

RESEARCH

Open Access

Meta-Qtest: meta-analysis of quadratic test for rare variants



Jieun Ka¹, Jaehoon Lee¹, Yongkang Kim¹, Bermseok Oh², T2D-GENES Consortium and Taesung Park^{1,3*}

From The 8th Annual Translational Bioinformatics Conference
Seoul, South Korea. 31 October - 2 November 2018

Abstract

Background: In genome-wide association studies (GWASs), meta-analysis has been widely used to improve statistical power by combining the results of different studies. Meta-analysis can detect phenotype associated variants that are failed to be detected in single studies. Especially, in biomedical sciences, meta-analysis is often necessary not only for improving statistical power, but also for reducing unavoidable limitation in data collection. As next-generation sequencing (NGS) technology has been developed, meta-analysis of rare variants is proceeding briskly along with meta-analysis of common variants in GWASs. However, meta-analysis on a single variant that is commonly used in common variant association test is improper for rare variants. A sparse signal of rare variant undermines the association signal and its large number causes multiple testing problem. To over-come these problems, we propose a meta-analysis method at the gene-level rather than variant level.

Results: Among many methods that have been developed, we used the unified quadratic tests (Q-tests); Q-test is more powerful than or as powerful as other tests such as Sequence Kernel Association Tests (SKAT). Since there are three different versions of Q-test (QTest1, QTest2, QTest3), each assumes different relationships among multiple rare variants, we extended them into meta-study accordingly. For meta-analysis, we consider two types of approaches, the one is to combine regression coefficients and the other is to combine test statistics from each single study. We extend the Q-test for meta-analysis, proposing Meta Quadratic Test (Meta-Qtest). Meta Q-test avoids the limitations of MetaSKAT. It does not only consider genetic heterogeneity among studies as MetaSKAT but also reflects diverse real situations; since we extend three different Q-tests into meta-analysis respectively, flexible Meta Q-test suggests way to deal with gene-level rare variant meta-analysis efficiently. From the results of real data analysis of blood pressure trait, our meta-analysis could successfully discovered genes, KCNA5 and CABIN1 that are already well known for relevance with hypertension disease and they are not detected in MetaSKAT.

Conclusion: As exemplified by an application to T2D Genes projects data set, Meta-Qtest more effectively identified genes associated with hypertension disease than MetaSKAT did.

Keywords: Meta-analysis, Rare variant analysis, Exome sequencing, Meta-Qtest

* Correspondence: tspark@stats.snu.ac.kr

¹Department of Statistics, Seoul National University, Seoul, South Korea

³Interdisciplinary program in Bioinformatics, Seoul National University, Seoul, South Korea

Full list of author information is available at the end of the article



Background

Genome-wide association studies (GWASs) have identified many loci that contributed to human complex traits. As genotyping technologies such as next-generation sequencing (NGS) technology evolve, we have been able to gain larger data and more accurate information on human genetics. Discovery of rare variants is one of the most valuable crops of the NGS technologies [1]. The subject of analysis naturally went over from common variants to relatively less studied rare variants, because GWASs on common variants could not entirely explain genetic-heritability, only explains small portion of expected heritability. Such phenomenon, known as “missing heritability”, posed the necessity of analyzing rare variant in human disease with a belief that rare variants play an important role in association study [2].

Persisting on same methods in common variant analyses is not appropriate for dealing with the rare variants [3]. Due to the fact that only few people share rare variants, we need a larger sample size than in common variant association test. Small sample size could markedly lower the power of a statistical test. Besides, if each variant effect is weak then single variant analysis has lower power to detect true weak signal. Therefore, in such situations instead of single variant association test, gene level test that handles multiple variants in a gene could be helpful in strengthening the signals by considering several weak ones at a time [4]. In addition to the benefit of increasing the power, gene-based multi-marker test mitigate the burden of multiple testing correction and easily interpret biological functional meaning of detected genes from the result of test. For these reasons, gene-based test is often used for rare variants analysis.

Over the past few years, various statistical methods for gene-level rare variant association test are developed. From collapsing based methods such as Combined Multivariate and Collapsing (CMC) and variable thresholds test (VT test) to variance component tests such as C-alpha and Sequence Kernel Association Tests (SKAT or SKAT-O), each method performs well in different situations [5–9]. CMC method is the one of most representative burden type tests. It unifies collapsing technique and multivariate t-test, Hotelling’s T-test. Based on variants’ minor allele frequency (MAF), variants are divided into several sub-groups, then their genotype values are summarized in 0 or 1. With collapsed genotype values, Hotelling’s T-test is conducted. Likewise, CMC or VT test is also well used burden test, but it is adaptive burden test. Compared to regular burden test like CMC, VT test allows flexible MAF threshold. Since appropriate threshold can impact on power, VT test can increase the power by choosing the optimal threshold that maximizes test statistic [6]. CMC and VT test are

powerful under the assumptions of high proportion of causal variants and same direction of their effects on certain disease; most single-nucleotide polymorphisms (SNPs) in a gene are causal and they are all protective or deleterious. When a small number of SNPs are causal and some of causal are protective and others are deleterious in a gene, burden test loses the power, and C-alpha or SKAT outperform them. Both C-alpha and SKAT are variance component test that test the variance components instead of means. C-alpha test is designed for case-control studies without covariates. Under null hypothesis that says that no variants are associated with a phenotype, for case-control data, the distribution of allele counts follows binomial distribution. It compares the observed variance of counts with expected variance. The test statistic for C-alpha includes squared terms that are observed sample variances, thus C-alpha is robust even in the presence of different directions of variants’ effects (because the signs of effects are canceled out and only their effect sizes remain in the test statistic). Despite of the advantage as noted earlier, C-alpha has some disadvantages too. *P* value is computed using permutation that requires intensive computation and covariate adjust is not available. The method proposed to solve these problems is SKAT. SKAT is variance component score test implemented in a regression framework. To test the null hypothesis of zero regression coefficients effect sizes of genetic variants in a gene, SKAT assumes that each regression coefficient follows arbitrary distribution with zero mean and the variance, product of weight and variance component. Then, testing original null hypothesis becomes the same as testing whether each variance component is zero or not. Variance component score statistic is employed in this process. Since SKAT is derived from regression model, it can include covariate terms easily. *P* value is calculated analytically and diverse kernel functions that can explain genetic similarity between individuals are introduced. Furthermore, optimal version of SKAT, SKAT-O is proposed to achieve robust power regardless of directions of variants. SKAT-O is combination of burden test and SKAT, weighted average of burden and SKAT. SKAT-O searches the optimal weight that is the weight obtaining the minimum *p* value of test statistics. However, even though robust power is substantially attained by SKAT-O test, there is still no uniformly most powerful test in all situations [10].

Gene-level Q-test is another powerful rare variant association test [11]. It also uses classical multiple linear regression as well as SKAT, but it takes Wald test based on an eigenvalue decomposition of regression coefficients. Quadratic form of test statistics in Q-test is efficiently implemented in diverse scenarios which embrace various cases in the relation of SNPs in a gene. First, Q-test considers proportion and effect sizes of causal

rare variants. Second, it considers the direction of causal variant effects. Finally, it can be used even in the presence of rare variants with common variants, together; this is the major advantage of Q-test. In other words, Q-test could achieve robust power in any case, and its exceeding power was verified in enormous simulation studies.

Meta-analysis is a popular approach to increase the power in GWASs [12]. Aggregated summary from diverse studies recovers as much information as individual-level data but without any exertion of pooling early stage data sets. In this respect, meta-study has an advantage of increasing the sample size and preserving computational efficiency [13]. Meta-analysis is sometimes essential in inevitable circumstances where individual-level data cannot be distributed although quickly advancing NGS technologies allow us to have sequencing data at smaller cost than before, producing data still requires considerable time and money. Not only because of this, but also because of releasing personal data to public is a sensitive issue, not all individual-level datasets are shared. Therefore, meta-analysis, which requires results only, becomes exceptionally useful.

The recently proposed MetaSKAT is extended SKAT for meta-analysis for gene-level rare variants. MetaSKAT aggregates the score statistics of each variants in a gene came from SKAT. Depending on the assumption of genetic effect, homogeneity or heterogeneity of genetic effects across studies, it aggregates the summary score statistics. When genetic effects are homogeneous, summary statistics are combined across the studies first and then combined across the variants in a gene, but when genetic effects are heterogeneous, the combining order is reversed. Optimal MetaSKAT that is weighted sum of test statistic of MetaSKAT and its burden for meta-analysis is also proposed to embrace the merits of burden and non-burden test together. However, simulation results of MetaSKAT show that type I error rates are somewhat uncontrolled. There is another limitation of MetaSKAT. When cohort specific genes are detected by MetaSKAT, it does not report that which cohort is highly associated with a phenotype [14].

In this paper, we extend the Q-test for meta-analysis, proposing Meta Quadratic Test (Meta-Qtest). Meta Q-test avoids the limitations of MetaSKAT. It does not only consider genetic heterogeneity among studies as MetaSKAT but also reflects diverse real situations; since we extend three different Q-tests into meta-analysis respectively, flexible Meta Q-test suggests way to deal with gene-level rare variant meta-analysis efficiently.

