



Using Genetic Risk Score Approaches to Infer Whether an Environmental Factor Attenuates or Exacerbates the Adverse Influence of a Candidate Gene

Wan-Yu Lin^{1,2*}, Yu-Shun Lin¹, Chang-Chuan Chan^{2,3}, Yu-Li Liu⁴, Shih-Jen Tsai^{5,6,7} and Po-Hsiu Kuo^{1,2}

¹ Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan, ² Department of Public Health, College of Public Health, National Taiwan University, Taipei, Taiwan, ³ Institute of Environmental and Occupational Health Sciences, College of Public Health, National Taiwan University, Taipei, Taiwan, ⁴ Center for Neuropsychiatric Research, National Health Research Institutes, Miaoli, Taiwan, ⁵ Department of Psychiatry, Taipei Veterans General Hospital, Taipei, Taiwan, ⁶ Division of Psychiatry, National Yang-Ming University, Taipei, Taiwan, ⁷ Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan

OPEN ACCESS

Edited by:

Mogens Fenger,
Capital Region of Denmark, Denmark

Reviewed by:

Christine Lary,
Maine Medical Center Research
Institute, United States
Samsiddhi Bhattacharjee,
National Institute of Biomedical
Genomics (NIBMG), India

*Correspondence:

Wan-Yu Lin
linwy@ntu.edu.tw

Specialty section:

This article was submitted to
Statistical Genetics and Methodology,
a section of the journal
Frontiers in Genetics

Received: 05 January 2020

Accepted: 20 March 2020

Published: 08 May 2020

Citation:

Lin W-Y, Lin Y-S, Chan C-C,
Liu Y-L, Tsai S-J and Kuo P-H (2020)
Using Genetic Risk Score Approaches
to Infer Whether an Environmental
Factor Attenuates or Exacerbates
the Adverse Influence of a Candidate
Gene. *Front. Genet.* 11:331.
doi: 10.3389/fgene.2020.00331

Some candidate genes have been robustly reported to be associated with complex traits, such as the *fat mass and obesity-associated (FTO)* gene on body mass index (BMI), and the *fibroblast growth factor 5 (FGF5)* gene on blood pressure levels. It is of interest to know whether an environmental factor (E) can attenuate or exacerbate the adverse influence of a candidate gene. To this end, we here evaluate the performance of “genetic risk score” (GRS) approaches to detect “gene-environment interactions” (G × E). In the first stage, a GRS is calculated according to the genotypes of variants in a candidate gene. In the second stage, we test whether E can significantly modify this GRS effect. This two-stage procedure can not only provide a *p*-value for a G × E test but also guide inferences on how E modifies the adverse effect of a gene. With systematic simulations, we compared several ways to construct a GRS. If E exacerbates the adverse influence of a gene, GRS formed by the elastic net (ENET) or the least absolute shrinkage and selection operator (LASSO) is recommended. However, the performance of ENET or LASSO will be compromised if E attenuates the adverse influence of a gene, and using the ridge regression (RIDGE) can be more powerful in this situation. Applying RIDGE to 18,424 subjects in the Taiwan Biobank, we showed that performing regular exercise can attenuate the adverse influence of the *FTO* gene on four obesity measures: BMI (*p* = 0.0009), body fat percentage (*p* = 0.0031), waist circumference (*p* = 0.0052), and hip circumference (*p* = 0.0001). As another example, we used RIDGE and found the *FGF5* gene has a stronger effect on blood pressure in Han Chinese with a higher waist-to-hip ratio [*p* = 0.0013 for diastolic blood pressure (DBP) and *p* = 0.0027 for systolic blood pressure (SBP)]. This study provides an evaluation on the GRS approaches, which is important to infer whether E attenuates or exacerbates the adverse influence of a candidate gene.

Keywords: body mass index, elastic net regression, gene-environment interaction, lasso, ridge regression

INTRODUCTION

The detection of “gene-environment interactions” ($G \times E$) is important and is even more challenging than the detection of main effects of genes (Greenland, 1983; Hunter, 2005; Aschard, 2016). Although some gene-based $G \times E$ methods have been developed (Jiao et al., 2013; Lin et al., 2013; Chen et al., 2014; Lin et al., 2016; Lin et al., 2019), very few $G \times E$ findings have reached the genome-wide significance level (i.e., $p < \frac{0.05}{20,000} = 2.5 \times 10^{-6}$. Because there are $\sim 20,000$ genes across the genome, 2.5×10^{-6} is the commonly used significance level in genome-wide gene-based analyses) (Chen et al., 2014; Lin et al., 2019). Most evidence of $G \times E$ was discovered by candidate gene analyses in which genome-wide association study (GWAS) hits were targeted. For example, physical activity has been found to attenuate the influence of the *fat mass and obesity-associated (FTO)* gene on obesity risk (Vimalaswaran et al., 2009; Kilpelainen et al., 2011). This means that the association of the *FTO* risk alleles with obesity measures is weaker in physically active subjects than in physically inactive subjects.

Another example is related to blood pressure levels. The *fibroblast growth factor 5 (FGF5)* gene is associated with blood pressure levels in Han Chinese (Liu et al., 2011; Lin et al., 2019). Obesity has been reported to exacerbate the adverse influence of *FGF5* on blood pressure. That is, the association of the *FGF5* risk alleles with blood pressure is stronger in obese subjects than in lean subjects (Li et al., 2015).

A candidate gene usually harbors multiple risk variants. It is of interest to know whether an environmental factor (E) attenuates or exacerbates the effects of these risk variants. However, most $G \times E$ methods provide only p -values without any inference for the direction of interaction (Lin et al., 2013, 2016, 2019; Chen et al., 2014). In contrast, a genetic risk score (GRS) approach aggregates the effects among an ensemble of single-nucleotide polymorphisms (SNPs) and can indicate whether the GRS interacts synergistically or antagonistically with E. A GRS is a linear combination of risk allele counts, where risk alleles and weights are usually retrieved from large published GWASs or meta-analyses. However, the vast majority of genetic studies have been performed on subjects of European ancestry (Sirugo et al., 2019). Appropriate external GWAS results may not be available for $G \times E$ studies in other ethnic populations.

To address this issue, GRS approaches using internal weights have been proposed for pathway-based $G \times E$ studies (Huls et al., 2017; Huls et al., 2017) and GWASs (Lin et al., 2018, 2019, 2020). The internal weights come from marginal effects of SNPs that can be estimated by a multivariate elastic net (ENET) regression (Huls et al., 2017; Huls et al., 2017).

In this study, we described how to use GRS approaches to infer whether E attenuates or exacerbates the adverse influence of a gene (therefore, GRS here aggregates the effects among an ensemble of SNPs in a gene). Moreover, we compared GRS approaches with the “Set-Based gene-EnviRonment InterAction test” (SBERIA) (Jiao et al., 2013), the “interaction sequence kernel association test” (iSKAT) (Lin et al., 2016), and our previously developed “adaptive combination of Bayes factors

method” (ADABF) (Lin et al., 2019). These methods were then applied to the Taiwan Biobank (TWB) data.

MATERIALS AND METHODS

Genetic Risk Score Approaches

Usually, a GRS combines information of multiple nearly independent SNPs across the genome (Ahmad et al., 2013; Rask-Andersen et al., 2017; Lin et al., 2019, 2020). However, there have been few applications of using GRS in gene-based $G \times E$ tests. To have a clear description, we first explain covariate adjustment in the GRS gene-based $G \times E$ approach. Suppose there are L SNPs in a gene or a pathway. Let $g[\cdot]$ be the link function, Y_i be the phenotype that can be either continuous or binary, G_{ij} be the number of minor alleles (0, 1, or 2, by additive genetic model) at the j -th SNP ($j = 1, \dots, L$), E_i be the environmental factor, X_i be the vector of non-genetic covariates such as the age and sex, and the subscript “ i ” represents data for the i -th subject ($i = 1, \dots, n$). In addition to the additive model (by counting the number of minor alleles), G_{ij} can also be coded according to the dominant or recessive genetic model.

Some GRS approaches involve SNP selection and then aggregate the information of selected SNPs. Therefore, we leave G_{ij} ($j = 1, \dots, L$) later and first regress Y_i on X_i by a linear model or a logistic regression model, as follows:

$$g[E(Y_i)] = \alpha_0 + \alpha'X_i, \quad i = 1, \dots, n. \quad (1)$$

Let $\hat{\mu}_{0i} = \hat{\alpha}_0 + \hat{\alpha}'X_i$ (for continuous Y_i) or $\hat{\mu}_{0i} = \frac{\exp(\hat{\alpha}_0 + \hat{\alpha}'X_i)}{1 + \exp(\hat{\alpha}_0 + \hat{\alpha}'X_i)}$ (for binary Y_i) be the predicted mean of Y_i under model (1). Therefore, the covariate-adjusted phenotype for the i -th subject is $\hat{\epsilon}_i = Y_i - \hat{\mu}_{0i}$.

Subsequently, we regress $\hat{\epsilon}_i$ on G_{i1}, \dots, G_{iL} , as follows:

$$g[E(\hat{\epsilon}_i)] = \beta_0 + \sum_{j=1}^L \beta_j G_{ij}, \quad i = 1, \dots, n. \quad (2)$$

Let $\beta' = [\beta_0 \dots \beta_L]$ be the vector of regression coefficients in model (2). The ordinary least squares (OLS) estimate of β is as follows:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left[\sum_{i=1}^n \left(\hat{\epsilon}_i - \beta_0 - \sum_{j=1}^L \beta_j G_{ij} \right)^2 \right] = (\mathbf{G}'\mathbf{G})^{-1} (\mathbf{G}'\hat{\epsilon}), \quad (3)$$

where n is the sample size, \mathbf{G} is the $n \times (L + 1)$ matrix with the i -th row of $[1 \ G_{i1} \dots G_{iL}]$, and $\hat{\epsilon}$ is the n -length vector of covariate-adjusted phenotypes. Because of linkage disequilibrium (LD), the SNPs in a gene are usually highly correlated with each other. In this situation, $\mathbf{G}'\mathbf{G}$ may be singular and not invertible.

