation with common variants

To evaluate the performance of each SNP-set method, we first conducted numerical studies to investigate their behaviors in type I error control and power with common SNPs (those with minor allele frequency (MAF) $\geq$ 0.05). To make our numerical studies as realistic as possible, we produced the phenotype ($\boldsymbol{Y}$) based on real genotypes of 4901 individuals available from the Wellcome Trust Case Control Consortium (WTCCC) study [86]. To this goal, we obtained a set of 550 genetic variants that were located within either 100 kb upstream of the transcription start site or 100 kb downstream of the transcription end site of the gene *CEPT1* on chr1. Note that, the selection of this gene was to some extent arbitrary. To generate the genotype matrix ($\mathbf{G}$), we randomly selected M (=50, 200 or 500) continuous SNPs to maintain LD structure, and simulated $\boldsymbol{Y}$ via a linear model $\boldsymbol{Y}=\mathbf{G}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$, with $\boldsymbol{\beta}$ the vector of effect sizes and $\boldsymbol{\varepsilon}$ the vector of normally distributed residual errors.

To assess power for every method, we made three diverse scenarios of modeling assumptions on effect sizes: (i) sparse case: among these M selected SNPs, only 5%, 20% or 50% were at random selected to have substantial impacts on $\boldsymbol{Y}$ while the remaining had zero effects, corresponding the sparse setting where only a fraction of genetic variants were causal; the non-zero effect sizes were distributed in terms of a standard normal distribution; (ii) polygenic case: all SNPs had non-zero effects on $\boldsymbol{Y}$ and their effects sizes following a standard normal distribution, or a standard double exponential distribution, or a standard t-distribution, corresponding the polygenic setting where the effect sizes of SNPs might have distinct distributions; (iii) mixed case: all SNPs had relatively small non-zero impacts on $\boldsymbol{Y}$ with their effects sizes following a standard normal distribution, but 5%, 20% or 50% SNPs were randomly selected to have additional influences, corresponding the hybrid modeling assumption made by Bayesian sparse linear mixed model [87] and latent Dirichlet process regression [64].

In all scenarios, we re-scaled the simulated SNP effect sizes on $\boldsymbol{Y}$ and residual errors so that the phenotypic variance explained (PVE) by genetic component was 0.3%, 0.5% or 1%; where $\mathrm{PVE} = \mathrm{var}(\mathbf{G}\boldsymbol{\beta})/(\mathrm{var}(\mathbf{G}\boldsymbol{\beta}) + \mathrm{var}(\boldsymbol{\varepsilon}))$ [31]. Afterwards, we performed the single-marker analysis on the phenotype with the selected M SNPs to obtain their marginal Z-scores or P values using a linear regression model [9]. These summary statistics would be taken as input to fit various SNP-set association methods, with corresponding genotypes of 503 European individuals from the 1000 Genomes Project as the reference panel to calculate LD if needed. We simply set $\boldsymbol{\beta} = 0$ and run 10^5 replications when assessing the type I error control, with the type I error primarily evaluated via the ratio between the empirical type I error and the given significance level. We repeated our numerical study 10^3 times for power evaluation, with the power calculated by the proportion of P values less than a given significance level α of 10^{-5}.

Simulation with rare variants

Some SNP-set methods (e.g., burden, SKAT and SKATO) were specially designed for analyzing rare variants although they were also often used for common variant association analysis as we assessed above; we hence performed a simulation to examine the performance of these methods in rare variant association study. First, we obtained a set of 759 rare variants ($\mathrm{MAF} < 0.05$) located within the gene *SUSD2* on chr22 from 337,198 independent individuals of European descent in the UK Biobank cohort [88]. Then, we randomly selected 15,000 individuals to generate phenotype and another 5,000 individuals as the reference panel to calculate LD. Note that these individuals were always fixed throughout this simulation. Like the same single-marker analysis in the first simulation, we conducted the simple linear regression to obtain marginal Z-scores or P values for each rare variant. Following previous work [25, 45], we calculated the weight via the beta distribution density function of MAF with the two shape parameters being 1 and 25, and further scaled these weights so that their summation was one. The parameters for type I error and power evaluations were set the same as those used in the first simulation.

Simulation by incorporating eQTL weights

For multilocus association analysis, it generally incorporates other types of omics data or functional annotation as weights, which is often more powerful than using GWAS summary statistics alone and can provide more biologically meaningful results [28, 31–33]. For example, the recently popular TWAS can be viewed as a linear weighted SNP-set analysis [89], which methodologically amounts to BT [31]; naturally, SKAT and SKATO can be considered a quadratically weighted version of TWAS [90]. The attractive property of TWAS is that it can prioritize causal genes in GWASs by incorporating eQTL weights in terms of the viewpoint of Mendelian randomization [91]. However, we do not recognize that other SNP-set methods could be interpreted in such a similar manner.

Therefore, we here conducted an additional simulation within the two-stage TWAS framework. The detail of simulation setting was described in our previous work [31]. For simplicity, in the first stage of TWAS, we only considered the polygenic case with PVE = 5% and selected 200 continuous genetic variants in the transcriptome data. Specifically, we generated eQTL weights ($\boldsymbol{w}$) and simulated gene expression ($\boldsymbol{e}$) using genotypes ($\mathbf{G}_1$) of 465 individuals from the Geuvadis project [65]; that is, $E(\boldsymbol{e}) = \mathbf{G}_1\boldsymbol{w}$ with $\mathbf{G}_1$ the genotypes of SNPs around the gene *CEPT1* on chr1. In the second stage of TWAS, we produced the phenotype ($\boldsymbol{Y}$) based on genotypes ($\mathbf{G}_2$) of *CEPT1* from WTCCC; that is $\boldsymbol{Y} = (\mathbf{G}_2\boldsymbol{w})\theta + \boldsymbol{\varepsilon}$, with $\boldsymbol{\varepsilon}$ the residual simulated from a standard normal distribution and $\theta = 0.10$ or 0.20.

The above simulation of TWAS explicitly assumed the absence of direct cis-SNP effects [92], which might be not true because of ubiquitous horizontal pleiotropy [31, 93–96]. Thus, we carried out another simulation under the case of horizontal pleiotropy by generating $\boldsymbol{Y} = (\mathbf{G}_2\boldsymbol{w})\theta + \mathbf{G}_2\boldsymbol{b} + \boldsymbol{\varepsilon}$, where $\boldsymbol{b}$ was considered random effect following a normal distribution with mean zero and variance 0.05. The setting of other parameters was the same as the case without horizontal pleiotropy.

We applied the maximum likelihood method through the computationally efficient PX-EM algorithm [96–101] to estimate joint effects (i.e., eQTL weights $\boldsymbol{w}$) for the simulated transcriptome data in the first stage, and used the linear regression model to obtain marginal *Z*-scores or *P* values for the GWAS data in the second stage [9]. Then, the estimated eQTL weights were included into these SNP-set methods via suitable transformations. Specifically, the squared eQTL weights were used for SKAT and SKATO, and the scaled absolute weights were applied in ACAT and HMP, while the original eQTL weights were employed in BT.

Real data applications

Common variant association analysis for psychiatric disorders

Psychiatric disorders are one of the most enigmatic maladies in medicine [102]; although their existence has been known for many years [60, 103] and their impact on the public health well-documented [104], relatively little remains known with regards to their causal factors and fundamental neurobiology in despite of a considerable corpus of genomic research [59, 105, 106]. Therefore, identifying potential genetic loci for early diagnosis and unraveling risk factors for prevention and treatment becomes critical in the clinic. To this goal, we applied the SNP-set methods that were demonstrated to be well-calibrated to European-only summary statistics of six psychiatric disorders yielded from PGC [59, 60] (Additional file 2: Table S1), including ADHD ($N = 53{,}293$), ASD ($N = 46{,}350$), BIP ($N = 51{,}710$), CU ($N = 184{,}765$), MDD ($N = 480{,}359$), and SCZ ($N = 77{,}096$).

We defined the set of SNPs located within a gene according to the annotation file provided by VAGIS [107], in which we considered 100 kb extension upstream of the transcription start site and 100 kb downstream of the transcription end site of that gene. Again, we leveraged genotypes of 503 European descents from the 1000 Genomes Project as the reference panel when the LD matrix was required. To avoid numerical instability, we only considered genes with at least ten SNPs following our prior work [34, 35], and further performed an enrichment analysis for all identified genes using FUMA [108].

Rare variant association analysis for four plasma lipid traits

Using these SNP-set methods, we here performed rare variant SNP-set association analysis for four plasma lipid traits (Additional file 2: Table S1), including HDL, LDL, TC, and TG. The summary statistics were publicly available from GLGC [61], which analyzed ~ 300,000 individuals of European ancestry genotyped with the HumanExome BeadChip (exome array). Following previous studies [61, 109], we considered 179,884 rare variants with MAF < 0.05 and defined the set of SNPs located within either 500 kb extension upstream of the transcription start site or 500 kb downstream of the transcription end site of a given gene in terms of the annotation file provided by GENCODE (version 12) [110]. We only analyzed 15,378 genes that contained at least two rare variants, and used genotypes from the UK Biobank [88] as the reference panel in this rare variant association analysis.

