Published in final edited form as:

Genet Epidemiol. 2014 December ; 38(8): 722–727. doi:10.1002/gepi.21848.

Adjusting Family Relatedness in Data-driven Burden Test of Rare Variants

Qunyu Zhang^{1,*}, Lihua Wang¹, Dan Koboldt², Ingrid B. Boreki¹, and Michael A. Province¹

¹Division of Statistical Genomics, Washington University School of Medicine, St. Louis, MO 63108

²The Genome Institute, Washington University School of Medicine, St. Louis, MO 63108

Abstract

Family data represents a rich resource for detecting association between rare variants (RVs) and human traits. However, most RV association analysis methods developed in recent years are data-driven burden tests which can adaptively learn weights from data but require permutation to evaluate significance, thus are not readily applicable to family data, because random permutation will destroy family structure. Direct application of these methods to family data may result in a significant inflation of false positives. To overcome this issue, we have developed a generalized, weighted sum mixed model (WSMM) and corresponding computational techniques that can incorporate family information into data-driven burden tests, and allow adaptive and efficient permutation test in family data. Using simulated and real datasets, we demonstrate that the WSMM method can be used to appropriately adjust for genetic relatedness among family members and has a good control of type I errors. We compare WSMM with a non-data-driven, family-based Sequence Kernel Association Test (famSKAT), showing that WSMM has significantly higher power in some cases. WSMM provides a generalized, flexible framework for adapting different data-driven burden tests to analyze data with any family structures, and it can be extended to binary and time-to-onset traits, with or without covariates.

Keywords

rare variants; burden test; family data; mixed model; permutation

Introduction: Advantage and Limitation of Data-driven Burden Tests

Recent advances in sequencing and genotyping technologies, together with the recognition of “missing heritability” in complex human traits have driven a rapidly-growing interest in detecting rare genetic variants that contribute to human diseases. One important hypothesis is that multiple rare variants (RVs) within a genetic unit (usually a gene) may collectively contribute to a significant portion of genetic variation that cannot be explained by common variants targeted in most genome-wide association studies (GWAS). Such collective effects due to enrichment or over-representation of multiple RVs have been successfully observed

*Correspondence to: Qunyu Zhang, Division of Statistical Genomics, Washington University School of Medicine, 4444 Forest Park Blvd., Campus Box 8506, St. Louis, MO 63108. qunyu@wustl.edu.

in many studies[Ahituv, et al. 2007; Buxbaum 2009; Cohen, et al. 2004; Cohen, et al. 2006; Haller, et al. 2009; Ingram, et al. 2009; Ji, et al. 2008; Knight, et al. 2009; Morris, et al. 2009; Nejentsev, et al. 2009; Romeo, et al. 2009; Sabatelli, et al. 2009]. Based on this hypothesis, and also because of the lack of power in individual variant tests, a strategy that tests multiple RVs as a group (not individually) has been widely accepted and, following this strategy, a variety of RV association test methods have been developed[Han and Pan 2010; Ionita-Laza, et al. 2011; Lee, et al. 2012; Li and Leal 2008; Lin and Tang 2011; Madsen and Browning 2009; Morgenthaler and Thilly 2007; Morris and Zeggini 2010; Neale, et al. 2011; Pan and Shen 2011; Price, et al. 2010; Sun, et al. 2013; Wu, et al. 2011; Zhang, et al. 2011]. According to whether multiple RVs are summarized into a single variable or kept as individual variables or treated in a hybrid way (i.e., some are summarized and some are not), these methods can be categorized into burden tests (e.g. CMC[Li and Leal 2008]), non-burden tests (e.g. SKAT[Wu, et al. 2011]) and unified tests (e.g. SKAT-O[Lee, et al. 2012]); according to whether the weights for individual variants are learned from both genotype and phenotype data, these methods can be categorized into data-driven and non-data-driven methods (Table S1). Most of these methods are data-driven burden test methods and require permutation to evaluate significance. They were originally developed for the data from unrelated subjects, and thus are not readily applicable to the data from genetically related subjects sampled from families.