Ridge regression (RIDGE) (Hoerl and Kennard, 1970) is used to address the collinearity problem in model (2), where the regression coefficients are estimated by minimizing the residual

sum of squares and an l_2 -norm penalty, as follows:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left[\sum_{i=1}^n \left(\hat{\varepsilon}_i - \beta_0 - \sum_{j=1}^L \beta_j G_{ij} \right)^2 + \lambda \sum_{j=1}^L \beta_j^2 \right]. \quad (4)$$

The regularization parameter λ controls the amount of shrinkage. When λ is close to 0, $\hat{\beta}$ from RIDGE will approximate the $\hat{\beta}$ from OLS. When λ is large, $\hat{\beta}_1, \dots, \hat{\beta}_L$ will approach 0, and model (2) will be reduced to an intercept-only model.

Least absolute shrinkage and selection operator (LASSO) was later proposed to estimate regression coefficients by minimizing the residual sum of squares and an l_1 -norm penalty (Tibshirani, 1996) as follows:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left[\sum_{i=1}^n \left(\hat{\varepsilon}_i - \beta_0 - \sum_{j=1}^L \beta_j G_{ij} \right)^2 + \lambda \sum_{j=1}^L |\beta_j| \right]. \quad (5)$$

The regularization parameter λ controls the amount of shrinkage. Moreover, variable selection can be performed by shrinking some β_j s to 0 ($j = 1, \dots, L$). Because equation (5) selects SNPs that are marginally associated with the covariate-adjusted phenotype ($\hat{\varepsilon}_i, i = 1, \dots, n$), this is called the marginal-association filtering by LASSO. The term “marginal” is used here because the E_i and $G_{ij} \times E_i$ interaction term have not been included in Eq. (5). If β_j is shrunk to 0, the j -th SNP is regarded as unassociated with the covariate-adjusted phenotype and will not be used for the construction of a GRS.

ENET (Zou and Hastie, 2005) strikes a balance between RIDGE and LASSO by estimating regression coefficients while minimizing the residual sum of squares and a mixture of an l_1 -norm and an l_2 -norm, as follows:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^n \left(\hat{\varepsilon}_i - \beta_0 - \sum_{j=1}^L \beta_j G_{ij} \right)^2 + \lambda \sum_{j=1}^L \left[\frac{1}{2} (1 - \alpha) \beta_j^2 + \alpha |\beta_j| \right] \right\}. \quad (6)$$

The regularization parameter λ controls the amount of shrinkage, whereas α is a penalty weight ranging from 0 (RIDGE) to 1 (LASSO). As suggested by Huls et al. (2017), $\alpha = 0.5$ is used for ENET throughout this work to achieve an optimal balance between RIDGE and LASSO. Eq. (6) is the marginal-association filtering by ENET.

Similar to genetic studies using penalized regression approaches (Waldmann et al., 2013; Huls et al., 2017; Huls et al., 2017), we used the R package “glmnet” (Friedman et al., 2010) to obtain $\hat{\beta}$. The optimal values of the regularization parameter λ in Eqs (4–6) were determined by 10-fold cross-validation. GWASs (Waldmann et al., 2013) and pathway-based G × E studies (Huls et al., 2017; Huls et al., 2017) have recommended choosing the largest λ such that the mean squared error (MSE) is within 1 standard error of the minimum MSE, to avoid selecting too

many SNPs in ENET or LASSO. However, most complex traits are polygenic, and a single gene usually explains little phenotypic variation. In our simulations and real data analyses, this criterion usually selected 0 SNPs for ENET or LASSO. Therefore, in gene-based G × E studies, we recommend using the λ that leads to the minimum MSE.

RIDGE, LASSO, and ENET are all techniques for regression models that suffer from multicollinearity. Therefore, the pruning stage to remove SNPs with LD is not required here. After obtaining $\hat{\beta}$ from RIDGE, LASSO, or ENET, the GRS of the i -th subject is constructed by $GRS'_i = \sum_{j=1}^L \hat{\beta}_j G_{ij}$, where G_{ij} is the number of minor alleles (0, 1, or 2) at the j -th SNP of the i -th subject. A positive $\hat{\beta}_j$ indicates that the minor allele is associated with an increase in phenotype. Therefore, a subject with more copies of the minor allele (more phenotype-increasing alleles) will obtain an increase in his/her GRS' . In contrast, a negative $\hat{\beta}_j$ indicates that the minor allele is associated with a decrease in phenotype, and a subject with more copies of the minor allele (more phenotype-decreasing alleles) will obtain a decrease in his/her GRS' . The GRS' is then transformed into a Z-score that represents how many standard deviations the GRS' is from the mean. The standardized GRS Z-score is denoted as GRS_i . A larger GRS is always associated with an increased phenotype.

Afterwards, we fit the following generalized linear model (GLM):

$$g[E(Y_i)] = \gamma_0 + \gamma_G GRS_i + \gamma_E E_i + \gamma_{Int} GRS_i \times E_i + \gamma'_C X_{i,i}, i = 1, \dots, n, \quad (7)$$

where $\gamma_G > 0$ because a larger GRS is always associated with an increased phenotype. Adding a constraint ($\gamma_G > 0$) is expected to improve power, although for simplicity we here perform a GLM without this constraint. By testing $H_0 : \gamma_{Int} = 0$ vs. $H_1 : \gamma_{Int} \neq 0$, we evaluate whether G × E exists. For continuous traits, each 1 standard deviation (SD) increase in GRS is associated with a $\widehat{\gamma}_{Int}$ change in phenotype for subjects exposed to $E = 1$ than for subjects exposed to $E = 0$. For binary traits, each 1 SD increase in GRS is associated with an odds ratio of $\exp(\widehat{\gamma}_{Int})$ for subjects exposed to $E = 1$ compared to subjects exposed to $E = 0$. A positive (negative) $\widehat{\gamma}_{Int}$ indicates that $E = 1$, or a larger E , exacerbates (attenuates) the adverse influence of a candidate gene.

The whole data can be used to estimate β_1, \dots, β_L (Eqs 4–6) and then to test $H_0 : \gamma_{Int} = 0$ vs. $H_1 : \gamma_{Int} \neq 0$ (model 7) without data splitting. Theoretically, $\hat{\beta}_j$ ($j = 1, \dots, L$) and $\widehat{\gamma}_{Int}$ are asymptotically independent under the null hypothesis of no SNP-by-environment interaction, proved by corollary 1 of Dai et al. (2012). A two-stage approach that first filters SNPs by a criterion independent of the test statistic ($\widehat{\gamma}_{Int}$ estimated from model 7) under the null hypothesis, and then only uses SNPs that pass the filter, can maintain type I error rates and boost power (Bourgon et al., 2010; Frost et al., 2016). Empirically, our following simulation studies confirmed that GRS using internal weights is a valid approach in the sense that type I error rates match the nominal significance level. Moreover, studies on genes (Jiao et al., 2013; Frost et al., 2016), pathways

(Huls et al., 2017; Huls et al., 2017), and GWASs (Lin et al., 2018) have presented the validity of using internally weighted GRSs to test for G × E.

Competing Methods

The abovementioned GRS approaches with some form of penalized regression (RIDGE, LASSO, or ENET) were compared with the SBERIA (Jiao et al., 2013), iSKAT (Lin et al., 2016), and ADABF (Lin et al., 2019) methods.

SBERIA is the first gene-based G × E test based on the GRS concept. The phenotype is first regressed on each SNP separately, while adjusting for covariates such as sex, age, and ancestry principal components (PCs), as follows:

$$g[E(Y_i)] = \beta_0 + \beta_j G_{ij} + \beta'_C X_i, i = 1, \dots, n; j = 1, \dots, L. \tag{8}$$

Let $\hat{\beta}_j$ be the estimated regression coefficient of the j -th SNP and p_j be the p -value of testing $H_0 : \beta_j \neq 0$ vs. $H_1 : \beta_j = 0$. The GRS of the i -th subject is constructed by $GRS'_i = \sum_{j=1}^L [I(p_j < 0.1) \text{sign}(\hat{\beta}_j) + v] G_{ij}$, where $I(p_j < 0.1)$ is an indicator variable with a value of 1 if $p_j < 0.1$ and 0 otherwise, $\text{sign}(\hat{\beta}_j)$ is either 1 or -1 depending on the sign of $\hat{\beta}_j$, and v is a very small value (e.g., 0.0001). Therefore, SBERIA GRS includes both SNP selection (only SNPs marginally associated with the phenotype are used to construct the GRS) and SNP weighting (only the directions of how SNPs influence the phenotype are used in the GRS). The standardized GRS Z-score is denoted as GRS_i . Then, the following GLM is fitted:

$$g[E(Y_i)] = \gamma_0 + \sum_{j=1}^L \gamma_{G_j} G_{ij} + \gamma_E E_i + \gamma_{Int} GRS_i \times E_i + \gamma'_C X_i, \tag{9}$$

$i = 1, \dots, n; j = 1, \dots, L.$

By testing $H_0 : \gamma_{Int} = 0$ vs. $H_1 : \gamma_{Int} \neq 0$, we can evaluate whether G × E exists (Jiao et al., 2013).

There are two fundamental differences between SBERIA and ENET (or LASSO). First, in the marginal-association filtering stage, SBERIA fits L regression models, respectively (Eq. 8), whereas ENET and LASSO fit a multivariate model incorporating L SNPs simultaneously (Eqs 5 and 6). Second, in the model for testing GRS × E, SBERIA includes the main effects of all the L SNPs ($\sum_{j=1}^L \gamma_{G_j} G_{ij}$ in Eq. 9), whereas ENET and LASSO incorporate only a GRS term as the aggregated genetic main effects ($\gamma_{G} GRS_i$ in Eq. 7).