TWAS analysis for nine immune-related diseases

We finally applied these SNP-set approaches under the TWAS context. Following our prior work [31, 64], we focused on 15,810 genes and estimated eQTL weights for every gene with BSLMM [87, 111] in the Geuvadis project [65]. Because the gene expression of Geuvadis was measured in lymphoblastoid cell line, which was an immune-related cell type, we thus only considered GWAS summary statistics of nine immune-related diseases (Additional file 2: Table S1), including inflammatory bowel disease (IBD: $N = 34{,}652$), ulcerative colitis (UC: $N = 27{,}432$), Crohn's disease (CD: $N = 20{,}883$), systemic lupus erythematosus (SLE: $N = 23{,}210$), PBC ($N = 13{,}239$), primary sclerosing cholangitis (PSC: $N = 24{,}751$), rheumatoid arthritis (RA: $N = 37{,}681$), multiple sclerosis (MS: $N = 68{,}379$), and OST ($N = 63{,}608$). Details with regards to these data can be found in the original papers and the quality control procedure for data processing was described in our previous studies [31, 63, 64]. We here focused only on common SNPs and applied genotypes from the 1000 Genomes Project as the reference panel.

Results

Results of numerical studies

Assessing the type I error rate

We first evaluated the performance of type I error control for all these compared methods with common variants (Table 2) and rare variants (Additional file 2: Table S2) under the simulated null scenarios. Note that, we here defined a type I error well-controlled method as the ratio of empirical type I errors (divided by the significance level α) between 0.8 and 1.2 as done in [112, 113]. Notably, the performance of type I error control (i.e., inflated, well-controlled, or conservative) of these methods was almost consistent regardless of using common or rare variants. Among the LD-free P value combination methods, we found that only HMP, ACAT, minP and Simes generated a well-calibrated type I error control. SimpleM was conservative; in contrast, FCP, TPM,

Table 2 Ratio between the empirical type I error and the given significance level estimated over 10^5 simulations under common variants

Method	Significance level α				Performance of type I error control		
	0.05	0.01	0.001	Average	Inflated	Well-controlled	Conservative
MLR	0.00	0.00	0.00	0.00			√
FLM	0.00	0.00	0.00	0.00			√
HC	1.33	1.82	2.33	1.83	√		
GHC	1.26	1.65	1.94	1.62	√		
BJ	1.29	1.64	1.97	1.63	√		
GBJ	0.85	1.32	1.71	1.29	√		
DOT	0.00	0.00	0.00	0.00			√
BT	1.04	1.07	1.10	1.07		√	
SKAT-O	1.08	1.18	1.11	1.12		√	
SKAT	1.02	1.08	1.08	1.06		√	
Simes	0.82	0.82	0.82	0.82		√	
FCP	5.29	21.88	174.81	67.33	√		
TPM	2.45	10.39	86.81	33.22	√		
RTP	3.76	14.71	110.07	42.85	√		
minP	0.88	0.82	0.77	0.82		√	
ART	4.15	16.51	126.91	49.19	√		
ART-A	1.17	3.05	12.97	5.73	√		
GM	2.01	7.43	52.03	20.49	√		
SimpleM	0.39	0.41	0.41	0.40			√
GATES	1.47	1.53	1.51	1.50	√		
HMP	0.87	1.01	1.06	0.98		√	
ACAT	1.04	1.08	1.07	1.06		√	

Determine whether a SNP-set method was inflated, well-controlled or conservative according to the average ratio between the empirical type I error and the given significance level over 10^5 simulations. inflated: ratio > 1.2; well-controlled: $0.8 \leq$ ratio ≤ 1.2; conservative: ratio < 0.8

RTP, ART, ART-A, GM, and gene-based association test that uses extended Simes procedure (GATES) were inflated.

We also observed that not all the LD-dependent methods could behave well in controlling type I error. For example, BJ and HC, as well as their generalized versions (i.e., GBJ and GHC), were inflated under our simulation scenarios, while DOT, multiple linear regression (MLP) and functional multiple linear regression model (FLM) were much conservative. Three methods (i.e., BT, SKAT and SKATO) could effectively maintain the control of type I error. Because some of these methods failed to control the type I error at a nominal level (inflation or much conservativeness), we therefore only considered seven well-calibrated methods, including BT, SKATO, SKAT, Simes, minP, HMP and ACAT in our subsequent analyses.

Estimated statistical power for common variants with no weights

When comparing the power of the rest seven methods (Additional file 2: Table S3), we primarily displayed their results obtained under the sparse simulation setting (Fig. 2), but relegated the results of the polygenic and mixed cases to Additional file 2: Fig. S1. Particularly, we observed several important patterns as follows. First, in general, when PVE was small (e.g., 0.3%), we found that HMP and ACAT had higher power compared

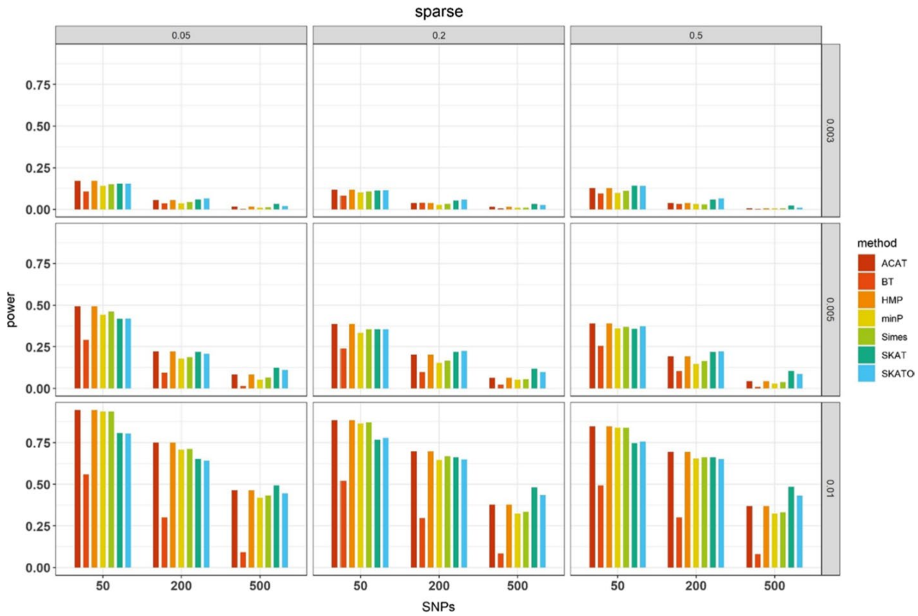

Fig. 2 Estimated power for the seven SNP-set methods under the sparse case with a significance level α of 10^{-5}. Here, PVE = 0.3%, 0.5% or 1% at the right side, the number of causal SNPs (prop) = 0.05, 0.20 or 0.50 on the top, the number of the total analyzed SNPs = 50, 200 or 500 on the x-axis. The power was estimated across 10^3 replications

to SKAT and SKATO when the number of analyzed SNPs (denoted by M) and (or) the proportion of causal SNPs (denoted by prop) were small; that is, HMP and ACAT outperform other methods when there were very less effective SNPs. However, SKAT and SKATO were better than HMP and ACAT as the increase in M and (or) prop. For example, the powers of SKAT and SKATO were 0.155 and 0.154 respectively when prop = 5% and $M = 50$, which were lower than HMP (0.171) and ACAT (0.171); whereas the powers of SKAT and SKATO were 0.020 and 0.033 respectively when prop = 5% and $M = 500$, which were more powerful than HMP (0.016) and ACAT (0.016). The similar patterns were consistently observed under the polygenic and mixed cases. Second, unlike prior studies [22, 114], as our simulations were relatively general and no very extreme settings were considered, we did not find there existed a consistent advantage of SKATO over SKAT, or vice versa; we also did not observe a substantial difference between HMP and ACAT.

Third, under the same simulation setting for causal SNPs, all these methods suffered from power loss as the number of null genetic variants increased. For example, when PVE = 1.0% and 5% of selected SNPs were causal, the power of ACAT reduced from 0.946 for 50 selected SNPs to 0.463 when the total number increased to 500. Such an observation is not unexpected because the increased noise SNPs diluted the true association signals. Fourth, both Simes and minP behaved well across all simulation settings; however, they were underpowered compared to SKAT, SKATO, HMP and ACAT even under the relatively sparse settings where only 5% of selected SNPs were causal.

Some studies previously stated that minP could exhibit higher power in the very extreme case where only one SNP showed an impact on the phenotype [51]. In order to validate such finding, we conducted an additional numerical study, in which one out of

200 SNPs was randomly causal. Under this case, we found that the power of minP was indeed higher (0.465) compared to BT (0.118), SKATO (0.267), SKAT (0.291) and Simes (0.455), but still slightly lower than HMP (0.494) or ACAT (0.495).

Fifth, as both positive and negative SNP effect sizes were simulated in all our simulation settings, BT had the lowest power across these scenarios, similar as that observed in prior work [21, 22, 25]. In order to assess the power of these methods under the situation that effect sizes of all the causal SNPs were in the same direction, we took the absolute value of simulated SNP effect sizes in the sparse case where PVE = 0.3% and prop = 5%, 20% or 50%. As expected, we observed that the power of BT was now considerably higher than that of other methods across these simulation scenarios (Table 3), in line with the prior finding [21, 25].

