Genetics: Early Online, published on September 11, 2017 as 10.1534/genetics.117.300270

A powerful framework for integrating eQTL and GWAS summary data

Zhiyuan Xu[1], Chong Wu[1], Peng Wei[2], Wei Pan[1,*]

[1]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN 55455, USA

[2]Department of Biostatistics, University of Texas MD Anderson Cancer Research Center, Houston, TX 77030, USA

* Corresponding author's email: weip@biostat.umn.edu

April 28, 2017; revised August 2, 2017

Short title: Integrating eQTL and GWAS data

Copyright 2017.

Abstract

Two new gene-based association analysis methods, called *PrediXcan* and *TWAS* for GWAS individual-level and summary data respectively, were recently proposed to integrate GWAS with eQTL data, alleviating two common problems in GWAS by boosting statistical power and facilitating biological interpretation of GWAS discoveries. Based on a novel reformulation of PrediXcan and TWAS, we propose a more powerful gene-based association test to integrate single set or multiple sets of eQTL data with GWAS individual-level data or summary statistics. The proposed test was applied to several GWAS datasets, including two lipid summary association datasets based on $\sim 100,000$ and $\sim 189,000$ samples respectively, and uncovered more known or novel trait-associated genes, showcasing much improved performance of our proposed method. The software implementing the proposed method is freely available as an R package.

Key words: aSPU test, statistical power, Sum test, transcriptome-wide association study (TWAS), weighted association testing.

Introduction

In spite of many successes, genome-wide association studies (GWAS) face two major challenges. The first is its limited statistical power even with tens to hundreds of thousands of individuals in a typical GWAS or mega-GWAS, thus missing many associated genetic variants, mostly single nucleotide polymorphisms (SNPs), due to the polygenic effects and small effect sizes. The second is that even for those few identified SNPs, since they often do not reside in protein-coding regions, it is difficult to interpret their function and thus biological mechanisms underlying complex traits. A new gene-based association test called *PrediXcan* was recently proposed to integrate GWAS individual-level data with an eQTL dataset, alleviating the above two problems in boosting statistical power of GWAS and facilitating biological interpretation of GWAS discoveries (Gamazon et al 2015). It was extended to GWAS summary association data (Torres et al 2017). A similar approach, called transcriptome-wide association study (TWAS), was proposed by another group for GWAS individual-level and summary data for one or more eQTL datasets (Gusev et al 2016). They are motivated by the key fact that many genetic variants influence complex traits through transcriptional regulation (He et al 2013). Focusing on the genetic component of expression excludes environmental factors influencing gene expression and complex traits, thus can increase statistical power. In addition, compared to standard GWAS, treating genes as analysis units reduces the number and thus burden of multiple tests, again leading to improved power. By applications to common diseases like T2D and complex traits like BMI, lipids and height, the authors have convincingly shown the power of integrating GWAS and eQTL data, gaining biological insights into complex traits. There are more follow-up studies applying TWAS to other diseases. For example, Gusev et al (2017) identified some new genes associated with schizophrenia; interestingly, they also confirmed a previous observation that, contrary to the usual GWAS practice, the nearest gene to a GWAS hit often is not the most likely susceptibility gene, highlighting the critical role of incorporating gene expression to unravel disease mechanisms that may not be achieved by GWAS alone. The current standard

and popular view is that PrediXcan and TWAS work because of their predicting or imputing *cis* genetic component of expression for a larger set of individuals in GWAS, facilitating the following expression-trait association testing. Based on this view, some new methods have been proposed to improve over TWAS by addressing some existing weaknesses in gene expression prediction (Bhutani et al 2017; Park et al 2017). In spite of its intuition and usefulness, the current view on PrediXcan and TWAS may not have told the whole story. Here we offer some new insights into PrediXcan and TWAS with a novel reformulation on their underlying association testing. Our key observation is that both PrediXcan and TWAS are based on the same weighted association test; the weights on a set of SNPs in a gene are the *cis*-effects of the SNPs on the gene's expression level (derived from an eQTL dataset). In other words, PrediXcan and TWAS put a higher weight on an SNP (eSNP) that is more strongly associated with the gene's expression level, in agreement with empirical evidence that eSNPs are more likely to be associated with complex traits and diseases (Nicolae et al 2010). This new formulation also points out the connection to existing weighted association analysis (Roeder et al 2006; Ho et al 2014). More importantly, since the same association test in PrediXcan and TWAS suffers from power loss under some common situations, we develop an alternative and more powerful association test with broader applications. Since there is no uniformly most powerful gene-based association test, any single non-adaptive test will lose power in some situations; it is important to develop and utilize adaptive tests to yield consistently high power (Li and Tseng 2011; Lee et al 2012; Pan et al 2014). We propose using such an adaptive and powerful test under a general and rigorous framework of generalized linear models (GLMs), which can accommodate various types of quantitative, categorical and survival phenotypes and can adjust for covariates. It is applicable to both individual-level genotypic, phenotypic data and GWAS summary statistics. It is flexible to incorporate a single or multiple sets of weights derived from eQTL data or other data sources.

Methods

PrediXcan and TWAS

We briefly review PrediXcan and TWAS for GWAS individual-level data before giving our new formulation. One first builds a prediction model for a gene's expression level, called "genetically regulated expression (GReX)", by using the genotypes around the gene based on an eQTL dataset. Next, for a GWAS dataset, one uses the prediction model to predict or "impute" the GReX of the gene using the SNPs around the gene for each subject in a main GWAS dataset. Specifically, for a given gene, suppose that in an eQTL dataset, Y^* and $X^* = (X_1^*, ..., X_p^*)'$ are the the expression level of of the gene and the p SNP genotype scores (with additive coding) around the gene respectively. A linear model is assumed: $Y^* = \sum_{j=1}^{p} w_j X_j^* + \epsilon$, where w_j is the *cis*-effect of SNP j on gene expression and ϵ is the noise. Based on the eQTL dataset, one can use a method, e.g. elastic net (Zou and Hastie 2005) or a Bayesian linear mixed model (Zhou et al 2013) as used in PrediXcan and TWAS respectively, to obtain the estimates $\hat{w}_j$'s. Now for a given GWAS dataset, for each subject i with the genotype scores $X_i = (X_{i,1}, ..., X_{i,p})'$ for the gene, the predicted GReX is $\widehat{\text{GReX}}_i = \sum_{j=1}^{p} \hat{w}_j X_{i,j}$. For a trait Y_i for subject i in the GWAS dataset, one simply applies a suitable GLM

$$g(E(Y_i)) = \beta_0 + \widehat{\text{GReX}}_i \beta_c = \beta_0 + \sum_{j=1}^{p} \hat{w}_j X_{i,j} \beta_c \tag{1}$$

to test for association between the trait and predicted/imputed expression with null hypothesis H_0: $\beta_c = 0$, where $g()$ is the canonical link function (e.g. the logit and the identity functions for binary and quantitative traits respectively), and $E(Y_i)$ is the mean of the trait. One of the asymptotically equivalent Wald, Score and likelihood ratio tests can be used.

