Detecting Epistasis in Genome-wide Association Studies with the Marginal EPIstasis Test

Lorin Crawford[1,†], Sayan Mukherjee[1,2,3], and Xiang Zhou[4,5,†]

1 Department of Statistical Science, Duke University, Durham, NC, USA
2 Department of Computer Science, Duke University, Durham, NC, USA
3 Department of Mathematics, Duke University, Durham, NC, USA
4 Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA
5 Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA

† **E-mail: lac55@stat.duke.edu; xzhousph@umich.edu**

Abstract

Epistasis, commonly defined as the interaction between multiple genes, is an important genetic component underlying phenotypic variation. Many statistical methods have been developed to model and identify epistatic interactions between genetic variants. However, because of the large combinatorial search space of interactions, most epistasis mapping methods face enormous computational challenges and often suffer from low statistical power. Here, we present a novel, alternative strategy for mapping epistasis: instead of directly identifying individual pairwise or higher-order interactions, we focus on mapping variants that have non-zero *marginal epistatic effects* — the combined pairwise interaction effects between the given variant and all other variants. By testing marginal epistatic effects, we can identify candidate variants that are involved in epistasis without the need to identify the exact partners with which the variants interact, thus potentially alleviating much of the statistical and computational burden associated with standard epistatic mapping procedures. Our method is based on the variance component model, and relies on a recently developed variance component estimation method for efficient parameter inference and p-value computation. We refer to our method as the "Marginal EPIstasis Test", or MEPIT. With simulations, we show how MEPIT can be used to robustly estimate marginal epistatic effects, produce calibrated test statistics under the null, and facilitate the detection of pairwise epistatic interactions. We further illustrate the benefits of MEPIT on several real datasets, including seven common diseases from the Wellcome Trust Case Control Consortium, as well as body composition traits from a swine genome-wide association study.

Introduction

Genome-wide association studies (GWASs) have identified thousands of genetic loci associated with many complex traits and common diseases, providing insights into the genetic basis of phenotypic variation [1]. Most of these existing GWASs look at one variant at a time and focus on identifying marginal genetic associations that exhibit either additive or dominant effects. However, it has long been hypothesized that effects beyond additivity could contribute to a large proportion of phenotypic variation. In particular, epistasis — the interaction between genetic loci — is thought to play a key role in defining the genetic architecture underlying complex traits [2] and constituting the genetic basis of evolution [3, 4]. Indeed, despite early controversies [5], studies have detected pervasive epistasis in many model organisms [6–11]. GWASs in humans have also identified several candidates of epistatic interactions that contribute to quantitative traits and diseases [12–15]. Consequently, modeling epistasis has been shown to increase phenotype prediction accuracy and facilitate genomic selection in animal breeding programs [16, 17]. Recently, epistasis has also been proposed as one of the main factors explaining missing heritability

— the proportion of heritability not explained by the top associated variants from GWASs [1, 18]. In particular, studies have hypothesized that epistasis could confound heritability estimation in pedigree studies and cause inflation of heritability estimates, creating the so-called "phantom heritability" [19, 20].

Because of the importance of epistasis, many statistical methods have been developed to identify epistasis in GWAS [21, 22]. Different statistical methods differ in their ways of selecting a testing unit (i.e. variants or genes [23]), their searching strategy (e.g. exhaustive search [24–26] or probabilistic search [27] or prioritization based on a candidate set [28]), and the calculation of test statistics (e.g. various frequentist tests [29] or Bayesian approaches [30, 31]). However, almost all of these statistical methods focus on explicitly searching for pairwise or higher-order interactions for identifying epistatic effects. Because of the extremely large search space (e.g. $\binom{p}{2}$ pairwise combinations for p variants), these methods often suffer from heavy computational burden and low statistical power. Despite various efficient computational implementations [26,32,33] and recently developed efficient search algorithms [27], exploring a large combinatory search space remains a daunting task for large GWASs. Statistically, because of a lack of *a priori* knowledge of epistatic loci, exploring all combinations of genetic variants could result in low statistical power, but restricting to a subset of prioritized combinations based on prior knowledge or marginal effects could also miss important genetic interactions.

Here, we present an alternative strategy for mapping epistasis. Instead of directly identifying individual pairwise or higher-order interactions, we focus on identifying variants that have a non-zero interaction effect with any other variants. To do so, we develop a novel statistical method, which we refer to as the the "Marginal EPIstasis Test" (MEPIT), to test each variant in turn on its *marginal epistatic effect* — the combined pairwise interaction effects between a given variant and all other variants. By testing marginal epistatic effects, we can identify candidate markers that are involved in epistasis without the need to identify the exact partners with which the variants interact, thus potentially alleviating much of the statistical and computational burden associated with standard epistatic mapping methods. Our method is based on variance component models, and importantly, by taking advantage of a recently developed variance component estimation method [34] for efficient parameter inference and p-value computation, is scalable to moderately sized genome-wide association studies. We illustrate how MEPIT can serve as a useful alternative to standard methods in mapping epistasis with both simulations and real data applications.

Materials and Methods

MEPIT Model

We describe the Marginal EPIstasis Test in detail here. Our goal is to identify variants that interact with other variants, and to avoid explicitly searching for pairwise interactions. Therefore, unlike standard epistatic tests, MEPIT works by examining one variant at a time. For the k^{th} variant, we consider the following linear model,

$$\mathbf{y} = \mu + \mathbf{x}_k\beta_k + \sum_{l\neq k}\mathbf{x}_l\beta_l + \sum_{l\neq k}(\mathbf{x}_k \circ \mathbf{x}_l)\alpha_l + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \text{MVN}(\mathbf{0}, \tau^2\mathbf{I}), \tag{1}$$

where $\mathbf{y}$ is an n-vector of phenotypes for n individuals; μ is an intercept term; $\mathbf{x}_k$ is an n-dimensional genotype vector for the k^{th} variant that is the focus of the model; β_k is the corresponding additive effect size; $\mathbf{x}_l$ is an n-dimensional genotype vector for the l^{th} variant, and l represents any of the p variants other than the k^{th}; β_l is the corresponding additive effect size; $\mathbf{x}_k \circ \mathbf{x}_l$ denotes an element-wise multiplication between genotype vectors, thus representing the interaction term between the k^{th} and l^{th} variants; α_l is the corresponding interaction effect size; $\boldsymbol{\varepsilon}$ is an n-vector of residual errors; τ^2 is the residual error variance; $\mathbf{I}$ is the identity matrix; and MVN denotes a multivariate normal distribution. In addition, we

assume that the genotype vector for each variant has been centered and standardized to have mean 0 and standard deviation 1.

The model in Equation (1) is an underdetermined linear system ($p > n$). Therefore, we have to make additional modeling assumptions on the effect sizes β_l and α_l to make the model identifiable. To do so, we follow standard approaches [45–47] and assume that each individual effect size follows a normal distribution, or $\beta_l \sim \mathrm{N}(0, \omega^2/(p-1))$ and $\alpha_l \sim \mathrm{N}(0, \sigma^2/(p-1))$ for $l \neq k$. With the normal assumption on effect sizes, the model in Equation (1) is equivalent to the following variance component model,

$$\mathbf{y} = \mu + \mathbf{x}_k\beta_k + \mathbf{m}_k + \mathbf{g}_k + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathrm{MVN}(\mathbf{0}, \tau^2\mathbf{I}), \tag{2}$$

where $\mathbf{m}_k = \sum_{l\neq k} \mathbf{x}_l\alpha_l$ is the combined additive effects from all other variants, and effectively represents the epistatic effect of the k^{th} variant under the polygenic background of all other variants; $\mathbf{m}_k \sim \mathrm{MVN}(0, \omega^2\mathbf{K}_k)$ with $\mathbf{K}_k = \mathbf{X}_{-k}\mathbf{X}_{-k}^T/(p-1)$ being the genetic relatedness matrix computed using genotypes from all variants other than the k^{th}; $\mathbf{g}_k = \sum_{l\neq k}(\mathbf{x}_k \circ \mathbf{x}_l)\alpha_l$ is the summation of all pairwise interaction effects between the k^{th} variant and all other variants; $\mathbf{g}_k \sim \mathrm{MVN}(0, \sigma^2\mathbf{G}_k)$ with $\mathbf{G}_k = \mathbf{D}_k\mathbf{K}_k\mathbf{D}_k$ representing a relatedness matrix computed based on pairwise interaction terms between the k^{th} variant and all other variants. Here, we denote $\mathbf{D}_k = \mathrm{diag}(\mathbf{x}_k)$ to be an $n \times n$ diagonal matrix with the genotype vector $\mathbf{x}_k$ as its diagonal elements. It is important to note that both $\mathbf{K}_k$ and $\mathbf{G}_k$ change with every new marker k that is considered.

Note that, in this work, we limit ourselves to only consider second order epistatic relationships between SNPs. However, the generalization of MEPIT to detect higher order interactions is straightforward and only involves the manipulation of $\mathbf{G}_k$. In addition, for case control studies, we can treat the binary case control labels as quantitative traits following the Taylor series approximation argument (e.g. [47]). Finally, we model β_k as a fixed effect here, but modeling it as a random effect is straightforward.

Point Estimates

Our goal is to identify variants that have non-zero interaction effects with any other variant. To do so, we can examine each variant in turn ($k = 1, \ldots, p$) and test the null hypothesis in Equation (1) that variant k has no interaction effect with any other variant, $\mathrm{H}_0 : \alpha_l' = 0 \; \forall \; l \neq k$. This same null hypothesis is specified in the variance component model stated in Equation (2) as $\mathrm{H}_0 : \sigma^2 = 0$. The variance component σ^2 effectively captures the total epistatic interaction effects between the k^{th} variant and all other variants — we call this the marginal epistatic effect for the k^{th} variant.

Testing the marginal epistatic effect σ^2 requires jointly estimating the variance component parameters (σ^2, ω^2, τ^2) in Equation (2). The standard method for variance component estimation is the restricted maximum likelihood estimation (REML) method. However, REML is computationally slow: it requires an iterative optimization procedure where the time complexity of each iteration scales cubically with the number of individuals [48, 49]. The slow computation speed of REML is further exacerbated by the fact that the variance component model changes for every variant k (i.e. both $\mathbf{K}_k$ and $\mathbf{G}_k$ are variant specific) and thus variance component parameters are required to be estimated over and over again across genome-wide variants. Therefore, we cannot use REML for marginal epistatic mapping. Instead, we follow the recently developed MQS method [34] for efficient variance component estimation and testing. MQS is based on the method of moments and produces estimates that are mathematically identical to the Haseman-Elston (HE) cross-product regression [35–41]. However, MQS is not only computationally more efficient than HE regression, but also provides a simple, analytic estimation form that allows for exact p-value computation, thus alleviating the need for jackknife re-sampling procedures [42] that are both computationally expensive and rely on incorrect individual independence assumptions [43, 44].

