A Joint Location-Scale Test Improves Power to Detect Associated SNPs, Gene Sets, and Pathways


To cite this version:
David Soave, Harriet Corvol, Naim Panjwani, Jiafen Gong, Weili Li, et al.. A Joint Location-Scale Test Improves Power to Detect Associated SNPs, Gene Sets, and Pathways. American Journal of Human Genetics, Elsevier (Cell Press), 2015, 97 (1), pp.125-138. <10.1016/j.ajhg.2015.05.015>. <hal-01172468>
A Joint Location-Scale Test Improves Power to Detect Associated SNPs, Gene Sets, and Pathways

David Soave,1,2 Harriet Corvol,3,4 Naim Panjwani,2 Jiafen Gong,2 Weili Li,1,2 Pierre-Yves Boelle,4,5 Peter R. Durie,6,7 Andrew D. Paterson,2,8 Johanna M. Rommens,2,9 Lisa J. Strug,1,2,* and Lei Sun1,10,*

Gene-based, pathway, and other multivariate association methods are motivated by the possibility of GxG and GxE interactions; however, accounting for such interactions is limited by the challenges associated with adequate modeling information. Here we propose an easy-to-implement joint location-scale (JLS) association testing framework for single-variant and multivariate analysis that accounts for interactions without explicitly modeling them. We apply the JLS method to a gene-set analysis of cystic fibrosis (CF) lung disease, which is influenced by multiple environmental and genetic factors. We identify and replicate an association between the constituents of the apical plasma membrane and CF lung disease (p = 0.0099 and p = 0.0180, respectively) and highlight a role for the SLC9A3-SLC9A3R1/2-EZR complex in contributing to CF lung disease. Many association studies could benefit from re-analysis with the JLS method that leverages complex genetic architecture for SNP, gene, and pathway identification. Analytical verification, simulation, and additional proof-of-principle applications support our approach.

Introduction

Identifying the genetic architecture of complex traits requires analytic strategies that move beyond single-variant association tests. Multivariate analyses such as gene-based, gene-by-gene interaction (GxG), gene-by-environment interaction (GxE), gene-set, and pathway analyses are now commonly implemented,1–5 yet, one rarely sees GxG or GxE explicitly accounted for within gene-set and pathway analyses.6 The specification of interacting variables that probably differ between genes in gene sets and pathways is not straightforward. Interacting exposure variables (termed E hereafter) could include contributing environmental factors or SNPs or haplotypes from the same region or at other susceptibility loci. Missing or incorrect information on interacting factors, as well as associated computational burden, might also limit more comprehensive surveys of the whole genome for disease association. Here we provide an easy-to-implement, straightforward solution to exploit potential GxG and GxE in gene-set and pathway analyses, and we illustrate the power of such an approach in a gene-modifier study of lung disease in cystic fibrosis (CF [MIM: 219700]).

CF is a life-limiting genetic disease, with the majority of mortality due to lung disease. CF is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR [MIM: 602421]). Individuals with the same loss-of-function CFTR mutations have variable disease severity across the lungs, intestine, and other CF-affected organs, suggesting that variation in other (modifier) genes might play a role in disease pathophysiology.7 A major source of pathophysiology in CF is impaired fluid and electrolyte flux in CF-affected organ epithelia. Sun et al.5 hypothesized that regulators of fluid, solute, and ion transport that physically co-localize with CFTR would play a role in determining CF disease severity. Using a composite tissue gene set that corresponds to 157 gene products localized to the apical plasma membrane of epithelia, a multivariate sum-test provided evidence that many of the constituent gene products contribute to early intestinal obstruction (called meconium ileus [MIM: 614665]) in CF. Meconium ileus is heritable (~88% estimated heritability) and presents at birth in ~15% of CF-affected individuals,8 with limited opportunity for environmental involvement.

Lung disease in CF is distinct, being progressive and notably affected by environmental exposures such as infection history, secondhand smoke, air pollution, and ambient air temperature,9–11 with a ~50% estimated heritability.12 Similar to the intestinal epithelia, the apical membrane constituents of lung epithelia should likewise contribute to disease. Testing this, or any hypothesis in CF lung disease, however, has the added difficulty of being able to specify and collect accurate information on the contributing exposure variables and model the potential GxE. Therefore, an alternative association testing methodology is needed to account for the effects of interacting variables that are either unknown or unmeasured.

1Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON MST 3M7, Canada; 2Program in Genetics and Genome Biology, Research Institute, The Hospital for Sick Children, Toronto, ON MSG 0A4, Canada; 3Assistance Publique-Hôpitaux de Paris (AP-HP), Tronche Hospital, Pediatric Pulmonology Department; Institut National de la Santé et la Recherche Médicale (INSERM), UMR S 938, CDR Saint-Antoine, 75012 Paris, France; 4Sorbonne Universités, Université Pierre et Marie Curie (UPMC) Paris 06, 75005 Paris, France; 5AP-HP, Saint-Antoine Hospital, Biostatistics Department, INSERM, UMR S 1136, 75012 Paris, France; 6Program in Physiology and Experimental Medicine, The Hospital for Sick Children, Toronto, ON MSG 0A4, Canada; 7Department of Pediatrics, University of Toronto, Toronto, ON M5S 1A1, Canada; 8Department of Statistical Sciences, University of Toronto, Toronto, ON M5S 3G3, Canada
*Correspondence: lisa.strug@utoronto.ca (L.J.S.), sun@utstat.toronto.edu (L.S.)
http://dx.doi.org/10.1016/j.ajhg.2015.05.015, ©2015 The Authors
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
For a biallelic SNP, a genetic interaction (either GxG or GxE) on a quantitative phenotype will lead to differences in phenotypic distribution across the three genotypes, leading to differences in phenotypic variance (scale). In light of this, Levene’s scale test of equality of variance has been proposed as a method of prioritizing SNPs for subsequent GxG and GxE studies, in contrast to the standard mean (location) test (i.e., testing for phenotypic differences in mean across genotypes). The advantage of using variance testing to incorporate GxG or GxE is that exposures need not be specified, and the enormous multiple testing burden of formally examining all possible pair-wise interactions is removed. Variance-only testing, however, has limited power to detect SNPs displaying main effects only.

Focusing on single-SNP analysis, Aschard et al. proposed a test that compares the percentiles of phenotypic values between genotypes, capturing mean, variance, and other differences. Although this approach comprehensively evaluates the phenotypic distribution between genotypes, it sacrifices statistical power when a (approximately) normally distributed trait is sufficiently summarized by its mean and variance. Furthermore, the distribution test statistic follows an asymptotic chi-square distribution under the null hypothesis of equal mean.

