Supplementary information for "Modeling regulatory network topology improves genome-wide analyses of complex human traits"

Xiang Zhu

Stanford University, Stanford, CA 94305, USA.

E-mail: xiangzhu@stanford.edu

Summary. This document contains Supplementary Notes, Supplementary Figures 1-8 and Supplementary Tables 1-19 associated with the manuscript entitled "Modeling regulatory network topology improves genome-wide analyses of complex human traits". This document was last edited by X.Z. on 2020-03-12.

Supplementary Notes

Computations underlying RSS-NET

RSS-NET is given by the following Bayesian hierarchical model:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &\sim \mathcal{N}(\widehat{\mathbf{S}}\widehat{\mathbf{R}}\widehat{\mathbf{S}}^{-1}\boldsymbol{\beta},\ \widehat{\mathbf{S}}\widehat{\mathbf{R}}\widehat{\mathbf{S}}), && (1)\\
\beta_j &\sim \pi_j\cdot\mathcal{N}(\mu_j,\ \sigma_0^2)+(1-\pi_j)\cdot\delta_0, && (2)\\
\pi_j &= (1+10^{-(\theta_0+a_j\cdot\theta)})^{-1}, && (3)\\
\mu_j &= \textstyle\sum_{g\in\mathbf{O}_j} w_{jg}\cdot\gamma_{jg}, && (4)\\
\gamma_{jg} &\sim \mathcal{N}(0,\ \sigma^2), && (5)
\end{aligned}$$

where $\widehat{\boldsymbol{\beta}} := (\hat{\beta}_1,\ldots,\hat{\beta}_p)'$ is a $p\times 1$ vector, $\widehat{\mathbf{S}} := \mathrm{diag}(\widehat{\mathbf{s}})$ is a $p\times p$ diagonal matrix with diagonal elements being $\widehat{\mathbf{s}} := (\hat{s}_1,\ldots,\hat{s}_p)'$, $\{\hat{\beta}_j,\hat{s}_j\}$ are single-SNP effect estimate and its standard error for each SNP j in a GWAS, $\widehat{\mathbf{R}}$ is a $p\times p$ LD matrix estimated from an external reference panel with ancestry matching the GWAS, $\boldsymbol{\beta} := (\beta_1,\ldots,\beta_p)'$ is a $p\times 1$ vector, β_j is the true effect of each SNP j, π_j denotes the prior probability that $\beta_j\neq 0$, μ_j denotes the prior expected value of β_j conditioning on $\beta_j\neq 0$, δ_0 denotes point mass at zero, $\{a_j,\mathbf{O}_j,w_{jg}\}$ are known annotations derived from a given regulatory network (see main text), γ_{jg} denotes the random effect of SNP j due to gene g, and $\{\theta_0,\theta,\sigma_0,\sigma\}$ are hyper-parameters.

To fit the RSS-NET model, we first integrate out γ_{jg} to obtain the following model:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &\sim \mathcal{N}(\widehat{\mathbf{S}}\widehat{\mathbf{R}}\widehat{\mathbf{S}}^{-1}\boldsymbol{\beta},\ \widehat{\mathbf{S}}\widehat{\mathbf{R}}\widehat{\mathbf{S}}), && (6)\\
\beta_j &\sim \pi_j\cdot\mathcal{N}(0,\ \sigma_j^2)+(1-\pi_j)\cdot\delta_0, && (7)\\
\pi_j &= (1+10^{-(\theta_0+a_j\theta)})^{-1}, && (8)\\
\sigma_j^2 &= \sigma_0^2+\sigma^2\cdot\textstyle\sum_{g\in\mathbf{O}_j} w_{jg}^2. && (9)
\end{aligned}$$

We then write the posterior distribution of $\boldsymbol{\beta}$ as

$$p(\boldsymbol{\beta}\mid\mathbf{D}) = \int p(\boldsymbol{\beta}\mid\mathbf{D},\theta_0,\theta,\sigma_0,\sigma)p(\theta_0,\theta,\sigma_0,\sigma\mid\mathbf{D})\mathrm{d}\theta_0\mathrm{d}\theta\mathrm{d}\sigma_0\mathrm{d}\sigma, \quad (10)$$

where $\mathbf{D}$ is a shorthand for the input data of RSS-NET including GWAS summary statistics $\{\widehat{\boldsymbol{\beta}},\widehat{\mathbf{S}}\}$, LD estimates $\widehat{\mathbf{R}}$ and network annotations $\{\mathbf{a},\mathbf{O},\mathbf{W}\}$ (see main text for definitions). For a given set of $\{\theta_0,\theta,\sigma_0,\sigma\}$, we use the mean-field variational approximation algorithm from RSS-E (Zhu and Stephens 2018) to estimate $p(\boldsymbol{\beta} \mid \mathbf{D},\theta_0,\theta,\sigma_0,\sigma)$:

$$p(\boldsymbol{\beta} \mid \mathbf{D},\theta_0,\theta,\sigma_0,\sigma) \approx \prod_{j=1}^{p} q_j^\star(\beta_j \mid \mathbf{D},\theta_0,\theta,\sigma_0,\sigma), \tag{11}$$

$$q_j^\star(\beta_j \mid \mathbf{D},\theta_0,\theta,\sigma_0,\sigma) = \alpha_j^\star \cdot \mathcal{N}(\beta_j;\, \nu_j^\star,(\tau_j^\star)^2) + \alpha_j^\star \cdot \delta_0(\beta_j), \tag{12}$$

where the optimal variational parameters $\{\alpha_j^\star,\nu_j^\star,\tau_j^\star\}$ are given by:

$$(\tau_j^\star)^2 = \frac{1}{\hat{s}_j^{-2}+\sigma_j^{-2}}, \tag{13}$$

$$\nu_j^\star = (\tau_j^\star)^2 \cdot \left(\frac{\hat{\beta}_j}{\hat{s}_j^2} - \sum_{i\neq j} \frac{\widehat{\mathbf{R}}_{ij}\alpha_i^\star\nu_i^\star}{\hat{s}_i\hat{s}_j}\right), \tag{14}$$

$$\frac{\alpha_j^\star}{1-\alpha_j^\star} = \frac{\pi_j}{1-\pi_j} \cdot \frac{\tau_j^\star}{\sigma_j} \cdot \exp\left\{\frac{(\nu_j^\star)^2}{2(\tau_j^\star)^2}\right\}. \tag{15}$$

Since we place independent uniform grid priors on $\{\theta_0,\theta,\sigma_0,\sigma\}$ (**Supplementary Table 19**), we approximate $p(\theta_0,\theta,\sigma_0,\sigma \mid \mathbf{D})$ by a discrete distribution $\omega^\star(\theta_0,\theta,\sigma_0,\sigma)$:

$$p(\theta_0,\theta,\sigma_0,\sigma \mid \mathbf{D}) \approx \omega^\star(\theta_0,\theta,\sigma_0,\sigma) \propto \exp\{F^\star(\mathbf{D},\theta_0,\theta,\sigma_0,\sigma)\}, \tag{16}$$

where $F^\star(\mathbf{D},\theta_0,\theta,\sigma_0,\sigma)$ is the variational lower bound corresponding to the optimal variational parameters $\{\alpha_j^\star,\nu_j^\star,\tau_j^\star\}$ in (13)-(15).

$$\begin{aligned} F^\star(\mathbf{D},\theta_0,\theta,\sigma_0,\sigma) = {} & \widehat{\boldsymbol{\beta}}'\widehat{\mathbf{S}}^{-2}\mathrm{E}_{q^\star}(\boldsymbol{\beta}) - \frac{1}{2}\mathrm{E}'_{q^\star}(\boldsymbol{\beta})\widehat{\mathbf{S}}^{-1}\widehat{\mathbf{R}}\widehat{\mathbf{S}}^{-1}\mathrm{E}_{q^\star}(\boldsymbol{\beta}) - \frac{1}{2}\sum_{j=1}^{p}\frac{\mathrm{Var}_{q^\star}(\beta_j)}{\hat{s}_j^2} - \sum_{j=1}^{p}\alpha_j^\star\log\left(\frac{\alpha_j^\star}{\pi_j}\right) \\ & - \sum_{j=1}^{p}(1-\alpha_j^\star)\log\left(\frac{1-\alpha_j^\star}{1-\pi_j}\right) + \sum_{j=1}^{p}\frac{\alpha_j^\star}{2}\left\{1+\log\left[\frac{(\tau_j^\star)^2}{\sigma_j^2}\right] - \frac{(\tau_j^\star)^2+(\nu_j^\star)^2}{\sigma_j^2}\right\}, \end{aligned} \tag{17}$$

$\mathrm{E}_{q^\star}(\boldsymbol{\beta}) = (\mathrm{E}_{q^\star}(\beta_1),\ldots,\mathrm{E}_{q^\star}(\beta_p))'$, $\mathrm{E}_{q^\star}(\beta_j) = \alpha_j^\star\nu_j^\star$ and $\mathrm{Var}_{q^\star}(\beta_j) = \alpha_j^\star[(\tau_j^\star)^2+(\nu_j^\star)^2] - (\alpha_j^\star\nu_j^\star)^2$. Of note, RSS-NET and RSS-E have the same posterior computation if we set $\sigma_j^2 = \sigma_\beta^2$ for all $j = 1,\ldots,p$ in (13)-(15) and (17).

Bayes factor for network enrichments

To assess whether a regulatory network is enriched for genetic associations with a trait, we evaluate the following Bayes factor (BF):

$$\mathrm{BF} = \frac{p(\widehat{\boldsymbol{\beta}} \mid \widehat{\mathbf{S}},\widehat{\mathbf{R}},\mathbf{a},\mathbf{O},\mathbf{W},M_1)}{p(\widehat{\boldsymbol{\beta}} \mid \widehat{\mathbf{S}},\widehat{\mathbf{R}},\mathbf{a},\mathbf{O},\mathbf{W},M_0)}, \tag{18}$$

where M_1 denotes the enrichment model where $\theta > 0$ or $\sigma^2 > 0$, and M_0 denotes the baseline model where $\theta = 0$ and $\sigma^2 = 0$. To compute BF (18), we approximate intractable

marginal likelihoods by corresponding variational lower bounds (17):

$$\begin{aligned} \mathrm{BF} &= \frac{\int p(\widehat{\boldsymbol{\beta}} \mid \widehat{\mathbf{S}}, \widehat{\mathbf{R}}, \mathbf{a}, \mathbf{O}, \mathbf{W}, \theta_0, \theta, \sigma_0, \sigma) p(\theta_0) p(\theta) p(\sigma_0) p(\sigma) \mathrm{d}\theta \mathrm{d}\theta_0 \mathrm{d}\sigma_0 \mathrm{d}\sigma}{\int p(\widehat{\boldsymbol{\beta}} \mid \widehat{\mathbf{S}}, \widehat{\mathbf{R}}, \mathbf{a}, \mathbf{O}, \mathbf{W}, \theta_0, \theta = 0, \sigma_0, \sigma = 0) p(\theta_0) p(\sigma_0) \mathrm{d}\theta_0 \mathrm{d}\sigma_0} \\ &\approx \frac{n_1^{-1} \sum_{s=1}^{n_1} \exp\{F^\star(\mathbf{D}, \theta_0^{(s)}, \theta^{(s)}, \sigma_0^{(s)}, \sigma^{(s)})\}}{n_0^{-1} \sum_{t=1}^{n_0} \exp\{F^\star(\mathbf{D}, \theta_0^{(t)}, \theta = 0, \sigma_0^{(t)}, \sigma = 0)\}}, \end{aligned} \tag{19}$$

where $\{\theta_0^{(s)}, \theta^{(s)}, \sigma_0^{(s)}, \sigma^{(s)}\}$ and $\{\theta_0^{(t)}, \sigma_0^{(t)}\}$ are pre-defined grids.

Posterior probability for genetic associations

To identify association between a locus and a trait, we compute P_1, the posterior probability that at least one SNP in the locus is associated with the trait:

$$P_1 = 1 - \Pr(\beta_j = 0, \forall j \in \text{locus} \mid \mathbf{D}, \text{model}), \tag{20}$$

where the "model" here can be the baseline model M_0 (yielding P_1^{base}), the enrichment model M_1 for the near-gene control network (yielding P_1^{near}) and a given network (yielding P_1^{net}). Given a grid $\{\theta_0^{(s)}, \theta^{(s)}, \sigma_0^{(s)}, \sigma^{(s)}\}$, P_1 is estimated as

$$\begin{aligned} P_1 &= 1 - \int \Pr(\beta_j = 0,\ \forall j \in \text{locus} \mid \mathbf{D}, \theta_0, \theta, \sigma_0, \sigma) p(\theta_0, \theta, \sigma_0, \sigma \mid \mathbf{D}) \mathrm{d}\theta \mathrm{d}\theta_0 \mathrm{d}\sigma_0 \mathrm{d}\sigma \\ &\approx 1 - \sum_{s=1}^{n_1} \prod_{j \in \text{locus}} \left[1 - \alpha_j^\star(\theta_0^{(s)}, \theta^{(s)}, \sigma_0^{(s)}, \sigma^{(s)})\right] \cdot \omega^\star(\theta_0^{(s)}, \theta^{(s)}, \sigma_0^{(s)}, \sigma^{(s)}). \end{aligned} \tag{21}$$

Methods compared in simulations

Three methods and their corresponding software packages are used in simulations of the present study: Pascal (Lamparter et al. 2016), LDSC (Finucane et al. 2015; Gazal et al. 2017) and RSS-E (Zhu and Stephens 2018). We downloaded the software package of Pascal and its associated database on October 5, 2017 from https://www2.unil.ch/cbg/index.php?title=Pascal. We downloaded the software package of LDSC and its associated database on November 27, 2018 from https://github.com/bulik/ldsc. To run LDSC with the baseline model (Finucane et al. 2015), we used the baseline model v1.1 files (https://data.broadinstitute.org/alkesgroup/LDSCORE/1000G_Phase3_baseline_v1.1_ldscores.tgz). To run LDSC with the baselineLD model (Gazal et al. 2017), we used the baselineLD model v2.1 files (https://data.broadinstitute.org/alkesgroup/LDSCORE/1000G_Phase3_baselineLD_v2.1_ldscores.tgz). Versions of all files were up-to-date at the time of analysis. We downloaded the software package of RSS-E on October 19, 2018 from https://github.com/stephenslab/rss/tree/master/src_vb, and we used rss_varbvsr_bigmem_squarem.m to fit the RSS-E model.

When applying Pascal, LDSC and RSS-E to analyze a given network, we create SNP-level annotations by only using the node information and ignoring the edge information. Specifically, for each SNP j we set its binary annotation $a_j = 1$ if it is within 100kb (both upstream and downstream) of any element in the network, and set $a_j = 0$ otherwise. We do not consider how these elements are connected in the network when defining a_j's.

Regulatory network as a bipartite graph

For simplicity we formulate a regulatory network as a bipartite graph $\{\mathbf{V}_{\mathrm{TF}}, \mathbf{V}_{\mathrm{TG}}, \mathbf{E}_{\mathrm{TF}\to\mathrm{TG}}\}$, where $\mathbf{V}_{\mathrm{TF}}$ denotes the node set of transcription factors (TFs), $\mathbf{V}_{\mathrm{TG}}$ denotes the node set of target genes (TGs), and $\mathbf{E}_{\mathrm{TF}\to\mathrm{TG}}$ denotes the set of directed TF-to-TG edges, summarizing how TFs regulate TGs via regulatory elements (REs), but not vice versa. Each edge has a weight between 0 and 1, measuring the relative regulation strength of a TF on a TG.

Below is a small subset of B cell regulatory network. In this example, $\mathbf{V}_{\mathrm{TF}}$ corresponds to the column `TF`, $\mathbf{V}_{\mathrm{TG}}$ corresponds to the column `TG`, $\mathbf{E}_{\mathrm{TF}\to\mathrm{TG}}$ corresponds to all rows of `TF` and `TG`, and the edge weights are specified by the column `Score`.

```
       TF     TG    Score
1:   RARG EEF1A1 0.723037
2:   SOX9   CD74 0.690921
3:   EBF1  FXYD5 0.659009
4:   PAX5   CD44 0.704366
5: TRIM28   GSAP 0.629798
6: POU6F1  RPS3A 0.632690
```

In this study, for a given TF, we only consider the downstream effects of all genes that are **directly** regulated by this TF (see main text Equation 4). For example, it is biologically possible that a TF 1 regulates a target gene which is also a TF, say TF 2; when specifying the RSS-NET SNP-level effect size distribution, we only incorporate the direct downstream effect of TF 1 on TF 2, and ignore all other indirect downstream effect of TF 1 on genes that are only directly regulated by TF 2.

Quantify *cis* impact of a SNP on a gene

To specify c_{jg}, the relative impact of a cis SNP j on a gene g (main text Equation 4), we use a simple approach based on published *cis* expression quantitative trait loci (eQTL). Specifically, if a pair of SNP j and gene g is available in an eQTL database, we define

$$c_{jg} = 1 + \sqrt{\frac{z_{jg}^2}{z_{jg}^2 + n_{jg}}}, \tag{22}$$

where n_{jg} is the number of individuals with both genotype of SNP j and expression of gene g available and z_{jg} is the single-SNP z-score measuring the marginal association between genotype of SNP j and expression of gene g across n_{jg} individuals; if the pair does not have eQTL association data available, we set $c_{jg} = 1$. When deriving c_{jg} from *cis*-eQTL data, we consider all available SNP-gene pairs, not just the significant ones.

In this study we specify c_{jg} in a context-matching manner. For example, if we analyze the regulatory network derived from liver samples, we also use *cis*-eQTL data from liver samples to define c_{jg}. For networks of 32 tissues, we use the tissue-matching *cis*-eQTL data from GTEx Analysis V7 (https://storage.googleapis.com/gtex_analysis_v7/single_tissue_eqtl_data/GTEx_Analysis_v7_eQTL_all_associations.tar.gz, accessed April 7, 2019). For networks of 5 immune cells, we use the cell-type-matching *cis*-eQTL data from DICE DB 1 (http://downloads.dice-database.org/downloads/DICE_DB_1/eqtl/unfiltered/, accessed July 3, 2019). For the "omnibus" network, we use the *cis*-eQTL data of blood samples from the eQTLGen Consortium (http://www.eqtlgen.org/

cis-eqtls.html, accessed May 17, 2019). Summary statistics of c_{jg} are provided in **Supplementary Tables 17-18**.

We acknowledge that our simple approach (22) for c_{jg} specification could be potentially improved, although we have not investigated it here. If one had improved ways to specify c_{jg} then these improvements could be incorporated into RSS-NET.

Explain SNP-level subscript j in γ_{jg}

In this work we model γ_{jg}, the random effect of SNP j due to gene g, as

$$\gamma_{jg} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\ \sigma^2).$$

We emphasize that the SNP-level subscript j in γ_{jg} ensures the exchangeability of β_j required by (2), as shown in the following example. Suppose there are two nearby SNPs 1 and 2, and their effects are sums of random effects of genes 1 and 2. If we ignore the SNP-level subscript j and use the γ_g notation, then

$$\begin{aligned}
\beta_1 &= 0.18 \cdot \gamma_1 + 0.36 \cdot \gamma_2, \\
\beta_2 &= 0.36 \cdot \gamma_1 + 0.72 \cdot \gamma_2, \\
\gamma_g &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\ \sigma^2),\ g = 1,2.
\end{aligned}$$

This is inconsistent with the independent prior (2) placed on β_j, because under this notation, $\beta_2 = 2\beta_1$ and the correlation of β_1 and β_2 is exactly 1. In contrast, this inconsistency disappears if we use the γ_{jg} notation:

$$\begin{aligned}
\beta_1 &= 0.18 \cdot \gamma_{11} + 0.36 \cdot \gamma_{12}, \\
\beta_2 &= 0.36 \cdot \gamma_{21} + 0.72 \cdot \gamma_{22}, \\
\gamma_{jg} &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\ \sigma^2),\ j = 1,2,\ g = 1,2.
\end{aligned}$$

Now β_1 and β_2 are independent, because γ_{jg}'s are all i.i.d for $j = 1,2$ and $g = 1,2$.

Induced prior distribution for σ_0^2 and σ^2

Here we assume a multiple regression model for the phenotype-genotype relationship:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{23}$$

where $\mathbf{y}$ is an $n \times 1$ centered vector of phenotype, $\mathbf{X}$ is an $n \times p$ column-centered matrix of genotype, $\boldsymbol{\beta}$ is the $p \times 1$ vector of multiple regression coefficients (i.e. SNP-level effect sizes), and $\boldsymbol{\epsilon}$ is the independent error term. Let $\sigma_{x,j}^2$ denote the population variance of genotype for SNP j, $j = 1,\ldots,p$, and let σ_y^2 denote the population variance of phenotype.

PROPOSITION 1. If SNP-level effect sizes $\boldsymbol{\beta}$ follow the prior distribution specified in RSS-NET (2), (4) and (5), then, for all $i, j = 1,\ldots,p$,

$$\mathrm{E}(\beta_j) = 0,\ \ \mathrm{Var}(\beta_j) = \pi_j \left(\sigma_0^2 + \sigma^2 \cdot \textstyle\sum_{g \in \mathbf{O}_j} w_{jg}^2\right),\ \ \mathrm{Cov}(\beta_i,\ \beta_j) = 0. \tag{24}$$

PROOF. This is a direct application of the law of total expectation. ■

PROPOSITION 2. $\mathrm{E}[V(\mathbf{X}\boldsymbol{\beta})] = \mathrm{E}^{\mathrm{base}}[V(\mathbf{X}\boldsymbol{\beta})] + \mathrm{E}^{\mathrm{net}}[V(\mathbf{X}\boldsymbol{\beta})]$, where

$$\mathrm{E}^{\mathrm{base}}[V(\mathbf{X}\boldsymbol{\beta})] = \sigma_0^2 \cdot \sum_{j=1}^{p} \pi_j \cdot \mathrm{E}[V(\mathbf{X}_j)], \tag{25}$$

$$\mathrm{E}^{\mathrm{net}}[V(\mathbf{X}\boldsymbol{\beta})] = \sigma^2 \cdot \sum_{j=1}^{p} \pi_j \cdot \left(\textstyle\sum_{g\in\mathbf{O}_j} w_{jg}^2\right) \cdot \mathrm{E}[V(\mathbf{X}_j)], \tag{26}$$

$\mathbf{X}_j$ is the jth column of $\mathbf{X}$, $j = 1,\ldots,p$, and $V(\mathbf{d})$ denotes the sample variance of a vector $\mathbf{d}$.

PROOF. Define a $p \times p$ diagonal matrix $\boldsymbol{\Sigma}_\beta$ where the jth diagonal entry is $\mathrm{Var}(\beta_j)$ in PROPOSITION 1. It suffices to notice that $V(\mathbf{X}\boldsymbol{\beta}) = n^{-1}\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$, and,

$$\begin{aligned}\mathrm{E}[V(\mathbf{X}\boldsymbol{\beta})] &= \mathrm{E}\{\mathrm{E}[V(\mathbf{X}\boldsymbol{\beta}) \mid \mathbf{X}]\} = \mathrm{E}[\mathrm{trace}(n^{-1}\mathbf{X}'\mathbf{X}\cdot\boldsymbol{\Sigma}_\beta)], \\ &= \mathrm{E}\left[\sum_{j=1}^{p} n^{-1}\mathbf{X}_j'\mathbf{X}_j \cdot \mathrm{Var}(\beta_j)\right] = \sum_{j=1}^{p} \mathrm{Var}(\beta_j)\cdot\mathrm{E}[V(\mathbf{X}_j)].\end{aligned} \tag{27}$$

Use PROPOSITION 1 to complete the proof. ■

REMARK 1. PROPOSITION 2 shows that the total genetic variation $\mathrm{E}[V(\mathbf{X}\boldsymbol{\beta})]$ can be decomposed into a component (25) that is shared by all p SNPs and a component (26) that is due to network annotations $\{\mathbf{O}_j, w_{jg}\}$.

PROPOSITION 3. Let $ns_j^2 = \sigma_y^2/\sigma_{x,j}^2$, for $j = 1,\ldots,p$. If $\{\sigma_0^2,\sigma^2\}$ are re-parameterized as

$$\sigma_0^2 = \eta\cdot(1-\rho)\cdot\left(\sum_{j=1}^{p}\frac{\pi_j}{ns_j^2}\right)^{-1}, \quad \sigma^2 = \eta\cdot\rho\cdot\left(\sum_{j=1}^{p}\frac{\pi_j\cdot\sum_{g\in\mathbf{O}_j} w_{jg}^2}{ns_j^2}\right)^{-1}, \tag{28}$$

then

$$\eta = \frac{\mathrm{E}[V(\mathbf{X}\boldsymbol{\beta})]}{\mathrm{E}[V(\mathbf{y})]}, \quad \rho = \frac{\mathrm{E}^{\mathrm{net}}[V(\mathbf{X}\boldsymbol{\beta})]}{\mathrm{E}[V(\mathbf{X}\boldsymbol{\beta})]}. \tag{29}$$

PROOF. It suffices to show that, for all $j = 1,\ldots,p$,

$$ns_j^2 = \frac{\sigma_y^2}{\sigma_{x,j}^2} = \frac{\mathrm{E}[V(\mathbf{y})]}{\mathrm{E}[V(\mathbf{X}_j)]}. \tag{30}$$

Use PROPOSITION 2 to complete the proof. ■

REMARK 2. PROPOSITION 3 provides the mathematical basis for interpretations of $\{\eta,\rho\}$. Specifically, η represents the proportion of the total phenotypic variation explained by p SNPs, and ρ represents the proportion of total genetic variation explained by network annotations $\{\mathbf{O}_j, w_{jg}\}$.

REMARK 3. Since $ns_j^2 = \sigma_y^2/\sigma_{x,j}^2$, where σ_y^2 is the population variance of phenotype and $\sigma_{x,j}^2$ is the population variance of genotype at SNP j, this re-parameterization based on $\{\eta,\rho\}$ ensures that the induced-prior (28) of $\{\sigma_0^2,\sigma^2\}$ does not depend on GWAS sample size n, and the resulting genetic effect sizes ($\boldsymbol{\beta}$) have the same measurement unit as the phenotype.

REMARK 4. The induced prior (28) of $\{\sigma_0^2,\sigma^2\}$ contains a quantity $ns_j^2 = \sigma_y^2/\sigma_{x,j}^2$ that depends on unknown population parameters $\{\sigma_y^2,\sigma_{x,j}^2\}$. As shown in Zhu and Stephens (2017), ns_j^2 can be reliably estimated from GWAS summary data: $ns_j^2 \approx n\hat{s}_j^2$. In practice, we replace the unknown ns_j^2 in (28) with known $n\hat{s}_j^2$ (see main text Equation 7).

Acknowledgments and data sources

This study makes use of data generated by the Wellcome Trust Case Control Consortium, the 1000 Genomes Project, the ENCODE Consortium, the GTEx Project, the DICE Project, the eQTLGen Consortium, and multiple GWAS consortia. We thank all these projects and consortia for making their data publicly available. Detailed acknowledgments and data sources are listed below.

- **Wellcome Trust Case Control Consortium (WTCCC).** Simulations of this study use individual-level genotype data generated by WTCCC. A full list of the investigators who contributed to the generation of the data is available from https://www.wtccc.org.uk/. The WTCCC data are available at the European Genome-phenome Archive (https://www.ebi.ac.uk/ega/).
- **1000 Genomes Project.** This study uses haplotypes of individuals with European ancestry from the 1000 Genomes Project, Phase 3 (1000 Genomes Project Consortium 2015) to estimate LD. The 1000 Genomes Phase 3 haplotype data have been contributed by 1000 Genomes investigators and have been downloaded from ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/.
- **Encyclopedia of DNA Elements (ENCODE) Consortium.** This study uses gene expression and chromatin openness data generated by the ENCODE Consortium and the ENCODE production laboratories. The ENCODE data are available at the ENCODE portal (https://www.encodeproject.org/).
- **Genotype-Tissue Expression (GTEx) Project.** This study uses *cis*-eQTL data generated by the GTEx Project. The GTEx Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The GTEx data are available at the GTEx Portal (https://gtexportal.org/home/).
- **Database of Immune Cell Expression, Expression quantitative trait loci and Epigenomics (DICE) Project.** This study uses *cis*-eQTL data generated by the DICE Project. The DICE Project was supported by NIH R24AI108564. The DICE data are available at https://dice-database.org/.
- **eQTLGen Consortium.** This study uses *cis*-eQTL data generated by the eQTLGen Consortium. The eQTLGen data are available at http://www.eqtlgen.org/.
- **Genetic Investigation of ANthropometric Traits (GIANT) Consortium.** GWAS summary statistics on adult human height (Wood et al. 2014), body mass index (Locke et al. 2015) and body fat distribution (Shungin et al. 2015) have been contributed by GIANT investigators and have been downloaded from http://portals.broadinstitute.org/collaboration/giant.
- **Psychiatric Genomics Consortium (PGC).** GWAS summary statistics on schizophrenia (Ripke et al. 2014) have been contributed by PGC investigators and have been downloaded from http://www.med.unc.edu/pgc.
- **International Inflammatory Bowel Disease Genetics Consortium (IIBDGC).** GWAS summary statistics on inflammatory bowel disease (Liu et al. 2015), including Crohn's disease and ulcerative colitis have been contributed by IIBDGC investigators and have been downloaded from https://www.ibdgenetics.org.

- **Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics (CARDIoGRAMplusC4D) Consortium.** GWAS summary statistics on coronary artery disease and myocardial infarction (Nikpay et al. 2015) have been contributed by CARDIoGRAMplusC4D investigators and downloaded from http://www.cardiogramplusc4d.org/.

- **GWAS summary statistics of heart rate.** GWAS summary statistics on heart rate have been contributed by authors of Den Hoed et al. (2013) and have been downloaded from https://walker05.u.hpc.mssm.edu/.

- **International Genomics of Alzheimer's Project (IGAP).** GWAS summary statistics on Alzheimer's disease (Lambert et al. 2013) have been contributed by IGAP investigators and have been downloaded from http://web.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php. We thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i-Select chips was funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de Lille, Universite de Lille 2 and the Lille University Hospital. GERAD was supported by the Medical Research Council (Grant no. 503480), Alzheimer's Research UK (Grant no. 503176), the Wellcome Trust (Grant no. 082604/2/07/Z) and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant no. 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01-AG-12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC-10-196728.

- **Social Science Genetic Association Consortium (SSGAC).** GWAS summary statistics on neuroticism (Okbay et al. 2016) have been contributed by SSGAC investigators and have been downloaded from https://www.thessgac.org. For financial support, the SSGAC thanks the U.S. National Science Foundation, the U.S. National Institutes of Health (National Institute on Aging, and the Office for Behavioral and Social Science Research), the Ragnar Soderberg Foundation, the Swedish Research Council, The Jan Wallander and Tom Hedelius Foundation, the European Research Council, and the Pershing Square Fund of the Foundations of Human Behavior.

- **GWAS summary statistics of rheumatoid arthritis.** GWAS summary statistics on rheumatoid arthritis have been contributed by authors of Okada et al. (2014) and have been downloaded from http://plaza.umin.ac.jp/yokada/datasource/software.htm.

- **DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium.** GWAS summary statistics on type 2 diabetes (Morris et al. 2012) have been contributed by DIAGRAM investigators and have been downloaded from http://www.diagram-consortium.org.

- **Global Lipids Genetics Consortium (GLGC).** GWAS summary statistics on low-density and high-density lipoprotein cholesterol (Teslovich et al. 2010) have been contributed by GLGC investigators and have been downloaded from http://csg.sph.umich.edu//abecasis/public/lipids2010.
- **GWAS summary statistics of atrial fibrillation.** GWAS summary statistics on atrial fibrillation have been contributed by authors of Christophersen et al. (2017) and are available from the corresponding author on reasonable request, as shown in the "Data availability" section of Christophersen et al. (2017).
- **Genetic Associations and Mechanisms in Oncology (GAME-ON), Discovery, Biology, and Risk of Inherited Variants in Breast Cancer (DRIVE) Project.** GWAS summary statistics on breast cancer have been contributed by the GAME-ON/DRIVE Project and have been downloaded from http://gameon.dfci.harvard.edu. The DRIVE project acknowledges the following GWASs and investigators that shared genome-wide summary data as part of the breast-cancer GWAS meta-analysis: the Australian Breast Cancer Family Study (ABCFS) (John L. Hopper, Melissa C. Southey, Enes Makalic, Daniel F. Schmidt), the British Breast Cancer Study (BBCS) (Olivia Fletcher, Julian Peto, Lorna Gibson, Isabel dos Santos Silva), the Breast and Prostate Cancer Cohort Consortium (BPC3) (David J. Hunter, Sara Lindstrom, Peter Kraft), the Breast Cancer Family Registries (BCFR) (Habib Ahsan, Alice Whittemore), the Dutch Familial Bilateral Breast Cancer Study (DFBBCS) (Quinten Waisfisz, Hanne Meijers-Heijboer, Muriel Adank, Rob B. van der Luijt, Andre G. Uitterlinden, Albert Hofman), German Consortium for Hereditary Breast and Ovarian Cancer (GC-HBOC) (Alfons Meindl, Rita K. Schmutzler, Bertram Muller-Myhsok, Peter Lichtner), the Helsinki Breast Cancer Study (HEBCS) (Heli Nevanlinna, Taru A. Muranen, Kristiina Aittomaki, Carl Blomqvist), the Mammary Carcinoma Risk Factor Investigation (MARIE) (Jenny Chang-Claude, Rebecca Hein, Norbert Dahmen, Lars Beckman), SardiNIA (Laura Crisponi), the Singapore and Sweden Breast Cancer Study (SASBAC) (Per Hall, Kamila Czene, Astrid Irwanto, Jianjun Liu), and the UK2 (Douglas F. Easton, Clare Turnbull, Nazneen Rahman).

Supplementary Figures

Supplementary Figure 1

Simulation details and additional results of Figure 2. Here we use real genotypes of 348,965 genome-wide SNPs on chromosomes 1-22 from 1,458 individuals in the UK Blood Service Control Group (Wellcome Trust Case Control Consortium 2007) to simulate phenotype data, and then compute single-SNP association summary statistics. On these summary data, we compare RSS-NET with existing enrichment methods.

We create a B cell regulatory network by applying PECA (Duren et al. 2017) to the paired gene expression and chromatin accessibility data of primary B cells from peripheral blood (ENCODE Project Consortium 2012; Luo et al. 2020). We use the B cell network to create SNP-level annotation for 348,965 SNPs. Specifically, we let $a_j = 1$ if SNP j is within 100 kb of either a cis regulatory element (RE) or the transcribed region of a gene in the B cell network, and let $a_j = 0$ otherwise. There are 121,308 SNPs mapped to the B cell network with annotation $a_j = 1$.

To ensure that baseline and enrichment datasets have roughly the same proportion of true genetic signals, we simulate SNP-level causal indicators for baseline and enrichment datasets in a paired way. Specifically, we first simulate the causal indicator ζ_j of each SNP j for an enrichment dataset as follows:

$$\begin{aligned} \zeta_j &\sim \text{Bernoulli}(\pi_j), \\ \pi_j &= (1+10^{-(\theta_0+a_j\theta)})^{-1}, \end{aligned}$$

where θ_0 and θ are the baseline and enrichment proportion parameters, respectively. We then count the number of causal SNPs in this enrichment dataset as $n_c := \sum_j \zeta_j$, and randomly choose n_c SNPs as causal SNPs for the corresponding baseline dataset.

