**Supplementary Note 1.**

Next we extend the logistic regression model (1) to the functional logistic regression for modeling main and interaction effects. We begin with reviewing some functional data analysis results. By Karhunen-Loeve expansion [15], we have

$$x_i(t) = \sum_{j=1}^{\infty} \xi_{ij}\phi_j(t) \text{ and}$$

$$x_i(s) = \sum_{k=1}^{\infty} \eta_{ik}\psi_k(s)\text{'} \qquad (3)$$

where $\phi_j(t)$ and $\psi_k(s)$ are sequences of orthonormal basis functions. The functional principal scores are defined by

$$\xi_{ij} = \int_T x_i(t)\phi_j(t)dt \text{ and}$$

$$\eta_{ik} = \int_S x_i(s)\psi_k(s)ds\,.$$

We extend the traditional logistic regression model (2) to the following functional logistic regression model for gene-gene interaction analysis:

$$\log\frac{\pi_i}{1-\pi_i} = \alpha_0 + \int_T \alpha(t)x_i(t)dt + \int_S \beta(s)x_i(s)ds + \int_T\int_S \gamma(t,s)x_i(t)x_i(s)dtds, \qquad (4)$$

where $\alpha(t)$ and $\beta(s)$ are the putative genetic additive effects of two SNPs located at the genomic positions $t$ and $s$, respectively, $\gamma(t,s)$ is the putative interaction effect between two SNPs located at the genomic positions $t$ and $s$. Thus, $\pi_i$ is given by

$$\pi_i = \frac{e^{\alpha_0 + \int_T \alpha(t)x_i(t)dt + \int_S \beta(s)x_i(s)ds + \int_T \int_S \gamma(t,s)x_i(t)x_i(s)dtds}}{1 + e^{\alpha_0 + \int_T \alpha(t)x_i(t)dt + \int_S \beta(s)x_i(s)ds + \int_T \int_S \gamma(t,s)x_i(t)x_i(s)dtds}}.$$

Substituting equation (3) into equation (4), we obtain

$$\begin{aligned}\log\frac{\pi_i}{1-\pi_i} &= \alpha_0 + \int_T \alpha(t)\sum_{j=1}^{\infty}\xi_{ij}\phi_j(t)dt + \int_S \beta(s)\sum_{j=1}^{\infty}\eta_{ik}\psi_k(s)ds \\ &+ \int_T\int_S \gamma(t,s)[\sum_{j=1}^{\infty}\xi_{ij}\phi_j(t)][\sum_{k=1}^{\infty}\eta_{ik}\psi_k(s)]dtds + \varepsilon_i \qquad (5)\\ &= \alpha_0 + \sum_{j=1}^{\infty}\xi_{ij}\int_T \alpha(t)\phi_j(t)dt + \sum_{j=1}^{\infty}\eta_{ik}\int_S \beta(s)\Psi_k(s)ds + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\xi_{ij}\eta_{ik}\int_T\int_S \gamma(t,s)\phi_j(t)\Psi_k(s)dtds\end{aligned}$$

Let $\alpha_j = \int_T \alpha(t)\phi_j(t)dt$, $\beta_k = \int_S \beta(s)\psi_k(s)ds$, and $\gamma_{jk} = \int_T\int_S \gamma(t,s)\phi_j(t)\psi_k(s)dtds$. Then, equation (5) can be reduced to

$$\log\frac{\pi_i}{1-\pi_i} = \alpha_0 + \sum_{j=1}^{\infty}\xi_{ij}\alpha_j + \sum_{k=1}^{\infty}\eta_{ik}\beta_k + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\xi_{ij}\eta_{ik}\gamma_{jk}. \qquad (6)$$

Let $\alpha = \begin{bmatrix}\alpha_1 \\ \vdots \\ \alpha_J\end{bmatrix}, \beta = \begin{bmatrix}\beta_1 \\ \vdots \\ \beta_K\end{bmatrix}, \gamma = \begin{bmatrix}\gamma_{11} \\ \gamma_{12} \\ \vdots \\ \gamma_{JK}\end{bmatrix}, b = \begin{bmatrix}\alpha_0 \\ \alpha \\ \beta \\ \gamma\end{bmatrix}, \xi_i = \begin{bmatrix}\xi_{i1} \\ \vdots \\ \xi_{iJ}\end{bmatrix}, \eta_i = \begin{bmatrix}\eta_{i1} \\ \vdots \\ \eta_{iK}\end{bmatrix}, \Gamma_i = \begin{bmatrix}\xi_{i1}\eta_{i1} \\ \vdots \\ \xi_{iJ}\eta_{iK}\end{bmatrix}$