The potential enrichment of the RVs that segregate with phenotypes of interest in pedigrees makes family studies particularly attractive for uncovering phenotype-associated RVs [Curtis 2011; De, et al. 2013; Fang, et al. 2012; Feng, et al. 2012; Guo and Shugart 2012; Ionita-Laza, et al. 2013; Kzama and Bailey 2011; Yip, et al. 2012; Zhu, et al. 2009] and a few family-based conditional tests have been developed [De, et al. 2013; Fang, et al. 2012; Ionita-Laza, et al. 2013]. Since these methods require conditioning on parental genotypes (or sufficient statistic for parental genotypes), when the observed genotypes in a pedigree cannot be sufficiently nuclearized, the total informative sample size of a given dataset could be substantially reduced, resulting in a significant loss of power. As extensions of a sequence kernel association test (SKAT)[Wu, et al. 2011], some methods that can incorporate genetic relatedness into an analysis for general pedigree data have also been developed [Chen, et al. 2013; Ouakacha, et al. 2013]. Unlike burden tests collapsing multiple variants into one variable (which is a strategy adopted by the majority of RV association test methods), SKAT is a non-burden score test based on a multivariate variance component model in which individual variants are treated as individual variables. Usually, SKAT is more powerful when the collective effect of a group of variants to be tested is small (i.e. close to the null), whereas burden tests are more powerful when the collective effect increases (especially when the portion of causal variants increases) [Lee, et al. 2012]. More importantly, as a major type of burden tests, data-driven burden tests can utilize the information (on individual variants) learned from observed data itself to construct a weighted test, which may significantly boost the power in some cases. These tests, however, require permutation to correctly assess statistical significance, and thus cannot be directly applied to family data, because direct permutation will destroy the family structure and usually results in an inflation of false positives. Here we report and describe statistical

techniques that we have developed to overcome this problem; these techniques allow us to extend most data-driven burden tests to the data with arbitrary family structures.

Techniques for Adapting Data-driven Burden Tests to Family Data

In order to incorporate family structure into burden tests, we borrow the mixed-model idea that has been widely used in common variant association analyses [Chen and Yang 2009; Kang, et al. 2010; Kang, et al. 2008; Yu, et al. 2006], and propose a flexible, generalized framework, referred to as Weighted Sum Mixed Model (WSMM), which is based on a mixed linear model that includes a weighted sum score as a predictor:

$$\mathbf{Y} = \alpha + \beta \sum_{i=1}^m w_i g_i + \mathbf{U}\tau + \mathbf{Z}\gamma + \varepsilon \quad (1)$$

where $\mathbf{Y}$ is the observed quantitative trait (it can be extended to other types of data through different link functions), α the intercept, β the collective effect coefficient of a set of m RVs of interest, w_i the weight of variant i , g_i the observed genotype of variant i (usually coded as 0, 1 and 2 for an additive genetic model), τ the effect vector of covariate(s) and $\mathbf{U}$ the covariate data matrix, γ the polygenic effect vector for individuals and $\mathbf{Z}$ the corresponding design matrix, ε the residual. In this model, β and τ are fixed effects whereas γ and ε are random effects; ε is assumed to follow an univariate normal distribution of $N(0, \sigma_e^2)$ and γ a multivariate normal distribution of $N(0, \mathbf{G})$; here $\mathbf{G}$ is the variance-covariance matrix of γ , defined as, $\mathbf{G} = 2\mathbf{K}\sigma_A^2$, where $\mathbf{K}$ is the kinship matrix and σ_A^2 the additive polygenic variance. $\mathbf{K}$ can be estimated from the pedigree data and/or molecular marker genotypes. By incorporating $\mathbf{K}$ in the likelihood calculation, genetic relatedness between individuals within families can be taken into account and therefore adjusted in a test.