Materials and methods

Q-test for single study

Q-test is established in a multiple linear regression framework for a quantitative trait.

$$y_{ki} = \beta_{k0} + \sum_{j=1}^m \beta_{kj} S_{kji} + \gamma Z_{ki} + \epsilon_{ki}, \quad \epsilon_{ki} \sim N(0, \sigma_k^2), \tag{1}$$

where y_{ki} is the phenotype of the i th individual in the k th study, S_{kji} is the genotype value coded 0, 1 or 2 under an additive genotype model (dominant or recessive model is also applicable), $\beta_k = (\beta_{k0}, \dots, \beta_{km})'$ is the vector of regression coefficients for genetic effects of m SNPs, Z_{ki} is the covariate value, and γ is the corresponding vector of regression coefficients. After fitting model, using the estimated regression coefficients, β_k , QTest1 creates the new variable, pooled coefficients, $\beta_{pooled, k}$. Then, the null hypothesis of interest is to test whether pooled genetic effect exists or not:

$$H_0 : \beta_{pooled, k} = 0.$$

However, for QTest2 we test the vector of regression coefficients, β_k , rather than pooled coefficient to allow for considering bidirectional variants effect. Corresponding null hypothesis is

$$H_0 : \beta_k = \mathbf{0} \text{ or } \beta_{k0} = \beta_{k1} = \dots = \beta_{km} = 0.$$

To test the hypothesis, Q-test constructs a Wald-type statistic which has the form of quadratic statistics. Depending on the different assumptions needed to collapse the estimates of effects size parameters, there are three versions of Q-test: QTest1, QTest2, and QTest3.

Burden Test: Quadratic Test1 (QTest₁)

QTest₁ is a burden type of test. A basic assumption is that all of SNPs in a gene have the same direction of effects on the phenotype. That is, we assume that all variants within a region are all deleterious or protective. If the assumption is true, the power of the burden test becomes higher by aggregated effects of each variant. When combining variants, Qtest₁ uses the inverse variance weighting method that gives more weight to SNPs that have smaller variances,

$$\alpha_{kj} = \frac{1 / \text{var}(\hat{\beta}_{kj})}{\sum_{j=1}^m (1 / \text{var}(\hat{\beta}_{kj}))}.$$

Also, like SKAT, QTest₁ introduces MAF based weight, $w_{kj} = 1 / \sqrt{MAF_j(1-MAF_j)}$, proposed by Madssen and Browning [15]. By using the MAF based weight, our research focus on the rare SNPs is justified (rarer variant has higher probability of being causal variant). With these two different weights, aggregated effects are expressed as $\hat{\beta}_{pooled, k} = \alpha_k^T W_k \hat{\beta}_k$, where α_k is the vector

of each variant's inverse variance weight and W_k is the diagonal matrix of each variant's MAF based weight,

$$W_k = \begin{bmatrix} w_{k1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{km} \end{bmatrix}.$$

$\hat{\beta}_k$ is the vector of estimated coefficients of variants. In the ordinal regression framework, $\hat{\beta}_{pooled,k}$ follows a normal distribution with zero mean and $\alpha_k^T W_k var(\hat{\beta}_k) W_k \alpha_k$ variance. Based on the distribution of $\hat{\beta}_{pooled,k}$, the Wald type of statistics Q_{Test_1} is given by

$$Q_1 = \left(\alpha_k^T W_k var(\hat{\beta}_k) W_k \alpha_k \right)^{-1} \hat{\beta}_{pooled,k}^2 \sim \chi_1^2.$$

Non-burden test: quadratic Test2 (QTest₂)

In case of mixed variants in direction of effect, Q_{test_1} would work poorly. Thus, Q_{Test_2} assumes that some of variants are protective, and the others are negative. Instead of aggregating the effects of variants, it directly constructs Wald statistics for $\hat{\beta}_k$,

$$Q_2^{Wald} = \hat{\beta}_k^T var(\hat{\beta}_k)^{-1} \hat{\beta}_k \sim \chi_m^2,$$

where m is the number of variants in a gene. However, the number of rare SNPs in a gene is usually large, so a large degree of freedom lowers the power. Thus, Q_{Test_2} proposes a gamma method that could lower the degree of freedom [16].

$$Q_2^{Wald} = \hat{\beta}_k^T var(\hat{\beta}_k)^{-1} \hat{\beta}_k = \hat{\beta}_k^T U_k \Lambda^{-1} U_k^T \hat{\beta}_k \sim \chi_m^2,$$

where, $var(\hat{\beta}_k) = U \Lambda U^T$ and $\Lambda = diag(\lambda_k)$.

$$p_{kj} = 2 \left(1 - \Phi \left(u_{kj}^T \hat{\beta}_k / \sqrt{\lambda_{kj}} \right) \right),$$

where u_{kj} is j^{th} column of U_k .

The final form of Q_{test_2} statistics is

$$Q_2 = \sum_{j=1}^m G_{a,1}^{-1} (1-p_{kj}) \sim 1/2 \chi_{2ma}^2.$$

Optimal test (unified test): quadratic Test3 (QTest₃)

Q_{test_3} is an optimal method, weighted average of burden type (Q_{test_1}) and non-burden type (Q_{test_2}) statistics. To combine two statistics that follow chi square distributions with different degrees of freedoms, there is a step for making Q_2 has the same degree of freedom as Q_1 (df = 1). To summarize the steps, first we define the new parameter, $\hat{\beta}_k^*$ that is independent with $\hat{\beta}_{pooled,k}$, and based on the new defined $\hat{\beta}_k^*$, we can get the Q_2 statis-

tics, $Q_{2|1}^*$. Next step is to use a gamma method to make $Q_{2|1}^*$ has 1 degree of freedom ($Q_{2|1}$). Using final two statistics, Q_1 and $Q_{2|1}$, we can get the optimal Q_{test_3} statistics through a grid search of weight. Final p value of Q_3 is calculated by empirically. Pre-calculated empirical distributions are employed in the final step, so the computational burden is reduced.

Since Q_{test_3} could accommodate both scenarios that assume same or different direction of SNPs in a gene, its result is usually robust. The followings are detailed steps of Q_{test_3} .

Algorithm 1: Homo-meta Q3

Step 1. Compute $\hat{\beta}_k^* = \hat{\beta}_k - E(\hat{\beta}_k | \hat{\beta}_{pooled,k}) = \hat{\beta}_k - \hat{\beta}_{pooled,k} (\alpha_k^T W_k V_k W_k \alpha_k)^{-1} W_k V_k W_k \alpha_k$

Step 2. Compute $Q_{2|1}^* = \hat{\beta}_k^{*T} V_k^{-1} \hat{\beta}_k^* = \hat{\beta}_k^{*T} U_k^* \Lambda_k^{-1} U_k^{*T} \hat{\beta}_k^* \sim \chi_{m-1}^2$

Step 3. Compute $Q_{2|1} = 2G_{0.5,1}^{-1}(1-p_{2|1})$

Step 4. Compute $Q_3^\pi = (1-\pi)Q_1 + \pi Q_{2|1}$, for $\pi = 0, 0.1, 0.2, \dots, 0.9, 1$

Step 5. Obtain p value p_3^π from mixture of χ_1^2

Step 6. Final p value for Q_3 is calculated from the empirical distribution,

$$\hat{\pi}^{optimal} = Argmin_{\pi} \{p_3^\pi, \pi = 0, 0.1, 0.2, \dots, 0.9, 1\}$$

In step 1 in algorithm 1, we compute $\hat{\beta}_k^*$ to make $\hat{\beta}_k^* \perp \hat{\beta}_{pooled,k}$. In step 2, we compute $Q_{2|1}^*$, where $V_k^* = var(\hat{\beta}_k^*)$ and U_k^* consists of eigenvalue vectors of V_k^* , and Λ_k^* is the diagonal matrix whose diagonal elements are eigenvalues of V_k^* . By step 3, we could obtain $Q_{2|1}$ which follows χ_1^2 distribution, where $p_{2|1}$ is obtained from $Q_{2|1}^*$.

MetaQ-test for Meta analysis

Meta-analysis requires summary statistics from each single study. P value and z-score have been conventional summary statistics that could combine the results across study, but in GWAS meta-analysis, the methods using p value and z-score are generally inferior to the model based meta-analysis; those methods cannot efficiently take account of the between-data-set heterogeneity [17]. In model based meta-analysis, there are two different approaches, fixed effect model and random effect model [18]. Under the fixed effect model, effect sizes of all studies are presumed to be same, in other case, to be different [19]. MetaQ-test considers both cases. Consequently, each Q_{Test} has fixed and random versions of meta-analysis.

Another significant feature in MetaQ-test is it is extended keeping the original statistical model structure, structure of Q_{Test} . Q_{test_1} is burden type and Q_{test_2} is non-burden type test. This fact is also applied to MetaQ-test. Meta Q_{test_1} is a burden and Meta Q_{test_2} is a non-burden type test. MetaQ-test keeps not only type of statistics but also the process derived the statistics. By doing that, it can consistently maximize merits of test statistics.

Summary statistics used and the way to synthesize them are essence in meta-analysis [20]. Depending on the type of model-based meta-analysis (fix or random) and the type of test statistics (burden or non-burden), meta-analysis requires different input values and takes different approaches for combining those values.