We compared the abovementioned GRS tests with iSKAT (Lin et al., 2016) and ADABF (Lin et al., 2019). In iSKAT, the following model is considered:

$$g[E(Y_i)] = \delta_0 + \sum_{j=1}^L \delta_{G_j} G_{ij} + \delta_E E_i + \sum_{j=1}^L \delta_{Int_j} G_{ij} E_i + \delta'_C X_i, i = 1, \dots, n, \tag{10}$$

where δ_{Int_j} is the interaction effect between the j -th SNP and E. Assuming $\delta_{Int_j} s(j = 1, \dots, L)$ follow a distribution with a mean of 0 and a variance of τ , the null hypothesis of all

$\delta_{Int_j} s = 0 (j = 1, \dots, L)$ can be reduced to $\tau = 0$. The iSKAT is a score statistic for testing the variance component, i.e., $H_0 : \tau = 0$ vs. $H_1 : \tau > 0$. The statistic can be referred to as Eq. (6) in Lin et al. (2016). The iSKAT method is regarded as optimal in the class of variance component tests (Lin et al., 2013; Chen et al., 2014). Therefore, we chose iSKAT as the representative of variance component tests.

In ADABF, we consider the following model for the j -th SNP ($j = 1, \dots, L$):

$$g[E(Y_i)] = \delta_0 + \delta_{G_j} G_{ij} + \delta_E E_i + \delta_{Int_j} G_{ij} E_i + \delta'_C X_i, \tag{11}$$

$i = 1, \dots, n.$

The SNP × E interaction is of interest, and therefore, $H_0 : \delta_{Int_j} = 0$ vs. $H_1 : \delta_{Int_j} \neq 0$ for the j -th SNP ($j = 1, \dots, L$). A Bayes factor (BF) was calculated for each SNP × E, where $BF = Pr(Data|H_1)/Pr(Data|H_0)$. A larger BF indicates that the relative evidence in favor of H_1 (SNP × E interaction exists) is stronger. Because the number of SNPs exhibiting interactions with E varies gene by gene, ADABF exhaustively searches for the evidence of G × E interaction by considering the largest BF, combining the largest 2 BFs, combining the largest 3 BFs, ..., to aggregating all the L BFs in the gene. The significance of G × E interaction is then determined by the efficient sequential resampling procedure (Liu et al., 2016). ADABF was selected for comparison because it was recommended as a powerful and robust gene-based G × E test in a recent study (Lin et al., 2019). The iSKAT and ADABF methods do not test for G × E through a GRS term. Therefore, these two methods do not make inference for the direction of G × E.

Simulation Study

To reflect the real LD structures of the human genome, we used the genotypes of 18,424 TWB subjects as our simulation material. Three genes (i.e., *TNNT3*, *RFX3*, and *FTO*) were drawn for simulations, including 48, 95, and 242 SNPs, respectively. It is common to see multiple trait-associated SNPs to be included in the same gene (Willer et al., 2007; Fawcett and Barroso, 2010). Therefore, we randomly specified 4 trait-associated SNPs in each simulation replicate. We considered binary and continuous exposures, respectively. For binary exposures, E was randomly sampled from 1 (exposed) or 0 (non-exposed), with $P(E = 1) = 0.2$ and $P(E = 1) = 0.5$, respectively. For continuous exposures, E was randomly selected from the normal distribution with a mean of 0 and a standard deviation of 0.5.

The continuous trait of the i -th subject was simulated as follows:

$$Y_i = \sum_{d=1}^4 \beta_{G_d} G_{id} + \beta_E E_i + \sum_{d=1}^D \beta_{Int_d} G_{id} E_i + \varepsilon_i, \tag{12}$$

where β_{G_d} is the SNP main effect of the d -th trait-associated SNP ($d = 1, \dots, 4$), β_{Int_d} is the effect size of interaction between the d -th trait-associated SNP and E ($d = 1, \dots, D$), D is the number of trait-associated SNPs that also exhibit interactions

with E ($D = 4$ or 2 in our simulation), and ε_i is the random error term following the standard normal distribution.

The binary trait of the i -th subject was simulated as follows:

$$\log \frac{P(Y_i = 1)}{1 - P(Y_i = 1)} = \beta_0 + \sum_{d=1}^4 \beta_{G_d} G_{id} + \beta_E E_i + \sum_{d=1}^D \beta_{Int_d} G_{id} E_i, \quad (13)$$

where the intercept β_0 was $\log(0.1/0.9) = -2.2$ or $\log(0.4/0.6) = -0.4$. $Y_i = 1$ represents that the i -th subject was diseased whereas $Y_i = 0$ indicates that he/she was non-diseased. This setting corresponds to a disease prevalence of 10% or 40%. A disease prevalence of 40% is also a reasonable setting for some complex diseases. For example, the worldwide prevalence of hypertension among adults aged ≥ 25 years was $\sim 40\%$ (Abebe et al., 2015).

The magnitudes of SNP main effects, $|\beta_{G_d}|$ ($d = 1, \dots, 4$), were uniformly sampled from 0.04 to 0.08 for continuous traits or uniformly sampled from $\log(1.05)$ to $\log(1.15)$ for binary traits. This simulation setting for binary trait (odds ratios uniformly sampled from 1.05 to 1.15) concurs with broad GWAS findings, where most odds ratios < 1.5 and many odds ratios < 1.2 (Baxter and Jordan, 2012; Hodge and Greenberg, 2016).

Unlike small effect sizes for SNPs, the effect size of E was assumed to be larger, because some E is critical to traits (e.g., regular exercise and dietary habits are important to obesity measures). The magnitude of E , $|\beta_E|$, was assigned to be 0.3 for continuous traits and $\log(1.3) = 0.2624$ for binary traits, respectively. When evaluating type I error rates, data have to be generated from the null hypothesis of no $SNP \times E$ interactions, and therefore we specified all $\beta_{Int_d} = 0$ ($d = 1, \dots, D$). When assessing power, the magnitudes of $SNP \times E$ interaction effects, $|\beta_{Int_d}|$ ($d = 1, \dots, D$), were uniformly sampled from 0.04 to 0.08 for continuous traits or uniformly sampled from $\log(1.05)$ to $\log(1.15)$ for binary traits. To not favor GRS-based approaches that construct GRSs with SNPs marginal effects, the sampling of $|\beta_{Int_d}|$ ($d = 1, \dots, D$) was independent of the sampling of $|\beta_{G_d}|$ ($d = 1, \dots, 4$). Power was evaluated under 14 simulation scenarios listed in **Table 1**, in which “+” indicates a positive effect and “-” means a negative effect. As summarized by **Table 1**, we always assumed 4 trait-associated SNPs, but the number of trait-associated SNPs that also exhibited interactions with E could be 2 or 4. This setting mimics the common situation that some trait-associated SNPs may also present interactions with E (Jonsson et al., 2009; Chen et al., 2014; Li et al., 2015).

Scenarios 1–7 describe that a larger E is associated with an increase in trait, whereas scenarios 8–14 represent that a larger E is associated with a decrease in trait. Suppose that a higher trait is linked to a less healthy situation, e.g., $Y_i = 1$ in binary traits. Scenarios 1, 3, 6, 8, 10, and 13 indicate that a larger E exacerbates the adverse effect of a candidate gene, whereas scenarios 2, 4, 7, 9, 11, and 14 present that a larger E attenuates the adverse effect of a candidate gene. Scenarios 5 and 12 are cross-over situations, representing that a larger E exacerbates the adverse effect of 50% of trait-associated SNPs while attenuating the adverse effect of the remaining 50% of trait-associated SNPs.

Application to the Taiwan Biobank Data

The TWB aims to build a research database that integrates the genomic profiles and lifestyle patterns of residents aged 30–70 years in Taiwan (Chen et al., 2016). Community-based volunteers provided blood samples and a range of information via a face-to-face interview and physical examination. This study included 20,287 TWB individuals and was approved by the Research Ethics Committee of the National Taiwan University Hospital (NTUH-REC No. 201805050RINB). To remove cryptic relatedness, we used PLINK 1.9 (Purcell et al., 2007) to calculate the genome-wide identity by descent (IBD) sharing coefficients between any two subjects. Similar to many genetic studies (Lowe et al., 2009; Mok et al., 2014; Ombrello et al., 2014), we excluded one subject from each pair with $PI-HAT \geq 0.125$, where $PI-HAT = \text{Probability}(\text{IBD} = 2) + 0.5 \times \text{Probability}(\text{IBD} = 1)$. Through this step, relatives within the third-degree of consanguinity were removed. Finally, 18,424 unrelated subjects (9,093 males and 9,331 females) were remained in the following analysis. **Table 2** shows basic characteristics of TWB participants stratified by sex.

The majority of TWB subjects were of Han Chinese ancestry (Chen et al., 2016). The Axiom Genome-Wide TWB genotyping array was designed for Han Chinese in Taiwan, which was run on the Axiom Genome-Wide Array Plate System (Affymetrix, Santa Clara, CA, United States). A total of 646,783 autosomal SNPs were genotyped in this TWB array. After removing 51,293 SNPs with genotyping rates $< 95\%$, 6,095 SNPs with Hardy-Weinberg test p -values $< 5.7 \times 10^{-7}$ (WTCCC, 2007), and 1,869 variants with minor allele frequencies (MAFs) $< 1\%$, 587,526 SNPs were used to construct ancestry PCs. Because variants with MAF $< 1\%$ have been excluded, no rare variants were included in our following analyses. Removing variants with MAFs $< 1\%$ is a commonly used quality control step in genetic association studies, because the chances of errors in genotype calling increase with decreasing MAFs (Goldstein et al., 2012; Coleman et al., 2016).

RESULTS

Type I Error Rates

To evaluate type I error rates, continuous traits and binary traits were simulated according to models (12) and (13), respectively. Scenarios 1 and 8 in **Table 1** were simulated, but all β_{Int_d} ($d = 1, \dots, 4$) have to be set at 0 in order to evaluate type I error rates. The main effects of SNPs, $|\beta_{G_d}|$ ($d = 1, \dots, 4$), and the environmental factor, $|\beta_E|$, have been described in section “Simulation study.” Based on 10,000 replications for each scenario, **Figure 1** and **Supplementary Figure S1** show that all methods were valid in the sense that their type I error rates matched the nominal significance level.

Power of 6 $G \times E$ Methods

We compared the power of the 4 GRS-based tests and 2 $G \times E$ methods: iSKAT (Lin et al., 2016) and ADABF (Lin et al., 2019). The power of the 4 GRS-based tests is defined as the probability of rejecting $H_0: \gamma_{Int} = 0$ (p -value < 0.05) and correctly specifying the sign of γ_{Int} (γ_{Int} can be found from models 7 and 9). The

TABLE 1 | The 14 simulation scenarios for power comparison, where “exacerbation” and “attenuation” mean that $E = 1$ (or a larger continuous E) exacerbates or attenuates the adverse effect of a candidate gene, respectively.