A novel reformulation and extensions

Here we first point out that PrediXcan and TWAS can be regarded as a special case of general association testing with multiple SNPs in a GLM:

$$g(E(Y_i)) = \beta_0 + \beta' X_i = \beta_0 + \sum_{j=1}^{p} X_{i,j}\beta_j. \quad (2)$$

The goal is to test $H_0 : \beta = (\beta_1, ..., \beta_p)' = 0$. It can be verified that both PrediXcan and TWAS are a weighted Sum test in the above general model (Pan 2009) with weights $\hat{w}_j$ on each SNP j; that is, PrediXcan and TWAS conduct the Sum test on H_0 with the genotype scores $X_{i,j}$ replaced by the weighted genotype scores $\hat{w}_j X_{i,j}$ in GLM (2). This new interpretation and formulation will facilitate our gaining insights into PrediXcan and TWAS, including their possible limitations, thus motivating some modifications for improvement. It offers a direct and intuitive justification for PrediXcan and TWAS: the two methods perform well due to their over-weighting on expression-associated SNPs (eSNPs), as supported by empirical evidence that eSNPs are more likely to be associated with complex traits and disease (Nicolae et al 2010). Obviously, it also suggests their extensions to other endophenotypes, and to incorporate prior knowledge and other data sources related to the GWAS trait of interest, such as previous linkage scans (Roeder et al 2006) and imaging endophenotypes (Xu et al 2017), though we do not pursue it here. More importantly, since the Sum test can be derived under the over-simplifying working assumption of $\beta_1 = \beta_2 = ... = \beta_p = \beta_c$ in (1) and (2) (i.e. all weighted SNPs have an equal effect size and the same effect direction, which is in general incorrect), we can see possible limitations of the Sum test and thus of PrediXcan and TWAS. As discussed in Pan (2009), Pan et al (2014) and others (Wu et al 2011), the Sum test may lose power if the effect directions of the (weighted) SNPs are different, or the effect sizes are sparse (i.e. with many 0s). Accordingly, one may apply other tests, e.g. the sum of squared score (SSU) test that is equivalent to a variance-component score test as used in kernel machine regression (also

known as SKAT in rare variant analysis) with a linear kernel and a nonparametric MANOVA (also called genomic distance-based regression) with the Euclidean distance metric (Wessel and Schork 2006), which may yield higher power under many situations (Pan 2011; Schaid 2010a, 2010b).

New method: aSPU

A class of the so-called sum of powered score (SPU) tests cover both the Sum and SSU tests as special cases. Specifically, we denote the unweighted and weighted score vectors for β in (1) as

$$U^* = (U_1^*, ..., U_p^*)' = \sum_{i=1}^{n} X_i'(Y_i - \hat{\mu}_i^0), \qquad U = (U_1, ..., U_p)' = WU^* = \sum_{i=1}^{n} WX_i'(Y_i - \hat{\mu}_i^0),$$

where $\hat{\mu}_i^0$ is the fitted mean of Y_i under H_0 (with $\beta = 0$) in (1), and $W = \text{Diag}(\hat{w}_1, ..., \hat{w}_p)$ are the weights of the SNPs derived from an eQTL dataset. The effects of the weights can be regarded as replacing the unweighted genotype scores $X_{i,j}$ by the weighted genotype scores $\hat{w}_j X_{i,j}$ in GLM (2). The Sum (i.e. PrediXcan and TWAS) and SSU tests based on the weighted genotypes are:

$$T_{\text{Sum}} = \sum_{j=1}^{p} U_j, \qquad T_{\text{SSU}} = U^T U = \sum_{j=1}^{p} U_j^2.$$

More generally, for an integer $\gamma \geq 1$, an SPU(γ) test is defined as

$$T_{\text{SPU}(\gamma)} = \sum_{j=1}^{p} U_j^{\gamma}.$$

It is clear SPU(1)=Sum and SPU(2)=SSU. Furthermore, for an even integer $\gamma \to \infty$, we have $T_{\text{SPU}(\gamma)} \propto \left(\sum_{j=1}^{p} |U_j|^{\gamma}\right)^{1/\gamma} \to \max_j |U_j| = T_{\text{SPU}(\infty)}$. The SPU($\infty$) is closely related to the UminP test (but ignoring possibly varying variances of U_j's); often they performed similarly (Pan 2009).

Since there is no uniformly most powerful test, for a given situation, any non-adaptive test

may or may not be powerful. By using various values of γ, we yield a class of SPU tests, one of which is expected to be more powerful in any given situation. For example, the Sum=SPU(1) test treats each SNP equally *a priori*, yielding high power if all the SNPs are associated with the trait with similar effect sizes and the same association direction. On the other hand, when only a smaller subset of SNPs are associated with the trait, or their association directions are different, the SSU=SPU(2) test is often more powerful. As γ increases, SPU(γ) relies more on the SNPs that are more strongly associated with the trait, and is thus more powerful for more sparse association signals (i.e. fewer associated SNPs). In the end, as γ approaches ∞ (as an even integer), it only considers the most significant SNP.

Since the optimal value of γ is unknown and data-dependent, we propose using an **adaptive SPU (aSPU)** test to data-adaptively approximate the most powerful SPU test among a set of versatile SPU(γ) tests with various values of γ, thus maintaining high power in a wide range of scenarios. Empirically we have found that using $\Gamma = \{1, 2, 3, \ldots, 6, \infty\}$ often performs well and thus adopt it; the aSPU test is defined as

$$T_{aSPU} = \min_{\gamma \in \Gamma} P_{\mathrm{SPU}(\gamma)}, \tag{3}$$

where $P_{\mathrm{SPU}(\gamma)}$ is the p-value of the SPU(γ) test.

P-value calculations: Although asymptotic p-values for the SPU(1)=Sum and SPU(2)=SSU tests can be calculated (Pan 2009) (with possible small-sample adjustments (Lee et al 2012; Chen et al 2015; Wang 2016)), in general, we can use *one layer of Monte Carlo simulations* to estimate the p-values for all the SPU and aSPU tests *simultaneously* (Pan et al 2014). Specifically, we simulate null score vectors $U^{(b)} \sim N(0, V)$ for $b = 1, \ldots, B$, from its asymptotic null distribution, a multivariate normal with mean 0 and covariance matrix V; there is a closed form solution for V (Pan et al 2014). Then the null statistics $T^{(b)}_{\mathrm{SPU}(\gamma)}$ are calculated from the null score vectors $U^{*(b)}$ for $b = 1, \ldots, B$, and the p-value of the SPU(γ) test is $P_{\mathrm{SPU}(\gamma)} = [\sum_{b=1}^{B} I(|T^{(b)}_{\mathrm{SPU}(\gamma)}| \geq |T_{\mathrm{SPU}(\gamma)}|) + 1]/(B + 1)$. Then the p-value for the aSPU test

is calculated as $P_{\text{aSPU}} = [\sum_{b=1}^{B} I(T_{\text{aSPU}}^{(b)} \leq T_{\text{aSPU}}) + 1]/(B+1)$ with $T_{\text{aSPU}}^{(b)} = \min_{\gamma \in \Gamma} p_{\gamma}^{(b)}$ and $p_{\gamma}^{(b_1)} = [\sum_{b \neq b_1} I(|T_{\text{SPU}(\gamma)}^{(b)}| \geq |T_{\text{SPU}(\gamma)}^{(b_1)}|) + 1]/B$.

Association testing with summary statistics

One practical way to increase the sample size is to form large consortia, aiming for meta analysis of multiple GWAS, for which often only summary statistics for single SNP-single trait associations, rather than individual-level genotypic and phenotypic data, are available (and practically feasible for many cohorts with possibly different study designs). Hence it is extremely useful to develop methods like TWAS that are applicable to GWAS summary statistics as well as to GWAS individual-level data. The aSPU test is easily extended to GWAS summary statistics without individual-level data. Suppose that $Z_j = \hat{\beta}_j/\text{SE}_j$ is the Z-statistic for association between the GWAS trait and SNP j, where $\hat{\beta}_j$ is the estimated (marginal and signed) association effect and SE_j is its standard error. We just need to simply redefine $U = WZ$ with $Z = (Z_1, Z_2, ..., Z_p)'$, then proceed as before. We use a reference sample (e.g. the 1000 Genome Project data) to estimate linkage disequilibrium (LD) among the SNPs and thus the correlation matrix for Z and U (Kwak and Pan 2016; Gusev et al 2016).