Then, we have

$$\begin{aligned}\log\frac{\pi_i}{1-\pi_i} &= \alpha_0 + \xi_i^T\alpha + \eta_i^T\beta + \Gamma_i^T\gamma \\ &= \begin{bmatrix}1 & \xi_i^T & \eta_i^T & \Gamma_i^T\end{bmatrix}\begin{bmatrix}\alpha_0 \\ \alpha \\ \beta \\ \gamma \\ \end{bmatrix} \qquad (7)\\ &= W_i^T b\end{aligned}$$

where $W_i = \begin{bmatrix} 1 \\ \xi_i \\ \eta_i \\ \Gamma_i \end{bmatrix}$.

**Supplementary Note 2.**

We consider two loci: G and H. Assume that the codes $G = 1(G = 0)$ and $H = 1(H = 0)$ denote whether an individual is a carrier (non-carrier) of the susceptible genotypes at the loci G and H, respectively. Let D denote disease status where $D = 1(D = 0)$ indicates an affected (unaffected) individual. The odds-ratio associated with G for nonsusceptible genotype at the locus H $(H = 0)$ is defined as

$$OR_G = \frac{P(D = 1 \mid G = 1, H = 0) / P(D = 0 \mid G = 1, H = 0)}{P(D = 1 \mid G = 0, H = 0) / P(D = 0 \mid G = 0, H = 0)}.$$

Similarly, the odds-ratio associated with H for nonsusceptible genotype at the locus G $(G = 0)$ is defined as

$$OR_H = \frac{P(D = 1 \mid G = 0, H = 1) / P(D = 0 \mid G = 0, H = 1)}{P(D = 1 \mid G = 0, H = 0) / P(D = 0 \mid G = 0, H = 0)}.$$

The odds-ratio associated with susceptibility at G and H compared to the baseline category $G = 0$ and $H = 0$ is then computed as

$$OR_{GH} = \frac{P(D = 1 \mid G = 1, H = 1) / P(D = 0 \mid G = 1, H = 1)}{P(D = 1 \mid G = 0, H = 0) / P(D = 0 \mid G = 0, H = 0)}.$$

Define a multiplicative interaction measure between two loci G and H as

$$I_{GH} = \log \frac{OR_{GH}}{OR_G OR_H}. \qquad \text{(A1)}$$

Equation (A1) can be reduced to

$$I_{GH} = \log \frac{P(G=1,H=1 \mid D=1)P(G=0,H=0 \mid D=1)}{P(G=1,H=0 \mid D=1)P(G=0,H=1 \mid D=1)} + \log \frac{P(G=1,H=0 \mid D=0)P(G=0,H=1 \mid D=0)}{P(G=1,H=1 \mid D=0)P((G=0,H=0 \mid D=0)}. \quad \text{(A2)}$$

Assume that we observe the polymorphism at the loci $G$ and $H$ in both cases and controls. In other words, we have $P(G=1,H=0 \mid D=1) \neq 0$, $P(G=0,H=1 \mid D=1) \neq 0$, $P(G=1,H=0 \mid D=0) \neq 0$ and $P(G=0,H=1 \mid D=0) \neq 0$. We assume that the rare variants are risk variants. In this case we have $P(G=0,H=0 \mid D=1) \neq 0$ and $P(G=0,H=0 \mid D=0) \neq 0$. Consider three scenarios:

(1) $P(G=1,H=1 \mid D=1) = 0$ and $P(G=1,H=1 \mid D=0) \neq 0$,

(2) $P(G=1,H=1 \mid D=1) \neq 0$ and $P(G=1,H=1 \mid D=0) = 0$, and

(3) $P(G=1,H=1 \mid D=1) = 0, P(G=1,H=1 \mid D=0) = 0$.

For scenario (1) it follows from equation (A2) that we have $I_{GH} = -\infty$. In other words, the risk alleles do not jointly appear in the cases, but are jointly presented in the controls. They work together to protect individuals from disease. For scenario (2) it again follows from equation (A2) that $I_{GH} = \infty$. The risk alleles at two loci are jointly present in cases, but never appeared in the controls. They strongly interact with each other to cause disease. For scenario (3) the risk alleles at two loci are not present in both cases and controls. They do not show interaction. Therefore, we have $I_{GH} = 0$.