Pare et al. showed that the conditional variance of a SNP under investigation can increase power but is more sensitive to model assumptions such as normality. Therefore, Cao et al. proposed a parametric bootstrap method to calculate “honest” p values at the cost of computational efficiency. For both the LRT and distribution methods, implementation difficulties arise for multivariate (e.g., gene-based, gene-set, and pathway) analyses (see Appendix A).

Here we propose an easy-to-implement joint location-scale (JLS) testing framework that simultaneously tests the null hypothesis of equal mean and equal variance between genotypes, by aggregating association evidence from the individual location-only and scale-only tests, focusing on Fisher’s method of combining information (JLS-Fisher). The proposed method detects association in the presence of underlying genetic main and/or interaction effects, without specifying the interactors; it allows any type of individual location and scale tests to be combined, making it particularly useful for gene-based, gene-set, and pathway analyses.

Using two proof-of-principle examples in the single-variant testing setting, where GxE interactions have been shown to exist from previous studies, we demonstrate the power of the JLS method. The first example pertains to a locus that determines glycosylated hemoglobin (HbA1c) levels based on a genome-wide association study (GWAS) of type 1 diabetes (T1D [MIM: 222100]) complications. The second pertains to determinants of CF lung disease severity. For these examples, the known environmental interactors are not explicitly modeled, but we leverage their existence to identify genetic contributors to HbA1c and CF lung disease.

We then use the multivariate JLS method to identify and replicate an association between the apical gene set, and the SLC9A3 complex within the set, and lung disease severity in CF; this finding would have been missed by the traditional multivariate gene-set tests that do not account for (heterogeneous) GxG and GxE interactions. Through extensive simulation analyses, we demonstrate that the proposed JLS method has good type 1 error control with improved power compared to other testing options, and, in contrast to the recently proposed distribution and LRT joint testing methods, can be easily implemented in the context of gene-set and pathway analyses.

Material and Methods

Genetic Model and Notation

Let $Y$ be a quantitative trait and $G$ be the minor allele count for the SNP under investigation ($G = 0, 1, or 2$); the additive assumption is not critical to the method development. Let $E$ be an exposure variable interacting with the genetic factor. This exposure $E$ could reflect continuous or categorical measures of environmental or genetic background. The underlying true genetic model can include main effects of both $G$ ($h_G$) and $E$ ($h_E$) on $Y$, as well as the interaction effect ($h_{GE}$):

$$Y \sim \beta_0 + \beta_G G + \beta_E E + \beta_{GE} GE + \epsilon.$$  \hspace{1cm} (Equation 1)

We assume that the trait $Y$ is normally distributed with unit variance conditional upon $G$ and $E$, in other words, $\text{Var}(Y | G = g, E = e) = 1$ and $\epsilon \sim \mathcal{N}(0, 1)$.

When considering only $G$, the working model would reduce to

$$Y \sim \beta_0 + \beta_G G + \epsilon_G.$$  \hspace{1cm} (Equation 2)

Pare et al. showed that the conditional variance of $Y$ conditional on $G$ alone could be expressed as $\sigma_Y^2 = \text{Var}(Y | G = g) = (\beta_E + \beta_{GE})^2 + 1$. Thus, if an interaction effect was present (i.e., $\beta_{GE} \neq 0$), the trait variance would differ between genotypes.

Joint Location-Scale Testing Procedure for Single-SNP Analysis

Our proposed JLS testing framework, based on the working model of Equation 2, tests the following null hypothesis:

$$H_0^{\text{joint}}: \beta_G = 0 \text{ and } \sigma_i = \sigma_j \text{ for all } i \neq j, i, j = 0, 1, 2.$$  

The alternative hypothesis of interest is

$$H_1^{\text{joint}}: \beta_G \neq 0 \text{ or } \sigma_i \neq \sigma_j \text{ for some } i \neq j.$$

For a SNP under study, different JLS test statistics can be considered. Let $p_1$ be the $p$ value for the location test of choice (i.e., testing $H_0^{\text{location}}: \beta_G = 0$ using, for example, ordinary least-squares regression), and $p_2$ be the $p$ value for the scale test of choice (i.e., testing $H_0^{\text{scale}}: \sigma_i = \sigma_j$ for all $i \neq j$ using, for example, Levene’s test). We first consider Fisher’s method (JLS-Fisher) to combine
the association evidence from the individual location and scale tests. The JLS-Fisher statistic is defined as

\[ W_F = -2 \left( \log(p_L) + \log(p_S) \right). \]

Large values of \( W_F \) correspond to small values of \( p_L \) and/or \( p_S \) and provide evidence against the null hypothesis for a normally distributed trait (Appendix A, Lemma 1), as well as empirically for approximately normally distributed traits in finite samples (Figures S1 and S2, Tables S1 and S2).

One can also consider the minimum p value (JLS-minP) approach, or various alternatives based on combining the individual test statistics themselves with or without weights.\(^5\) The JLS-minP statistic is defined as

\[ W_M = \min(p_L, p_S). \]

If \( p_L \) and \( p_S \) are independent under \( H_{0,j}^{\text{joint}} \), \( W_M \) is distributed as a Beta random variable (with shape parameters 1 and 2) where small values of \( W_M \) correspond to small values of \( p_L \) and/or \( p_S \) and evidence against the null.

Joint Location-Scale Testing Procedure for Gene-Set Analysis

The chosen JLS test statistic (e.g., \( W_F \)) for single-SNP analysis can then be used for implementing gene-based, gene-set, or pathway analysis in a direct fashion.

Assume that \( J \) SNPs have been annotated to a gene or gene-set of interest. For each SNP \( j \), the JLS-Fisher test statistic (e.g., \( W_{F,j} \)) is first obtained and then the association evidence can be aggregated across the SNPs by considering, for example, the sum statistic,\(^5\)

\[ \sum_j W_{F,j}. \]

To account for LD between SNPs, the overall association evidence can be evaluated by a phenotype-permutation approach where the empirical p value is the proportion of \( K \) permutation replicates with sum statistics more extreme than the observed value. Because this multivariate method analyzes all \( J \) SNPs simultaneously, the number of permutations need not be exceedingly large and \( K = 10,000 \) provides accurate estimates for p values in the range of 0.05. If multiple genes or gene-sets are of interest, more replicates would be required to adjust for the corresponding number of hypothesis tests.

To compare strength of association evidence between sets of variants within the same gene or across different genes, an extension of the gene-set approach can be implemented. Sum statistics are obtained as previously described for each group of variants, then calibrated by the respective number of variants. The difference between the two proportional sum statistics is the test statistic of interest,

\[ D_F = \frac{1}{J} \sum_j W_{F,j} - \frac{1}{K} \sum_k W_{F,k}, \]

where the \( j = 1, \ldots, J \) and \( k = 1, \ldots, K \) subscripts index the competing sets of variants, and the \( F \) subscript indicates that the variant-specific joint location-scale statistics are obtained by Fisher’s method (although other methods of combining such as minP could be used as well). The significance of \( D_F \) can be evaluated with the phenotype-permutation approach as described above.