To estimate the variance components with MQS, we first multiply a projection matrix $\mathbf{M}_k$ on both sides of the model in Equation (2) to remove the influence of μ and $\mathbf{x}_k$. Here, $\mathbf{M}_k = \mathbf{I} - \mathbf{b}_k(\mathbf{b}_k^T\mathbf{b}_k)^{-1}\mathbf{b}_k^T$, where $\mathbf{b}_k = [\mathbf{1}_n, \mathbf{x}_k]$ with $\mathbf{1}_n$ denoting an n-vector of ones. Thus, $\mathbf{M}_k$ is a variant specific projection matrix

onto both the null space of the intercept and the corresponding genotypic vector $\mathbf{x}_k$. By multiplying $\mathbf{M}_k$, we obtain the following simplified modeling specification

$$\mathbf{y}_k^* = \mathbf{g}_k^* + \mathbf{m}_k^* + \boldsymbol{\varepsilon}_k^*, \quad \mathbf{g}_k^* \sim \text{MVN}(\mathbf{0}, \sigma^2 \mathbf{G}_k^*), \quad \mathbf{m}_k^* \sim \text{MVN}(\mathbf{0}, \omega^2 \mathbf{K}_k^*), \quad \boldsymbol{\varepsilon}_k^* \sim \text{MVN}(\mathbf{0}, \tau^2 \mathbf{M}_k). \tag{3}$$

where $\mathbf{y}_k^* = \mathbf{M}_k\mathbf{y}$; $\mathbf{g}_k^* = \mathbf{M}_k\mathbf{g}_k$; $\mathbf{G}_k^* = \mathbf{M}_k\mathbf{G}_k\mathbf{M}_k$; $\mathbf{m}_k^* = \mathbf{M}_k\mathbf{m}$; $\mathbf{K}_k^* = \mathbf{M}_k\mathbf{K}\mathbf{M}_k$; and $\boldsymbol{\varepsilon}_k^* = \mathbf{M}_k\boldsymbol{\varepsilon}$, respectively. Note that Equation (3) also changes with every new marker k that is considered.

To simplify notation, we use $\boldsymbol{\delta} = (\sigma^2, \omega^2, \tau^2)$ to denote the variance components. We use the notation $\boldsymbol{\Sigma}_k = [\boldsymbol{\Sigma}_{k,1}, \boldsymbol{\Sigma}_{k,2}, \boldsymbol{\Sigma}_{k,3}] = [\mathbf{G}_k^*, \mathbf{K}_k^*, \mathbf{M}_k]$. We use indices $i, j, l \in \{1, 2, 3\}$ to represent the corresponding variance component or covariance matrix. Given estimates $\widehat{\boldsymbol{\Sigma}}_k$, we can obtain the MQS estimates for the variance components of each variant $(\hat{\boldsymbol{\delta}}_{k,1}, \hat{\boldsymbol{\delta}}_{k,2}, \hat{\boldsymbol{\delta}}_{k,3}) = (\hat{\sigma}_k^2, \hat{\omega}_k^2, \hat{\tau}_k^2)$ via the following simple analytic formula

$$\hat{\boldsymbol{\delta}}_{k,i} = \mathbf{y}_k^{*T} \mathbf{H}_{k,i} \mathbf{y}_k^*. \tag{4}$$

Here, we define $\mathbf{H}_{k,i} = (\mathbf{S}^{-1})_{ii}\widehat{\boldsymbol{\Sigma}}_{k,i} + (\mathbf{S}^{-1})_{ij}\widehat{\boldsymbol{\Sigma}}_{k,j} + (\mathbf{S}^{-1})_{il}\widehat{\boldsymbol{\Sigma}}_{k,l}$, where $\mathbf{S}$ is a 3×3 matrix in which $\mathbf{S}_{ij} = \text{trace}(\widehat{\boldsymbol{\Sigma}}_{k,i}\widehat{\boldsymbol{\Sigma}}_{k,j})$ for every $i, j, l = 1, 2, 3$.

Hypothesis Testing

MEPIT provides two options to compute p-values. The first option is approximate and is based on a normal test that only requires variance component estimate $\hat{\sigma}^2$ and its corresponding standard error. In particular, the variance of the MQS estimates in Equation (4) are given via a previously suggested and computationally efficient approximation [34]

$$\text{V}(\hat{\boldsymbol{\delta}}_{k,i}) \approx 2\mathbf{y}_k^{*T} \mathbf{H}_{k,i}^T \mathbf{V}_k \mathbf{H}_{k,i} \mathbf{y}_k^*, \tag{5}$$

where $\mathbf{V}_k = \hat{\sigma}_k^2 \mathbf{G}_k^* + \hat{\omega}_k^2 \mathbf{K}_k^* + \hat{\tau}_k^2 \mathbf{M}_k$. Given the estimate in Equation (4) and its standard error in Equation (5), we can rely on asymptotic normality and perform a normal test (or z-test) to compute p-values. More specifically, we use a two sided test since the MQS estimates can be either positive or negative. The normal test is computationally efficient, but when the sample size is small it is not appropriate.

We also provide a second, exact option to compute p-values which is valid in the case of small sample sizes. The second option relies on the fact that the MQS variance component estimate in Equation (4) follows a mixture of chi-square distributions under the null hypothesis. This is because $\mathbf{y}^*$ is assumed to follow a multivariate normal distribution under the modeling assumptions. In particular, $\hat{\sigma}^2 \sim \sum_{i=1}^n \lambda_i \chi_{1,i}^2$, where $\chi_{1,i}^2$ are chi-square random variables with one degree of freedom and $(\lambda_1, \ldots, \lambda_n)$ are the corresponding eigenvalues of the matrix

$$\left(\hat{\omega}_0^2 \mathbf{K}_k^* + \hat{\tau}_0^2 \mathbf{M}_k\right)^{1/2} \mathbf{H}_{k,1} \left(\hat{\omega}_0^2 \mathbf{K}_k^* + \hat{\tau}_0^2 \mathbf{M}_k\right)^{1/2},$$

with $(\hat{\omega}_0^2, \hat{\tau}_0^2)$ being the MQS estimates of (ω^2, τ^2) under the null hypothesis. We can use the Davies method [50] to compute exact p-values.

While the Davies method is expected to produce calibrated p-values, it is also computationally demanding (Table S1). Therefore, in practice, we advertise a hybrid p-value computation procedure that uses the normal test by default and applies the Davies method when the p-value from the normal test is below the threshold of 0.05. The hybrid procedure combines the advantages of the two different tests and produces calibrated p-values while remaining computationally efficient (Table S1). As we will also show in the results section, MQS estimation procedures allow for both accurate and efficient marginal epistatic mapping in moderately sized genome-wide association studies.

Real Data Sets

We used two data sets in the present study: the WTCCC data and a swine GWAS data. The WTCCC data set is from the Wellcome Trust Case Control Consortium (WTCCC) 1 study [51] (`http://www.wtccc.org.uk/`). The data set consists of about 14,000 cases of seven common diseases, including 1,868 cases of bipolar disorder (BD), 1,926 cases of coronary artery disease (CAD), 1,748 cases of Crohn's disease (CD), 1,952 cases of hypertension (HT), 1,860 cases rheumatoid arthritis (RA), 1,963 cases of type 1 diabetes (T1D) and 1,924 cases of type 2 diabetes (T2D), as well as 2,938 shared controls. We selected a total of 458,868 shared SNPs following a previous study [47]. In the analysis, we mapped SNPs to the closest neighboring gene(s) using the databases dbSNP, ImmunoBase, and UCSC Genome Browser, which can be found at the following:

- dbSNP: `http://www.ncbi.nlm.nih.gov/SNP/`
- ImmunoBase: `http://www.immunobase.org/`
- UCSC Genome Browser: `http://ucscbrowser.genap.ca/`

The swine genome-wide association data set consists of $n = 820$ female pigs. The sows were from a commercial operation which utilized breeding stock from Newsham Choice Genetics (West Des Moines, IA, USA) [52]. More specifically, these animals belonged to a Large White grandparent maternal line and a Large White $\times$ Landrace parent maternal line [53]. Using Illumina's PorcineSNP60 BeadChip, the pigs were genotyped at exactly 64,232 SNPs. Only polymorphic SNPs that were able to be mapped to a genomic location using *Sus scrofa* (SSC) Build 10.2, those with call rates greater than 90%, and those with minor allele frequency above 5% were used for analyzation. These quality control measures left a final dataset with $p = 55{,}393$ markers. For interpretation of the analysis, we first mapped each significant SSC marker to its corresponding reference SNP, and used those to find the closest neighboring gene(s) according to dbSNP.

Other Methods

Single-SNP trait additive association analyses and pairwise epistatic analyses were fit with a linear regression model by using the -lm argument in the GEMMA software [47, 49]. This software is publicly available at `http://www.xzlab.org/software.html`.

Software Availability

The software implementing MEPIT is freely available at `https://github.com/lorinanthony/MEPIT`. We use the CompQuadForm R package to compute p-values from the Davies method. The Davies method can sometimes yield a p-value equal exactly to 0 when the true p-value is extremely small [55]. In this case, we report p-values as $P \approx 0$. If this is of concern, one can compute the p-values for MEPIT using Kuonen's saddlepoint method [55, 56] or Satterthwaite's approximation equation [57].

Results

Simulations: Type I Error Control

To validate MEPIT and our hybrid testing procedure in terms of controlling type I error, we carried out a simulation study. Specifically, we utilize chromosome 22 of the control cases from the WTCCC 1 study [51] to generate continuous phenotypes. Exclusively considering this group of individuals and SNPs leaves us with an initial dataset consisting of $n = 2{,}938$ control samples and $p = 5{,}747$ markers.

In order to investigate the type I error control, we first subsample from the genotypes for n = 1,000, 1,750, and 2,500 subjects. Next, we randomly select 1,000 causal SNPs and simulate continuous phenotypes by using a linear model, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. We simulate the additive effect sizes of each causal SNP and random noise term both from a standard normal distribution, and then we scale the two terms further to ensure a heritability of 50%. Note that the idea of the null model holds because MEPIT does not consider additive effects and solely searches for significant marginal epistatic effects that are a summation of pairwise interactions. We then assess the calibration of MEPIT under both the normal test and the Davies method for each sample size n. Figure 1 shows the quantile-quantile (QQ) plots based on application of MEPIT to these null datasets under both hypothesis testing strategies. The normal test heavily relies on the assumption of asymptotic normality — therefore, it is expected to see improvement of performance as the sample size increases. However, as one also expects, the normal test is inaccurate in the extreme tails of the test even for larger sample sizes. On the other hand, utilizing the Davies method via a mixture of chi-squares allows MEPIT to robustly control for type I error across all sample sizes. Table 1 shows the empirical type I error rates estimated for MEPIT at significance levels $\alpha = 0.05$, 0.01, and 0.001, respectively. As expected based on the QQ plots, under the Davies method, MEPIT controls the type I error rate for reasonably sized datasets, and can be slightly liberal when the sample size is small. Presumably, the liberal behavior of p-values in small samples arises from the fact that frequentist tests do not account for uncertainty in the variance component estimates in the null model. Based on the null simulation results, we use a hybrid p-value computation procedure (detailed in Material and Methods) that recalibrates p-value for a SNP using the Davies method when the z-test p-value for the SNP is below the threshold of 0.05.

Simulations: Estimating and Identifying Marginal Epistatic Effects

In this section, we use simulation studies to illustrate the advantages of MEPIT in identifying marginal epistatic associations. In addition to correctly detecting marginal epistatic associations, we will show that MEPIT can also estimate the marginal epistatic effects reasonably well. Therefore, analogous to SNP heritability estimation settings [45, 47], these variance component estimates can serve as a measurement of the marginal interaction phenotypic variance explained (PVE) by each epistatic causal variant.