The component $\sum w_i g_i$ in the WSMM model is a weighted sum score (denoted by S) of m variants, which is a generalized feature shared by many different burden tests. In terms of collapsing of RVs, most existing burden test methods can be viewed as special instances of this model with different weighting methods (i.e., how to calculate w_i). For example, Li and Leal's CMC method [Li and Leal 2008] is equivalent to setting $w_i = 1$ for all RVs and $S = 1$ if $S > 1$. Madsen and Browning's weighted sum statistic (WSS) approach [Madsen and Browning 2009] calculates w_i based on allele frequency estimated from controls. Han and Pan's adaptive sum (aSum) test [Han and Pan 2010] recodes genotypes (equivalent to choosing $w_i = 1$ or -1) according to the direction of estimated regression coefficient and a pre-defined p-value cutoff. Ionita-Laza et al.'s replication-based method [Ionita-Laza, et al. 2011] defines w_i as the $-\log_{10}$ transformed probability of a variant occurring at most k times in controls and at least k times in cases, under the null hypothesis of no association between the variant and disease. Zhang et al.'s p-value weighted sum test (PWST) [Zhang, et al. 2011] calculates weights as rescaled left-tailed p-values obtained from individual variant tests. Due to this feature, most burden tests can be modified and incorporated into the WSMM model.

According to what information is used to determine w_i , burden tests can be classified into two categories, data-driven methods and non-data-driven methods. When data-driven

methods learn w_i from the observed data itself, non-data-driven methods determine w_i based on other prior knowledge (such as variants' deleteriousness predicted by bioinformatics tools). More precisely, the essential difference between data-driven and non-data-driven methods is whether observations of a phenotype are involved in the calculation of w_i . For example, CMC is a non-data-driven method, because its variant collapsing relies on only genotypes; WSS, aSum and PWST are data-driven methods, because phenotype data are used in the calculation of w_i .

A major advantage of data driven methods over non-data-driven methods is that they requires no specific biological assumptions for weighting variants. For example, it is not required to assume the association between RV frequency and disease risk, or the same direction of effects of causal variants. Instead, data-driven methods estimate importance and effect direction each variant from observed data, and then use that information to construct weights for individual variants in the calculation of sum scores. Due to this feature, data-driven methods are usually more powerful than non-data-driven methods [Han and Pan 2010; Ionita-Laza, et al. 2011; Lin and Tang 2011; Liu and Leal 2010; Zhang, et al. 2011], especially when sample size is large enough for weight learning.

We have previously reported that the WSMM model can be directly applied with non-data-driven methods (e.g. CMC) to family data and has a good control for the inflation of false positives [Zhang, et al. 2012]; when combined with data-driven methods, however, WSMM becomes computationally intensive and needs a modification. Since w_i in data-driven tests already contains information extracted from the observed data, re-fitting it into a WSMM model may result in a significant inflation of false positives if a regular test is used; usually, permutation is required to obtain the correct distribution of statistic under the null for a significance test. Coupled with family-based mixed model, permutation makes data-driven tests more challenging, not only because of more computer resources demanded by mixed model, but also due to the requirement of permuting data without destroying family structure. To overcome these issues, we propose to utilize the techniques below to build an efficient WSMM test procedure.

Mixed model based permutation

Since a regular permutation that randomly shuffles samples will destroy family structure, we suggest a mixed model based permutation. According to model (1), family relatedness between subjects can be adjusted by incorporating a polygenic variance (σ_A^2) and a kinship matrix ($\mathbf{K}$) in the calculation of the covariance matrix of $\mathbf{Y}$, therefore, as long as subject IDs are kept consistent between $\mathbf{Y}$ and $\mathbf{K}$, family relatedness will be maintained in the model even with permutation. This can be done by permuting $\mathbf{Y}$ and $\mathbf{K}$ simultaneously or permuting genotype data g_i (Table S2). Since permuting $\mathbf{Y}$ and $\mathbf{K}$ requires more computer time (especially when sample size and $\mathbf{K}$ are large), we choose to permute g_i . In model (1), because only g_i is permuted, $\mathbf{Y}$, $\mathbf{U}$ and $\mathbf{Z}$ are not permuted, both covariate effect(s) and family relatedness are retained in the model.