To be more intuitive to compare the power difference in diverse SNP-set methods, we ranked their estimated powers in each setting and averaged the rank across simulation scenarios (Additional file 2: Fig. S2). Totally, except BT, we found that SKAT, SKATO, HMP, ACAT, Simes and minP were robust and powerful under distinct simulation cases, while the SKAT, SKATO, HMP and ACAT were much better than Simes and minP. Particularly, SKAT and SKATO had a remarkable advantage under the polygenic and mixed situations, whereas HMP and ACAT seemed to outperform others in the sparse setting.

Estimated statistical power for rare variants

Finally, as can be anticipated, SKAT and SKATO, two specially designed methods for rare variant association analysis, showed evident advantage over other approaches especially when PVE was low (Fig. 3 and Additional file 2: Fig. S3). Despite not originally designing for rare variants, the two LD-free methods including ACAT and HMP also behaved satisfactorily although they were inferior relative to SKAT and SKATO across most simulation settings. For example, under the sparse case, when the number of SNPs was 200 and 50% of them were causal, the power gain of SKAT over ACAT increased from 1.7% to 56.9% when PVE reduced from 1 to 0.3% (Fig. 3). In addition, we observed that BT, Simes and minP were generally underpowered in our simulated scenarios for identifying significant association of rare variants with phenotype.

Estimated statistical power under the TWAS framework

Within the simulation of TWAS framework, we primarily performed methods which could take eQTL weights as input (i.e., ACAT, HMP, BT, SKAT and SKATO). Under the case of no horizontal pleiotropy, ACAT, HMP, SKAT and SKATO behaved comparably (Fig. 4A); while BT, analogously to the original TWAS method [32, 33], had much higher

Table 3 Estimated powers of the seven methods under sparse case where PVE = 0.3%, and prop = 5%, 20% or 50% of SNPs were randomly selected to be causal with the same direction of effect sizes

Prop	BT	SKATO	SKAT	Simes	minP	HMP	ACAT
0.05	0.350	0.065	0.059	0.044	0.037	0.054	0.054
0.20	0.379	0.058	0.062	0.039	0.035	0.051	0.051
0.50	0.363	0.066	0.058	0.038	0.038	0.047	0.047

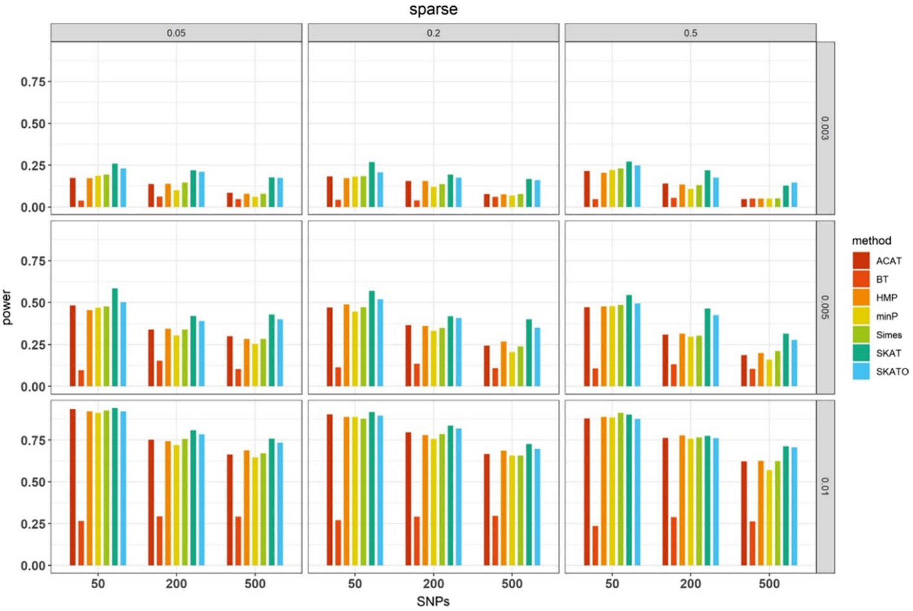

Fig. 3 Estimated power for the seven SNP-set methods in the case of rare variant association study under the sparse case with a significance level α of 10^{-5}. Here, PVE = 0.3%, 0.5% or 1% at the right side, the number of causal SNPs = 0.05, 0.20 or 0.50 on the top, the number of the total analyzed SNPs = 50, 200 or 500 on the x-axis. The power was estimated across 10^3 replications

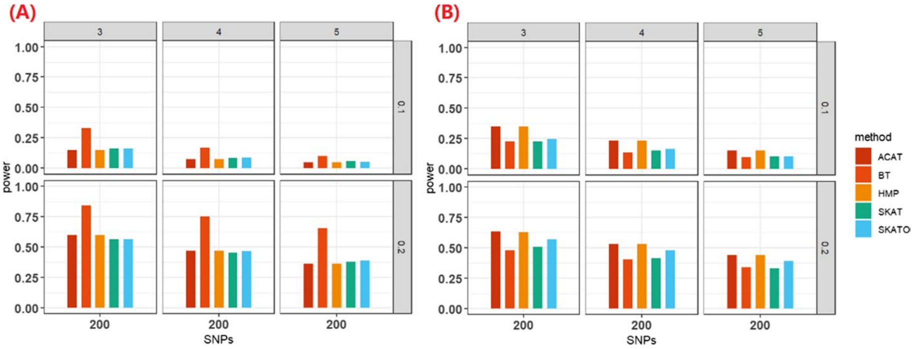

Fig. 4 (**A**) Estimated power for SNP-set methods under the polygenic TWAS framework of no horizontal pleiotropy. (**B**) Estimated power for SNP-set methods under the TWAS polygenic framework of horizontal pleiotropy. Here, $\theta = 0.1$ or 0.2 at the right side, the $-\log_{10}(\alpha) = 3$, 4, or 5 on the top, the number of the total analyzed SNPs = 50, 200 or 500 on the x-axis. The power was estimated across 10^3 replications

power, and the advantage became more pronounced as the increase of genetic effect. For example, the power advantage of BT over SKAT increased from 0.044 to 0.276 if the effect size changed from 0.1 to 0.2. In contrast, under the case of horizontal pleiotropy, BT suffered from substantial power reduction compared to other multilocus methods (Fig. 4B), in line with our previous finding [31]. Furthermore, ACAT and HMP behaved better than SKATO, and SKAT had a relatively low power among these approaches. However, as mentioned before, unlike SKAT and SKATO, ACAT and HMP cannot be explained from the perspective of TWAS analysis. To facilitate comparison, we further summarized the power performance of these methods evaluated under distinct simulation scenarios in Table 4.

Table 4 Summary performance of these SNP-set based association methods in the power evaluation of simulation studies and in real-data applications to distinct fields

	Common variants			**Rare variants**
	Unweighted	**TWAS with eQTL weights**		
		No horizontal pleiotropy	**Horizontal pleiotropy**	
Simulation	HMP ACAT	BT	HMP ACAT	SKAT SKATO
Application	HMP ACAT	HMP SKATO		SKAT SKATO

The methods listed in the table were selected in terms of their power in the simulation studies or based on the number of identified genes in the real-data applications

Table 5 Total computation times (second) of 10^3 simulations under the sparse case with PVE = 0.5% and only 20% of simulated SNPs were selected to have substantial impacts on phenotype

M	BT	SKATO	SKAT	Simes	minP	HMP	ACAT
50	4.10	483.83	7.16	3.98	4.48	4.12	3.83
200	58.59	1234.77	70.84	58.35	60.55	58.62	52.71
500	297.73	1561.44	524.48	296.46	312.76	295.70	279.46

Comparing the computing time

We here compared the running time of the seven SNP-set methods based on an Intel(R) Xeon(R) Gold 5118 CPU (2.30 GHz) with 125 GB of RAM. The total computation times across 10^3 replications are shown in Table 5 and Additional file 2: Table S4. As anticipated, it is found that the number of SNPs had substantial impact on computation time, while other simulation parameters had a negligible influence. For example, under the sparse case when the number of SNPs was 50, the average computation time of ACAT was only 3.98 s; it increased to 279.46 s when the number of SNPs was 500. Overall, LD-free methods (e.g., HMP and ACAT) were much faster than those LD-dependent ones (e.g., SKAT and SKATO). Except SKATO which was an optimization-search method, all other methods were computationally quick under various simulation scenarios, with ACAT the fastest one under the same settings.