Scenario	E	β_{G_1}	β_{G_2}	β_{G_3}	β_{G_4}	β_{Int_1}	β_{Int_2}	β_{Int_3}	β_{Int_4}
1 Exacerbation	+	+	+	+	+	+	+	+	+
2 Attenuation	+	+	+	+	+	-	-	-	-
3 Exacerbation	+	+	+	+	+	+	+	0	0
4 Attenuation	+	+	+	+	+	-	-	0	0
5 Cross-over	+	+	+	+	+	+	+	-	-
6 Exacerbation	+	+	+	-	-	+	+	-	-
7 Attenuation	+	+	+	-	-	-	-	+	+
8 Exacerbation	-	+	+	+	+	+	+	+	+
9 Attenuation	-	+	+	+	+	-	-	-	-
10 Exacerbation	-	+	+	+	+	+	+	0	0
11 Attenuation	-	+	+	+	+	-	-	0	0
12 Cross-over	-	+	+	+	+	+	+	-	-
13 Exacerbation	-	+	+	-	-	+	+	-	-
14 Attenuation	-	+	+	-	-	-	-	+	+

TABLE 2 | Basic characteristics of TWB participants stratified by sex.

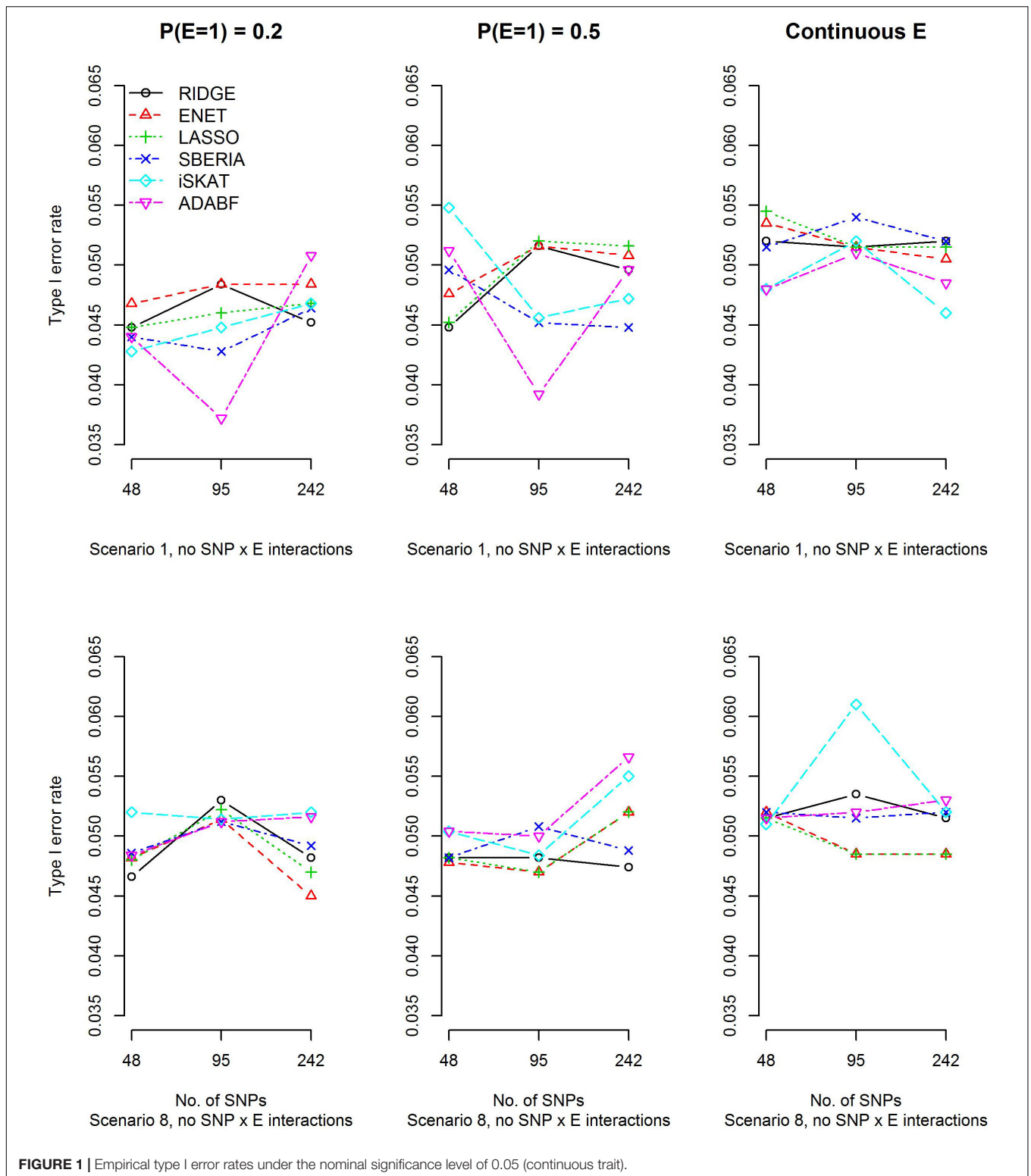
	Overall	Males	Females
Total, n (%)	18,424	9,093	9,331
Age (years), mean (SD)	48.9 (11.0)	49.0 (11.0)	48.9 (10.9)
Smoking, n (%)	2,134 (11.6)	1,882 (20.7)	252 (2.7)
Drinking, n (%)	1,345 (7.3)	1,178 (13.0)	167 (1.8)
Regular exercise, n (%)	7,652 (41.5)	3,896 (42.8)	3,756 (40.3)
Educational attainment, mean (SD)	5.46 (0.99)	5.62 (0.92)	5.29 (1.02)
BMI (kg/m^2), mean (SD)	24.31 (3.66)	25.2 (3.4)	23.5 (3.7)
Body fat %, mean (SD)	27.29 (7.38)	22.9 (5.5)	31.5 (6.5)
Waist circumference (cm), mean (SD)	83.93 (10.03)	87.4 (9.1)	80.5 (9.7)
Hip circumference (cm), mean (SD)	96.34 (6.90)	97.6 (6.5)	95.2 (7.0)
Waist-hip ratio, mean (SD)	0.87 (0.068)	0.90 (0.06)	0.85 (0.07)
Diastolic blood pressure (mmHg), mean (SD)	73.11 (11.10)	76.9 (10.6)	69.4 (10.3)
Systolic blood pressure (mmHg), mean (SD)	117.62 (17.37)	121.9 (16.1)	113.5 (17.6)

iSKAT (Lin et al., 2016) and ADABF (Lin et al., 2019) methods can only provide a p -value for testing $G \times E$, without an inference of how E modifies the genetic effects. Therefore, their power is defined as the probability of rejecting the null hypothesis of no $G \times E$ (p -value < 0.05). **Figure 2** presents the power based on 1,000 simulation replications under each scenario, for continuous traits and $P(E = 1) = 0.2$. If E exacerbates the adverse effect of a gene, ENET and LASSO outperform the other methods (scenarios 1, 3, 6, 8, 10, and 13). If E attenuates the adverse effect of a gene, RIDGE is more powerful (scenarios 2, 4, 7, 9, 11, and 14). Moreover, ADABF was the optimal test under the cross-over scenario (scenarios 5 and 12).

Among the 4 GRS-based tests, ENET, LASSO, and SBERIA first select SNPs marginally associated with the phenotype. ENET and LASSO select SNPs according to the multivariate ENET regression (Eq. 6) and multivariate LASSO regression (Eq. 5), respectively. SBERIA selects SNPs based on regressions considering one SNP at a time (Eq. 8). The performance of these three methods to detect $G \times E$ depends on their ability to find the true trait-associated SNPs. **Supplementary Figures S2, S3**

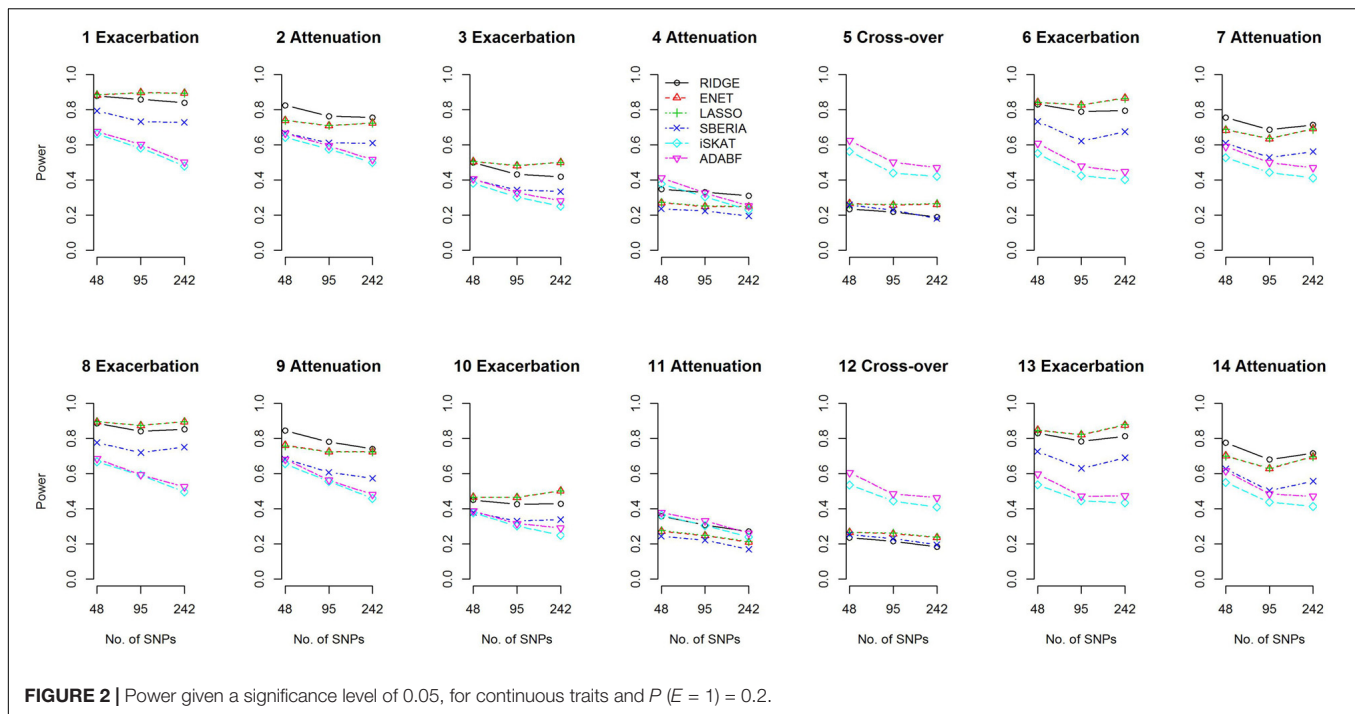
summarize the sensitivity (SEN) and positive predictive value (PPV) of the marginal-association filtering in ENET, LASSO, and SBERIA. As described in Eq. (12) and **Table 1**, 4 trait-associated SNPs were randomly assigned in each simulation replication. SEN is defined as the percentage of true positives among the 4 trait-associated SNPs, whereas PPV is defined as the percentage of true positives among the total findings.