To test the power of MEPIT, we again consider simulation designs similar to those proposed by previous epistatic analysis studies [32]. First, we assume that the broad-sense heritability is known ($H^2 = 0.6$) [47,51,58,59]. Next, we use the 22$^{\text{nd}}$ chromosome of all control cases from the WTCCC 1 study $\mathbf{X}$ (i.e. $n \approx 3{,}000$ and $p \approx 6{,}000$) to simulate continuous phenotypes that mirror genetic architectures affected by a combination of additive and pairwise epistatic effects. Specifically, we randomly choose 1,000 causal SNPs to directly affect the phenotype and classify the causal variants into three groups: (1) a small set of interaction SNPs, (2) a larger set of interaction SNPs, and (3) a large set of additive SNPs. In the simulations carried out in this study, SNPs interact between sets, so that SNPs in the first group interact with SNPs in the second group, but do not interact with variants in their own group (the same applies to the second group). One may view the SNPs in the first set as the "hubs" in an interaction map. We are reminded that interaction (epistatic) effects are different from additive effects. All causal SNPs in both the first and second groups have additive effects and are involved in pairwise interactions, while causal SNPs in the third set only have additive effects.

The additive effect sizes of all causal SNPs again come from a standard normal distribution or $\boldsymbol{\beta} \sim \text{MVN}(\mathbf{0}, \mathbf{I})$. Next we create a separate matrix $\mathbf{W}$ which holds the pairwise interactions of all the causal SNPs between groups 1 and 2. These SNPs have effect sizes also drawn as $\boldsymbol{\alpha} \sim \text{MVN}(\mathbf{0}, \mathbf{I})$. We scale both the additive and pairwise genetic effects so that collectively they explain a fixed proportion of genetic variance. Namely, the additive effects make up ρ%, while the pairwise interactions make up the remaining $(1-\rho)$%. Once we obtain the final effect sizes for all causal SNPs, we draw errors to achieve the target H^2. The phenotypes are then created using the model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\alpha}$ are the simulated effect sizes of the pairwise epistatic effects.

We consider a few scenarios that depend on two parameters:

- $(1-\rho)$, which measures the portion of H^2 that is contributed by the interaction effects of the first and second groups of causal SNPs. Specifically, the phenotypic variance explained (PVE) by the additive genetic effects is said to be $V(\mathbf{X}\beta) = \rho H^2$, while the PVE of the pairwise epistatic genetic effects is given as $V(\mathbf{W}\boldsymbol{\alpha}) = (1-\rho)H^2$.
- $p_1/p_2/p_3$, which are the number of causal SNPs in each of the three groups, respectively.

Specifically, we set $\rho = \{0.5, 0.8\}$ and choose $p_1/p_2/p_3 = 10/10/980$ (scenario I), 10/20/970 (scenario II), 10/50/940 (scenario III), and 10/100/890 (scenario IV). Note that scenarios III and IV assume a larger number of interactions than scenario I and II do, and are thus likely to be closer to reality. For ρ, the particular case where $\rho = 0.5$ assumes that additive and epistatic effects equally contribute to the broad-sense heritability of the simulated phenotypes. The alternative case in which $\rho = 0.8$ is a case where the PVE of the simulated complex traits are dominated by additive effects. We analyze 100 different simulated datasets for each value of ρ, in each scenario. All of the results described in this section are based on the cases in which $\rho = 0.8$, as this case is a more realistic setting where epistatic effects only make up a small percentage of the broad-sense heritability. The results for $\rho = 0.5$ can be found in Supporting Information (see S1 Fig.).

Figure 2(a) shows the power results for MEPIT's ability to detect both group 1 and 2 causal variants, respectively, compared across each simulation scenario. Empirical power of MEPIT was estimated as the proportion of p-values < 0.05. We can see MEPIT's ability to detect both groups of causal markers depends on the pairwise interaction PVE explained by each variant. For example in Figure 2(a), each causal variant in group 1 is expected to explain $V(\mathbf{W}\boldsymbol{\alpha})/p_1 = 1.2\%$ of the true interaction PVE since in every scenario $p_1 = 10$. In these situations, the cumulative PVE of these markers is great and MEPIT's power is large for all four scenarios (approximately 40% power). Note that the power is similar to MEPIT's ability to detect the group 2 causal markers under Scenario I (i.e. $p_2 = 10$), where each epistatic variant is also expected to explain $V(\mathbf{W}\boldsymbol{\alpha})/p_2 = 1.2\%$ of the interaction PVE. Alternatively, MEPIT exhibits half of the power when detecting the group 2 SNPs in the case of Scenario II (i.e. $p_2 = 20$), as each SNP explains only $V(\mathbf{W}\boldsymbol{\alpha})/p_2 = 0.6\%$ of the PVE (approximately 20% power). In addition, MEPIT's power to identify group 1 variants is independent of the number of variants in group 2 (i.e. p_2), suggesting that MEPIT's power depends on the total interaction effects rather than individual pairwise effect or the number of interaction pairs. The results based on the genome-wide significance threshold are similar and can be found in Supporting Information (see S2 Fig.).

Next, we assess MEPIT's ability to accurately estimate the expected contribution of each group 1 and 2 causal SNP to the interaction PVE. Figure 2(b) show boxplots of these estimates. The true expected interaction PVE explained by each causal SNP is depicted as the grey dashed lines. These plots show that even though MEPIT's power is directly affected by the epistatic contribution to the phenotypic variation, its ability to rightly estimate the effects of causal interacting SNPs is robust and unbiased. It is important to note that we see MEPIT maintain its estimation ability even when the portion of PVE explained by a set of causal SNPs is very small (i.e. group 2 SNPs in scenario IV). The estimation results are consistent with the well-known robustness of variance component models in estimating PVE in other settings (e.g. estimation of SNP heritability) [45, 47].

Simulation Comparisons

Here, we compare the performance of MEPIT with a standard exhaustive search procedure that examines all pairwise interactions to explicitly identify the exact pairs of variants involved in epistatic interactions [25, 26, 60, 61]. Specifically, in the exhaustive search, we consider the linear model $\mathbf{y} = \mu + \mathbf{x}_i\beta_i + \mathbf{x}_j\beta_j + (\mathbf{x}_i \circ \mathbf{x}_j)\alpha_{ij} + \boldsymbol{\varepsilon}$ and test $H_0 : \alpha_{ij} = 0$ for every marker combination of i and j in turn. Keeping notation consistent, $\mathbf{x}_i \circ \mathbf{x}_j$ denotes element wise multiplication between genotypes i and j, and α_{ij} represents the

effect size of their interaction. Note that the exhaustive search procedure is computationally feasible here because we only have $p \approx 6,000$ markers in the simulations.

It is helpful here to point out that we do not view MEPIT as a replacement for the exhaustive search procedure, nor do we view MEPIT as a superior method. Rather, we view MEPIT as an attractive alternative strategy for mapping epistasis: MEPIT not only can perform a significance test to detect variants involved in epistasis, but also can be used to obtain a prioritized set of variants that are further used to identify pairwise interactions. Our simulation comparisons are thus targeted to illustrate how MEPIT can be used in these two tasks, and how its performance differs from the exhaustive search procedure in different scenarios.

Identifying variants involved in epistasis. We first compare MEPIT with the exhaustive search method in identifying variants that are involved in epistasis. For this task, MEPIT can directly perform a significance test and produce a p-value. Here, we note that the power of MEPIT and the exhaustive search method are determined by different factors: the power of the linear interaction method depends on each individual epistatic interaction effect size α_{ij}, while the power of MEPIT, as we have shown in the previous section, depends on the marginal epistatic effects — the summation of interaction effects. Therefore, we would expect MEPIT and the exhaustive search method to be advantageous in different situations (if the exhaustive search method is computationally feasible). In particular, we would expect the exhaustive search method to be more powerful in scenario I (and II) where each individual interaction effect is large, and MEPIT to be more powerful in scenario (III and) IV where each individual interaction effect is small but the marginal epistatic effect remains large. To validate our expectations, we again generate continuous outcomes using the same previously described $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ simulation scheme. Once again, all results described in this section are based on cases in which $\rho = 0.8$, and results for $\rho = 0.5$ can be found in Supporting Information (see S3 Fig.).

We evaluate MEPIT's and the exhaustive search's ability to accurately identify marginal epistatic effects for markers in each of the two causal groups. The criteria we use compares the false positive rate (FPR) with the rate at which true variants are identified for each model (TPR). Figure 3 depicts the ability of MEPIT and the exhaustive search to detect causal variants in groups 1 and 2. In particular, these plots depict the portion of causal markers discovered after prioritizing all of those considered in order of their significance. We assess the marginal epistatic detection in the exhaustive search by first running the previously described pairwise linear model, ordering the resulting p-values for each possible interaction, and drawing a power curve for identifying the SNPs that are members of simulated causal groups 1 and 2. For example, if the top p-values from the exhaustive search are interactions SNP1-SNP2, SNP2-SNP3, SNP4-SNP5, and only SNP2 is the true causal epistatic variant, then the top three pairs only marginally identify 1 true variant and 4 false variants.

As expected, while the power of MEPIT depends on the pairwise interaction PVE explained by each SNP, the power of the exhaustive search depends on the individual interaction effect size. For example, the power of the exhaustive search to detect group 1 causal epistatic SNPs is dependent on the number of group 2 causal SNPs, which determines the interaction effect size in simulations. Therefore, while the exhaustive search exhibits higher power in the sparse scenario where there are only a small number of interactions each with a large effect size (e.g. scenarios I and II), its power quickly decays in the more polygenic scenario where there is a large number of interactions each with a small effect size (e.g. scenarios III and IV). MEPIT is able to perform well in the more realistic polygenic scenarios (III and IV) by modeling the marginal epistatic effects of each variant, allowing the detection of epistatic variants not to be dependent on the individual pairwise interaction effect size.

We are reminded that another advantage to MEPIT is the reduced space it must search over. For p genetic markers, the exhaustive search must run $p(p-1)/2$ tests while MEPIT only has to perform p tests. As a result, even in our simulation study with only 6,000 variants, the exhaustive search approach is already slower than MEPIT. The computational advantage of MEPIT over the exhaustive search approach can be much larger in moderate size GWASs with millions of markers.

Identifying pairwise interactions. Although we have focused on identifying marginal epistatic effects so far, MEPIT can also be used to facilitate the identification of pairwise (or high-order) epistatic interactions. In addition to comparing MEPIT with the exhaustive search method, we also consider another common approach to identify these epistatic pairs — the two-step filtering association mapping procedures detailed in [21, 25, 28, 62]. These methods often apply a marginal (additive) single-SNP test [49, 63] first to identify associated genetic variants with non-zero additive effects, then focus on the identified variants to test all pairwise interactions between them. Depending on the correlation between the marginal additive effect size and the probability of being involved in epistasis across genome-wide SNPs, these filtering methods can be more powerful than the exhaustive search strategy in the previous section, and is certainly much more efficient computationally. Here, instead of using the additive test, we propose using MEPIT as the initial filter. We hypothesize that the initial list of associated SNPs from MEPIT will be more robust and more likely capture epistatic effects as MEPIT directly prioritizes SNPs based on marginal epistatic effects. By using marginal epistatic evidence in the initial filtering step, we expect MEPIT to outperform the previous common procedure of using a linear model for filtering.