Association testing with multiple sets of weights

Now we extend the aSPU test to the case with multiple sets of eQTL datasets, or more generally, multiple sets of weights. This is important because of the existence of multiple eQTL datasets measured from different populations or different tissues; it is in general unclear which one is most suitable. After applications with each eQTL dataset separately, it may gain statistical power and biological insights to combine the results across multiple eQTL datasets. Suppose we have K sets of weights, $W^{(k)} = \text{Diag}(w_1^{(k)}, \ldots, w_p^{(k)})$ for $k = 1, 2, \ldots, K$, each estimated from a separate eQTL dataset. To avoid the results depending on the varying scales of the sets of weights, we first standardize the weights to have $\sum_{j=1}^{p} |w_j^{(k)}| = 1$ for each k. Based on

the score vector U^* (with individual-level data) or Z-statistics Z (with GWAS summary data) and the weights $W^{(k)}$, we define $U^{(k)} = W^{(k)}U^*$ or $U^{(k)} = W^{(k)}Z$ accordingly. As before, for a fixed γ, we first apply SPU(γ) to $U^{(k)} = (U_1^{(k)}, ..., U_p^{(k)})'$, yielding its test statistic $T_{SPU(\gamma;k)} = \sum_{j=1}^{p}(U_j^{(k)})^\gamma$ and p-value $P_{SPU(\gamma;k)}$. We then Z-transform each p-value to a Z-statistic $z^*(\gamma;k) = 1 - \Phi^{-1}(P_{SPU(\gamma;k)}/2)$, where $\Phi(.)$ is the CDF of a standard normal distribution. To recover the sign of each statistic, for an odd γ, we have $z(\gamma;k) = \text{sign}(T_{SPU(\gamma;k)})z^*(\gamma;k)$; for an even γ or $\gamma = \infty$, we use $z(\gamma;k) = \text{sign}(T_{SPU(1;k)})z^*(\gamma;k)$. We combine the K sets of weights through combining the K statistics $z(\gamma) = (z(\gamma;1), ..., z(\gamma;K))'$ to form an omnibus SPU(γ) test:

$$\text{SPU}(\gamma)\text{-O} = [z(\gamma) - \mu_0(\gamma)]'V^{-1}(\gamma)[z(\gamma) - \mu_0(\gamma)],$$

where $\mu_0(\gamma)$ and $V^{-1}(\gamma)$ are the sample mean vector and covariance matrix of $z(\gamma)$ under H_0, which can be calculated along with other p-values inside the single layer of simulations. Then, as usual, we combine the omnibus SPU(γ)-O tests into an omnibus aSPU test:

$$T_{\text{aSPU-O}} = \min_{\gamma\in\Gamma} P_{\text{SPU}(\gamma)\text{-O}},$$

where $P_{\text{SPU}(\gamma)\text{-O}}$ is the p-value of SPU(γ)-O. As before, the p-values of all the SPU(γ)-O and aSPU-O can be calculated in a single layer of Monte Carlo simulations.

The omnibus test is not sensitive to weight standardization. For example, scaling by $\sum_j (w_j^\kappa)^{(1/\kappa)}$ yielded similar results in our experiments.

It is easy to verify that SPU(1)-O is equivalent to the omnibus TWAS, denoted TWAS-O. Again, by combining SPU(1)-O and other SPU(γ)-O tests, we obtain the adaptive and omnibus aSPU-O test that may be more powerful across a wide range of scenarios.

Data availability

The WTCCC data can be found at https://www.wtccc.org.uk. The original 2010 lipid GWAS summary data can be downloaded at http://csg.sph.umich.edu/abecasis/public/lipids2010/, while the original 2013 lipid GWAS summary data at http://csg.sph.umich.edu/abecasis/public/lipids2013/; the two pre-processed datasets that we used are downloadable from https://figshare.com/articles/Lipid_2010_summary_data/5373370 and https://figshare.com/articles/Lipid_2013_summary_data/5373382. The LD reference data and eQTL-based weights can be obtained from http://gusevlab.org/projects/fusion/ and https://github.com/hakyimlab/PrediXcan. The related computer scripts and examples can be found at https://github.com/ChongWu-Biostat/TWAS, and the online manual about how to use our proposed methods can be also found at www.wuchong.org/TWAS.html or https://github.com/ChongWu-Biostat/TWAS.

Results

Application to the WTCCC data

We first applied the aSPU test and PrediXcan to the WTCCC individual-level data with the weights downloaded from the PrediXcan database, demonstrating the equivalence of the SPU(1) test and PrediXcan, and more importantly, that the aSPU test could identify more associated genes than PrediXcan in many cases. Specifically, first, following the same procedure of quality control (Burton et al 2007), we lifted the annotation of the WTCCC genotype data from hg18 to hg19 via the UCSC browser (http://genome.ucsc.edu/cgi-bin/hgLiftOver); second, we imputed the genotype data via the Michigan Imputation Server with the following specifications: 1000G Phase 1 v3 as the reference panel, SHAPEIT as the phasing algorithm and EUR (European) as the target population. After imputation, the variants with a minor allele frequency (MAF) > 0.05, the HWE exact test P-value > 0.05

and $R^2 > 0.8$ were kept. As Gamazon et al (2015), we kept only the HapMap Phase 2 subset of SNPs. We considered 7 traits/diseases: bipolar disorder (BD), coronary artery disease (CAD), inflammatory bowel disease (CD), rheumatoid arthritis (RA), hypertension (HT), type 1 diabetes (T1D) and type 2 diabetes (T2D). The weights based on the DGN whole blood expression were downloaded from the PrediXcan database (`https://app.box.com/s/gujt4m6njqjfqqc9tu0oqgtjvtz9860w`). There were 8917 genes whose expression levels could be predicted by elastic net with a cross-validated $R^2 > 0.01$; we thus tested on these 8917 genes with a conservative Bonferroni adjustment with a genome-wide significance level at $0.05/9000 = 5.56 \times 10^{-6}$.

As most of the genes were not expected to be significantly associated with a trait, we used a step-up procedure to increase the number of simulations when calculating the p-values of aSPU and aSPU-O in the subsequent data analysis. We started with a relatively small $B = 10^3$, and re-ran the tests with $B = 10^4$ for the genes with p-values $< 5 \times 10^{-3}$ (but stopped otherwise); we repeated this process by increasing B to 10 times of its previous value for the genes with p-values $< 5/B$ up to $B = 1 \times 10^7$; finally, to be more accurate for a p-value around the significance cut-off, we re-ran the tests on the genes with p-values between 10^{-5} and 10^{-6} with $B = 1 \times 10^8$.