Adaptive permutation

In the application of WSMM, most computer burden comes from permutation test. To reduce it, instead of choosing a fixed number of permutations for all tests, we propose to use an adaptive permutation test (APT) procedure in which the number of permutations for a test is determined by monitoring and evaluating the permutation result in a dynamical way. In the APT procedure for a group of variants, an initial p-value is obtained first by a small number (e.g., 30) of permutations and compared with a pre-defined precision criterion. If the p-value satisfies the criterion, the permutation will be terminated; otherwise, the p-value will be updated with more permutations. This procedure is repeated until the obtained p-value satisfies the criterion. The criterion used in APT reflects the expected precision of p-value and can be defined by users in different ways. We choose to use the criterion whether the confidence interval of an estimated p-value satisfies a pre-defined cutoff (Supplemental Methods). A basic feature of the APT procedure is that less permutations are used for larger p-values and more permutations for smaller p-values, which can save large amount of computer time without significant loss of accuracy (Figure S1), especially when there are many genes to be tested and most of them are under the null.

Sparse matrix

Even with adaptive permutation, the computational burden of fitting mixed model can still be substantial, primarily due to the mathematical manipulation of the $\mathbf{K}$ matrix, especially when sample size (N) is large, because $\mathbf{K}$ is a $N \times N$ matrix and could be very huge and demand large amount of both computer memory and CPU time. However, if $\mathbf{K}$ is calculated based on pedigree information from multiple families (not estimated from genotype data), it will be a sparse matrix with many 0 elements, because the kinship coefficient between any two subjects from different families is usually set to 0. The non-zero elements in $\mathbf{K}$ can be organized into many small blocks, stored in a linearly indexed vector and then efficiently processed block by block. This technique can substantially reduce computer memory and time used in the mixed model analysis for the data with large sample sizes.

Re-using estimated parameter

In the procedure of fitting the WSMM model to observed data, a significant portion of time is spent on iterative estimation of polygenic variance σ_A^2 . Since the estimation of σ_A^2 depends mainly on $\mathbf{Y}$ and $\mathbf{K}$ and the permutation of g_i has very weak effect on it when the heritability of g_i is low (Figure S2), we propose to first estimate σ_A^2 using non-permuted data, then apply the estimated value of σ_A^2 in all permutations. This allows us to further save a substantial amount of computer time and obtain very similar results (Figure S3).

Applications to Simulated and Real Data

We incorporated these WSMM-related techniques with a data driven burden test method, p-value weighed sum test (PWST)[Zhang, et al. 2011] (Supplemental Methods), and applied them to three data sets with family structures (Table S3), including one simulated data set from Genetic Analysis Workshop 17 (GAW17)[Almasy, et al. 2012] and two real data sets from the NHLBI Family Heart Study (FamHS)[Higgins, et al. 1996] and the NIA Long Life

Family Study (LLFS)[An, et al. 2013; Newman, et al. 2011]. The quantile-quantile (QQ) plots of p-values obtained by WSMM for all three data sets are very close to the expected uniform distribution (under the null), showing no significant inflation (Figure 1a, 1b, 1c). As a comparison, direct application of PWST to the data without using WSMM (i.e., ignoring family structure and treating the data as from unrelated subjects) produces significantly inflated p-values (Figure 1d, 1e, 1f).

We compared the power of WSMM with a recently published, non-burden, non-data-driven, family-based sequence kernel association test (famSKAT) method[Chen, et al. 2013] using the GAW17 simulated data of 12 genes with effects (Table S4). According to the estimated receiver operating characteristic (ROC) curves (Figure 2), WSMM shows significantly higher power for some genes (*LPL*, *PDGFD*, *PLAT* and *SIRT1*) and similar or slightly lower power for other genes.