In this set of simulations, we utilize the same subset of real genotypes used for the marginal epistatic simulations in the last section [51], and again generate phenotypes under the same four simulation scenarios where pairwise interactions are well defined. After randomly selecting the three sets of causal SNPs and creating their pairwise interactions, we run MEPIT and the single-SNP linear model using all variants. We also reuse the exhaustive search, again as a baseline comparison. For MEPIT and the single-SNP linear model, we rank each variant according to their marginal p-values. The top 100 SNPs identified by both models are then selected, and all pairwise interactions among them are tested using a linear model that controls for the two main effects. For the exhaustive search, we simply rank the top 100^2 interactions to assess pairwise power. The results based on the top 250 SNPs are similar and can be found in Supporting Information (see S5 and S6 Fig.).

Figure 5 compares the power of the filtering procedure using the two different methods as the initial step. Phenotypes used to create this figure were generated under each scenario with broad-sense heritability $H^2 = 0.6$ and $\rho = 0.8$. All results for $\rho = 0.5$ can be found in Supporting Information (see S4 Fig.). Compared with the single-SNP test, filtering SNPs using MEPIT provides more power in finding true pairwise epistatic interactions. In fact, even for the cases in which marginal additive effects contribute to majority of the broad-sense heritability (i.e. $\rho = 0.8$), using MEPIT as the initial filtration procedure (as opposed to the single-SNP linear model) provides more power to finding exact causal epistatic pairs. This improvement comes from the fact that MEPIT allows the ranking of variants to be based on their marginal epistatic effects, rather than their marginal additive effects. Therefore, the set of SNPs identified by MEPIT in the first step already contains variants that capture epistatic effects, thus resulting in higher power in the second step to identify epistatic interaction pairs. In addition, similar to the simulation comparison in the previous sub-section, MEPIT and the exhaustive search procedure are advantageous in different settings: the exhaustive search procedure is again more powerful in the sparse setting where each individual pairwise interaction is large (scenarios I and II) while MEPIT gains advantage in the polygenic setting with a large number of interactions each with small effects (scenarios III and IV). Therefore, MEPIT also represents an attractive alternative to identifying pairwise interactions.

Detecting Epistasis in WTCCC Data

In addition to the our simulation study, we further assess MEPIT's ability to detect epistasis by analyzing two real GWAS datasets. The first dataset we consider consists of 7 common diseases from the Wellcome Trust Case Control Consortium (WTCCC) 1 study [51]. Specifically, this data includes about 14,000 cases from seven common diseases and about 3,000 shared controls, typed at a total of about 450,000 SNPs. The seven common diseases are bipolar disorder (BD), coronary artery disease (CAD), Crohn's disease (CD), hypertension (HT), rheumatoid arthritis (RA), type 1 diabetes (T1D), and type 2 diabetes (T2D). Overall, for each disease we have a dataset that consists of $n \approx 5,000$ subjects.

For each disease, we provide a summary table in Supporting Information which lists the marginal epistatic p-values for all significant SNPs as computed by MEPIT (see S2 Table). For this analysis, strong significance of association for a particular SNP or locus was determined by using the same marginal genome-wide significance threshold used in the original study ($P < 5 \times 10^{-7}$) [51]. Similarly, we deem an association to be moderate if a SNP has a p-value that falls between $5 \times 10^{-7} < P < 1 \times 10^{-5}$. In Table 2, we list the number of strong and moderate marginal epistatic associations according to MEPIT across each of the seven diseases.

Figure 5 displays manhattan plots of our epistatic genome-wide scan for two of the seven diseases, RA and T1D, where the only strong and moderately significant marginal epistatic effects were detected (see S7 Fig. for genome-wide scans of the other diseases). We stress that the interpretation of these images is slightly different than what is used for traditional manhattan plots. Specifically, in these figures spikes across chromosomes suggest loci where members involved in epistatic interactions can be found. Overall, MEPIT most noticeably identified the major histocompatibility complex (MHC) region on chromosome 6 in RA and T1D as having variants involved in prominent pairwise interactions. It is well known that the MHC region holds significant clinical relevance in complex traits and diseases with respect to infection, inflammation, autoimmunity, and transplant medicine [65,66]. Since this region has also been consistently implicated by other epistatic analyses, we restrict our focus to these two phenotypic traits. In Table 3, we list all MEPIT-discovered loci with at least one variant whose marginal p-value is below the predefined moderate genome-wide significance threshold.

We took all of the markers identified as having at least moderately significant marginal epistatic associations in T1D (= 27) and analyzed all 351 pairwise interactions between them. The last row in Table 2 lists the number of epistatic interactions with pairwise p-values below the Bonferroni-corrected threshold $P < 8 \times 10^{-13}$, which would have been used if we examined all genome-wide SNP pairs. A summary table with all of the significant joint pairs from T1D can be found in the Supporting Information (see S3 Table). Similar to what was shown in the numerical experiments and simulation studies in previous sections, MEPIT exhibits great ability to detect the "hub" SNPs of interactions. For instance in T1D, the SNP rs9270986 (MEPIT $P \approx 0$) is part of many of the top significant epistatic pairs. All these interaction effects have been verified in previous epistatic analyses [25, 32, 67].

Detecting Epistasis in Body Composition Traits in Pigs

The second real GWAS dataset we analyze comes from a swine genome-wide association study which focused on the development of a high density porcine SNP array that has made GWAS feasible in pigs [53]. Using Illumina's PorcineSNP60 BeadChip, a pilot study in $n = 820$ commercial female pigs were genotyped for approximately 64,000 SNPs and phenotyped for two backfat thickness traits: 10^{th} rib backfat and the last rib backfat. Specifically, these two highly selected for traits are important indicators of lean meat content. Only polymorphic SNPs that were able to be mapped to a genomic location using *Sus scrofa* (SSC) Build 10.2, those with call rates greater than 90%, and those with minor allele frequency above 5% were used for analysis. These quality control measures left a final dataset with $p \approx 55{,}000$ markers.

For this analysis, significance of association for a particular SNP or locus was determined by using a genome-wide significance threshold. Here, we follow a previous GWA animal study [69] and use the Dunn-Šidák correction, an exact form of the Bonferroni-correction, to account for the number of tests performed. Specifically, for p variants, the corrected 5% genome-wide significance threshold for both traits is given as $P < 1 - (1 - 0.05)^{1/p} \approx 9 \times 10^{-7}$. Markers that pass this threshold are considered to have strong marginal epistatic associations with the backfat traits. Once again, we deem an association to be moderate if a SNP has a p-value that falls between 5×10^{-7} and 9×10^{-7}.

Analyses for the two backfat traits revealed the exact same epistatically associated locus on chromosome 1 of the SSC1 (SSC1) between 274.1 Mb-276.6 Mb marked by the same 5 markers (INRA0006067, INRA0006074, DRGA0002004, INRA0006186, and INRA0006192). This result comes as no surprise since

the genetic correlation between the 10^{th} and last rib backfat thicknesses has been estimated to be nearly 0.90 [53,70]. Table 4 lists the variants located on the MEPIT-discovered locus, all of which at least satisfy the moderately significant genome-wide threshold. In the Supporting Information there are manhattan plots of the marginal epistatic genome-wide scan for both traits (see S8 Fig.), as well as a table summary of the pairwise p-values between the moderately and strongly associated marginal epistatic variants for both traits (see S4 Table).

The novel epistatic region on SSC1 contains a couple of potentially interesting interacting genes specifically within the context of porcine fat, growth rates, and metabolism. These genes are *KANK1* and *KLF9*. In particular, studies have shown that *KANK1* is associated with fasting proinsulin or insulinogenic index, and may have an influence on abnormal insulin production [71]. This is contextually important because insulin indirectly regulates and acts upon blood sugar levels and the production of fat tissue. Similarly, *KLF9* has been shown to promote porcine adipocyte differentiation, a process that is responsible for establishing the mature fat-cell phenotype [72–74]. These validating findings from previous studies, particularly within the porcine literature, lead us to believe that the analysis conducted here may contain true positives.

Discussion

We have presented MEPIT for detecting variants that are involved in epistasis in genome-wide association studies. For each variant in turn, MEPIT estimates and tests its marginal epistatic effect — the combined epistatic effect between the examined variant and all other variants. By modeling and inferring the marginal epistatic effects, MEPIT can identify variants that exhibit non-zero epistatic interactions with any other variant without the need to identify the specific marker combinations that drive the epistatic association. Therefore, MEPIT represents an attractive alternative to standard methods [24–26, 32, 33] for mapping epistasis. With both simulations and real data applications, we have illustrated the benefits of MEPIT.

In the present study, we have focused on estimating and testing marginal epistatic effects in the presence of pairwise interactions with MEPIT. MEPIT can also be easily extended to detect variants that are involved in higher-order interactions. Specifically, in the presence of higher-order interactions, we can introduce extra random effects terms to represent the combined higher-order interaction effects between the examined variant and all other variants — meaning, one extra random effects term for each extra order of interactions. Under the normality assumption of the interaction effect sizes, the introduced random effects terms all follow multivariate normal distributions, with the covariance matrices determined as a function of the Hadamard product of the additive genetic relatedness matrix [16, 68, 75–77]. Therefore, we can use a multiple variance component model with additional variance components to map epistatic variants in the presence of higher-order interactions. From there, we can test the variance components jointly to identify variants that are involved in any order of epistatic interactions. We can test each variance component separately to identify variants that are involved in a particular order epistatic interaction. Or, better still, we can perform variable selection on the variance components to identify which higher order interaction a particular variant of interest is involved in. Extending MEPIT to mapping high-order interactions will likely provide further insights into the epistatic genetic architecture of various traits and diseases.

Here, we have focused on demonstrating MEPIT with a variance component model. The variance component model in MEPIT effectively assumes that the interaction effect between the examined variant and every other variant follows a normal distribution. The normality assumption and the resulting variance component model have been widely used in many areas of genetics. For example, variance component models are used in rare variant tests to combine the additive effects of multiple rare variants to improve association mapping power [46, 78–81]. Similarly, variance component models are used to jointly model all genome-wide SNPs at once for estimating SNP heritability [45, 47]. Studies have already shown

that variance component models produce unbiased estimates regardless of whether or not the underlying effect sizes follow a normal distribution, and are robust even when the model is severely misspecified [45, 47]. In addition, MEPIT can easily be extended to incorporate other effect size assumptions. Indeed, the main idea in MEPIT of mapping marginal epistatic effects is not restricted to the particular variance component model we examine here, nor is it restricted to the normality assumption of the interaction effect sizes. Therefore, we can incorporate sparsity-inducing priors for effect sizes if the number of interaction pairs is known to be small *a priori*. Alternatively, we can use the recently developed hybrid effect size prior that has been shown to work well under a variety of effect size distributions [47]. Different interaction effect size assumptions can be advantageous under different genetic architectures and incorporating them in different scenarios will likely improve the power of MEPIT further.