Here are the main results. First, as shown in Supplementary Figure 1, as expected, PrediXcan gave essentially the same results (i.e. p-values) as those of the SPU(1) test for each of the seven traits. Hence, we confirmed and would treat the SPU(1) test to be equivalent to PrediXcan. Second, as shown in Supplementary Figure 2, the aSPU test identified more significant genes than the SPU(1) test (or equivalently, PrediXcan) for traits CD, BD and T1D (i.e. (10, 3, 38) versus (8, 2, 29)), while it was the opposite for HT (i.e. 0 versus 1), and they were tied (with (1, 4, 0)) for CAD, RA and T2D; note the large difference for T1D, which was statistically significant with an exact p-value=0.0039 by McNemar's test (Fagerland et al 2013) (while other differences were not). Table 1 lists the significant genes identified by the aSPU test but not by the SPU(1) test (and PrediXcan) at the genome-wide significance level; some of the significant

genes were confirmed in later studies. In total, we identified 15 new genes, five out of which have been reported by other studies. For CD, we identified 4 new genes, all of which contained some genome-wide significant SNPs identified by other studies, constituting a highly significant validation of our discoveries. Specifically, genes IRGM, P4HA2, PTGER4, and RBM22 contain significant SNPs rs11741861 ($p = 2 \times 10^{-19}$, de Lange et al 2017), rs2188962 ($p = 6 \times 10^{-36}$, de Lange et al 2017), rs11742580 ($p = 7 \times 10^{-36}$, Franke et al 2010), and rs11741861 ($p = 2 \times 10^{-19}$, de Lange et al 2017), respectively. Importantly, gene IRGM is related to CD pathogenesis with the involvement in the process of autophagy (Liu et al 2015). However, our newly identified genes JAKMIP1 and PDK1, associated with BD and CAD respectively, have not yet been reported elsewhere. For T1D, we identified 9 new genes, of which one contains some genome-wide significant SNPs as reported by another study (Plagnol et al 2011). In summary, these 15 significant genes identified by the aSPU test would have been missed by PrediXcan, some of which have been confirmed by other studies while the remaining ones are to be validated in the future.

Application to the lipid GWAS summary data

We next applied our new methods and TWAS to a 2010 lipid GWAS summary dataset (∼100,000 samples, Teslovich et al 2010), while using its follow-up with a larger sample size (∼189,000) for partial validation. To facilitate comparison, we used the three sets of weights and the 1000 Genomes Project data as the reference sample, all downloaded from the TWAS database (`http://gusevlab.org/projects/fusion/#reference-functional-data`) (on *Jan 11th, 2017*). The three sets of weights were based on three eQTL datasets: microarray gene expression data of peripheral blood from 1,245 unrelated subjects from the Netherlands Twin Registry (NTR) (Wright et al 2014), microarray expression data of blood from 1,264 subjects from the Young Finns Study (YFS), and RNA-seq measured in adipose tissue from 563 individuals from the Metabolic Syndrome in Men study (METSIM); for each pair of gene-eQTL dataset, we used the

Table 1. Significant genes identified by the aSPU test, but not by the SPU(1) test (or PrediXcan) at the genome-wide significance threshold of 5.56×10^{-6}. The validated gene-trait associations appeared in the following references: [1] Franke et al (2010); [2] Kenny et al (2012); [3] de Lange et al (2017); [4] Plagnol et al (2011).

Trait	Gene	Chr.	SNPs in #SNPs	aSPU	SPU(1)	SPU(2)	Reported Valid. ref
CD	IRGM	5	34	5.0E-07	1.7E-04	3.8E-08	CD [1]
	P4HA2	5	7	2.5E-06	1.7E-01	8.4E-07	CD [2]
	PTGER4	5	22	1.1E-06	1.0E-05	2.6E-07	CD [1]
	RBM22	5	15	7.0E-07	1.4E-02	5.5E-06	CD [3]
BD	JAKMIP1	4	21	1.0E-07	1.8E-05	3.2E-09	-
CAD	PDK1	2	33	5.5E-06	5.8E-03	2.7E-05	-
T1D	ALDH2	12	19	1.0E-07	6.6E-05	1.1E-07	-
	BCL2L15	1	21	1.0E-07	5.1E-04	2.1E-06	T1D [4]
	HFE*	6	61	7.0E-07	5.8E-01	5.7E-08	-
	MPHOSPH10	2	20	6.0E-07	1.2E-01	1.4E-07	-
	PGBD1*	6	33	1.0E-07	3.0E-04	2.4E-09	-
	PRSS16*	6	32	1.5E-06	6.3E-05	2.0E-07	-
	TMEM116	12	23	2.0E-06	5.0E-04	6.9E-07	-
	ZNF193*	6	64	1.0E-07	2.3E-02	5.0E-11	-
	ZSCAN12*	6	14	4.0E-07	8.1E-06	2.5E-06	-

* The genes fall in the HLA region.

set of the optimal weights estimated by TWAS. For each trait, there were 1264, 3555 and 2295 significant cis-heritable genes with weights drawn from the NTR, YFS and METSIM eQTL datasets respectively, resulting in a total of 7114 genes being tested; when combining across three sets of weights, there were 1223 genes being tested by the omnibus TWAS and thus considered as the candidate genes by any omnibus test. Thus, we used a conservative Bonferroni adjustment with $0.05/8500 = 5.88 \times 10^{-6}$ as the genome-wide significance level. The GWAS Z-scores were imputed for any missing SNPs using the IMPG algorithm (Pasaniuc et al 2014).

The new test identified more associations

We numerically confirmed the equivalence between the SPU(1) test and TWAS (Supplementary Figure 3). Hence, we used the results of the SPU(1) test to represent those of TWAS in the following. More importantly, the aSPU test could identify a larger number of significant genes

Table 2. The numbers of the significant genes identified by analyzing the 2010 lipid data for each single set of the weights and the combined one (i.e. with the omnibus aSPU and TWAS tests). The numbers a/b/c in each cell indicate the numbers of (a) the significant genes; (b) the significant genes that covered a genome-wide significant SNPs in the 2010 lipid data; (c) the significant genes that covered a genome-wide significant SNPs in the 2013 lipid data.

Trait	Test	NTR	YFS	METSIM	Combined
HDL	aSPU	19/16/17	29/27/29	22/19/22	21/17/17
	TWAS	16/14/15	25/22/24	19/15/19	20/16/17
LDL	aSPU	15/15/15	19/18/18	17/16/17	14/13/13
	TWAS	8/7/8	10/9/9	7/7/7	7/7/7
TG	aSPU	17/16/17	33/30/32	15/14/14	20/19/19
	TWAS	9/9/9	17/16/17	8/7/7	12/11/11
TC	aSPU	26/25/26	28/26/27	28/28/28	20/20/20
	TWAS	15/14/15	18/16/17	15/14/15	14/13/13

than TWAS in every case across the four traits (HDL, LDL, TC and TG) and three sets of weights (NTR, YFS and METSIM); the same conclusion holds for the omnibus aSPU and omnibus TWAS tests (Table 2). As a partial validation, a high proportion of the identified genes covered at least one genome-wide significant SNP in the 2010 ($\sim$ 100,000 samples) or the larger 2013 data ($\sim$ 189,000 samples, Global Lipids Genetics Consortium 2013). In Supplementary Tables 1-7, we list the significant genes identified by aSPU and TWAS based on analyzing the 2010 lipid data, including novel ones not overlapping with any known risk loci covering any significant SNPs ($P < 5 \times 10^{-8}$) in the 2010 data (Supplementary Tables 1-2) or in the larger 2013 lipid GWAS dataset (Supplementary Tables 3).

Compared to TWAS, the aSPU test can still maintain high power if many of the SNPs in a gene are not associated with a trait. For example, based on the 2010 data, for trait HDL and gene DR1 with the YFS-based weights, among the 17 SNPs with non-zero weights, there were only one SNP with a p-value less than 5×10^{-7} (but larger than the genome-wide significance level 5×10^{-8}), resulting in a non-significant p-value ($= 1.4 \times 10^{-4}$) by TWAS, or equivalently by SPU(1). Since an SPU(γ) test with a larger $\gamma > 1$ relied more on the SNPs with the smaller p-values (i.e. more strongly associated with the trait), it yielded a more significant p-value with a larger γ: the SPU(2) and SPU(5) tests gave p-value $= 2.9 \times 10^{-6}$ and 3.5×10^{-6} respectively,

Table 3. The numbers of the significant genes identified by analyzing the 2013 lipid data for each single set of the weights and the combined one (i.e. with the omnibus aSPU and TWAS tests). The numbers a/b/c in each cell indicate the numbers of (a) the significant genes; (b) the significant genes that covered a genome-wide significant SNPs in the 2010 lipid data; (c) the significant genes that covered a genome-wide significant SNPs in the 2013 lipid data.