For enrichment datasets, we simulate the true effect β_j of each SNP j as

$$\begin{aligned} \beta_j \mid \zeta_j = 0 &\sim \delta_0, \\ \beta_j \mid \zeta_j = 1 &\sim \mathcal{N}(0,\ \sigma_0^2 + \sigma^2 \cdot \textstyle\sum_{g\in\mathbf{O}_j} w_{jg}^2), \end{aligned}$$

where δ_0 denotes point mass at zero, σ_0^2 and σ^2 are the baseline and enrichment magnitude parameter respectively, $\mathbf{O}_j$ denotes a set of genes contributing to the total genetic effect of SNP j, and w_{jg} measures the relevance between SNP j and gene g. For baseline datasets, we simulate β_j from the same model above with $\sigma^2 = 0$.

To ensure that baseline and enrichment datasets have roughly the same magnitude of total genetic signals, we simulate phenotypes by matching signal-to-noise ratio. Specifically, we simulate phenotype y_i of individual i as follows:

$$\begin{aligned} y_i &= \sum_{j=1}^{p} x_{ij} \cdot \beta_j + \epsilon_i, \\ \epsilon_i &\sim \mathcal{N}(0, \tau^{-1}), \end{aligned}$$

where p is 348,965, x_{ij} is the genotype of SNP j for individual i. The true value of residual variance τ^{-1} is determined by the true PVE, that is, the total proportion of variance in

phenotype $\mathbf{y}$ explained by effects of all available SNPs in $\mathbf{X}$:

$$\text{PVE} = \frac{V(\mathbf{X}\boldsymbol{\beta})}{\tau^{-1} + V(\mathbf{X}\boldsymbol{\beta})},$$

where $V(\mathbf{X}\boldsymbol{\beta})$ is the sample variance of $\mathbf{X}\boldsymbol{\beta}$. For the same simulation scenario, baseline and enrichment datasets have the same PVE values.

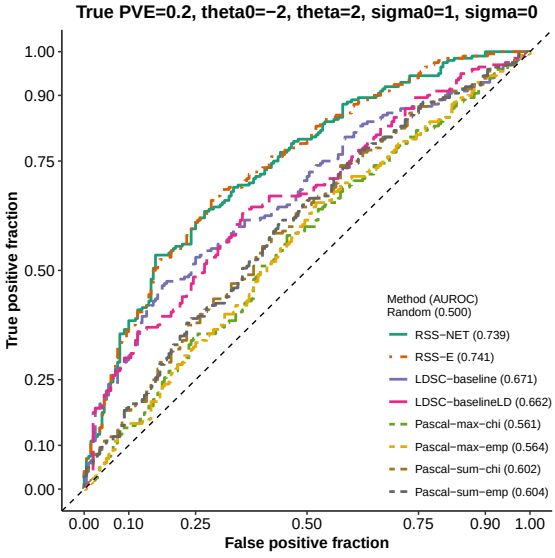
Here the true values of baseline proportion parameter θ_0 are $\{-2,-3,-4\}$, the true values of enrichment proportion parameter θ are $\{0,2\}$, the true value of baseline magnitude parameter σ_0 is 1, the true values of enrichment magnitude parameter σ are $\{0,2\}$, and the true values of PVE are $\{0.2,0.6\}$ For each possible combination of $\{\theta_0,\theta,\sigma_0,\sigma,\text{PVE}\}$ (excluding the case where $\theta=\sigma=0$), we simulate 200 baseline and 200 enrichment independent datasets. Captions of the following plots show the true parameter values for enrichment datasets. The "sparse scenario" in **Figure 2** corresponds to simulations with true $\theta_0=-4$ and PVE = 0.6. The "polygenic scenario" in **Figure 2** corresponds to simulations with true $\theta_0=-2$ and PVE = 0.6.

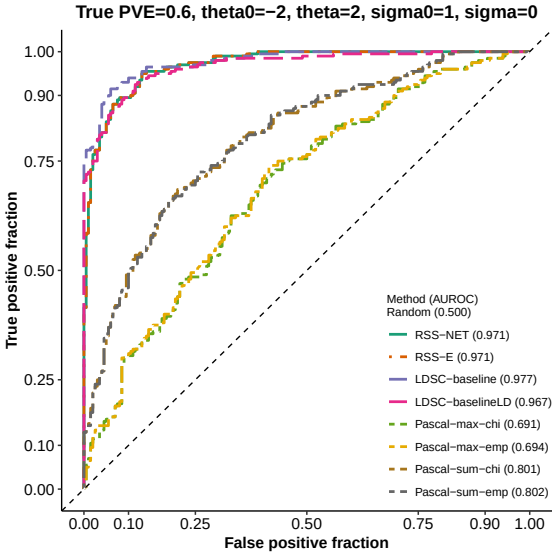
We apply RSS-NET to simulated datasets with the following hyper-parameter grids. The baseline parameters θ_0 and σ_0 are set as their true values. The grids on the enrichment parameters θ and σ are `(0:0.25:1)` if their true values are zero; otherwise they are `[0 ((truth-0.5):0.25:(truth+0.5))]`. For the same simulation scenario, we use the same hyper-parameter grid in RSS-NET analyses of baseline and enrichment datasets.

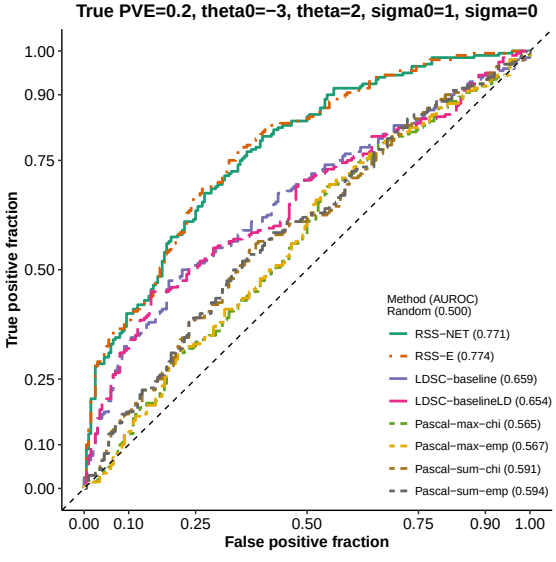
For each simulated dataset, we also perform enrichment analysis using existing approaches with their default settings: RSS-E (Zhu and Stephens 2018), LDSC-baseline (Finucane et al. 2015), LDSC-baselineLD (Gazal et al. 2017) and Pascal (Lamparter et al. 2016). Pascal includes two gene scoring options: maximum-of-χ^2 (`-max`) and sum-of-χ^2 (`-sum`), and two pathway scoring options: χ^2 approximation (`-chi`) and empirical sampling (`-emp`). For each dataset, Pascal and LDSC methods produce p-values, whereas RSS-E and RSS-NET produce BFs; these statistics are used to rank the significance of enrichments. Additional details are provided in **Supplementary Notes**.

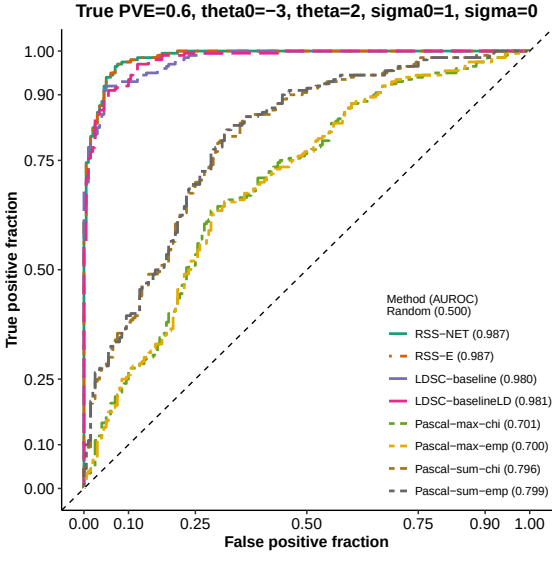
For all datasets the candidate of enrichment testing is the B cell network. A false positive occurs if a method identifies a baseline dataset as enriched. A true positive occurs if a method identifies an enrichment dataset as enriched.

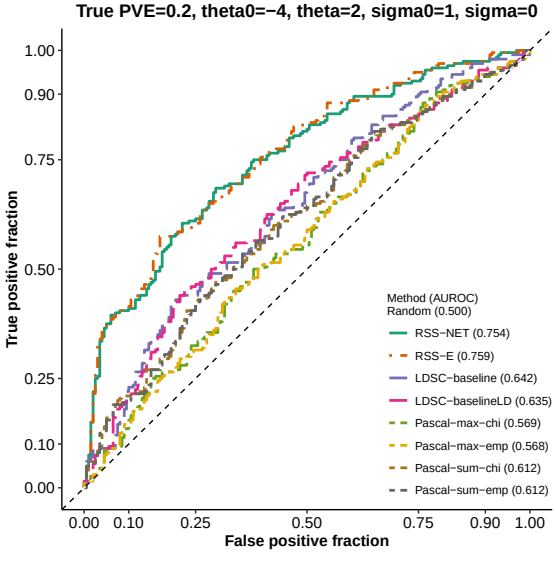
We evaluate the performance of these enrichment methods by plotting the receiver operating characteristic (ROC) curve and computing the area under the ROC curve (AUROC) for each method. Both metrics are implemented in the R package `plotROC`.

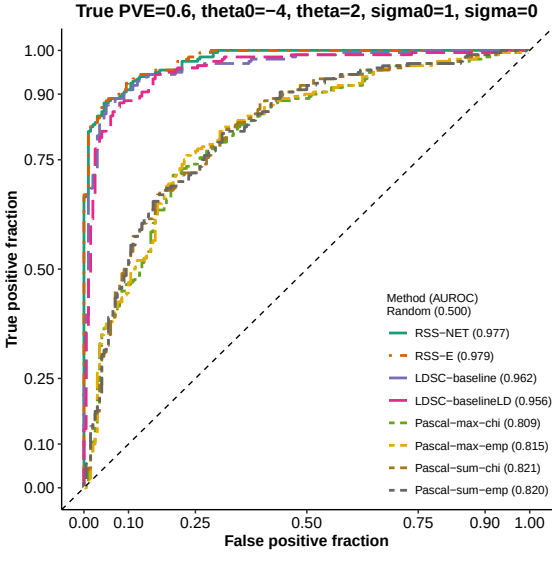

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.739)
RSS−E (0.741)
LDSC−baseline (0.671)
LDSC−baselineLD (0.662)
Pascal−max−chi (0.561)
Pascal−max−emp (0.564)
Pascal−sum−chi (0.602)
Pascal−sum−emp (0.604)


True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.971)
RSS−E (0.971)
LDSC−baseline (0.977)
LDSC−baselineLD (0.967)
Pascal−max−chi (0.691)
Pascal−max−emp (0.694)
Pascal−sum−chi (0.801)
Pascal−sum−emp (0.802)

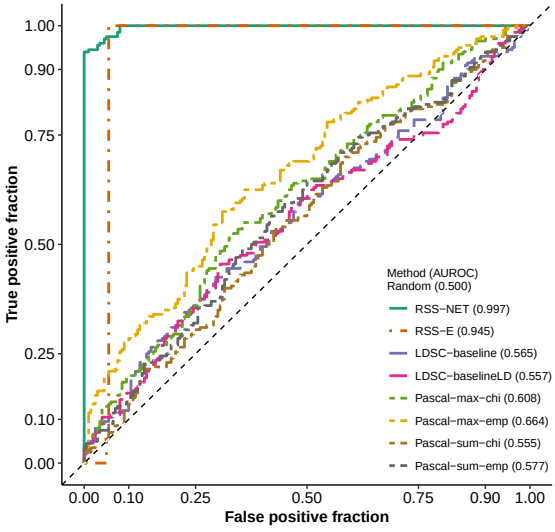

True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.771)
RSS−E (0.774)
LDSC−baseline (0.659)
LDSC−baselineLD (0.654)
Pascal−max−chi (0.565)
Pascal−max−emp (0.567)
Pascal−sum−chi (0.591)
Pascal−sum−emp (0.594)


True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.987)
RSS−E (0.987)
LDSC−baseline (0.980)
LDSC−baselineLD (0.981)
Pascal−max−chi (0.701)
Pascal−max−emp (0.700)
Pascal−sum−chi (0.796)
Pascal−sum−emp (0.799)

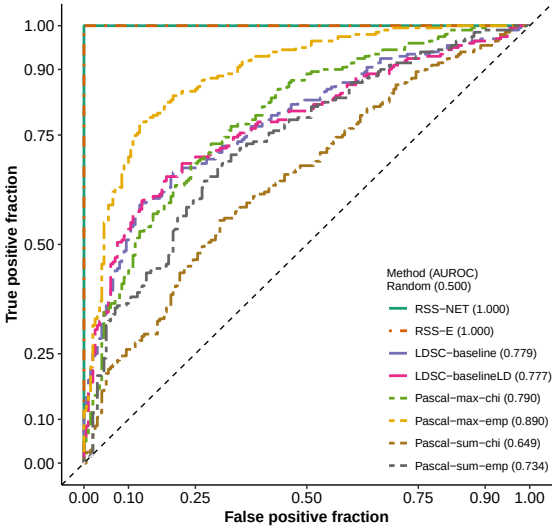

True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.754)
RSS−E (0.759)
LDSC−baseline (0.642)
LDSC−baselineLD (0.635)
Pascal−max−chi (0.569)
Pascal−max−emp (0.568)
Pascal−sum−chi (0.612)
Pascal−sum−emp (0.612)


True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.977)
RSS−E (0.979)
LDSC−baseline (0.962)
LDSC−baselineLD (0.956)
Pascal−max−chi (0.809)
Pascal−max−emp (0.815)
Pascal−sum−chi (0.821)
Pascal−sum−emp (0.820)

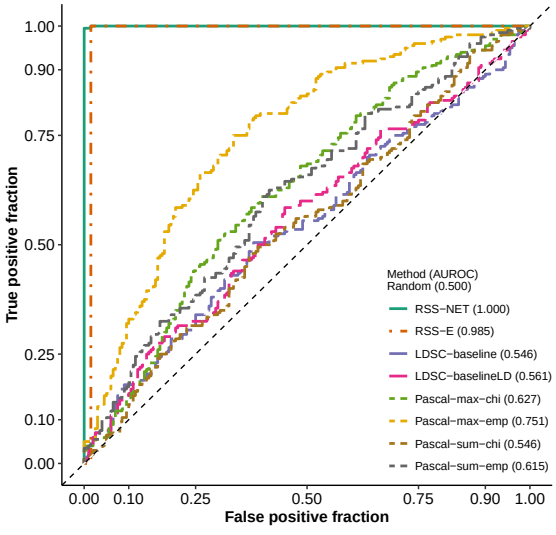
True PVE=0.2, theta0=−2, theta=0, sigma0=1, sigma=2



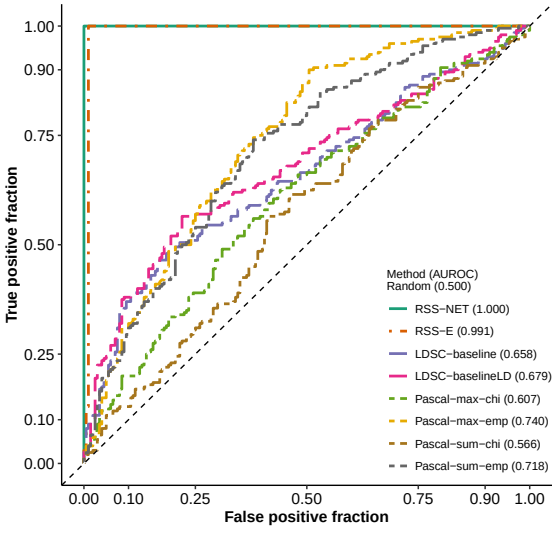
True PVE=0.6, theta0=−2, theta=0, sigma0=1, sigma=2



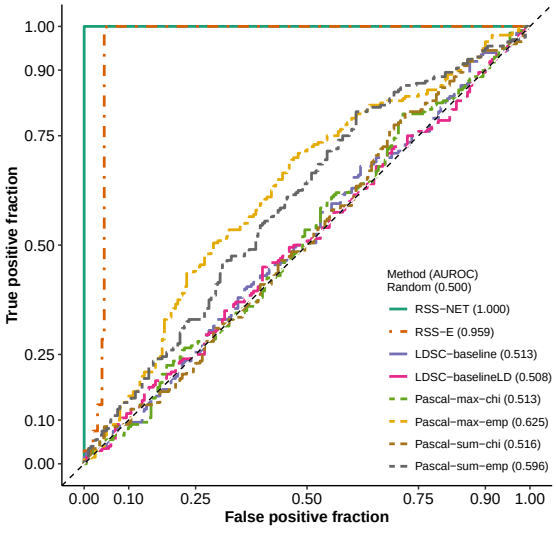
True PVE=0.2, theta0=−3, theta=0, sigma0=1, sigma=2



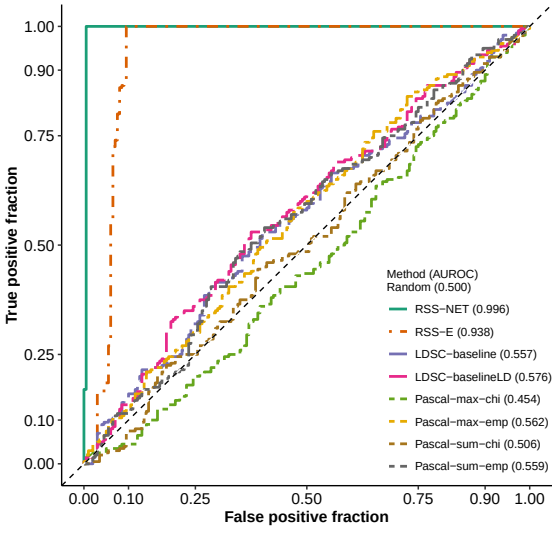
True PVE=0.6, theta0=−3, theta=0, sigma0=1, sigma=2

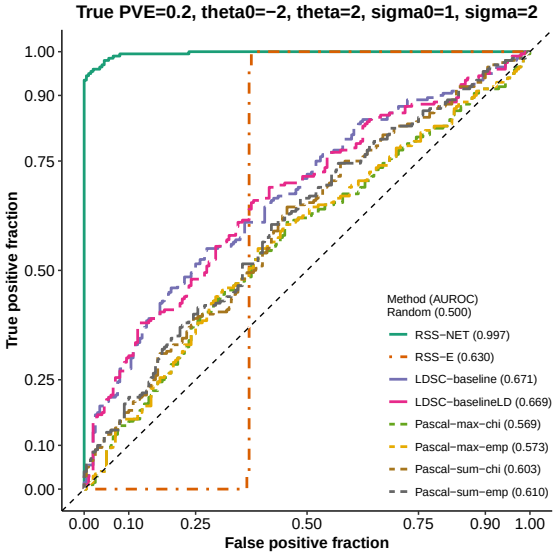


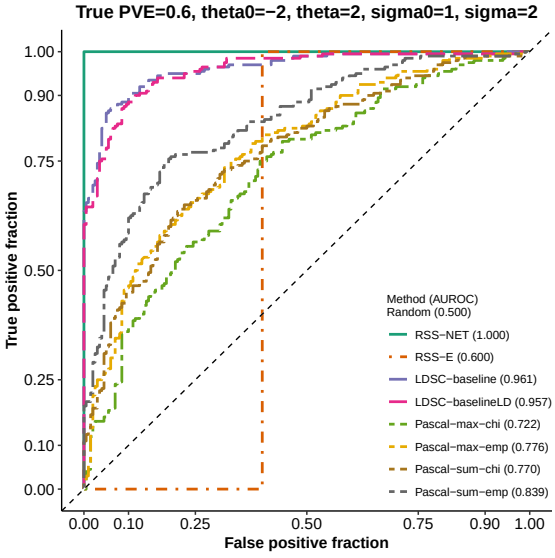
True PVE=0.2, theta0=−4, theta=0, sigma0=1, sigma=2

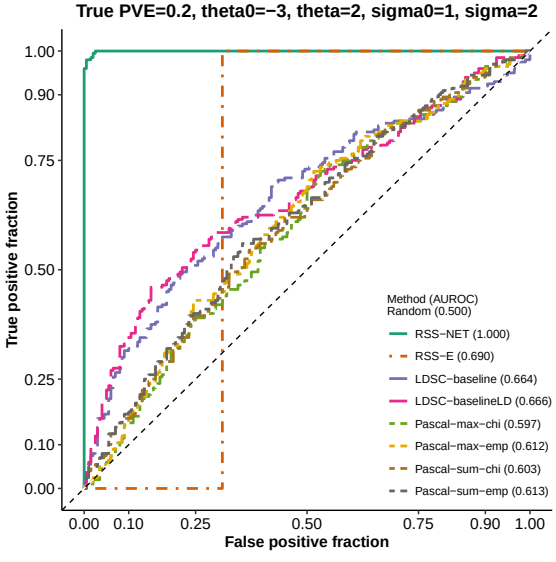


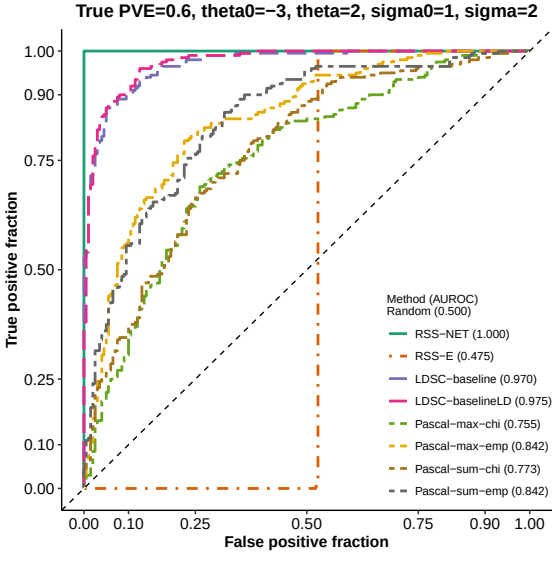
True PVE=0.6, theta0=−4, theta=0, sigma0=1, sigma=2

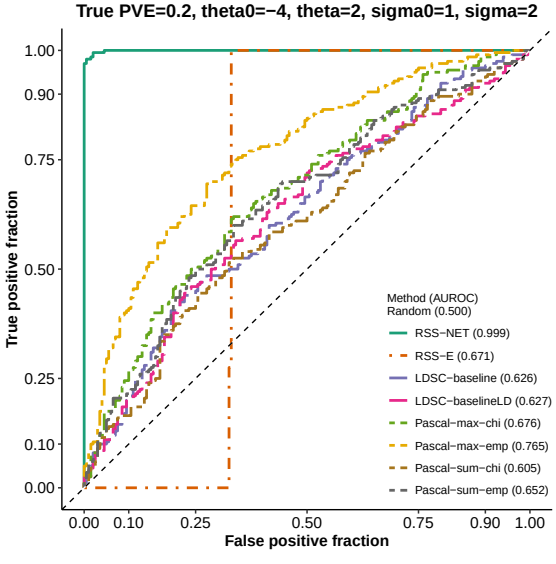


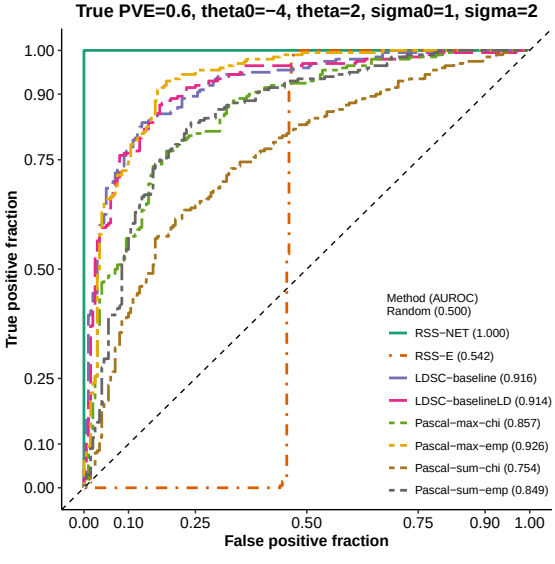

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.997)
RSS−E (0.630)
LDSC−baseline (0.671)
LDSC−baselineLD (0.669)
Pascal−max−chi (0.569)
Pascal−max−emp (0.573)
Pascal−sum−chi (0.603)
Pascal−sum−emp (0.610)


True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
RSS−E (0.600)
LDSC−baseline (0.961)
LDSC−baselineLD (0.957)
Pascal−max−chi (0.722)
Pascal−max−emp (0.776)
Pascal−sum−chi (0.770)
Pascal−sum−emp (0.839)


True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
RSS−E (0.690)
LDSC−baseline (0.664)
LDSC−baselineLD (0.666)
Pascal−max−chi (0.597)
Pascal−max−emp (0.612)
Pascal−sum−chi (0.603)
Pascal−sum−emp (0.613)


True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
RSS−E (0.475)
LDSC−baseline (0.970)
LDSC−baselineLD (0.975)
Pascal−max−chi (0.755)
Pascal−max−emp (0.842)
Pascal−sum−chi (0.773)
Pascal−sum−emp (0.842)


True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.999)
RSS−E (0.671)
LDSC−baseline (0.626)
LDSC−baselineLD (0.627)
Pascal−max−chi (0.676)
Pascal−max−emp (0.765)
Pascal−sum−chi (0.605)
Pascal−sum−emp (0.652)


True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
RSS−E (0.542)
LDSC−baseline (0.916)
LDSC−baselineLD (0.914)
Pascal−max−chi (0.857)
Pascal−max−emp (0.926)
Pascal−sum−chi (0.754)
Pascal−sum−emp (0.849)

Supplementary Figure 2

Simulation details and additional results of Figure 3(a). This part aims to assess the robustness of RSS-NET to model mis-specification where a random set of "near-gene" SNPs is enriched for association in baseline datasets. Details of this part are almost identical to those in **Supplementary Figure 1**. Here we only highlight the differences.

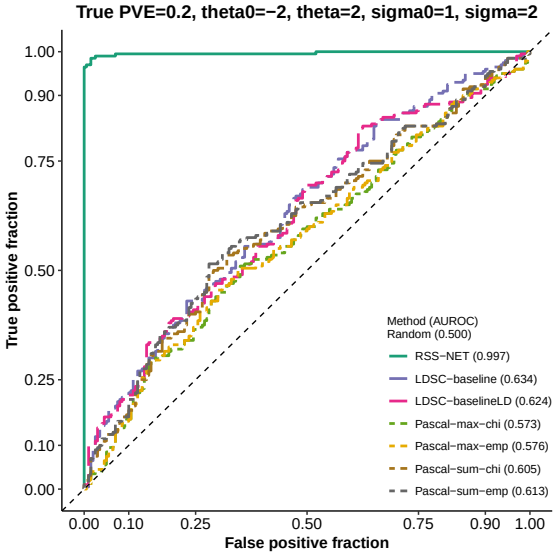
We define a SNP as "near-gene" if this SNP is within 100 kb of the transcribed region of any autosomal protein-coding gene. In total, there are 254,296 near-gene SNPs among 348,965 genome-wide SNPs from Wellcome Trust Case Control Consortium (2007).

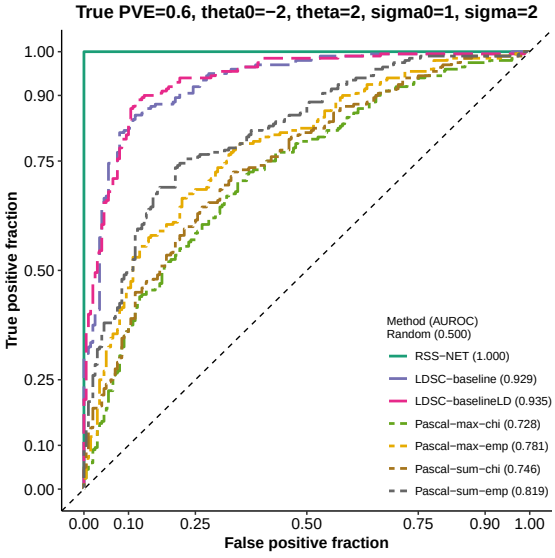
Here we simulate baseline and enrichment datasets in a paired way. We first simulate an enrichment dataset as in **Supplementary Figure 1**. For this enrichment dataset, we count the total number of causal SNPs as $n_c = \sum_j \zeta_j$, and count the number of causal SNPs in the target network as $n_p = \sum_j \zeta_j \alpha_j$ (recall that $\alpha_j = 1$ indicates that SNP j is in the network). We then randomly choose n_p SNPs from the 254,296 near-gene SNPs and $(n_c - n_p)$ SNPs from the remaining SNPs, and use them as causal SNPs for the corresponding baseline dataset. With causal indicators $\{\zeta_j\}$ in place, we simulate the true genetic effects $\boldsymbol{\beta}$ in a baseline dataset as follows

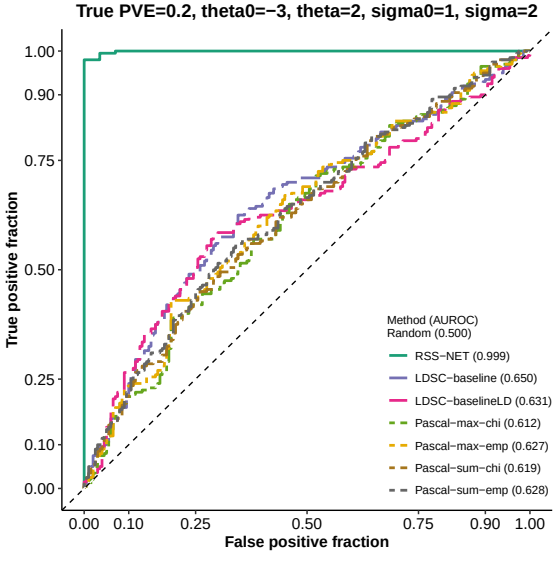
$$\begin{aligned} \beta_j \mid \zeta_j = 0 &\sim \delta_0, \\ \beta_j \mid \zeta_j = 1,\ \alpha_j = 0 &\sim \mathcal{N}(0,\ \sigma_0^2), \\ \beta_j \mid \zeta_j = 1,\ \alpha_j = 1 &\sim \mathcal{N}(0,\ (2 \cdot \sigma_0)^2), \end{aligned}$$

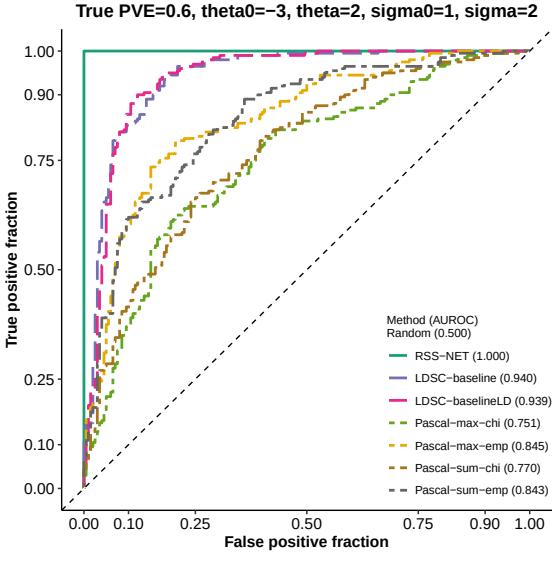
where $\zeta_j = 1$ if SNP j is causal and 0 otherwise, $\alpha_j = 1$ if SNP j is near-gene and 0 otherwise. The rest of simulations is the same as **Supplementary Figure 1**.

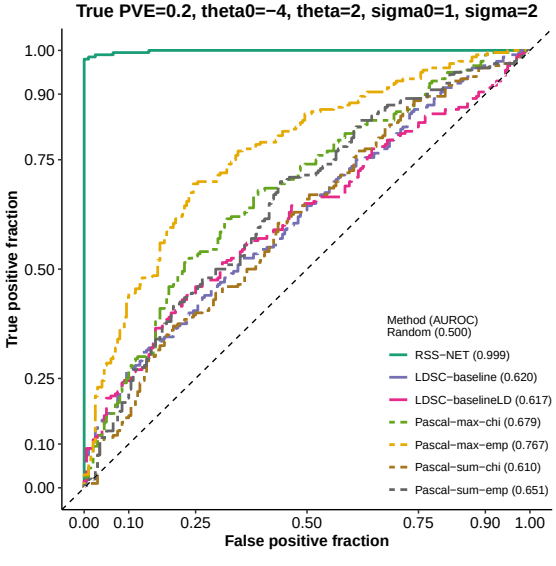
Captions of the following plots show the true parameter values for enrichment datasets. The "sparse scenario" in **Figure 3(a)** corresponds to simulations with true $\theta_0 = -4$ and PVE = 0.6. The "polygenic scenario" in **Figure 3(a)** corresponds to simulations with true $\theta_0 = -2$ and PVE = 0.6.

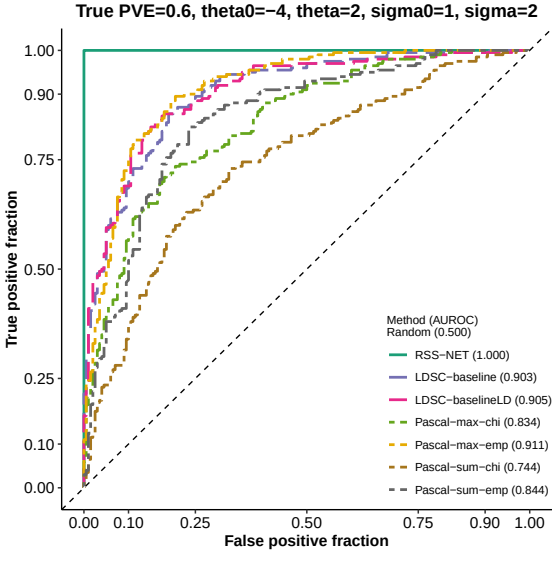

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.997)
LDSC−baseline (0.634)
LDSC−baselineLD (0.624)
Pascal−max−chi (0.573)
Pascal−max−emp (0.576)
Pascal−sum−chi (0.605)
Pascal−sum−emp (0.613)


True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
LDSC−baseline (0.929)
LDSC−baselineLD (0.935)
Pascal−max−chi (0.728)
Pascal−max−emp (0.781)
Pascal−sum−chi (0.746)
Pascal−sum−emp (0.819)


True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.999)
LDSC−baseline (0.650)
LDSC−baselineLD (0.631)
Pascal−max−chi (0.612)
Pascal−max−emp (0.627)
Pascal−sum−chi (0.619)
Pascal−sum−emp (0.628)


True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
LDSC−baseline (0.940)
LDSC−baselineLD (0.939)
Pascal−max−chi (0.751)
Pascal−max−emp (0.845)
Pascal−sum−chi (0.770)
Pascal−sum−emp (0.843)


True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.999)
LDSC−baseline (0.620)
LDSC−baselineLD (0.617)
Pascal−max−chi (0.679)
Pascal−max−emp (0.767)
Pascal−sum−chi (0.610)
Pascal−sum−emp (0.651)


True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
LDSC−baseline (0.903)
LDSC−baselineLD (0.905)
Pascal−max−chi (0.834)
Pascal−max−emp (0.911)
Pascal−sum−chi (0.744)
Pascal−sum−emp (0.844)

Supplementary Figure 3

Simulation details and additional results of Figure 3(b). This part aims to assess the robustness of RSS-NET to model mis-specification where a random set of "near-RE" SNPs is enriched for association in baseline datasets. Details of this part are almost identical to those in **Supplementary Figure 1**. Here we only highlight the differences.