There are many other potential extensions of MEPIT. We have only focused on analyzing one phenotype at a time in this study. However, it has been extensively shown that modeling multiple phenotypes can often dramatically increase power [49, 82]. Therefore, it would be interesting to extend MEPIT to take advantage of phenotype correlations to identify pleiotropic epistasis effects. Modeling epistasis in the context of multiple phenotypes could be highly non-trivial, as we need to properly model the shared epistatic components between phenotypes, in addition to the shared additive effects between phenotypes. Modeling strategies based on the multivariate linear mixed model (mvLMM) [49] could be helpful here. In addition to its use in genetic association studies, MEPIT can also be applied to other association studies, such as eQTL mapping studies [83]. Often times, eQTL studies deal with SNP effect sizes (on gene expression levels) that are orders of magnitude larger than that (on organism-level traits) from GWASs. Thus, eQTL studies requires a much smaller sample size than GWASs and subsequently allow for the efficient application of many sophisticated models. Recent studies have started to reveal an initial set of epistatic interactions that underlie gene expression variation [84–86]. By adapting MEPIT to eQTL studies, we expect to better understand the genetic architecture that underlie gene expression variation.

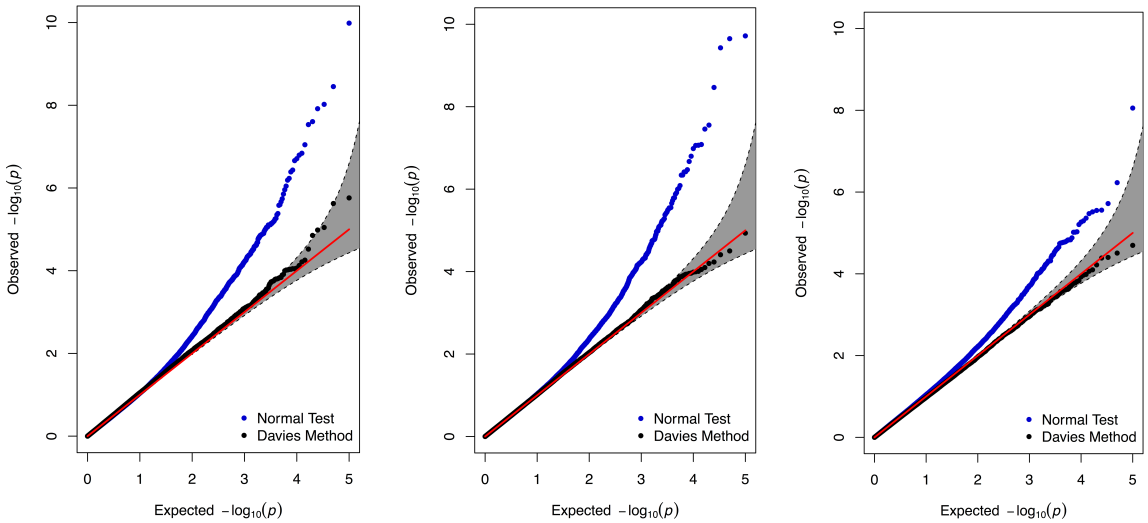
MEPIT is not without its limitations. Perhaps the most noticeable limitation is that MEPIT cannot be used to directly identify the interaction pairs that drive individual variant association. In particular, after identifying a variant involved in epistasis, it is still unclear which variants it interacts with. Thus, despite being able to identify SNPs that are involved in epistasis, MEPIT is unable to directly identify the detailed interaction pairs. However, we argue that being able to identify variants that are involved in epistasis is often an important first step towards identifying and understanding the detailed epistatic associations. In addition, being able to identify SNPs involved in epistasis allows us to come up with an initial likely set of variants that are worth further exploration. Indeed, we advertise a two-step *ad hoc* epistasis association mapping procedure. First, we identify individual SNP associations with MEPIT. Then, we focus on the most significant associations from the first step to further test all of the pairwise interactions among them to identify specific epistatic interactions. Unlike the previous filtering strategies that are commonly used in epistatic mapping, our two-step procedure is unique in the sense that the SNP set identified in our first step contains SNPs that already display strong epistatic effects with other variants. Therefore, our two-step procedure outperforms alternative filtering strategies in simulations. However, we caution that the two-step procedure is nevertheless *ad hoc* in nature and could miss important epistatic associations. Therefore, exploring statistical approaches that can unify the two steps would be an interesting area for future research. Besides this main limitation, we also note that MEPIT can be computationally expensive. MEPIT requires fitting a variance component model for every SNP in turn, and fitting variance component models are known to be computationally challenging [48, 49]. In this study, we use the recently developed MQS method for variance component estimation and testing. Compared with the standard REML method, MQS is computationally efficient, allows for exact p-value computation based on the Davies method, and is statistically more efficient than the REML estimates when the variance component is small [34] — a property that is particularly relevant here considering the marginal epistatic effect size is often small. MQS allows us to apply MEPIT to moderately sized genome-wide association studies with thousands of samples and millions of variants, which is otherwise

impossible using any other variance component estimation methods. However, new algorithms are likely needed to scale MEPIT up to data sets that orders of magnitude larger in size.

Acknowledgements

LC is supported by the National Science Foundation Graduate Research Program under Grant No. DGF-1106401. SM would like to acknowledge the support of grants SM would like to acknowledge the support of grants NSF IIS-1546331, NSF DMS-1418261, NSF IIS-1320357, NSF DMS-1045153, and NSF DMS-1613261. XZ would like to acknowledge the support of NIH Grants R01HL117626 (PI Abecasis), R21ES024834 (PI Pierce), and a grant from the Foundation for the National Institutes of Health through the Accelerating Medicines Partnership (BOEH15AMP, co-PIs Boehnke and Abecasis). We thank Kris C. Wood for helpful comments on a previous version of the manuscript. We thank Professor Xiaolei Liu at the HuaZhong Agriculture University for providing us with the pig data. This study also makes use of data generated by the Wellcome Trust Case Control Consortium (WTCCC). A full list of the investigators who contributed to the generation of the data is available from `www.wtccc.org.uk`. Funding for the WTCCC project was provided by the Wellcome Trust under award 076113 and 085475. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of any of the funders or supporters.

Figures and Tables

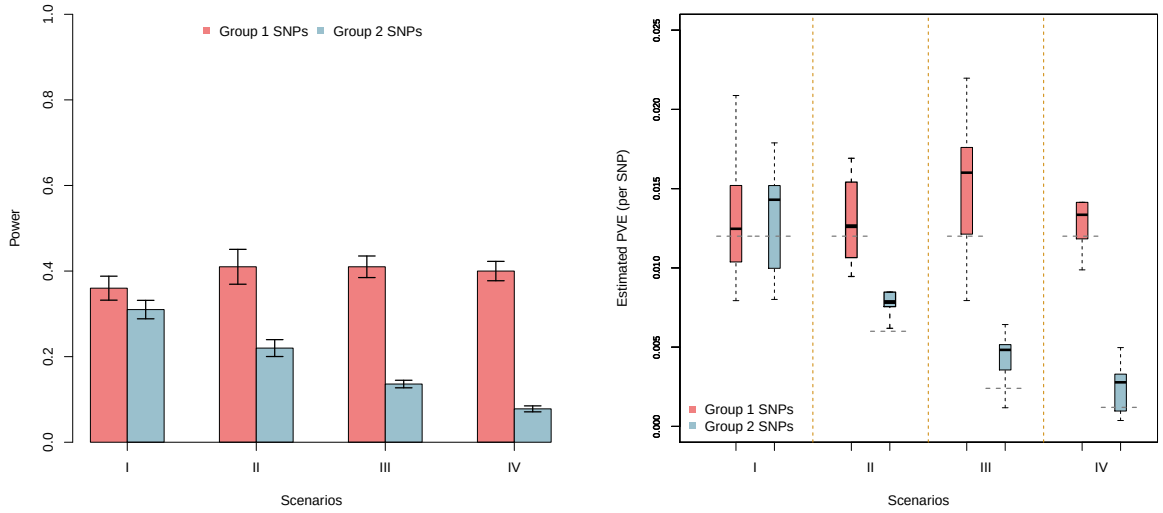


(a) Sample Size $n = 1,000$ (b) Sample Size $n = 1,750$ (c) Sample Size $n = 2,500$

Figure 1. Calibration of p-values produced by MEPIT via QQ plots. The QQ plots applying MEPIT to 100 simulated null datasets assuming sample sizes (a) 1,000, (b) 1,750, and (c) 2,500. Blue dots are p-values produced by under the normal test (or z-test), while the black dots represent p-values tested using the Davies method via a mixture of chi-square distributions. The 95% confidence intervals for the null hypothesis of no association are shown in grey.

Table 1. Empirical type I error estimates of MEPIT. Each entry represents type I error rate estimates as the proportion of p-values a under the null hypothesis based on 100 simulated continuous phenotypes for the normal test (or z-test) and the Davies method. Empirical size for the analyses used significance thresholds of $\alpha = 0.05$, 0.01, and 0.001. Sample sizes were set to 1,000, 1,750, and 2,500. Values in the parentheses are the standard deviations of the estimates.

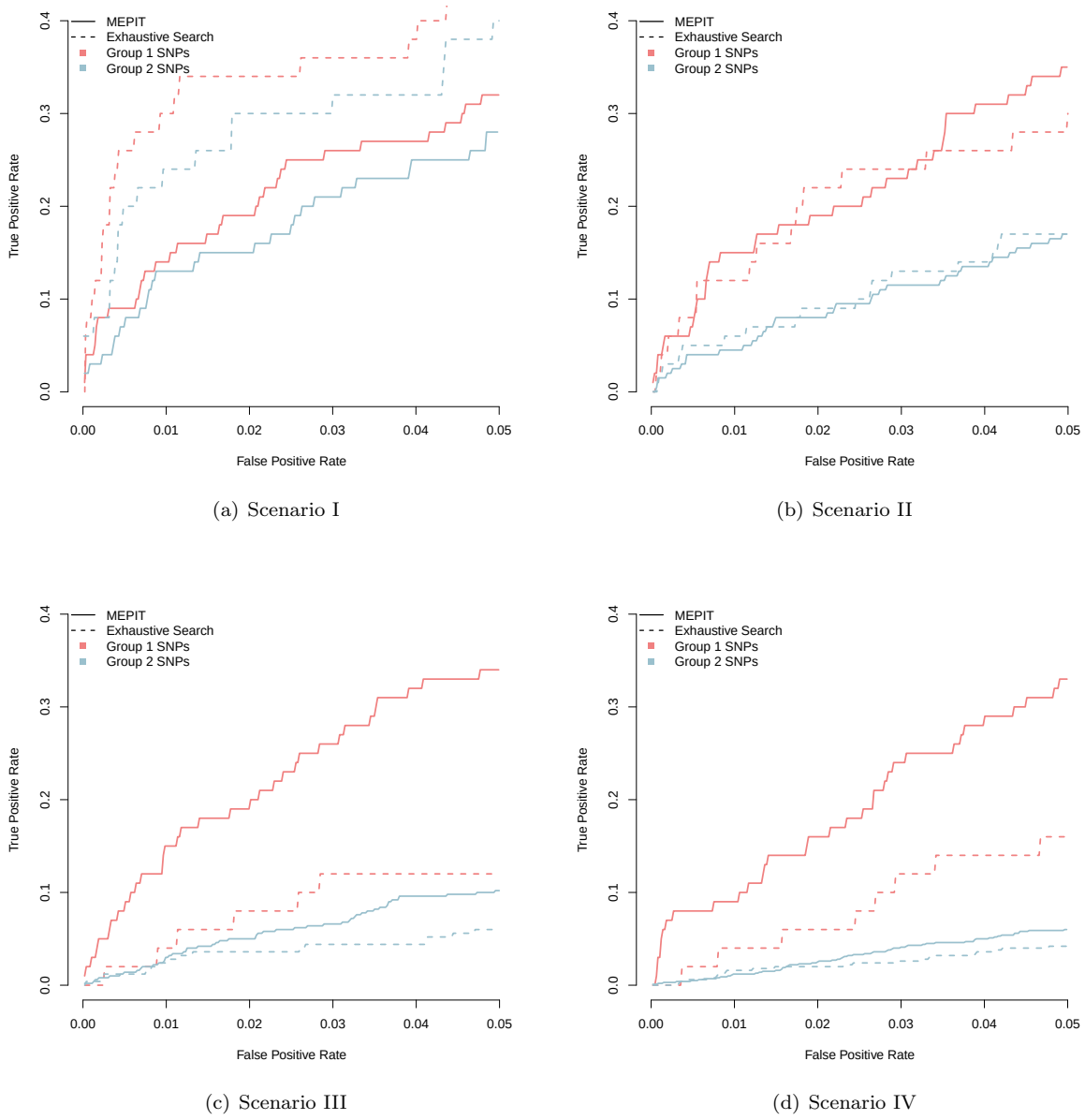
Test	Total Sample Size	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.001$
Normal Test	$n = 1{,}000$	0.0598 (0.0061)	0.0180 (0.0031)	0.0047 (0.0013)
	$n = 1{,}750$	0.0584 (0.0066)	0.0172 (0.0039)	0.0040 (0.0009)
	$n = 2{,}500$	0.0576 (0.0063)	0.0147 (0.0025)	0.0028 (0.0006)
Davies Method	$n = 1{,}000$	0.0563 (0.0104)	0.0121 (0.0042)	0.0012 (0.0008)
	$n = 1{,}750$	0.0528 (0.0083)	0.0108 (0.0023)	0.0011 (0.0004)
	$n = 2{,}500$	0.0469 (0.0073)	0.0093 (0.0024)	0.0009 (0.0005)