From the applications above, it can be seen that the WSMM model not only can adjust relatedness in family data and therefore has a good control for the inflation of false positives, but also can benefit from data-driven techniques and increase statistical power in detecting rare-variant association in some cases. Besides, the proposed computational techniques allow us to perform the WSMM analysis in an efficient way, making it feasible to apply WSMM to large datasets with thousands of subjects and thousands of genes (e.g., the FamHS data).

More importantly, WSMM is not limited to a single method, it provides a flexible, generalized framework that can extend most data-driven burden tests for the data with any family structures. Although we demonstrated WSMM only for quantitative traits and one data-driven test method (i.e., PWST), this framework can be combined with different data-driven burden tests (by choosing different weighting methods) and extended for binary and time-to-onset traits, with or without covariates (Table S5).

There are a few limitations in the WSMM model and techniques. First, WSMM is essentially an association test and the proposed permutation test only works for association test when the heritability of the tested gene is low. It may produce biased results when the heritability is high or if the model is modified for linkage test using the IBD information (which will be destroyed by permutation). Second, although we have substantially reduced the computational burden of WSMM using multiple techniques, it is still more time-consuming than most non-permutation based methods, and more computational time should be expected for the tests with larger effect sizes, because their p-values could be extremely small, thus requiring more permutations. Finally, due to many factors (such as sample size, gene size, portion and number of causal variants in a gene and their effect sizes, as well as family size and family structure) and their interactions, it is still not clear for us in what situations WSMM can produce the best power. Further investigations on statistical power of WSMM are still warranted, including family-data based comparisons between WSMM and non-burden tests and between different WSMM-adapted data-driven burden tests.

Software

The R program for the WSMM analysis of quantitative trait is available for download at <https://dsgweb.wustl.edu/qunyuan/software/wsmm/>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grants from the US National Institute of Health (NIH): 5U01AG02374607, 1R01DK8925601 and R01DK075681.

References

- Ahituv N, Kavaslar N, Schackwitz W, Ustaszewska A, Martin J, Hebert S, Doelle H, Ersoy B, Kryukov G, Schmidt S, et al. Medical sequencing at the extremes of human body mass. *Am J Hum Genet.* 2007; 80(4):779–91. [PubMed: 17357083]
- Almasy L, Dyer TD, Peralta JM, Kent JW Jr, Charlesworth JC, Curran JE, Blangero J. Genetic Analysis Workshop 17 mini-exome simulation. *BMC Proc.* 2012; 5(Suppl 9):S2. [PubMed: 22373155]
- An P, Miljkovic I, Thyagarajan B, Kraja AT, Daw EW, Pankow JS, Selvin E, Kao WH, Maruthur NM, Nalls MA, et al. Genome-wide association study identifies common loci influencing circulating glycosylated hemoglobin (HbA) levels in non-diabetic subjects: The Long Life Family Study (LLFS). *Metabolism.* 2013
- Buxbaum JD. Multiple rare variants in the etiology of autism spectrum disorders. *Dialogues Clin Neurosci.* 2009; 11(1):35–43. [PubMed: 19432386]
- Chen H, Meigs JB, Dupuis J. Sequence kernel association test for quantitative traits in family samples. *Genet Epidemiol.* 2013; 37(2):196–204. [PubMed: 23280576]
- Chen MH, Yang Q. GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics.* 2009; 26(4):580–1. [PubMed: 20040588]
- Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science.* 2004; 305(5685):869–72. [PubMed: 15297675]
- Cohen JC, Pertsemlidis A, Fahmi S, Esmail S, Vega GL, Grundy SM, Hobbs HH. Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proc Natl Acad Sci U S A.* 2006; 103(6):1810–5. [PubMed: 16449388]
- Curtis D. Assessing the contribution family data can make to case-control studies of rare variants. *Ann Hum Genet.* 2011; 75(5):630–8. [PubMed: 21675965]
- De G, Yip WK, Ionita-Laza I, Laird N. Rare variant analysis for family-based design. *PLoS ONE.* 2013; 8(1):e48495. [PubMed: 23341868]
- Fang S, Sha Q, Zhang S. Two adaptive weighting methods to test for rare variant associations in family-based designs. *Genet Epidemiol.* 2012; 36(5):499–507. [PubMed: 22674630]
- Feng T, Elston RC, Zhu X. A novel method to detect rare variants using both family and unrelated case-control data. *BMC Proc.* 2012; 5(Suppl 9):S80. [PubMed: 22373319]
- Guo W, Shugart YY. Detecting rare variants for quantitative traits using nuclear families. *Hum Hered.* 2012; 73(3):148–58. [PubMed: 22699804]
- Haller G, Torgerson DG, Ober C, Thompson EE. Sequencing the IL4 locus in African Americans implicates rare noncoding variants in asthma susceptibility. *J Allergy Clin Immunol.* 2009; 124(6):1204–9 e9. [PubMed: 19910025]
- Han F, Pan W. A data-adaptive sum test for disease association with multiple common or rare variants. *Hum Hered.* 2010; 70(1):42–54. [PubMed: 20413981]