Results of real data applications

Identified genes associated with psychiatric disorders

Applying the seven methods to psychiatric disorders, we identified a total of 588 (531 unique) genes associated with these disorders (Bonferroni-corrected P value < 0.05) (Fig. 5A), including 172 novel genes simply defined as loci not including SNPs with P value > 5×10^{-8}. More results were given in Additional file 2: Fig. S4 and Table S5. Particularly, there were 305 schizophrenia (SCZ)-associated genes but only 2 major depression disorder (MDD)-associated genes. In addition, we found that approximately 10.7% of identified genes showed pleiotropic association with at least two disorders. For example, there were 43 genes showing simultaneously significant association with SCZ and bipolar disorder (BIP), which was consistent with the highly common genetic foundation underlying the two disorders [34, 35, 105, 115–119]. We discovered that HMP identified

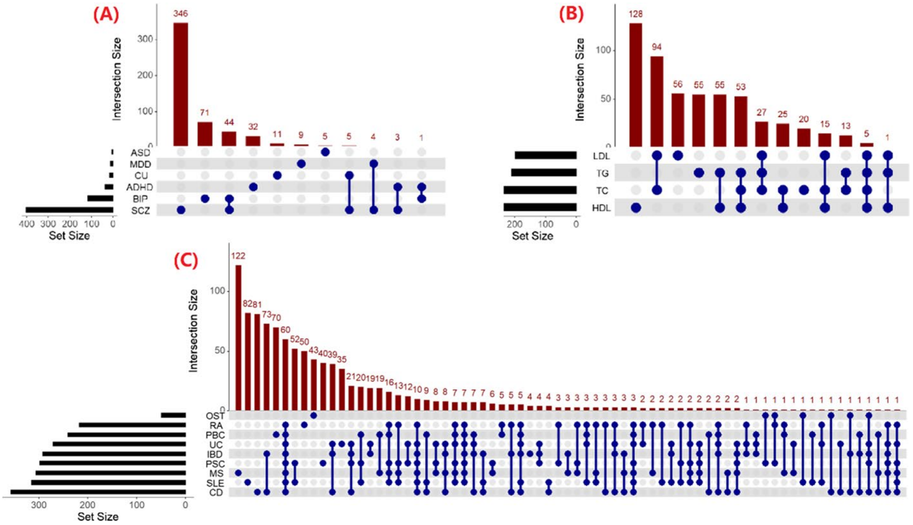

Fig. 5 Upset plot to illustrate the number of identified genes shared across distinct SNP-set methods for six psychiatric disorders (**A**), four plasma lipid traits (**B**), and nine immune-related diseases (**C**)

the most associated genes for four disorders including attention-deficit/hyperactivity disorder (ADHD: 27 genes), cannabis use (CU: 14 genes), BIP (81 genes) and SCZ (307 genes), while SKAT detected more associated genes for the remaining two disorders including autism spectrum disorder (ASD: 4 genes) and MDD (10 genes) (Table 6). The enrichment analysis demonstrated that some of these detected genes were significantly enriched in the pancreas, brain, and liver tissues (Additional file 2: Fig. S5), consistent with prior findings [34, 35].

In order to further compare HMP and SKAT in our application of psychiatric disorders, we created a bar plot for the proportion of significant cis-SNPs ($P < 5 \times 10^{-8}$) for each of the 531 unique genes (Additional file 2: Fig. S6). It was observed that the P value obtained by SKAT became more significant (smaller) than that of HMP as the increase of the proportion of significant *cis*-SNPs of an associated gene (the genetic architecture of a gene becomes from sparsity to polygenicity), which is consistent with the finding described in the simulation study above.

Identified genes associated with plasma lipid traits

When applying these SNP-set methods to rare variants of four plasma lipid traits (Fig. 5B), we found that SKAT and SKATO identified more genes associated with three lipids (except high-density-lipoprotein cholesterol (HDL)) than other approaches, and BT detected the minimal genes among all compared methods, consistent with the results given in the simulation of rare variant association analysis. Specifically, we identified 282 associated genes for HDL (Bonferroni-corrected P value < 0.05), 198 for low-density-lipoprotein cholesterol (LDL), 209 for triglyceride (TG), and 252 for total cholesterol (TC), respectively (Table 6), which involved a total of 547 unique genes (496 novel) (Table S6). Among these, 288 (52.7%) were shared in as least two lipids, and five genes (*MAP3K2*, *IMP4*, *ITGB1BP1*, *TP53I3*, and *MLLT4-AS1*) were associated with all the four lipid traits, which were confirmed by previous studies [120, 121]. In terms of

Table 6 Identified genes associated with six psychiatric disorders, four plasma lipid traits and nine immune-related diseases under various real-data applications

Phenotype	BT	SKATO	SKAT	Simes	minP	HMP	ACAT	Total
Six psychiatric disorders under the context of common variant association analysis								
ADHD	6	25	25	24	25	**27**	26	36
ASD	1	**4**	3	2	2	3	3	5
BIP	7	65	74	57	59	**81**	80	116
CU	0	10	12	10	11	**14**	**14**	16
MDD	2	**10**	9	1	3	5	5	13
SCZ	11	282	299	295	299	**307**	298	402
Four plasma lipid traits within the framework of rare variant association analysis								
HDL	22	221	222	193	192	**239**	215	282
LDL	65	147	**152**	146	147	150	144	198
TG	78	203	**205**	168	168	168	168	209
TC	146	**219**	218	198	198	197	198	252
Nine immune-related diseases under the setting of TWAS analysis								
IBD	22	146	94			**253**	249	292
UC	13	175	129			**222**	219	271
CD	22	222	144			**282**	272	357
SLE	101	180	104			**267**	266	315
PBC	102	**149**	65			122	121	240
PSC	92	210	138			**223**	221	298
RA	46	157	139			**165**	141	217
MS	106	137	183			**205**	201	306
OST	10	**36**	0			5	5	49

The maximum number of associated genes is highlighted in bold for each disease. Methods including Simes and minP which cannot incorporate eQTL weights were excluded from the TWAS analysis of the nine immune-related diseases

ADHD Attention-deficit/hyperactivity disorder, *ASD* Autism spectrum disorder, *BIP* Bipolar disorder, *CU* Cannabis use, *MDD* Major depression disorder, *SCZ* Schizophrenia, *HDL* High-density-lipoprotein cholesterol, *LDL* Low-density-lipoprotein cholesterol, *TG* Triglycerides, *TC* Total cholesterol, *IBD* Inflammatory bowel disease, *UC* Ulcerative colitis, *CD* Crohn's disease, *SLE* Systemic lupus erythematosus, *PBC* Primary biliary cirrhosis, *PSC* Primary sclerosing cholangitis, *RA* Rheumatoid arthritis, *MS* Multiple sclerosis, *OST* Osteoarthritis

the enrichment analysis, we did not find these identified genes were significantly differentially expressed in any GTEx tissues (Additional file 2: Fig. S7), which can be expected as FUMA only included common genetic variants [108]. Nevertheless, we observed suggestive evidence that these genes likely enriched in the liver, pancreas, and lymphocytes tissues, supporting by prior work [122–125].

Identified genes associated with immune-related diseases

When applying these SNP-set methods to nine immune-related diseases by incorporating eQTL information under the TWAS context, we discovered a total of 1,029 genes (446 novel) (Bonferroni-corrected P value < 0.05) (Table 6, Additional file 2: Table S7 and Fig. 5C), approximately half (48.8%) of which showed pleiotropic association with at least two diseases. It was observed that HMP identified the most associated genes (except primary biliary cirrhosis (PBC) and osteoarthritis (OST)), followed by ACAT. This observation was consistent with the simulation result. In addition, SKATO discovered more genes compared to SKAT, again in line with the corresponding simulation result. In contrast, BT detected much less significant genes. These findings further implied that these SNP loci likely showed widespread horizontal pleiotropy on the analyzed immune-related diseases. Furthermore, these detected genes were significantly enriched in the lymphocytes tissue (Additional file 2: Fig. S8), consistent with the pathological mechanism that the immune system was se associated with these diseases [126–128]. Based on the number of identified genes, we finally summarized the performance of the seven SNP-set methods in various real-data applications in Table 4.

Discussion

As part of great efforts to explain more heritability of phenotypes and enhance power in association studies by integrating other types of omics data [5], SNP-set analysis has already become a powerful alternative to single-marker analysis. In the present study, we performed a comprehensive comparison for 22 SNP-set methods that can be applied with only summary statistics. Through extensive simulation studies, we demonstrated that some LD-free methods were inflated in controlling type I error, which might be a direct consequence of not accounting for correlation between SNPs. The similar inflation pattern was also observed for some of conventional LD-free *P* value combination methods (i.e., Fisher's method) in TWAS when multiple gene expression prediction models were employed to construct weights for expression quantitative trait loci [63]. In addition, as the number of SNPs in a gene might be very large up to hundreds of thousands and often highly correlated due to pervasive LD, it was discovered that fixed-effect based methods (e.g., MLP and FLM) were generally conservative because of the loss of degree of freedom in these methods.

Particularly, among these compared methods, we only identified seven methods which could correctly control type I error, including BT, SKATO, SKAT, Simes, minP, HMP and ACAT. In total, these well-calibrated methods had varying performance in power evaluation. For example, prior studies showed minP was powerful in the case in which the association signals were extremely sparse [51, 129]. However, because of only considering the top signal across genetic variants, minP would add little to our knowledge of the association at the gene level when the top signal was genome-wide significant. In fact, minP cannot solve the primary task of SNP-set analysis because it did not consider every locus in a region and thus cannot effectively combine all available information. As a result, minP often had limited power as demonstrated in our simulations and real data applications.

By contrast, in many cases we found that integrating individual genetic variants (e.g., BT, SKAT and SKATO) together might be a more suitable manner for SNP-set analysis [21, 22, 130–132]. For instance, BT used the weighted or unweighted sum of linear test statistics [133, 134], which would have high power if all SNPs had the same effect size and the same effect direction. We also discovered that BT had better performance in eQTL integrative TWAS analysis in the absence of horizontal pleiotropy; however, BT suffered from a great power loss if the effect sizes were directionally different as demonstrated in both common and rare variant association analyses. On the other hand, SKAT and SKATO, two variance component score methods that were established with the sum of quadratic test statistics [21], were robust and particularly powerful in the presence of protective, deleterious and null variants. We demonstrate that SKAT and SKATO showed a significant advantage under the polygenic and mixed genetic architecture in common variant association study; we also confirmed the superiority of these two methods in detecting the association of rare variants with complex phenotype [21, 22, 41, 114].