**Collapsing Method**

For the collapsing method, we consider two scenarios to define indicator variables:

(1) common variants

$$x_i = \begin{cases} 1 & \text{At least one variant within a gene with MAF} \geq 0.01 \\ 0 & \text{otherwise} \end{cases}$$

(2) rare variants

$$x_i = \begin{cases} 1 & \text{At least one variant within a gene with MAF} \leq 0.01 \\ 0 & \text{otherwise} \end{cases}$$

Table S1. Type 1 error rates of functional logistic regression for testing interaction between two genes with rare variants

| Sample Size | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0498 | 0.0101 | 0.0010 | 0.0524 | 0.0093 | 0.0010 | 0.0521 | 0.0100 | 0.0011 |
| 1500 | 0.0501 | 0.0099 | 0.0010 | 0.0500 | 0.0103 | 0.0010 | 0.0514 | 0.0103 | 0.0010 |
| 2000 | 0.0498 | 0.0098 | 0.0010 | 0.0493 | 0.0100 | 0.0010 | 0.0522 | 0.0103 | 0.0010 |
| 2500 | 0.0502 | 0.0099 | 0.0010 | 0.0498 | 0.0099 | 0.0010 | 0.0499 | 0.0102 | 0.0010 |
| 3000 | 0.0499 | 0.0099 | 0.0010 | 0.0507 | 0.0101 | 0.0010 | 0.0515 | 0.0102 | 0.0010 |

Table S2. Type 1 error rates of functional logistic regression for testing interaction between two genes with all variants

| | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0474 | 0.0103 | 0.0010 | 0.0524 | 0.0092 | 0.0010 | 0.0520 | 0.0101 | 0.0011 |
| 1500 | 0.0512 | 0.0102 | 0.0010 | 0.0487 | 0.0106 | 0.0010 | 0.0523 | 0.0106 | 0.0010 |
| 2000 | 0.0518 | 0.0099 | 0.0010 | 0.0493 | 0.0100 | 0.0010 | 0.0553 | 0.0105 | 0.0011 |
| 2500 | 0.0501 | 0.0098 | 0.0010 | 0.0530 | 0.0098 | 0.0010 | 0.0526 | 0.0105 | 0.0010 |
| 3000 | 0.0488 | 0.0102 | 0.0010 | 0.0520 | 0.0103 | 0.0010 | 0.0526 | 0.0106 | 0.0010 |

Table S3. Type 1 error rates of collapsing method for testing interaction between two genes with common variants

| Sample Size | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0739 | 0.0166 | 0.0014 | 0.0828 | 0.0148 | 0.0016 | 0.0815 | 0.0180 | 0.0018 |
| 1500 | 0.0725 | 0.0151 | 0.0014 | 0.0779 | 0.0130 | 0.0016 | 0.0880 | 0.0152 | 0.0017 |
| 2000 | 0.0729 | 0.0166 | 0.0015 | 0.0703 | 0.0133 | 0.0015 | 0.0808 | 0.0178 | 0.0018 |
| 2500 | 0.0733 | 0.0145 | 0.0013 | 0.0703 | 0.0147 | 0.0016 | 0.0896 | 0.0156 | 0.0019 |
| 3000 | 0.0612 | 0.0139 | 0.0016 | 0.0680 | 0.0148 | 0.0013 | 0.0849 | 0.0175 | 0.0018 |

Table S4. Type 1 error rates of pair-wise logistic regression for testing interaction between two genes with common variants

| Sample Size | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0360 | 0.0084 | 0.0009 | 0.0411 | 0.0083 | 0.0008 | 0.0466 | 0.0079 | 0.0010 |
| 1500 | 0.0364 | 0.0083 | 0.0008 | 0.0389 | 0.0073 | 0.0009 | 0.0385 | 0.0082 | 0.0008 |
| 2000 | 0.0400 | 0.0088 | 0.0009 | 0.0411 | 0.0088 | 0.0009 | 0.0402 | 0.0093 | 0.0008 |
| 2500 | 0.0374 | 0.0080 | 0.0009 | 0.0421 | 0.0074 | 0.0009 | 0.0452 | 0.0082 | 0.0009 |
| 3000 | 0.0384 | 0.0079 | 0.0008 | 0.0371 | 0.0076 | 0.0007 | 0.0411 | 0.0087 | 0.0009 |