Burden test: Meta quadratic Test1 (metaQtest₁)
Meta-analysis assuming homogeneous genetic effects
across studies: Homo-meta-Q₁

First, we extend Qtest1 to meta-analysis maintaining burden type test. We assume that all variants in a gene across studies are from one single study. In Qtest1, to aggregate the effects of variants, $\hat{\beta}_{pooled.k} = \alpha_k^T W_k \hat{\beta}_k$ is introduced. Similarly, we introduce $\hat{\beta}_{pooled}$ that adds up all of variants throughout studies. The way of combining each variant's effect is exactly same as Qtest₁, weighted sum of effect sizes of all variants. When we suppose that there are K studies, the estimated pooled effect size of regression coefficients is $\hat{\beta}_{pooled} = \alpha^T W \hat{\beta} = \sum_{j=1}^m \alpha_j w_j \hat{\beta}_j$, where j is the index for variants, $\hat{\beta}$ is the column vector composed of $\hat{\beta}_j$, α is the column vector that includes α_j s (j = 1, 2, ..., m) as components, and W is the diagonal matrix with w_j s as diagonal elements.

$$\beta_j = \sum_{k=1}^K \frac{n_k}{I_1 * n_1 + I_2 * n_2 + \dots + I_K * n_K} \hat{\beta}_{kj}, \text{ for } j = 1, 2, \dots, m,$$

$$\beta = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_m)^T,$$

$$\alpha_j = \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + I_2 * n_2 + \dots + I_K * n_K} \right)^2 \alpha_{kj}, \text{ for } j = 1, 2, \dots, m,$$

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)^T,$$

$$w_j = \sum_{k=1}^K \frac{n_k}{I_1 * n_1 + I_2 * n_2 + \dots + I_K * n_K} w_{kj}, \text{ for } j = 1, 2, \dots, m,$$

$$W = \begin{bmatrix} w_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & w_m \end{bmatrix}.$$

It is natural that we collapse the variants in the same locus with appropriate weights, such as sample size. Testing hypothesis in this case is $H_0: \beta_{pooled} = 0$. Since under the null assumption, $\hat{\beta}_{pooled}$ is normally distributed with zero mean and variance $\alpha^T W V W \alpha$, the Wald type of test statistics is

$$Q_{hom-meta-q1} = \left(\alpha^T W var(\hat{\beta}) W \alpha \right)^{-1} \hat{\beta}_{pooled}^2.$$

Since the statistics for meta-analysis is constructed based on the preexisting result of individual level data analysis, $var(\hat{\beta})$ cannot be estimated by the raw data. However, if we assume the independence among studies,

we can define $var(\hat{\beta})$ as the block diagonal matrix that has diagonal with $var(\hat{\beta}_k)$,

$$V = var(\hat{\beta}) = var\left((\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_m)^T \right) = \begin{bmatrix} \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + \dots + I_K * n_K} \right)^2 cov(\hat{\beta}_{k1}, \hat{\beta}_{k1}) & \dots & \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + \dots + I_K * n_K} \right)^2 cov(\hat{\beta}_{k1}, \hat{\beta}_{km}) \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + \dots + I_K * n_K} \right)^2 cov(\hat{\beta}_{km}, \hat{\beta}_{k1}) & \dots & \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + \dots + I_K * n_K} \right)^2 cov(\hat{\beta}_{km}, \hat{\beta}_{km}) \end{bmatrix}.$$

We can easily check that $Q_{hom-meta-q1}$ follows chi-square distribution with 1 degree of freedom, because it is the square of single standard normal random variable, $\hat{\beta}_{pooled}$. To draw the statistics, $Q_{hom-meta-q1}$, we need estimated $\hat{\beta}_{kj}$ s from each study, its variance and MAF weights as inputs. Thus, we call this approach beta-based approach, also because this is burden type, it is like assuming homogeneous effect (same regression coefficients) across the studies.

Meta-analysis assuming heterogeneous genetic effects
across studies: Het-meta-Q₁

Assuming a variant may have a different effect across studies, we can consider the case of heterogeneous effects over studies. This assumption is in accordance with a meta model with random effects. Since we allow the heterogeneity of effect, we can use the results of each study's regression fit. The model fittings were carried out separately, so the test statistics for association are made of different estimated regression coefficients. Accordingly, the outcome of model fitting in each study that takes account of heterogeneity is appropriate for summary statistics for meta-analysis.

The fact that Q_1 follows the chi-square distribution also makes us combine Q_1 s easily. Thus, test statistics of each study would be a naive and handy summary statistic for meta-analysis

$$Q_{het-meta-q1} = \sum_{k=1}^K Q_1 = \sum_{k=1}^K \left(\alpha_k^T W_k var(\hat{\beta}_k) W_k \alpha_k \right)^{-1} \hat{\beta}_{pooled.k}^2$$

However, $Q_{het-meta-q1}$ is a statistic for burden test in respect of sum of single burden statistics. Wald type of statistics for a single variant is distributed as the chi-square distribution with one degrees of freedom. When we assume the independence between studies, then the sum of test statistics also follows a chi-square distribution with K degree freedom, where K is the number of studies in meta-analysis, $Q_{het-meta-q1} \sim \chi_K^2$. In the process of deriving $Q_{het-meta-q1}$, we only require the test statistics. Therefore, we call this approach statistics based meta.

Non-burden test: Meta quadratic Test2 (metaQtest₂)
Meta-analysis assuming homogeneous genetic effects
across studies: Homo-meta-Q2

We extend Qtest2 to meta-analysis in the same manner as the Meta-Qtest1: fixed(homo) and random(hetero) version of meta. However, the main difference between Meta-Qtest1 is that we do not combine the effects of different variants, we dose only collapse the same variants. The basic idea of Meta-Qtest2 is that we regard SNPs on the same locus as the single SNP. For this reason, this approach is different with burden meta-analysis.

It is possible to collapse variants of same locus, because of the assumption of homogeneity in genetic effects size throughout studies. In order to make representative genetic effects size that combines SNPs on the same locus, we use weighted sum. Weights are given as proportional to sample size of study. Representative of regression coefficient of jth SNP is expressed as below,

$$\hat{\beta}_j = \sum_{k=1}^K \frac{n_k}{I_1 * n_1 + I_2 * n_2 + \dots + I_K * n_K} \hat{\beta}_{kj}, \text{ for } j = 1, 2, \dots, m.$$

For the sake of simplicity, we assume that the number of variants in a gene is the same across the study, $m = m_1 = m_2 = \dots = m_K$.

Since all populations do not share the same variants commonly, some variants exit only in a specific population. We thus put an indicator function in front of index of sample size. To maintain the framework of Qtest₂, we also need variance-covariance matrix of estimated regression coefficients. Under the independence assumption of studies, we derive the variance-covariance matrix, V, analogous to $\hat{\beta}_j$ using the sample size proportional weights,

$$V = \text{var}(\hat{\beta}) = \text{var}(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_m)^T = \begin{bmatrix} \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + \dots + I_K * n_K}\right)^2 \text{cov}(\hat{\beta}_{k1}, \hat{\beta}_{k1}) & \dots & \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + \dots + I_K * n_K}\right)^2 \text{cov}(\hat{\beta}_{k1}, \hat{\beta}_{km}) \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + \dots + I_K * n_K}\right)^2 \text{cov}(\hat{\beta}_{km}, \hat{\beta}_{k1}) & \dots & \sum_{k=1}^K \left(\frac{n_k}{I_1 * n_1 + \dots + I_K * n_K}\right)^2 \text{cov}(\hat{\beta}_{km}, \hat{\beta}_{km}) \end{bmatrix}$$

The following step is identical to QTest2. Using eigen-decomposed $V = U\Lambda U^T$, we construct the Wald type statistic,

$$Q_{2-meta}^{Wald} = \hat{\beta}^T V^{-1} \hat{\beta} = \hat{\beta}^T U \Lambda^{-1} U^T \hat{\beta} \sim \chi_m^2,$$

where $\Lambda = \text{diag}(\lambda)$.

$$p_j = 2(1 - \Phi\left(\frac{u_j^T \hat{\beta}_j}{\sqrt{\lambda_j}}\right)),$$

where $u_j = j^{th}$ column of U,

$$Q_{hom-meta-q2} = \sum_{j=1}^m G_{a,1}^{-1}(1-p_j) \sim 1/2 \chi_{2ma}^2.$$

Meta-analysis assuming heterogeneous genetic effects
across studies: Het-meta-Q2

For rare variant analysis, overlapped variants with other studies are not many. Thus, the approach of homo-meta-Q2 that aggregates the overlapped SNPs in the same locus could not be appropriate for the rare variant analysis; the estimate of variance-covariance matrix is unstable. An easy alternative way is to assume not homogeneity but heterogeneity of genetic effects. To satisfy this assumption and to accommodate different directional effects among variants in a region, we use the result of single study.

Combining single study results using p value is conventional, but this method has a limitation upon achievement of heterogeneity among studies; it just assumes independence among studies. Instead of p-value based method, we combine test statistics of each study for Qtest2. Q₂ follows a mixture of chi-square distributions and the summation of the Q₂ then also follows a mixture of chi-square distributions.

$$Q_{het-meta-q2} = \sum_{k=1}^K Q_2 = \sum_{j=1}^m G_{a,1}^{-1}(1-p_{kj}) \sim 1/2 \sim 1/2 \chi_{2maK}^2.$$

In Homo/Hetero-meta-Q2, the value of a determines the shape of the mixture of chi-square distribution and this value depends on the gene size (number of SNPs in a gene). Therefore, the choice of a should be careful.