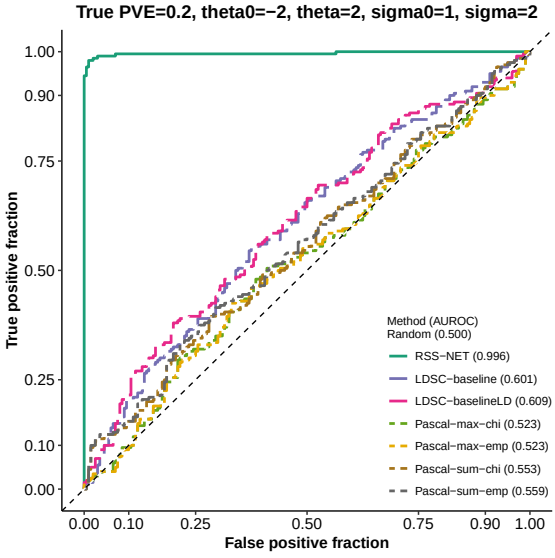
We define a SNP as "near-RE" if this SNP is within 10 kb of any RE. There are 3,895,021 unique autosomal REs among PECA networks. In total, there are 169,100 near-RE SNPs among 348,965 genome-wide SNPs from Wellcome Trust Case Control Consortium (2007).

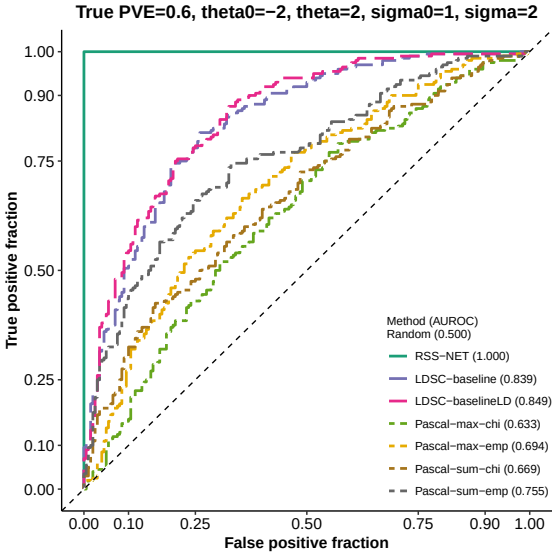
Here we simulate baseline and enrichment datasets in a paired way. We first simulate an enrichment dataset as in **Supplementary Figure 1**. For this enrichment dataset, we count the total number of causal SNPs as $n_c = \sum_j \zeta_j$, and count the number of causal SNPs in the target network as $n_p = \sum_j \zeta_j a_j$ (recall that $a_j = 1$ indicates that SNP j is in the network). We then randomly choose n_p SNPs from the 169,100 near-RE SNPs and $(n_c - n_p)$ SNPs from the remaining SNPs, and use them as causal SNPs for the corresponding baseline dataset. With causal indicators $\{\zeta_j\}$ in place, we simulate the true genetic effects $\boldsymbol{\beta}$ in a baseline dataset as follows

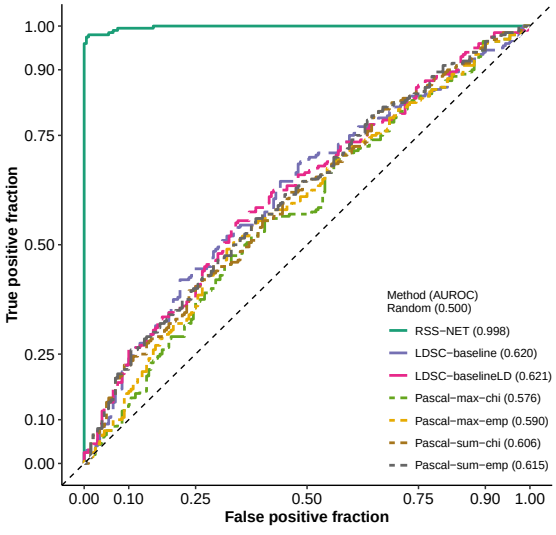
$$\begin{aligned} \beta_j \mid \zeta_j = 0 &\sim \delta_0, \\ \beta_j \mid \zeta_j = 1,\ a_j = 0 &\sim \mathcal{N}(0,\ \sigma_0^2), \\ \beta_j \mid \zeta_j = 1,\ a_j = 1 &\sim \mathcal{N}(0,\ (2 \cdot \sigma_0)^2), \end{aligned}$$

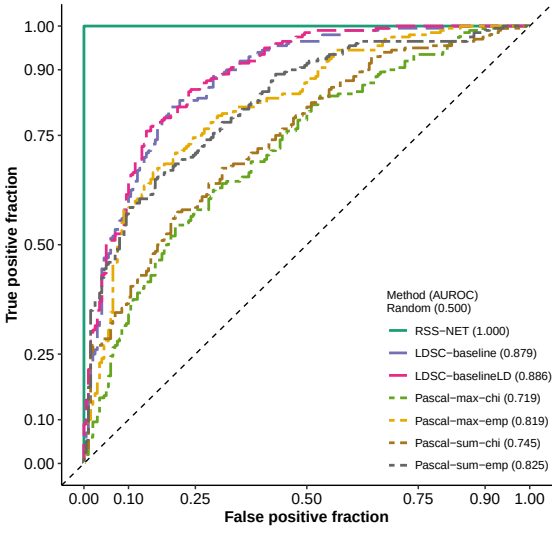
where $\zeta_j = 1$ if SNP j is causal and 0 otherwise, $a_j = 1$ if SNP j is near-RE and 0 otherwise. The rest of simulations is the same as **Supplementary Figure 1**

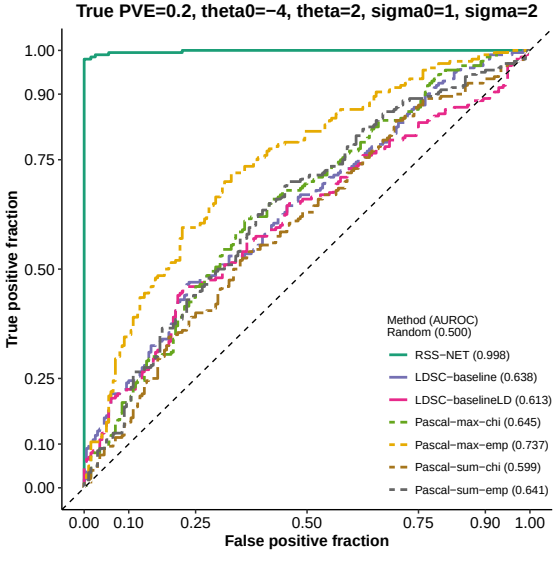
Captions of the following plots show the true parameter values for enrichment datasets. The "sparse scenario" in **Figure 3(b)** corresponds to simulations with true $\theta_0 = -4$ and PVE = 0.6. The "polygenic scenario" in **Figure 3(b)** corresponds to simulations with true $\theta_0 = -2$ and PVE = 0.6.

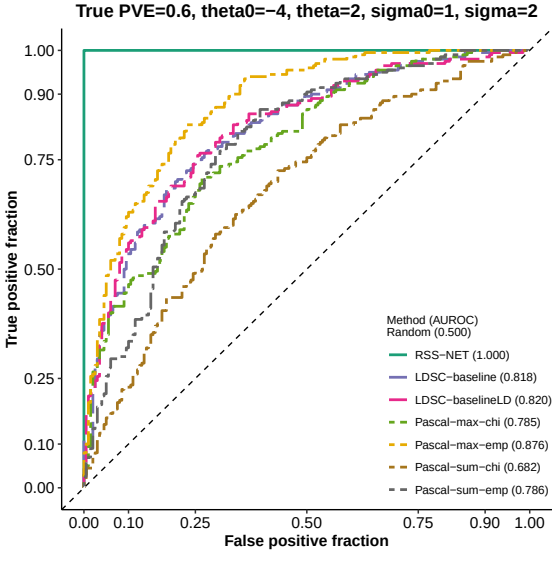

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.996)
LDSC−baseline (0.601)
LDSC−baselineLD (0.609)
Pascal−max−chi (0.523)
Pascal−max−emp (0.523)
Pascal−sum−chi (0.553)
Pascal−sum−emp (0.559)


True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
LDSC−baseline (0.839)
LDSC−baselineLD (0.849)
Pascal−max−chi (0.633)
Pascal−max−emp (0.694)
Pascal−sum−chi (0.669)
Pascal−sum−emp (0.755)


True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.998)
LDSC−baseline (0.620)
LDSC−baselineLD (0.621)
Pascal−max−chi (0.576)
Pascal−max−emp (0.590)
Pascal−sum−chi (0.606)
Pascal−sum−emp (0.615)


True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
LDSC−baseline (0.879)
LDSC−baselineLD (0.886)
Pascal−max−chi (0.719)
Pascal−max−emp (0.819)
Pascal−sum−chi (0.745)
Pascal−sum−emp (0.825)


True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.998)
LDSC−baseline (0.638)
LDSC−baselineLD (0.613)
Pascal−max−chi (0.645)
Pascal−max−emp (0.737)
Pascal−sum−chi (0.599)
Pascal−sum−emp (0.641)


True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
LDSC−baseline (0.818)
LDSC−baselineLD (0.820)
Pascal−max−chi (0.785)
Pascal−max−emp (0.876)
Pascal−sum−chi (0.682)
Pascal−sum−emp (0.786)

Supplementary Figure 4

Simulation details and additional results of Figure 3(c). This part aims to assess the robustness of RSS-NET to model mis-specification where an edge-altered version of the true network is enriched for genetic association in baseline datasets. Details of this part are almost identical to those in **Supplementary Figure 1**. Here we only highlight the differences.

We simulate random edge-altered networks on the basis of the B cell regulatory network defined in **Supplementary Figure 1**. Specifically, we keep all REs and genes of the B cell network, remove the actual connections between TFs and TGs, and then create fake edges by randomly connecting TFs to TGs with random edge weights. We ensure that the edge-altered network has the same number of edges as the actual B cell network, and their edge weights have the same range. For each baseline dataset, we simulate a new edge-altered network.

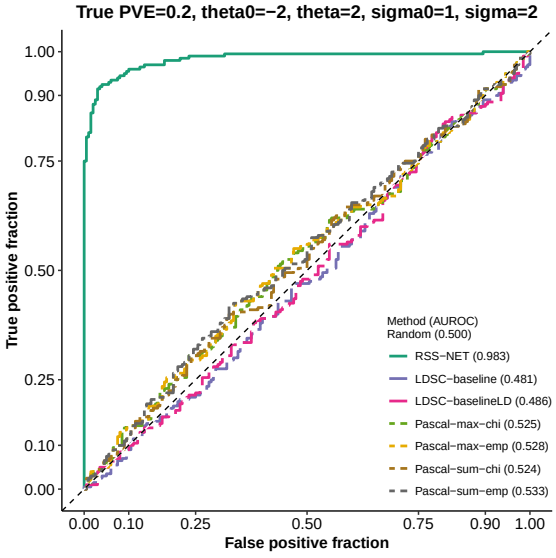
We first simulate enrichment datasets as in **Supplementary Figure 1**. Since the true and edge-altered networks have the same REs and genes, we let a baseline dataset and its corresponding enrichment dataset have the same causal indicators $\{\zeta_j\}$. We then simulate true genetic effects $\boldsymbol{\beta}$ in a baseline dataset as:

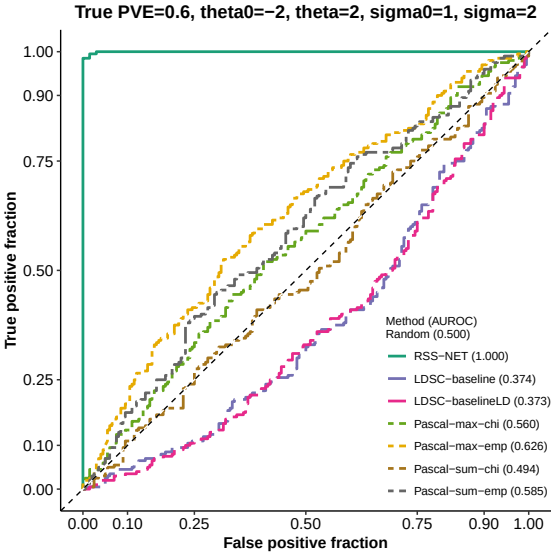
$$\begin{aligned} \beta_j \mid \zeta_j = 0 &\sim \delta_0, \\ \beta_j \mid \zeta_j = 1 &\sim \mathcal{N}(0,\ \sigma_0^2 + \sigma^2 \cdot \textstyle\sum_{g \in \mathbf{O}_j^\dagger} (w_{jg}^\dagger)^2), \end{aligned}$$

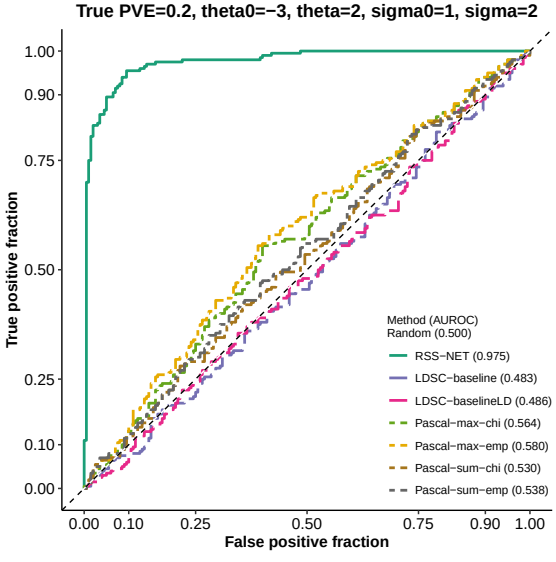
where $\{\mathbf{O}_j^\dagger, w_{jg}^\dagger\}$ are given by the random edge-altered network. For all datasets the candidate of enrichment testing is the B cell network.

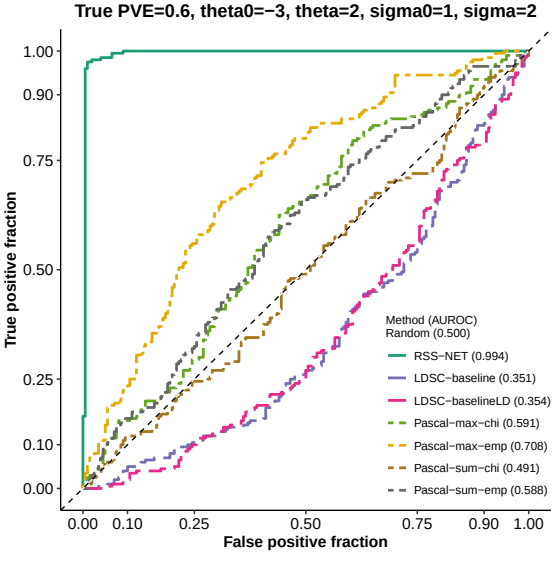
Captions of the following plots show the true parameter values for enrichment datasets. The "sparse scenario" in **Figure 3(c)** corresponds to simulations with true $\theta_0 = -4$ and PVE = 0.6. The "polygenic scenario" in **Figure 3(c)** corresponds to simulations with true $\theta_0 = -2$ and PVE = 0.6.

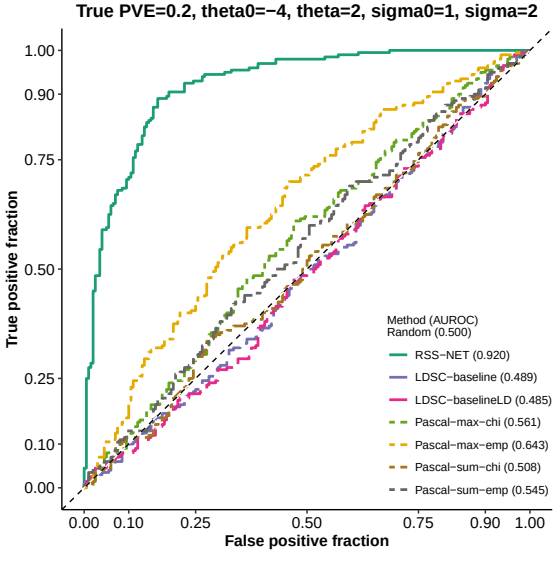
The true and random edge-altered networks have the same REs and genes but totally different topology (TF-TG edges and edge weights). Existing methods like LDSC and Pascal only use node information (REs and/or genes), and thus they cannot distinguish the true and random edge-altered networks. In contrast, because of the edge-enrichment parameter σ^2, RSS-NET is able to capture the topological differences among networks, and thus it can reliably distinguish true enrichments of the B cell network from enrichments of its edge-altered counterparts.

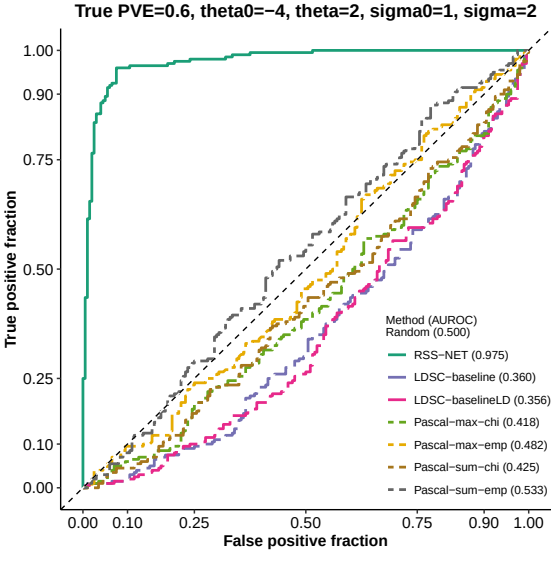

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.983)
LDSC−baseline (0.481)
LDSC−baselineLD (0.486)
Pascal−max−chi (0.525)
Pascal−max−emp (0.528)
Pascal−sum−chi (0.524)
Pascal−sum−emp (0.533)


True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (1.000)
LDSC−baseline (0.374)
LDSC−baselineLD (0.373)
Pascal−max−chi (0.560)
Pascal−max−emp (0.626)
Pascal−sum−chi (0.494)
Pascal−sum−emp (0.585)


True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.975)
LDSC−baseline (0.483)
LDSC−baselineLD (0.486)
Pascal−max−chi (0.564)
Pascal−max−emp (0.580)
Pascal−sum−chi (0.530)
Pascal−sum−emp (0.538)


True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.994)
LDSC−baseline (0.351)
LDSC−baselineLD (0.354)
Pascal−max−chi (0.591)
Pascal−max−emp (0.708)
Pascal−sum−chi (0.491)
Pascal−sum−emp (0.588)


True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.920)
LDSC−baseline (0.489)
LDSC−baselineLD (0.485)
Pascal−max−chi (0.561)
Pascal−max−emp (0.643)
Pascal−sum−chi (0.508)
Pascal−sum−emp (0.545)


True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.975)
LDSC−baseline (0.360)
LDSC−baselineLD (0.356)
Pascal−max−chi (0.418)
Pascal−max−emp (0.482)
Pascal−sum−chi (0.425)
Pascal−sum−emp (0.533)

Supplementary Figure 5

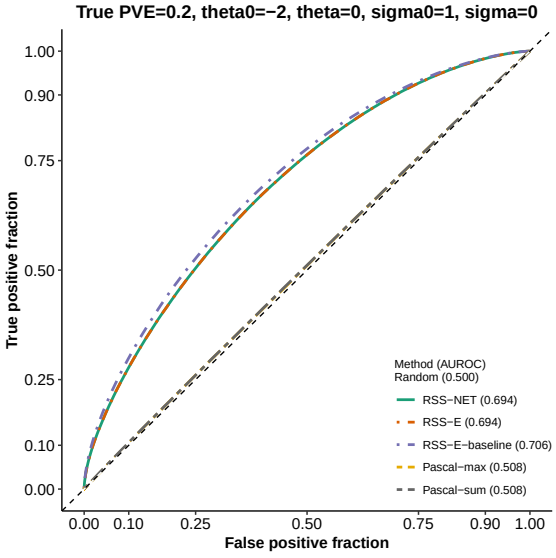
Simulation details and additional results of Figure 4. Here we compare RSS-NET with existing gene-level association testing methods on the same simulated summary data used in **Supplementary Figure 1**. Details of this part are almost identical to those in **Supplementary Figure 1**. Here we only highlight the differences.

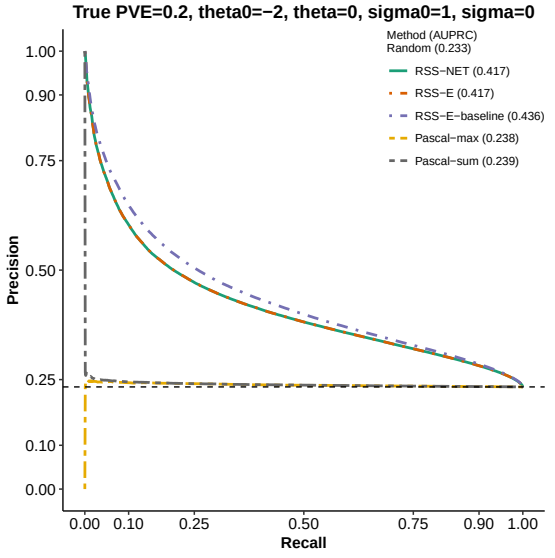
For each simulated dataset, we define a gene as "trait-associated" if at lease one SNP j within 100 kb of the transcribed region of this gene has non-zero effect ($\beta_j \neq 0$). For each gene in each simulated dataset, RSS-NET and RSS-E (Zhu and Stephens 2018) methods produce P_1, the posterior probability that the gene is trait-associated, whereas Pascal (Lamparter et al. 2016) methods produce p-value with the null hypothesis that the gene is not trait-associated; these statistics are used to rank the significance of gene-level associations. If a method identifies association between a non-trait-associated gene and the trait, then it is a "false positive". If a method identifies association between a trait-associated gene and the trait, then it is a "true positive".

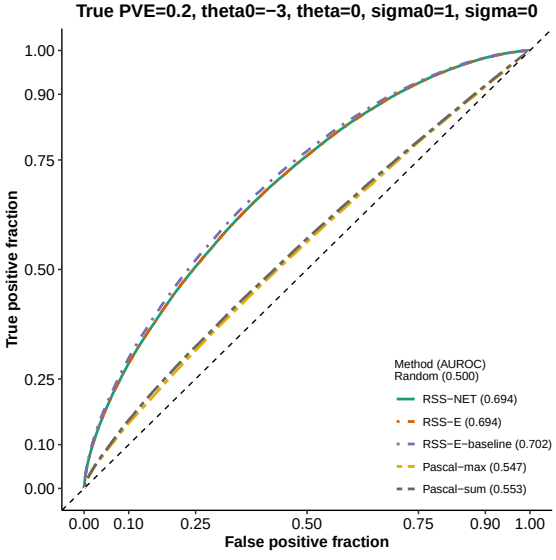
Here the true values of baseline proportion parameter θ_0 are $\{-2,-3,-4\}$, the true values of enrichment proportion parameter θ are $\{0,2\}$, the true value of baseline magnitude parameter σ_0 is 1, the true values of enrichment magnitude parameter σ are $\{0,2\}$, and the true values of PVE are $\{0.2,0.6\}$. In total there are 24 ($3\times2\times2\times2$) simulation scenarios. Each scenario contains 200 independent datasets. For all datasets we test associations of 16,954 autosomal protein-coding genes that are available in Pascal-associated database.

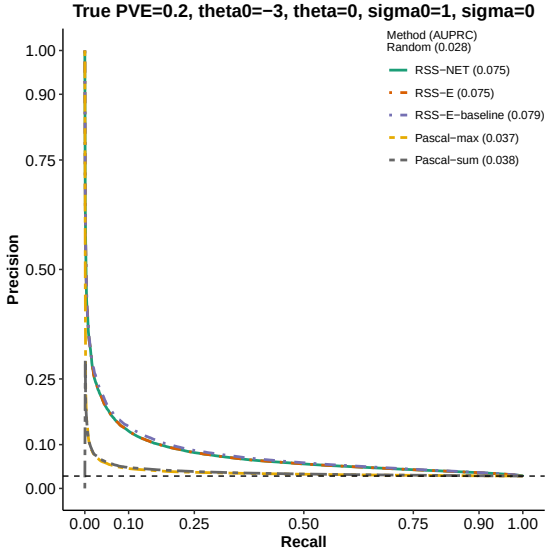
Captions of the following plots show the true parameter values for simulations. All panels of **Figure 4** correspond to simulations with true $\theta_0 = -2$, $\sigma_0 = 1$ and PVE = 0.6. **Figure 4(a)** corresponds to simulations with true $\theta = 0$ and $\sigma = 0$. **Figure 4(b)** corresponds to simulations with true $\theta = 0$ and $\sigma = 2$. **Figure 4(c)** corresponds to simulations with true $\theta = 2$ and $\sigma = 0$. **Figure 4(d)** corresponds to simulations with true $\theta = 2$ and $\sigma = 2$.

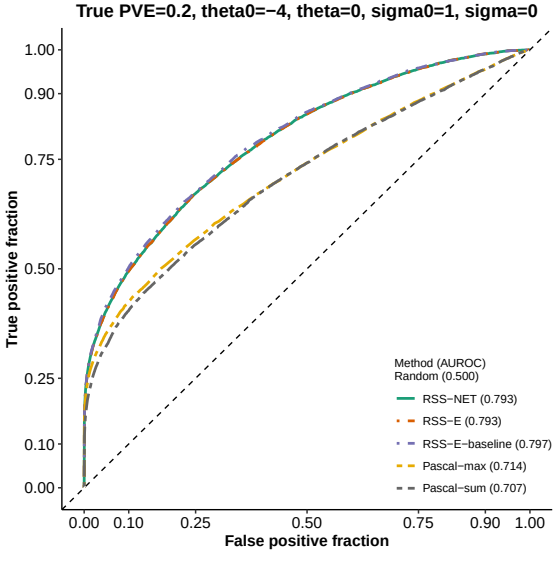
We evaluate the performance of these gene-level association methods by plotting the receiver operating characteristic (ROC) curve and computing the area under the ROC curve (AUROC) for each method. Both metrics are implemented in the R package `plotROC`. Unlike the balanced enrichment simulations, the number of trait-associated genes (true labels) and the number of non-trait-associated genes (false labels) are often different in all simulated datasets. Due to this imbalance nature, we also plot the precision-recall (PRC) curve and compute the area under the PRC curve (AUPRC) for each method. Both metrics are implemented in the R package `precrec`.

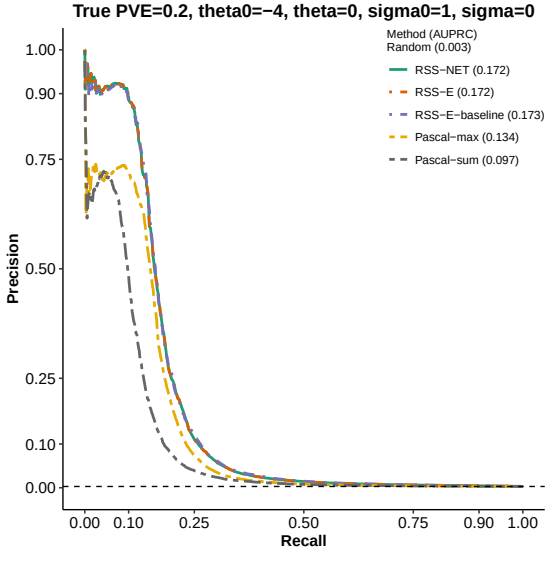

True PVE=0.2, theta0=−2, theta=0, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.694)
RSS−E (0.694)
RSS−E−baseline (0.706)
Pascal−max (0.508)
Pascal−sum (0.508)

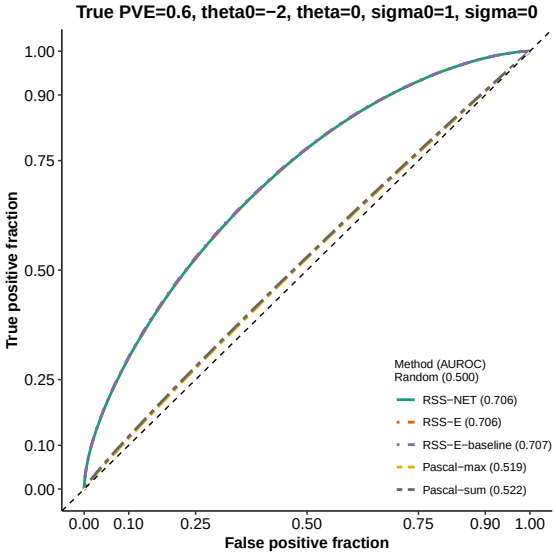

True PVE=0.2, theta0=−2, theta=0, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.233)
RSS−NET (0.417)
RSS−E (0.417)
RSS−E−baseline (0.436)
Pascal−max (0.238)
Pascal−sum (0.239)

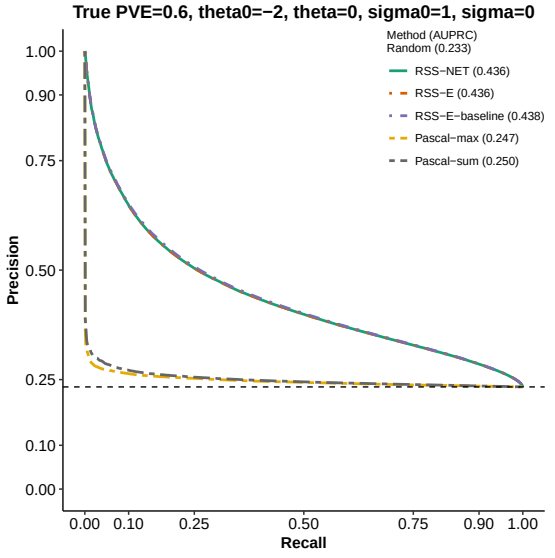

True PVE=0.2, theta0=−3, theta=0, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.694)
RSS−E (0.694)
RSS−E−baseline (0.702)
Pascal−max (0.547)
Pascal−sum (0.553)

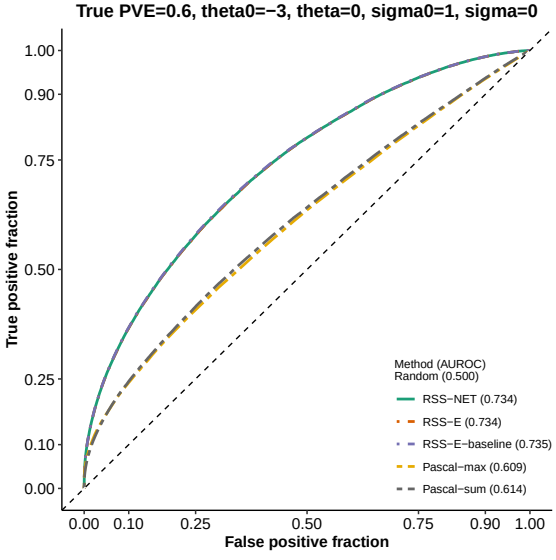

True PVE=0.2, theta0=−3, theta=0, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.028)
RSS−NET (0.075)
RSS−E (0.075)
RSS−E−baseline (0.079)
Pascal−max (0.037)
Pascal−sum (0.038)

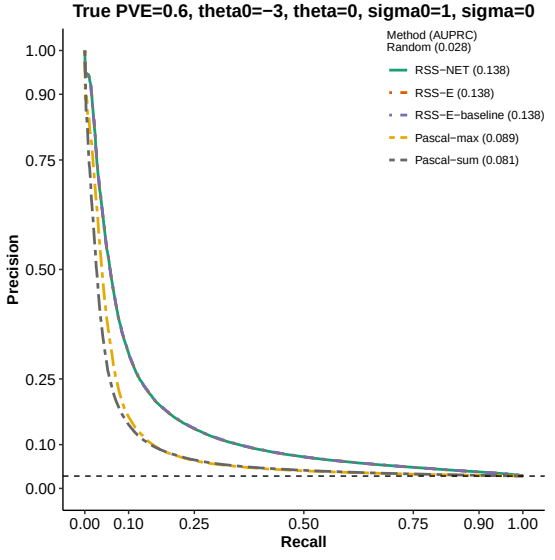

True PVE=0.2, theta0=−4, theta=0, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.793)
RSS−E (0.793)
RSS−E−baseline (0.797)
Pascal−max (0.714)
Pascal−sum (0.707)

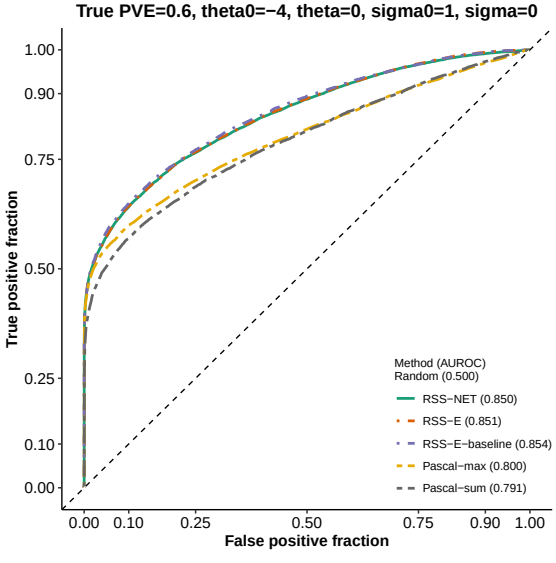

True PVE=0.2, theta0=−4, theta=0, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.003)
RSS−NET (0.172)
RSS−E (0.172)
RSS−E−baseline (0.173)
Pascal−max (0.134)
Pascal−sum (0.097)

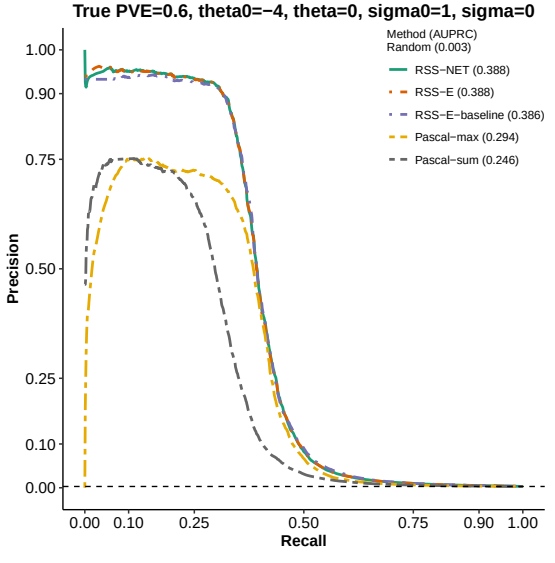

True PVE=0.6, theta0=−2, theta=0, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.706)
RSS−E (0.706)
RSS−E−baseline (0.707)
Pascal−max (0.519)
Pascal−sum (0.522)