Table S5. Type 1 error rates of PCA logistic regression for testing interaction between two genes with common variants

| Sample Size | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0477 | 0.0091 | 0.0010 | 0.0485 | 0.0092 | 0.0009 | 0.0512 | 0.0105 | 0.0010 |
| 1500 | 0.0504 | 0.0093 | 0.0010 | 0.0476 | 0.0096 | 0.0009 | 0.0499 | 0.0105 | 0.0010 |
| 2000 | 0.0484 | 0.0096 | 0.0010 | 0.0508 | 0.0103 | 0.0011 | 0.0473 | 0.0101 | 0.0009 |
| 2500 | 0.0481 | 0.0091 | 0.0010 | 0.0460 | 0.0097 | 0.0009 | 0.0477 | 0.0100 | 0.0010 |
| 3000 | 0.0470 | 0.0095 | 0.0010 | 0.0501 | 0.0096 | 0.0009 | 0.0504 | 0.0099 | 0.0010 |

Table S6. Type 1 error rates of collapsing method for testing interaction between two genes with rare variants

| | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0738 | 0.0160 | 0.0016 | 0.0821 | 0.0160 | 0.0016 | 0.0911 | 0.0180 | 0.0017 |
| 1500 | 0.0743 | 0.0162 | 0.0016 | 0.0773 | 0.0156 | 0.0016 | 0.0890 | 0.0174 | 0.0017 |
| 2000 | 0.0740 | 0.0166 | 0.0016 | 0.0772 | 0.0157 | 0.0016 | 0.0895 | 0.0183 | 0.0018 |
| 2500 | 0.0740 | 0.0163 | 0.0016 | 0.0781 | 0.0156 | 0.0016 | 0.0902 | 0.0177 | 0.0018 |
| 3000 | 0.0741 | 0.0161 | 0.0016 | 0.0764 | 0.0157 | 0.0015 | 0.0899 | 0.0182 | 0.0018 |

Table S7. Type 1 error rates of pair-wise logistic regression for testing interaction between two genes with rare variants

| | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0343 | 0.0077 | 0.0008 | 0.0379 | 0.0071 | 0.0008 | 0.0399 | 0.0072 | 0.0009 |
| 1500 | 0.0342 | 0.0079 | 0.0008 | 0.0374 | 0.0076 | 0.0008 | 0.0390 | 0.0084 | 0.0008 |
| 2000 | 0.0349 | 0.0077 | 0.0008 | 0.0376 | 0.0078 | 0.0008 | 0.0403 | 0.0082 | 0.0008 |
| 2500 | 0.0355 | 0.0078 | 0.0008 | 0.0371 | 0.0076 | 0.0008 | 0.0399 | 0.0078 | 0.0008 |
| 3000 | 0.0357 | 0.0078 | 0.0008 | 0.0379 | 0.0076 | 0.0007 | 0.0403 | 0.0079 | 0.0008 |

Table S8. Type 1 error rates of PCA logistic regression for testing interaction between two genes with rare variants

| | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0472 | 0.0095 | 0.0010 | 0.0475 | 0.0089 | 0.0009 | 0.0522 | 0.0099 | 0.0010 |
| 1500 | 0.0470 | 0.0096 | 0.0010 | 0.0450 | 0.0096 | 0.0009 | 0.0510 | 0.0103 | 0.0010 |
| 2000 | 0.0480 | 0.0094 | 0.0010 | 0.0483 | 0.0098 | 0.0010 | 0.0478 | 0.0100 | 0.0009 |
| 2500 | 0.0471 | 0.0096 | 0.0010 | 0.0476 | 0.0097 | 0.0009 | 0.0479 | 0.0096 | 0.0010 |
| 3000 | 0.0467 | 0.0095 | 0.0010 | 0.0479 | 0.0095 | 0.0009 | 0.0483 | 0.0098 | 0.0010 |