Optimal test (unified test): Meta quadratic Test3 (MetaQtest₃)

We develop an optimal quadratic meta-analysis for robust power gain. A burden type test is known to be more powerful when most variants in a region are causal and their effect directions are the same, but a non-burden type is opposite to this case. Therefore, applying one of the two tests can lose the power to detect the associated variants with a phenotype, if every single gene dose not satisfy the same assumption. A unified approach is to use a weighted average of burden and non-burden test. Since this approach allows two conflicted assumptions about direction among variants (burden or non-burden), an optimal test usually achieves the robust power regardless of directional assumptions.

Meta-analysis assuming homogeneous genetic effects
across studies: Homo-meta-Q3

When studies are homogeneous, we suggest Q_{hom-meta-q3} that is weighted average of Q_{hom-meta-q1} and Q_{hom-meta-q2}. We build a test for meta with an adjustment for homogeneous case. The following are steps for

constructing $Q_{hom-meta-q3}$. The steps are identical to Qtest3 but for substituting collapsed variables, $(\hat{\beta}, \hat{\beta}_{pooled}, \alpha, W, V)$ instead of vector of variables from each study, $(\hat{\beta}_k, \hat{\beta}_{pooled,k}, \alpha_k, W_k, V_k)$. The collapsed variables are defined in homo MetaQ-test. Like Qtest3, the empirical distribution of suggested test statistic is calculated from pre-calculated distribution in the step 5.

Algorithm 2: Homo-meta Q3

Step 1. Compute $\hat{\beta}^* = \hat{\beta} - E(\hat{\beta}^* | \hat{\beta}_{pooled}) = \hat{\beta} - \hat{\beta}_{pooled}(\alpha^T W V W \alpha)^{-1} W V W \alpha$

Step 2. Compute $Q_{2|1}^* = \hat{\beta}^{*T} V^{*-1} \hat{\beta}^* = \hat{\beta}^{*T} U^* \Lambda^{*-1} U^{*T} \hat{\beta}^* \sim \chi_{m-1}^2$,

Step 3. Compute $Q_{2|1} = 2G_{0.5,1}^{-1}(1 - p_{2|1})$, where $p_{2|1}$ is obtained from $Q_{2|1}^*$

Step 4. Compute $Q_{hom-meta-q3}^\pi = (1 - \pi)Q_{hom-meta-q1} + \pi Q_{2|1}$

Step 5. Obtain p value $p_{hom-meta-q3}^\pi$ from mixture of χ^2

Step 6. Compute final p value for $Q_{hom-meta-q3}$ from the empirical distribution,

$$\hat{\pi}^{optimal} = \text{Argmin}_\pi \{p_{hom-meta-q3}^\pi; \pi = 0, 0.1, 0.2, \dots, 0.9, 1\}$$

Meta-analysis Assuming Heterogeneous Genetic Effects across Studies: Het-meta-Q3.

When studies are heterogeneous, we suggest $Q_{het-meta-q3}$ that is weighted average of $Q_{het-meta-q1}$ and $Q_{het-meta-q2}$. The following are steps for $Q_{het-meta-q3}$. The steps are similar to Qtest3 and MetaQTest3 but, in step 2 we use the sum of Wald type statistics, $Q_{2|1}^*$ from each study. For the sake of simplicity, we assume that the number of variants in a gene is the same across the study, thus the degree of freedom is just multiplication K by m .

Algorithm 3: Hetero-meta Q3

Step 1. Compute $\hat{\beta}_k^* = \hat{\beta}_k - E(\hat{\beta}_k^* | \hat{\beta}_{pooled,k}) = \hat{\beta}_k - \hat{\beta}_{pooled,k}(\alpha_k^T W_k V_k W_k \alpha_k)^{-1} W_k V_k W_k \alpha_k$

Step 2. Compute $\sum_{k=1}^K Q_{2|1}^* = \sum_{k=1}^K \hat{\beta}_k^{*T} V_k^{*-1} \hat{\beta}_k^* = \sum_{k=1}^K \hat{\beta}_k^{*T} U_k^* \Lambda_k^{*-1} U_k^{*T} \hat{\beta}_k^* \sim \chi_{K m-1}^2$

Step 3. Compute $Q_{2|1} = 2G_{0.5, K, 1}^{-1}(1 - p_{2|1})$

Step 4. Compute $Q_{het-meta-q3}^\pi = (1 - \pi)Q_{het-meta-q1} + \pi Q_{2|1}$

Step 5. Obtain p value $p_{het-meta-q3}^\pi$ from mixture of χ_K^2

Step 6. Final p value for Q_3 is calculated from the empirical distribution,

$$\hat{\pi}^{optimal} = \text{Argmin}_\pi \{p_{het-meta-q3}^\pi; \pi = 0, 0.1, 0.2, \dots, 0.9, 1\}$$

Numerical simulations

To evaluate the properties of the proposed methods, we performed the simulation study. We generated the genotype data using COSI program that implements coalescent model, and we obtained both length 200 kb and 10,000 European-like haplotypes and African-American-like haplotypes [21]. By randomly mating with haplotypes, we could obtain genotype data for analysis. The 3 kb regions are also randomly selected for each gene and rare variants under the threshold are analyzed [22]. We filtered out

common variants with the threshold of MAF larger than 0.03 and singleton SNPs. Left figures consider MAF < 0.05 and right figures consider MAF < 0.01. In European population, 82.93% variants have MAF < 0.01 and 67.51% variants have MAF < 0.001, this frequency is different from that of generated simulation data set in MetaSKAT (86 and 76% respectively). Since COSI program gives randomness in generating haplotypes, this difference could happen even using the same input parameters. Table 1 shows the simulation study settings which we borrowed the basic idea of them in MetaSKAT. For precise comparison of our methods with MetaSKAT, we brought their setting here and tried to prove better performance of MetaQ-test even in the setting for MetaSKAT. Giving variety to kind of population such as sample size and number of covariates, we considered six different scenarios like MetaSKAT. Using COSI program, MetaSKAT generated genotype data set for different populations and created 3 single studies. The number of different scenarios of simulation that they considered is 6, and half of them have the same study sample size and rest of them have different sample size. To give heterogeneity among the studies, in addition to different size, they also allocated different population groups and different number of covariates to each study. We compared the proposed methods with other existing meta-analysis methods. Meta-analysis methods that we considered are as follows: MetaQ-test (Hom--meta-Q1, Het-meta-Q1, Hom-meta-Q2, Het-meta-Q2, Hom-meta-Q3, Het-meta-Q3), MetaSKAT (Hom--meta-SKAT, Het-meta-SKAT, Hom-meta-SKAT-O, Het-meta-SKAT-O), Meta-Burden, Fisher's combined probability test and Stouffer's Z-score method [23, 24].

Type I error and power simulations

For type I error simulations, we generated 100 phenotypes for 100 each gene under the null hypothesis; there is no association between a gene and a phenotype.

$$y_{ki} = 0.5X_{k1i} + \dots + 0.5X_{kq_k i} + \varepsilon_{ki}, \quad \varepsilon_{ki} \sim N(0, 1), \quad (2)$$

As the type I error simulation of MetaSKAT, X_{k1i} is the covariate taking 0 or 1 value with equal probability 0.5. The rest of covariates, from X_{k2i} to $X_{kq_k i}$ are taken from a standard normal distribution. The information of covariates for each study is in the Table 1. The index, q_k indicates the number of covariates for k th study. Since we generated 100 genes and 100 phenotypes, we carried out 10,000 times association tests, and the level of significance, we gave $\alpha = 0.01$ and 0.001.

Table 1 Simulation study settings

Scenario	Population	Sample Size			Covariates		
		Study 1	Study 2	Study 3	Study 1	Study 2	Study 3
1	EUR	1600	2200	3200	(x ₁ , x ₂)	(x ₁ , x ₂)	(x ₁ , x ₂)
2	EUR	1600	2200	3200	(x ₁)	(x ₁ , x ₂)	(x ₁ , x ₂ , x ₃)
3	EUR + AA	1600	2200	3200	(x ₁)	(x ₁ , x ₂)	(x ₁ , x ₂ , x ₃)
4	EUR	2400	2400	2400	(x ₁ , x ₂)	(x ₁ , x ₂)	(x ₁ , x ₂)
5	EUR	2400	2400	2400	(x ₁)	(x ₁ , x ₂)	(x ₁ , x ₂ , x ₃)
6	EUR + AA	2400	2400	2400	(x ₁)	(x ₁ , x ₂)	(x ₁ , x ₂ , x ₃)

EUR + AA denotes that population of first and second study is European and the population of Study3 is African-American

Unlike type I error simulations, we generated 1000 genotype datasets and made phenotypes using the below model for power simulations,

$$y_{ki} = 0.5X_{k1i} + \dots + 0.5X_{kq_i} + \mathbf{G}'_{ki,causal}\boldsymbol{\beta}_{k,causal} + \varepsilon_{ki}, \quad \varepsilon_{ki} \sim N(0, 1) \quad (3)$$