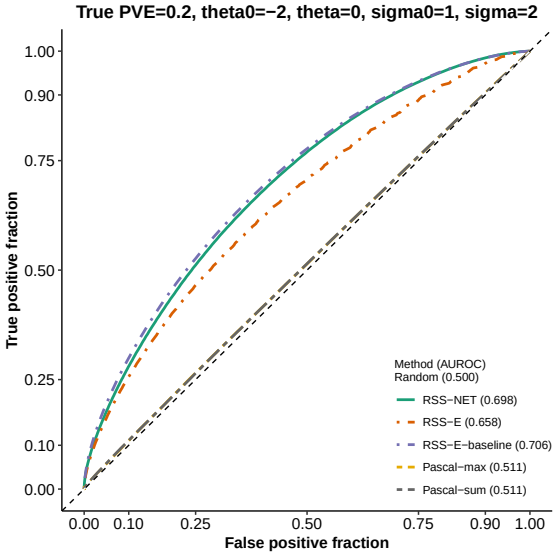

True PVE=0.6, theta0=−2, theta=0, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.233)
RSS−NET (0.436)
RSS−E (0.436)
RSS−E−baseline (0.438)
Pascal−max (0.247)
Pascal−sum (0.250)

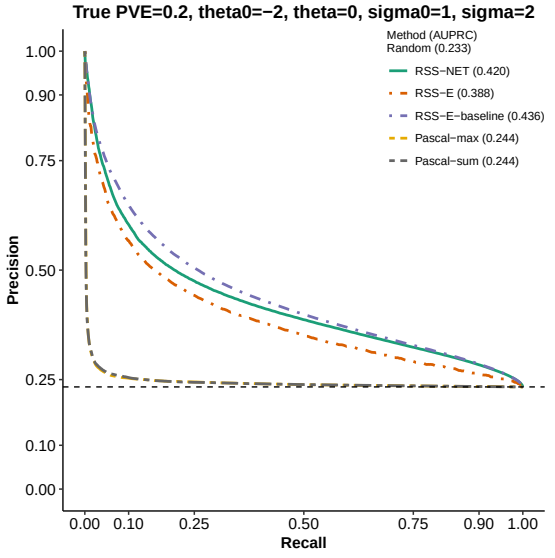

True PVE=0.6, theta0=−3, theta=0, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.734)
RSS−E (0.734)
RSS−E−baseline (0.735)
Pascal−max (0.609)
Pascal−sum (0.614)

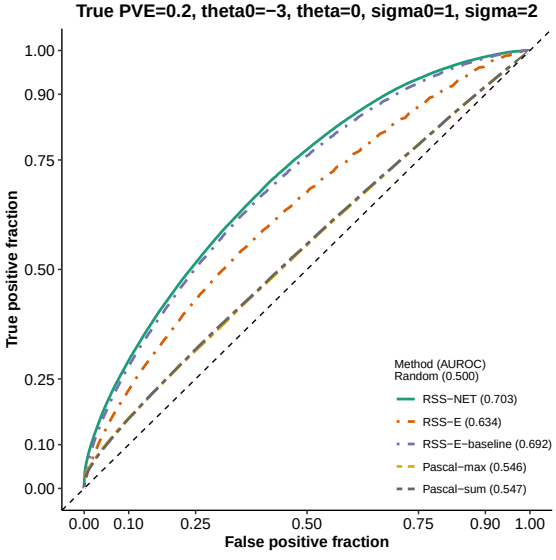

True PVE=0.6, theta0=−3, theta=0, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.028)
RSS−NET (0.138)
RSS−E (0.138)
RSS−E−baseline (0.138)
Pascal−max (0.089)
Pascal−sum (0.081)

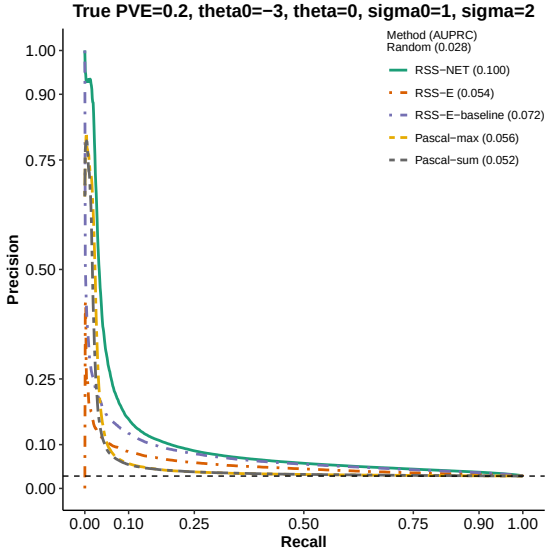

True PVE=0.6, theta0=−4, theta=0, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.850)
RSS−E (0.851)
RSS−E−baseline (0.854)
Pascal−max (0.800)
Pascal−sum (0.791)

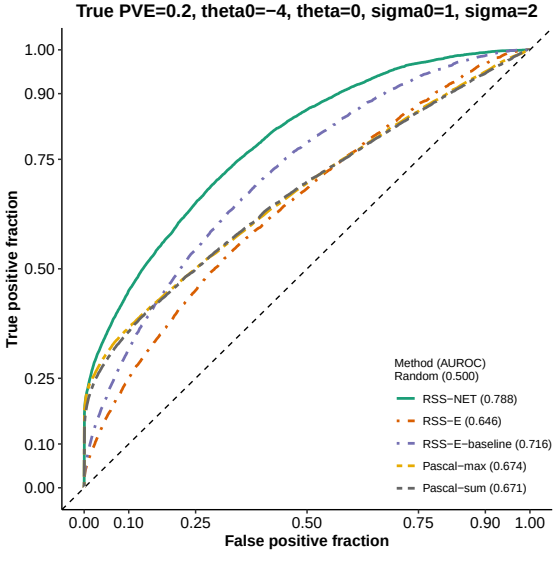

True PVE=0.6, theta0=−4, theta=0, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.003)
RSS−NET (0.388)
RSS−E (0.388)
RSS−E−baseline (0.386)
Pascal−max (0.294)
Pascal−sum (0.246)

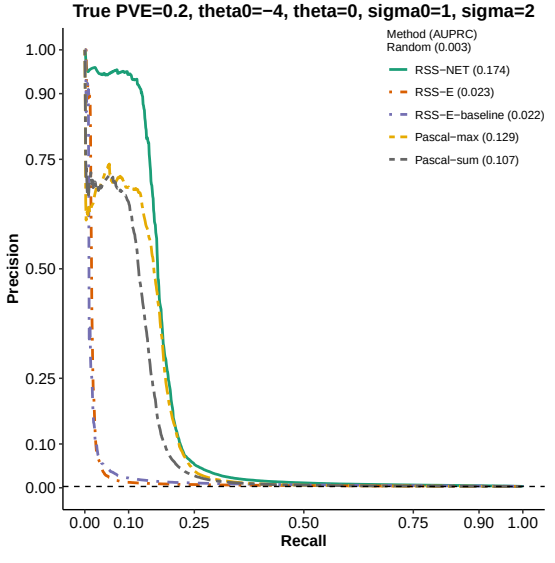

True PVE=0.2, theta0=−2, theta=0, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.698)
RSS−E (0.658)
RSS−E−baseline (0.706)
Pascal−max (0.511)
Pascal−sum (0.511)

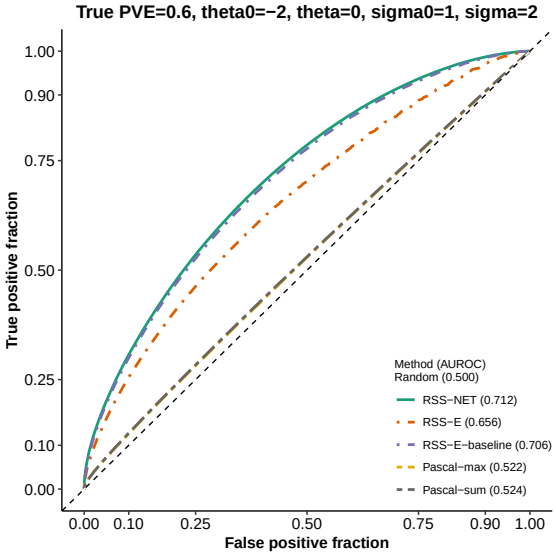

True PVE=0.2, theta0=−2, theta=0, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.233)
RSS−NET (0.420)
RSS−E (0.388)
RSS−E−baseline (0.436)
Pascal−max (0.244)
Pascal−sum (0.244)

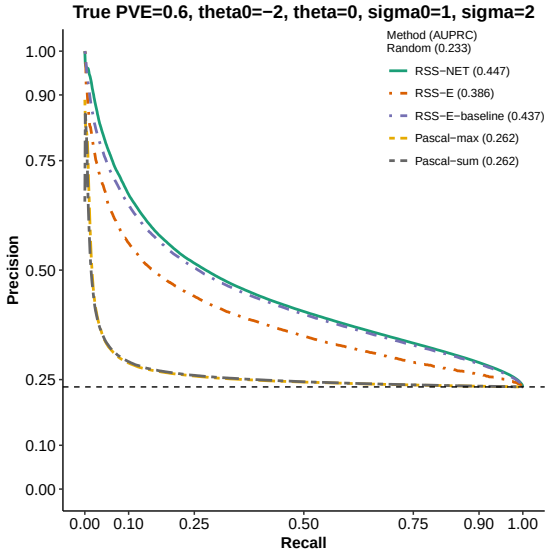

True PVE=0.2, theta0=−3, theta=0, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.703)
RSS−E (0.634)
RSS−E−baseline (0.692)
Pascal−max (0.546)
Pascal−sum (0.547)

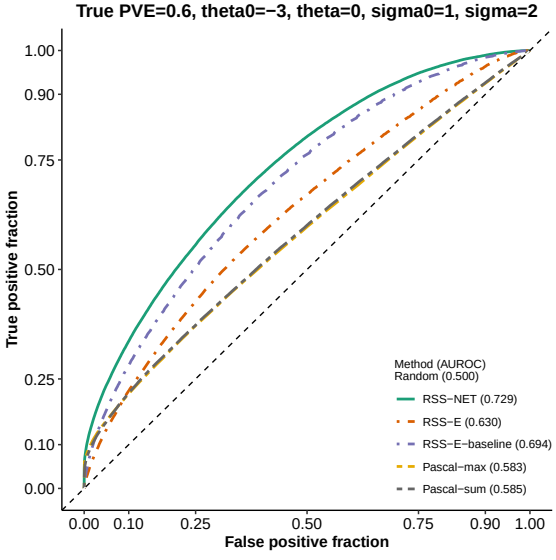

True PVE=0.2, theta0=−3, theta=0, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.028)
RSS−NET (0.100)
RSS−E (0.054)
RSS−E−baseline (0.072)
Pascal−max (0.056)
Pascal−sum (0.052)

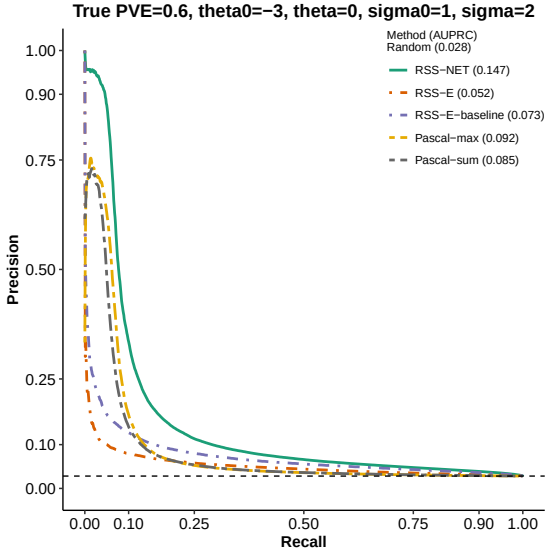

True PVE=0.2, theta0=−4, theta=0, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.788)
RSS−E (0.646)
RSS−E−baseline (0.716)
Pascal−max (0.674)
Pascal−sum (0.671)

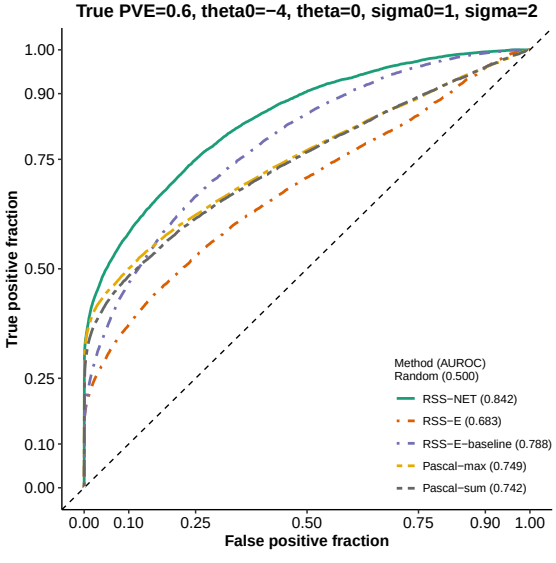

True PVE=0.2, theta0=−4, theta=0, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.003)
RSS−NET (0.174)
RSS−E (0.023)
RSS−E−baseline (0.022)
Pascal−max (0.129)
Pascal−sum (0.107)

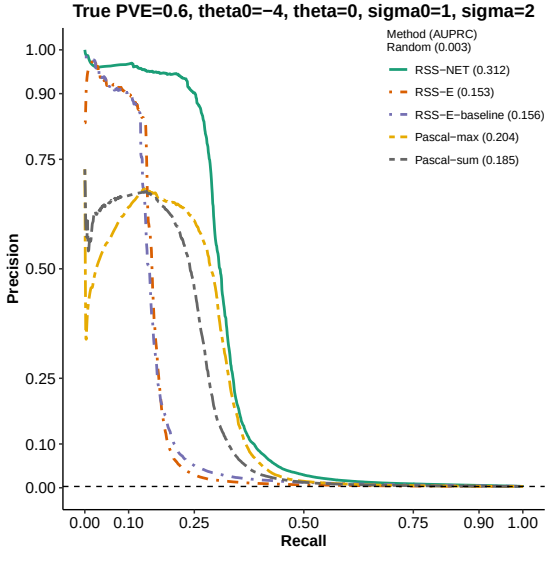

True PVE=0.6, theta0=−2, theta=0, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.712)
RSS−E (0.656)
RSS−E−baseline (0.706)
Pascal−max (0.522)
Pascal−sum (0.524)

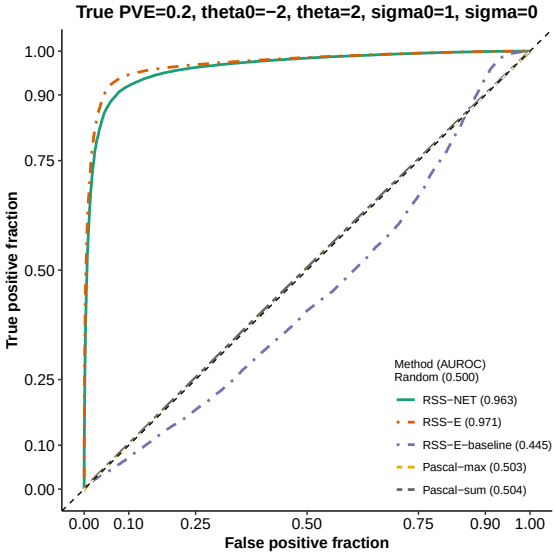

True PVE=0.6, theta0=−2, theta=0, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.233)
RSS−NET (0.447)
RSS−E (0.386)
RSS−E−baseline (0.437)
Pascal−max (0.262)
Pascal−sum (0.262)

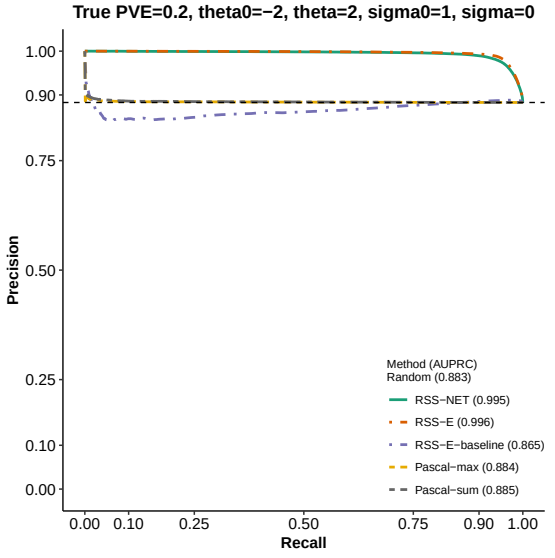

True PVE=0.6, theta0=−3, theta=0, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.729)
RSS−E (0.630)
RSS−E−baseline (0.694)
Pascal−max (0.583)
Pascal−sum (0.585)

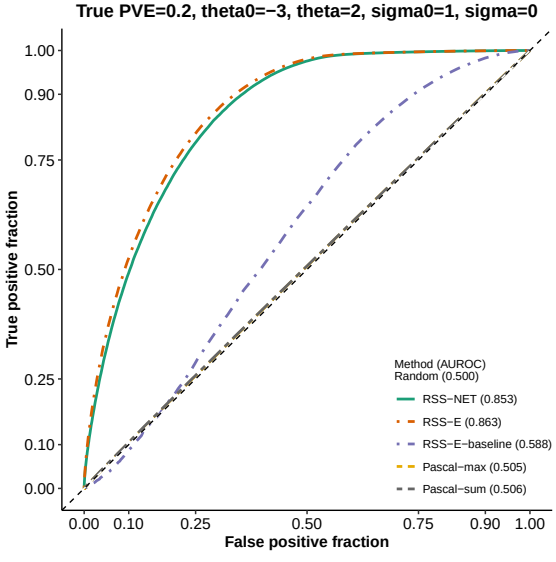

True PVE=0.6, theta0=−3, theta=0, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.028)
RSS−NET (0.147)
RSS−E (0.052)
RSS−E−baseline (0.073)
Pascal−max (0.092)
Pascal−sum (0.085)

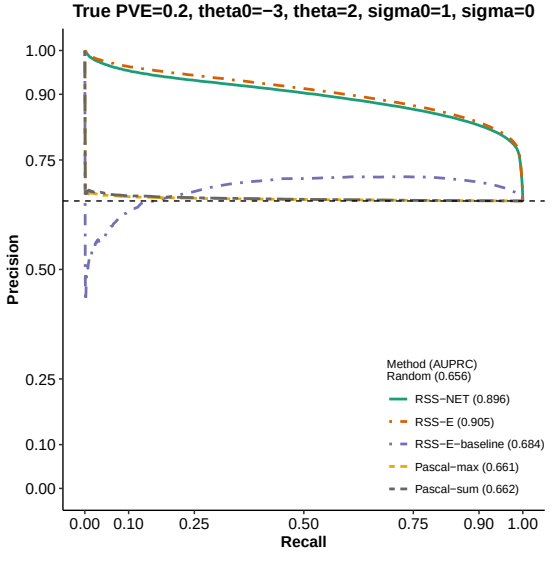

True PVE=0.6, theta0=−4, theta=0, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.842)
RSS−E (0.683)
RSS−E−baseline (0.788)
Pascal−max (0.749)
Pascal−sum (0.742)

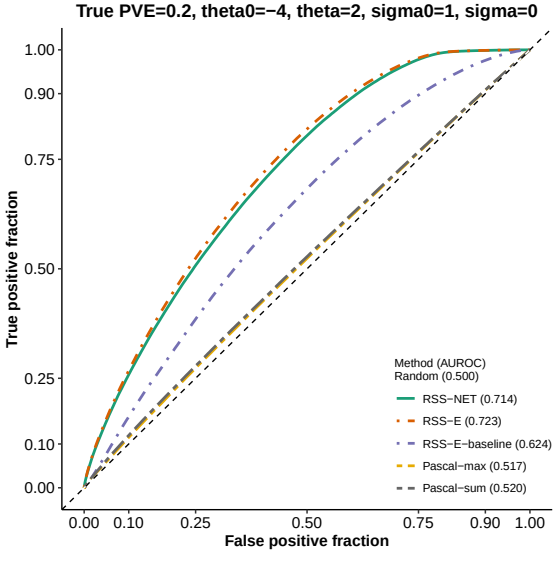

True PVE=0.6, theta0=−4, theta=0, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.003)
RSS−NET (0.312)
RSS−E (0.153)
RSS−E−baseline (0.156)
Pascal−max (0.204)
Pascal−sum (0.185)

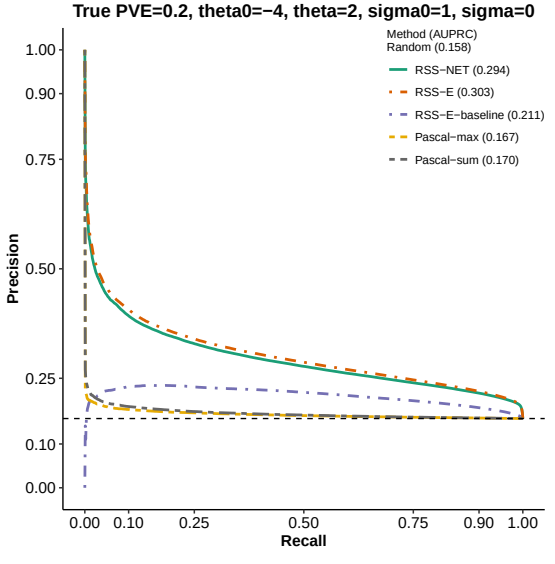

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.963)
RSS−E (0.971)
RSS−E−baseline (0.445)
Pascal−max (0.503)
Pascal−sum (0.504)

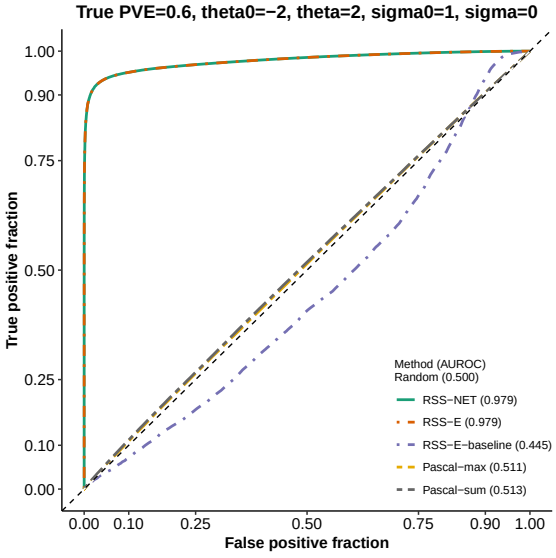

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.883)
RSS−NET (0.995)
RSS−E (0.996)
RSS−E−baseline (0.865)
Pascal−max (0.884)
Pascal−sum (0.885)

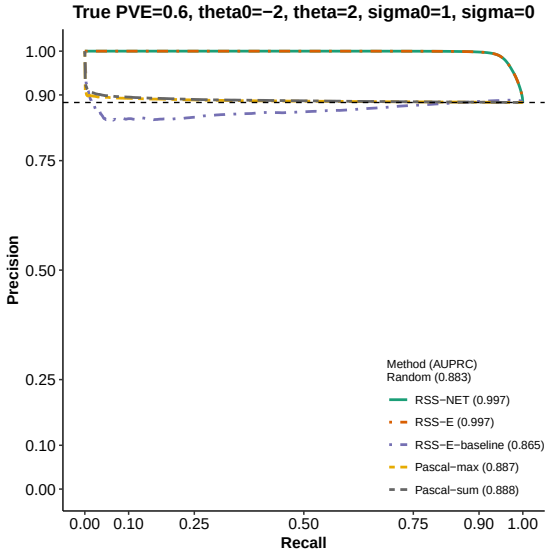

True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.853)
RSS−E (0.863)
RSS−E−baseline (0.588)
Pascal−max (0.505)
Pascal−sum (0.506)

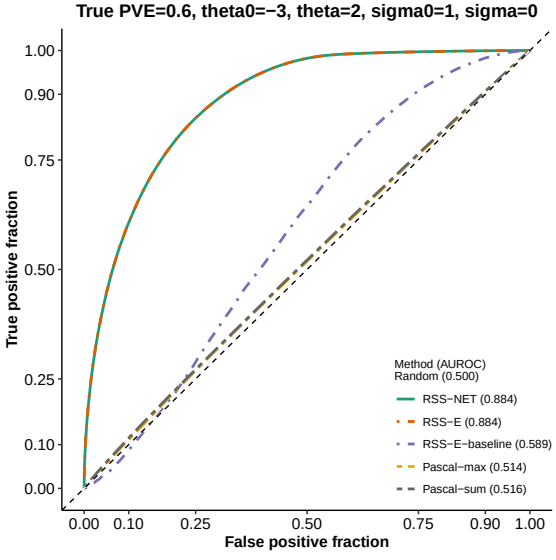

True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.656)
RSS−NET (0.896)
RSS−E (0.905)
RSS−E−baseline (0.684)
Pascal−max (0.661)
Pascal−sum (0.662)

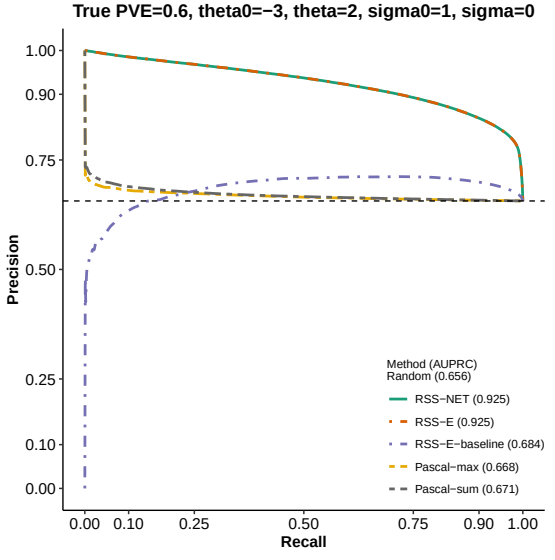

True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.714)
RSS−E (0.723)
RSS−E−baseline (0.624)
Pascal−max (0.517)
Pascal−sum (0.520)

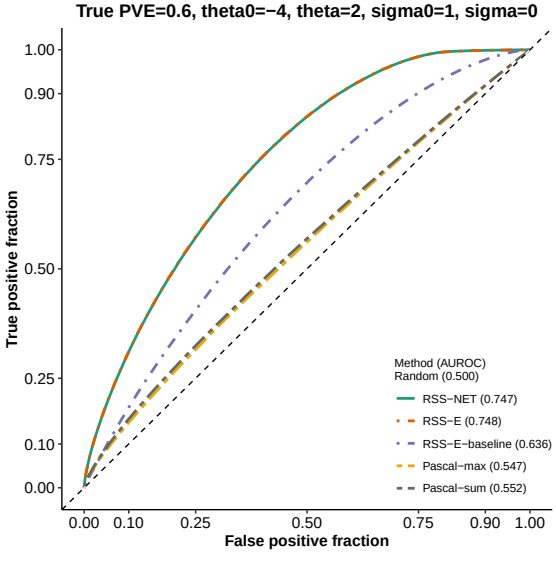

True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.158)
RSS−NET (0.294)
RSS−E (0.303)
RSS−E−baseline (0.211)
Pascal−max (0.167)
Pascal−sum (0.170)

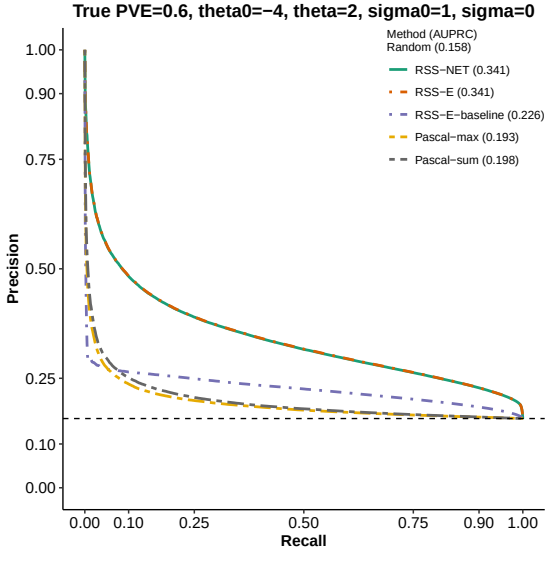

True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.979)
RSS−E (0.979)
RSS−E−baseline (0.445)
Pascal−max (0.511)
Pascal−sum (0.513)

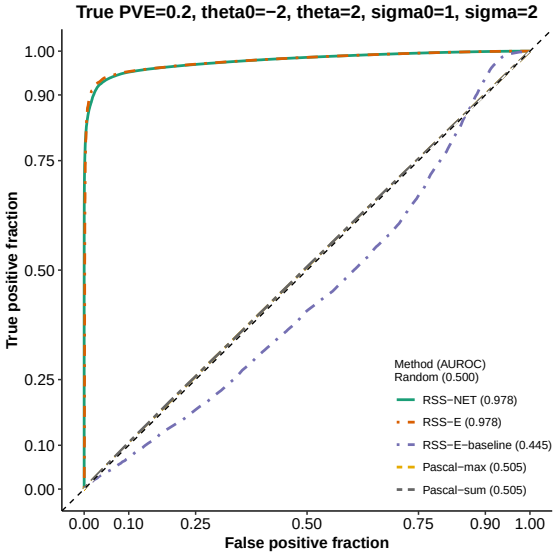

True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.883)
RSS−NET (0.997)
RSS−E (0.997)
RSS−E−baseline (0.865)
Pascal−max (0.887)
Pascal−sum (0.888)

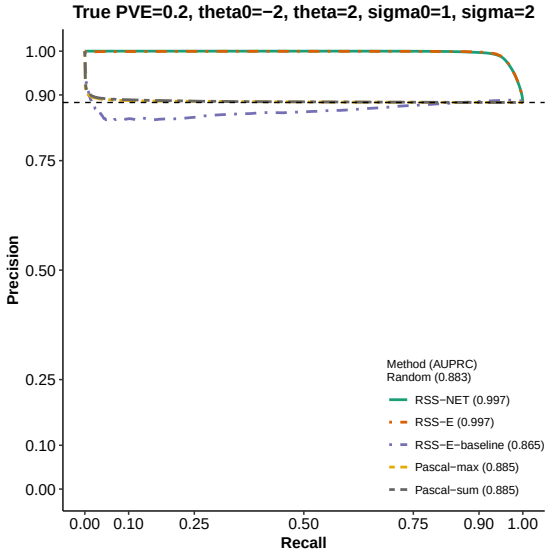

True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.884)
RSS−E (0.884)
RSS−E−baseline (0.589)
Pascal−max (0.514)
Pascal−sum (0.516)

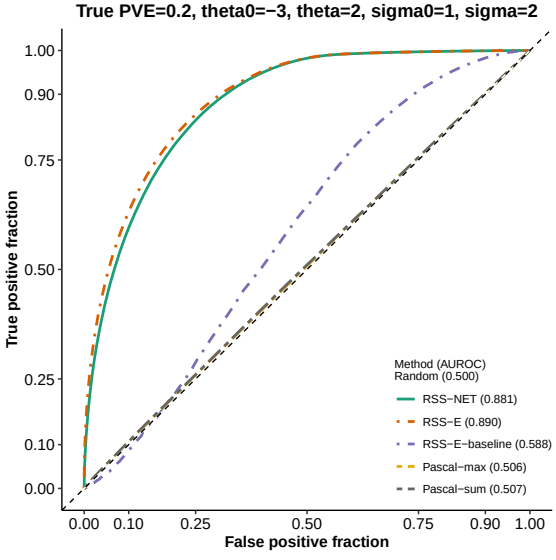

True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.656)
RSS−NET (0.925)
RSS−E (0.925)
RSS−E−baseline (0.684)
Pascal−max (0.668)
Pascal−sum (0.671)

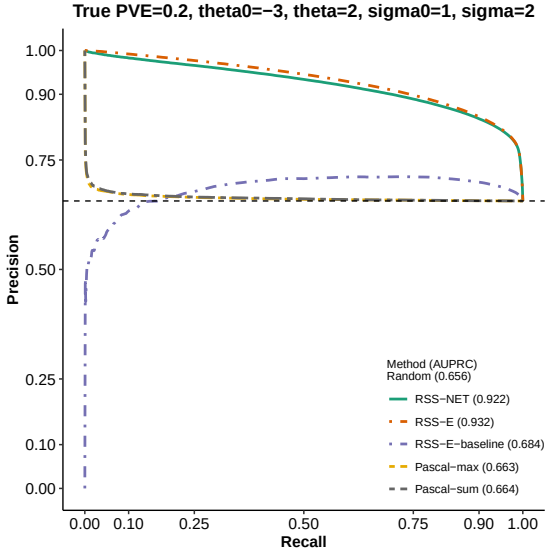

True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=0
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.747)
RSS−E (0.748)
RSS−E−baseline (0.636)
Pascal−max (0.547)
Pascal−sum (0.552)

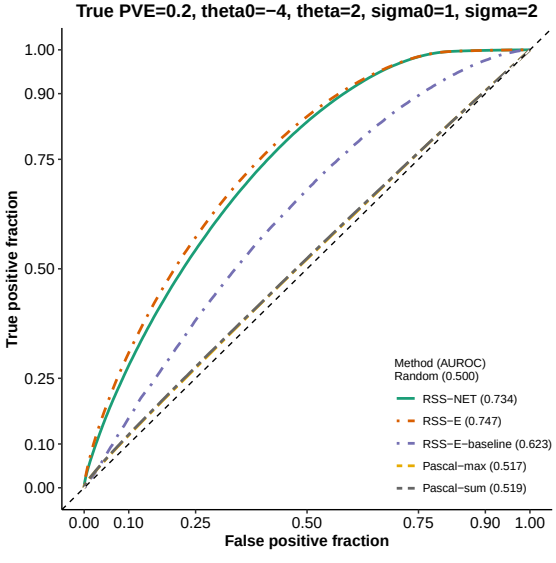

True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=0
Precision
Recall
Method (AUPRC)
Random (0.158)
RSS−NET (0.341)
RSS−E (0.341)
RSS−E−baseline (0.226)
Pascal−max (0.193)
Pascal−sum (0.198)

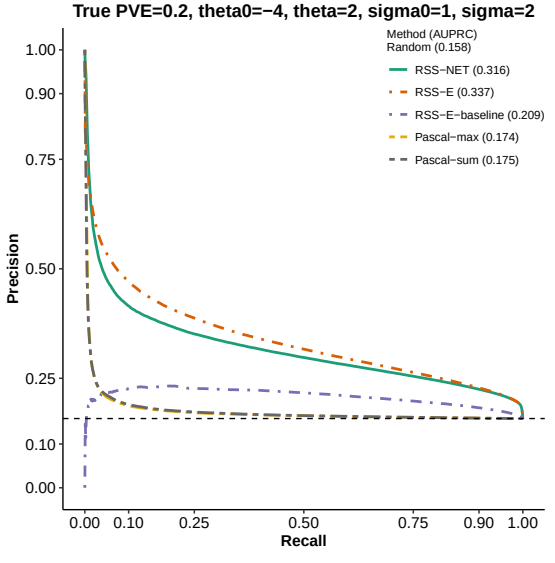

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.978)
RSS−E (0.978)
RSS−E−baseline (0.445)
Pascal−max (0.505)
Pascal−sum (0.505)