Furthermore, we revealed that two LD-free methods (i.e., HMP [75] and ACAT [54]) appeared to be superior to other methods under the sparse genetic architecture in common variant association analysis. Despite not especially developing for rare variants, based on our limited experience of simulations with common variants and real-data applications, we demonstrated that ACAT and HMP also likely had the potential to be powerful methods for rare variant association analysis. In addition, the two approaches also showed better behavior in the two-stage TWAS analysis relative to other methods; unfortunately, they cannot be interpreted from the perspective of TWAS. Compared to other SNP-set aggregation methods, an important feature of ACAT and HMP is that their test statistics approximately or asymptotically follow certain null distributions (e.g., Landau distribution for HMP [75]) regardless of correlation structure between these SNPs and such an approximation is rather accurate even at very small tail area of the distribution. Consequently, one can obtain the *P* values of HMP and ACAT based on the right tail area of the respective approximate null distributions. Under regularity conditions, their performance is robust with respect to the number of SNPs, the weights, as well as the correlation among SNPs [54, 75]. Moreover, because of without requiring the knowledge of explicit correlation, compared to these LD-dependent methods (e.g., SKAT), HMP and ACAT have a wider applicability to many other cases where the correlation is too complicated to fit or reference panels cannot be available, such as multiple-tissue or multiple-model TWAS [62, 63] and spatial expression pattern identification in transcriptomic studies with multiple candidate kernels [135].

Finally, although we showed that some of these methods might be relatively slow, as all methods can be applied using GWAS summary statistics, they can be thus scalable to biobank-scale data. In summary, we evaluated 22 SNP-set methods using simulations and real data applications, and compared the robustness and effectiveness of these methods under diverse genetic architectures of phenotypes. However, our study had several limitations. First, in the real data analysis of six psychiatric disorders, we detected a number of significant genes and further showed that the identified genes may be functionally important for these disorders. However, there is no gold standard to accurately assess these methods in our real data application as the true associations of these discovered genes with the disorders are unknown; further follow-up studies are needed. Second, because of being extremely computationally expensive, we did not compared some computation-intensive SNP-set approaches (e.g., aSPUs [136] and VEGAS [107]) that utilized permutation testing rather than analytical solutions to obtain *P* values. For example, at least 10^7 samplings would be needed to calculate a sufficiently accurate *P* value for aSPUs or VEGAS if the significance level was set to 10^{-5} in each test. In fact, according to our limited simulations we found that both aSPUs (performed with the R aSPU package (Version 1.50, Updated in 2021-06-28)) and VEGAS (performed with the COMBAT package (Version 0.04, Updated in 2018-01-14)) did not have much advantage over other methods. Third, because there were too many distinct genetic backgrounds needed to study; to be simple we only focused limited settings in our simulations. Some methods might be powerful in other uncovered scenarios. For instance, GBJ exhibited excellent single-gene effect separation but showed slight inflation in our simulation settings. In addition, DOT [67] was expected to gain power as the number of SNPs increases in scenarios where effect sizes varied markedly from SNP to SNP. However, if effect sizes for all SNPs were in fact very close to each other, the power of DOT decreased and behaved conservative. Fourth, since our work focused only on summary-level data, we cannot guarantee that our conclusions could be completely generalize to the setting with individual-level data. For example, we showed that summary-statistics based SKAT outperformed summary-statistics based minP in most simulation cases of the common variant association analysis with no weights; we were however not fully clear whether this conclusion remained true in individual-level data. Nevertheless, due to the concern of privacy in individual-data sharing and widespread availability of summary-level data, our finding was certainly more important and meaningful in practice. Fifth, we did not discuss how to further pinpoint these responsible ones after discovering the overall significance for a set of SNPs with the disease or trait. The step-down inference procedure introduced in [51] may be a promising strategy that can be employed to discriminate which specific SNPs likely drive the observed association signal. We reserve this as an interesting direction for future investigation.

Abbreviations

SNPs	Single nucleotide polymorphisms
eQTL	Expression quantitative trait loci
GWASs	Genome-wide association studies
PGC	The Psychiatric Genomics consortium
GLGC	The Global Lipids Genetics consortium
TWAS	Transcriptome-wide association study
LD	Linkage disequilibrium
SKAT	Sequence kernel association test
SKATO	Optimal SKAT
ACAT	Aggregated Cauchy association test
GATES	Gene-based association test that uses extended Simes procedure
MLP	Multiple linear regression
FLM	Functional multiple linear regression model
MAF	Minor allele frequency
PVE	The phenotypic variance explained
ADHD	Attention-deficit/hyperactivity disorder
ASD	Autism spectrum disorder
BIP	Bipolar disorder
CU	Cannabis use
MDD	Major depression disorder
SCZ	Schizophrenia
HDL	High-density-lipoprotein cholesterol
LDL	Low-density-lipoprotein cholesterol
TC	Total cholesterol
TG	Triglyceride
IBD	Inflammatory bowel disease
UC	Ulcerative colitis
CD	Crohn's disease
SLE	Systemic lupus erythematosus
PBC	Primary biliary cirrhosis
PSC	Primary sclerosing cholangitis
RA	Rheumatoid arthritis
MS	Multiple sclerosis
OST	Osteoarthritis

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12859-022-04897-3.

Additional file 1. Various gene-based association analysis methods.

Additional file 2. Figure S1. Estimated power for the seven SNP-set methods under the polygenic case (**A**) and the mixed case (**B**) with a significance level α of 10^{-5}. Here, PVE = 0.3%, 0.5% or 1% at the right side, the number of causal SNPs (prop) = 0.05, 0.20 or 0.50 or the distribution of effect size including double, normal and *t* on the top, the number of the total analyzed SNPs = 50, 200 or 500 on the x-axis. The power was estimated across 10^3 replications. **Figure S2.** Rank of power for the seven SNP-set methods under the sparse case (**A**), the polygenic case (**B**), and the mixed case (**C**) with a significance level α of 10^{-5}. The number in each cell represents − log(P). normal: SNP effect sizes followed a standard normal distribution; double: SNP effect sizes followed a standard double exponential distribution; t: SNP effect sizes followed a standard t-distribution. **Figure S3.** Estimated power for the seven SNP-set methods in the case of rare variant association study under the polygenic case (**A**) and the mixed case (**B**) with a significance level α of 10^{-5}. Here, PVE = 0.3%, 0.5% or 1% at the right side, the number of causal SNPs (prop) = 0.05, 0.20 or 0.50 or the distribution of effect size including double, normal and *t* on the top, the number of the total analyzed SNPs = 50, 200 or 500 on the x-axis. The power was estimated across 10^3 replications. **Figure S4.** Upset plot to illustrate the number of identified genes shared across seven SNP-set methods for six psychiatric disorders. **Figure S5. A** Enrichment of differentially expressed pleiotropic genes associated with the six psychiatric disorders in terms of expression level across the 54 GTEx tissues. *P* values are shown in the y-axis with a scale of − log10. The bar in red represents significant enrichment after Bonferroni's adjustment for multiple hypothesis tests; **B** Top 10 significant types of pathways in terms of the GO and KEGG enrichment analyses. *BP* Biological process, *CC* Cellular component, *MF* Molecular function. **Figure S6.** Bar plot of 531 unique genes associated with the six psychiatric disorders. The red color in the heatmap represents the rank of P values of SKAT and HMP; prop: the proportion of significant cis-SNPs ($P < 5 \times 10^{-8}$) within each associated gene. **Figure S7. A** Enrichment of differentially expressed pleiotropic genes related to the four plasma lipid traits in terms of expression level across the 54 GTEx tissues. *P* values are shown in the y-axis with a scale of − log10. The bar in red represents significant enrichment after Bonferroni's adjustment for multiple hypothesis tests; **B** Top 10 significant types of pathways in terms of the GO and KEGG enrichment analyses. *BP* Biological process, *CC* Cellular component, *MF* Molecular function. **Figure S8. A** Enrichment of differentially expressed pleiotropic genes associated with the nine immune-related diseases in terms of expression level across the 54 GTEx tissues. *P* values are shown in the y-axis with a scale of − log10. The bar in red represents significant enrichment after Bonferroni's adjustment for multiple hypothesis tests; **B** Top 10 significant types of pathways in terms of the GO and KEGG enrichment analyses. *BP* Biological process, *CC* Cellular component, *MF* Molecular function. **Table S1.** Summary information of the six psychiatric disorders, four plasma lipid traits and nine immune-related diseases. **Table S2.** Ratio between the empirical type I error and the given significance level estimated over 10^5 simulations under rare variants. **Table S3.** Estimated power over 10^3 simulations with common variants. **Table S4.** Total running time of 10^3 simulations for the seven SNP-set methods under various simulation settings. **Table S5.** Identified genes associated with the six psychiatric disorders. **Table S6.** Identified genes associated with the four plasma lipid traits. **Table S7.** Identified genes associated with the nine immune-related diseases.