$\mathbf{G}_{ki,causal}$ is a vector of causal variants in a gene, and $\boldsymbol{\beta}_{k,causal}$ is their effect size. For the proportion of causal variants, we set four cases, 10, 20, 30 and 50% of variants are causal like in MetaSKAT. To illustrate the effect of burden or non-burden type test, we also assumed that all variants are positive or 80% are positive and rest of 20% are negative. Since the number of causal variants or positive variants could not be integer in some cases because of small number of variants in a gene, we generated number of causal variants in Bernoulli distribution. The regression coefficient of genetic effect is given same as the MetaSKAT, $\boldsymbol{\beta}_{k,causal} = c|\log_{10}(MAF)|$. However, we used study specific MAF rather than population MAF, because in reality we hardly get the population MAF even in Meta-analysis. Defining the coefficient in this way reflects the assumption of rare variant study; the rarer SNPs have the larger effects on the phenotype. We set $c = 0.475$ in 5% of causal variants, 0.375 for 10%, 0.25 for 20%, and 0.175 for 50% of causal variants respectively. Since the effect size of regression coefficient depends on the MAF, the case for different MAF across the studies assumes heterogeneous effect and the case for the same MAF assumes homogeneous effect. These assumptions rely on the belief that different populations could have different MAF, and this phenomenon is called population stratification. In the simulation work, there are some differences with MetaSKAT. First, we used haplotypes that have different distribution of allele frequencies with that of MetaSKAT and second, we used study specific MAF rather than population MAF that was used in simulation of MetaSKAT. Finally, we used Bernoulli distribution to assign causal variants rather than exact given proportion percentages. We expected that the differences might cause the slightly different

power results of MetaSKAT calculated here with MetaSKAT paper.

Materials

Meta-analysis for gene-level rare variants association studies

We analyzed whole exome sequencing data from Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) consortium. Sequencing was performed at the Broad Institute with Agilent v2 capture reagent on HiSeq platform. The consortia have exome sequencing data of 13,000 individuals from 5 different ancestry groups: African Americans (AJ, AW), American Hispanics (HA, HS), East Asian (EK, ES), South Asians (SL, SS), and European (UA, UF, UG, US, and UB). The words in parenthesis stand for the populations in the ancestry group. For meta-analysis, we selected East Asian ancestry group, EK from Korean samples and ES from Singaporean samples. Total number of samples in EK is 1086 and in ES is 1078. The primary goal of two consortia is to identify novel type 2 diabetes (T2D) related genetic factors, but data also includes several diabetes-related quantitative traits such as lipid and blood pressure traits. We applied the proposed meta methods to lipid and blood pressure traits: total cholesterol (CHOL), HDL cholesterol (HDL), LDL cholesterol (LDL), triglycerides (TG), systolic blood pressure (SBP) and diastolic blood pressure (DBP). Subjects having medication that might affect the traits were excluded for the analysis. Sample sizes considered missing count and medication are described in Table 2. For covariates, we used age, sex information (BMI information was used only for blood pressure phenotypes).

Results

Type I error

Empirical type I error rates under the setting of Scenario 1 and 2 are calculated. Since first scenario consider homogeneous case and second scenario considers heterogeneous case, we only calculate type I error in these two scenarios. The rates are calculated under each

Table 2 Sample Size of Asian Population Groups for Seven Quantitative Traits

Traits	CHOL	HDL	LDL	TG	SBP	DBP
EK (1086)	1078	1078	1031	1078	1086	1086
ES (1078)	627	628	621	627	1077	1077

Missing samples and samples who receive medication are excluded

significant level, 0.01 and 0.001. Empirical type I error rates under the setting of Scenario 1 are given in Table 3. For both MetaQTest and MetaSKAT, type I error rates were well controlled, but in Het-meta-SKAT-O, type I error rate was somewhat inflated. Empirical type I error rates under the setting of Scenario 2 are given in Table 4. For both MetaQ and MetaSKAT, type I error rates were also well controlled, but in Hom-meta-Q2, type I error rate was somewhat inflated.

Power

To compare the power, we performed the simulation work. First, we compared power of our meta method with those of joint analysis to check efficiency of meta method. When meta-analysis yields highly comparable results with joint analysis, the meta method serves well as an alternative of joint analysis. Since joint-analysis cannot handle heterogeneity between studies such as different number of covariates directly, we only performed joint-analysis in homogeneous scenario settings (scenario 1 and 4). In Fig. 1, the blue 8 dots indicate each case for scenario 1 (causal 5 to 50% and all same direction of variants to different direction) and the pink dots are for scenario 4. Joint-analysis is consistently powerful, but Hom-meta-Q also achieves highly comparable power compared with joint-analysis. Especially, for scenario 1, our meta methods recovers almost power of joint-analysis.

Figure 2 shows power comparisons of eight different meta methods when all causal variants are risk increasing. In the figure, first 4 bars indicate our proposed methods, Hom-meta-Q2, Hom-meta-Q3, Het-meta-Q2 and Het-meta-Q3 in order. The next 4 bars indicate competing MetaSKAT methods, Hom-meta-SKAT, Hom-meta-SKAT-O, Het-meta-SKAT and Het-meta-SKAT-O. In the first scenarios, since we assume homogeneous genetic effects, Hom-meta-Q are more powerful than Het-meta-Q. Up to 20% causal, Hom-meta-Q2 is the most powerful but in 50% causal case, as the power

of burden test increases, Hom-meta-Q3, optimal type, becomes the most powerful. In these scenarios, our methods outperform in 5%~20% percentages of causal variants. However, when a causal percentage reaches 50%, then MetaSKAT (Hom-meta-SKAT-O) overtakes the power (Hom-meta-Q3: 74.2% and Hom-meta-SKAT-O: 76.4%). The reversal occurs, because Hom-meta-Q1 works poorly even Hom-meta-Q2 is more powerful than Hom-meta-SKAT when it is expected to outperform in a case of high causal percentage. But, the power differences are not so big, about 2% difference. In scenario 2, we gave heterogeneity between studies by giving different number of covariates. The powers of Het-meta-Q are higher than those of Hom-meta-Q and their power differences are clearly different, about twice higher in 5% of causal variants. Het-meta-Q outperform except in the case of 50% of casual variants and this trend also appears in the rest of scenarios: as the number of causal variants increases, MetaSKAT is more powerful than our method. The reason why this trend is maintained is that our burden type test has lower power than Meta-Burden and even our non-burden type test when burden type test should have the highest power is more powerful. In scenario 3, population of studies is different and the number of covariate is also different. In this more heterogeneous case, Het-meta-Q has higher power than Hom-meta-Q, and the value of power itself is higher than in scenario 2. Our method outperform up to 20% of causal variants. As the percentage of casual variants increases the power of our method becomes lower than MetaSKAT. One of the distinguishable features of our method is that the power of our methods varies sensitively depending on the range of heterogeneity. The gap of power differences of homo and heterogeneous cases are larger than those of MetaSKAT. All methods in MetaSKAT have the robust power but in 50% of casual variants. Thus, when there are information about heterogeneity prior to analysis and the percentage of causal variants is expected to be lower, then our method is more powerful than MetaSKAT. For the rest of scenarios, we gained similar power results with previous scenarios; only in the case of 50% causal variants, MetaSKAT has higher power than MetaQ, but in scenario 4, our methods outperformed regardless of causal variants percentages and in scenario 6, in the case of 20% causal variants with 50% causal, MetaSKAT outperforms.

Table 3 Type I Error Rates Estimates at in Scenario 1

α	Hom-meta-Q1	Het-meta-Q1	Hom-meta-Q2	Het-meta-Q2	Hom-meta-Q3	Het-meta-Q3	Hom-meta-SKAT	Het-meta-SKAT	Hom-meta-SKAT-O	Het-meta-SKAT-O
10^{-2}	9.30E-03	9.00E-03	1.13E-02	1.26E-02	1.13E-02	1.21E-02	1.51E-02	1.05E-02	1.38E-02	1.23E-02
10^{-3}	1.30E-03	1.10E-03	1.10E-03	1.30E-03	1.30E-03	9.00E-04	1.60E-03	1.00E-03	2.10E-03	1.80E-03