(a) Power ($\rho = 0.8$) with Significance Level ($\alpha = 0.08$)

(b) PVE Estimates ($\rho = 0.8$)

Figure 2. Empirical power to detect simulated causal interacting makers and estimating their marginal PVE. Groups 1 and 2 causal markers are colored in light red and light blue, respectively. These figures are based on a broad-sense heritability level of $H^2 = 0.6$ and parameter $\rho = 0.8$, estimated with 100 replicates. Here, $\rho = 0.8$ was used to determine the portion of broad-sense heritability contributed by interaction effects. Figure (a) shows the power of MEPIT to identify SNPs in each causal group under significance level $\alpha = 0.05$. The lines represent 95% variability due to resampling error. Figure (b) shows boxplots of the marginal PVE estimates for the group 1 and 2 causal SNPs from MEPIT for the four simulation scenarios. The true PVEs per causal SNP (0.012 for the group 1 SNPs; 0.012, 0.006, 0.0024, and 0.0012 for the Group 2 SNPs) are shown as dashed grey horizontal lines.



(a) Scenario I

(b) Scenario II

(c) Scenario III

(d) Scenario IV

Figure 3. Power analysis for detecting group 1 (light red) and group 2 (light blue) causal SNPs. We compare the mapping abilities of MEPIT (solid line) to an exhaustive search linear model (dotted line) in all scenarios (alternating panels), under broad-sense heritability level $H^2 = 0.6$ and $\rho = 0.8$. Here, $\rho = 0.8$ was used to determine the portion of broad-sense heritability contributed by interaction effects. The x-axis shows the false positive rate, while the y-axis gives the rate at which true causal variants were identified. Results are based on 100 replicates in each case.

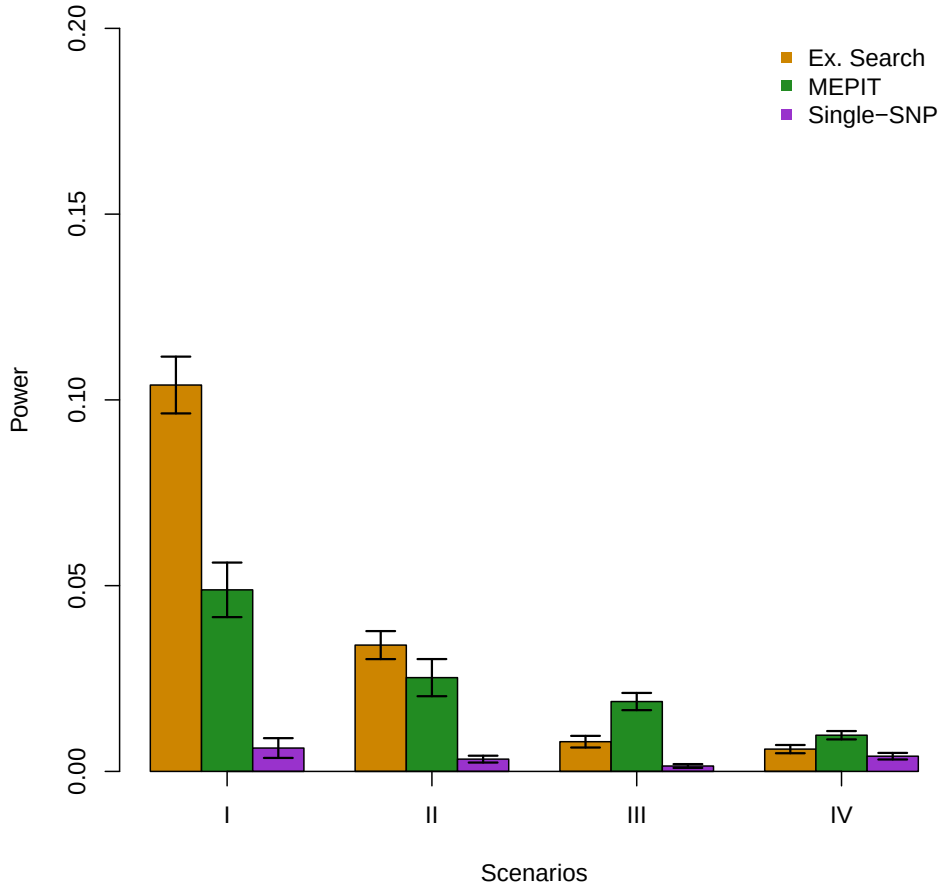


Figure 4. Analysis that compares the different proposed two-step procedures that check every pairwise SNP interaction to identify causal epistatic pairs. Specifically, we test the effectiveness of our method as an initial step in the pairwise detection filtration process by searching between the top 100 marginal significant SNPs identified by MEPIT (green) and a single-SNP linear model (purple). We use an exhaustive search linear model (orange) as a baseline comparison. We compare the three methods in all scenarios (x-axis), under broad-sense heritability level $H^2 = 0.6$. Here, $\rho = 0.8$ was used to determine the portion of broad-sense heritability contributed by interaction effects. The y-axis gives the rate at which true causal epistatic pairs were identified. Results are based on 100 replicates in each case. The lines represent 95% variability due to resampling error.

Table 2. Summary of the number of significant marginal and pairwise epistatic effects identified in the seven diseases from the WTCCC Dataset. Abbreviations are as follows: BD, bipolar disorder; CAD, coronary artery disease; CD, Crohn's disease; HT, hypertension; RA, rheumatoid arthritis; T1D, type 1 diabetes; T2D, type 2 diabetes. "Strong" and "Moderate" reflect the genome-wide significance thresholds for MEPIT p-values set to $P < 5 \times 10^{-7}$ and $1 \times 10^{-5} < P < 5 \times 10^{-7}$, respectively. G×G represents the number of interactions between the marginally strong and moderately associated markers with pairwise p-values less than $P < 8 \times 10^{-13}$.

Association	BD	CAD	CD	HT	RA	T1D	T2D
Strong	0	0	0	0	0	19	0
Moderate	0	0	0	0	1	8	0
G×G	-	-	-	-	-	210	-

Table 3. Regions of the genome with more than one SNP marginal epistatic p-value satisfying the moderate genome-wide significance threshold $P < 1 \times 10^{-5}$. Listed for all regions are the SNPs with the lowest marginal p-value. The marginal p-values reported are found via MEPIT. The reference column gives literature that have previously suggested some level of association between a given region and disease. ⋆The MHC region on chromosome 6 are divided into three classes: class I (29.8 Mb-31.6 Mb), class II (32.3 Mb-33.4 Mb), and class III (31.6 Mb-32.3 Mb). We treat each of these classes as independent loci. ‡Additional SNPs in its neighborhood also show strong marginal epistatic associations.

Disease	Chromosome	Region	SNP	MEPIT P	Nearest Gene	Reference
RA	6	MHC II⋆	rs7775228	2.0×10^{-6}	*HLA-DQA1*	[25, 32]
T1D	6	MHC I	rs2596437	2.4×10^{-6}	*HLA-DRB5*	[25, 32, 67]
T1D	6	MHC II‡	rs9270986	≈ 0	*HLA-DRB1*	[25, 32, 67]
T1D	6	MHC III‡	rs3131294	≈ 0	*NOTCH4*	[25, 32, 67]