Table S9. Type 1 error rates of collapsing method for testing interaction between two genes with all variants

| | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0724 | 0.0153 | 0.0016 | 0.0766 | 0.0151 | 0.0016 | 0.0869 | 0.0173 | 0.0016 |
| 1500 | 0.0745 | 0.0158 | 0.0017 | 0.0759 | 0.0154 | 0.0015 | 0.0895 | 0.0170 | 0.0016 |
| 2000 | 0.0721 | 0.0154 | 0.0015 | 0.0747 | 0.0153 | 0.0017 | 0.0856 | 0.0173 | 0.0018 |
| 2500 | 0.0727 | 0.0154 | 0.0015 | 0.0785 | 0.0153 | 0.0017 | 0.0908 | 0.0164 | 0.0019 |
| 3000 | 0.0741 | 0.0151 | 0.0016 | 0.0730 | 0.0150 | 0.0014 | 0.0854 | 0.0187 | 0.0019 |

Table S10. Type 1 error rates of pair-wise logistic regression for testing interaction between two genes with all variants

| | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0363 | 0.0079 | 0.0008 | 0.0387 | 0.0074 | 0.0008 | 0.0418 | 0.0072 | 0.0009 |
| 1500 | 0.0327 | 0.0080 | 0.0008 | 0.0384 | 0.0075 | 0.0008 | 0.0401 | 0.0086 | 0.0008 |
| 2000 | 0.0359 | 0.0080 | 0.0008 | 0.0395 | 0.0079 | 0.0008 | 0.0418 | 0.0087 | 0.0009 |
| 2500 | 0.0366 | 0.0081 | 0.0008 | 0.0363 | 0.0077 | 0.0009 | 0.0422 | 0.0082 | 0.0009 |
| 3000 | 0.0371 | 0.0082 | 0.0008 | 0.0416 | 0.0078 | 0.0007 | 0.0422 | 0.0080 | 0.0009 |

Table S11. Type 1 error rates of PCA logistic regression for testing interaction between two genes with all variants

| | $\beta_G=\beta_H=0$ | | | $\beta_G=2,\beta_H=0$ | | | $\beta_G=\beta_H=2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 | 0.05 | 0.01 | 0.001 |
| 1000 | 0.0498 | 0.0099 | 0.0010 | 0.0482 | 0.0087 | 0.0009 | 0.0524 | 0.0100 | 0.0010 |
| 1500 | 0.0454 | 0.0094 | 0.0010 | 0.0467 | 0.0099 | 0.0009 | 0.0525 | 0.0102 | 0.0010 |
| 2000 | 0.0491 | 0.0093 | 0.0010 | 0.0469 | 0.0098 | 0.0010 | 0.0477 | 0.0104 | 0.0009 |
| 2500 | 0.0466 | 0.0097 | 0.0010 | 0.0491 | 0.0096 | 0.0009 | 0.0460 | 0.0094 | 0.0010 |
| 3000 | 0.0484 | 0.0096 | 0.0010 | 0.0488 | 0.0099 | 0.0009 | 0.0491 | 0.0099 | 0.0010 |

Table S12. Log odds in disease interaction models.

| Models | First locus | Second locus | | |
|---|---|---|---|---|
| | | $H_1H_1$ | $H_1H_2$ | $H_2H_2$ |
| Dom $\cup$ Dom | $G_1G_1$ | $\alpha+\beta_G+\beta_H+\beta_{GH}$ | $\alpha+\beta_G+\beta_H+\beta_{GH}$ | $\alpha+\beta_G$ |
| | $G_1G_2$ | $\alpha+\beta_G+\beta_H+\beta_{GH}$ | $\alpha+\beta_G+\beta_H+\beta_{GH}$ | $\alpha+\beta_G$ |
| | $G_2G_2$ | $\alpha+\beta_H$ | $\alpha+\beta_H$ | $\alpha$ |
| Add $\cup$ Add | $G_1G_1$ | $\alpha+2\beta_G+2\beta_H+4\beta_{GH}$ | $\alpha+2\beta_G+\beta_H+2\beta_{GH}$ | $\alpha+2\beta_G$ |
| | $G_1G_2$ | $\alpha+\beta_G+2\beta_H+2\beta_{GH}$ | $\alpha+\beta_G+\beta_H+\beta_{GH}$ | $\alpha+\beta_G$ |
| | $G_2G_2$ | $\alpha+2\beta_H$ | $\alpha+\beta_H$ | $\alpha$ |
| Rec $\cup$ Rec | $G_1G_1$ | $\alpha+\beta_G+\beta_H+\beta_{GH}$ | $\alpha+\beta_G$ | $\alpha+\beta_G$ |
| | $G_1G_2$ | $\alpha+\beta_H$ | $\alpha$ | $\alpha$ |
| | $G_2G_2$ | $\alpha+\beta_H$ | $\alpha$ | $\alpha$ |