Table 4 Type I Error Rates Estimates at in Scenario 2

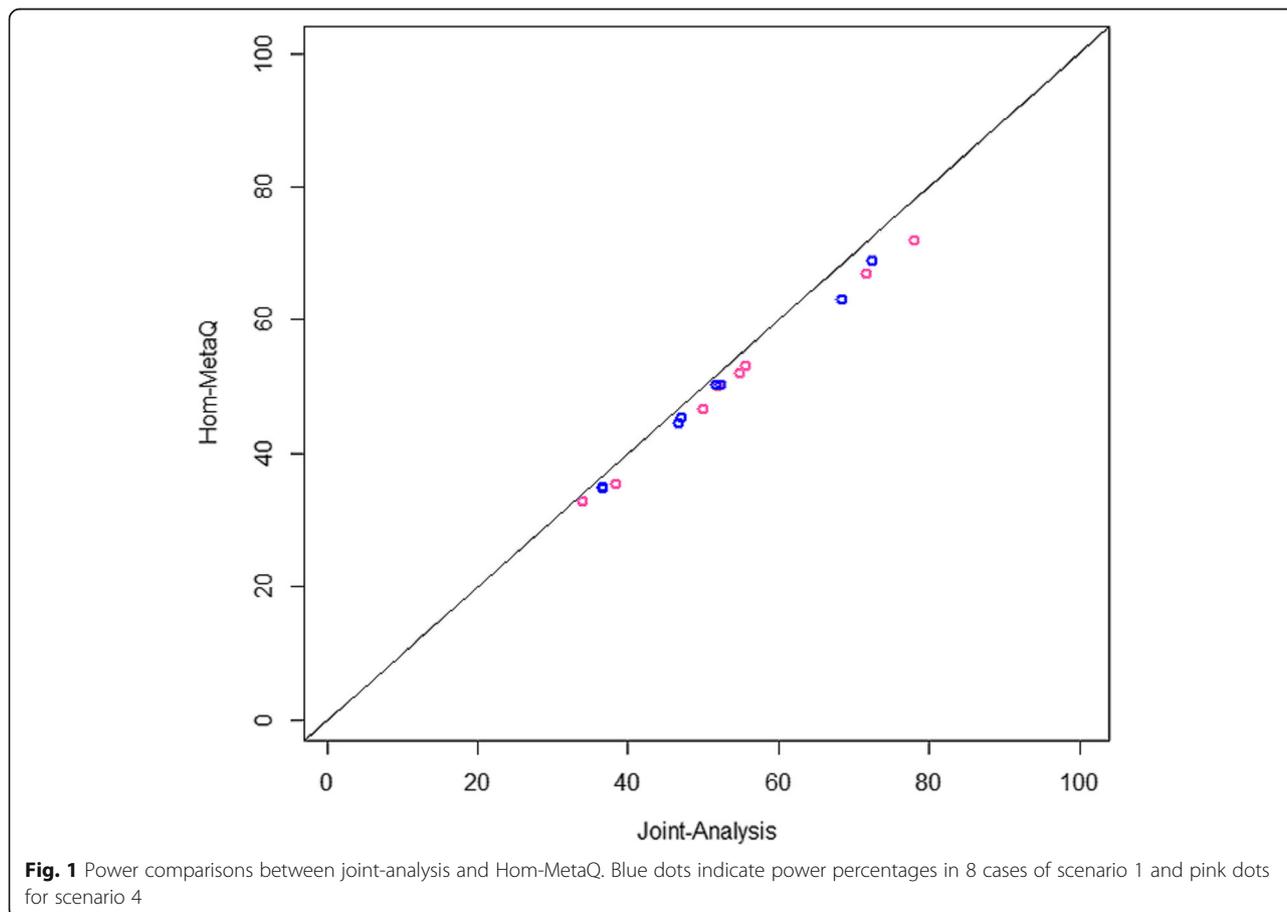
α	Hom-meta-Q1	Het-meta-Q1	Hom-meta-Q2	Het-meta-Q2	Hom-meta-Q3	Het-meta-Q3	Hom-meta-SKAT	Het-meta-SKAT	Hom-meta-SKAT-O	Het-meta-SKAT-O
10^{-2}	1.08E-02	9.90E-03	1.05E-02	9.80E-03	9.80E-03	1.00E-02	9.80E-03	1.03E-02	9.80E-03	1.04E-02
10^{-3}	1.40E-03	9.00E-04	1.40E-03	1.10E-03	9.00E-04	1.10E-03	1.00E-03	9.00E-04	1.10E-03	1.20E-03

Figure 3 shows power comparisons of meta methods when 20% of causal variants are risk decreasing and 80% are risk increasing. In scenario 1, Hom-meta-Q have higher power than MetaSKAT regardless of percentage of causal variants. However, the power differences become smaller as casual percentage increases. Compared to the case of all risk causal variants, in 50% of causal variants our method still have higher power because power loss of burden type is not large. In scenario 2 and 3, Het-meta-Q have higher power than MetaSKAT. Our methods show significant difference between Hom-meta-Q and Het-meta-Q, but MetaSKAT has little difference between their methods. In scenario 5, up to 20% causal, our best method is powerful than MetaSKAT, but when causal variants are 50%, then MetaSKAT outperforms. Moreover, in scenario 6, in only 20% causal, MetaSKAT outperforms, but the power differences

between our proposed methods and MetaSKAT are smaller than those observed in the previous, because in this setting non-burden type test is more powerful than burden, so our burden type test that is less powerful is less involved here. The reason why our burden type test, Hom-meta-Q, have little power in heterogeneous case is that the assumption used in burden type test is not kept well in heterogeneous case. In heterogeneous case, we can hardly say that all of variants in a gene across the different population groups are detected in any population and have same direction. Thus, we do not include the result of Hom-meta-Q in simulation and real data analysis.

Real data analysis results

We applied suggested methods to the real lipid and blood pressure data. All considered meta methods did



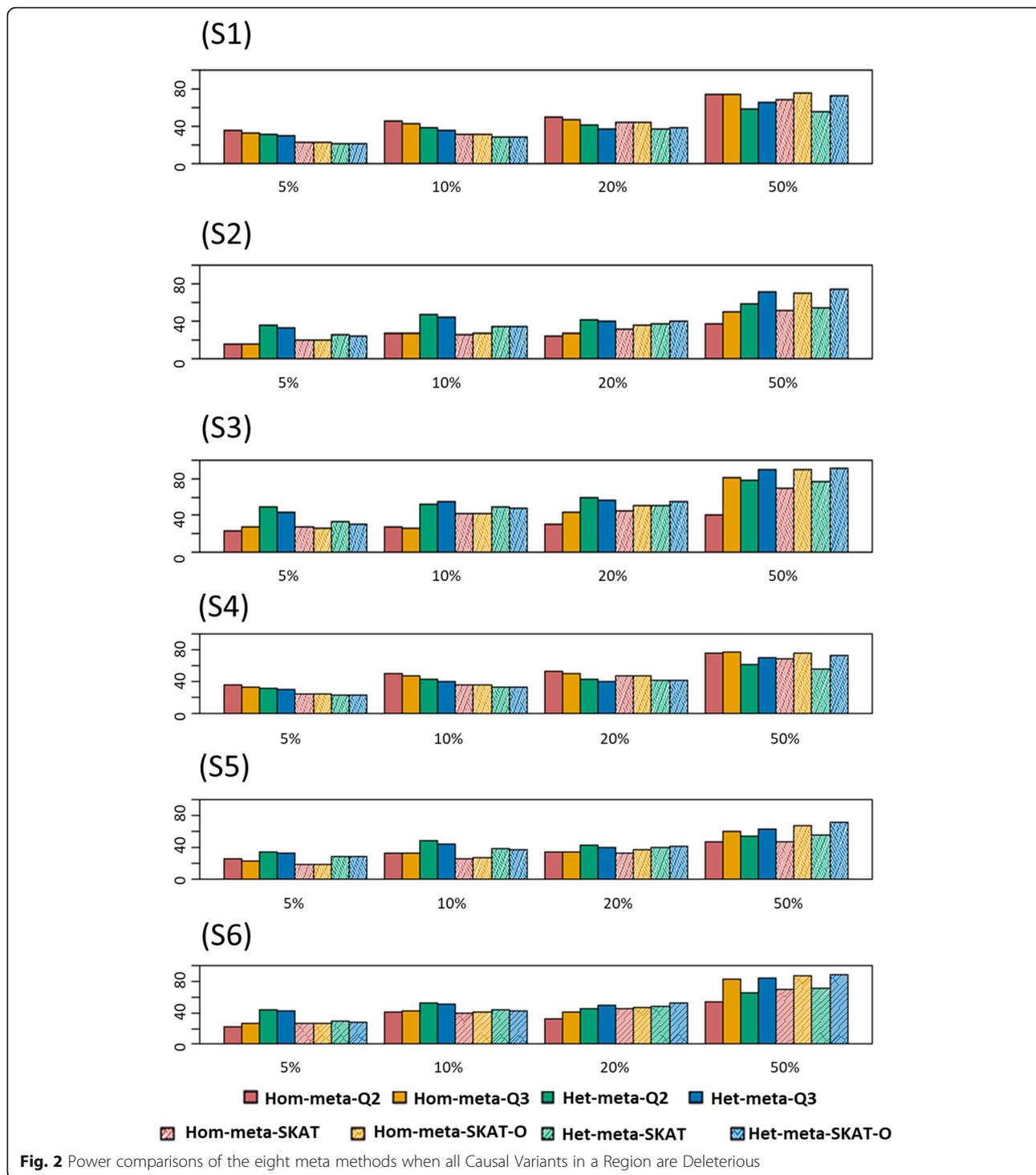


Fig. 2 Power comparisons of the eight meta methods when all Causal Variants in a Region are Deleterious

not detect any significant genes for lipid traits. But, for blood pressure traits we gained the different association results with lipid traits. Before analysis, we filtered common SNPs. Thus, remaining number of genes for the analysis of SBP is 9303 and SNPs in each gene have $MAF < 0.05$ and $MAC > 4$. Unlike MetaSKAT methods that failed to detect any significant genes as well in the

lipid traits analysis (at the bonferroni's significant level and arbitrary level, $1.00e-04$), MetaQ-test detected PCDHA9 gene in Het-meta-Q2 (p value = $4.75e-05$) at the bonferroni's significant level ($\alpha = 5.37e-06$), but this gene is also not known to be related with blood pressure. According to GO annotations, it is involved in calcium ion binding. However, the gene that has the second