In all applications, genotypes were coded additively (\( G = 0, 1 \) or 2), and for X chromosome SNPs, female and male genotypes were analyzed together and coded as \( G = 0, 1 \) or 2 and \( G = 0 \) or 2, respectively.

Application Data: HbA1c Levels in Type 1 Diabetes and Cystic Fibrosis Lung Disease

We tested the proposed JLS approach with applications to genetic association studies of HbA1c levels in type 1 diabetes and lung disease in cystic fibrosis.

The T1D application used the Diabetes Control and Complications Trial (DCCT) sample in which 667 individuals were conventionally treated and 637 intensively treated.\(^{19-21}\) In this sample, an earlier GWAS of HbA1c levels in type 1 diabetes\(^{15}\) showed that rs1358030 near SORCS1 (10q25.1 [MIM: 606283]) interacts with treatment type (conventional versus intensive) on HbA1c levels (quarterly measured values spanning 6.5 years). To demonstrate that the JLS testing framework could leverage the interaction effect without knowledge of the interacting variable, we analyzed the association of rs1358030 with the average, inverse normal transformed HbA1c value, assuming the treatment variable was not available.

The CF application involves association studies of an averaged lung function measure, forced expiratory volume in 1 s, adjusted for sex, age, height, and mortality, and normalized (SaKnorm),\(^{24}\) using data from the Canadian Cystic Fibrosis Gene Modifier Study (CGS).\(^{25}\) To reduce the duration of heterogeneous environmental exposures that were not measured, and recognizing that age can serve as a surrogate for these exposures, we\(^{19}\) previously restricted our (location-only) analysis to lung function measures from pediatric ages (\( < 18 \) years, \( n = 815 \) subjects from 753 unique families), analyzing eight SNPs in three genes, SLC9A3 (MIM: 182307), SLC6A14 (MIM: 300444), and SLCZ6A9 (MIM: 608481), previously identified as associated with meconium ileus in a hypothesis-driven GWAS (GWAS-HD).\(^{5}\) This approach involved a 46% reduction in sample size, and that the age restriction be fixed for all variants, despite the possibility that the optimal exposure (here, age) is gene specific. Here we re-analyzed the SNPs with the individual location-only and scale-only tests, as well as the joint JLS-Fisher, JLS-minP, LRT,\(^{17}\) and distribution\(^{16}\) tests, removing the age restriction and using the full CGS sample (\( n = 1,409 \) unrelated subjects). For comparison, the location-only analyses restricted to the pediatric population using different age cut-off points were also investigated.

With the full CGS sample, we further tested the hypothesis that multiple proteins present on the apical plasma membrane contribute to lung disease severity as measured by SaKnorm; the hypothesis was considered previously for meconium ileus susceptibility in CF.\(^{3}\) In total, 3,814 GWAS SNPs (MAF > 0.02) were annotated to within \( \pm 10 \) kb of 155 apical genes obtained from the Gene Ontology project.\(^{3}\) The JLS-Fisher test was first applied to each SNP, and the SNP-specific test statistic was then aggregated across all SNPs to perform the multivariate apical gene-set analysis. We then used an independent French sample (\( n = 1,232 \)) for replication. Imputation based on 1000 Genomes\(^{26,27}\) (as outlined in the Online Methods of Sun et al.) was used to assess regional association within the SLC9A3R1, SLC9A3R2, and EZR binding sites of SLC9A3.

Institutional review committees at all participating institutions for the DCCT T1D study as well as all Canadian CF clinics approved this study. Consent was also obtained for participants from France with procedural approval (CPR 2004/15) and information collection approval by CNIL (04.404). Data collection, genotyping, and quality-control procedures are reported elsewhere for the T1D\(^{15}\) and CF (Canadian and French) studies.\(^{2,25}\)
JLS Method Evaluation by Simulation

We conducted extensive simulation analyses to evaluate the performance of the proposed JLS-Fisher and JLS-minP tests for single-variant analysis and compared them with the individual location-only and scale-only tests, as well as the distribution test14 and the LRT.15 We also conducted a sensitivity analysis on the various tests, studying the impact of poorly captured genotypes that can be expected from imputed datasets. Simulation details are provided in Appendix A.

Results

Type 1 Diabetes Complications

Traditional analysis using the individual location-only test did not yield a genome-wide significant association of p < 5 × 10⁻⁸ (as previously proposed by Dudbridge and Gusnanto28) between rs1358030 and HbA1c (p = 2.3 × 10⁻⁷), whereas the JLS-Fisher method did (p = 4.9 × 10⁻⁸). Genome-wide significance was also not achieved by the scale-only test (p = 0.01), the distribution test (p = 1.7 × 10⁻⁷, estimated from 10⁸ permutation replicates), the LRT (p = 2.0 × 10⁻⁷), or the JLS-minP test (p = 4.6 × 10⁻⁷).

Knowing the treatment information, Paterson et al.18 was able to explicitly model this interactor and identified rs1358030 as genome-wide significantly associated with HbA1c levels and interacting with treatment type (p = 3.8 × 10⁻¹⁰ and p = 0.013 for the main and interaction effect, respectively).

Cystic Fibrosis Lung Disease: Single-Variant Association Analysis

We first noted in Table 1 that all methods (the joint and individual tests using the full sample or various subsamples) consistently highlighted that variants in both SLC9A3 and SLC6A14 are associated with the lung phenotype, but not those in SLC26A9. For the individual location-only test, although the age cut-off of <18 years appears to be ideal for rs17563161 in SLC9A3 resulting in the smallest p value (p = 3.3 × 10⁻⁵) even with the smaller sample size (n = 753), a different cut-off (<20 years, n = 830) yields the most significant location-only test result for rs3788766 in SLC6A14 (p = 0.0002). This illustrates the challenge of specifying and modeling interacting exposures in the context of multiple SNPs and genes of interest.

With the entire sample of CF-affected individuals (n = 1,409) encompassing a wider age range and thus greater variability of environmental exposures, we indeed observed evidence for scale differences in the lung phenotype (Levene’s test p = 0.01–0.08) for variants in SLC9A3 and SLC6A14, revealing the possibility of GxG or GxE interaction. (Table S3 provides evidence for SNP-by-age interaction effects from regression models directly incorporating age.) Compared with the distribution and LRT joint tests, the JLS-Fisher test consistently provides the smallest p values for variants in SLC9A3 and SLC6A14.