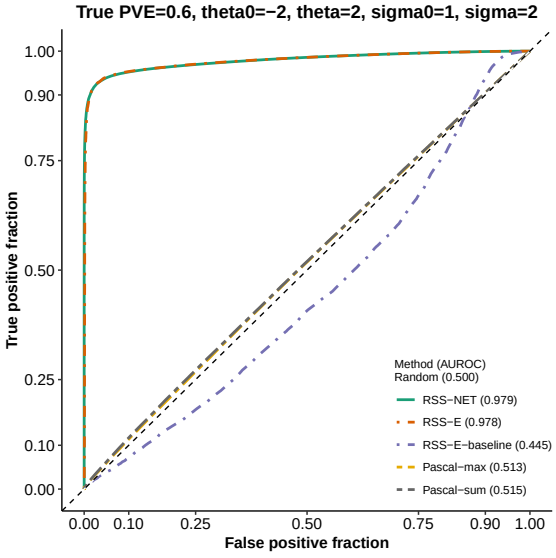

True PVE=0.2, theta0=−2, theta=2, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.883)
RSS−NET (0.997)
RSS−E (0.997)
RSS−E−baseline (0.865)
Pascal−max (0.885)
Pascal−sum (0.885)

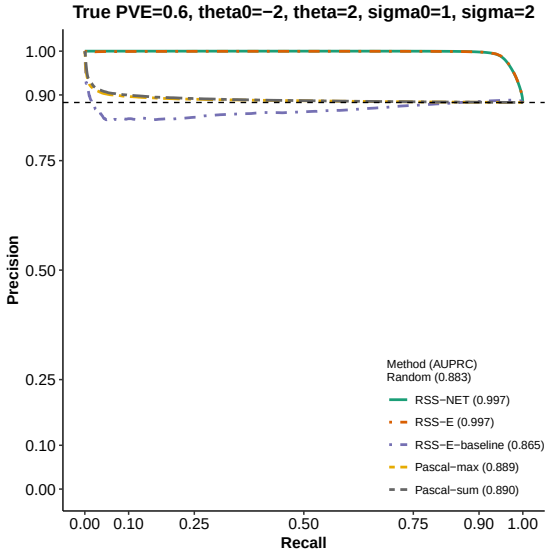

True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.881)
RSS−E (0.890)
RSS−E−baseline (0.588)
Pascal−max (0.506)
Pascal−sum (0.507)

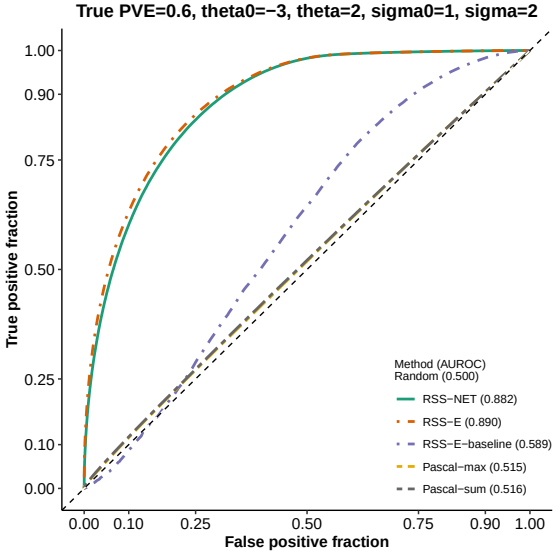

True PVE=0.2, theta0=−3, theta=2, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.656)
RSS−NET (0.922)
RSS−E (0.932)
RSS−E−baseline (0.684)
Pascal−max (0.663)
Pascal−sum (0.664)

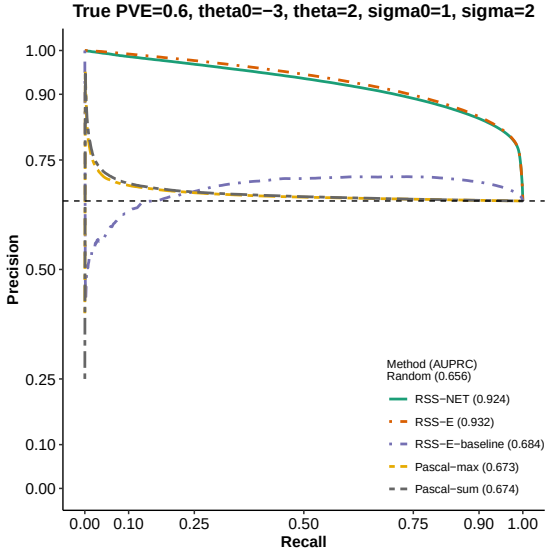

True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
Method (AUROC)
Random (0.500)
RSS−NET (0.734)
RSS−E (0.747)
RSS−E−baseline (0.623)
Pascal−max (0.517)
Pascal−sum (0.519)

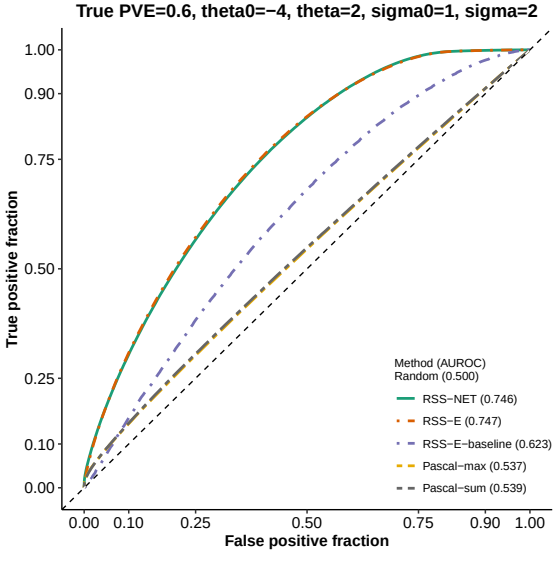

True PVE=0.2, theta0=−4, theta=2, sigma0=1, sigma=2
Precision
Recall
Method (AUPRC)
Random (0.158)
RSS−NET (0.316)
RSS−E (0.337)
RSS−E−baseline (0.209)
Pascal−max (0.174)
Pascal−sum (0.175)

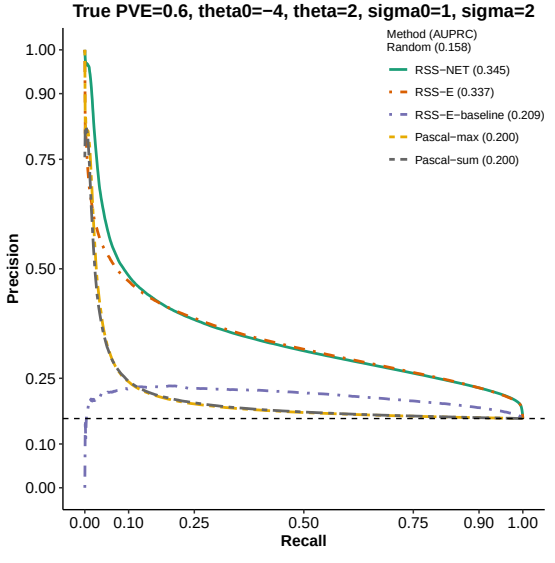

True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.979)
RSS−E (0.978)
RSS−E−baseline (0.445)
Pascal−max (0.513)
Pascal−sum (0.515)


True PVE=0.6, theta0=−2, theta=2, sigma0=1, sigma=2
Precision
Recall
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUPRC)
Random (0.883)
RSS−NET (0.997)
RSS−E (0.997)
RSS−E−baseline (0.865)
Pascal−max (0.889)
Pascal−sum (0.890)


True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.882)
RSS−E (0.890)
RSS−E−baseline (0.589)
Pascal−max (0.515)
Pascal−sum (0.516)


True PVE=0.6, theta0=−3, theta=2, sigma0=1, sigma=2
Precision
Recall
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUPRC)
Random (0.656)
RSS−NET (0.924)
RSS−E (0.932)
RSS−E−baseline (0.684)
Pascal−max (0.673)
Pascal−sum (0.674)


True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=2
True positive fraction
False positive fraction
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUROC)
Random (0.500)
RSS−NET (0.746)
RSS−E (0.747)
RSS−E−baseline (0.623)
Pascal−max (0.537)
Pascal−sum (0.539)


True PVE=0.6, theta0=−4, theta=2, sigma0=1, sigma=2
Precision
Recall
0.00
0.10
0.25
0.50
0.75
0.90
1.00
Method (AUPRC)
Random (0.158)
RSS−NET (0.345)
RSS−E (0.337)
RSS−E−baseline (0.209)
Pascal−max (0.200)
Pascal−sum (0.200)

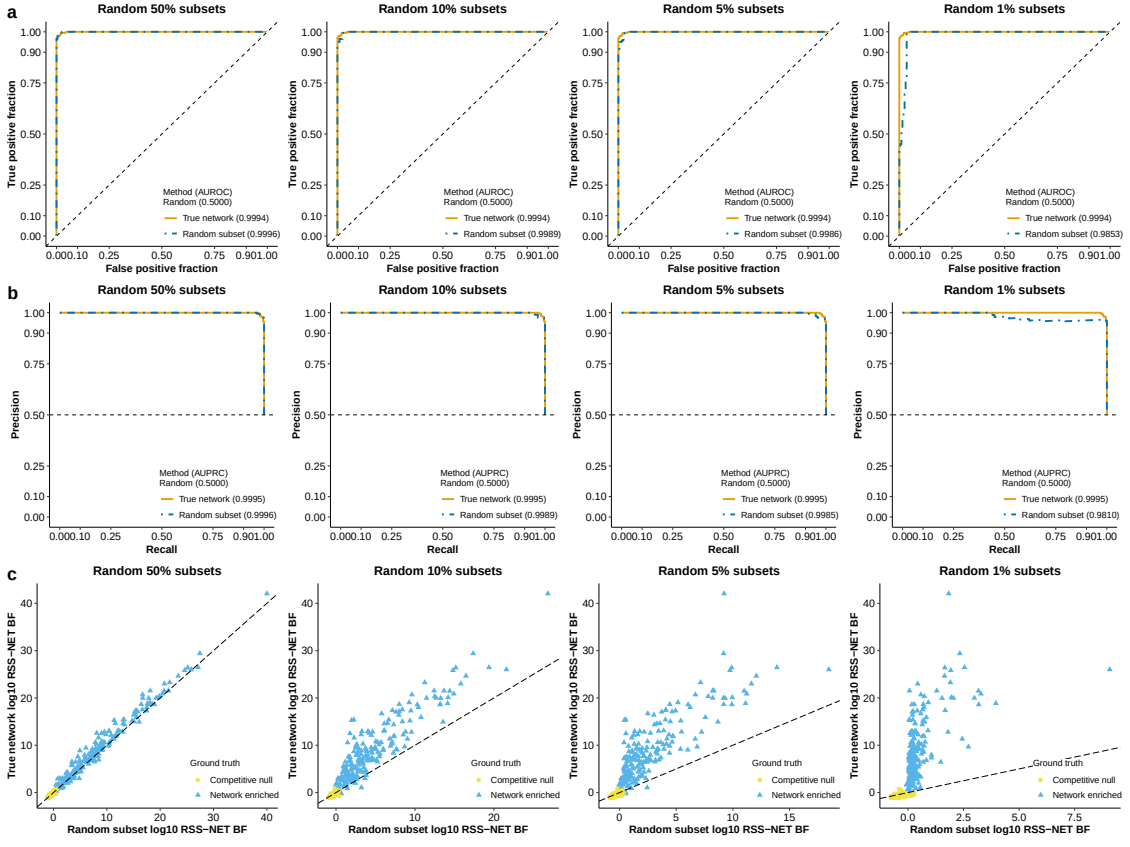
Supplementary Figure 6

Robustness of RSS-NET under noisy networks. This part aims to assess the robustness of RSS-NET to model mis-specification where the GWAS summary data are simulated from a real target network and the enrichment testing is performed on a noisy version of this real target. Details of this part are almost identical to those in **Supplementary Figure 1**. Here we only highlight the differences.

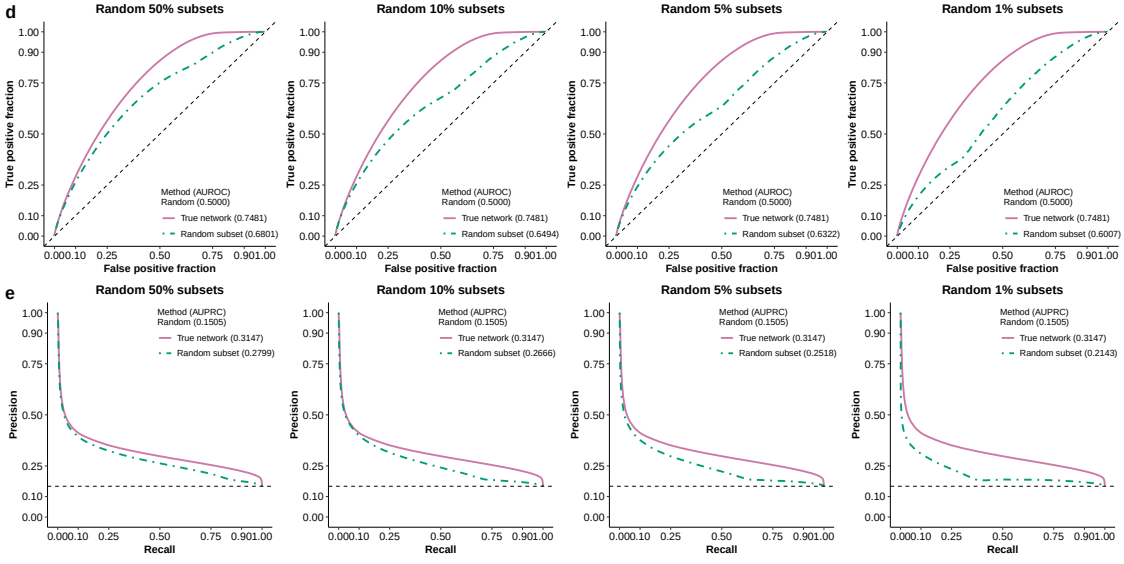
The real target network used to simulate GWAS summary data is the B cell regulatory network in **Supplementary Figure 1**. We create a noisy network for this target by randomly subsetting a given fraction of its TF-TG edges and the associated REs. Here the fraction of TF-TG edges to subset are $\{1\%, 5\%, 10\%, 50\%\}$. For each paired baseline-enrichment GWAS dataset, we simulate a new noisy network and use it for RSS-NET analysis.

The GWAS summary data (both baseline and enrichment) used here are the same as those used in **Supplementary Figure 1**, with true $\theta_0 = -4$, $\theta = 2$, $\sigma_0 = 1$, $\sigma = 2$ and PVE = 0.2. Unlike **Supplementary Figure 1**, here we do not test the enrichment of B cell network (i.e. the real target network), but test its random subsets instead.

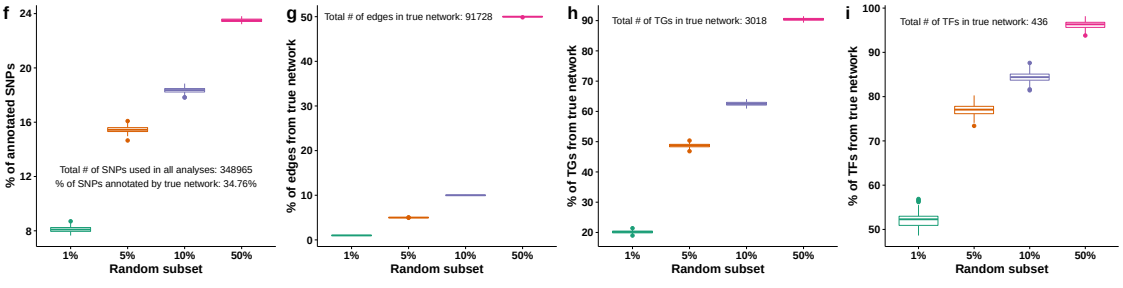
Panels **a-c** Comparison of network enrichment results between using the real target network and using its noisy versions in RSS-NET.



Panels **d-e** Comparison of gene-level association results between using the real target network and using its noisy versions in RSS-NET. Note that the AUROC based on true network shown below is slightly different from the AUROC shown in **Supplementary Figure 5**, although they are generated from the same simulated datasets. This is because AUROC shown in **Supplementary Figure 5** is evaluated based on 16,954 autosomal protein-coding genes that are available in Pascal-associated database, and AUROC shown below is evaluated based on all available autosomal protein-coding genes.



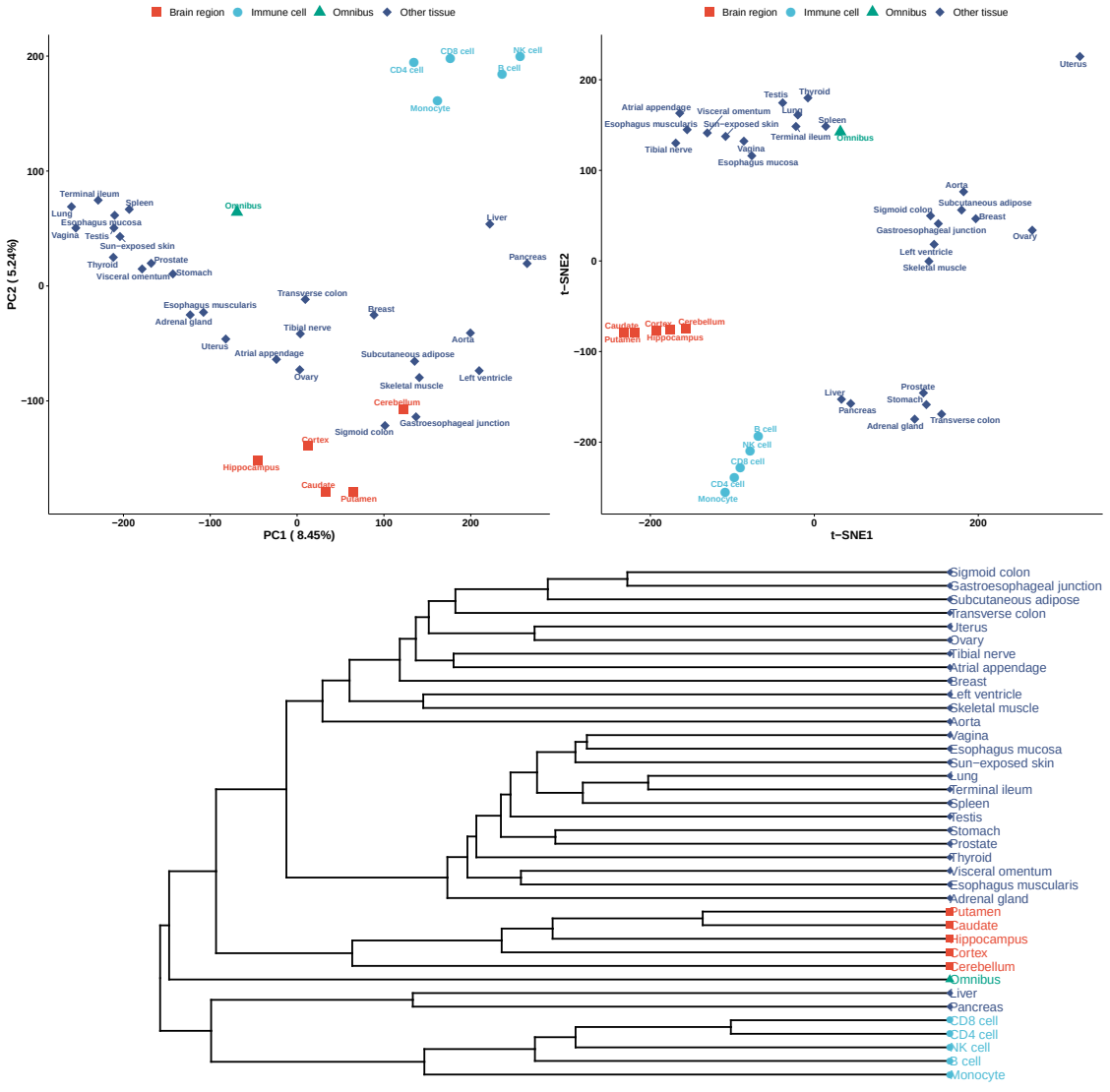
Panels **f-i** Descriptive statistics of noisy networks. Each box plot below is based on 200 noisy networks.



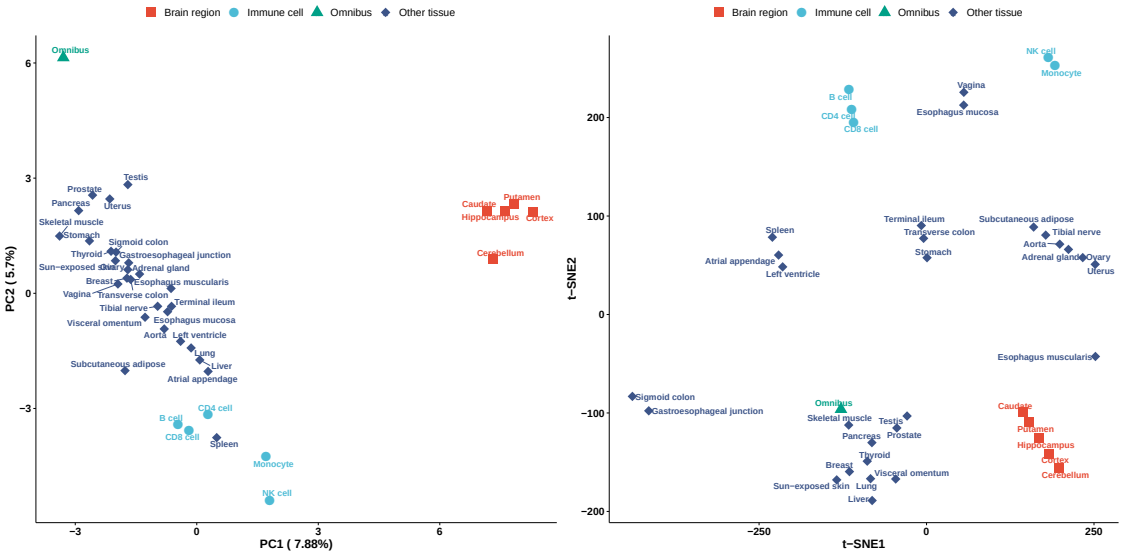
Supplementary Figure 7

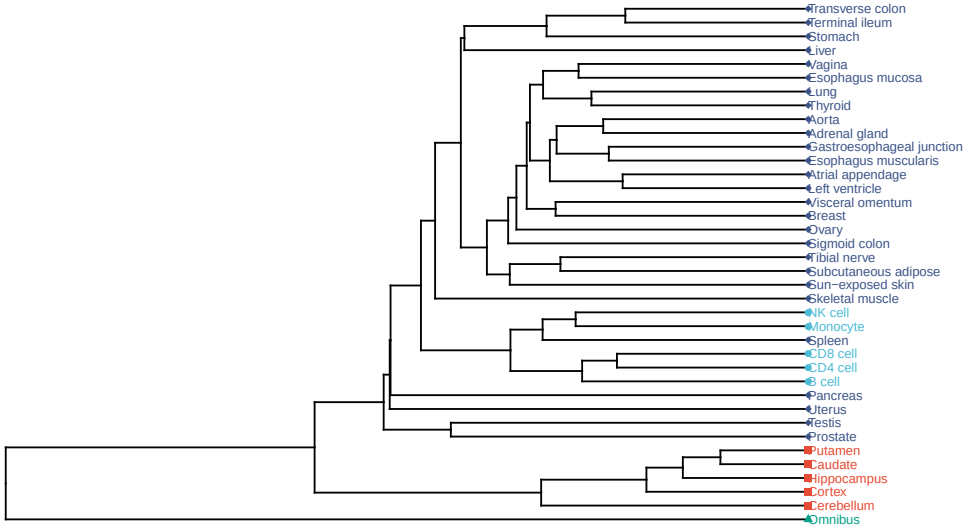
Clustering analysis of 38 regulatory networks based on principal component analysis (PCA), *t*-distributed stochastic neighbor embedding (t-SNE), and hierarchical clustering. To perform PCA, we used the R built-in function `prcomp`. To perform t-SNE, we used the R package `Rtsne`. To perform hierarchical clustering, we used the R built-in function `hclust`, with distance = 1 – Pearson correlation, and average method. **Figure 5(a)** corresponds to the t-SNE plot in Panel **e**.

Panel **a** Here the input data matrix has 1,289,786 rows (SNPs) and 38 columns (networks), where the (i, j)-th entry is 1 if SNP i is inside network j (i.e. SNP j is within 100 kb of any element in network j), and it is 0 otherwise.

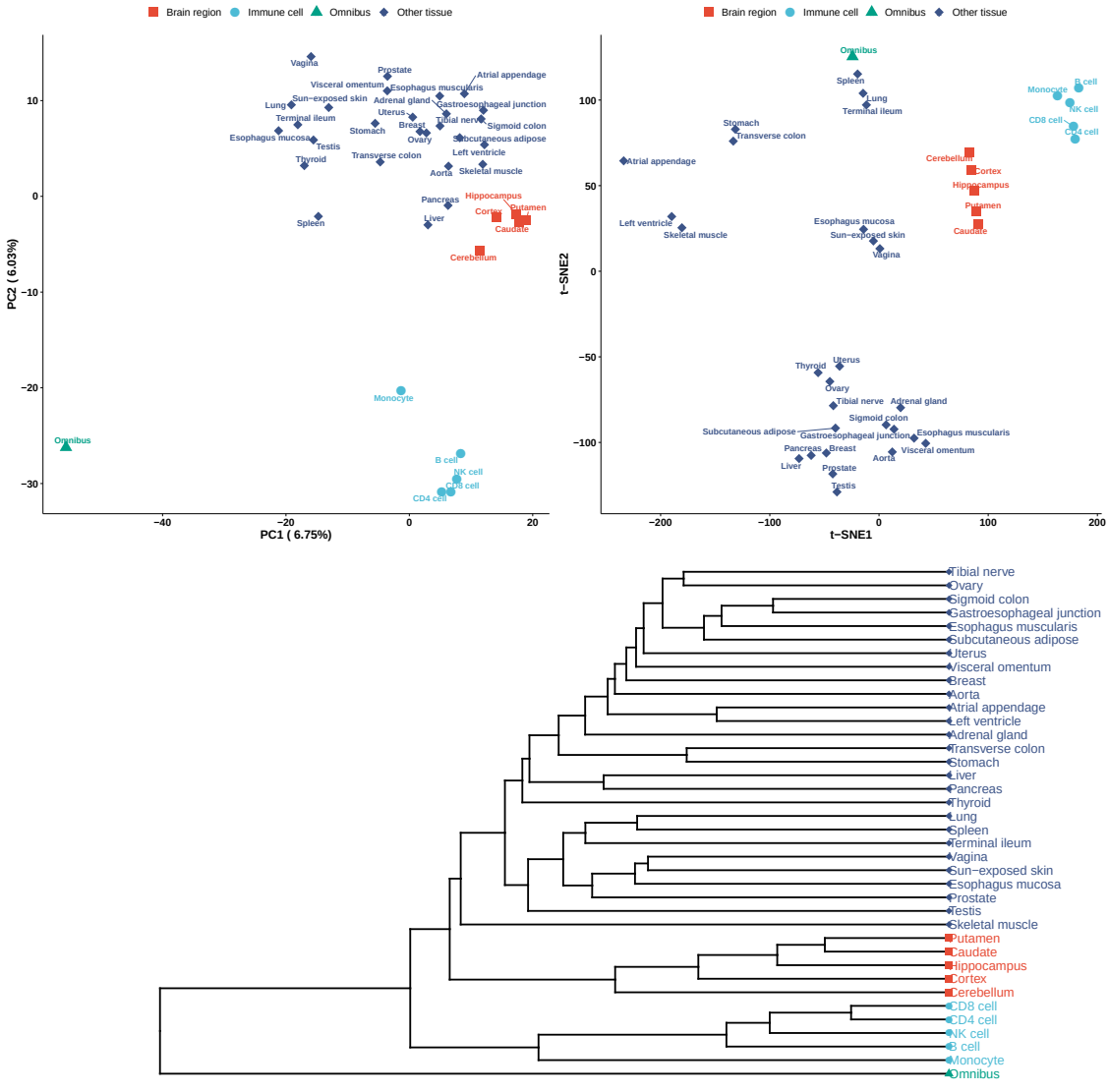


Panel **b** Here the input data matrix has 605 rows (TFs) and 38 columns (networks), where the (i, j)-th entry is 1 if TF i belongs to network j, and it is 0 otherwise.

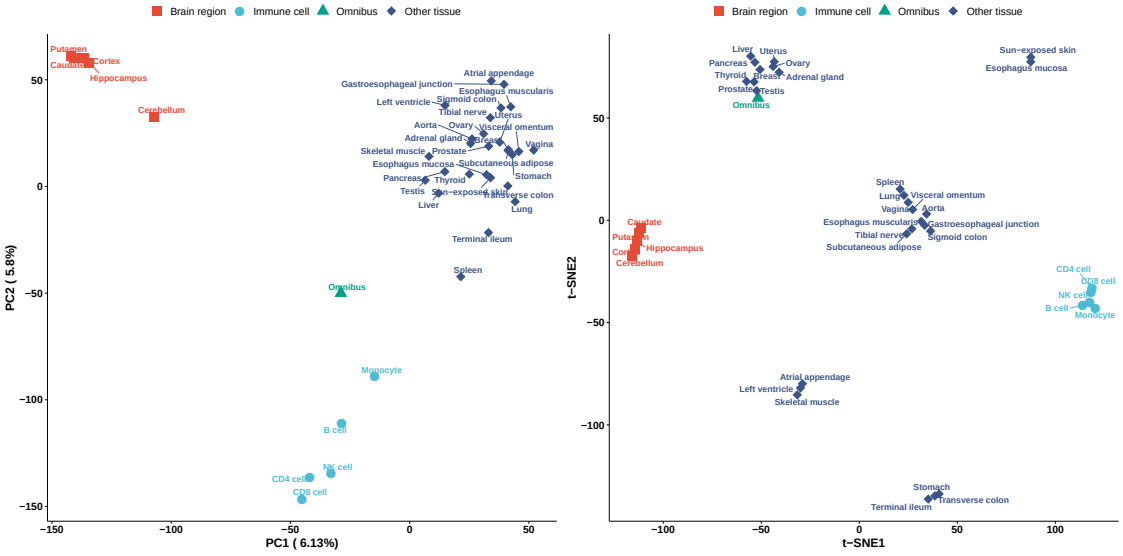


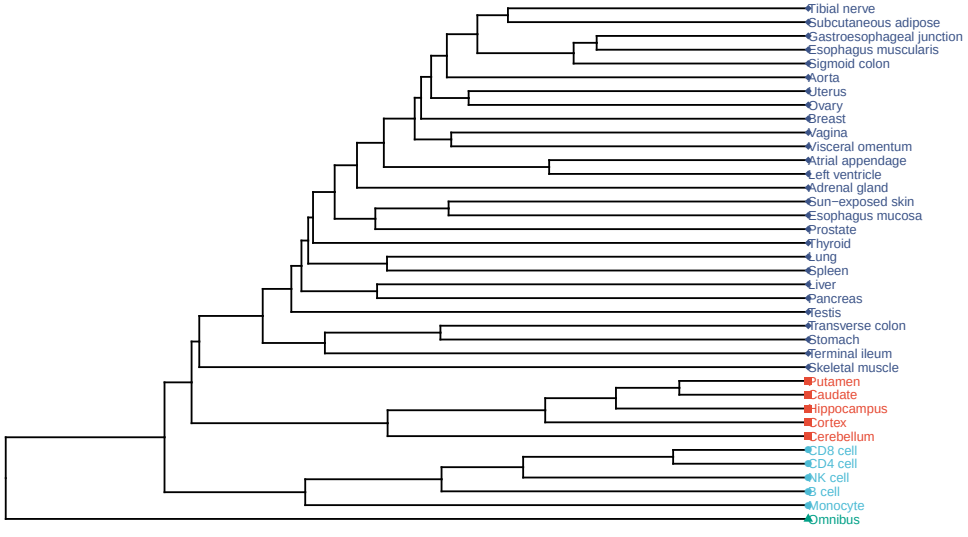


Panel **c** Here the input data matrix has 12,963 rows (TGs) and 38 columns (networks), where the (i, j)-th entry is 1 if TG i belongs to network j, and it is 0 otherwise.

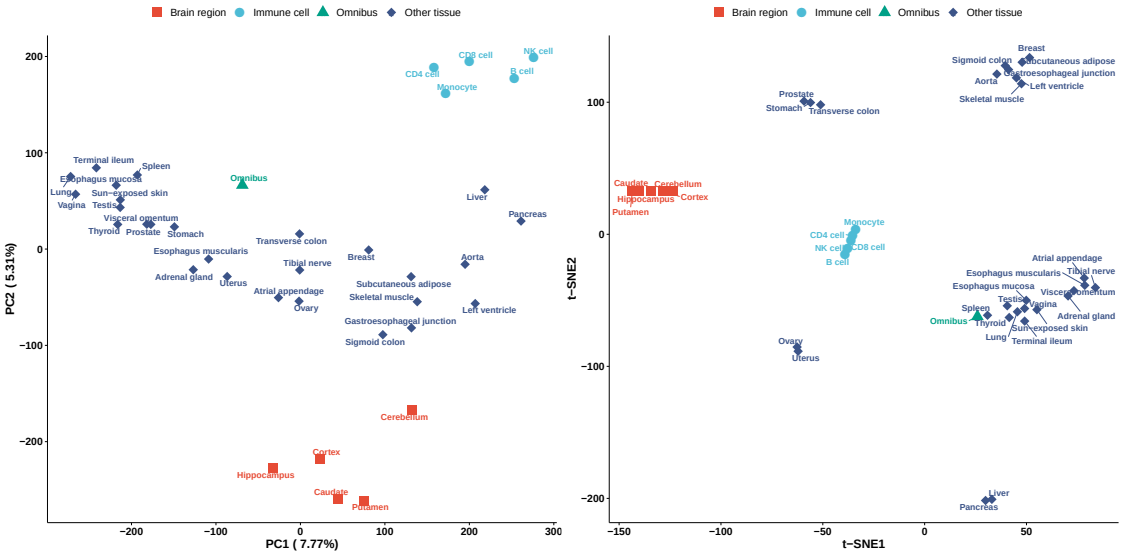


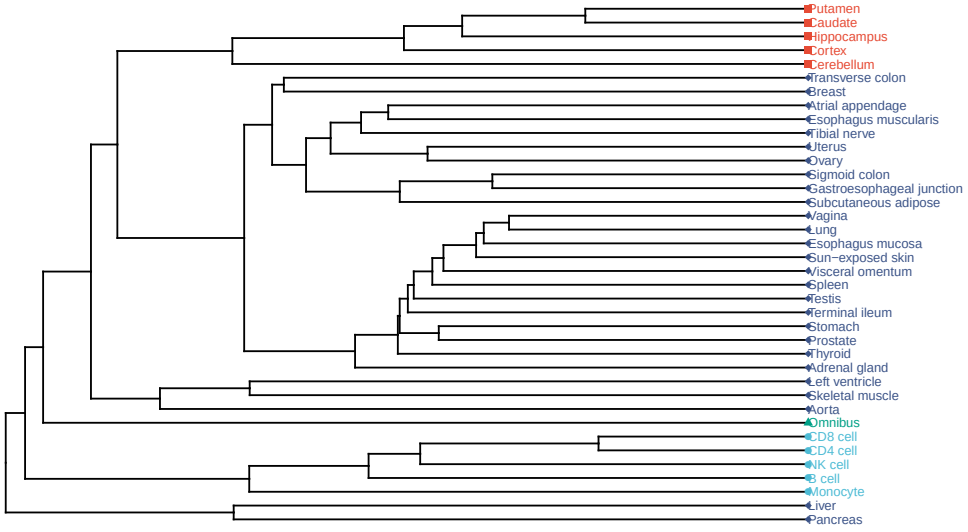
Panel **d** Here the input data matrix has 880,777 rows (TF-TG edges) and 38 columns (networks), where the (i,j)-th entry is the weight of TF-TG edge i in network j, and it is 0 if TF-TG edge i is not available in network j.