Because SBERIA takes a liberal p -value threshold of 0.1, in the filtering stage it usually selects more SNPs than ENET and LASSO do. Therefore, as shown in **Supplementary Figures S2, S3**, SBERIA generally has a higher SEN but a lower PPV, compared with ENET and LASSO. For each method, SEN was much lower in attenuation scenarios than in exacerbation scenarios. This means that pinpointing true trait-associated SNPs is more difficult in attenuation scenarios. In the filtering stage (Eqs 5, 6, and 8), $G_{id} \times E_i$ cannot be included; therefore, a β_{Int_d} in the opposite direction to β_{G_d} will weaken the magnitude of the marginal effect of the d -th trait-associated SNP. In contrast, a β_{Int_d} in the same direction to β_{G_d} will strengthen the magnitude of the marginal effect of the d -th trait-associated SNP. Therefore,



for each method, the ability to pinpoint true trait-associated SNPs is inferior in attenuation scenarios than in exacerbation scenarios (as shown in **Supplementary Figures S2, S16** for continuous traits and binary traits, respectively).

Because there is no SNP selection in RIDGE, true trait-associated SNPs are all reserved for constructing GRS. Therefore, RIDGE is the optimal GRS-based test in attenuation scenarios (**Figure 2**). In exacerbation scenarios, ENET and LASSO are the



best two methods because of their superior ability in finding trait-associated SNPs (**Supplementary Figures S2, S16**).

We then investigated the percentage of sign-misspecifications for each GRS-based method. In the presence of $G \times E$, we calculated the percentage of wrongly specifying the sign of γ_{Int} (in models 7 or 9) among all rejections of $H_0: \gamma_{Int} = 0$ (p -value < 0.05). With 1,000 simulation replications under each scenario, **Figure 3** presents the percentages of sign-misspecifications, for continuous traits and $P(E = 1) = 0.2$. The true γ_{Int} is positive under scenarios 1, 3, 6, 8, 10, and 13; negative under scenarios 2, 4, 7, 9, 11, and 14. The cross-over scenarios (5 and 12) were not considered here, because the true sign of γ_{Int} was unclear when $E = 1$ (or a larger continuous E) exacerbates the adverse effect of 50% of trait-associated SNPs but attenuates the adverse effect of the remaining 50% of trait-associated SNPs.

Figure 3 shows that the signs of γ_{Int} were usually correctly specified except scenarios 4 and 11. SBERIA generally makes more mistakes in indicating the true direction for γ_{Int} , because it filters SNPs by considering only one SNP at a time (Eq. 8). **Supplementary Figures S4–S7** present the power and percentages of sign-misspecifications in continuous-trait simulations, given $P(E = 1) = 0.5$ and a continuous E , respectively. Each method under $P(E = 1) = 0.5$ was slightly more powerful than that under $P(E = 1) = 0.2$, but the comparisons across methods were similar to those shown in **Figures 2, 3**.

The performance of each method for binary-trait simulations can be found in **Supplementary Figures S8–S21**. Regarding different prevalence values, each method was more powerful under $P(Y = 1) = 0.4$ than under $P(Y = 1) = 0.1$. Concerning different probabilities of being exposed, each method was more powerful under $P(E = 1) = 0.5$ than under $P(E = 1) = 0.2$.

Nonetheless, the comparisons across methods were similar to those described for continuous traits.

Regarding the 14 simulation scenarios listed in **Table 1**, only 2 SNP $\times E$ interactions (rather than 4) were simulated in scenarios 4 and 11, and β_{Int} s were in the opposite direction to β_G s. Therefore, detecting $G \times E$ and correctly specifying the sign for γ_{Int} were more challenging under scenarios 4 and 11.

Regarding the computation time (**Supplementary Figures S22–S29**), SBERIA is the fastest method, followed by the three penalized regression methods: RIDGE, ENET, and LASSO. ADABF uses the sequential resampling procedure (Liu et al., 2016), and therefore, it takes a longer time in the presence of $G \times E$. A recent study concluded that ADABF is a computationally feasible method in a GWAS context (Lin et al., 2019). This is because in a real GWAS usually very few genes can be detected to interact with E . Therefore, a genome-wide analysis using ADABF takes a much shorter time than the power evaluation performed here (power is always evaluated in the presence of $G \times E$). On average, the iSKAT method requires the longest time in most situations (**Supplementary Figures S22–S29**).

***FTO* \times Exercise Interaction on Five Obesity Measures**

The *FTO* gene is obesity-associated and has been replicated by many studies (Frayling et al., 2007; Scuteri et al., 2007; Fawcett and Barroso, 2010). It spans from 53, 737, 875 to 54, 148, 379 base pairs on chromosome 16, according to the human genome GRCh37/hg19 assembly. Similar to many gene-based tests (Liu et al., 2010; Wray et al., 2012; Alonso-Gonzalez et al., 2019; Lin et al., 2019), 50 kb in the 3' and 5' regions that might regulate

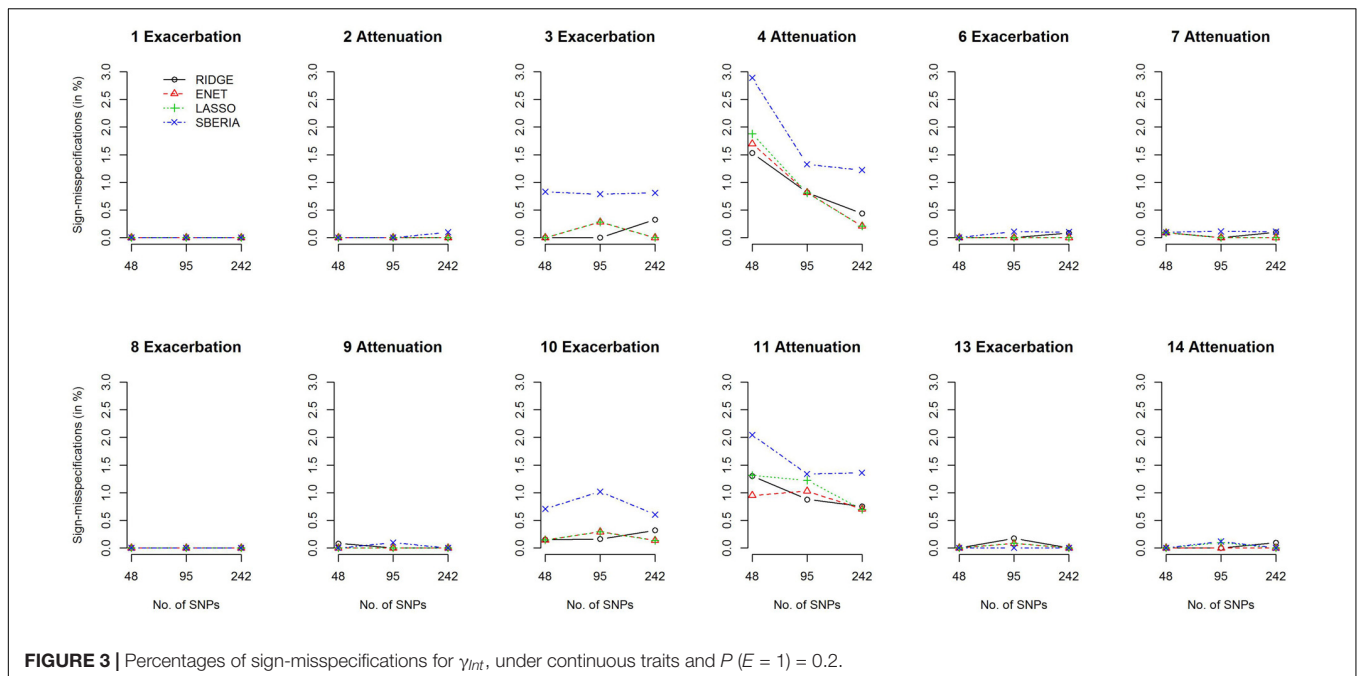


FIGURE 3 | Percentages of sign-misspecifications for γ_{Int} , under continuous traits and $P(E = 1) = 0.2$.

a gene were also incorporated in our analysis. Specifying a large boundary would be difficult to pinpoint the exact gene interacting with E, whereas a small boundary may not fully capture the regulatory regions (Liu et al., 2010). A total of 242 SNPs (all with $MAF > 1\%$) were included in this chromosomal region. Here, we used the 6 $G \times E$ tests to investigate whether the influence of *FTO* can be modified by performing regular physical exercise. A total of five obesity measures were analyzed: body mass index (BMI), body fat percentage (BFP), waist circumference (WC), hip circumference (HC), and the waist-to-hip ratio (WHR).