Cystic Fibrosis Lung Disease: Gene-Set Association Analysis

We observed evidence of an association between the apical gene set and SakNorm (JLS-Fisher permutation p = 0.0099; Figures 1A and 1B and Table 2). Comparison with the individual location-only and scale-only tests showed that this association does not reach statistical significance using the conventional gene-set location test (regression permutation p = 0.0876; Figures 1C and 1D) whereas the variance component contributed a significant result (Levene’s permutation p = 0.0222; Figures 1E and 1F).

Examination of the SNP- and gene-specific JLS-Fisher p values and rankings showed SLC9A3 to be the top ranked (Table S4), with three of the top four ranked SNPs being annotated to SLC9A3. This provided consistent support for the relationship between SLC9A3 and lung function as previously reported (Table 1).19,29 In the independent French sample (n = 1,232), we observed replication of the apical hypothesis via the JLS-Fisher test (permutation p = 0.0180; Figures 2A and 2B and Table 2). Again, the standard location testing approach, by itself, was not powerful enough to detect the association (regression permutation p = 0.2058; Figures 2C and 2D), and the added contribution from the scale-testing component was beneficial (Levene’s permutation p = 0.0077; Figures 2E and 2F). After excluding all ten genotyped SNPs annotated to SLC9A3, the apical gene-set test in both the CGS and French samples remained significant (Table 2), suggesting that the JLS-Fisher method identified multiple additional associations within the gene set, beyond the known SLC9A3 contribution.

Three of the five top-ranked genes in the apical gene set were SLC9A3, SLC9A3R2 (MIM: 606553), and EZR (MIM: 123900), with ten genes in total displaying JLS-Fisher p < 0.05 (Table S4). Of these top five, protein product interactions are known between (1) SLC9A3 (also known as NHE3) and SLC9A3R2 (“SLC9A3 regulator 2,” also known as E3KARP or NHERF2), (2) SLC9A3 and EZR, and (3) SLC9A3R2 and EZR. SLC9A3R1, or “SLC9A3 regulator 1” (also known as EBPS0 or NHERF1), is recognized as a paralog of SLC9A3R2 with comparable binding sites to SLC9A3 from the Golgi to the cell periphery, with eventual “hand-off” to the SLC9A3 regulators. The SLC9A3 regulators help to maintain SLC9A3 at the apical membrane and facilitate its dynamic shuffling with endosomes and internal vesicles.30

Notably, variants within the SLC9A3 complex component genes are significantly associated as a set in the CGS sample using the JLS-Fisher or regression tests (p < 0.0001 and p < 0.0001, respectively; Table 2). The French sample provided replicated support for the SLC9A3
Table 1. Candidate-Gene Study of Cystic Fibrosis Lung Disease Severity as Measured by SaKnorm in the Canadian Gene Modifier Study Sample

<table>
<thead>
<tr>
<th>Chr</th>
<th>Gene</th>
<th>SNP</th>
<th>BP*</th>
<th>MAFb</th>
<th>Location/Regression</th>
<th>Scale/Levene</th>
<th>Joint Tests</th>
<th>CGS Pediatric Subsample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;16 years (n = 653)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;18 yearsd (n = 753)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;20 years (n = 830)</td>
</tr>
<tr>
<td>1</td>
<td>SLC26A9</td>
<td>rs7512462</td>
<td>204,166,218</td>
<td>0.41</td>
<td>0.30</td>
<td>0.58</td>
<td>0.48</td>
<td>0.52</td>
</tr>
<tr>
<td>1</td>
<td>SLC26A9</td>
<td>rs4077468</td>
<td>204,181,380</td>
<td>0.42</td>
<td>0.53</td>
<td>0.61</td>
<td>0.69</td>
<td>0.78</td>
</tr>
<tr>
<td>1</td>
<td>SLC26A9</td>
<td>rs12047830</td>
<td>204,183,322</td>
<td>0.49</td>
<td>0.55</td>
<td>0.15</td>
<td>0.29</td>
<td>0.28</td>
</tr>
<tr>
<td>1</td>
<td>SLC26A9</td>
<td>rs7419153</td>
<td>204,183,932</td>
<td>0.37</td>
<td>0.50</td>
<td>0.06</td>
<td>0.14</td>
<td>0.12</td>
</tr>
<tr>
<td>S*</td>
<td>SLC9A3</td>
<td>rs17563161</td>
<td>550,624</td>
<td>0.26</td>
<td>0.0004</td>
<td>0.02</td>
<td>0.0001</td>
<td>0.0008</td>
</tr>
<tr>
<td>X</td>
<td>SLC6A14</td>
<td>rs12839137</td>
<td>115,479,578</td>
<td>0.24</td>
<td>0.02</td>
<td>0.08</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>X*</td>
<td>SLC6A14</td>
<td>rs5905283</td>
<td>115,479,909</td>
<td>0.49</td>
<td>0.009</td>
<td>0.07</td>
<td>0.005</td>
<td>0.02</td>
</tr>
<tr>
<td>X*</td>
<td>SLC6A14</td>
<td>rs3788766</td>
<td>115,480,867</td>
<td>0.40</td>
<td>0.001</td>
<td>0.01</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

SaKnorm (normalized, averaged sex-, age-, height-, mortality-adjusted forced expiratory volume in 1 s; FEV1) was used as the lung function measure. Analysis was performed on eight SNPs in three candidate genes previously identified as associated with meconium ileus and subsequently studied for association with SaKnorm in a pediatric subsample. Individual location-only and scale-only tests, joint location-scale Fisher (JLS-Fisher), joint location-scale minimum-P (JLS-minP), the likelihood ratio test (LRT) of Cao et al., and the distribution test of Aschard et al. were performed in the full sample of 1,409 unrelated individuals. For comparison, the location-only test was performed in three CGS pediatric subsamples using different age cutoffs of 16, 18, and 20 years.

ahg18 assembly (March 2006; NCBI36).
bMAF is similar across the pediatric subsets (see Table S8).
cDistribution test p values were estimated from 10^5 permutation replicates.
dPediatric subset cutoff (<18 years) used in Li et al.; subset here includes only unrelated subjects (n = 753).
eSignificant variants (p < 0.05 corrected for multiple testing) reported in Li et al.
complex using the JLS-Fisher test \((p = 0.0415)\), although in this latter sample, \(SLC9A3R1\) is more highly ranked than \(SLC9A3R2\) and \(EZR\). Removal of the four-gene set, \(SLC9A3\), \(SLC9A3R1/2\), and \(EZR\), and the re-testing of the apical gene set with the remaining 151 genes suggests that association(s) beyond the \(SLC9A3\) regulator complexes also exist (JLS-Fisher \(p = 0.0329\) and 0.0201 in the CGS and French samples, respectively).