Panel **e** Here the input data matrix is obtained by concatenating the previous 4 data matrices vertically, with a total of 2,184,131 rows and 38 columns (networks).

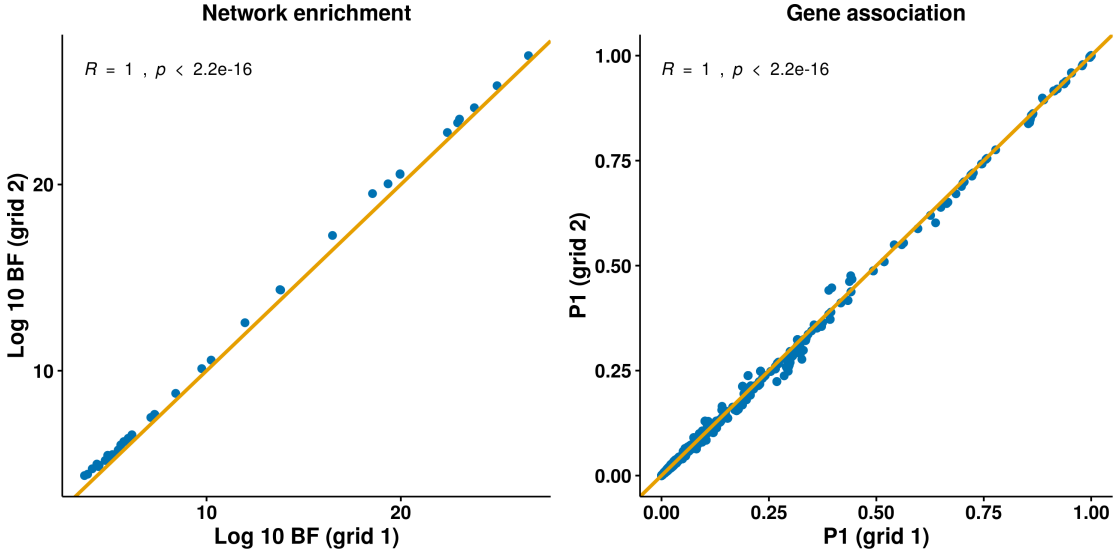




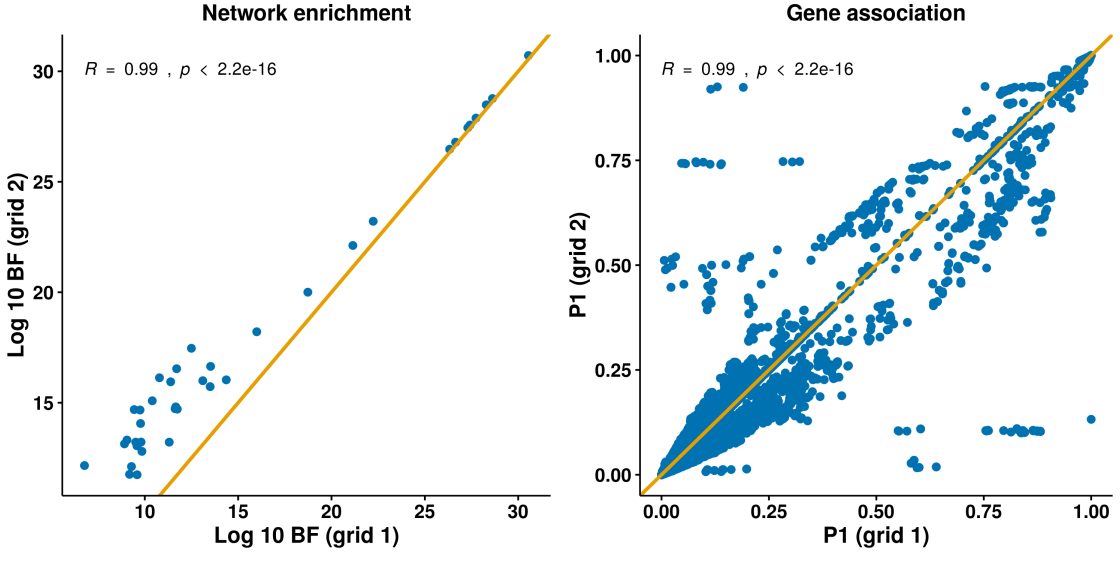
Supplementary Figure 8

Robustness of RSS-NET results to hyper-parameter grid choice. For three GWAS datasets, we ran RSS-NET for 38 networks on two different hyper-parameter grids, and then compared results of network enrichments (BF, left plot) and gene associations (P_1^{net}, right plot). In the network enrichment comparison plot, each dot represents a network. In the gene association comparison plot, each dot represent a gene-network pair. All diagonal lines have slope 1 and intercept 0.

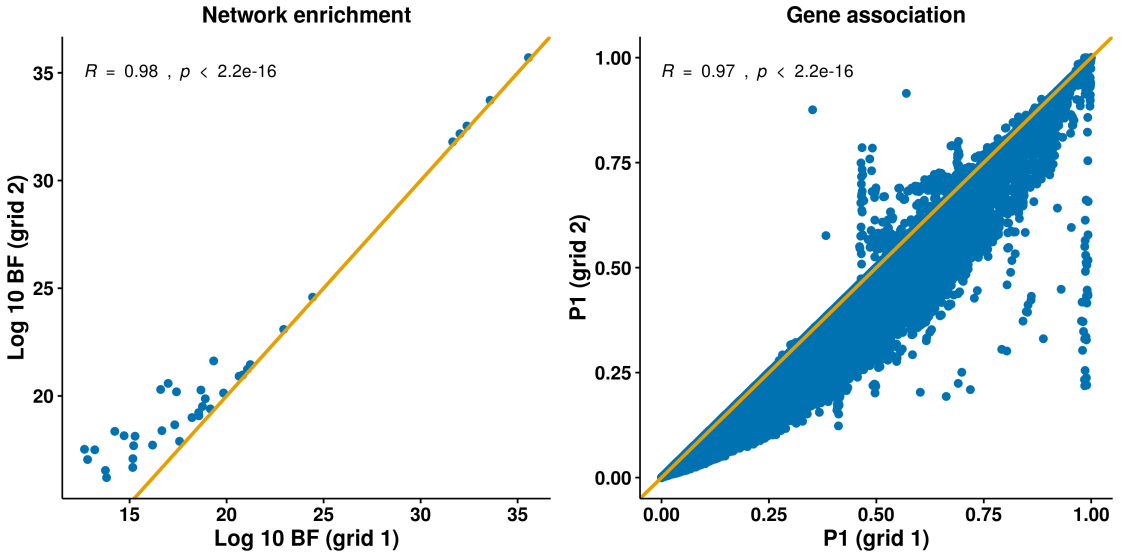
Panel **a** Alzheimer's disease (Lambert et al. 2013). In grid 1, $\theta \in (0 : 0.5 : 4)$. In grid 2, $\theta \in (0 : 0.25 : 1)$. For both grids, $\eta = 0.6$, $\rho \in (0 : 0.2 : 0.8)$ and $\theta_0 \in (-5.25 : 0.1 : -4.75)$.



Panel **b** low-density lipoprotein (Teslovich et al. 2010). In grid 1, $\theta \in (0 : 0.5 : 3)$. In grid 2, $\theta \in (0 : 0.25 : 1)$. For both grids, $\eta = 0.3$, $\rho \in (0 : 0.2 : 0.8)$ and $\theta_0 \in (-3.75 : 0.05 : -3.5)$.



Panel **c** inflammatory bowel disease (Liu et al. 2015). In grid 1, $\theta \in (0 : 0.5 : 3)$. In grid 2, $\theta \in (0 : 0.25 : 1)$. For both grids, $\eta = 0.3$, $\rho \in (0 : 0.2 : 0.8)$ and $\theta_0 \in (-3 : 0.05 : -2.8)$.



Supplementary Tables

Supplementary Table 1

Information of paired high-throughput sequencing data of gene expression and chromatin accessibility that are used to construct 38 regulatory networks. Please see the attached spreadsheet `rss_net_supp_tab_1.xlsx`.

Supplementary Table 2

GWAS publications, total sample sizes and numbers of genetic variants analyzed by RSS-NET for 18 human traits. Click blue links to view publications online.

Trait (abbreviation)	PMID	# of SNPs (analyzed)	Sample size (cases+controls)
Alzheimer's disease (LOAD)	24162737	1,136,997	17,008+37,154
Neuroticism (NEU)	27089181	1,119,108	170,911
Schizophrenia (SCZ)	25056061	1,113,442	152,805
Body mass index (BMI)	25673413	1,012,465	234,069
Height (HEIGHT)	25282103	1,064,575	253,288
Waist-to-hip ratio (WAIST)	25673412	1,008,898	142,762
Crohn's disease (CD)	26192919	1,064,533	5,956+14,927
Inflammatory bowel disease (IBD)	26192919	1,081,481	12,882+21,770
Rheumatoid arthritis (RA)	24390342	1,158,064	14,361+43,923
Ulcerative colitis (UC)	26192919	1,092,170	6,968+20,464
Breast cancer (BC)	23535729	1,188,902	15,863+40,022
Atrial fibrillation (AF)	28416818	1,175,059	17,931+115,142
Coronary artery disease (CAD)	26343387	1,121,322	60,801+123,504
High-density lipoprotein (HDL)	20686565	1,032,214	99,900
Heart rate (HR)	23583979	1,066,168	92,355
Low-density lipoprotein (LDL)	20686565	1,030,397	95,454
Myocardial infarction (MI)	26343387	1,111,568	42,561+123,504
Type 2 diabetes (T2D)	22885922	1,047,618	12,171+56,862

Supplementary Table 3

To confirm that RSS-NET network enrichment results are unlikely to be driven by generic regulatory enrichments harbored in the vicinity of genes, we perform a "near-gene" control analysis. Specifically, we create a "near-gene" control network with 18,334 protein-coding autosomal genes as nodes and no edge, and then analyze this control with RSS-NET on the same GWAS data for 18 traits. The near-gene control analysis is formalized as:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &\sim \mathcal{N}(\widehat{\mathbf{S}}\widehat{\mathbf{R}}\widehat{\mathbf{S}}^{-1}\boldsymbol{\beta},\ \widehat{\mathbf{S}}\widehat{\mathbf{R}}\widehat{\mathbf{S}}),\\
\beta_j &\sim \pi_j \cdot \mathcal{N}(\mu_j,\ \sigma_0^2) + (1-\pi_j)\cdot\delta_0,\\
\pi_j &= (1+10^{-(\theta_0+a_j\theta)})^{-1},\\
\mu_j &= \textstyle\sum_{g\in G_j} c_{jg}\cdot\gamma_{jg},\\
\gamma_{jg} &\sim \mathcal{N}(0,\ \sigma^2),
\end{aligned}$$

where $a_j = 1$ if SNP j is within 100 kb of any protein-coding autosomal gene and $a_j = 0$ otherwise, G_j denotes the set of all protein-coding genes within 1 Mb window of SNP j, and c_{jg} measures the relative impact of SNP j on gene g (which is derived from eQTLGen cis eQTL results; see **Supplementary Notes**).

Using the same GWAS data of each trait and the same hyper-parameter grid, we compare BFs between each network and the near-gene control under four enrichment models: (1) M_{11}: $\theta > 0$ and $\sigma^2 = 0$; (2) M_{12}: $\theta = 0$ and $\sigma^2 > 0$; (3) M_{13}: $\theta > 0$ and $\sigma^2 > 0$; (4) M_1: $\theta > 0$ or $\sigma^2 > 0$. For all 38 networks and the near-gene control, BF computations are based on the same baseline model (M_0: $\theta = 0$ and $\sigma^2 = 0$). Each row of the following table reports the number of networks with BFs greater than near-gene control BF for each trait and each enrichment model. Trait abbreviations are defined in **Supplementary Table 2**.

	# of networks			
Trait	M_{11}	M_{12}	M_{13}	M_1
AF	32	38	38	38
BC	33	38	38	38
BMI	9	38	38	38
CAD	38	38	38	38
HDL	38	38	38	38
HR	32	38	38	38
LDL	38	38	38	38
MI	38	38	38	38
NEU	33	38	38	38
RA	38	38	38	38
SCZ	9	38	38	38
WAIST	38	38	38	38
UC	38	22	34	34
LOAD	38	13	11	11
IBD	38	0	6	6
T2D	2	0	21	5
CD	38	0	5	0
HEIGHT	38	0	0	0

Supplementary Table 4

We compute Pearson correlations between five network features and log 10 enrichment BFs, either across 512 trait-network pairs that pass the near-gene control (**Supplementary Table 3**), or across all 684 trait-network pairs (18 traits and 38 networks). We use R built-in function `cor.test` for all correlation analyses. All tests are two-sided.

	Pearson r (p-value)	
Network feature	512 pairs	684 pairs
% of SNPs in a network	-0.0304 (0.4925)	-0.0403 (0.2924)
# of TF-TG edges	-0.0092 (0.8353)	0.0025 (0.9479)
# of nodes (TF or TG)	-0.0535 (0.2272)	-0.0422 (0.2702)
# of TGs	-0.0531 (0.2308)	-0.0424 (0.2686)
# of TFs	-0.0403 (0.3625)	-0.0184 (0.6312)

Supplementary Table 5

We compute Pearson correlations between 73 binnary functional annotations in LDSC baselineLD v2.1 (Gazal et al. 2017) and log 10 enrichment BFs, either across 512 trait-network pairs that pass the near-gene control (**Supplementary Table 3**), or across all 684 trait-network pairs (18 traits and 38 networks). Specifically, for a given functional annotation, we estimate the correlation between log 10 BFs and proportion of SNPs falling into both a network and this functional category, across all trait-network pairs. Rows are ranked by Pearson p-values based on 512 trait-network pairs. The Bonferroni cutoff is $0.05/73 = 6.8 \times 10^{-4}$. The rest is the same as **Supplementary Table 4**.

	Pearson r (- log 10 p-value)	
Functional annotation	512 pairs	684 pairs
BivFlnk	-0.1295 (2.4786)	-0.1252 (2.9849)
TSS_Hoffman	-0.1256 (2.3543)	-0.1547 (4.3146)
BivFlnk_500	-0.1195 (2.1668)	-0.1166 (2.6468)
TSS_Hoffman_500	-0.1178 (2.1193)	-0.1465 (3.9183)
H3K9ac_peaks_Trynka	-0.0795 (1.1410)	-0.1041 (2.1907)
Promoter_UCSC_500	-0.0777 (1.1022)	-0.1137 (2.5363)
Promoter_UCSC	-0.0772 (1.0926)	-0.1134 (2.5254)
PromoterFlanking_Hoffman_500	-0.0742 (1.0296)	-0.1125 (2.4941)
Enhancer_Hoffman	-0.0707 (0.9579)	-0.0907 (1.7518)
H3K9ac_Trynka	-0.0704 (0.9521)	-0.0939 (1.8543)
PromoterFlanking_Hoffman	-0.0696 (0.9357)	-0.1144 (2.5616)
H3K4me3_peaks_Trynka	-0.0681 (0.9063)	-0.0941 (1.8597)
Enhancer_Hoffman_500	-0.0671 (0.8879)	-0.0891 (1.7044)
H3K4me3_Trynka	-0.0647 (0.8428)	-0.0903 (1.7393)
WeakEnhancer_Hoffman	-0.0641 (0.8313)	-0.0798 (1.4331)
H3K9ac_Trynka_500	-0.0606 (0.7666)	-0.0852 (1.5871)
SuperEnhancer_Hnisz_500	-0.0596 (0.7490)	-0.0806 (1.4540)
SuperEnhancer_Hnisz	-0.0595 (0.7481)	-0.0804 (1.4496)
Enhancer_Andersson	-0.0590 (0.7386)	-0.0747 (1.2932)
WeakEnhancer_Hoffman_500	-0.0564 (0.6929)	-0.0750 (1.3024)
UTR_5_UCSC_500	-0.0562 (0.6897)	-0.0933 (1.8339)
Enhancer_Andersson_500	-0.0557 (0.6807)	-0.0701 (1.1735)
H3K4me3_Trynka_500	-0.0536 (0.6465)	-0.0796 (1.4284)
TFBS_ENCODE	-0.0522 (0.6230)	-0.0767 (1.3473)
UTR_5_UCSC	-0.0497 (0.5823)	-0.0787 (1.4009)
UTR_3_UCSC_500	-0.0480 (0.5559)	-0.0813 (1.4738)
MAFbin1	-0.0474 (0.5461)	-0.0508 (0.7331)
H3K4me1_peaks_Trynka	-0.0460 (0.5249)	-0.0704 (1.1819)
H3K27ac_PGC2	-0.0453 (0.5144)	-0.0712 (1.2030)
CTCF_Hoffman_500	-0.0447 (0.5043)	-0.0655 (1.0606)
DGF_ENCODE	-0.0432 (0.4828)	-0.0676 (1.1124)
H3K27ac_PGC2_500	-0.0429 (0.4775)	-0.0692 (1.1512)

(continued)

Functional annotation	Pearson r (- log 10 p-value) 512 pairs	684 pairs
Coding_UCSC_500	-0.0424 (0.4710)	-0.0753 (1.3093)
CTCF_Hoffman	-0.0421 (0.4667)	-0.0618 (0.9732)
MAFbin2	-0.0420 (0.4642)	-0.0530 (0.7797)
H3K27ac_Hnisz	-0.0412 (0.4531)	-0.0665 (1.0860)
TFBS_ENCODE_500	-0.0408 (0.4474)	-0.0659 (1.0700)
H3K27ac_Hnisz_500	-0.0402 (0.4386)	-0.0658 (1.0682)
H3K4me1_Trynka	-0.0377 (0.4040)	-0.0643 (1.0314)
FetalDHS_Trynka	-0.0375 (0.4013)	-0.0615 (0.9672)
DHS_peaks_Trynka	-0.0365 (0.3869)	-0.0610 (0.9557)
Transcr_Hoffman	-0.0359 (0.3794)	-0.0728 (1.2426)
MAFbin3	-0.0353 (0.3710)	-0.0557 (0.8365)
DGF_ENCODE_500	-0.0350 (0.3675)	-0.0609 (0.9532)
FetalDHS_Trynka_500	-0.0347 (0.3635)	-0.0589 (0.9067)
DHS_Trynka	-0.0347 (0.3629)	-0.0594 (0.9191)
H3K4me1_Trynka_500	-0.0336 (0.3491)	-0.0606 (0.9464)
Intron_UCSC_500	-0.0330 (0.3412)	-0.0661 (1.0752)
Intron_UCSC	-0.0330 (0.3408)	-0.0662 (1.0768)
non_synonymous	-0.0323 (0.3313)	-0.0639 (1.0221)
DHS_Trynka_500	-0.0318 (0.3259)	-0.0575 (0.8757)
MAFbin4	-0.0310 (0.3156)	-0.0557 (0.8374)
Vertebrate_phastCons46way_500	-0.0309 (0.3136)	-0.0583 (0.8935)
Mammal_phastCons46way_500	-0.0305 (0.3084)	-0.0577 (0.8810)
Transcr_Hoffman_500	-0.0297 (0.2992)	-0.0602 (0.9356)
LindbladToh_500	-0.0296 (0.2972)	-0.0554 (0.8312)
Primate_phastCons46way_500	-0.0293 (0.2933)	-0.0565 (0.8541)
Vertebrate_phastCons46way	-0.0285 (0.2838)	-0.0529 (0.7776)
MAFbin5	-0.0268 (0.2635)	-0.0561 (0.8447)
Mammal_phastCons46way	-0.0267 (0.2618)	-0.0511 (0.7392)
MAFbin6	-0.0263 (0.2581)	-0.0576 (0.8780)
LindbladToh	-0.0248 (0.2402)	-0.0482 (0.6828)
MAFbin9	-0.0244 (0.2348)	-0.0580 (0.8876)
Primate_phastCons46way	-0.0244 (0.2348)	-0.0498 (0.7131)
MAFbin8	-0.0241 (0.2314)	-0.0580 (0.8862)
MAFbin7	-0.0241 (0.2313)	-0.0558 (0.8396)
MAFbin10	-0.0237 (0.2267)	-0.0567 (0.8579)
Coding_UCSC	-0.0236 (0.2266)	-0.0542 (0.8045)
GERP.RSsup4	-0.0227 (0.2163)	-0.0461 (0.6402)
UTR_3_UCSC	-0.0220 (0.2079)	-0.0531 (0.7807)
Repressed_Hoffman_500	-0.0176 (0.1599)	-0.0429 (0.5803)
Repressed_Hoffman	-0.0158 (0.1418)	-0.0411 (0.5472)
synonymous	-0.0135 (0.1191)	-0.0425 (0.5734)

Supplementary Table 6

For 512 network-trait pairs passing the near-gene enrichment control (**Supplementary Table 3**), we compute BFs comparing the baseline model (M_0: $\theta = 0$ and $\sigma^2 = 0$) against three disjoint enrichment models: (1) M_{11}: $\theta > 0$ and $\sigma^2 = 0$; (2) M_{12}: $\theta = 0$ and $\sigma^2 > 0$; (3) M_{13}: $\theta > 0$ and $\sigma^2 > 0$. For each network-trait pair, we compare BFs based on M_{11}, M_{12} and M_{13}, and define the "best" enrichment model as the one with the largest BF. Each row below shows, for each trait, the number of networks with M_{11}, M_{12} and M_{13} being the "best" model, respectively. Trait abbreviations are defined in **Supplementary Table 2**. **Figure 5(b)** is based on the table below.

	# of networks			
Trait	M_{11}	M_{12}	M_{13}	Total
AF	0	0	38	38
BC	0	10	28	38
BMI	0	14	24	38
CAD	0	0	38	38
HDL	0	0	38	38
HR	0	0	38	38
LDL	0	0	38	38
MI	0	0	38	38
NEU	0	0	38	38
RA	0	0	38	38
SCZ	0	38	0	38
WAIST	0	37	1	38
UC	0	0	34	34
LOAD	0	0	11	11
IBD	0	0	6	6
T2D	2	0	3	5
Total	2	99	411	512

Supplementary Table 7

Correlations of RSS-NET enrichment log 10 BFs between 18 traits. Correlation are computed over all 38 networks. Rows are ranked by Pearson p-values. The Bonferroni cutoff is $0.05/153 = 3.3 \times 10^{-4}$. Trait abbreviations are defined in **Supplementary Table 2**. **Figure 5(c)** is based on the table below.

Trait 1	Trait 2	Pearson r	- Log 10 p-value
CD	IBD	0.9551	19.8723
IBD	UC	0.9204	15.5240
LDL	LOAD	0.8969	13.5933
CAD	MI	0.8954	13.4826
CD	UC	0.8342	10.1112
CD	LDL	0.7969	8.6674
RA	UC	0.7853	8.2783
CD	LOAD	0.7791	8.0795
RA	SCZ	0.7676	7.7266
IBD	LOAD	0.7548	7.3575
CAD	HR	0.7526	7.2965
NEU	SCZ	0.7335	6.7908
BMI	SCZ	0.7289	6.6741
IBD	LDL	0.7237	6.5469
BMI	MI	0.7131	6.2964
IBD	RA	0.7014	6.0293
BMI	CAD	0.6858	5.6952
AF	BMI	0.6827	5.6303
HR	MI	0.6743	5.4617
BMI	RA	0.6671	5.3205
CD	RA	0.6562	5.1122
AF	CAD	0.6520	5.0354
AF	MI	0.6491	4.9826
CAD	HEIGHT	0.6390	4.8016
HEIGHT	MI	0.6269	4.5958
HDL	LDL	0.6244	4.5539
NEU	RA	0.6050	4.2420
LOAD	UC	0.5880	3.9847
LDL	UC	0.5587	3.5744
AF	HEIGHT	0.5499	3.4575
AF	RA	0.5291	3.1956
LDL	RA	0.5131	3.0062
BMI	HR	0.4912	2.7597
CAD	SCZ	0.4684	2.5202
BMI	NEU	0.4586	2.4222
HDL	MI	0.4563	2.3999
MI	SCZ	0.4541	2.3784

(continued)

Trait 1	Trait 2	Pearson r	- Log 10 p-value
AF	SCZ	0.4516	2.3544
SCZ	UC	0.4371	2.2164
BC	IBD	-0.4332	2.1810
AF	HR	0.4170	2.0364
BMI	LDL	0.4158	2.0256
LDL	MI	0.4149	2.0181
NEU	UC	0.4034	1.9198
BMI	HDL	0.4029	1.9161
HEIGHT	HR	0.4012	1.9018
HDL	LOAD	0.4008	1.8985
CAD	HDL	0.3992	1.8848
LOAD	RA	0.3956	1.8554
CD	BC	-0.3949	1.8495
BMI	WAIST	0.3773	1.7091
BMI	UC	0.3739	1.6832
CD	HDL	0.3616	1.5904
BC	SCZ	0.3603	1.5807
BMI	HEIGHT	0.3542	1.5359
HR	WAIST	0.3500	1.5051
MI	RA	0.3476	1.4883
HR	SCZ	0.3396	1.4321
SCZ	WAIST	0.3370	1.4140
BMI	IBD	0.3311	1.3735
CD	MI	0.3303	1.3683
AF	UC	0.3275	1.3496
MI	WAIST	0.3259	1.3389
AF	IBD	0.3241	1.3270
BMI	CD	0.3204	1.3025
IBD	MI	0.3115	1.2443
BC	LOAD	-0.3000	1.1721
NEU	WAIST	0.2963	1.1493
BC	HR	0.2956	1.1455
HDL	IBD	0.2956	1.1450
HDL	HEIGHT	0.2921	1.1240
BC	UC	-0.2857	1.0859
CAD	WAIST	0.2847	1.0801
IBD	SCZ	0.2804	1.0545
CAD	T2D	-0.2705	0.9982
HEIGHT	LDL	0.2692	0.9905
HDL	RA	0.2680	0.9841
IBD	NEU	0.2674	0.9804
MI	UC	0.2648	0.9661

(continued)

Trait 1	Trait 2	Pearson r	- Log 10 p-value
HR	LOAD	-0.2619	0.9500
AF	NEU	0.2586	0.9320
HR	UC	-0.2579	0.9284
HR	IBD	-0.2560	0.9178
AF	CD	0.2526	0.8997
AF	LDL	0.2501	0.8865
CD	NEU	0.2468	0.8687
AF	HDL	0.2466	0.8680
CD	SCZ	0.2461	0.8653
RA	WAIST	0.2434	0.8511
CAD	BC	0.2379	0.8228
CAD	NEU	0.2361	0.8139
CAD	RA	0.2355	0.8107
MI	T2D	-0.2351	0.8089
BC	NEU	0.2308	0.7868
BMI	BC	0.2249	0.7578
HEIGHT	NEU	-0.2242	0.7544
BC	LDL	-0.2177	0.7230
MI	NEU	0.2138	0.7046
LOAD	MI	0.2129	0.7003
CD	HR	-0.2046	0.6620
AF	WAIST	0.2024	0.6517
LDL	SCZ	0.2008	0.6445
HEIGHT	LOAD	0.1980	0.6317
HDL	SCZ	0.1973	0.6289
CD	WAIST	0.1951	0.6190
BC	HDL	-0.1847	0.5735
HDL	HR	0.1833	0.5674
HR	T2D	-0.1819	0.5615
HDL	T2D	-0.1722	0.5211
CAD	LDL	0.1717	0.5191
LDL	WAIST	0.1622	0.4805
LDL	T2D	-0.1611	0.4764
BC	WAIST	0.1553	0.4536
LOAD	T2D	-0.1488	0.4289
UC	WAIST	0.1412	0.4003
HEIGHT	WAIST	-0.1389	0.3919
HDL	UC	0.1318	0.3662
SCZ	T2D	-0.1309	0.3631
AF	LOAD	0.1305	0.3618
HEIGHT	T2D	-0.1298	0.3591
HR	NEU	0.1289	0.3560

(continued)

Trait 1	Trait 2	Pearson r	- Log 10 p-value
IBD	WAIST	0.1254	0.3438
HEIGHT	IBD	0.1239	0.3386
BMI	LOAD	0.1200	0.3251
AF	BC	0.1062	0.2793
BC	MI	0.0984	0.2543
HEIGHT	RA	0.0973	0.2509
CD	HEIGHT	0.0949	0.2434
HEIGHT	SCZ	0.0869	0.2190
BMI	T2D	-0.0833	0.2082
T2D	WAIST	0.0793	0.1965
CAD	LOAD	-0.0752	0.1846
AF	T2D	0.0708	0.1722
BC	T2D	-0.0689	0.1669
NEU	T2D	-0.0648	0.1554
HEIGHT	UC	0.0626	0.1495
BC	RA	-0.0485	0.1121
HR	LDL	-0.0477	0.1101
HR	RA	-0.0457	0.1051
HDL	NEU	0.0456	0.1048
T2D	UC	0.0447	0.1024
RA	T2D	0.0391	0.0885
LOAD	NEU	-0.0390	0.0881
IBD	T2D	0.0251	0.0549
HDL	WAIST	-0.0236	0.0515
LOAD	SCZ	0.0212	0.0460
LDL	NEU	0.0210	0.0455
CD	T2D	-0.0159	0.0341
CAD	CD	0.0112	0.0238
BC	HEIGHT	0.0099	0.0209
CAD	IBD	0.0055	0.0116
CAD	UC	0.0050	0.0103
LOAD	WAIST	0.0001	0.0002

Supplementary Table 8

For each trait or trait-network pair, we compute the proportion of genes with higher enrichment P_1 estimates (P_1^{bma} or P_1^{net}) than reference estimates (P_1^{base} or P_1^{near}), among genes with reference P_1 estimates higher than a given cutoff.

Panel **a** Median proportion of genes with P_1^{bma} higher than reference estimates (P_1^{base} or P_1^{near}), among genes with reference estimates higher than a given cutoff. Medians shown below are across 16 traits that have multiple networks more enriched than the near-gene enrichment control (**Supplementary Table 3**). **Figure 5(d)** is based on this panel.

	All genes		Only TFs		Only TGs	
Cutoff	P_1^{base}	P_1^{near}	P_1^{base}	P_1^{near}	P_1^{base}	P_1^{near}
0	0.9828	0.9401	0.9556	0.8829	0.9842	0.9423
0.1	0.9531	0.8811	0.9837	0.8593	0.9499	0.8920
0.2	0.9180	0.8590	0.9514	0.8333	0.9166	0.8707
0.3	0.8856	0.8342	0.9564	0.8000	0.8898	0.8360
0.4	0.8754	0.8061	0.9667	0.7500	0.8784	0.8084
0.5	0.8695	0.7871	0.9444	0.7321	0.8771	0.7979
0.6	0.8499	0.7815	0.9348	0.7321	0.8607	0.7872
0.7	0.8647	0.7795	0.9583	0.7321	0.8624	0.7831
0.8	0.8399	0.7578	1.0000	0.6667	0.8521	0.7657
0.9	0.8167	0.7443	1.0000	0.6870	0.8259	0.7245

Panel **b** Median proportion of genes with P_1^{net} higher than reference estimates (P_1^{base} or P_1^{near}), among genes with reference estimates higher than a given cutoff. Medians shown below are across 512 network-trait pairs passing the near-gene enrichment control (**Supplementary Table 3**).

	All genes		Only TFs		Only TGs	
Cutoff	P_1^{base}	P_1^{near}	P_1^{base}	P_1^{near}	P_1^{base}	P_1^{near}
0	0.9637	0.9046	0.8926	0.7058	0.9730	0.9325
0.1	0.9203	0.8613	0.8549	0.6667	0.9340	0.8898
0.2	0.9081	0.8494	0.8571	0.6667	0.9114	0.8696
0.3	0.8623	0.8276	0.8313	0.6667	0.8745	0.8537
0.4	0.8622	0.7971	0.8000	0.6000	0.8750	0.8225
0.5	0.8756	0.7891	0.7656	0.6079	0.8827	0.8053
0.6	0.8707	0.7895	0.7895	0.6154	0.8757	0.8000
0.7	0.8750	0.7906	0.7895	0.6206	0.8788	0.8077
0.8	0.8462	0.7679	0.7471	0.5862	0.8583	0.7863
0.9	0.8333	0.7500	0.7230	0.5000	0.8468	0.7619

Supplementary Table 9

For each trait we compute the correlation between RSS-NET enrichment statistic (log10 BF) and the percentage of network genes with P_1^{net} higher than the reference estimates (P_1^{base} or P_1^{near}), across all networks. Here we only consider 12 traits where all 38 networks show stronger enrichment than the near-gene control (**Supplementary Table 3**). Rows are ranked by Pearson p-values. The Bonferroni cutoff is $0.05/12 = 4.2 \times 10^{-3}$. Trait abbreviations are defined in **Supplementary Table 2**.

Panel **a** The reference P_1 is P_1^{base}.

	Pearson r (- log 10 p-value)		
Trait	All genes	Only TFs	Only TGs
LDL	0.8992 (13.7584)	0.9419 (17.9148)	0.8947 (13.4362)
HR	0.8166 (9.3903)	0.7301 (6.7046)	0.8090 (9.0992)
BMI	0.6810 (5.5959)	0.6628 (5.2378)	0.7271 (6.6306)
AF	0.6687 (5.3518)	0.7041 (6.0910)	0.6492 (4.9850)
MI	0.5851 (3.9420)	0.5452 (3.3972)	0.6125 (4.3594)
SCZ	0.5722 (3.7584)	0.6440 (4.8916)	0.6173 (4.4368)
BC	0.5579 (3.5637)	0.6122 (4.3543)	0.5412 (3.3468)
RA	0.3848 (1.7679)	0.3427 (1.4540)	0.4013 (1.9022)
WAIST	0.3672 (1.6316)	0.4298 (2.1496)	0.3946 (1.8468)
HDL	0.2629 (0.9555)	0.4662 (2.4981)	0.2658 (0.9714)
CAD	0.1948 (0.6175)	0.0261 (0.0574)	0.3409 (1.4413)
NEU	0.1461 (0.4184)	0.0523 (0.1219)	0.2634 (0.9581)

Panel **b** The reference P_1 is P_1^{near}. This panel is reported in the main text.