Acknowledgements

We thank the Psychiatric Genomics consortium, the Global Lipids Genetics consortium, and the Geuvadis project for making their GWAS summary statistics or transcriptome data publicly available for us, and thank the consortia of nine immune-related diseases for the data availability. We are grateful of all the investigators and participants contributed to these studies. This study was conducted using the UK Biobank resource under the application number 88159. The UK Biobank cohort was established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government, and the Northwest Regional Development Agency. This study also made use of data generated by the Wellcome Trust Case Control Consortium under the application number 2806. A full list of the investigators who contributed to the generation of the data was available from http://www.wtccc.org.uk; funding for the WTCCC project was provided by the Wellcome Trust under award 076113 and 085475. Finally, the data analyses in the present study were carried out with the high-performance computing cluster that was supported by the special central finance project of local universities for Xuzhou Medical University. We thank the Editor and the referees for their thorough and useful comments which substantially improved our manuscript.

Author contributions

PZ conceived the idea for the study. PZ, SH and ZS obtained and cleared the datasets; PZ, ZS, TW and JQ performed the data analyses. PZ, ZS, SH, TW and YZ interpreted the results of the data analyses. PZ and ZS wrote the manuscript with the help from other authors. All authors read and approved the final manuscript.

Funding

The research of Ping Zeng was supported in part by the National Natural Science Foundation of China (82173630 and 81402765), the Youth Foundation of Humanity and Social Science funded by Ministry of Education of China (18YJC910002), the Natural Science Foundation of Jiangsu Province of China (BK20181472), the China Postdoctoral Science Foundation (2018M630607 and 2019T120465), the QingLan Research Project of Jiangsu Province for Young and Middle-aged Academic Leaders, the Six-Talent Peaks Project in Jiangsu Province of China (WSN-087), the Training Project for Youth Teams of Science and Technology Innovation at Xuzhou Medical University (TD202008), the Postdoctoral Science Foundation of Xuzhou Medical University, and the Statistical Science Research Project from National Bureau of Statistics of China (2014LY112). The research of Ting Wang was supported in part by the Social Development Project of Xuzhou City (KC20062).

Availability of data and materials

All data generated or analyzed during this study are included in this article and its Additional files 1 and 2.

Declarations

Ethics approval and consent to participate

All methods were carried out in accordance with relevant guidelines and regulations (declaration of Helsinki).

Consent for publication

Not applicable.

Competing interests

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

Received: 27 May 2022 Accepted: 22 August 2022
Published online: 30 August 2022

References

1. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, Yang J. 10 Years of GWAS discovery: biology, function, and translation. Am J Hum Genet. 2017;101(1):5–22.
2. Klein RJ, Xu X, Mukherjee S, Willis J, Hayes J. Successes of genome-wide association studies. Cell. 2010;142(3):350–1.
3. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019;47(D1):D1005–12.
4. Loos RJF. 15 years of genome-wide association studies and no signs of slowing down. Nat Commun. 2020;11(1):1–3.
5. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al. Finding the missing heritability of complex diseases. Nature. 2009;461(7265):747–53.
6. Eichler E, Flint J, Gibson G, Kong A, Leal S, Moore J, Nadeau J. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet. 2010;11(6):446–50.
7. Gusev A, Bhatia G, Zaitlen N, Vilhjalmsson BJ, Diogo D, Stahl EA, Gregersen PK, Worthington J, Klareskog L, Raychaudhuri S. Quantifying missing heritability at known GWAS loci. PLoS Genet. 2013;9(12): e1003993.
8. Eskin E. Discovering genes involved in disease and the mystery of missing heritability. Commun ACM. 2015;58(10):80–7.
9. Zeng P, Zhao Y, Qian C, Zhang L, Zhang R, Gou J, Liu J, Liu L, Chen F. Statistical analysis for genome-wide association study. J Biomed Res. 2015;29(4):285–97.
10. Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. Am J Hum Genet. 2004;75(3):353–62.
11. Hormozdiari F, Zhu A, Kichaev G, Ju CJT, Segre AV, Joo JWJ, Won HJ, Sankararaman S, Pasaniuc B, Shifman S, et al. Widespread allelic heterogeneity in complex traits. Am J Hum Genet. 2017;100(5):789–802.
12. Zhao H, Nyholt DR. Gene-based analyses reveal novel genetic overlap and allelic heterogeneity across five major psychiatric disorders. Hum Genet. 2017;136(2):263–74.
13. Pers TH. Gene set analysis for interpreting genetic studies. Hum Mol Genet. 2016;25(R2):R133-r140.
14. Elbers CC, van Eijk KR, Franke L. Using genome-wide pathway analysis to unravel the etiology of complex diseases. Genet Epidemiol. 2009;33:419–31.
15. Zhong H, Yang X, Kaplan LM, Molony C, Schadt EE. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. Am J Hum Genet. 2010;86(4):581–91.
16. Evangelou M, Smyth DJ, Fortune MD, Burren OS, Walker NM, Guo H, Onengut-Gumuscu S, Chen W-M, Concannon P, Rich SS, et al. A method for gene-based pathway analysis using genomewide association study summary statistics reveals nine new type 1 diabetes associations. Genet Epidemiol. 2014;38(8):661–70.
17. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015;11(4): e1004219.
18. Leiserson MDM, Vandin F, Wu HT, Dobson JR, Eldridge JV, Thomas JL, Papoutsaki A, Kim Y, Niu BF, McLellan M, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. Nat Genet. 2015;47(2):106–14.
19. Yoon S, Nguyen HCT, Yoo YJ, Kim J, Baik B, Kim S, Kim J, Kim S, Nam D. Efficient pathway enrichment and network analysis of GWAS summary data using GSA-SNP2. Nucleic Acids Res. 2018;46(10):e60–e60.
20. Adolphe C, Xue A, Fard AT, Genovesi LA, Yang J, Wainwright BJ. Genetic and functional interaction network analysis reveals global enrichment of regulatory T cell genes influencing basal cell carcinoma susceptibility. Genome Med. 2021;13(1):19.
21. 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(1):82–93.
22. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, Christiani David C, Wurfel Mark M, 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(2):224–37.
23. Ionita-Laza I, Lee S, Makarov V, Buxbaum Joseph D, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. Am J Hum Genet. 2013;92(6):841–53.
24. Lee S, Abecasis Gonçalo R, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet. 2014;95(1):5–23.