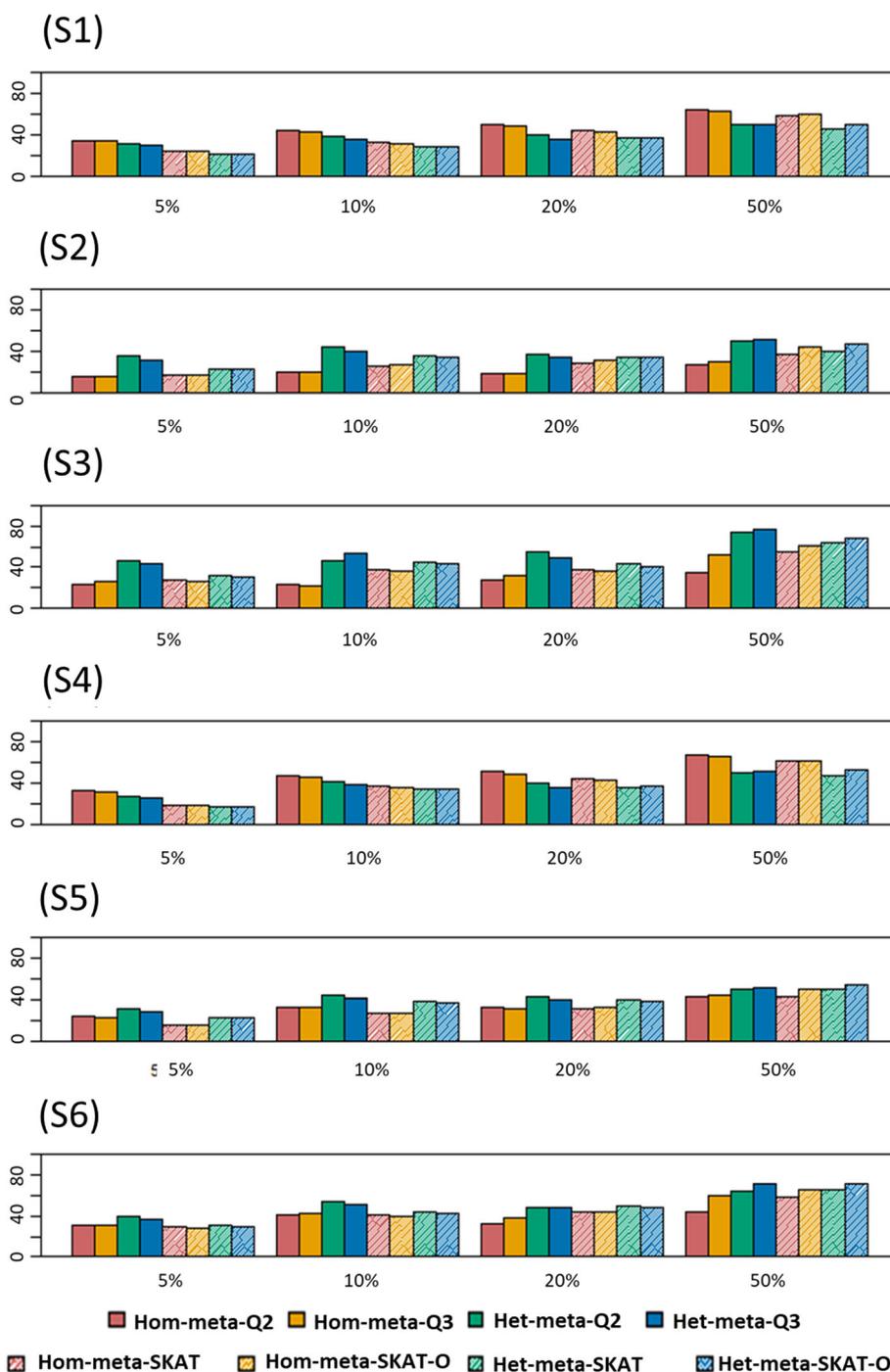


Fig. 3 Power comparisons of the eight meta methods when 20% causal variants in a region are protective and 80% are deleterious

smallest p value in both Het-meta-Q2 (p value = $1.19e-05$) and Het-meta-Q3 (p value = $2.41e-05$) is KCNA5 and this gene is already known for having strong relevance to hypertension [25, 26]. Table 5 shows that p values of genes that were detected at the threshold, $1.00e-04$ have the smallest value in MetaQ-test. Moreover, we compared our meta method results with that of single QTest analysis to

verify the improved power of meta methods. Table 6 shows that PCDHA9 gene is also detected in QTest2 in EK single study analysis and its p value (p value = $8.22e-07$) is smaller than in Het-meta-Q2. However, its p value in ES single study analysis has very small value, so p value from MetaQ-test is the compromised these two p values from EK and ES single QTest analysis. But,

Table 5 Meta-Analysis Results for Testing the Rare Variants Effects on Systolic Blood Pressure

Meta-Analysis		GENE	
Method	Test type	PCDHA9 (CHR 5)	KCNA5 (CHR 12)
meta Q tests	Het-meta-Q1	6.18E-01	1.03E-01
	Het-meta-Q2	4.75E-06	1.19E-05
	Het-meta-Q3	7.34E-06	2.41E-05
meta SKAT	Hom-meta-SKAT	3.93E-03	5.03E-04
	Het-meta-SKAT	3.72E-03	1.68E-04
	Hom-meta-SKAT-O	7.65E-03	1.13E-03
	Het-meta-SKAT-O	7.60E-03	4.67E-04
Other methods	Meta-burden	7.97E-01	2.59E-01
	Fisher's method (Q3)	8.18E-06	3.06E-05
	Score method (Q3)	2.23E-04	1.31E-05

Bonferroni corrected significant level for Meta-analysis is $\alpha = 6.82e-06$

p values of KCNA5 from each single QTest are both larger than p value from MetaQ-test. This fact can say that KCNA5 which is known for one of causal genes of hypertension arterial was only detected by MetaQ-test and MetaQ-test improved the statistical power of QTest.

For the analysis of DBP, after filtering common variants, the number of genes is 10,528 and SNPs in each gene have $MAF < 0.05$ and $MAC > 3$. In the meta-analysis of DBP, two genes, DTYMK and PCCA were detected at the Bonferroni's significant level ($\alpha = 4.75e-06$). Table 7 shows that Meta-Analysis results for testing the rare variants effects on DBP. DTYMK was detected in Het-meta-Q3 (p value = $1.17e-06$) and PCCA was detected in Hom-meta-SKAT-O (p value = $3.18e-06$). DTYMK, however, is only found in EK samples, thus MetaQ-test result is only stemmed from EK QTest result (Table 8 shows that p values of QTest are the same with that of MetaQ-test). When a gene exists in a single study, then our MetaQ-test result reflects the result of single study. DTYMK is not currently identified as related with blood pressure, but as related with thymidylate kinase activity. Another detected gene in MetaSKAT,

Table 6 Single QTest Analysis Results for Testing the Rare Variants Effects on Systolic Blood Pressure

Single QTest		GENE	
Population	Test type	PCDHA9 (CHR 5)	KCNA5 (CHR 12)
EK	EK-Q1	6.82E-01	6.66E-01
	EK-Q2	8.22E-07	1.83E-03
	EK-Q3	1.33E-06	1.53E-03
ES	ES-Q1	3.73E-01	3.68E-02
	ES-Q2	2.48E-01	4.69E-04
	ES-Q3	3.97E-01	1.43E-03

EK GWAS significant level is $\alpha = 6.30e-06$ and ES GWAS significant level is $\alpha = 6.82e-06$

Bonferroni corrected significant level for Meta-analysis is $\alpha = 6.82e-06$

PCCA is not known to be associated with blood pressure function, but it is related with propionic academia and pcca-related propionic academia.

PCCA gene was also detected using QTest1 of ES sample analysis (p value = $2.76e-06$, Table 8). But EK QTest results offset the results of MetaQ-test, so MetaQ-test could not detect this gene.

Although CABIN1 gene was not discovered at the Bonferroni's significant level, but it is discovered at the FDR adjusted p value. CABIN1 which was detected using Het-meta-Q2 and Het-meta-Q3 (p value = $8.67e-06$ and $6.91e-06$ respectively) is known to be associated hypertension arterial and purpura thrombotic thrombocytopenic that is closely related blood pressure. Table 8 shows that QTest results of EK and ES population groups. CABIN1 was not detected using single QTest.

Discussion and conclusion

We propose MetaQ-test for meta-analysis of gene-level rare variant association studies. The basis of MetaQ-test is preserving the prior phase of association studies that is Q-test. MetaQ-test retains the relationship among multiple variants in a region. Further, MetaQ-test considers whether the genetic effects on the phenotype are same across studies or not. Assuming same effects corresponds to fixed effect meta-analysis model or else random effect model. By considering direction of variant effects and their equivalence in effect size at the same time, MetaQ-test can cover broad range of realistic meta-analysis cases.

We investigated the performance of MetaQ-test through simulation and real data analysis. Simulation studies showed that type I error rates were controlled well and MetaQ-test, particularly Het-meta-Q achieved the higher than or as powerful as MetaSKAT in various scenarios. However, when causal variants are over than 50%, then our methods are not powerful as MetaSKAT.