Trait	Test	NTR	YFS	METSIM	Combined
HDL	aSPU	21/18/21	52/39/48	33/24/32	31/21/29
	TWAS	21/17/20	35/26/33	24/17/23	26/16/24
LDL	aSPU	24/23/24	40/34/37	27/24/26	20/17/18
	TWAS	16/14/16	28/23/25	16/14/15	17/15/16
TG	aSPU	21/19/21	43/37/42	26/20/25	23/21/23
	TWAS	15/13/15	27/22/25	15/12/15	16/15/16
TC	aSPU	36/27/35	70/52/66	42/36/42	37/32/36
	TWAS	25/18/24	38/28/35	26/21/26	25/20/23

leading to the significant p-value $= 2.3 \times 10^{-6}$ of aSPU in the end. Furthermore, based on the 2013 data, several SNPs in the locus were genome-wide significant, and both aSPU and TWAS confirmed the significance of gene DR1. Two locusZoom plots for the locus based on the 2010 and 2013 data are shown in Figure 1. Another similar example is with the 2010 data for trait TG and gene NEIL2 with the NTR-based weights: although all the five weighted score elements for the SNPs were negative, their absolute values varied from 0.07 to 2.89, leading to the p-values $< 5 \times 10^{-6}$ for any SPU(γ) with $\gamma > 1$, more significant than the p-value $= 8.8 \times 10^{-6}$ of SPU(1).

An SPU(γ) test with an even γ could be more powerful than SPU(1) with varying SNP association directions. In the 2010 data, for trait LDL and gene NTN5 with the METSIM-based weights, among the 450 weighted score elements of the individual SNPs, 43% were positive while the remaining 57% were negative, leading to a much more significant p-value of SPU(2) over that of SPU(1): 8.2×10^{-8} versus 4.9×10^{-5}.

Generally, as expected from statistical theory and as shown in Supplementary Figures 4-5, since the SPU(1) test (i.e. TWAS) might not be powerful for a given gene, the aSPU test could gain statistical power through other more powerful SPU tests like SPU(2).

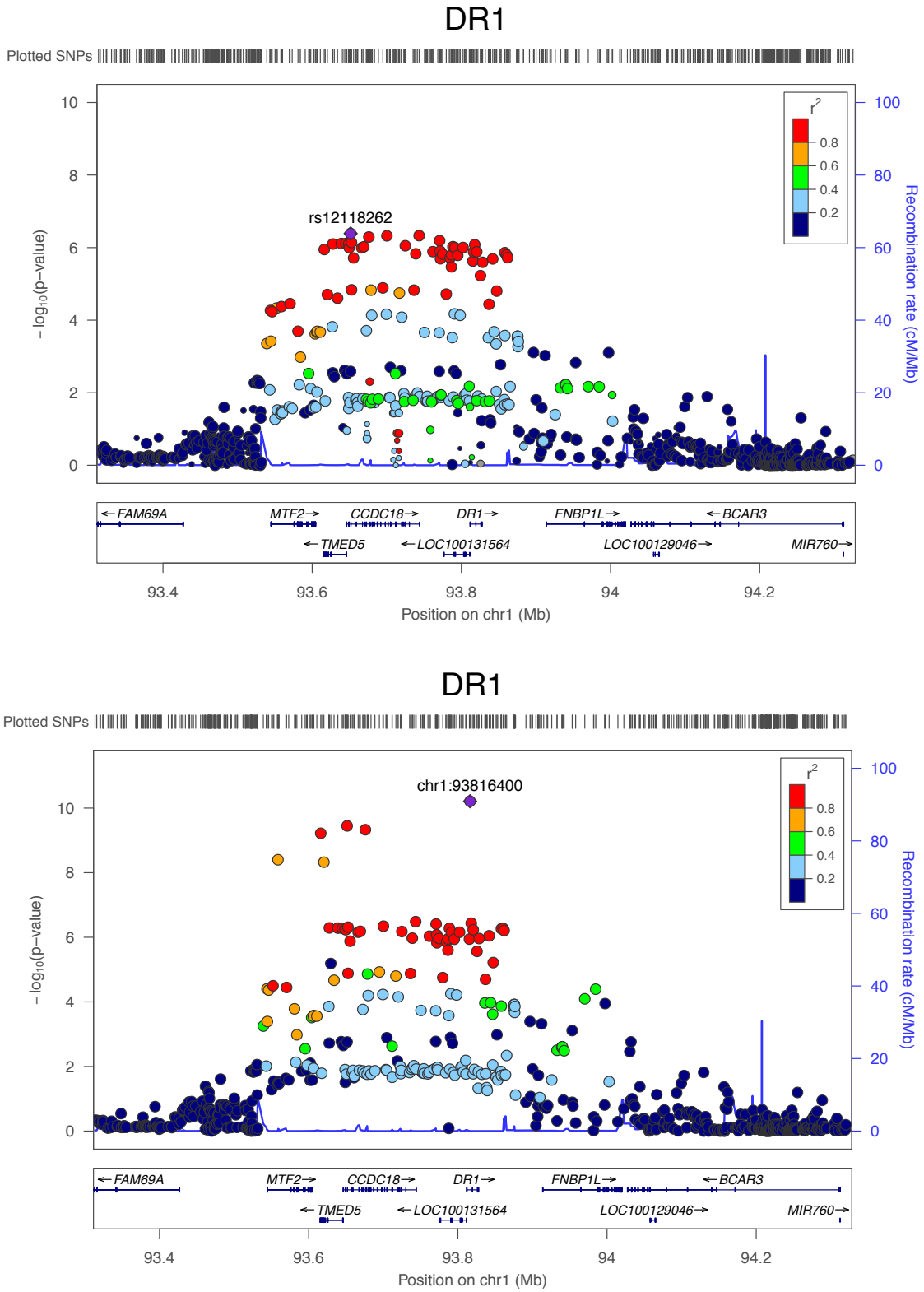

Figure 1. LocusZoom plots of SNP-HDL associations at the locus around gene DR1 based on the 2010 lipid data (top) and 2013 data (bottom).

The new test identified novel associations

Finally, we applied the aSPU and TWAS (and their omnibus versions) to the larger 2013 lipid dataset (Global Lipids Genetics Consortium 2013), listing the numbers of the significant genes identified by each method in Table 3. Again the aSPU test identified a much larger number of significant associations. The Manhattan plots for the pooled results of aSPU for each set of the weights and of aSPU-O combining the three sets of the weights for each trait are shown in Figures 2-3; a comparison between aSPU/aSPU-O versus TWAS/TWAS-O for trait LDL is shown in Supplementary Figures 6-7. In total, aSPU and TWAS identified 17 and 14 new associations not overlapping with known risk loci respectively; among the 6 new associations uniquely identified by aSPU test, gene PFAS was reported to be associated with LDL in a later meta-analysis (Below et al 2016). The new associations identified by aSPU or/and TWAS are listed in Table 4, while all other ones are in Supplementary Tables 8-11. It is noteworthy that in Table 4, with the p-values close to the significance cut-off, the aSPU test barely missed the three significant genes uniquely identified by the SPU(1) test (i.e. TWAS); in contrast, the SPU(1) test gave the much larger p-values for several significant genes uniquely identified by aSPU.