\(SLC9A3\) contains multi-membrane spanning motifs in its amino portion to facilitate transport function, with a large cytosolic carboxyl portion that affords regulation with binding sites for multiple interactors. Although there is significant LD that extends throughout the gene (Figure 3), there was evidence suggesting greater association for variants as a group from the region corresponding to the regulatory portion, compared to the transporting portion (regression, Levene’s, and JLS-Fisher permutation \(p = 0.0391, 0.312,\) and 0.1254, respectively). Amino acids 586–660 of \(SLC9A3\) bind the \(SLC9A3\) regulators where an exonic nucleotide variant (rs2230437) is associated with lung function (JLS-Fisher \(p = 7.6 \times 10^{-6}\)). This synonymous change is in high LD \((r^2 > 0.8)\) with four other variants, all with similar association evidence (rs11743825, rs41282625, 5:475625:GC_G, and rs11745923 with JLS-Fisher \(p = 1.4 \times 10^{-6}, 2.2 \times 10^{-6}, 2.3 \times 10^{-6},\) and 2.7 \(\times 10^{-6}\), respectively), and all in noncoding positions. There are also no associated coding variants in the major \(EZR\) binding site from amino acids 519–595. Similarly, there are no coding variants in the respective PDZ domains of the \(SLC9A3\)
To study the type 1 error at the genome-wide level, we conducted association analysis of all 866,995 GWAS SNPs available in the T1D HbA1c example with permuted phenotype. Results showed that all joint testing methods provide correct type 1 error control in a finite sample (Figures S1A, S1D, and S1G, respectively, for the JLS-Fisher, JLS-minP, and LRT methods; the distribution test by design has the correct type 1 error control). However, when stratified by minor allele frequency (MAF), we observed that, whereas the JLS-Fisher and JLS-minP tests are slightly conservative for SNPs with MAF < 0.1 (Figures S1C and S1F), the LRT method has increased type 1 error rate (Figure S1I) and would probably not be appropriate when the size of the genotype group is limited (n < 20). Permutation analysis of all of the 565,884 GWAS SNPs available in the CF application example provided type 1 error results similar to those of the T1D HbA1c example (Figure S2).

For the sensitivity analysis, we observed that the individual location and scale tests and the JLS-Fisher and JLS-minP tests were not affected by poorly captured genotypes; they maintained correct type 1 error level irrespective of the proportion of incorrectly assigned genotypes (Table S5).

### Discussion

Interactions between genes in gene-set and pathway analyses are assumed to exist, but they are, in general, not accounted for in current genetic analysis methods. Here we proposed a joint location-scale (mean-variance) testing method to account for these relationships as well as main effects, while avoiding explicit specification of the contributing interactions. We used this method to identify the apical membrane constituent gene set as contributing to CF lung disease and replicated this association in an independent sample of similarly ascertained CF-affected subjects. Further, we illustrated how the gene-set association notably benefits from consideration of phenotypic variance differences that might have been induced by GxG and/or GxE relationships. This apical gene-set association was not initially detected via the classical location-only testing as used in a typical GWAS and provides insight into interacting genes and exposure variables affecting lung function.

SNPs in SLC9A3 have been reported to associate with CF lung disease in two candidate gene studies. However, a broader role for the SLC9A3/SLC9A3R2/EZR complex was not recognized. The JLS-Fisher test ranked these players as three of the top five contributing apical plasma membrane genes. Variation at SLC9A3R2 or EZR loci and possibly also SLC9A3R1 given its paralogous function (and supporting evidence in the French sample) could be expected to influence CF lung disease severity, presumably by affecting changes in net SLC9A3 activity. For example, modest changes in levels of either SLC9A3R2 or EZR could alter functional apical membrane SLC9A3 levels, just as regulators or in the FERM domain of EZR that bind SLC9A3. The imputation analysis would have captured the major variation in this gene. Collectively, because the constituents or their direct physical interactions would not appear to be affected, disturbed expression with altered stoichiometry of components or dynamic positioning of the SLC9A3 complex might be contributing to lung disease severity.

### Simulation

Results showed that, under the normal model and in most scenarios considered, the LRT and JLS-Fisher methods had similar power and were more powerful than the JLS-minP and distribution tests (Figures 4 and S3) and the individual location-only and scale-only tests (Figures S4 and S5). Investigation of type 1 error (100,000 replicates) under the normal model showed that all individual and joint tests considered here maintained the nominal error rates at the 0.05, 0.005, and 0.0005 levels when MAF was at least 0.1 and there were ≥ 20 individuals within each genotype group (Tables S1 and S2). However, as MAF or the smallest genotype group size decreased, the LRT method demonstrated inflated type 1 error (Tables S1 and S2). Departure from normality also resulted in LRT having inflated type 1 error, as previously discussed in Cao et al.

### Table 2. Apical Gene-Set and SLC9A3 Complex Association Analysis with CF Lung Disease Severity as Measured by SaKnorm in the Canadian Gene Modifier Study Discovery Sample and in the French Replication Sample

<table>
<thead>
<tr>
<th>Test</th>
<th>Location-only Test</th>
<th>Scale-only Test</th>
<th>JLS-Fisher Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Apical Gene-Set (3,814 SNPs)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canadian Discovery Sample</td>
<td>0.0876</td>
<td>0.0222</td>
<td>0.0099</td>
</tr>
<tr>
<td>French replication sample</td>
<td>0.2058</td>
<td>0.0077</td>
<td>0.0180</td>
</tr>
<tr>
<td><strong>SLC9A3 Gene-Based (10 SNPs)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canadian Discovery Sample</td>
<td>&lt;0.0001</td>
<td>0.0238</td>
<td>0.0001</td>
</tr>
<tr>
<td>French replication sample</td>
<td>0.0953</td>
<td>0.0415</td>
<td>0.0286</td>
</tr>
<tr>
<td><strong>Apical Gene-Set Excluding SLC9A3 (3,804 SNPs)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canadian Discovery Sample</td>
<td>0.1319</td>
<td>0.0248</td>
<td>0.0161</td>
</tr>
<tr>
<td>French replication sample</td>
<td>0.2139</td>
<td>0.0076</td>
<td>0.0182</td>
</tr>
<tr>
<td><strong>SLC9A3 Complex: SLC9A3, SLC9A3R1, SLC9A3R2, and EZR (38 SNPs)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canadian Discovery Sample</td>
<td>&lt;0.0001</td>
<td>0.0016</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>French replication sample</td>
<td>0.2390</td>
<td>0.0245</td>
<td>0.0415</td>
</tr>
<tr>
<td><strong>Apical Gene-Set Excluding the SLC9A3 Complex (3,776 SNPs)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canadian Discovery Sample</td>
<td>0.178</td>
<td>0.0329</td>
<td>0.0329</td>
</tr>
<tr>
<td>French replication sample</td>
<td>0.2173</td>
<td>0.0099</td>
<td>0.0201</td>
</tr>
</tbody>
</table>