Table 7 Meta-Analysis Results for Testing the Rare Variants Effects on Diastolic Blood Pressure

Meta-Analysis		GENE		
Method	Test type	DTYMK (CHR 2)	CABIN1 (CHR 22)	PCCA (CHR 13)
meta Q tests	Het-meta-Q1	8.87E-01	5.95E-01	1.59E-01
	Het-meta-Q2	1.21E-06	8.67E-06	1.92E-04
	Het-meta-Q3	1.17E-06	6.91E-06	7.92E-05
meta SKAT	Hom-meta-SKAT	1.06E-01	3.93E-03	8.63E-06
	Het-meta-SKAT	1.43E-02	4.97E-03	8.11E-06
	Hom-meta-SKAT-O	1.40E-01	7.74E-03	3.18E-06
	Het-meta-SKAT-O	2.42E-02	1.03E-02	3.76E-06
Other methods	Meta-burden	1.51E-01	5.09E-01	4.78E-01
	Fisher's method (Q3)	NA	1.65E-05	1.05E-04
	Score method (Q3)	NA	1.50E-05	1.33E-02

GWAS significant level is $\alpha = 4.75E-06$

The minimum p value in each gene among the meta methods is marked in bold
Fisher's method and Z-score method used p values of optimal version of QTest, QTest3
NA denotes that p value is not calculated because of very low valued statistics

Thus, when there are small percentage of causal variants, our method is more powerful in all scenario settings. Since Hom-meta test uses estimating regression coefficients, satisfaction of model assumption affects the power of MetaQ-test sensitively. If there are no many overlapped variants in a gene across single studies, the assumption of Hom-meta is broken and Hom-meta-Q hardly perform well. For this reason, in the result of real data analysis (EK and ES sample have the small proportion of shared SNPs in a gene), Hom-meta-Q have inflated p values and we thought that most of them are false positive. Thus, in the result of real data analysis, we excluded the Hom-meta-Q results. Therefore, we expect that the prior test of heterogeneity of genetic effects can help us to determine appropriate model. In real data analysis, we have shown that Het-meta-Q searched out some known genes associated blood pressure trait. However, it also discovered some novel genes that are not at known to be associated with blood pressure at least for now, thus to validate biological relationship between them, more research and experiment in biology field are

needed. For the future research, we can combine adaptive test that can determine degree of heterogeneity of genetic effects to improve the power further.

Abbreviations

AJ: African American Jackson Heart Study Candidate Gene Association Resource; AW: African American Wake Forest Study; BMI: Body mass index; CHOL: Cholesterol; DBP: Diastolic blood pressure; EK: East Asian Korea Association Research Project (KARE) and Korean National Institute of Health (KNIH); ES: East Asian Singapore Diabetes Cohort Study and Singapore Prospective Study Program; GWAS: Genome wide association test; HA: Hispanic San Antonio Mexican American Family Studies, Texas; HDL: High-density lipoprotein cholesterol; HS: Hispanic Starr County, Texas; LDL: Low-density lipoprotein cholesterol; QTest: Quadratic test; SBP: Systolic blood pressure; SKAT: Sequence kernel association test; SL: South Asian London Life Sciences Population (LOLIPOP); SS: South Asian Singapore Indian Eye Study; T2D-GENES: A consortium for Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples; TG: Triglycerides; UA: European Longevity Genes in Founder Populations (Ashkenazi); UB: European Malmo-Botnia Study; UF: European Metabolic Syndrome in Men Study (Finnish); UG: European Kooperative Gesundheitsforschung in der Region Augsburg (KORA); US: UK Type 2 Diabetes Genetics Consortium

Funding

Publication costs are funded by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) grant (HI16C2037). Also, this work was supported by the Bio & Medical Technology

Table 8 Single QTest Analysis Results for Testing the Rare Variants Effects on Diastolic Blood Pressure

Single QTest		GENE		
Population	Test type	DTYMK (CHR 2)	CABIN1 (CHR 22)	PCCA (CHR 13)
EK	EK-Q1	8.87E-01	7.84E-01	7.28E-01
	EK-Q2	1.21E-06	2.22E-02	7.84E-01
	EK-Q3	1.17E-06	2.33E-02	8.72E-01
ES	ES-Q1	NA	3.26E-01	2.76E-06
	ES-Q2	NA	3.00E-05	1.52E-05
	ES-Q3	NA	4.81E-05	9.45E-06

Bonferroni corrected significant level for Meta-analysis is $\alpha = 6.82e-06$

Development Program of the National Research Foundation of Korea (NRF) grant (2013M3A9C4078158) and by grants of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI16C2037, HI15C2165, HI16C2048).

Availability of data and materials

Not applicable.

About this supplement

This article has been published as part of *BMC Medical Genomics Volume 12 Supplement 5, 2019: Selected articles from the 8th Translational Bioinformatics Conference: Medical Genomics*. The full contents of the supplement are available online at <https://bmcmedgenomics.biomedcentral.com/articles/supplements/volume-12-supplement-5>.

Authors' contributions

JK and TP developed statistical method. JK performed statistical analysis. TP conceived and planned the experiments. JK, JL, YK and TP planned and carried out the simulations. JK, JL, BO, TP contributed to the interpretation of the results. JK and TP took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

We used the exome sequencing data of 1,086 samples from KARE and 1,078 from LOLIPOP. KARE study is a part of Korean Genome Epidemiology Study (KoGES), and the dataset was used under the partnership of T2D-GENES. The dataset of LOLIPOP study was also used under the partnership of T2D-GENES. All participants of KARE study provided written informed consent. The study using KARE and LOLIPOP samples was approved by two independent institutional review boards at Seoul National University.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹Department of Statistics, Seoul National University, Seoul, South Korea. ²Department of Biochemistry and Molecular Biology, School of Medicine, Kyung Hee University, Seoul, South Korea. ³Interdisciplinary program in Bioinformatics, Seoul National University, Seoul, South Korea.

Published: 11 July 2019

References

- Asimit J, Zeggini E. Rare variant association analysis methods for complex traits. *Annu Rev Genet.* 2010;44:293–308.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy ML, Ramos EM, Cardon LR, Chakravarti A, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461:747–53.
- Hu YJ, Berndt SI, Gustafsson S, Ganna A. Genetic investigation of, a.T.C., Hirschhorn, J, north, K.E., Ingelsson, E, and Lin, DY : Meta-analysis of gene-level associations for rare variants based on single-variant statistics. *Am J Hum Genet.* 2013;93:236–48.
- Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. *Am J Hum Genet.* 2013;75:353–62.
- 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;83:311–21.
- 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;86:832–8.
- 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.
- 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.
- Lee S, Emond, Bamshad MJ, Barnes MJ, Rieder KC, Nickerson MJ, Team DA, 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;91:224–37.
- Ladouceur M, Dastani Z, Aulchenko YS, Greenwood CM, Richards JB. The empirical power of rare variant association methods: results from sanger sequencing in 1,998 individuals. *PLoS Genet.* 2012;8:e1002496.
- Lee J, Kim YJ, Lee JY, T2D-Genes Consortium, Lee S, Park T. Gene-set association test for next-generation sequencing data. *Bioinformatics.* 2016; 32:i611–19.
- Lin DY, Zeng D. Meta-analysis of genome-wide association studies: no efficiency gain in using individual participant data. *Genet Epidemiol.* 2010; 34:60–6.
- Evangelou E, Ioannidis JPA. Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet.* 2013;14:379–89.
- Lee S, Teslovich TM, Boehnke M, Lin X. General framework for meta-analysis of rare variants in sequencing association studies. *Am J Hum Genet.* 2013; 93:42–53.
- Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genet.* 2009;5:e1000384.
- Biernacka JM, Jenkins GD, Wang L, Moyer AM, Fridley BL. Use of the gamma method for self-contained gene-set analysis of SNP data. *Eur J Hum Genet.* 2012;20:565–71.
- Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010;26:2190–1.
- Zeggini E, Ioannidis JP. Meta-analysis in genome-wide association studies. *Pharmacogenomics.* 2009;10:191–201.
- Nikolakopoulou A, Mavridis D, Salanti G. How to interpret meta-analysis models: fixed effect and random effects meta-analyses. *Evid Based Ment Health.* 2014;17:64.
- Mosteller F, Colditz GA. Understanding research synthesis (meta-analysis). *Annu Rev Public Health.* 1996;17:1–23.
- Schaffner SF, Foo C, Gabriel S, Reich D, Daly MJ, Altshuler D. Calibrating a coalescent simulation of human genome sequence variation. *Genome Res.* 2005;15:1576–83.
- Pruitt KD, Tatusova T, Brown GR, Maglott DR. NCBI reference sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* 2012;40:130–5.
- Fisher RA, Genetiker S, Genetician S, Britain G, Ge'ne'ticien S. Statistical methods for Researc workers. Edinburgh: Oliver and Boyd; 1970.
- Stouffer SA, Suchman EA, DeVinney LC, Star SA, Williams RM Jr. The American soldier: adjustment during army life. In: *Studies in Social Psychology in World War II*, vol. 1; 1949.
- Remillard CV, Tigno DD, Platoshyn O, Burg ED, Brevnova EE, Conger D, Nicholson A, Rana BK, Channick RN, Rubin LJ, et al. Function of Kv1.5 channels and genetic variations of KCNA5 in patients with idiopathic pulmonary arterial hypertension. *Am J Phys Cell Phys.* 2007;292:1837–53.
- Pousada G, Balloira A, Vilarino C, Cifrian JM, Valverde D. Novel mutations in BMPR2, ACVRL1 and KCNA5 genes and hemodynamic parameters in patients with pulmonary arterial hypertension. *PLoS One.* 2014;9:e100261.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