Regular physical exercise was defined as engaging in 30 min of “exercise” three times a week. “Exercise” indicated leisure-time activities such as jogging, yoga, mountain climbing, or playing basketball. According to the TWB questionnaire, activities during work were not counted in “exercise.” Among the 18,424 subjects, 7,652 (41.5%) reported performing regular exercise, while 10,764 reported no regular exercise. A total of eight subjects did not respond to this question. Covariates adjusted in all models included sex, age (in years), educational attainment, drinking status, smoking status, and the first 10 ancestry PCs.

Table 3 presents the results of six methods. We may miss possibly important findings due to a harsh penalty of multiple testing. Because this study focuses on $G \times E$ detection for candidate genes, we set the significance level at 0.05. The results of RIDGE show that regular physical exercise attenuates the adverse effect of the *FTO* gene on BMI, BFP, WC, and HC. The other three GRS-based tests (ENET, LASSO, and SBERIA) provided significant results for three obesity measures, respectively. All $\widehat{\gamma}_{Int}$ s were negative, indicating that regularly performing exercise attenuates the adverse influence of *FTO* on obesity measures.

Genetic risk score-based tests are more powerful than iSKAT and ADABF because performing regular exercise generally blunted the effects of the trait-increasing alleles in *FTO*. Take

BMI-associated SNPs as examples. A total of 12 out of the 242 SNPs reached the genome-wide significance (p -value $< 5 \times 10^{-8}$) when testing $H_0 : \beta_j = 0$ vs. $H_1 : \beta_j \neq 0$ in the following model:

$$E(Y_i) = \beta_0 + \beta_j G_{ij} + \beta'_C X_i, \quad (14)$$

Where Y_i is BMI, G_{ij} is the number of minor alleles (0, 1, or 2) at the j -th SNP ($j = 1, \dots, 242$), and X_i is the vector of covariates of the i -th subject ($i = 1, \dots, 18,424$). All $\widehat{\beta}_j$ s of the 12 genome-wide significant SNPs were positive, representing that their minor alleles were associated with increased BMIs.

When exercise ($E_i = 1$ if yes; $E_i = 0$ if no) and the SNP \times exercise interaction were included in the 12 models, as follows:

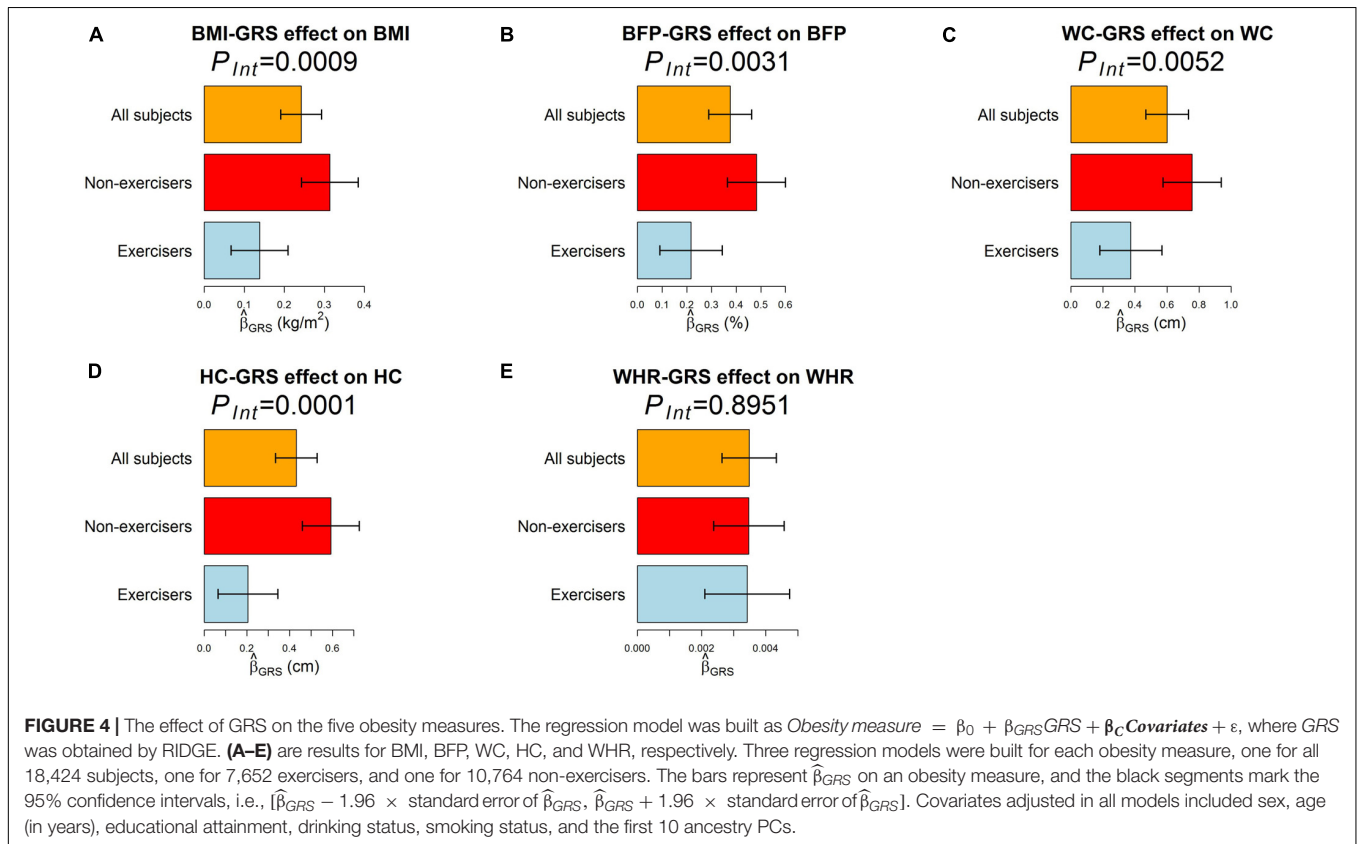
$$E(Y_i) = \beta_0 + \beta_j G_{ij} + \beta_E E_i + \beta_{Intj} G_{ij} \times E_i + \beta'_C X_i, \quad j = 1, \dots, 12, \quad (15)$$

all $\widehat{\beta}_j$ s were positive, and all $\widehat{\beta}_{Intj}$ s were negative ($j = 1, \dots, 12$). This indicated that the minor alleles of the 12 SNPs were associated with increased BMIs, but their effects were blunted by performing regular exercise. As mentioned above, a β_{Intj} in the opposite direction to β_j weakens the magnitude of marginal effect of this SNP, and therefore $\widehat{\beta}_j$ from model (15) $>$ $\widehat{\beta}_j$ from model (14). This is a situation similar to our simulation scenario 9 in **Table 1**, because exercise is associated with a decrease in obesity measures. According to **Figure 2** (scenario 9, 242 SNPs), RIDGE is the most powerful test. In real data, the *FTO* gene harbors more BMI-associated SNPs than our simulation scenario 9, and these SNPs interact with E in the same direction. Therefore, GRS-based tests are more powerful than iSKAT and ADABF.

TABLE 3 | *FTO* × exercise interaction on five obesity measures.

Trait		RIDGE	ENET	LASSO	SBERIA	iSKAT	ADABF
BMI (kg/m ²)	$\widehat{\gamma}_{Int}$	-0.1743	-0.0821	-0.0964	-0.1482		
	P_{Int}	0.0009	0.1192	0.0671	0.0067	0.2043	0.1700
Body fat %	$\widehat{\gamma}_{Int}$	-0.2661	-0.2069	-0.2081	-0.2259		
	P_{Int}	0.0031	0.0212	0.0205	0.0160	0.2430	0.2200
Waist circumference (cm)	$\widehat{\gamma}_{Int}$	-0.3854	-0.3719	-0.3760	-0.2786		
	P_{Int}	0.0052	0.0069	0.0063	0.0512	0.5369	0.3700
Hip circumference (cm)	$\widehat{\gamma}_{Int}$	-0.3868	-0.3286	-0.3291	-0.2902		
	P_{Int}	0.0001	0.0011	0.0011	0.0055	0.5061	0.3300
Waist-to-hip ratio	$\widehat{\gamma}_{Int}$	-0.000116	-0.000775	-0.000374	-0.000314		
	P_{Int}	0.8951	0.3773	0.6699	0.7308	0.7994	0.3100

Significant results with $P_{Int} < 0.05$ are highlighted in bold values. Negative $\widehat{\gamma}_{Int}$ s indicate that regular physical exercise attenuates the adverse influence of *FTO* on obesity measures.



According to RIDGE, each 1 SD increase in BMI-GRS was associated with a 0.1743 kg/m² lower BMI in exercisers than in non-exercisers ($p = 0.0009$, Table 3). Each 1 SD increase in BFP-GRS was associated with a 0.2661% lower BFP in exercisers than in non-exercisers ($p = 0.0031$). Each 1 SD increase in WC-GRS was associated with a 0.3854 cm lower WC in exercisers than in non-exercisers ($p = 0.0052$). Each 1 SD increase in HC-GRS was associated with a 0.3868 cm lower HC in exercisers than in non-exercisers ($p = 0.0001$). As shown by Figure 4, except WHR, GRS effect on each obesity measure was smaller in exercisers than in non-exercisers. For each obesity measure, $\widehat{\gamma}_{Int}$ in Table 3 is approximately equal to the length of blue bar – the length of red

bar in Figure 4. In the whole study population, each 1 SD increase in BMI-GRS was associated with a 0.2420 kg/m² (the length of orange bar in Figure 4A) higher BMI ($P = 1.1 \times 10^{-20}$). This association was stronger in non-exercisers than in exercisers ($P_{Int} = 0.0009$, Table 3). In non-exercisers, each 1 SD increase in BMI-GRS was associated with a 0.3136 kg/m² (the length of red bar in Figure 4A) higher BMI ($P = 3.4 \times 10^{-18}$); in exercisers, each 1 SD increase in BMI-GRS was associated with a 0.1382 kg/m² (the length of blue bar in Figure 4A) higher BMI ($P = 1.4 \times 10^{-4}$).

Performing regular exercises is associated with an attenuation of the adverse influence of the *FTO* gene on both WC and HC,

but not on WHR because it is a ratio of WC to HC. This negative result for WHR is in line with the result from polygenic analyses (i.e., aggregating information of multiple nearly independent SNPs across the genome) (Lin et al., 2019).