- Higgins M, Province M, Heiss G, Eckfeldt J, Ellison RC, Folsom AR, Rao DC, Sprafka JM, Williams R. NHLBI Family Heart Study: objectives and design. *Am J Epidemiol.* 1996; 143(12):1219–28. [PubMed: 8651220]
- Ingram CJ, Raga TO, Tarekegn A, Browning SL, Elamin MF, Bekele E, Thomas MG, Weale ME, Bradman N, Swallow DM. Multiple rare variants as a cause of a common phenotype: several different lactase persistence associated alleles in a single ethnic group. *J Mol Evol.* 2009; 69(6): 579–88. [PubMed: 19937006]
- Ionita-Laza I, Buxbaum JD, Laird NM, Lange C. A new testing strategy to identify rare variants with either risk or protective effect on disease. *PLoS Genet.* 2011; 7(2):e1001289. [PubMed: 21304886]
- Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Family-based association tests for sequence data, and comparisons with population-based association tests. *Eur J Hum Genet.* 2013
- Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D, Lifton RP. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. *Nat Genet.* 2008; 40(5):592–9. [PubMed: 18391953]
- Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet.* 2010; 42(4):348–54. [PubMed: 20208533]
- Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E. Efficient control of population structure in model organism association mapping. *Genetics.* 2008; 178(3):1709–23. [PubMed: 18385116]
- Kazma R, Bailey JN. Population-based and family-based designs to analyze rare variants in complex diseases. *Genet Epidemiol.* 2011; 35(Suppl 1):S41–7. [PubMed: 22128057]
- Knight HM, Pickard BS, Maclean A, Malloy MP, Soares DC, McRae AF, Condie A, White A, Hawkins W, McGhee K, et al. A cytogenetic abnormality and rare coding variants identify ABCA13 as a candidate gene in schizophrenia, bipolar disorder, and depression. *Am J Hum Genet.* 2009; 85(6):833–46. [PubMed: 19944402]
- Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson 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(2):224–37. [PubMed: 22863193]
- Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet.* 2008; 83(3):311–21. [PubMed: 18691683]
- Lin DY, Tang ZZ. A general framework for detecting disease associations with rare variants in sequencing studies. *Am J Hum Genet.* 2011; 89(3):354–67. [PubMed: 21885029]
- Liu DJ, Leal SM. A novel adaptive method for the analysis of next-generation sequencing data to detect complex trait associations with rare variants due to gene main effects and interactions. *PLoS Genet.* 2010; 6(10):e1001156. [PubMed: 20976247]
- Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genet.* 2009; 5(2):e1000384. [PubMed: 19214210]
- Morgenthaler S, Thilly WG. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). *Mutat Res.* 2007; 615(1-2):28–56. [PubMed: 17101154]
- Morris AP, Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. *Genet Epidemiol.* 2010; 34(2):188–93. [PubMed: 19810025]
- Morris AP, Zeggini E, Lindgren CM. Identification of novel putative rheumatoid arthritis susceptibility genes via analysis of rare variants. *BMC Proc.* 2009; 3(Suppl 7):S131. [PubMed: 20017998]
- 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(3):e1001322. [PubMed: 21408211]
- Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science.* 2009; 324(5925):387–9. [PubMed: 19264985]