	Pearson r (- log 10 p-value)		
Trait	All genes	Only TFs	Only TGs
LDL	0.9100 (14.6053)	0.9432 (18.0902)	0.9086 (14.4901)
HR	0.8392 (10.3326)	0.8105 (9.1548)	0.8323 (10.0295)
BMI	0.7247 (6.5719)	0.7711 (7.8321)	0.7873 (8.3440)
MI	0.7079 (6.1755)	0.7797 (8.0986)	0.7372 (6.8854)
AF	0.7069 (6.1523)	0.7894 (8.4116)	0.6956 (5.9022)
NEU	0.5618 (3.6150)	0.6441 (4.8921)	0.6092 (4.3072)
CAD	0.5268 (3.1675)	0.5975 (4.1260)	0.6206 (4.4905)
SCZ	0.5218 (3.1084)	0.6433 (4.8791)	0.5812 (3.8855)
BC	0.5124 (2.9978)	0.5914 (4.0355)	0.4724 (2.5613)
RA	0.5110 (2.9809)	0.5191 (3.0761)	0.5098 (2.9673)
WAIST	0.2877 (1.0977)	0.4315 (2.1654)	0.3785 (1.7184)
HDL	0.2831 (1.0705)	0.4966 (2.8189)	0.3151 (1.2679)

Supplementary Table 10

Here we quantify overlaps between RSS-NET prioritized genes ($P_1^{\text{bma}} \geq 0.9$) and genes implicated in the GWAS Catalog (Buniello et al. 2019), for each of 16 traits that pass the near-gene enrichment control (**Supplementary Table 3**). The "Same GWAS" column reports the overlap between RSS-NET prioritized genes and genes implicated in the same GWAS. The "Later GWAS" column reports the overlap between RSS-NET prioritized genes and genes implicated in a later GWAS with increased sample size. The "No GWAS" column reports the number of RSS-NET prioritized genes that were not reported in either the same or the later GWAS. For example, there are 38 genome-wide significant genes in GWAS Catalog for the same GWAS of HDL (Teslovich et al. 2010) that are analyzed by RSS-NET, and 32 of them have $P_1^{\text{bma}} \geq 0.9$; there are 20 new genome-wide significant genes in GWAS Catalog for a later GWAS of HDL (Global Lipids Genetics Consortium 2013), and 3 of them have $P_1^{\text{bma}} \geq 0.9$ based on RSS-NET analysis of previous GWAS (Teslovich et al. 2010); there are 142 genes with $P_1^{\text{bma}} \geq 0.9$ that are not reported in either the same or the later GWAS of HDL. All p-values are generated by R built-in function `fisher.test`. Rows are sorted by p-values in the "Same GWAS" column, then p-values in the "Later GWAS" column. Trait abbreviations are defined in **Supplementary Table 2**.

It is important to note that some published GWAS hits of certain traits are not necessarily (genome-wide) significant in the corresponding summary data file. For example, rs34856868 shows $p = 9.80 \times 10^{-9}$ for association with IBD in Table 2 of Liu et al. (2015); however, the p value of the same SNP is 0.27 in the corresponding summary data file (`EUR.IBD.gwas.assoc.gz`). For this SNP, the result in Table 2 of Liu et al. (2015) was indeed obtained from a combined analysis of data on both GWAS and Immunochip arrays, whereas the result in the summary data file was only based on GWAS arrays. Because of this type of potential discrepancy between summary data files and corresponding publications, the overlap fractions shown in the "Same GWAS" column are not very high for certain traits.

	Fraction (- log 10 p-value)		
Trait	Same GWAS	Later GWAS	No GWAS
RA	44 / 61 (63.75)	4 / 7 (5.79)	190
HDL	32 / 38 (54.66)	3 / 20 (2.87)	142
LDL	26 / 27 (48.62)	1 / 18 (0.75)	141
SCZ	47 / 58 (45.93)	138 / 330 (82.51)	638
AF	16 / 18 (36.05)	10 / 83 (12.79)	35
CAD	18 / 20 (35.64)	21 / 140 (21.75)	72
IBD	33 / 80 (35.50)	4 / 12 (4.60)	101
WAIST	16 / 33 (29.73)	5 / 233 (3.11)	36
HR	10 / 15 (22.08)	2 / 43 (2.33)	29
BC	8 / 9 (21.03)	4 / 107 (4.82)	15
BMI	13 / 59 (19.61)	4 / 461 (1.39)	33
MI	8 / 9 (17.45)	NA	61
T2D	6 / 16 (12.26)	5 / 110 (6.29)	8
LOAD	6 / 18 (11.12)	11 / 29 (22.14)	18
UC	6 / 17 (7.92)	2 / 5 (3.02)	126
NEU	3 / 5 (6.58)	9 / 122 (10.40)	28

Supplementary Table 11

Examples of RSS-NET highlighted genes that were reported in GWAS of the same data. The "mouse trait" column is based on the Mouse Genome Informatics (Bult et al. 2019). The "therapeutic/clinical evidence" column is based on the Online Mendelian Inheritance in Man (Amberger et al. 2019) and Therapeutic Target Database (Wang et al. 2020). Click blue links to view details online. Drugs are highlighted in yellow. CS: cardiovascular system; NS: nervous system.

Trait	Gene (Role)	P_1^{base}	P_1^{near}	P_1^{bma}	P_1^{net} (Network, BF)	Mouse trait	Therapeutic/clinical evidence
BMI	*ADCY3* (TG)	1.00	1.00	1.00	1.00 (Ileum, 4.99×10^{12})	Growth	Severe obesity
	NPC1 (TG)	0.72	0.93	0.97	0.97 (Colon, 5.49×10^{12})	Growth	Niemann-Pick disease
	RPL27A (TG)	0.66	0.70	0.86	0.90 (Pancreas, 2.07×10^{13})	Growth	
WAIST	*PIGC* (TG)	1.00	1.00	1.00	1.00 (Esophagus, 6.78×10^{239})		GPIBD16
	TBX15 (TF)	1.00	1.00	1.00	1.00 (Pancreas, 2.72×10^{212})	Growth	Cousin syndrome
	PPARG (TF)	0.94	0.94	0.96	0.96 (Esophagus, 6.78×10^{239})	Metabolism	Insulin resistance, Obesity
BC	*FGFR2* (TG)	1.00	1.00	1.00	1.00 (Lung, 7.14×10^{7})	Growth	Pfeiffer syndrome
	PTHLH (TG)	1.00	1.00	1.00	1.00 (Aorta, 8.27×10^{8})	Growth	Brachydactyly E2
	MKL1 (TG)	0.80	0.80	0.91	0.92 (Spleen, 6.19×10^{7})		AMKL
RA	*CD40* (TG)	1.00	1.00	1.00	1.00 (B cell, 3.31×10^{57})	Immune	HIGM3
	CTLA4 (TG)	1.00	1.00	1.00	1.00 (CD8, 6.96×10^{56})	Immune	ALPS5
	ICOS (TG)	1.00	1.00	1.00	1.00 (CD8, 6.96×10^{56})	Immune	Immunodeficiency CV1
	IL2RA (TG)	1.00	1.00	1.00	1.00 (NK cell, 2.95×10^{60})	Immune	Immunodeficiency 41
	IL6R (TG)	1.00	1.00	1.00	1.00 (Monocyte, 8.91×10^{53})	Immune	
	RASGRP1 (TG)	1.00	1.00	1.00	1.00 (NK cell, 2.95×10^{60})	Immune	Immunodeficiency 64
	STAT4 (TF)	1.00	1.00	1.00	1.00 (NK cell, 2.95×10^{60})	Immune	
	TYK2 (TG)	1.00	1.00	1.00	1.00 (NK cell, 2.95×10^{60})	Immune	Immunodeficiency 35
IBD	*ATG16L1* (TG)	1.00	1.00	1.00	1.00 (NK cell, 5.07×10^{35})	Immune	
	BACH2 (TF)	1.00	1.00	1.00	1.00 (NK cell, 5.07×10^{35})	Immune	Immunodeficiency 60
	FCGR2A (TG)	1.00	1.00	1.00	1.00 (Monocyte, 6.28×10^{31})	Immune	Lupus nephritis
	JAK2 (TG)	1.00	1.00	1.00	1.00 (Monocyte, 6.28×10^{31})	Immune	Baricitinib
	STAT3 (TF)	1.00	1.00	1.00	1.00 (NK cell, 5.07×10^{35})	Immune	ADMIO1
	TYK2 (TG)	0.96	0.98	0.98	0.98 (NK cell, 5.07×10^{35})	Immune	Immunodeficiency 35
	STAT4 (TF)	0.77	0.82	0.91	0.91 (NK cell, 5.07×10^{35})	Immune	Susceptibility to SLE
HDL	*ABCA1* (TG)	1.00	1.00	1.00	1.00 (Liver, 2.81×10^{21})	Metabolism	Tangier disease, Probucol
	APOB (TG)	1.00	1.00	1.00	1.00 (Liver, 2.81×10^{21})	Metabolism	FCHL2, FHBL1, Mipomersen
	GALNT2 (TG)	1.00	1.00	1.00	1.00 (Liver, 2.81×10^{21})	Metabolism	
	LIPG (TG)	1.00	1.00	1.00	1.00 (Liver, 2.81×10^{21})	Metabolism	GSK-264220A
	LPL (TG)	1.00	1.00	1.00	1.00 (Omentum, 8.70×10^{13})	Metabolism	Clofibrate Gemfibrozil
	SCARB1 (TG)	1.00	1.00	1.00	1.00 (Liver, 2.81×10^{21})	Metabolism	
LDL	*APOB* (TG)	1.00	1.00	1.00	1.00 (Liver, 7.66×10^{27})	Metabolism	FCHL2, FHBL1, Mipomersen
	HMGCR (TG)	1.00	1.00	1.00	1.00 (CD8, 5.86×10^{28})	Metabolism	
	HNF1A (TF)	1.00	1.00	1.00	1.00 (CD8, 5.86×10^{28})	Metabolism	MODY3, Hepatic adenomas
	LDLR (TG)	1.00	1.00	1.00	1.00 (Pancreas, 3.04×10^{28})	Metabolism	FHCL1
	NPC1L1 (TG)	1.00	1.00	1.00	1.00 (Liver, 7.66×10^{27})	Metabolism	Response to ezetimibe
T2D	*TCF7L2* (TF)	1.00	1.00	1.00	1.00 (NK cell, 1.49×10^{77})	Metabolism	Susceptibility to T2D
	IGF2BP2 (TG)	0.99	1.00	1.00	1.00 (Ileum, 4.52×10^{62})	Metabolism	Susceptibility to T2D
	PPARG (TF)	0.98	1.00	1.00	1.00 (Prostate, 5.64×10^{66})	Metabolism	Insulin resistance, Obesity
	PROX1 (TG)	0.32	0.92	0.92	0.92 (Thyroid, 3.61×10^{61})	CS, Metabolism	
HR	*MYH6* (TG)	1.00	1.00	1.00	1.00 (Heart, 2.12×10^{7})	CS, Muscle	Cardiomyopathy D1EE, H14
	GJA1 (TG)	1.00	1.00	1.00	1.00 (Uterus, 1.51×10^{6})	CS, Muscle	AVSD3, HLHS1
	EPHB4 (TG)	1.00	1.00	1.00	1.00 (Aorta, 2.43×10^{7})	CS, Muscle	CMAVM2, LMPHM7
	TTN (TG)	1.00	1.00	1.00	1.00 (Aorta, 2.43×10^{7})	CS, Muscle	Cardiomyopathy D1G, H9
	GNB4 (TG)	0.55	0.64	0.91	0.91 (Aorta, 2.43×10^{7})	CS	CMTDIF
CAD	*SMAD3* (TF)	0.99	1.00	1.00	1.00 (Adipose, 1.67×10^{29})	CS, Blood	Loeys-Dietz syndrome 3
	APOB (TG)	0.94	0.98	0.99	0.99 (Liver, 2.99×10^{25})	CS, Metabolism	FCHL2, FHBL1, Mipomersen
	ATP2B1 (TG)	0.91	0.96	0.98	0.99 (Heart, 1.93×10^{28})	CS, Muscle	
	MRAS (TG)	0.89	0.96	0.98	0.98 (Adipose, 1.67×10^{29})	Immune	Noonan syndrome 11
	APOE (TG)	0.80	0.92	0.96	0.96 (Adipose, 1.67×10^{29})	CS, Metabolism	Hyperlipoproteinemia 3
	ABHD2 (TG)	0.81	0.92	0.95	0.95 (Adipose, 1.67×10^{29})	CS	
	IL6R (TG)	0.77	0.91	0.93	0.94 (Heart, 1.93×10^{28})	Immune	Serum level of IL6
	FURIN (TG)	0.69	0.86	0.93	0.95 (Heart, 1.93×10^{28})	CS	
AF	*TBX5* (TF)	1.00	1.00	1.00	1.00 (Heart, 2.15×10^{14})	CS, Muscle	Holt-Oram syndrome
	PITX2 (TF)	1.00	1.00	1.00	1.00 (Muscle, 8.55×10^{14})	CS, Muscle	
	CAV1 (TG)	1.00	1.00	1.00	1.00 (Muscle, 8.55×10^{14})	CS, Muscle	PP hypertension 3
	SH3PXD2A (TG)	1.00	1.00	1.00	1.00 (Muscle, 8.55×10^{14})	CS, Muscle	

(continued)

Trait	Gene (Role)	P_1^{base}	P_1^{near}	P_1^{bma}	P_1^{net} (Network, BF)	Mouse trait	Therapeutic/clinical evidence
	ASAH1 (TG)	1.00	1.00	1.00	1.00 (Muscle, 8.55×10^{14})	Metabolism	Farber lipogranulomatosis
	TTN (TG)	0.95	0.97	0.99	0.99 (Muscle, 8.55×10^{14})	CS, Muscle	Cardiomyopathy D1G, H9
LOAD	*APOE* (TG)	1.00	1.00	1.00	1.00 (Liver, 1.09×10^{20})	Liver, NS	Hyperlipoproteinemia 3
	BIN1 (TG)	1.00	1.00	1.00	1.00 (CD8, 8.31×10^{26})	Muscle	Centronuclear myopathy 2
	CLU (TG)	1.00	1.00	1.00	1.00 (Aorta, 3.24×10^{23})	CS, Muscle	
	EPHA1 (TG)	0.99	1.00	1.00	1.00 (CD8, 8.31×10^{26})	Immune	
	CD2AP (TG)	0.70	0.95	0.98	0.98 (Pancreas, 3.53×10^{20})	Metabolism	FSGS3
SCZ	*CACNA1C* (TG)	1.00	1.00	1.00	1.00 (Spleen, 1.44×10^{141})	Immune, NS	Timothy syndrome
	TCF4 (TF)	1.00	1.00	1.00	1.00 (Colon, 1.20×10^{144})	Immune, NS	Pitt-Hopkins syndrome
	DPYD (TG)	1.00	1.00	1.00	1.00 (Spleen, 1.44×10^{141})	Neuron	DPD deficiency
NEU	*LINGO1* (TG)	0.99	1.00	1.00	1.00 (Putamen, 2.12×10^{19})	NS	Mental retardation
	SBF2 (TG)	0.94	0.97	0.99	0.99 (Lung, 1.42×10^{18})	Neuron, NS	CMT4B2
	PAFAH1B1 (TG)	0.77	0.93	1.00	1.00 (Cerebellum, 4.85×10^{18})	Neuron, NS	Lissencephaly

Supplementary Table 12

For each of 14 GWAS traits, we quantify overlap between RSS-NET prioritized genes ($P_1^{bma} \geq 0.9$) and genes implicated in 27 categories of knockout mouse phenotypes (Bult et al. 2019). These 14 traits are obtained by removing 2 disease subtypes (MI for CAD, UC for IBD) from the 16 traits that pass the near-gene control (**Supplementary Table 3**). Both odds ratios and p-values are generated by R built-in function `fisher.test`. Within a trait, rows are sorted by p-values, then odds ratios. Trait abbreviations are defined in **Supplementary Table 2**.

Trait	Category	Odds ratio	- Log 10 p-value
AF	Muscle	3.8780	3.3553
AF	Skeleton	3.1715	2.7380
AF	Cardiovascular system	2.9168	2.6480
AF	Neoplasm	4.1098	2.2991
AF	Respiratory system	3.2481	2.2458
AF	Renal/urinary system	3.0490	1.9599
AF	Growth/size/body region	2.3452	1.9315
AF	Craniofacial	2.9981	1.7941
AF	Nervous system	2.3314	1.7879
AF	Behavior/neurological	2.2534	1.7081
AF	Reproductive system	2.2865	1.4005
AF	Endocrine/exocrine gland	2.2613	1.3426
AF	Liver/biliary system	2.5714	1.3407
AF	Digestive/alimentary	2.5213	1.3217
AF	Mortality/aging	1.9200	1.2984
AF	Homeostasis/metabolism	1.9021	1.1774
AF	Embryo	2.0787	1.0617
AF	Hematopoietic system	1.8645	1.0576
AF	Integument	1.9662	0.8450
AF	Immune system	1.7301	0.7664
AF	Limbs/digits/tail	2.0915	0.7406
AF	Vision/eye	1.8578	0.6876
AF	Adipose tissue	1.8256	0.6560

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
AF	Taste/olfaction	3.6424	0.5955
AF	Cellular	1.5843	0.5860
AF	Hearing/vestibular/ear	1.7927	0.3800
AF	Pigmentation	0.9310	0.0000
BC	Cardiovascular system	4.5988	2.7953
BC	Skeleton	4.1864	2.2045
BC	Renal/urinary system	5.0749	2.1235
BC	Limbs/digits/tail	5.2364	1.9563
BC	Neoplasm	5.8486	1.8542
BC	Integument	3.8241	1.7701
BC	Adipose tissue	4.5710	1.7457
BC	Mortality/aging	2.9204	1.6103
BC	Respiratory system	4.0477	1.5656
BC	Muscle	3.4434	1.3391
BC	Hearing/vestibular/ear	4.4885	1.2709
BC	Vision/eye	2.9016	1.1170
BC	Liver/biliary system	3.2077	1.1120
BC	Immune system	2.4704	1.0162
BC	Growth/size/body region	2.2887	0.9258
BC	Homeostasis/metabolism	2.1587	0.8607
BC	Craniofacial	2.8003	0.8402
BC	Nervous system	2.2612	0.8282
BC	Hematopoietic system	2.1727	0.8004
BC	Behavior/neurological	2.3670	0.7950
BC	Cellular	2.1116	0.7540
BC	Digestive/alimentary	2.3570	0.7021
BC	Embryo	2.3058	0.6086
BC	Reproductive system	2.0735	0.5682
BC	Endocrine/exocrine gland	1.8794	0.5221
BC	Pigmentation	2.3310	0.4158
BC	Taste/olfaction	0.0000	0.0000
BMI	Renal/urinary system	4.7663	4.1616
BMI	Liver/biliary system	3.8716	3.0521
BMI	Embryo	3.2442	2.5972
BMI	Vision/eye	2.7942	2.0096
BMI	Growth/size/body region	2.2426	1.7571
BMI	Adipose tissue	2.9300	1.7459
BMI	Reproductive system	2.4959	1.6468
BMI	Mortality/aging	2.0040	1.4467
BMI	Taste/olfaction	7.3601	1.4321
BMI	Respiratory system	2.5938	1.3484
BMI	Homeostasis/metabolism	1.9021	1.1774

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
BMI	Endocrine/exocrine gland	2.0718	1.1354
BMI	Nervous system	1.9407	1.0905
BMI	Behavior/neurological	1.8957	1.0441
BMI	Hematopoietic system	1.7396	0.7679
BMI	Cardiovascular system	1.6827	0.6951
BMI	Integument	1.7466	0.6635
BMI	Cellular	1.5843	0.5860
BMI	Pigmentation	1.8667	0.5007
BMI	Immune system	1.4819	0.4984
BMI	Limbs/digits/tail	1.6712	0.4942
BMI	Digestive/alimentary	1.5716	0.4248
BMI	Hearing/vestibular/ear	1.7927	0.3800
BMI	Craniofacial	1.4927	0.2872
BMI	Skeleton	1.2990	0.2019
BMI	Muscle	1.0993	0.1098
BMI	Neoplasm	0.5813	0.0000
CAD	Cardiovascular system	2.4480	3.6550
CAD	Liver/biliary system	2.7126	2.8632
CAD	Mortality/aging	1.8468	2.2416
CAD	Embryo	2.1494	2.0684
CAD	Renal/urinary system	2.2472	1.9298
CAD	Hematopoietic system	1.8149	1.7038
CAD	Cellular	1.7295	1.5700
CAD	Respiratory system	2.0067	1.4745
CAD	Limbs/digits/tail	2.0375	1.3155
CAD	Integument	1.8362	1.3120
CAD	Digestive/alimentary	1.9504	1.2725
CAD	Immune system	1.6391	1.1899
CAD	Neoplasm	2.0683	1.1477
CAD	Growth/size/body region	1.5289	1.0838
CAD	Endocrine/exocrine gland	1.6640	1.0655
CAD	Muscle	1.7048	1.0016
CAD	Reproductive system	1.6517	0.9868
CAD	Nervous system	1.5425	0.9853
CAD	Homeostasis/metabolism	1.4500	0.9143
CAD	Vision/eye	1.5385	0.7714
CAD	Pigmentation	2.0683	0.7389
CAD	Craniofacial	1.6511	0.6966
CAD	Adipose tissue	1.4506	0.5316
CAD	Skeleton	1.3115	0.3763
CAD	Taste/olfaction	1.6013	0.3246
CAD	Behavior/neurological	1.2007	0.3078

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
CAD	Hearing/vestibular/ear	0.7881	0.0000
HDL	Liver/biliary system	2.2276	2.8889
HDL	Digestive/alimentary	2.1834	2.8337
HDL	Homeostasis/metabolism	1.7244	2.7650
HDL	Embryo	1.7717	1.8381
HDL	Cardiovascular system	1.6423	1.7857
HDL	Immune system	1.5138	1.3889
HDL	Hematopoietic system	1.4900	1.2939
HDL	Respiratory system	1.6534	1.1598
HDL	Mortality/aging	1.3819	1.0779
HDL	Vision/eye	1.5409	1.0602
HDL	Endocrine/exocrine gland	1.4882	1.0448
HDL	Reproductive system	1.4287	0.8866
HDL	Cellular	1.3458	0.8616
HDL	Adipose tissue	1.4901	0.8217
HDL	Growth/size/body region	1.2956	0.6989
HDL	Renal/urinary system	1.3761	0.5650
HDL	Limbs/digits/tail	1.3843	0.5225
HDL	Muscle	1.2611	0.4131
HDL	Behavior/neurological	1.1740	0.3378
HDL	Integument	1.1653	0.2255
HDL	Hearing/vestibular/ear	0.7559	0.1683
HDL	Neoplasm	1.1862	0.1652
HDL	Skeleton	1.0853	0.1503
HDL	Nervous system	1.0836	0.1306
HDL	Pigmentation	0.7078	0.0983
HDL	Craniofacial	0.9469	0.0000
HDL	Taste/olfaction	0.9184	0.0000
HR	Endocrine/exocrine gland	4.4457	2.7298
HR	Muscle	4.7326	2.4456
HR	Immune system	3.7127	2.3428
HR	Reproductive system	4.0095	2.2113
HR	Nervous system	3.6067	2.1380
HR	Mortality/aging	3.2209	2.1176
HR	Cardiovascular system	3.6111	2.0878
HR	Vision/eye	3.9871	2.0672
HR	Behavior/neurological	3.3019	2.0200
HR	Cellular	3.1725	1.9161
HR	Respiratory system	4.1665	1.8440
HR	Hematopoietic system	3.1975	1.8419
HR	Liver/biliary system	4.1310	1.8292
HR	Digestive/alimentary	4.0504	1.7953

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
HR	Renal/urinary system	3.6209	1.4562
HR	Homeostasis/metabolism	2.5926	1.3434
HR	Limbs/digits/tail	3.5860	1.2746
HR	Growth/size/body region	2.6174	1.2431
HR	Hearing/vestibular/ear	3.8463	1.1488
HR	Neoplasm	3.7538	1.1256
HR	Adipose tissue	3.1306	1.1183
HR	Integument	2.8074	1.0377
HR	Skeleton	2.3881	0.9234
HR	Embryo	2.4717	0.8052
HR	Craniofacial	2.3997	0.7347
HR	Pigmentation	1.9977	0.3715
HR	Taste/olfaction	0.0000	0.0000
IBD	Hematopoietic system	2.3478	4.4776
IBD	Neoplasm	3.4435	4.4083
IBD	Immune system	2.2653	4.1384
IBD	Digestive/alimentary	2.8482	4.1302
IBD	Homeostasis/metabolism	2.0810	3.4697
IBD	Adipose tissue	2.7403	3.0439
IBD	Endocrine/exocrine gland	2.1975	2.9543
IBD	Cardiovascular system	2.0274	2.6414
IBD	Nervous system	2.0217	2.6062
IBD	Renal/urinary system	2.3600	2.3661
IBD	Cellular	1.7366	1.9069
IBD	Liver/biliary system	2.0181	1.7387
IBD	Respiratory system	1.9313	1.5639
IBD	Behavior/neurological	1.6893	1.5322
IBD	Limbs/digits/tail	2.0844	1.5313
IBD	Skeleton	1.7443	1.5251
IBD	Reproductive system	1.7703	1.4224
IBD	Mortality/aging	1.5626	1.4125
IBD	Growth/size/body region	1.5397	1.2589
IBD	Embryo	1.6135	1.1067
IBD	Integument	1.5728	0.9934
IBD	Vision/eye	1.5525	0.9134
IBD	Muscle	1.6051	0.8612
IBD	Craniofacial	1.5265	0.7194
IBD	Pigmentation	1.2941	0.2561
IBD	Hearing/vestibular/ear	1.1066	0.0957
IBD	Taste/olfaction	0.0000	0.0000
LDL	Liver/biliary system	2.1639	2.7494
LDL	Digestive/alimentary	1.9657	2.1606

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
LDL	Skeleton	0.4825	1.6260
LDL	Homeostasis/metabolism	1.2564	0.6275
LDL	Growth/size/body region	1.2124	0.4771
LDL	Renal/urinary system	1.2877	0.4428
LDL	Embryo	1.2092	0.3625
LDL	Respiratory system	1.2331	0.3445
LDL	Nervous system	1.1386	0.2792
LDL	Cellular	1.1316	0.2585
LDL	Cardiovascular system	0.8711	0.1941
LDL	Hematopoietic system	1.0936	0.1715
LDL	Behavior/neurological	1.0986	0.1685
LDL	Hearing/vestibular/ear	1.1376	0.1606
LDL	Limbs/digits/tail	0.7923	0.1378
LDL	Endocrine/exocrine gland	1.0740	0.0936
LDL	Immune system	1.0578	0.0797
LDL	Neoplasm	0.8296	0.0747
LDL	Mortality/aging	1.0514	0.0701
LDL	Craniofacial	0.8861	0.0606
LDL	Adipose tissue	1.0413	0.0593
LDL	Integument	1.0379	0.0475
LDL	Muscle	0.9810	0.0000
LDL	Vision/eye	0.9928	0.0000
LDL	Reproductive system	0.9861	0.0000
LDL	Pigmentation	0.8856	0.0000
LDL	Taste/olfaction	0.8594	0.0000
LOAD	Pigmentation	3.8236	1.2047
LOAD	Muscle	2.5444	1.1586
LOAD	Homeostasis/metabolism	1.9630	1.0599
LOAD	Integument	2.3395	1.0035
LOAD	Immune system	1.9411	0.8937
LOAD	Vision/eye	2.1761	0.8177
LOAD	Hearing/vestibular/ear	2.4519	0.7973
LOAD	Renal/urinary system	2.2619	0.7740
LOAD	Taste/olfaction	5.2412	0.7254
LOAD	Cardiovascular system	1.8327	0.6507
LOAD	Behavior/neurological	1.6202	0.5774
LOAD	Embryo	0.3004	0.4990
LOAD	Digestive/alimentary	1.6660	0.4852
LOAD	Liver/biliary system	1.6610	0.4837
LOAD	Growth/size/body region	1.5015	0.4136
LOAD	Limbs/digits/tail	1.6488	0.3602
LOAD	Craniofacial	1.4943	0.3319

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
LOAD	Skeleton	0.4926	0.2725
LOAD	Reproductive system	1.3954	0.2502
LOAD	Hematopoietic system	1.2992	0.1978
LOAD	Cellular	1.2496	0.1862
LOAD	Endocrine/exocrine gland	1.2446	0.1093
LOAD	Nervous system	0.7067	0.1047
LOAD	Mortality/aging	1.1014	0.0809
LOAD	Respiratory system	0.8581	0.0000
LOAD	Adipose tissue	0.9856	0.0000
LOAD	Neoplasm	0.7428	0.0000
NEU	Endocrine/exocrine gland	2.5688	1.3726
NEU	Mortality/aging	2.0481	1.2050
NEU	Muscle	2.6313	1.1846
NEU	Cellular	2.1622	1.1716
NEU	Growth/size/body region	2.0823	1.1508
NEU	Reproductive system	2.2652	0.9882
NEU	Skeleton	2.0270	0.9043
NEU	Embryo	2.2034	0.8214
NEU	Respiratory system	2.2060	0.7619
NEU	Digestive/alimentary	2.1447	0.7470
NEU	Taste/olfaction	4.9690	0.7054
NEU	Nervous system	1.7623	0.6061
NEU	Hematopoietic system	1.6933	0.5955
NEU	Immune system	1.6841	0.5936
NEU	Behavior/neurological	1.6136	0.5753
NEU	Liver/biliary system	1.7482	0.5094
NEU	Cardiovascular system	1.6679	0.4985
NEU	Vision/eye	1.5815	0.4267
NEU	Limbs/digits/tail	1.7084	0.3692
NEU	Craniofacial	1.5261	0.3383
NEU	Adipose tissue	1.4917	0.3314
NEU	Hearing/vestibular/ear	0.0000	0.2102
NEU	Neoplasm	1.5890	0.1965
NEU	Renal/urinary system	1.3790	0.1455
NEU	Integument	1.1888	0.1180
NEU	Homeostasis/metabolism	1.1764	0.0828
NEU	Pigmentation	0.0000	0.0000
RA	Hematopoietic system	2.1432	5.6619
RA	Immune system	2.0536	5.0232
RA	Endocrine/exocrine gland	2.1380	4.4808
RA	Neoplasm	2.8695	4.2020
RA	Cellular	1.7482	3.1870

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
RA	Digestive/alimentary	2.1046	3.1067
RA	Liver/biliary system	1.9467	2.4249
RA	Homeostasis/metabolism	1.4940	1.8938
RA	Respiratory system	1.7623	1.7867
RA	Cardiovascular system	1.4759	1.4185
RA	Nervous system	1.4548	1.3930
RA	Behavior/neurological	1.4026	1.2454
RA	Skeleton	1.4177	1.0676
RA	Reproductive system	1.4169	0.9419
RA	Vision/eye	1.3981	0.9018
RA	Renal/urinary system	1.4234	0.7822
RA	Pigmentation	1.6929	0.7460
RA	Growth/size/body region	1.2644	0.7288
RA	Integument	1.3131	0.6916
RA	Mortality/aging	1.2365	0.6722
RA	Limbs/digits/tail	1.4266	0.6570
RA	Muscle	1.2688	0.5139
RA	Embryo	1.2479	0.4702
RA	Adipose tissue	1.2429	0.4129
RA	Hearing/vestibular/ear	1.1959	0.2334
RA	Taste/olfaction	0.0000	0.1891
RA	Craniofacial	1.0441	0.0548
SCZ	Nervous system	2.0448	13.9729
SCZ	Behavior/neurological	1.8956	11.4605
SCZ	Growth/size/body region	1.7744	9.8894
SCZ	Muscle	2.0881	9.1153
SCZ	Mortality/aging	1.5110	5.6120
SCZ	Cardiovascular system	1.6237	5.5988
SCZ	Craniofacial	1.8433	4.9785
SCZ	Homeostasis/metabolism	1.4799	4.9427
SCZ	Hematopoietic system	1.4807	4.1263
SCZ	Skeleton	1.5519	4.1013
SCZ	Cellular	1.4534	4.0948
SCZ	Reproductive system	1.5459	3.7741
SCZ	Hearing/vestibular/ear	1.8184	3.3949
SCZ	Adipose tissue	1.6532	3.3917
SCZ	Neoplasm	1.8095	3.3450
SCZ	Endocrine/exocrine gland	1.4635	3.0824
SCZ	Vision/eye	1.5030	3.0335
SCZ	Digestive/alimentary	1.5493	2.9284
SCZ	Respiratory system	1.5340	2.6693
SCZ	Renal/urinary system	1.5406	2.6429

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
SCZ	Embryo	1.4188	2.2963
SCZ	Immune system	1.3215	2.2105
SCZ	Taste/olfaction	2.5060	2.0788
SCZ	Liver/biliary system	1.3755	1.6197
SCZ	Integument	1.3093	1.5395
SCZ	Limbs/digits/tail	1.3155	1.0486
SCZ	Pigmentation	0.6986	0.5870
T2D	Liver/biliary system	10.7822	3.6438
T2D	Digestive/alimentary	9.2688	3.0391
T2D	Integument	7.5090	2.7724
T2D	Respiratory system	8.1430	2.5400
T2D	Taste/olfaction	35.4154	2.4666
T2D	Cardiovascular system	5.8080	2.3365
T2D	Muscle	7.0758	2.2827
T2D	Adipose tissue	7.8634	2.2367
T2D	Growth/size/body region	5.2016	2.1631
T2D	Renal/urinary system	7.2051	2.0983
T2D	Endocrine/exocrine gland	5.6144	2.0270
T2D	Vision/eye	5.9885	1.9928
T2D	Embryo	5.6009	1.8816
T2D	Mortality/aging	4.2136	1.7310
T2D	Neoplasm	7.1503	1.5540
T2D	Homeostasis/metabolism	3.8693	1.3285
T2D	Behavior/neurological	3.6817	1.2693
T2D	Immune system	3.6582	1.2626
T2D	Craniofacial	4.5869	1.1244
T2D	Hematopoietic system	3.1525	1.0049
T2D	Nervous system	2.7917	0.7970
T2D	Cellular	2.7035	0.7387
T2D	Reproductive system	2.8471	0.7309
T2D	Skeleton	3.1172	0.6819
T2D	Hearing/vestibular/ear	2.5070	0.4050
T2D	Limbs/digits/tail	1.7634	0.3010
T2D	Pigmentation	0.0000	0.0000
WAIST	Limbs/digits/tail	4.4650	3.6275
WAIST	Muscle	3.1734	2.4107
WAIST	Embryo	2.6553	1.8770
WAIST	Vision/eye	2.4646	1.6438
WAIST	Skeleton	2.3001	1.6207
WAIST	Integument	2.3166	1.5492
WAIST	Respiratory system	2.5759	1.5028
WAIST	Immune system	2.0747	1.4633

(continued)

Trait	Category	Odds ratio	- Log 10 p-value
WAIST	Craniofacial	2.6443	1.3660
WAIST	Pigmentation	3.3105	1.3238
WAIST	Behavior/neurological	1.9875	1.2686
WAIST	Mortality/aging	1.7675	1.0824
WAIST	Growth/size/body region	1.7971	1.0408
WAIST	Renal/urinary system	2.0877	0.9982
WAIST	Nervous system	1.7117	0.8647
WAIST	Cardiovascular system	1.7555	0.7912
WAIST	Digestive/alimentary	1.9443	0.7654
WAIST	Hematopoietic system	1.6446	0.7152
WAIST	Neoplasm	2.0609	0.5768
WAIST	Liver/biliary system	1.6980	0.5735
WAIST	Cellular	1.4910	0.5472
WAIST	Reproductive system	1.4641	0.4480
WAIST	Adipose tissue	1.6102	0.4317
WAIST	Endocrine/exocrine gland	1.4935	0.4310
WAIST	Hearing/vestibular/ear	1.5812	0.3508
WAIST	Homeostasis/metabolism	1.2183	0.2212
WAIST	Taste/olfaction	0.0000	0.0000

Supplementary Table 13

For each of 14 GWAS traits shown in **Supplementary Table 12**, we quantify overlap between RSS-NET prioritized genes ($P_1^{\text{bma}} \geq 0.9$) and genes causing 19 categories of Mendelian disorders (Freund et al. 2018; Amberger et al. 2019). The rest is the same as **Supplementary Table 12**.