Individual location-only and scale-only tests and the joint location-scale fisher (JLS-Fisher) test were performed in the full CGS (n = 1,409) and French (n = 1,232) samples using unrelated individuals (see Material and Methods and Results). p values reported were estimated from 10,000 permutation replicates.
SLC9A3 locus variants might alter the amount of SLC9A3 transcripts. It should also be noted that the SLC9A3 regulators, with their proximal localization at the apical membrane, could contribute a tethering function for a number of other apical membrane proteins, including CFTR.31 Given loss of function in CF, SLC9A3 regulators might not be mediating their effects via CFTR; however, it is possible that some of the remaining significant apical membrane components, beyond SLC9A3, utilize the EZR and/or the SLC9A3 regulator scaffold.32,33

We demonstrated the potential advantage of using the JLS framework over the conventional location-only or scale-only approach in practice. As proof of principle, we used examples from an association study of a complex trait (glycemic control in type 1 diabetes) and a modifier study of a Mendelian disease (lung function in CF). Both of these datasets provided examples of a genetic effect on a quantitative phenotype that varied based on another variable (i.e., treatment regimen and age, respectively). Our results showed that if the interacting exposure(s) were not known or available, the JLS method could have improved power over conventional approaches in detecting the association.

Recent literature cautions the investigation and interpretation of significant differences in phenotypic variability across genotypes as well as observed GxE and GxG interaction effects.17,34–37 First, variance heterogeneity across genotypes can result from a mean-variance relationship induced by an inappropriate choice of measurement scale for the phenotype under investigation. Second, dependence between a causal locus and an exposure (G or E) can create synthetic interactions between the exposure and other tagged SNPs in incomplete LD with the causal locus.36,37 Consequently, we might observe phenotypic variance heterogeneity across genotypes of a tagged SNP.37 In these scenarios, however, because there is a true underlying genotype-phenotype relationship (main effect of G), it does not matter whether we detect this association through a location or a scale test, or both; the goal of our approach is to leverage complex genetic architecture to efficiently identify the associated G for follow-up investigation of causal interactors.

Challenges in understanding and being able to replicate original findings, particularly in the GxG, GxE, gene-set,
and pathway analysis setting, is well documented. Given the potential heterogeneity between the Canadian discovery sample and the French replication sample (e.g., clinical treatment, climate, and air quality), interaction models could differ between the two samples. Therefore, even if the exposure variables were precisely measured, a significant multivariate interaction effect found in the discovery sample may not be observed in the replication sample. The proposed method allows different exposure variables between genes and between samples, as long as the corresponding underlying phenotype-genotype association mechanism results in phenotypic mean and/or variance differences between genotypes of the variants of interest. A step-by-step guide for application of the JLS framework, and a discussion of its advantages compared to alternative approaches, is included as additional considerations in Appendix A.

In conclusion, we have provided a robust joint location-scale (JLS) testing framework for the detection of single-variant, gene-set, or pathway associations involving either main or interaction effects, or both. This methodology applies to analysis of a quantitative trait, regardless of the biological interpretation of the chosen scale of phenotypic measurement. Application of the JLS approach identified the SLC9A3 regulatory complex as an important contributor to CF lung disease, which was completely missed in previous analyses of the same data. The JLS association testing method should be considered for future gene-set, pathway, and whole-genome association scans. Further, JLS should be employed to re-examine datasets previously analyzed via the conventional location-only or scale-only testing approaches, for complex traits or complex secondary phenotypes of Mendelian diseases. The method will help researchers pinpoint susceptibility loci for additional analysis toward the understanding of the genetic architecture of a complex trait.

Appendix A

Implementation: Steps for Applying the JLS Testing Framework

Note that the following implementation is appropriate for testing association of a phenotype with a genotyped genetic variant (e.g., SNP) or an imputed variant with a “hard call” (i.e., assign individual to the genotype with the highest posterior imputation probability; high genotype uncertainty does not affect type 1 error but will decrease power), using a sample of unrelated subjects.

1. Check the phenotype of interest for fit to a normal distribution. If required, adjust the phenotype using a suitable transformation, e.g., inverse normal transform. If the researcher proceeds using a non-normal
phenotype, only the permutation (resampling-based) p value analysis will be valid (see step 4b).

2. Choose the individual location and scale tests based on the distribution of phenotype (normal or non-normal) or preference (for example, parametric or non-parametric versions of each test). In the present paper, our phenotypes were normally distributed (after transformation) and we chose linear regression and Levene’s test for the location and scale tests, respectively.

3. Choose a JLS testing method of combining information from the individual location and scale tests and calculate the JLS test statistic. We acknowledge that there is no “most powerful” method for all situations in practice. Based on our experience, we recommend the use of Fisher’s method (JLS-Fisher) of combining the association evidence:

$$W_F = -2 \left( \log(p_L) + \log(p_S) \right),$$

where \(p_L\) and \(p_S\) are the individual location and scale test p values, respectively.

4. Chose the p value estimation method for the JLS statistic.

(a) Based on the approximate asymptotic distribution of the JLS test statistic: For the JLS-Fisher example, \(W_F\) is distributed as a \(\chi^2_f\) random variable, if the chosen individual location and scale tests are independent of each other under the null hypothesis. This assumption is correct if the trait is normally distributed and if the location-only test statistic is a function of the complete sufficient statistic (e.g., linear regression t-statistic, ANOVA F-statistic) and the distribution of the scale-only test statistic does not depend on the model parameters (e.g., Levene’s test or the F-test for equality of variances).

(b) Based on resampling methods such as permutation:
- Calculate the observed JLS test statistic, e.g., \(W_F\)
- Choose the number of permutation replicates, \(K\), based on the desired p value accuracy.
- Permute the phenotype independently \(K\) times (not valid if subjects are correlated with each other), and for each replicate \(k\), recalculate the JLS test statistic, \(W_F^k\), \(k = 1, ..., K\).
- Obtain the permutation p value as [the number of \(W_F^k > W_F\)]/\(K\).

Figure 4. Power Comparison under Simulation Model 1
Four different testing methods are examined: the proposed JLS-Fisher (red) and JLS-minP (purple) methods, the distribution test (blue) of Aschard et al., and the LRT (black) of Cao et al. Phenotype values for 2,000 independent subjects were simulated under Model 1, \(E[Y] = \beta_G G + \beta_E E_1 + \beta_{GE} G \cdot E_1\), where the MAF of \(G\) was 0.3 and the exposure variable \(E_1\) was simulated as a Bernoulli variable with frequency = 0.3. The effect of the exposure \(E_1\), \(\beta_E\), was fixed at 0.3 while the other effects varied.