25. Zeng P, Zhao Y, Liu J, Liu L, Zhang L, Wang T, Huang S, Chen F. Likelihood ratio tests in rare variant detection for continuous phenotypes. Ann Hum Genet. 2014;78(5):320–32.
26. Wang X, Zhang Z, Morris N, Cai T, Lee S, Wang C, Yu TW, Walsh CA, Lin X. Rare variant association test in family-based sequencing studies. Brief Bioinform. 2016;18:954–61.
27. Jurgens SJ, Choi SH, Morrill VN, Chaffin M, Pirruccello JP, Halford JL, Weng L-C, Nauffal V, Roselli C, Hall AW, et al. Analysis of rare genetic variation underlying cardiometabolic diseases and traits among 200,000 individuals in the UK Biobank. Nat Genet. 2022;54(3):240–50.
28. Su Y, Di C, Bien S, Huang L, Dong X, Abecasis G, Berndt S, Bezieau S, Brenner H, Caan B, et al. A mixed-effects model for powerful association tests in integrative functional genomics. Am J Hum Genet. 2018;102(5):904–19.
29. Lu H, Wei Y, Jiang Z, Zhang J, Wang T, Huang S, Zeng P. Integrative eQTL-weighted hierarchical Cox models for SNP-set based time-to-event association studies. J Transl Med. 2021;19(1):418.
30. Sun J, Zheng Y, Hsu L. A unified mixed-effects model for rare-variant association in sequencing studies. Genet Epidemiol. 2013;37(4):334–44.
31. Wang T, Qiao J, Zhang S, Wei Y, Zeng P. Simultaneous test and estimation of total genetic effect in eQTL integrative analysis through mixed models. Brief Bioinform. 2022;23:bbac08.
32. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, Eyler AE, Denny JC, Consortium GT, Nicolae DL, et al. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 2015;47(9):1091–1098.
33. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, Jansen R, de Geus EJC, Boomsma DI, Wright FA, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. 2016;48(3):245–52.
34. Lu H, Qiao J, Shao Z, Wang T, Huang S, Zeng P. A comprehensive gene-centric pleiotropic association analysis for 14 psychiatric disorders with GWAS summary statistics. BMC Med. 2021;19(1):314.
35. Wang T, Lu H, Zeng P. Identifying pleiotropic genes for complex phenotypes with summary statistics from a perspective of composite null hypothesis testing. Brief Bioinform. 2021. https://doi.org/10.1093/bib/bbab1389.
36. Luningham JM, Chen J, Tang S, De Jager PL, Bennett DA, Buchman AS, Yang J. Bayesian genome-wide TWAS method to leverage both cis- and trans-eQTL information through summary statistics. Am J Hum Genet. 2020;107(4):714–26.
37. Kim-Hellmuth S, Aguet F, Oliva M, Muñoz-Aguirre M, Kasela S, Wucher V, Castel SE, Hamel AR, Viñuela A, Roberts AL, et al. Cell type-specific genetic regulation of gene expression across human tissues. Science (New York, NY). 2020;369(6509):eaaz8528.
38. Wang R, Lin D-Y, Jiang Y. EPIC: inferring relevant cell types for complex traits by integrating genome-wide association studies and single-cell RNA sequencing. *bioRxiv* 2021:2021.2006.2009.447805.
39. Zhang MJ, Hou K, Dey KK, Jagadeesh KA, Weinand K, Sakaue S, Taychameekiatchai A, Rao P, Pisco AO, Zou J, et al. Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *bioRxiv* 2021:2021.2009.2024.461597.
40. Kwee LC, Liu D, Lin X, Ghosh D, Epstein MP. A powerful and flexible multilocus association test for quantitative traits. Am J Hum Genet. 2008;82(2):386–97.
41. Wu MC, Kraft P, Epstein MP, Taylor DM, Chanock SJ, Hunter DJ, Lin X. Powerful SNP-set analysis for case-control genome-wide association studies. Am J Hum Genet. 2010;86(6):929–42.
42. Lin X, Cai T, Wu MC, Zhou Q, Liu G, Christiani DC, Lin X. Kernel machine SNP-set analysis for censored survival outcomes in genome-wide association studies. Genet Epidemiol. 2011;35(7):620–31.
43. Schifano ED, Epstein MP, Bielak LF, Jhun MA, Kardia SLR, Peyser PA, Lin X. SNP set association analysis for familial data. Genet Epidemiol. 2012;36(8):797–810.
44. Chen H, Wang C, Conomos Matthew P, Stilp Adrienne M, Li Z, Sofer T, Szpiro Adam A, Chen W, Brehm John M, Celedón Juan C, et al. Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. Am J Hum Genet. 2016;98(4):653–66.
45. Svishcheva GR, Belonogova NM, Zorkoltseva IV, Kirichenko AV, Axenovich TI. Gene-based association tests using GWAS summary statistics. Bioinformatics. 2019;35(19):3701–8.
46. Zeng P, Zhao Y, Li H, Wang T, Chen F. Permutation-based variance component test in generalized linear mixed model with application to multilocus genetic association study. BMC Med Res Methodol. 2015;15:37.
47. Zeng P, Wang T. Bootstrap restricted likelihood ratio test for the detection of rare variants. Curr Genomics. 2015;16:194–202.
48. Wu MC, Maity A, Lee S, Simmons EM, Harmon QE, Lin X, Engel SM, Molldrem JJ, Armistead PM. Kernel machine SNP-set testing under multiple candidate kernels. Genet Epidemiol. 2013;37(3):267–75.
49. Barnett I, Mukherjee R, Lin X. The generalized higher criticism for testing SNP-set effects in genetic association studies. J Am Stat Assoc. 2017;112(517):64–76.
50. Guo B, Wu B. Powerful and efficient SNP-set association tests across multiple phenotypes using GWAS summary data. Bioinformatics (Oxford, England). 2019;35(8):1366–72.
51. Sun R, Hui S, Bader GD, Lin X, Kraft P. Powerful gene set analysis in GWAS with the Generalized Berk–Jones statistic. PLoS Genet. 2019;15(3): e1007530.
52. Liu Y, Chen S, Li Z, Morrison AC, Boerwinkle E, Lin X. ACAT: a fast and powerful *p* value combination method for rare-variant analysis in sequencing studies. Am J Hum Genet. 2019;104(3):410–21.
53. Zhang J, Zhao Z, Guo X, Guo B, Wu B. Powerful statistical method to detect disease-associated genes using publicly available genome-wide association studies summary data. Genet Epidemiol. 2019;43(8):941–51.
54. Liu Y, Xie J. Cauchy combination test: a powerful test with analytic *p*-value calculation under arbitrary dependency structures. J Am Stat Assoc. 2020;115(529):393–402.
55. Pasaniuc B, Price AL. Dissecting the genetics of complex traits using summary association statistics. Nat Rev Genet. 2016;18(2):117–27.
56. Jia P, Wang L, Meltzer HY, Zhao Z. Pathway-based analysis of GWAS datasets: effective but caution required. Int J Neuropsychopharmacol. 2011;14(4):567–72.

57. Holmans P. 7—Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. In: Dunlap JC, Moore JH, editors. Advances in genetics, vol. 72. Academic Press; 2010. p. 141–79.
58. Moskvina V, Schmidt KM, Vedernikov A, Owen MJ, Craddock N, Holmans P, O'Donovan MC. Permutation-based approaches do not adequately allow for linkage disequilibrium in gene-wide multi-locus association analysis. Eur J Hum Genet. 2012;20(8):890–6.
59. Smoller JW, Andreassen OA, Edenberg HJ, Faraone SV, Glatt SJ, Kendler KS. Psychiatric genetics and the structure of psychopathology. Mol Psychiatry. 2019;24(3):409–20.
60. Sullivan PF, Agrawal A, Bulik CM, Andreassen OA, Børglum AD, Breen G, Cichon S, Edenberg HJ, Faraone SV, Gelernter J, et al. Psychiatric genomics: an update and an agenda. Am J Psychiatry. 2018;175(1):15–27.
61. Liu DJ, Peloso GM, Yu H, Butterworth AS, Wang X, Mahajan A, Saleheen D, Emdin C, Alam D, Alves AC, et al. Exome-wide association study of plasma lipids in >300,000 individuals. Nat Genet. 2017;49(12):1758–66.
62. Xiao L, Yuan Z, Jin S, Wang T, Huang S, Zeng P. Multiple-tissue integrative transcriptome-wide association studies discovered new genes associated with amyotrophic lateral sclerosis. Front Genet. 2020;11: 587243.
63. Zeng P, Dai J, Jin S, Zhou X. Aggregating multiple expression prediction models improves the power of transcriptome-wide association studies. Hum Mol Genet. 2021;30(10):939–51.
64. Zeng P, Zhou X. Non-parametric genetic prediction of complex traits with latent Dirichlet process regression models. Nat Commun. 2017;8(1):456.
65. Lappalainen T, Sammeth M, Friedländer MR, t Hoen PAC, Monlong J, Rivas MA, Gonzàlez-Porta M, Kurbatova N, Griebel T, Ferreira PG, et al. Transcriptome and genome sequencing uncovers functional variation in humans. Nature. 2013;501(7468):506–11.
66. Donoho D, Jin J. Higher criticism for detecting sparse heterogeneous mixtures. Ann Stat. 2004;32(3):962–94.
67. Vsevolozhskaya OA, Shi M, Hu F, Zaykin DV. DOT: gene-set analysis by combining decorrelated association statistics. PLoS Comput Biol. 2020;16(4): e1007819.
68. Simes J. An improved Bonferroni procedure for multiple tests of significance. Biometrika. 1986;73:751–4.
69. Fisher RA. Statistical methods for research workers. In: Kotz S, Johnson NL, editors. Breakthroughs in statistics: methodology and distribution. New York: Springer; 1992. p. 66–70.
70. Zaykin DV, Zhivotovsky LA, Westfall PH, Weir BS. Truncated product method for combining *P*-values. Genet Epidemiol. 2002;22(2):170–85.
71. Dudbridge F, Koeleman BPC. Rank truncated product of *P*-values, with application to genomewide association scans. Genet Epidemiol. 2003;25(4):360–6.
72. Conneely KN, Boehnke M. So many correlated tests, so little time! Rapid adjustment of *P* values for multiple correlated tests. Am J Hum Genet. 2007;81(6):1158–68.
73. Zaykin DV, Zhivotovsky LA, Czika W, Shao S, Wolfinger RD. Combining *p*-values in large-scale genomics experiments. Pharm Stat. 2007;6(3):217–26.
74. Gao X, Stamier J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol. 2008;32(4):361–9.
75. Wilson D. The harmonic mean *p*-value for combining dependent tests. Proc Natl Acad Sci USA. 2019;116(4):1195–200.
76. Chow GC. Tests of equality between sets of coefficients in two linear regressions. Econometrica. 1960;28(3):591–605.
77. Wang K, Abbott D. A principal components regression approach to multilocus genetic association studies. Genet Epidemiol. 2008;32(2):108–18.
78. Vsevolozhskaya OA, Hu F, Zaykin DV. Detecting weak signals by combining small *P*-values in genetic association studies. Front Genet. 2019;10:1051.
79. Li M-X, Gui H-S, Kwan JSH, Sham PC. GATES: a rapid and powerful gene-based association test using extended Simes procedure. Am J Hum Genet. 2011;88(3):283–93.
80. Bakshi A, Zhu Z, Vinkhuyzen AA, Hill WD, McRae AF, Visscher PM, Yang J. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. Sci Rep. 2016;6:32894.
81. The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526(7571):68–74.
82. Ledoit O, Wolf M. A well-conditioned estimator for large-dimensional covariance matrices. J Multivar Anal. 2004;88(2):365–411.
83. Fan J, Liao Y, Liu H. An overview of the estimation of large covariance and precision matrices. Economet J. 2016;19(1):C1–32.
84. Cheng Q, Yang Y, Shi X, Yang C, Peng H, Liu J. MR-LDP: a two-sample Mendelian randomization for GWAS summary statistics accounting linkage disequilibrium and horizontal pleiotropy. *bioRxiv* 2019:684746.
85. Yang Y, Shi X, Jiao Y, Huang J, Chen M, Zhou X, Sun L, Lin X, Yang C, Liu J. CoMM-S2: a collaborative mixed model using summary statistics in transcriptome-wide association studies. *bioRxiv* 2019:652263.
86. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447(7145):661–78.
87. Zhou X, Carbonetto P, Stephens M. Polygenic modeling with Bayesian sparse linear mixed models. PLoS Genet. 2013;9(2): e1003264.
88. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3): e1001779.
89. Wu C, Pan W. Integrating eQTL data with GWAS summary statistics in pathway-based analysis with application to schizophrenia. Genet Epidemiol. 2018;42(3):303–16.
90. Tang S, Buchman AS, De Jager PL, Bennett DA, Epstein MP, Yang J. Novel variance-component TWAS method for studying complex human diseases with applications to Alzheimer's dementia. PLoS Genet. 2021;17(4): e1009482.