Simulations

It was shown previously that the aSPU test could control its type I error rate effectively in the context of unweighted association testing (Pan et al 2014), which is expected to hold in the current context. Nevertheless, we conducted a simulation study to confirm it. We used the individual-level (imputed) genotypic data of the WTCCC control and T2D samples with a combined sample size of $n = 4862$. We randomly generated a binary trait with an equal probability 0.5 (of being in either category) for each subject, and calculated a summary Z statistic for each SNP. We then applied the aSPU test along with the asymptotic SPU(1) and SPU(2) tests to the individual-level data with the same PrediXcan-constructed weights

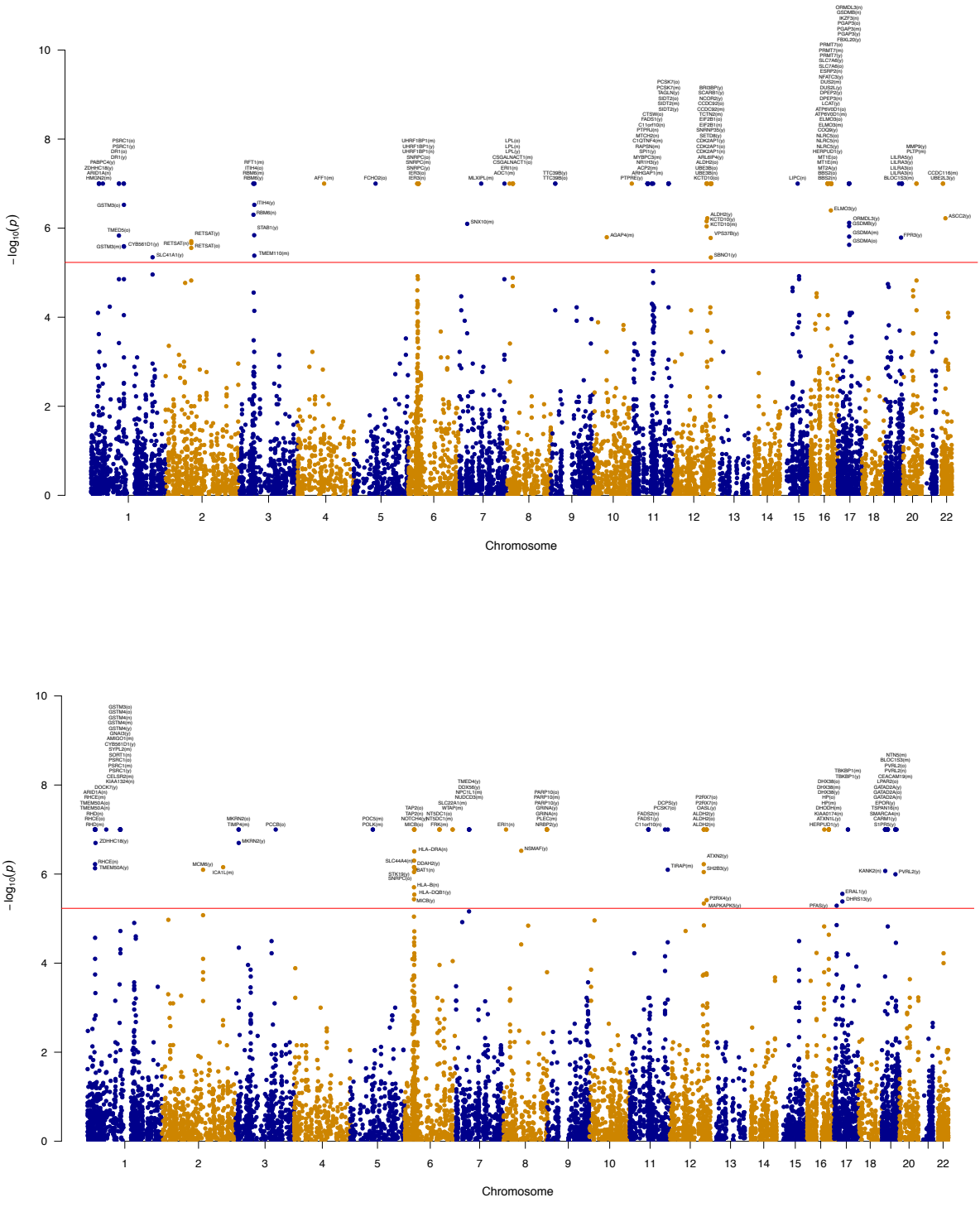

Figure 2. Manhattan plots for the pooled results of aSPU and aSPU-O for traits HDL (top) and LDL (bottom) based on the 2013 lipid data. The letters "(n)", "(y)", "(m)" and "(o)" following a gene's name indicate the result of aSPU based on the NTR, YFS and METSIM weights and that of aSPU-O respectively.

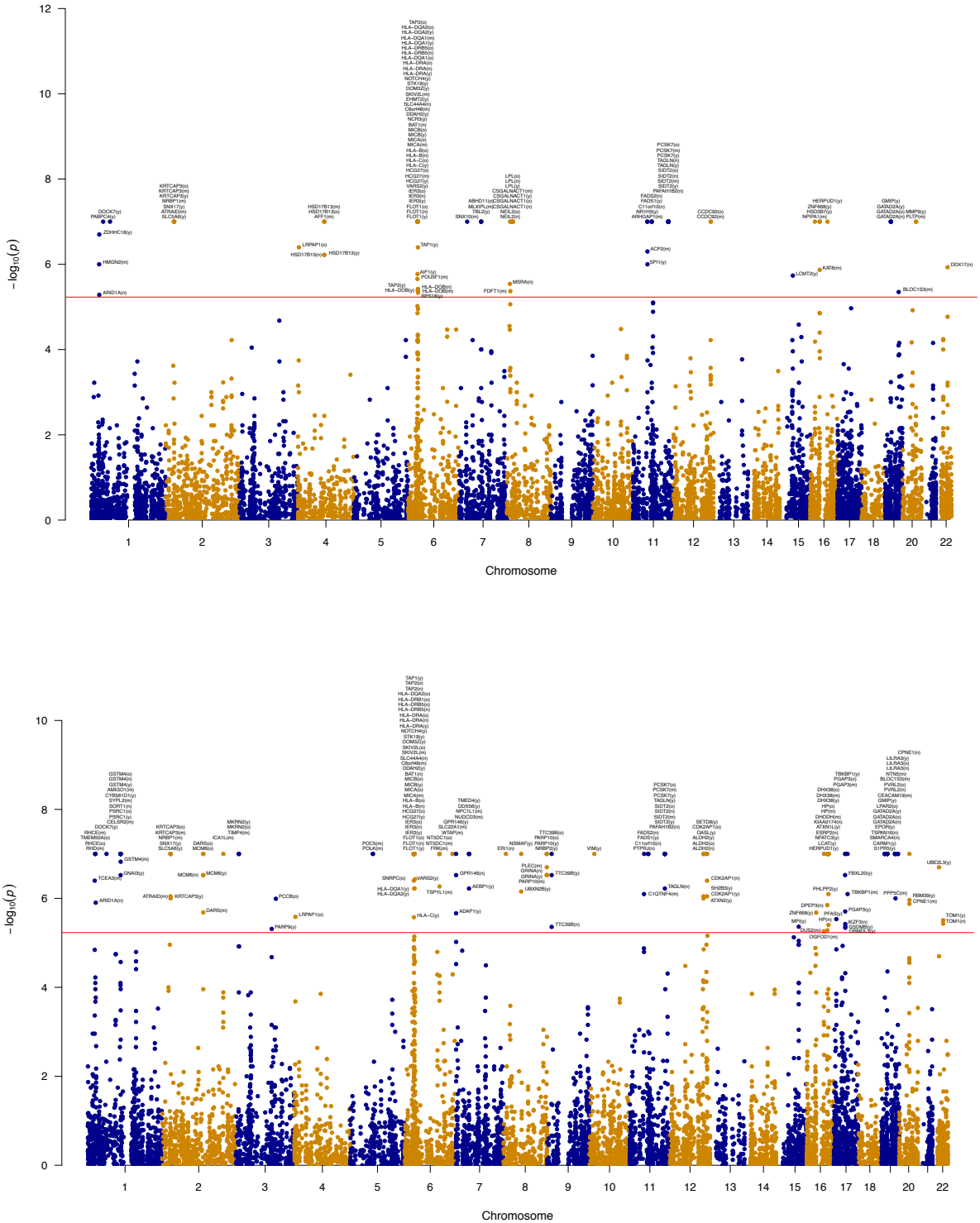

Figure 3. Manhattan plots for the pooled results of aSPU and aSPU-O for traits TG (top) and TC (bottom) based on the 2013 lipid data.