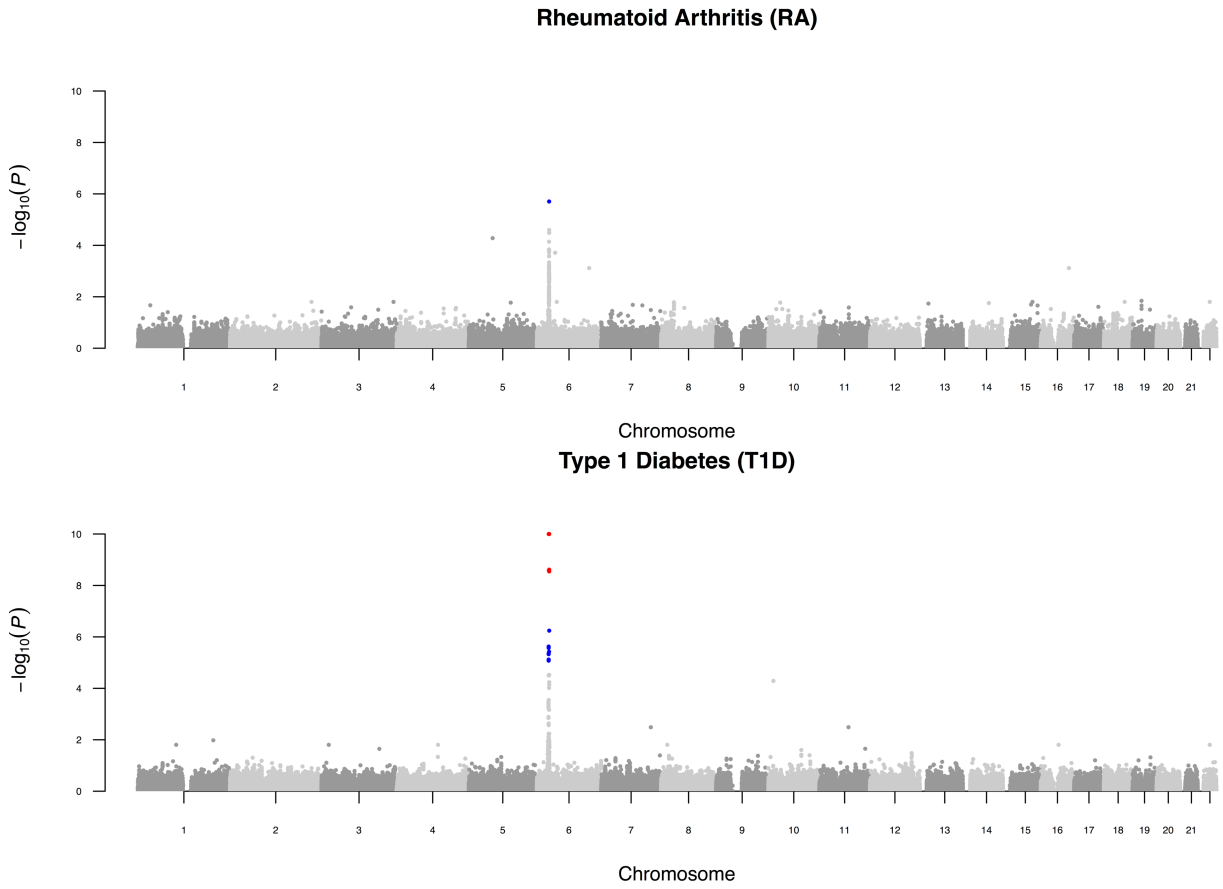


Figure 5. Genome-wide scan for rheumatoid arthritis (RA) and type 1 diabetes (T1D) in the WTCCC dataset. For both of the diseases, $-\log_{10}(P)$ of the MEPIT marginal epistatic p-value for quality-control-positive SNPs are plotted against position on each chromosome. Chromosomes are shown in alternating colors for clarity. All panels are truncated at $-\log_{10}(P) = 10$ for consistency and presentation, although the strongly associated markers in T1D had p-values $P \approx 0$. Markers exceeding the strong genome-wide significance threshold from the original analysis ($P < 5 \times 10^{-7}$) are colored in red. Markers colored in blue are those that are considered to have moderate associations with the phenotypes and have p-values greater than 5×10^{-7} and less than 1×10^{-5}.

Table 4. Moderate and strong associations in the two backfat quantitative traits within the pig dataset. Listed are the SNPs located in SSC1 with MEPIT p-values satisfying the moderate significant genome-wide threshold $P < 5 \times 10^{-7}$. The second column lists the positions (Mb) of the markers on the first chromosome. The third and fourth columns list the p-values for each variant corresponding to the 10^{th} and last rib backfat traits, respectively. The last column details the genes nearest to the significant locus. *Genes that are most likely to associated with both traits.

SNP	Position (Mb)	10^{th} Rib (P)	Last Rib (P)	Nearest Gene(s)
INRA0006067	274.1	1.2×10^{-6}	≈ 0	*APBA1*, *DOCK8*, *FAM189A2*, *KANK1**, *KLF9**, *PGM5*, *PIP5K1B*, *PTAR1*, *TJP2*
INRA0006074	274.7			
DRGA0002004	276.2			
INRA0006186	276.5			
INRA0006192	276.6			

References

1. Visscher PM, Brown MA, McCarthy MI, Yang J. Five Years of GWAS Discovery. The American Journal of Human Genetics. 2012 01;90(1):7–24. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3257326/.

2. Phillips PC. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet. 2008 10;9(11):855–867. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2689140/.

3. Carlborg O, Jacobsson L, Ahgren P, Siegel P, Andersson L. Epistasis and the release of genetic variation during long-term selection. Nat Genet. 2006 04;38(4):418–420. Available from: http://dx.doi.org/10.1038/ng1761.

4. Martin G, Elena SF, Lenormand T. Distributions of epistasis in microbes fit predictions from a fitness landscape model. Nat Genet. 2007;39:555–560.

5. Hill WG, Goddard ME, Visscher PM. Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. PLoS Genet. 2008 02;4(2):e1000008. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.1000008.

6. Bloom JS, Ehrenreich IM, Loo WT, Lite TLV, Kruglyak L. Finding the sources of missing heritability in a yeast cross. Nature. 2013 02;494(7436):234–237. Available from: http://dx.doi.org/10.1038/nature11867.

7. Shao H, Burrage LC, Sinasac DS, Hill AE, Ernest SR, O'Brien W, et al. Genetic architecture of complex traits: Large phenotypic effects and pervasive epistasis. Proceedings of the National Academy of Sciences. 2008 12;105(50):19910–19914. Available from: http://www.pnas.org/content/105/50/19910.abstract.

8. He X, Qian W, Wang Z, Li Y, Zhang J. Prevalent positive epistasis in Escherichia coli and Saccharomyces cerevisiae metabolic networks. Nat Genet. 2010 03;42(3):272–276. Available from: http://dx.doi.org/10.1038/ng.524.

9. Chari S, Dworkin I. The Conditional Nature of Genetic Interactions: The Consequences of Wild-Type Backgrounds on Mutational Interactions in a Genome-Wide Modifier Screen. PLoS Genet. 2013 08;9(8):e1003661. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.1003661.

10. Monnahan PJ, Kelly JK. Epistasis Is a Major Determinant of the Additive Genetic Variance in *Mimulus guttatus*. PLoS Genet. 2015 05;11(5):e1005201. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.1005201.

11. Mackay TFC. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. Nat Rev Genet. 2014 01;15(1):22–33. Available from: http://dx.doi.org/10.1038/nrg3627.

12. Brown AA, Buil A, Viñuela A, Lappalainen T, Zheng HF, Richards JB, et al. Genetic interactions affecting human gene expression identified by variance association mapping. eLife. 2014;3:e01381. Available from: https://dx.doi.org/10.7554/eLife.01381.

13. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet. 2010 11;42(11):985–990. Available from: http://dx.doi.org/10.1038/ng.694.

14. Hemani G, Shakhbazov K, Westra HJ, Esko T, Henders AK, McRae AF, et al. Detection and replication of epistasis influencing transcription in humans. Nature. 2014 04;508(7495):249–253. Available from: http://dx.doi.org/10.1038/nature13005.

15. Evans DM, Spencer CCA, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat Genet. 2011 08;43(8):761–767. Available from: http://dx.doi.org/10.1038/ng.873.

16. Jiang Y, Reif JC. Modeling Epistasis in Genomic Selection. Genetics. 2015 10;201(2):759–768. Available from: http://www.genetics.org/content/201/2/759.abstract.

17. Muñoz PR, Resende MFR, Gezan SA, Resende MDV, de los Campos G, Kirst M, et al. Unraveling Additive from Non-Additive Effects Using Genomic Relationship Matrices. Genetics. 2014 10;Available from: http://www.genetics.org/content/early/2014/10/15/genetics.114.171322.abstract.

18. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet. 2010 06;11(6):446–450. Available from: http://dx.doi.org/10.1038/nrg2809.

19. Hemani G, Knott S, Haley C. An Evolutionary Perspective on Epistasis and the Missing Heritability. PLoS Genet. 2013 02;9(2):e1003295. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.1003295.

20. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. Proceedings of the National Academy of Sciences. 2012 01;109(4):1193–1198. Available from: http://www.pnas.org/content/109/4/1193.abstract.

21. Wei WH, Hemani G, Haley CS. Detecting epistasis in human complex traits. Nat Rev Genet. 2014 11;15(11):722–733. Available from: http://dx.doi.org/10.1038/nrg3747.

22. Cordell HJ. Detecting gene-gene interactions that underlie human diseases. Nat Rev Genet. 2009 06;10(6):392–404. Available from: http://dx.doi.org/10.1038/nrg2579.

23. Ma L, Clark AG, Keinan A. Gene-Based Testing of Interactions in Association Studies of Quantitative Traits. PLoS Genet. 2013 02;9(2):e1003321. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.1003321.

24. Zhang X, Huang S, Zou F, Wang W. TEAM: efficient two-locus epistasis tests in human genome-wide association study. Bioinformatics. 2010 06;26(12):217–227. Available from: http://bioinformatics.oxfordjournals.org/content/26/12/i217.abstract.

25. Lippert C, Listgarten J, Davidson RI, Baxter J, Poon H, Kadie CM, et al. An Exhaustive Epistatic SNP Association Analysis on Expanded Wellcome Trust Data. Scientific Reports. 2013 01;3:1099 EP –. Available from: http://dx.doi.org/10.1038/srep01099.

26. Hemani G, Theocharidis A, Wei W, Haley C. EpiGPU: exhaustive pairwise epistasis scans parallelized on consumer level graphics cards. Bioinformatics. 2011 06;27(11):1462–1465.

27. Prabhu S, Pe'er I. Ultrafast genome-wide scan for SNP-SNP interactions in common complex disease. Genome Research. 2012 11;22(11):2230–2240. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3483552/.

28. Lewinger JP, Morrison JL, Thomas DC, Murcray CE, Conti DV, Li D, et al. Efficient two-step testing of gene-gene interactions in genome-wide association studies. Genetic Epidemiology. 2013 07;37(5):440–451.

29. Ueki M, Cordell HJ. Improved Statistics for Genome-Wide Interaction Analysis. PLoS Genet. 2012 04;8(4):e1002625. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.1002625.

30. Zhang Y, Liu JS. Bayesian inference of epistatic interactions in case-control studies. Nat Genet. 2007 09;39(9):1167–1173. Available from: http://dx.doi.org/10.1038/ng2110.

31. Tang W, Wu X, Jiang R, Li Y. Epistatic Module Detection for Case-Control Studies: A Bayesian Model with a Gibbs Sampling Strategy. PLoS Genet. 2009 05;5(5):e1000464. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.1000464.

32. Wan X, Yang C, Yang Q, Xue H, Fan X, Tang NLS, et al. BOOST: A Fast Approach to Detecting Gene-Gene Interactions in Genome-wide Case-Control Studies. The American Journal of Human Genetics. 2010 09;87(3):325–340. Available from: http://www.sciencedirect.com/science/article/pii/S0002929710003782.

33. Gyenesei A, Moody J, Laiho A, Semple CA, Haley CS, Wei WH. BiForce Toolbox: powerful high-throughput computational analysis of gene–gene interactions in genome-wide association studies. Nucleic Acids Research. 2012 07;40(Web Server issue):W628–W632. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3394281/.

34. Zhou X. A Unified Framework for Variance Component Estimation with Summary Statistics in Genome-wide Association Studies. bioRxiv. 2016 03;Available from: http://biorxiv.org/content/early/2016/03/08/042846.abstract.

35. Haseman JK, Elston RC. The investigation of linkage between a quantitative trait and a marker locus. Behavior Genetics. 1972;2(1):3–19. Available from: http://dx.doi.org/10.1007/BF01066731.

36. Drigalenko E. How Sib Pairs Reveal Linkage. The American Journal of Human Genetics. 1998;63(4):1243–1245. Available from: http://dx.doi.org/10.1086/302055.

37. Elston RC, Buxbaum S, Jacobs KB, Olson JM. Haseman and Elston revisited. Genetic Epidemiology. 2000;19(1):1–17. Available from: http://dx.doi.org/10.1002/1098-2272(200007)19:1<1::AID-GEPI1>3.0.CO;2-E.