(A–C) Results when the main genetic effect \(\beta_G\) and the interaction effect \(\beta_{GE}\) are in the same direction.

(D–F) Results when \(\beta_G\) and \(\beta_{GE}\) are in opposing direction.

Power was calculated at the $5 \times 10^{-8}$ level based on 500 replicates. For other simulation details, see Appendix A. Additional power results are in Figures S3–S5.
Simulation Models
The following three models, as previously considered by Aschard et al., were used to simulate the data:

Model 1: \[ E[Y] = \beta_0G + \beta_1E_1 + \beta_{GE1}G \cdot E_1 \]
Model 2: \[ E[Y] = \beta_1E_1 + \beta_{GE2}E_2 + \beta_{GE1}G \cdot E_1 + \beta_{GE2}G \cdot E_2 \]
Model 3: \[ E[Y] = \beta_{GE1}G \cdot E_1 \]

For all three models, the observed genetic variant \((G)\) was coded additively with minor allele frequency (MAF) of 0.3. \(Y\) was simulated from models with varying effects \((\beta)\) and residual variation \((\varepsilon)\) following a standard normal distribution (mean = 0, standard deviation = 1).

Model 1 is analogous to Equation 1 where \(Y\) depended on the main effects of both \(G\) and \(E_1\) and an interaction effect between \(G\) and \(E_2\). The unobserved exposure variable \(E_1\) was binary with frequency 0.3. The main genetic effect \(\beta_{GE1}\) was always positive and less than 1, whereas \(\beta_{GE2}\) was always positive and less than 1, whereas \(\beta_{GE1}\) and \(\beta_{GE2}\) were negatively correlated.

For Model 2, \(Y\) was a function of main effects due to two unobserved exposures \((E_1\) and \(E_2)\) both binary with frequency 0.3) and interaction effects between the exposures and \(G\). \(\beta_{GE1}\) was always positive and less than 1, whereas \(\beta_{GE2}\) was fixed at 0.3, whereas \(\beta_{GE2}\) was fixed at 0.3 when \(\beta_{GE1}\) was positive and \(-0.3\) when \(\beta_{GE1}\) was negative.

For Model 3, \(Y\) depended only on the interaction between \(G\) and \(E_1\). For this model, the interaction effect \(\beta_{GE1}\) and exposure frequency were chosen such that the marginal effect of \(G\) was fixed at 10% of the trait standard deviation.

In all cases, the working association model corresponded to Equation 2 because information on \(E_1\) and \(E_2\) was assumed to be unavailable.

To assess the type 1 error level of the joint location-scale methods at 0.05, 0.005, and 0.0005 levels, we simulated 100,000 replicate samples of \(n = 2,000\) subjects each from the null model with no genetic association (i.e., \(\beta_0 = 0\) and \(\beta_{GE} = 0\)). The unobserved genotype data using a Dirichlet distribution not depending on \(\beta_{GE1}\) and interaction effects between the exposures and \(G\). \(\beta_{GE1}\) and \(\beta_{GE2}\) were both binary with frequency 0.3 when \(\beta_{GE1}\) was positive and \(-0.3\) when \(\beta_{GE1}\) was negative.

For power evaluation, as in Aschard et al., the results presented were from Aschard et al., kindly provided by Drs. Aschard and Kraft.

Lemma 1: Independence of Location-only and Scale-only Test Statistics under the Null Hypothesis for Normally Distributed Traits
Let \(T_{Location} = \frac{\tilde{\beta}_1/(S/\sqrt{S_{xx}})}{b}\) be our location-only test statistic, testing the linear effect of \(x\) on \(Y\) in a sample of size \(n\), where \(S^2 = ((n - 2))\sum(y - \tilde{\beta}_0 - \tilde{\beta}_1x_i)^2\), \(\tilde{\beta}_0 = \bar{Y} - \tilde{\beta}_1\bar{x}, \quad \tilde{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \quad S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{Y}), \quad S_{xx} = \sum(x_i - \bar{x})^2\). (\(x = \mathbf{G}\) in Equation 2), and let \(T_{Scale}\) be our scale-only test statistic, here defined as Levene’s test statistic for equality of variances. Note that the proof of independence holds regardless of the version of Levene’s test statistic chosen, provided that the approximation to the \(F\) distribution (or some other distribution not depending on \(\theta\)) is justifiable. Similar
In the most extreme scenario of simulation, i.e., Model 3, when there were no main $G$ or $E$ effects, the interaction effect was large ($\hat{\beta}_{GE} = 2$), and the unobserved exposure was rare (prob($E_1 = 1) = 0.05$), the distribution test was observed to be more powerful than the JLS-Fisher test (0.916 versus 0.406) (row 1 of Table S6). This is because the resulting phenotype distributions across genotypes differed in shape and their differences were not well captured by only mean and variance parameters. (The JLS-Fisher test could, in theory, be extended to include the skewness parameter or higher moments of such a distribution. The gain in power for this particular setting, however, would be associated with power loss for other approximately normally distributed traits.) In this scenario, when the phenotype deviated from normality, the LRT method appeared to be most powerful (power = 0.996). However, further analysis using permutation estimation of $p$ values showed that power of the LRT under asymptotic analysis was greatly inflated, whereas the proposed JLS methods were robust (Table S7).

The alternative distribution and LRT joint testing methods were proposed for analysis of single SNPs. In principle, they can be extended to gene-set or pathway analysis, but multiple issues arise in implementation. The distribution test statistic depends on the size of the genotype groups for each variant, so it is not clear what the best strategy is to combine the statistics across SNPs with different MAFs. The LRT method is sensitive to the normality assumption of the phenotype distribution and to small group size of the genotype distribution (Tables S1 and S2).

The proposed multivariate JLS testing method is also extremely relevant for single-variant association analysis. In GWAS or Next-Generation Sequencing (NGS) settings where millions or tens of millions of SNPs are investigated, rapid screening of the whole genome to correctly prioritize SNPs for further examination demands methods that are powerful, yet computationally efficient. The proposed JLS testing method is robust and easy-to-implement, suitable for large-scale whole-genome scans, and can reveal individual genetic variants with main and/or interaction effects without the need to explicitly specify the interacting genetic and/or environmental variables. Compared with the distribution test and LRT alternatives, our method combines both simplicity of implementation and robustness to small size of the rare homozygous genotype group.

Our simulation analyses also demonstrated that the location (regression) and scale (Levene’s) $t$-statistic, and consequently our JLS-Fisher and JLS-minP tests, are robust to poorly captured genotype data. These findings agree with the results of Kutalik et al. where minimal to no bias in statements of independence with analogous proofs can be obtained for other choices of location test statistics such as the analysis of variance (ANOVA) $F$-statistic.