Table 4. Significant gene-trait associations identified by aSPU or/and TWAS with no known risk loci within 500kb. The column "Sig.test" indicates the corresponding association was detected by aSPU or SPU(1) or both.

Trait	Weight	Gene	Chr.	Locus start	Locus end	aSPU	SPU(1)	SPU(2)	Sig.test
HDL	YFS	SLC41A1	1	205758221	205782876	4.5E-06	7.9E-04	2.2E-04	aSPU
HDL	YFS	ASCC2	22	30184597	30234271	6.0E-07	9.7E-04	6.2E-06	aSPU
LDL	YFS	PFAS	17	8150936	8173809	5.1E-06	4.2E-03	1.1E-06	aSPU
TG	METSIM	ARHGAP1	11	46698630	46722165	1.0E-07	3.8E-04	1.2E-08	aSPU
TC	YFS	ZNF668	16	31072164	31085641	2.1E-06	3.7E-05	1.4E-05	aSPU
TC	YFS	PFAS	17	8150936	8173809	2.9E-06	5.4E-03	1.8E-06	aSPU
LDL	YFS	HSPA6	1	161494036	161496681	2.8E-05	4.6E-06	3.2E-02	SPU(1)
TG	YFS	PACS1	11	65837834	66012218	8.2E-06	5.7E-06	1.2E-06	SPU(1)
TC	YFS	ADCY3	2	25042038	25142708	1.1E-05	4.1E-06	2.5E-02	SPU(1)
HDL	NTR	RETSAT	2	85569078	85581848	2.2E-06	1.3E-06	2.3E-06	Both
HDL	YFS	RETSAT	2	85569211	85581743	2.0E-06	2.3E-06	1.5E-06	Both
HDL	YFS	PTPRE	10	129705325	129884119	1.0E-07	3.7E-08	1.4E-01	Both
HDL	METSIM	SNX10	7	26331541	26413949	8.0E-07	5.6E-07	8.7E-07	Both
LDL	YFS	ERAL1	17	27181956	27188085	2.8E-06	1.7E-06	2.2E-06	Both
LDL	YFS	DHRS13	17	27224799	27230089	4.1E-06	2.2E-06	2.4E-06	Both
LDL	METSIM	ICA1L	2	203640690	203736708	7.0E-07	2.1E-07	5.9E-07	Both
TG	YFS	LCMT2	15	43619974	43622803	1.8E-06	1.6E-06	3.2E-06	Both
TC	NTR	ARID1A	1	27022521	27109023	1.2E-06	9.7E-07	1.4E-06	Both
TC	YFS	PARP9	3	122246771	122283424	4.9E-06	2.4E-06	3.7E-06	Both
TC	YFS	MPI	15	75182346	75191798	4.3E-06	6.1E-07	6.8E-07	Both

based on the DGN whole blood gene expression data; in addition, we also applied the tests to the summary Z-statistics with the TWAS-constructed weights based on the NTR, YFS and METSIM gene expression data respectively; finally, we applied the omnibus aSPU-O test to the summary statistics to combine results across the three sets of NTR, YFS and METSIM weights. Similar to that in the real data analysis, we used a step-up procedure to adaptively determine the number of simulations when calculating the p-values of aSPU and aSPU-O. We started with a relatively small $B = 10^3$; we re-ran the tests with $B = 10^4$ for the genes with p-values $< 5 \times 10^{-3}$, but stopped otherwise; we repeated this process by increasing B to 10 times of its previous value for the genes with p-values $< 5/B$) up to $B = 1 \times 10^7$. As shown in the Q-Q plots in Supplementary Figure 8, in each case each test controlled the type I error rate satisfactorily.

Discussion

We note that the algorithm in PrediXcan and TWAS (for quantitative GWAS traits) is similar to two-stage least squares (2SLS) with the SNPs as instrumental variables, which is related to the Mendelian Randomization (MR) approach of Zhu et al (2016). However, since only *cis*-eQTLs are considered in the above approaches (while *trans*-eQTLs are ignored), the used SNPs may be weak instrumental varables and the assumptions underlying MR are likely to be violated. Hence we have so far avoided a causal interpretation of detected associations.

In summary, we have developed a powerful adaptive test (aSPU) to integrate GWAS and eQTL data. We have demonstrated its improved power over the existing methods; in fact, the same association test underlying the two existing methods, PrediXcan and TWAS, can be regarded as a special case of our proposed test, explaining why our proposed test may have improved power. Importantly, our new formulation of PrediXcan and TWAS in the general framework of weighted multi-SNP association testing suggests other possible extensions, e.g. applications not only to other informative endophenotypes (Xu et al 2017), but also to

incorporate other sources of information like SNP functional annotation, previous linkage scans (Roeder et al 2006) and multiple phenotypes (Kim et al 2015; Zhu et al 2015), which may be worth to be investigated in the future.

The proposed statistical tests are implemented in R package `aSPU2`, which along with some example R code is publicly available at `www.wuchong.org/TWAS.html`.

Supplemental information

Supplemental Data include 8 Supplementary Figures and 11 Supplementary Tables.

Acknowledgment

The authors are grateful to the reviewers and the editor for helpful and constructive comments. This research was supported by NIH grants R21AG057038, R01HL116720, R01GM113250 and R01HL105397, and by the Minnesota Supercomputing Institute; ZX was supported by a University of Minnesota MnDRIVE Fellowship and CW by a University of Minnesota Dissertation Fellowship. The authors thank Dr. Sasha Gusev for providing and helping with the use of the TWAS database.

References

1000 Genomes Project Consortium, 2010 A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-1073.

Below, J. E., E. J. Parra, E. R. Gamazon, J. Torres, S. Krithika, S. Candille *et al.*, 2016 Meta-analysis of lipid-traits in Hispanics identifies novel loci, population-specific effects, and tissue-specific enrichment of eQTLs. *Scientific Reports* **6**.

345 Bhutani, K., A. Sarkar, Y. Park, M. Kellis, and N. J. Schork, 2017 Modeling prediction error
346 improves power of transcriptome-wide association studies. bioRxiv.
347 doi: http://dx.doi.org/10.1101/108316.

348 Burton, P. R., D. G. Clayton, L. R. Cardon, N. Craddock, P. Deloukas *et al.*, 2007 Genome-
349 wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.
350 *Nature* **447**, 661-678.

351 Chen, J., W. Chen, N. Zhao, M. C. Wu, and D. J. Schaid, 2015 Small sample kernel association
352 tests for human genetic and microbiome association studies. *Genet Epidemiol.* **40**, 5-19.

353 de Lange, K. M., L. Moutsianas, J. C. Lee, C. A. Lamb, Y. Luo *et al.*, 2017 Genome-wide
354 association study implicates immune activation of multiple integrin genes in inflammatory
355 bowel disease. *Nat Genet.* **49**, 256-261.

356 Fagerland, M. W., S. Lydersen, and P. Laake, 2013 The McNemar test for binary matched-
357 pairs data: mid-p and asymptotic are better than exact conditional. *BMC Medical Re-*
358 *search Methodology* **13**, 91.