FGF5 × WHR Interaction on Blood Pressure Levels

In addition to binary exposures, we also provide an example for continuous exposures. The *FGF5* gene is associated with blood pressure levels in Han Chinese (Liu et al., 2011; Lin et al., 2019). *FGF5* spans from 81, 187, 742 to 81, 212, 171 base pairs on chromosome 4, according to the human genome GRCh37/hg19 assembly. Similar to many gene-based tests (Liu et al., 2010), our analysis region also included 50 kb in the 3' and 5' regions that might regulate the gene. A total of 38 SNPs (all with MAF > 1%) were included in this chromosomal region. It has been known that central obesity is a risk factor for hypertension (Parrinello et al., 1996). Therefore, we used the 6 G × E tests to investigate whether the influence of *FGF5* can be modified by an indicator of central obesity, WHR. Covariates adjusted in all models included sex, age (in years), drinking status, smoking status, and the first 10 ancestry PCs. As described above, a total of 18,424 unrelated TWB subjects were included in our analysis, where 9,093 were males and 9,331 were females.

Table 4 presents the results of six methods. All six tests suggested the significance of *FGF5* × WHR interaction on DBP and SBP ($P_{Int} < 0.05$). Positive $\widehat{\gamma}_{Int}$ s from GRS-based tests represent that a higher WHR exacerbates the adverse influence of *FGF5*.

The above two examples demonstrate using GRS-based methods to infer whether E attenuates or exacerbates the adverse influence of a candidate gene (*FTO* and *FGF5*, respectively).

DISCUSSION

Detecting G × E has been an important but challenging issue. Although set-based G × E tests have been evaluated in a genome-wide context, very few genes have been found to interact with some exposures at the genome-wide significance level (i.e., $p < \frac{0.05}{20000} = 2.5 \times 10^{-6}$. Because there are ~20,000 genes across the genome, 2.5×10^{-6} is the commonly used significance level in genome-wide gene-based analyses) (Chen et al., 2014; Lin et al., 2019). Searching for evidence of G × E for candidate genes may still be a more feasible strategy.

It is common to see a GRS that aggregates information of multiple nearly independent SNPs across the genome (Ahmad et al., 2013; Rask-Andersen et al., 2017; Lin et al., 2019, 2020). However, there have been few applications of using GRS in gene-based G × E tests. In this work, we evaluate the performance of GRS gene-based G × E approach, with simulations and real applications. GRS-based tests can outperform other methods if interactions tend to go in the same direction (Aschard, 2016). The filtering stage of SBERIA overlooks LD among SNPs because it considers L SNPs in L respective regressions (Eq. 8). RIDGE has the advantage of accommodating LD among SNPs and providing regression coefficients that lead to minimum MSE. Although RIDGE has been used to address LD in genetic association analysis (Malo et al., 2008), our study is the first attempt to use the ridge regression to construct a GRS for G × E analyses. ENET and LASSO can not only accommodate LD among SNPs, but they also select SNPs by shrinking some regression coefficients to 0.

Although ENET has been suggested as a powerful G × E test in pathway analyses (Huls et al., 2017; Huls et al., 2017), in this study, we showed that its performance to find trait-associated SNPs can be compromised if E attenuates the adverse effect of a gene (simulation scenarios 2, 4, 7, 9, 11, and 14). As a result, RIDGE is recommended if E attenuates the adverse effects of most SNPs in a gene, as shown in the application of *FTO* × exercise interaction on obesity measures (**Table 3**). In contrast, if E exacerbates the adverse effects of most SNPs in a gene, ENET and LASSO will be more powerful, as shown in the application of *FGF5* × WHR interaction on SBP (**Table 4**).

In this work, we described how to use GRS approaches to infer whether E attenuates or exacerbates the adverse effect of a gene. GRS approaches can not only provide a p -value for G × E but also infer how E interacts with the gene. The signs of γ_{Int} were usually correctly specified except scenarios 4 and 11, where only 2 SNP × E interactions were specified and their interaction effects were in the opposite direction to the SNP main effects. SBERIA usually makes more mistakes because its filtering stage overlooks LD among SNPs by considering L SNPs in L respective regressions (Eq. 8).

In the cross-over situations (simulation scenarios 5 and 12), where E exacerbates the adverse effects of some SNPs but attenuates the adverse effects of others, GRS approaches were underpowered and ADABF became the most useful. None of the six methods could outperform the others across all situations. In this work, we performed simulations to explore the relative performances of these six tests in detecting G × E and correctly indicating the interaction direction.

TABLE 4 | *FGF5* × WHR interaction on blood pressure levels.

Trait		RIDGE	ENET	LASSO	SBERIA	iSKAT	ADABF
DBP (mmHg)	$\widehat{\gamma}_{Int}$	0.2419	0.1980	0.2141	0.2378		
	P_{Int}	0.0013	0.0082	0.0042	0.0014	0.0154	0.0096
SBP (mmHg)	$\widehat{\gamma}_{Int}$	0.3396	0.3548	0.3551	0.3261		
	P_{Int}	0.0027	0.0017	0.0017	0.0039	0.0482	0.0480

Significant results with $P_{Int} < 0.05$ are highlighted in bold values. Positive $\widehat{\gamma}_{Int}$ s indicate that a higher WHR exacerbates the adverse influence of *FGF5* on blood pressure levels.

To show the utility of GRS-based approaches, we provided an example for binary exposure and an example for continuous exposure: (1) *FTO* × exercise interaction on five obesity measures (Table 3); (2) *FGF5* × WHR interaction on blood pressure levels (Table 4). The γ_{IntS} in Table 3 were consistently negative, representing that performing regular exercise attenuates the adverse influence of the *FTO* gene on four obesity measures. Moreover, γ_{IntS} in Table 4 were consistently positive, indicating that a higher WHR exacerbates the adverse effect of the *FGF5* gene on blood pressure levels.

This work provides contributions to an important issue: identify whether E attenuates or exacerbates the adverse influence of a candidate gene. Genes exhibiting interactions with E are usually difficult to detect and replicate (McAllister et al., 2017), especially at the genome-wide significance level (2.5×10^{-6}). When testing for many genes simultaneously, even true positive G × Es would have difficulty in standing out among all the noise (Lin and Lee, 2010). Therefore, our discussions here are restricted to candidate genes. As TWB keeps recruiting more subjects, in the future we look forward to performing genome-wide G × E analyses on a larger TWB cohort.

DATA AVAILABILITY STATEMENT

Individual-level Taiwan Biobank data are available upon application to Taiwan Biobank (https://www.twbiobank.org.tw/new_web/).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of the National Taiwan University Hospital (NTUH-REC No. 201805050RINB). The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Abebe, S. M., Berhane, Y., Worku, A., and Getachew, A. (2015). Prevalence and associated factors of hypertension: a cross-sectional community based study in northwest ethiopia. *PLoS One* 10:e0125210. doi: 10.1371/journal.pone.0125210
- Ahmad, S., Rukh, G., Varga, T. V., Ali, A., Kurbasic, A., Shungin, D., et al. (2013). Gene x physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. *PLoS Genet.* 9:e1003607. doi: 10.1371/journal.pgen.1003607
- Alonso-Gonzalez, A., Calaza, M., Rodriguez-Fontenla, C., and Carracedo, A. (2019). Novel gene-based analysis of ASD GWAS: insight into the biological role of associated genes. *Front. Genet.* 10:733. doi: 10.3389/fgene.2019.00733
- Aschard, H. (2016). A perspective on interaction effects in genetic association studies. *Genet. Epidemiol.* 40, 678–688. doi: 10.1002/gepi.21989
- Baxter, A. G., and Jordan, M. A. (2012). From markers to molecular mechanisms: type 1 diabetes in the post-GWAS era. *Rev. Diabet. Stud.* Winter 9, 201–223. doi: 10.1900/RDS.2012.9.201
- Bourgon, R., Gentleman, R., and Huber, W. (2010). Independent filtering increases detection power for high-throughput experiments. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9546–9551. doi: 10.1073/pnas.0914005107
- Chen, C. H., Yang, J. H., Chiang, C. W. K., Hsiung, C. N., Wu, P. E., Chang, L. C., et al. (2016). Population structure of Han Chinese in the modern Taiwanese

AUTHOR CONTRIBUTIONS

W-YL developed the methods and the analysis tools, designed and performed the simulation studies, analyzed the TWB data, and wrote the manuscript. Y-SL contributed to the programming for the simulations and data analyses. C-CC, Y-LL, S-JT, and P-HK contributed to the writing of the manuscript. P-HK, Y-LL, and S-JT provided the TWB data. All authors reviewed the manuscript.

FUNDING

This study was supported by the Ministry of Science and Technology of Taiwan (MOST, grant number MOST 107-2314-B-002-195-MY3 to W-YL). The acquisition of TWB data was supported by a MOST grant (Grant number MOST 102-2314-B-002-117-MY3 to P-HK) and a collaboration grant (National Taiwan University Hospital: grant number UN106-050 to Shyr-Chyr Chen and P-HK).