**Additional Considerations for the JLS Framework**

In the most extreme scenario of simulation, i.e., Model 3, when there were no main $G$ or $E$ effects, the interaction effect was large ($\hat{\beta}_{GE} = 2$), and the unobserved exposure was rare (prob($E_1 = 1) = 0.05$), the distribution test was observed to be more powerful than the JLS-Fisher test (0.916 versus 0.406) (row 1 of Table S6). This is because the resulting phenotype distributions across genotypes differed in shape and their differences were not well captured by only mean and variance parameters. (The JLS-Fisher test could, in theory, be extended to include the skewness parameter or higher moments of such a distribution. The gain in power for this particular setting, however, would be associated with power loss for other approximately normally distributed traits.) In this scenario, when the phenotype deviated from normality, the LRT method appeared to be most powerful (power = 0.996). However, further analysis using permutation estimation of $p$ values showed that power of the LRT under asymptotic analysis was greatly inflated, whereas the proposed JLS methods were robust (Table S7).

The alternative distribution and LRT joint testing methods were proposed for analysis of single SNPs. In principle, they can be extended to gene-set or pathway analysis, but multiple issues arise in implementation. The distribution test statistic depends on the size of the genotype groups for each variant, so it is not clear what the best strategy is to combine the statistics across SNPs with different MAFs. The LRT method is sensitive to the normality assumption of the phenotype distribution and to small group size of the genotype distribution (Tables S1 and S2).

The proposed multivariate JLS testing method is also extremely relevant for single-variant association analysis. In GWAS or Next-Generation Sequencing (NGS) settings where millions or tens of millions of SNPs are investigated, rapid screening of the whole genome to correctly prioritize SNPs for further examination demands methods that are powerful, yet computationally efficient. The proposed JLS testing method is robust and easy-to-implement, suitable for large-scale whole-genome scans, and can reveal individual genetic variants with main and/or interaction effects without the need to explicitly specify the interacting genetic and/or environmental variables. Compared with the distribution test and LRT alternatives, our method combines both simplicity of implementation and robustness to small size of the rare homozygous genotype group.

Our simulation analyses also demonstrated that the location (regression) and scale (Levene’s) $t$-statistic, and consequently our JLS-Fisher and JLS-minP tests, are robust to poorly captured genotype data. These findings agree with the results of Kutalik et al. where minimal to no bias in false positive rates and type 1 error was found for location-only association testing of quantitative phenotypes with uncertain genotypes, specifically when imputed genotype probabilities were converted to most-likely genotype categories and analyzed without explicitly accounting for the uncertainty.

Violations of the normal data assumption can affect the type 1 error level of the proposed JLS test method as it does the LRT approach (increased type 1 error rate is more severe with LRT). This is largely due to the assumption of independence between the individual location-only and scale-only tests, which is required for the $\chi^2$ approximation of the JLS-Fisher statistic. To circumvent this issue, investigators might choose to rely on a permutation distribution when estimating $p$ values. However, this aspect can impose computation challenges at the genome-wide level. An alternative approach is to model the dependency between the individual location and scale test statistics and obtain an adjusted distribution for the JLS statistic. These adjusted distributions have been explored elsewhere under complete specification of the dependency structure between the test statistics being combined. However, it remains to be explored how to model the dependency between the individual location and scale association test statistics at different loci across the genome.

In consideration of the choices available for the individual location and scale tests and the methods of combining information from these individual components for each variant, we recognize that there is no single, most powerful method for all circumstances. We showed that for normally distributed traits, the JLS-Fisher test statistic $W_F$ follows a $\chi^2$ distribution under the null hypothesis, as long as the location test statistic is a function of the complete sufficient statistic (e.g., linear regression $t$-statistic, ANOVA $F$-statistic) and the distribution of the scale test statistic does not depend on the model parameters (e.g., Levene’s test or the $F$-test for equality of variances). In practice, when the normality assumption is violated or other JLS tests are preferred, permutation-based $p$ value evaluation can be used with increased computational cost.

The proposed framework can be easily extended for meta-analysis, where the sample- or study-specific association test statistics or $p$ values to be combined across samples are obtained from the JLS testing application instead of the typical location-only testing method. The sample-specific choices of the individual location and scale tests need not be identical across different studies, as long as $p$ values of the JLS tests are valid within each study. However, choice of optimal weighting factors assigned to individual samples requires further investigation. Analyzing imputed SNPs (explicitly incorporating the genotype probabilities) or correlated subjects (i.e., pedigree family data) using the proposed JLS framework would require development of appropriate scale-only testing methods; location-only methods for these more complex settings are already available.
Supplemental Data

Supplemental Data include five figures and eight tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2015.05.015.

Acknowledgments

The authors are grateful to the subjects in the CGS, French CF, and DCCT studies for their participation. The authors also thank Drs. Hughes Aschard and Peter Kraft for providing their simulation results; the CF Gene Modifier Consortium and Drs. Michael Knowles, Garry Cutting, and Mitchell Drumh for helpful discussions; and the Canadian and US CF Foundations for their generous support of the genotyping. This data resource was also supported in part by Genome Canada, through the Ontario Genomics Institute per research agreement 2004-OGI-3-05 (to P.R.D.), with the Ontario Research Fund, Research Excellence Program. This work was funded by the Canadian Institutes of Health Research (CIHR; 201309MOP-310732-G-CEAA-117978 to L.S. and MOP-258916 to L.J.S.); the Natural Sciences and Engineering Research Council of Canada (NSERC; 250053-2013 to L.S. and 371399-2009 to L.J.S.); Cystic Fibrosis Canada to L.S. and MOP-258916 to L.J.S.); the Natural Sciences and Health Research (CIHR; 201309MOP-310732-G-CEAA-117978 Program. This work was funded by the Canadian Institutes of P.R.D.), with the Ontario Research Fund, Research Excellence Genomics Institute per research agreement 2004-OGI-3-05 (to L.J.S.); Training grant GET-101831; Université Pierre et Marie Curie Paris; Agence Nationale de la Recherche (R09186DS to H.C.), Direction Générale de la Santé; Association et Marie Curie Paris; Agence Nationale de la Recherche (R09186DS to H.C.), Direction Générale de la Santé; Association Vaincre la Mucoviscidose, Chancellerie des Universités (Legs Poix); Association Agir Informer Contre la Mucoviscidose; and Groupement d’Intérêt Scientifique (GIS)-Institut des Maladies Rares. A.D.P held a Canada Research Chair in the Genetics of Complex Diseases. D.S. is a trainee of CIHR STAGE (Strategic Training for Advanced Genetic Epidemiology) program at the University of Toronto.

Received: February 28, 2015
Accepted: May 26, 2015
Published: July 2, 2015

Web Resources

The URLs for data presented herein are as follows:


References