Trait	Category	Odds ratio	- Log 10 p-value
AF	Arrythmia	8.2819	4.7706
AF	Cardiovascular	3.3123	2.0595
AF	Diabetes	3.2124	0.8635
AF	Positive mood	5.0229	0.7279
AF	Renal	1.7262	0.6412
AF	Immune	1.5268	0.3428
AF	MODY	1.5138	0.3405
AF	Psychiatric	1.1300	0.2269
AF	Reproductive	1.0313	0.2034
AF	Hematologic	1.0427	0.1437
AF	Growth	0.4044	0.1356
AF	Insulin	0.9136	0.0000
AF	Platelet	0.6298	0.0000
AF	Uric acid	0.0000	0.0000

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
AF	Neurologic	0.0000	0.0000
AF	Development	0.5145	0.0000
AF	Autism	0.0000	0.0000
AF	Weight	0.0000	0.0000
AF	Microalbumin	0.0000	0.0000
BC	Insulin	9.7085	4.3523
BC	Cardiovascular	7.0883	2.7668
BC	Platelet	5.7120	1.6533
BC	Renal	4.1481	1.5907
BC	Hematologic	4.7137	1.4476
BC	Uric acid	7.4760	1.4452
BC	Immune	4.5921	1.4201
BC	MODY	4.5534	1.4111
BC	Reproductive	6.2296	1.3059
BC	Growth	3.6616	1.1895
BC	Positive mood	15.0984	1.1573
BC	Weight	6.3976	0.8111
BC	Development	3.1020	0.8083
BC	Microalbumin	5.6364	0.7615
BC	Diabetes	4.7989	0.6994
BC	Psychiatric	3.3987	0.5705
BC	Arrythmia	0.0000	0.0000
BC	Autism	0.0000	0.0000
BC	Neurologic	0.0000	0.0000
BMI	Insulin	2.5398	1.2229
BMI	Weight	4.7461	1.1383
BMI	Platelet	2.1001	0.7298
BMI	Psychiatric	2.5105	0.6996
BMI	Autism	3.0009	0.5363
BMI	Renal	1.5248	0.4594
BMI	Microalbumin	2.0723	0.4084
BMI	Hematologic	1.7331	0.3715
BMI	Cardiovascular	1.5573	0.3479
BMI	Neurologic	1.5270	0.3118
BMI	Growth	1.3460	0.3052
BMI	Uric acid	1.3671	0.2792
BMI	Arrythmia	1.1112	0.2222
BMI	Development	1.1403	0.1572
BMI	Immune	1.1231	0.1550
BMI	MODY	1.1136	0.1537
BMI	Reproductive	0.0000	0.0000
BMI	Diabetes	0.0000	0.0000

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
BMI	Positive mood	0.0000	0.0000
CAD	Cardiovascular	3.7000	4.0128
CAD	Platelet	3.1890	2.4394
CAD	Insulin	2.8159	2.4386
CAD	Uric acid	4.1942	2.3464
CAD	Hematologic	2.6244	1.9345
CAD	Neurologic	3.8932	1.9182
CAD	Renal	2.3052	1.8685
CAD	Development	2.3019	1.5005
CAD	Microalbumin	3.1461	1.1113
CAD	MODY	1.9620	0.9924
CAD	Reproductive	2.3018	0.9722
CAD	Autism	3.0241	0.8268
CAD	Immune	1.6927	0.7356
CAD	Growth	1.5751	0.6435
CAD	Positive mood	2.7668	0.5081
CAD	Diabetes	1.7695	0.4953
CAD	Arrythmia	1.6745	0.3642
CAD	Weight	1.1717	0.2376
CAD	Psychiatric	1.2506	0.1698
HDL	Cardiovascular	2.6916	3.2435
HDL	Diabetes	4.3920	3.0738
HDL	Insulin	2.4627	2.5952
HDL	Weight	4.3817	2.4423
HDL	MODY	2.3718	2.2229
HDL	Microalbumin	3.1728	1.5943
HDL	Renal	1.4890	0.7790
HDL	Growth	1.4838	0.7174
HDL	Hematologic	1.5624	0.6987
HDL	Immune	1.5215	0.6792
HDL	Platelet	1.4688	0.4625
HDL	Positive mood	1.6474	0.3357
HDL	Neurologic	1.3745	0.3131
HDL	Arrythmia	1.3350	0.2604
HDL	Development	1.1965	0.1749
HDL	Uric acid	1.2291	0.1317
HDL	Autism	0.8900	0.0000
HDL	Psychiatric	0.7446	0.0000
HDL	Reproductive	0.6793	0.0000
HR	Arrythmia	7.1605	2.9312
HR	Cardiovascular	2.6375	1.0982
HR	Neurologic	1.9357	0.3848

(continued)

Trait	Category	Odds ratio	- Log 10 p-value
HR	Renal	1.4476	0.3273
HR	Psychiatric	1.5838	0.3217
HR	Immune	0.0000	0.1912
HR	MODY	0.0000	0.1905
HR	Development	1.4455	0.1863
HR	Growth	0.5668	0.0000
HR	Diabetes	0.0000	0.0000
HR	Platelet	0.8828	0.0000
HR	Hematologic	0.7292	0.0000
HR	Insulin	0.6391	0.0000
HR	Autism	0.0000	0.0000
HR	Weight	0.0000	0.0000
HR	Uric acid	0.0000	0.0000
HR	Reproductive	0.0000	0.0000
HR	Microalbumin	0.0000	0.0000
HR	Positive mood	0.0000	0.0000
IBD	Immune	4.3157	6.2205
IBD	MODY	3.9192	4.9489
IBD	Psychiatric	5.1647	3.7971
IBD	Uric acid	4.4576	3.0596
IBD	Renal	2.4535	2.4167
IBD	Growth	2.4471	2.1182
IBD	Neurologic	3.4789	1.4632
IBD	Insulin	1.9937	1.1753
IBD	Development	1.9060	1.0714
IBD	Arrythmia	2.1069	0.8687
IBD	Autism	3.1029	0.8340
IBD	Cardiovascular	1.7466	0.8266
IBD	Hematologic	1.7204	0.8150
IBD	Platelet	1.6724	0.5609
IBD	Diabetes	1.6888	0.4664
IBD	Reproductive	1.4599	0.3307
IBD	Weight	1.2484	0.2532
IBD	Microalbumin	0.8571	0.0000
IBD	Positive mood	0.0000	0.0000
LDL	Cardiovascular	3.4028	4.9572
LDL	MODY	2.3471	2.0577
LDL	Hematologic	2.2401	1.8306
LDL	Uric acid	2.6674	1.5063
LDL	Platelet	2.0280	1.2829
LDL	Insulin	1.7907	1.1697
LDL	Renal	1.5904	0.9586

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
LDL	Reproductive	0.0000	0.9300
LDL	Neurologic	1.9690	0.8055
LDL	Development	1.6503	0.7295
LDL	Immune	1.6251	0.7223
LDL	Microalbumin	2.0008	0.7018
LDL	Arrythmia	1.4259	0.4283
LDL	Weight	1.5036	0.4083
LDL	Diabetes	1.6993	0.3668
LDL	Psychiatric	0.3957	0.2818
LDL	Positive mood	0.0000	0.0000
LDL	Autism	0.9506	0.0000
LDL	Growth	0.8571	0.0000
LOAD	Immune	7.3247	3.7702
LOAD	Renal	5.7313	3.4911
LOAD	Platelet	7.7806	3.4797
LOAD	MODY	6.1917	2.9835
LOAD	Cardiovascular	5.7936	2.8427
LOAD	Arrythmia	6.5840	1.8439
LOAD	Hematologic	4.2961	1.6845
LOAD	Microalbumin	7.6902	1.4847
LOAD	Insulin	3.6737	1.4768
LOAD	Neurologic	5.8814	1.2798
LOAD	Uric acid	5.0557	1.1671
LOAD	Weight	4.4554	0.6768
LOAD	Diabetes	3.2146	0.5557
LOAD	Psychiatric	2.4327	0.4574
LOAD	Growth	1.6702	0.4486
LOAD	Development	1.1033	0.2174
LOAD	Reproductive	0.0000	0.0000
LOAD	Autism	0.0000	0.0000
LOAD	Positive mood	0.0000	0.0000
NEU	Arrythmia	4.7359	1.5033
NEU	Positive mood	7.8232	0.8993
NEU	Platelet	1.9673	0.5455
NEU	Development	1.6066	0.4336
NEU	Immune	1.5824	0.4257
NEU	Neurologic	2.1514	0.4188
NEU	Renal	1.6089	0.3559
NEU	Psychiatric	1.7603	0.3537
NEU	Cardiovascular	1.4599	0.1873
NEU	Growth	1.2621	0.1711
NEU	Hematologic	0.8104	0.0000

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
NEU	MODY	0.7828	0.0000
NEU	Diabetes	0.0000	0.0000
NEU	Insulin	0.7103	0.0000
NEU	Autism	0.0000	0.0000
NEU	Weight	0.0000	0.0000
NEU	Uric acid	0.0000	0.0000
NEU	Reproductive	0.0000	0.0000
NEU	Microalbumin	0.0000	0.0000
RA	Immune	3.1132	5.0387
RA	MODY	2.2506	2.4678
RA	Hematologic	2.0490	1.8464
RA	Uric acid	2.6130	1.7991
RA	Platelet	1.8113	1.1918
RA	Renal	1.6124	1.1833
RA	Positive mood	2.6130	0.7276
RA	Insulin	1.4227	0.6579
RA	Psychiatric	0.2876	0.5730
RA	Growth	1.3661	0.5013
RA	Cardiovascular	1.2119	0.3207
RA	Development	1.1995	0.2427
RA	Neurologic	1.0670	0.1186
RA	Arrythmia	1.0364	0.0991
RA	Reproductive	0.7944	0.0000
RA	Microalbumin	0.9618	0.0000
RA	Diabetes	0.8178	0.0000
RA	Autism	0.6909	0.0000
RA	Weight	0.5416	0.0000
SCZ	Development	2.2261	6.2565
SCZ	Psychiatric	2.2240	3.2105
SCZ	Positive mood	2.2484	1.1808
SCZ	Autism	1.7932	0.9620
SCZ	Uric acid	0.5078	0.9201
SCZ	Arrythmia	1.3782	0.6777
SCZ	Weight	1.5432	0.6495
SCZ	Immune	1.2360	0.6318
SCZ	Reproductive	1.3365	0.5863
SCZ	Hematologic	1.2309	0.5734
SCZ	Diabetes	1.3787	0.5319
SCZ	Platelet	1.1693	0.3580
SCZ	Insulin	1.1383	0.3478
SCZ	Renal	1.1261	0.3414
SCZ	Growth	0.8808	0.2301

(continued)

Trait	Category	Odds ratio	- Log 10 *p*-value
SCZ	Microalbumin	0.7806	0.1511
SCZ	MODY	1.0328	0.0745
SCZ	Neurologic	0.8692	0.0603
SCZ	Cardiovascular	0.9570	0.0347
T2D	Immune	11.1690	4.0706
T2D	Weight	28.0348	3.4862
T2D	Diabetes	19.2678	3.0299
T2D	Neurologic	16.0917	2.8130
T2D	Psychiatric	11.6802	2.4318
T2D	MODY	7.4238	2.3581
T2D	Cardiovascular	6.8905	2.2496
T2D	Insulin	6.4877	2.1623
T2D	Microalbumin	12.8519	1.8510
T2D	Reproductive	8.0723	1.4849
T2D	Positive mood	20.4317	1.2714
T2D	Renal	3.6966	1.1726
T2D	Platelet	4.7481	1.0877
T2D	Development	3.5809	0.8903
T2D	Growth	2.9908	0.7709
T2D	Uric acid	4.8519	0.6969
T2D	Hematologic	1.8633	0.3604
T2D	Arrythmia	0.0000	0.0000
T2D	Autism	0.0000	0.0000
WAIST	Platelet	3.3419	1.6388
WAIST	Diabetes	5.0941	1.5987
WAIST	Cardiovascular	2.4750	1.1895
WAIST	Renal	2.1787	0.9203
WAIST	Uric acid	2.6103	0.7253
WAIST	Insulin	1.9263	0.5552
WAIST	Growth	1.7075	0.5131
WAIST	Weight	2.2342	0.4338
WAIST	Microalbumin	1.9683	0.3917
WAIST	Development	0.0000	0.3743
WAIST	Immune	1.6035	0.3551
WAIST	MODY	1.5898	0.3530
WAIST	Neurologic	1.4503	0.2967
WAIST	Psychiatric	1.1867	0.2399
WAIST	Reproductive	1.0831	0.2157
WAIST	Arrythmia	1.0554	0.2091
WAIST	Hematologic	1.0950	0.1512
WAIST	Autism	0.0000	0.0000
WAIST	Positive mood	0.0000	0.0000

Supplementary Table 14

Descriptive statistics of proportion of GWAS SNPs inside a network ($a_j = 1$) over 18 traits for each of 38 regulatory networks. Rows are sorted by the "Median" column.

	Summary of $\sum_{j=1}^{p} a_j/p$				
Network	Min	Q1	Median	Q3	Max
Vagina	0.6400	0.6423	0.6438	0.6481	0.6505
Lung	0.6377	0.6399	0.6414	0.6460	0.6485
Thyroid	0.6287	0.6305	0.6324	0.6368	0.6394
Terminal ileum	0.6145	0.6170	0.6188	0.6230	0.6256
Testis	0.6109	0.6129	0.6149	0.6193	0.6220
Esophagus mucosa	0.6068	0.6090	0.6107	0.6153	0.6180
Sun-exposed skin	0.6014	0.6037	0.6056	0.6101	0.6129
Spleen	0.5965	0.5988	0.6003	0.6056	0.6085
Visceral omentum	0.5875	0.5901	0.5912	0.5963	0.5989
Prostate	0.5738	0.5762	0.5788	0.5828	0.5856
Stomach	0.5632	0.5655	0.5677	0.5725	0.5753
Adrenal gland	0.5576	0.5605	0.5621	0.5673	0.5699
Uterus	0.5508	0.5534	0.5551	0.5602	0.5631
Omnibus	0.5475	0.5494	0.5516	0.5567	0.5597
Esophagus muscularis	0.5442	0.5469	0.5489	0.5535	0.5561
Hippocampus	0.5429	0.5450	0.5463	0.5516	0.5545
Cortex	0.5068	0.5083	0.5101	0.5164	0.5197
Ovary	0.4999	0.5028	0.5049	0.5103	0.5135
Atrial appendage	0.4976	0.5001	0.5017	0.5070	0.5099
Transverse colon	0.4930	0.4956	0.4981	0.5034	0.5067
Caudate	0.4914	0.4930	0.4946	0.5001	0.5033
Tibial nerve	0.4876	0.4900	0.4918	0.4971	0.5002
Putamen	0.4701	0.4721	0.4743	0.4793	0.4824
CD4 cell	0.4468	0.4487	0.4521	0.4578	0.4614
Breast	0.4426	0.4451	0.4477	0.4532	0.4564
Sigmoid colon	0.4397	0.4425	0.4446	0.4501	0.4533
Cerebellum	0.4366	0.4385	0.4405	0.4467	0.4503
CD8 cell	0.4198	0.4213	0.4249	0.4307	0.4345
Subcutaneous adipose	0.4170	0.4193	0.4219	0.4275	0.4309
Gastroesophageal junction	0.4150	0.4180	0.4202	0.4260	0.4294
Skeletal muscle	0.4116	0.4145	0.4169	0.4225	0.4258
Monocyte	0.4093	0.4111	0.4148	0.4202	0.4236
Aorta	0.3766	0.3786	0.3817	0.3869	0.3901
Left ventricle	0.3713	0.3742	0.3768	0.3826	0.3859
B cell	0.3669	0.3691	0.3730	0.3786	0.3822
Liver	0.3611	0.3640	0.3676	0.3730	0.3763
NK cell	0.3494	0.3508	0.3549	0.3605	0.3643
Pancreas	0.3241	0.3262	0.3300	0.3355	0.3390

Supplementary Table 15

Descriptive statistics of 38 gene regulatory networks. Rows are sorted by the "# of TF-TG edges" column.

Network	# of TF-TG edges	# of TGs	# of TFs
Stomach	96135	3727	450
Omnibus	96008	9317	571
Esophagus mucosa	95797	4891	419
Terminal ileum	95596	3700	408
Pancreas	95446	2941	481
Transverse colon	95446	3696	435
Lung	95361	4123	415
Breast	95355	3260	442
Left ventricle	95030	2546	424
Sun-exposed skin	94988	4144	448
Gastroesophageal junction	94838	2380	443
Atrial appendage	94749	2564	399
Vagina	94592	4325	431
Sigmoid colon	94549	2545	460
Esophagus muscularis	94408	2993	429
Visceral omentum	94334	3701	428
Tibial nerve	94133	3214	431
Subcutaneous adipose	94005	2718	434
Ovary	93822	3659	430
Liver	93706	3398	412
Spleen	93494	4358	397
Aorta	93408	3295	414
Uterus	93057	3569	453
Skeletal muscle	92832	2308	490
NK cell	92399	3105	376
Thyroid	92263	5369	450
CD8 cell	91925	3439	424
B cell	91728	3018	436
Prostate	91722	3165	474
CD4 cell	91569	3643	433
Monocyte	91401	3702	384
Testis	90804	4646	474
Cortex	90579	2972	333
Adrenal gland	90401	2758	434
Caudate	89754	2547	374
Cerebellum	89371	3301	336
Hippocampus	88948	2592	366
Putamen	88634	2380	365

Supplementary Table 16

Descriptive statistics of relative trans impact of TFs on TGs $\{v_{gt}\}$ over all TF-TG edges available in each of 38 regulatory networks. Rows are sorted by the "Median" column.

	Summary of v_{gt}				
Network	Min	Q1	Median	Q3	Max
Cortex	0.6946	0.7105	0.7315	0.7636	1.0000
Putamen	0.6797	0.6972	0.7201	0.7560	0.9944
Cerebellum	0.6817	0.6970	0.7170	0.7484	1.0000
Caudate	0.6704	0.6880	0.7113	0.7469	1.0000
Hippocampus	0.6678	0.6843	0.7063	0.7403	0.9977
Atrial appendage	0.6489	0.6674	0.6920	0.7314	1.0000
CD8 cell	0.6385	0.6553	0.6775	0.7125	1.0000
Subcutaneous adipose	0.6339	0.6520	0.6769	0.7160	1.0000
Omnibus	0.6278	0.6484	0.6748	0.7143	1.0000
Transverse colon	0.6353	0.6518	0.6740	0.7099	1.0000
Terminal ileum	0.6362	0.6522	0.6736	0.7082	1.0000
Adrenal gland	0.6296	0.6445	0.6650	0.6987	1.0000
CD4 cell	0.6266	0.6430	0.6647	0.6989	1.0000
Gastroesophageal junction	0.6233	0.6413	0.6645	0.7020	1.0000
Stomach	0.6276	0.6431	0.6640	0.6984	1.0000
Skeletal muscle	0.6187	0.6366	0.6609	0.6998	1.0000
Lung	0.6243	0.6393	0.6597	0.6927	1.0000
NK cell	0.6138	0.6324	0.6568	0.6949	1.0000
B cell	0.6101	0.6278	0.6515	0.6885	1.0000
Breast	0.6087	0.6263	0.6497	0.6870	1.0000
Spleen	0.6122	0.6279	0.6492	0.6831	1.0000
Monocyte	0.6088	0.6259	0.6488	0.6844	1.0000
Sigmoid colon	0.6079	0.6250	0.6482	0.6849	1.0000
Prostate	0.6047	0.6210	0.6427	0.6780	1.0000
Left ventricle	0.6006	0.6185	0.6418	0.6788	1.0000
Esophagus muscularis	0.5972	0.6157	0.6405	0.6797	1.0000
Uterus	0.5977	0.6142	0.6369	0.6728	1.0000
Sun-exposed skin	0.5984	0.6149	0.6368	0.6724	1.0000
Ovary	0.5938	0.6122	0.6368	0.6763	1.0000
Tibial nerve	0.5920	0.6104	0.6350	0.6740	1.0000
Liver	0.5867	0.6060	0.6312	0.6707	1.0000
Pancreas	0.5847	0.6046	0.6305	0.6722	1.0000
Vagina	0.5852	0.6010	0.6228	0.6581	1.0000
Visceral omentum	0.5814	0.5977	0.6198	0.6552	1.0000
Aorta	0.5646	0.5826	0.6067	0.6438	1.0000
Esophagus mucosa	0.5643	0.5809	0.6034	0.6400	1.0000
Thyroid	0.5621	0.5766	0.5966	0.6288	0.9587
Testis	0.5512	0.5675	0.5896	0.6252	1.0000

Supplementary Table 17

Descriptive statistics of cis eQTL databases used in this study. Rows are sorted by the "# of SNP-gene pairs" column.

Source of eQTL	# of SNP-gene pairs	# of genes	# of SNPs
Testis, GTEx (V7)	15,759,533	17,366	1,230,746
Prostate, GTEx (V7)	14,585,826	16,085	1,219,002
Terminal ileum, GTEx (V7)	14,546,128	16,041	1,215,324
Lung, GTEx (V7)	14,483,585	15,973	1,214,950
Sun-exposed skin, GTEx (V7)	14,404,353	15,869	1,213,959
Transverse colon, GTEx (V7)	14,402,835	15,875	1,215,596
Breast, GTEx (V7)	14,392,706	15,872	1,212,389
Vagina, GTEx (V7)	14,380,972	15,869	1,217,316
Thyroid, GTEx (V7)	14,336,449	15,810	1,214,645
Tibial nerve, GTEx (V7)	14,288,824	15,761	1,217,709
Cortex, GTEx (V7)	14,261,947	15,745	1,226,087
Hippocampus, GTEx (V7)	14,233,460	15,719	1,224,980
Stomach, GTEx (V7)	14,220,507	15,684	1,213,739
Caudate, GTEx (V7)	14,209,653	15,689	1,226,945
Sigmoid colon, GTEx (V7)	14,202,126	15,668	1,219,231
Cerebellum, GTEx (V7)	14,192,621	15,679	1,224,520
Visceral omentum, GTEx (V7)	14,151,283	15,611	1,210,453
Uterus, GTEx (V7)	14,144,127	15,599	1,208,996
Spleen, GTEx (V7)	14,088,656	15,568	1,203,063
Ovary, GTEx (V7)	14,083,563	15,543	1,211,776
Subcutaneous adipose, GTEx (V7)	14,074,828	15,528	1,213,026
Esophagus mucosa, GTEx (V7)	14,050,266	15,504	1,201,919
Putamen, GTEx (V7)	14,036,948	15,503	1,226,334
Gastroesophageal junction, GTEx (V7)	13,998,654	15,446	1,217,148
Esophagus muscularis, GTEx (V7)	13,940,389	15,390	1,211,409
Adrenal gland, GTEx (V7)	13,918,735	15,368	1,215,183
Aorta, GTEx (V7)	13,913,807	15,356	1,207,409
Pancreas, GTEx (V7)	13,707,801	15,139	1,200,808
Atrial appendage, GTEx (V7)	13,683,429	15,122	1,207,966
Liver, GTEx (V7)	13,481,989	14,908	1,190,839
Left ventricle, GTEx (V7)	13,120,834	14,507	1,195,904
Skeletal muscle, GTEx (V7)	13,077,954	14,458	1,193,409
CD8 cell, DICE (DB1)	12,467,472	18,258	944,172
CD4 cell, DICE (DB1)	12,464,676	18,258	943,869
NK cell, DICE (DB1)	12,379,934	18,258	937,629
Omnibus, eQTLGen	12,372,754	13,588	1,173,301
B cell, DICE (DB1)	12,344,429	18,258	934,954
Monocyte, DICE (DB1)	12,344,429	18,258	934,954

Supplementary Table 18

Descriptive statistics of relative cis impact of SNPs on genes $\{c_{jg}\}$ over all SNP-gene pairs available in cis eQTL databases. Rows are sorted by the "Median" column.

	Summary of $(c_{jg} - 1)$				
Source of eQTL	Min	Q1	Median	Q3	Max
Uterus, GTEx (V7)	0	0.033	0.069	0.118	0.935
Vagina, GTEx (V7)	0	0.032	0.067	0.115	0.932
Putamen, GTEx (V7)	0	0.032	0.067	0.114	0.923
Hippocampus, GTEx (V7)	0	0.031	0.066	0.114	0.912
CD8 cell, DICE (DB1)	0	0.024	0.065	0.118	0.774
CD4 cell, DICE (DB1)	0	0.023	0.064	0.118	0.771
Monocyte, DICE (DB1)	0	0.023	0.064	0.118	0.792
Terminal ileum, GTEx (V7)	0	0.030	0.064	0.109	0.935
NK cell, DICE (DB1)	0	0.023	0.064	0.117	0.762
Ovary, GTEx (V7)	0	0.030	0.063	0.108	0.931
B cell, DICE (DB1)	0	0.021	0.063	0.117	0.777
Cortex, GTEx (V7)	0	0.029	0.061	0.105	0.915
Prostate, GTEx (V7)	0	0.029	0.061	0.104	0.943
Spleen, GTEx (V7)	0	0.028	0.059	0.101	0.942
Caudate, GTEx (V7)	0	0.028	0.059	0.101	0.908
Cerebellum, GTEx (V7)	0	0.027	0.058	0.101	0.952
Liver, GTEx (V7)	0	0.027	0.057	0.098	0.921
Adrenal gland, GTEx (V7)	0	0.026	0.054	0.093	0.945
Sigmoid colon, GTEx (V7)	0	0.024	0.050	0.087	0.942
Gastroesophageal junction, GTEx (V7)	0	0.023	0.049	0.085	0.948
Pancreas, GTEx (V7)	0	0.023	0.049	0.084	0.939
Testis, GTEx (V7)	0	0.023	0.049	0.084	0.927
Stomach, GTEx (V7)	0	0.022	0.046	0.080	0.942
Transverse colon, GTEx (V7)	0	0.022	0.046	0.079	0.943
Breast, GTEx (V7)	0	0.021	0.045	0.078	0.949
Atrial appendage, GTEx (V7)	0	0.021	0.045	0.077	0.943
Aorta, GTEx (V7)	0	0.021	0.045	0.077	0.948
Left ventricle, GTEx (V7)	0	0.021	0.044	0.076	0.943
Visceral omentum, GTEx (V7)	0	0.019	0.041	0.071	0.939
Esophagus muscularis, GTEx (V7)	0	0.019	0.040	0.070	0.955
Tibial nerve, GTEx (V7)	0	0.019	0.040	0.069	0.964
Esophagus mucosa, GTEx (V7)	0	0.018	0.039	0.068	0.938
Subcutaneous adipose, GTEx (V7)	0	0.018	0.038	0.066	0.952
Thyroid, GTEx (V7)	0	0.018	0.038	0.065	0.955
Lung, GTEx (V7)	0	0.018	0.038	0.065	0.928
Sun-exposed skin, GTEx (V7)	0	0.017	0.037	0.064	0.948
Skeletal muscle, GTEx (V7)	0	0.016	0.034	0.058	0.918
Omnibus, eQTLGen	0	0.003	0.007	0.014	0.855

Supplementary Table 19

RSS-NET hyper-parameter grids used in the present study. For all traits and networks, the grid for θ is $(0 : 0.25 : 1)$ and the grid for ρ is $(0 : 0.2 : 0.8)$. For a given trait, analyses of all 38 networks and the near-gene control use the same hyper-parameter grid. Here $(j : i : k)$ denotes a regularly-spaced vector that starts at j, uses i as the increment, and (roughly) stops at k. Trait abbreviations are defined in **Supplementary Table 2**.

Trait	θ_0	η
LOAD	$(-5.25 : 0.1 : -4.75)$	0.6
NEU	$(-4.4 : 0.05 : -4)$	0.3
SCZ	$(-2.175 : 0.025 : -2.05)$	0.3
BMI	$(-4.1 : 0.025 : -3.95)$	0.3
HEIGHT	$(-2.05 : 0.05 : -1.95)$	0.3
WAIST	$(-3 : 0.025 : -2.95)$	0.3
CD	$(-3 : 0.05 : -2.9)$	0.3
IBD	$(-3 : 0.05 : -2.8)$	0.3
RA	$(-3.25 : 0.025 : -3.125)$	0.3
UC	$(-3.05 : 0.025 : -2.95)$	0.3
BC	$(-4.25 : 0.05 : -4)$	0.2
AF	$(-4.5 : 0.25 : -4)$	$0.2, 0.3$
CAD	$(-4 : 0.025 : -3.85)$	0.3
HDL	$(-3.575 : 0.025 : -3.45)$	0.3
HR	$(-4.45 : 0.05 : -4.25)$	$0.3, 0.4$
LDL	$(-3.75 : 0.05 : -3.5)$	0.3
MI	$(-4.3 : 0.025 : -4.1)$	0.3
T2D	$(-4.65 : 0.05 : -4.55)$	$0.4, 0.5, 0.6$

References

1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature*, **526**, 68–74.

Amberger, J., Bocchini, C., Scott, A. and Hamosh, A. (2019) OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Research*, **47**, D1038–D1043.

Bult, C. J., Blake, J. A., Smith, C. L., Kadin, J. A. and Richardson, J. E. (2019) Mouse genome database (MGD) 2019. *Nucleic Acids Research*, **47**, D801–D806.

Buniello, A., MacArthur, J., Cerezo, M., Harris, L., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E. et al. (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, **47**, D1005–D1012.

Christophersen, I. E., Rienstra, M., Roselli, C., Yin, X., Geelhoed, B., Barnard, J., Lin, H., Arking, D. E., Smith, A. V., Albert, C. M. et al. (2017) Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nature Genetics*, **49**, 946.

Den Hoed, M., Eijgelsheim, M., Esko, T., Brundel, B. J., Peal, D. S., Evans, D. M., Nolte, I. M., Segrè, A. V., Holm, H., Handsaker, R. E. et al. (2013) Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nature Genetics*, **45**, 621–631.

Duren, Z., Chen, X., Jiang, R., Wang, Y. and Wong, W. H. (2017) Modeling gene regulation from paired expression and chromatin accessibility data. *Proceedings of the National Academy of Sciences*, **114**, E4914–E4923.

ENCODE Project Consortium (2012) An integrated encyclopedia of dna elements in the human genome. *Nature*, **489**, 57–74.

Finucane, H. K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., Anttila, V., Xu, H., Zang, C., Farh, K. et al. (2015) Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*, **47**, 1228–1235.

Freund, M. K., Burch, K. S., Shi, H., Mancuso, N., Kichaev, G., Garske, K. M., Pan, D. Z., Miao, Z., Mohlke, K. L., Laakso, M. et al. (2018) Phenotype-specific enrichment of Mendelian disorder genes near GWAS regions across 62 complex traits. *The American Journal of Human Genetics*, **103**, 535–552.

Gazal, S., Finucane, H. K., Furlotte, N. A., Loh, P.-R., Palamara, P. F., Liu, X., Schoech, A., Bulik-Sullivan, B., Neale, B. M., Gusev, A. et al. (2017) Linkage disequilibrium–dependent architecture of human complex traits shows action of negative selection. *Nature Genetics*, **49**, 1421–1427.

Global Lipids Genetics Consortium (2013) Discovery and refinement of loci associated with lipids levels. *Nature Genetics*, **45**, 1274–1283.

Lambert, J.-C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., Jun, G., DeStefano, A. L., Bis, J. C., Beecham, G. W. et al. (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics*, **45**, 1452–1458.

Lamparter, D., Marbach, D., Rueedi, R., Kutalik, Z. and Bergmann, S. (2016) Fast and rigorous computation of gene and pathway scores from SNP-based summary statistics. *PLoS Computational Biology*, **12**, e1004714.

Liu, J. Z., van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., Ripke, S., Lee, J. C., Jostins, L., Shah, T. et al. (2015) Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics*, **47**, 979–986.

Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., Powell, C., Vedantam, S., Buchkovich, M. L., Yang, J. et al. (2015) Genetic studies of body mass index yield new insights for obesity biology. *Nature*, **518**, 197–206.

Luo, Y., Hitz, B. C., Gabdank, I., Hilton, J. A., Kagda, M. S., Lam, B., Myers, Z., Sud, P., Jou, J., Lin, K. et al. (2020) New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. *Nucleic Acids Research*, **48**, D882–D889.

Morris, A. P., Voight, B. F., Teslovich, T. M., Ferreira, T., Segre, A. V., Steinthorsdottir, V., Strawbridge, R. J., Khan, H., Grallert, H., Mahajan, A. et al. (2012) Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature Genetics*, **44**, 981–990.

Nikpay, M., Goel, A., Won, H.-H., Hall, L. M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C. P., Hopewell, J. C. et al. (2015) A comprehensive 1000 genomes–based genome-wide association meta-analysis of coronary artery disease. *Nature Genetics*, **47**, 1121–1130.

Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S. et al. (2014) Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*, **506**, 376–381.

Okbay, A., Baselmans, B., De Neve, J., Turley, P., Nivard, M., Fontana, M., Meddens, S., Linnér, R., Rietveld, C., Derringer, J. et al. (2016) Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*, **48**, 624–633.

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K.-H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H. et al. (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, **511**, 421–427.

Shungin, D., Winkler, T. W., Croteau-Chonka, D. C., Ferreira, T., Locke, A. E., Mägi, R., Strawbridge, R. J., Pers, T. H., Fischer, K., Justice, A. E. et al. (2015) New genetic loci link adipose and insulin biology to body fat distribution. *Nature*, **518**, 187–196.

Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J. P., Ripatti, S., Chasman, D. I., Willer, C. J. et al. (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, **466**, 707–713.

Wang, Y., Zhang, S., Li, F., Zhou, Y., Zhang, Y., Wang, Z., Zhang, R., Zhu, J., Ren, Y., Tan, Y. et al. (2020) Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics. *Nucleic acids research*, **48**, D1031–D1041.

Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, **447**, 661–678.

Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., Chu, A. Y., Estrada, K., Luan, J., Kutalik, Z. et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, **46**, 1173–1186.

Zhu, X. and Stephens, M. (2017) Bayesian large-scale multiple regression with summary statistics from genome-wide association studies. *Annals of Applied Statistics*, **11**, 1561–1592.

— (2018) Large-scale genome-wide enrichment analyses identify new trait-associated genes and pathways across 31 human phenotypes. *Nature Communications*, **9**, 4361.