- Newman AB, Glynn NW, Taylor CA, Sebastiani P, Perls TT, Mayeux R, Christensen K, Zmuda JM, Barral S, Lee JH, et al. Health and function of participants in the Long Life Family Study: A comparison with other cohorts. *Aging* (Albany NY). 2011; 3(1):63–76. [PubMed: 21258136]
- Ouakacha K, Dastani Z, Li R, Cingolani PE, Spector TD, Hammond CJ, Richards JB, Ciampi A, Greenwood CM. Adjusted sequence kernel association test for rare variants controlling for cryptic and family relatedness. *Genet Epidemiol*. 2013; 37(4):366–76. [PubMed: 23529756]
- Pan W, Shen X. Adaptive tests for association analysis of rare variants. *Genet Epidemiol*. 2011
- 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(6):832–8. [PubMed: 20471002]
- Romeo S, Yin W, Kozlitina J, Pennacchio LA, Boerwinkle E, Hobbs HH, Cohen JC. Rare loss-of-function mutations in ANGPTL family members contribute to plasma triglyceride levels in humans. *J Clin Invest*. 2009; 119(1):70–9. [PubMed: 19075393]
- Sabatelli M, Eusebi F, Al-Chalabi A, Conte A, Madia F, Luigetti M, Mancuso I, Limatola C, Trettel F, Sobrero F, et al. Rare missense variants of neuronal nicotinic acetylcholine receptor altering receptor function are associated with sporadic amyotrophic lateral sclerosis. *Hum Mol Genet*. 2009; 18(20):3997–4006. [PubMed: 19628475]
- Sun J, Zheng Y, Hsu L. A unified mixed-effects model for rare-variant association in sequencing studies. *Genet Epidemiol*. 2013; 37(4):334–44. [PubMed: 23483651]
- Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet*. 2011; 89(1):82–93. [PubMed: 21737059]
- Yip WK, De G, Raby BA, Laird N. Identifying causal rare variants of disease through family-based analysis of Genetics Analysis Workshop 17 data set. *BMC Proc*. 2012; 5(Suppl 9):S21. [PubMed: 22373204]
- Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet*. 2006; 38(2):203–8. [PubMed: 16380716]
- Zhang Q, Chung D, Kraja A, Borecki II, Province MA. Methods for adjusting population structure and familial relatedness in association test for collective effect of multiple rare variants on quantitative traits. *BMC Proc*. 2012; 5(Suppl 9):S35. [PubMed: 22373066]
- Zhang Q, Irvin MR, Arnett DK, Province MA, Borecki I. A data-driven method for identifying rare variants with heterogeneous trait effects. *Genet Epidemiol*. 2011; 35(7):679–85. [PubMed: 21818776]
- Zhu X, Feng T, Li Y, Lu Q, Elston RC. Detecting rare variants for complex traits using family and unrelated data. *Genet Epidemiol*. 2009; 34(2):171–87. [PubMed: 19847924]

Figure 1. Q-Q plots of p-values from the analyses with and without applying WSMM to family data

In each plot, x-axis represents expected, uniformly distributed p-values and y-axis represents observed p-values, both in the $-\log_{10}$ scale. The observed p-values were obtained by applying WSMM (incorporated with PWST) to a) the GAW17 data simulated under the null, b) the LLFS data and c) the FHS data, and by applying only PWST (ignoring family structure and without using WSMM) to d) the GAW17 data, e) the LLFS data and f) the FHS data. All observed p-values are estimated using an APT procedure.

Figure 2. ROC curves of WSMM and famSKAT

Receiver operator characteristic (ROC) curves were estimated through 200 replications of a quantitative trait and genotypes of 12 genes (labeled at the top of each plot) with known effects. Both trait and genotypes are simulated data from GAW17.