359 Franke, A., D. P. McGovern, J. C. Barrett, K. Wang, G. L. Radford-Smith *et al.*, 2010 Genome-
360 wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility
361 loci. *Nat Genet.* **42**, 1118-1125.

362 Fuchsberger, C., G. R. Abecasis, and D. A. Hinds, 2015 minimac2: faster genotype imputation.
363 *Bioinformatics* **31**, 782-784.

364 Gamazon, E. R., H. E. Wheeler, K. P. Shah, S. V. Mozaffari, K. Aquino-Michaels, R. J.
365 Carroll, *et al.*, 2015 A gene-based association method for mapping traits using reference
366 transcriptome data. *Nat. Genet.* **47**, 1091-1098.

367 Global Lipids Genetics Consortium, 2013 Discovery and refinement of loci associated with
368 lipid levels. *Nat Genet* **45**, 1274-1283.

GTEx Consortium, 2015 The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-660.

Gusev, A., A. Ko, H. Shi, G. Bhatia, W. Chung, B. W. Penninx *et al.*, 2016 Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet.* **48**, 245-252.

Gusev, A., N. Mancuso, H.K. Finucane, Y. Reshef, L. Song *et al.*, 2017 Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. bioRxiv. doi: https://doi.org/10.1101/067355.

He., X., C. K. Fuller, Y. Song, Q. Meng, B. Zhang, X. Yang, and H. Li., 2013 Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. *Am J Hum Genet* **92**, 667-680.

Ho., Y. Y., E. C. Baechler, W. Ortmann, T. W. Behrens, R. R. Graham *et al.*, 2014 Using gene expression to improve the power of genome-wide association analysis. *Hum Hered* **78**, 94-103.

Howie., B., C. Fuchsberger, M. Stephens, J. Marchini, and G. R. Abecasis, 2012 Fast and accurate genotype imputation in genome-wide association studies through prephasing. *Nat. Genet.* **44**, 955-959.

Kenny., E. E., I. Pe'er, A. Karban, L. Ozelius, A. A. Mitchell, *et al.*, 2012 A genome-wide scan of Ashkenazi Jewish Crohn's disease suggests novel susceptibility loci. *PLoS Genet.* **8(3)**, e1002559.

Kim, J., Y. Bai, and W. Pan, 2015 An adaptive association test for multiple phenotypes with GWAS summary statistics. *Genet Epidemiol.* **39**, 651-663.

Kwak, I., and W. Pan, 2016 Adaptive gene- and pathway-trait association testing with gwas summary statistics. *Bioinformatics* **32**, 1178-1184.

Lee., S., M. J. Emond, M. J. Bamshad, K. C. Barnes, M. J. Rieder, *et al.*, 2012 Optimal Unified Approach for Rare-Variant Association Testing with Application to Small-Sample Case-Control Whole-Exome Sequencing Studies. *Am J Hum Genet.* **91**, 224-237.

Li, J., and G. C. Tseng, 2011 An adaptively weighted statistic for detecting differential gene expression when combining multiple transcriptomic studies. *Ann Appl Stat* **5**, 994-1019.

Liu, J. Z., S. van Sommeren, H. Huang H, S. C. Ng, R. Alberts *et al.*, 2015 Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* **47**, 979-986.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff et al., 2009 Finding the missing herita- bility of complex diseases. *Nature* **461**, 747–753.

Nicolae, D. L., E. Gamazon, W. Zhang, S. Duan, M. E. Dolan *et al.*, 2010 Trait-Associated SNPs Are More Likely to Be eQTLs: Annotation to Enhance Discovery from GWAS. *PLoS Genet* **6**, e1000888.

Pan, W., 2009 Asymptotic tests of association with multiple SNPs in linkage disequilibrium. *Genet Epidemiol* **33**, 497–507.

Pan, W., 2011 Relationship between genomic distance-based regression and kernel machine regression for multi-marker association testing. *Genet Epidemiol* **35** (4), 211–216.

Pan, W., J. Kim, Y. Zhang, X. Shen, and P. Wei, 2014 A powerful and adaptive association test for rare variants. *Genetics* **197**, 1081-1095.

Pan, W., I. Kwak, and P. Wei., 2015 A Powerful Pathway-Based Adaptive Test for Genetic Association with Common or Rare Variants. *American Journal of Human Genetics* **97**, 86-98.

Park, Y., A. Sarkar, K. Bhutani, and M. Kellis., 2017 Multi-tissue polygenic models for transcriptome-wide association studies. bioRxiv. doi: http://dx.doi.org/10.1101/107623.

Pasaniuc, B., N. Zaitlen, H. Shi, G. Bhatia, A. Gusev *et al.*, 2014 Fast and accurate imputation of summary statistics enhances evidence of functional enrichment. *Bioinformatics* **30**, 2906-2914.

Plagnol, V., J. M. Howson, D. J. Smyth, N. Walker, J. P. Hafler *et al.*, 2011 Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases. *PLoS Genet.* **7**, e1002216.

Roeder, K., S. A. Bacanu, L. Wasserman, and B. Devlin, 2006 Using linkage genome scans to improve power of association in genome scans. *Am J Hum Genet.* **78**, 243-252.

Schaid, D. J., 2010a Genomic similarity and kernel methods I: advancements by building on mathematical and statistical foundations. *Hum Hered.* **70**, 109-131.

Schaid, D. J., 2010b Genomic similarity and kernel methods II: methods for genomic information. *Hum Hered.* **70**, 132-140.

Teslovich, T. M., K. Musunuru, A. V. Smith, A. C. Edmondson, I. M. Stylianou *et al.*, 2010 Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-713.

Torres, J. M., A. N. Barbeira, R. Bonazzola, A. P. Morris, K. P. Shah *et al.*, 2017 Integrative cross tissue analysis of gene expression identifies novel type 2 diabetes genes. bioRxiv. doi: https://doi.org/10.1101/108134

Wang, K., 2016 Boosting the Power of the Sequence Kernel Association Test by Properly Estimating Its Null Distribution. *Am J Hum Genet* **99**, 104-114.

Wessel, J., and N. J. Schork, 2006 Generalized genomic distance-based regression methodology for multilocus association analysis. *Am J Hum Genet* **79**, 792–806.

Wright, F., P. Sullivan, A. Brooks, F. Zou, W. Sun *et al.*, 2014 Heritability and genomics of gene expression in peripheral blood. *Nat Genet* **46**, 430-437.

441 Wu, M. C., S. Lee, T. Cai, Y. Li, M. Boehnke, and X. Lin., 2011 Rare-variant association
442 testing for sequencing data with the sequence kernel association test. *Am J Hum Genet*
443 **89**, 82-93.

444 Xu, Z., C. Wu, and W. Pan; Alzheimer's Disease Neuroimaging Initiative, 2017 Imaging-
445 wide association study: Integrating imaging endophenotypes in GWAS. *Neuroimage* **159**,
446 159-169.

447 Zhou, X., P. Carbonetto, and M. Stephens, 2013 Polygenic modeling with bayesian sparse
448 linear mixed models. *PLoS Genetics* **9**, e1003264.

449 Zhu, X., T. Feng, B. O. Tayo, J. Liang, J. H. Young *et al.*, 2015 Meta-analysis of correlated
450 traits via summary statistics from GWASs with an application in hypertension. *Am J*
451 *Hum Genet* **96**, 21-36.

452 Zhu, Z., F. Zhang, H. Hu, A. Bakshi, M. R. Robinson *et al.*, 2016 Integration of summary
453 data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genet*
454 **48**, 481-487.

455 Zou, H.,and T. Hastie, 2005 Regularization and variable selection via the elastic net. *J. R.*
456 *Stat. Soc. Series B* **67**, 301-320.