38. Sham PC, Purcell S. Equivalence between Haseman-Elston and Variance-Components Linkage Analyses for Sib Pairs. American Journal of Human Genetics. 2001 06;68(6):1527–1532. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1226141/.

39. Sham PC, Purcell S, Cherny SS, Abecasis GR. Powerful Regression-Based Quantitative-Trait Linkage Analysis of General Pedigrees. American Journal of Human Genetics. 2002 08;71(2):238–253. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC379157/.

40. Chen WM, Broman KW, Liang KY. Quantitative trait linkage analysis by generalized estimating equations: Unification of variance components and Haseman-Elston regression. Genetic Epidemiology. 2004;26(4):265–272. Available from: http://dx.doi.org/10.1002/gepi.10315.

41. Chen GB. Estimating heritability of complex traits from genome-wide association studies using IBS-based Haseman-Elston regression. Frontiers in Genetics. 2014;5. Available from: http://www.frontiersin.org/Journal/Abstract.aspx?s=1187&name=statistical_genetics_and_methodology&ART_DOI=10.3389/fgene.2014.00107.

42. Golan D, Lander ES, Rosset S. Measuring missing heritability: Inferring the contribution of common variants. Proceedings of the National Academy of Sciences. 2014 12;111(49):E5272–E5281. Available from: http://www.pnas.org/content/111/49/E5272.abstract.

43. Churchill GA, Doerge RW. Naive Application of Permutation Testing Leads to Inflated Type I Error Rates. Genetics. 2008 01;178(1):609–610. Available from: http://www.genetics.org/content/178/1/609.abstract.

44. Abney M. Permutation Testing in the Presence of Polygenic Variation. Genetic epidemiology. 2015 05;39(4):249–258. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4634896/.

45. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010 07;42(7):565–569. Available from: http://dx.doi.org/10.1038/ng.608.

46. 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. American Journal of Human Genetics. 2011 07;89(1):82–93. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3135811/.

47. Zhou X, Carbonetto P, Stephens M. Polygenic Modeling with Bayesian Sparse Linear Mixed Models. PLoS Genet. 2013 02;9(2):e1003264. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.1003264.

48. Zhou X, Stephens M. Genome-wide Efficient Mixed Model Analysis for Association Studies. Nature genetics. 2012 07;44(7):821–824. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3386377/.

49. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nat Meth. 2014 04;11(4):407–409. Available from: http://dx.doi.org/10.1038/nmeth.2848.

50. Davies RB. Algorithm AS 155: The Distribution of a Linear Combination of χ^2 Random Variables. Journal of the Royal Statistical Society Series C (Applied Statistics). 1980;29(3):323–333. Available from: http://www.jstor.org/stable/2346911.

51. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007 06;447(7145):661–678. Available from: http://dx.doi.org/10.1038/nature05911.

52. Onteru SK, Fan B, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF. A whole-genome association study for pig reproductive traits. Animal Genetics. 2012;43(1):18–26. Available from: http://dx.doi.org/10.1111/j.1365-2052.2011.02213.x.

53. Fan B, Onteru SK, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF. Genome-Wide Association Study Identifies Loci for Body Composition and Structural Soundness Traits in Pigs. PLoS ONE. 2011 02;6(2):e14726–. Available from: http://dx.doi.org/10.1371%2Fjournal.pone.0014726.

54. Schaffner SF, Foo C, Gabriel S, Reich D, Daly MJ, Altshuler D. Calibrating a coalescent simulation of human genome sequence variation. Genome Research. 2005 11;15(11):1576–1583. Available from: http://genome.cshlp.org/content/15/11/1576.abstract.

55. Chen H, Meigs JB, Dupuis J. Sequence Kernel Association Test for Quantitative Traits in Family Samples. Genetic epidemiology. 2013 02;37(2):196–204. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3642218/.

56. Kuonen D. Saddlepoint Approximations for Distributions of Quadratic Forms in Normal Variables. Biometrika. 1999;86(4):929–935. Available from: http://www.jstor.org/stable/2673596.

57. Satterthwaite FE. An Approximate Distribution of Estimates of Variance Components. Biometrics Bulletin. 1946;2(6):110–114. Available from: http://www.jstor.org/stable/3002019.

58. Pilia G, Chen WM, Scuteri A, Orrú M, Albai G, Dei M, et al. Heritability of Cardiovascular and Personality Traits in 6,148 Sardinians. PLoS Genet. 2006 08;2(8):e132. Available from: http://dx.doi.org/10.1371%2Fjournal.pgen.0020132.

59. Feng S. Design and Association Methods for Next-generation Sequencing Studies for Quantitative Traits. University of Michigan; 2015.

60. Kam-Thong T, Pütz B, Karbalai N, MüllerMyhsok B, Borgwardt K. Epistasis detection on quantitative phenotypes by exhaustive enumeration using GPUs. Bioinformatics. 2011 07;27(13):i214–i221. Available from: http://bioinformatics.oxfordjournals.org/content/27/13/i214.abstract.

61. Kam-Thong T, Czamara D, Tsuda K, Borgwardt K, Lewis CM, Erhardt-Lehmann A, et al. EPIBLASTER-fast exhaustive two-locus epistasis detection strategy using graphical processing units. Eur J Hum Genet. 2011 04;19(4):465–471. Available from: http://dx.doi.org/10.1038/ejhg.2010.196.

62. Wu J, Devlin B, Ringquist S, Trucco M, Roeder K. Screen and Clean: a tool for identifying interactions in genome-wide association studies. Genetic epidemiology. 2010 04;34(3):275–285. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2915560/.

63. Moser G, Lee SH, Hayes BJ, Goddard ME, Wray NR, Visscher PM. Simultaneous Discovery, Estimation and Prediction Analysis of Complex Traits Using a Bayesian Mixture Model. PLoS Genet. 2015;11(4):e1004969. Available from: doi:10.1371/journal.pgen.1004969.

64. Zhang L, Liu R, Wang Z, Culver DA, Wu R. Modeling Haplotype-Haplotype Interactions in Case-Control Genetic Association Studies. Frontiers in Genetics. 2012;3:2. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3260479/.

65. Fernando MMA, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM, et al. Defining the Role of the MHC in Autoimmunity: A Review and Pooled Analysis. PLoS Genetics. 2008 04;4(4):e1000024. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2291482/.

66. Mosaad YM. Clinical Role of Human Leukocyte Antigen in Health and Disease. Scandinavian Journal of Immunology. 2015;82(4):283–306. Available from: http://dx.doi.org/10.1111/sji.12329.

67. Zhang Y, Zhang J, Liu JS. Block-based Bayesian epistasis association mapping with application to WTCCC type 1 diabetes data. Annals of Applied Statistics. 2011;5(3):2052–2077. Available from: http://projecteuclid.org/euclid.aoas/1318514295.

68. Crawford L, Wood KC, Zhou X, Mukherjee S. Bayesian Approximate Kernel Models with Variable Selection; 2016. ArXiv. Available from: http://arxiv.org/abs/1508.01217.

69. Valdar W, Solberg LC, Gauguier D, Cookson WO, Rawlins JNP, Mott R, et al. Genetic and Environmental Effects on Complex Traits in Mice. Genetics. 2006 10;174(2):959–984. Available from: http://www.genetics.org/content/174/2/959.abstract.

70. Nikkilä M, Stalder KJ, Mote BE, Lampe J, Thorn B, Rothschild MF, et al. Heritabilities and genetic correlations of body composition and structural soundness traits in commercial gilts. Animal Industry Report. 2008;654(1):98.

71. Huyghe JR, Jackson AU, Fogarty MP, Buchkovich ML, Stančáková A, Stringham HM, et al. Exome array analysis identifies novel loci and low-frequency variants for insulin processing and secretion. Nature genetics. 2013 02;45(2):197–201. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3727235/.

72. Farmer SR. Transcriptional control of adipocyte formation. Cell metabolism. 2006 10;4(4):263–273. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1958996/.

73. Wei S, Zhang L, Zhou X, Du M, Jiang Z, Hausman GJ, et al. Emerging roles of zinc finger proteins in regulating adipogenesis. Cellular and molecular life sciences : CMLS. 2013 12;70(23):4569–4584. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4100687/.

74. Cristancho AG, Lazar MA. Forming functional fat: a growing understanding of adipocyte differentiation. Nat Rev Mol Cell Biol. 2011 11;12(11):722–734. Available from: http://dx.doi.org/10.1038/nrm3198.

75. Rönnegård L, Pong-Wong R, Carlborg Ö. Defining the Assumptions Underlying Modeling of Epistatic QTL Using Variance Component Methods. Journal of Heredity. 2008 07;99(4):421–425. Available from: http://jhered.oxfordjournals.org/content/99/4/421.abstract.

76. Young AI, Durbin R. Estimation of Epistatic Variance Components and Heritability in Founder Populations and Crosses. Genetics. 2014 12;198(4):1405–1416. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4256760/.

77. Martini JWR, Wimmer V, Erbe M, Simianer H. Epistasis and covariance: how gene interaction translates into genomic relationship. Theoretical and Applied Genetics. 2016;129(5):963–976. Available from: http://dx.doi.org/10.1007/s00122-016-2675-5.

78. Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics (Oxford, England). 2012 09;13(4):762–775. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3440237/.

79. Broadaway KA, Cutler DJ, Duncan R, Moore JL, Ware EB, Jhun MA, et al. A Statistical Approach for Testing Cross-Phenotype Effects of Rare Variants. The American Journal of Human Genetics. 2016/06/02;98(3):525–540. Available from: http://dx.doi.org/10.1016/j.ajhg.2016.01.017.

80. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence Kernel Association Tests for the Combined Effect of Rare and Common Variants. The American Journal of Human Genetics. 2013 6;92(6):841–853. Available from: http://www.sciencedirect.com/science/article/pii/S0002929713001766.

81. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-Variant Association Analysis: Study Designs and Statistical Tests. The American Journal of Human Genetics. 2014 7;95(1):5–23. Available from: http://www.sciencedirect.com/science/article/pii/S0002929714002717.

82. Stephens M. A Unified Framework for Association Analysis with Multiple Related Phenotypes. PLoS ONE. 2013 07;8(7):e65245–. Available from: http://dx.doi.org/10.1371%2Fjournal.pone.0065245.

83. Tung J, Zhou X, Alberts SC, Stephens M, Gilad Y. The genetic architecture of gene expression levels in wild baboons. eLife. 2015;4:e04729. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4383332/.

84. Huang Y, Wuchty S, Przytycka TM. eQTL Epistasis –Challenges and Computational Approaches. Frontiers in Genetics. 2013;4:51. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3668133/.

85. Becker J, Wendland JR, Haenisch B, Nöthen MM, Schumacher J. A systematic eQTL study of cis–trans epistasis in 210 HapMap individuals. European Journal of Human Genetics. 2012 01;20(1):97–101. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3234520/.

86. Kapur K, Schüpbach T, Xenarios I, Kutalik Z, Bergmann S. Comparison of Strategies to Detect Epistasis from eQTL Data. PLoS ONE. 2011 12;6(12):e28415–. Available from: http://dx.doi.org/10.1371%2Fjournal.pone.0028415.