ACKNOWLEDGMENTS

We would like to thank the reviewers for their insightful and constructive comments, and Mr. Ya-Chin Lee for assisting with the acquisition of the TWB data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00331/full#supplementary-material>

- population based on 10,000 participants in the Taiwan Biobank project. *Hum. Mol. Genet.* 15, 5321–5331. doi: 10.1093/hmg/ddw346
- Chen, H., Meigs, J. B., and Dupuis, J. (2014). Incorporating gene-environment interaction in testing for association with rare genetic variants. *Hum. Hered.* 78, 81–90. doi: 10.1159/000363347
- Coleman, J. R., Euesden, J., Patel, H., Folarin, A. A., Newhouse, S., and Breen, G. (2016). Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Brief Funct. Genomics* 15, 298–304. doi: 10.1093/bfgp/elv037
- Dai, J. Y., Kooperberg, C., Leblanc, M., and Prentice, R. L. (2012). Two-stage testing procedures with independent filtering for genome-wide gene-environment interaction. *Biometrika* 99, 929–944. doi: 10.1093/biomet/ass044
- Fawcett, K. A., and Barroso, I. (2010). The genetics of obesity: FTO leads the way. *Trends Genet.* 26, 266–274. doi: 10.1016/j.tig.2010.02.006
- Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., et al. (2007). A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 11, 889–894.
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22.
- Frost, H. R., Shen, L., Saykin, A. J., Williams, S. M., Moore, J. H., Alzheimer's Disease, et al. (2016). Identifying significant gene-environment interactions

- using a combination of screening testing and hierarchical false discovery rate control. *Genetic Epidemiol.* 40, 544–557. doi: 10.1002/gepi.21997
- Goldstein, J. I., Crenshaw, A., Carey, J., Grant, G. B., Maguire, J., Fromer, M., et al. (2012). zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics* 1, 2543–2545. doi: 10.1093/bioinformatics/bts479
- Greenland, S. (1983). Tests for interaction in epidemiologic studies: a review and a study of power. *Stat. Med.* 2, 243–251. doi: 10.1002/sim.4780020219
- Hodge, S. E., and Greenberg, D. A. (2016). How can we explain very low odds ratios in GWAS? I. Polygenic Models. *Hum. Hered.* 81, 173–180. doi: 10.1159/000454804
- Hoerl, A. E., and Kennard, R. W. (1970). Ridge regression - biased estimation for nonorthogonal problems. *Technometrics* 12, 55–67. doi: 10.1080/00401706.1970.10488634
- Huls, A., Ickstadt, K., Schikowski, T., and Kramer, U. (2017). Detection of gene-environment interactions in the presence of linkage disequilibrium and noise by using genetic risk scores with internal weights from elastic net regression. *BMC Genet.* 12:55. doi: 10.1186/s12863-017-0519-1
- Huls, A., Kramer, U., Carlsten, C., Schikowski, T., Ickstadt, K., and Schwender, H. (2017). Comparison of weighting approaches for genetic risk scores in gene-environment interaction studies. *BMC Genet.* 16:115. doi: 10.1186/s12863-017-0586-3
- Hunter, D. J. (2005). Gene-environment interactions in human diseases. *Nat. Rev. Genet.* 6, 287–298. doi: 10.1038/nrg1578
- Jiao, S., Hsu, L., Bezieau, S., Brenner, H., Chan, A. T., Chang-Claude, J., et al. (2013). SBERIA: set-based gene-environment interaction test for rare and common variants in complex diseases. *Genet. Epidemiol.* 37, 452–464. doi: 10.1002/gepi.21735
- Jonsson, A., Renstrom, F., Lyssenko, V., Brito, E. C., Isomaa, B., Berglund, G., et al. (2009). Assessing the effect of interaction between an FTO variant (rs9939609) and physical activity on obesity in 15,925 Swedish and 2,511 Finnish adults. *Diabetologia* 52, 1334–1338. doi: 10.1007/s00125-009-1355-2
- Kilpelainen, T. O., Qi, L., Brage, S., Sharp, S. J., Sonestedt, E., Demerath, E., et al. (2011). Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med.* 8:e1001116. doi: 10.1371/journal.pmed.1001116
- Li, J., Shi, J. X., Huang, W., Sun, J. L., Wu, Y., Duan, Q., et al. (2015). Variant near FGF5 has stronger effects on blood pressure in Chinese with a higher body mass index. *Am. J. Hypertens.* 28, 1031–1037. doi: 10.1093/ajh/hp u263
- Lin, W. Y., Chan, C. C., Liu, Y. L., Yang, A. C., Tsai, S. J., and Kuo, P. H. (2019). Performing different kinds of physical exercise differentially attenuates the genetic effects on obesity measures: evidence from 18,424 Taiwan Biobank participants. *PLoS Genet.* 15:e1008277. doi: 10.1371/journal.pgen.1008277
- Lin, W. Y., Chan, C. C., Liu, Y. L., Yang, A. C., Tsai, S. J., and Kuo, P. H. (2020). Sex-specific autosomal genetic effects across 26 human complex traits. *Hum. Mol. Genet.* doi: 10.1093/hmg/ddaa040
- Lin, W. Y., Huang, C. C., Liu, Y. L., Tsai, S. J., and Kuo, P. H. (2018). Polygenic approaches to detect gene-environment interactions when external information is unavailable. *Brief Bioinform.* 20, 2236–2252. doi: 10.1093/bib/bby086
- Lin, W. Y., Huang, C. C., Liu, Y. L., Tsai, S. J., and Kuo, P. H. (2019). Genome-wide gene-environment interaction analysis using set-based association tests. *Front. Genet.* 9:715. doi: 10.3389/fgene.2018.00715
- Lin, W. Y., and Lee, W. C. (2010). Incorporating prior knowledge to facilitate discoveries in a genome-wide association study on age-related macular degeneration. *BMC Res. Notes* 3:26. doi: 10.1186/1756-0500-3-26
- Lin, X. Y., Lee, S., Christiani, D. C., and Lin, X. H. (2013). Test for interactions between a genetic marker set and environment in generalized linear models. *Biostatistics* 14, 667–681. doi: 10.1093/biostatistics/kxt006
- Lin, X. Y., Lee, S., Wu, M. C., Wang, C. L., Chen, H., Li, Z. L., et al. (2016). Test for rare variants by environment interactions in sequencing association studies. *Biometrics* 72, 156–164. doi: 10.1111/biom.12368
- Liu, C., Li, H., Qi, Q., Lu, L., Gan, W., Loos, R. J., et al. (2011). Common variants in or near FGF5, CYP17A1 and MTHFR genes are associated with blood pressure and hypertension in Chinese Hans. *J. Hypertens.* 29, 70–75. doi: 10.1097/HJH.0b013e32833f60ab
- Liu, J. Z., Mcrae, A. F., Nyholt, D. R., Medland, S. E., Wray, N. R., Brown, K. M., et al. (2010). A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* 9, 139–145. doi: 10.1016/j.ajhg.2010.06.009
- Liu, Q., Chen, L. S., Nicolae, D. L., and Pierce, B. L. (2016). A unified set-based test with adaptive filtering for gene-environment interaction analyses. *Biometrics* 72, 629–638. doi: 10.1111/biom.12428
- Lowe, J. K., Maller, J. B., Pe'er, I., Neale, B. M., Salit, J., Kenny, E. E., et al. (2009). Genome-wide association studies in an isolated founder population from the Pacific Island of Kosrae. *PLoS Genet.* 5:e1000365. doi: 10.1371/journal.pgen.1000365
- Malo, N., Libiger, O., and Schork, N. J. (2008). Accommodating linkage disequilibrium in genetic-association analyses via ridge regression. *Am. J. Hum. Genet.* 82, 375–385. doi: 10.1016/j.ajhg.2007.10.012
- McAllister, K., Mechanic, L. E., Amos, C., Aschard, H., Blair, I. A., Chatterjee, N., et al. (2017). Current challenges and new opportunities for gene-environment interaction studies of complex diseases. *Am. J. Epidemiol.* 1, 753–761.
- Mok, K. Y., Schneider, S. A., Trabzuni, D., Stamelou, M., Edwards, M., Kasperaviciute, D., et al. (2014). Genomewide association study in cervical dystonia demonstrates possible association with sodium leak channel. *Mov. Disord.* 29, 245–251. doi: 10.1002/mds.25732
- Ombrello, M. J., Kirino, Y., de Bakker, P. I., Gul, A., Kastner, D. L., and Remmers, E. F. (2014). Behcet disease-associated MHC class I residues implicate antigen binding and regulation of cell-mediated cytotoxicity. *Proc. Natl. Acad. Sci. U.S.A.* 117, 8867–8872. doi: 10.1073/pnas.1406575111
- Parrinello, G., Scaglione, R., Pinto, A., Corrao, S., Cecala, M., DiSilvestre, G., et al. (1996). Central obesity and hypertension – the role of plasma endothelin. *Am. J. Hypertens.* 9, 1186–1191. doi: 10.1016/s0895-7061(96)00259-2
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Rask-Andersen, M., Karlsson, T., Ek, W. E., and Johansson, A. (2017). Gene-environment interaction study for BMI reveals interactions between genetic factors and physical activity, alcohol consumption and socioeconomic status. *PLoS Genet.* 13:e1006977. doi: 10.1371/journal.pgen.1006977
- Scuteri, A., Sanna, S., Chen, W. M., Uda, M., Albai, G., Strait, J., et al. (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* 3:e115. doi: 10.1371/journal.pgen.0030115
- Sirugo, G., Williams, S. M., and Tishkoff, S. A. (2019). The missing diversity in human genetic studies. *Cell* 21, 26–31. doi: 10.1016/j.cell.2019.02.048
- Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *J. Roy. Stat. Soc. B Met.* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
- Vimalaswaran, K. S., Li, S., Zhao, J. H., Luan, J., Bingham, S. A., Khaw, K. T., et al. (2009). Physical activity attenuates the body mass index-increasing influence of genetic variation in the FTO gene. *Am. J. Clin. Nutr.* 90, 425–428. doi: 10.3945/ajcn.2009.27652
- Waldmann, P., Meszaros, G., Gredler, B., Fuerst, C., and Solkner, J. (2013). Evaluation of the lasso and the elastic net in genome-wide association studies. *Front. Genet.* 4:270. doi: 10.3389/fgene.2013.00270
- Willer, C. J., Bonnycastle, L. L., Conneely, K. N., Duren, W. L., Jackson, A. U., Scott, L. J., et al. (2007). Screening of 134 single nucleotide polymorphisms (SNPs) previously associated with type 2 diabetes replicates association with 12 SNPs in nine genes. *Diabetes* 56, 256–264. doi: 10.2337/db06-0461
- Wray, N. R., Pergadia, M. L., Blackwood, D. H., Penninx, B. W., Gordon, S. D., Nyholt, D. R., et al. (2012). Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Mol. Psychiatry* 17, 36–48. doi: 10.1038/mp.2010.109
- WTCCC (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 7, 661–678. doi: 10.1038/nature05911
- Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. Roy. Stat. Soc. B* 67, 301–320. doi: 10.1111/j.1467-9868.2005.00503.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Lin, Lin, Chan, Liu, Tsai and Kuo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.