91. Zhu H, Zhou X. Transcriptome-wide association studies: a view from Mendelian randomization. Quant Biol. 2021;9(2):107–21.
92. Zeng P, Wang T, Zheng J, Zhou X. Causal association of type 2 diabetes with amyotrophic lateral sclerosis: new evidence from Mendelian randomization using GWAS summary statistics. BMC Med. 2019;17(1):225.
93. Corradin O, Saiakhova A, Akhtar-Zaidi B, Myeroff L, Willis J, Cowper-Sal R, Lupien M, Markowitz S, Scacheri PC. Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. Genome Res. 2014;24(1):1–13.
94. Ghoussaini M, French JD, Michailidou K, Nord S, Beesley J, Canisus S, Hillman KM, Kaufmann S, Sivakumaran H, Marjaneh MM, et al. Evidence that the 5p12 variant rs10941679 confers susceptibility to estrogen-receptor-positive breast cancer through FGF10 and MRPS30 regulation. Am J Hum Genet. 2016;99(4):903–11.
95. Pai AA, Pritchard JK, Gilad Y. The genetic and mechanistic basis for variation in gene regulation. PLoS Genet. 2015;11(1): e1004857.
96. Yuan Z, Zhu H, Zeng P, Yang S, Sun S, Yang C, Liu J, Zhou X. Testing and controlling for horizontal pleiotropy with probabilistic Mendelian randomization in transcriptome-wide association studies. Nat Commun. 2020;11(1):3861.
97. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Ser B (Stat Methodol). 1977;39(1):1–22.
98. Liu C, Rubin DB, Wu YN. Parameter expansion to accelerate EM: The PX-EM algorithm. Biometrika. 1998;85(4):755–70.
99. Meng X. Dyk Dv: fast EM-type implementations for mixed effects models. J R Stat Soc Ser B (Stat Methodol). 1998;60(3):559–78.
100. Liu L, Zeng P, Xue F, Yuan Z, Zhou X. Multi-trait transcriptome-wide association studies with probabilistic Mendelian randomization. Am J Hum Genet. 2021;108(2):240–56.
101. Yang C, Wan X, Lin X, Chen M, Zhou X, Liu J. CoMM: a collaborative mixed model to dissecting genetic contributions to complex traits by leveraging regulatory information. Bioinformatics. 2018;35(10):1644–52.
102. Paus T, Keshavan M, Giedd JN. Why do many psychiatric disorders emerge during adolescence? Nat Rev Neurosci. 2008;9(12):947–57.
103. Geschwind DH, Flint J. Genetics and genomics of psychiatric disease. Science. 2015;349(6255):1489–94.
104. Walker ER, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. JAMA Psychiat. 2015;72(4):334–41.
105. Lee PH, Anttila V, Won H, Feng Y-CA, Rosenthal J, Zhu Z, Tucker-Drob EM, Nivard MG, Grotzinger AD, Posthuma D, et al. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell. 2019;179(7):1469–82.
106. Sullivan PF, Geschwind DH. Defining the genetic, genomic, cellular, and diagnostic architectures of psychiatric disorders. Cell. 2019;177(1):162–83.
107. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, Hayward NK, Montgomery GW, Visscher PM, Martin NG, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet. 2010;87(1):139–45.
108. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826.
109. Luo L, Shen J, Zhang H, Chhibber A, Mehrotra DV, Tang ZZ. Multi-trait analysis of rare-variant association summary statistics using MTAR. Nat Commun. 2020;11(1):2850.
110. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012;22(9):1760–74.
111. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012;44(7):821–4.
112. Ma C, Shan G, Liu S. Homogeneity test for correlated binary data. PLoS ONE. 2015;10(4): e0124337.
113. Tang N-S, Tang M-L, Qiu S-F. Testing the equality of proportions for correlated otolaryngologic data. Comput Stat Data Anal. 2008;52(7):3719–29.
114. Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics. 2012;13(4):762–75.
115. Lichtenstein P, Yip BH, Björk C, Pawitan Y, Cannon TD, Sullivan PF, Hultman CM. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. Lancet. 2009;373(9659):234–9.
116. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P, Ruderfer DM, McQuillin A, Morris DW. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009;460(7256):748–52.
117. Ruderfer DM, Ripke S, McQuillin A, Boocock J, Stahl EA, Pavlides JMW, Mullins N, Charney AW, Ori APS, Loohuis LMO, et al. Genomic dissection of bipolar disorder and schizophrenia, including 28 subphenotypes. Cell. 2018;173(7):1705-1715.e1716.
118. Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013;381:1371–9.
119. Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, Schork AJ, Appadurai V, Buil A, Werge TM, et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science. 2018;359(6376):693–7.
120. Mo Y, Sun Y-Y, Yue E, Liu Y, Liu K-Y. MicroRNA-379-5p targets MAP3K2 to reduce autophagy and alleviate neuronal injury following cerebral ischemia via the JNK/c-Jun signaling pathway. Kaohsiung J Med Sci. 2022;38(3):230–43.
121. Yi J, An Y. Circulating miR-379 as a potential novel biomarker for diagnosis of acute myocardial infarction. Eur Rev Med Pharmacol Sci. 2018;22(2):540–6.
122. Li Z, Chiang YP, He M, Zhang K, Zheng J, Wu W, Cai J, Chen Y, Chen G, Chen Y, et al. Effect of liver total sphingomyelin synthase deficiency on plasma lipid metabolism. Biochim Biophys Acta. 2021;1866(5): 158898.
123. Saito K, Kagawa T, Tsuji K, Kumagai Y, Sato K, Sakisaka S, Sakamoto N, Aiso M, Hirose S, Mori N, et al. Plasma lipid profiling of three types of drug-induced liver injury in Japanese patients: a preliminary study. Metabolites. 2020;10(9):355.

124. Kuo PT, Huang NN. The effect of medium chain triglyceride upon fat absorption and plasma lipid and depot fat of children with cystic fibrosis of the pancreas. J Clin Investig. 1965;44(11):1924–33.
125. Zhang AY, Mysore N, Vali H, Koenekoop J, Cao SN, Li S, Ren H, Keser V, Lopez-Solache I, Siddiqui SN, et al. Choroideremia is a systemic disease with lymphocyte crystals and plasma lipid and RBC membrane abnormalities. Invest Ophthalmol Vis Sci. 2015;56(13):8158–65.
126. Suzuki K, Hayano Y, Nakai A, Furuta F, Noda M. Adrenergic control of the adaptive immune response by diurnal lymphocyte recirculation through lymph nodes. J Exp Med. 2016;213(12):2567–74.
127. Furuncuoğlu Y, Tulgar S, Dogan AN, Cakar S, Tulgar YK, Cakiroglu B. How obesity affects the neutrophil/lymphocyte and platelet/lymphocyte ratio, systemic immune-inflammatory index and platelet indices: a retrospective study. Eur Rev Med Pharmacol Sci. 2016;20(7):1300–6.
128. Schnellhardt S, Hirneth J, Büttner-Herold M, Daniel C, Haderlein M, Hartmann A, Fietkau R, Distel L. The prognostic value of FoxP3+ tumour-infiltrating lymphocytes in rectal cancer depends on immune phenotypes defined by CD8+ cytotoxic T cell density. Front Immunol. 2022;13: 781222.
129. Sun R, Lin X. Genetic variant set-based tests using the generalized Berk–Jones statistic with application to a genome-wide association study of breast cancer. J Am Stat Assoc. 2020;115(531):1079–91.
130. Fan R, Wang Y, Mills JL, Wilson AF, Bailey-Wilson JE, Xiong M. Functional linear models for association analysis of quantitative traits. Genet Epidemiol. 2013;37(7):726–42.
131. Gao Q, He Y, Yuan Z, Zhao J, Zhang B, Xue F. Gene- or region-based association study via kernel principal component analysis. BMC Genet. 2011;12:75.
132. Wu B, Guan W, Pankow JS. On efficient and accurate calculation of significance *P*-values for sequence kernel association testing of variant set. Ann Hum Genet. 2016;80(2):123–35.
133. Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. PLoS Genet. 2009;5(2): e1000384.
134. Guo B, Wu B. Statistical methods to detect novel genetic variants using publicly available GWAS summary data. Comput Biol Chem. 2018;74:76–9.
135. Sun S, Zhu J, Zhou X. Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. Nat Methods. 2020;17(2):193–200.
136. Kwak I-Y, Pan W. Adaptive gene- and pathway-trait association testing with GWAS summary statistics. Bioinformatics. 2016;32(8):1178–84.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.