Table S14. P-values of 25 pairs of SNPs between genes TMX4 and C200rf7.

| Gene 1 | | Gene 2 | | P-value |
|---|---|---|---|---|
| TMX4 | | C20orf7 | | 1.09E-18 |
| SNP1 | | SNP2 | | |
| Position | MAF | Position | MAF | Pair-wise Test |
| 7990831 | 0.0671 | 13767943 | 0.0181 | 6.25E-05 |
| 7990831 | 0.0671 | 13763660 | 0.0189 | 8.32E-05 |
| 7980390 | 0.2991 | 13763660 | 0.0189 | 1.91E-03 |
| 7980390 | 0.2991 | 13767943 | 0.0181 | 2.21E-03 |
| 7967980 | 0.1566 | 13763660 | 0.0189 | 8.29E-03 |
| 7967980 | 0.1566 | 13767943 | 0.0181 | 9.64E-03 |
| 7980390 | 0.2991 | 13763553 | 0.0030 | 2.64E-02 |
| 7990831 | 0.0671 | 13765897 | 0.0017 | 3.73E-02 |
| 7967980 | 0.1566 | 13763590 | 0.0013 | 5.37E-02 |
| 7964476 | 0.0581 | 13767943 | 0.0181 | 1.48E-01 |
| 7964476 | 0.0581 | 13763660 | 0.0189 | 1.59E-01 |
| 7967980 | 0.1566 | 13765897 | 0.0017 | 1.99E-01 |
| 7980390 | 0.2991 | 13765897 | 0.0017 | 2.00E-01 |
| 7964476 | 0.0581 | 13763553 | 0.0030 | 2.43E-01 |
| 7990831 | 0.0671 | 13763553 | 0.0030 | 2.83E-01 |
| 7967980 | 0.1566 | 13763553 | 0.0030 | 2.85E-01 |
| 7980390 | 0.2991 | 13763590 | 0.0013 | 2.98E-01 |
| 7964476 | 0.0581 | 13765897 | 0.0017 | 4.35E-01 |
| 7964476 | 0.0581 | 13763631 | 0.0013 | 9.97E-01 |
| 7990831 | 0.0671 | 13765704 | 0.0013 | 9.97E-01 |
| 7967980 | 0.1566 | 13765859 | 0.0013 | 9.98E-01 |
| 7967980 | 0.1566 | 13765704 | 0.0013 | 9.98E-01 |
| 7967980 | 0.1566 | 13765721 | 0.0009 | 9.98E-01 |
| 7980390 | 0.2991 | 13765859 | 0.0013 | 9.99E-01 |
| 7980390 | 0.2991 | 13765704 | 0.0013 | 9.99E-01 |
| 7980390 | 0.2991 | 13765721 | 0.0009 | 9.99E-01 |

Table S15. P-values of 13 pairs of SNPs between genes TMX4 and C200rf7 by extended logistic regression analysis.

| Gene 1 | | Gene 2 | | P-value |
|---|---|---|---|---|
| TMX4 | | C20orf7 | | 1.09E-18 |
| SNP1 | | SNP2 | | |
| Position | MAF | Position | MAF | Pair-wise Test |
| 7963021 | 0.0004 | 13767943 | 0.01807 | < 1.0E-20 |
| 7990808 | 0.0004 | 13763660 | 0.01893 | < 1.0E-20 |
| 7990831 | 0.0671 | 13765721 | 0.00086 | < 1.0E-20 |
| 7967980 | 0.1566 | 13763192 | 0.00129 | < 1.0E-20 |
| 7967980 | 0.1566 | 13763631 | 0.00129 | < 1.0E-20 |
| 7980390 | 0.2991 | 13763601 | 0.00086 | < 1.0E-20 |
| 7980478 | 0.0004 | 13782229 | 0.00043 | < 1.0E-20 |
| 7976739 | 0.0004 | 13763192 | 0.00129 | < 1.0E-20 |
| 7980478 | 0.0004 | 13763553 | 0.00301 | < 1.0E-20 |
| 7964476 | 0.0581 | 13763192 | 0.00129 | 1 |
| 7967980 | 0.1566 | 13763601 | 0.00086 | 1 |
| 7980390 | 0.2991 | 13763631 | 0.00129 | 1 |
| 7963021 | 0.0004 | 13763314 | 0.00043 | 1 |

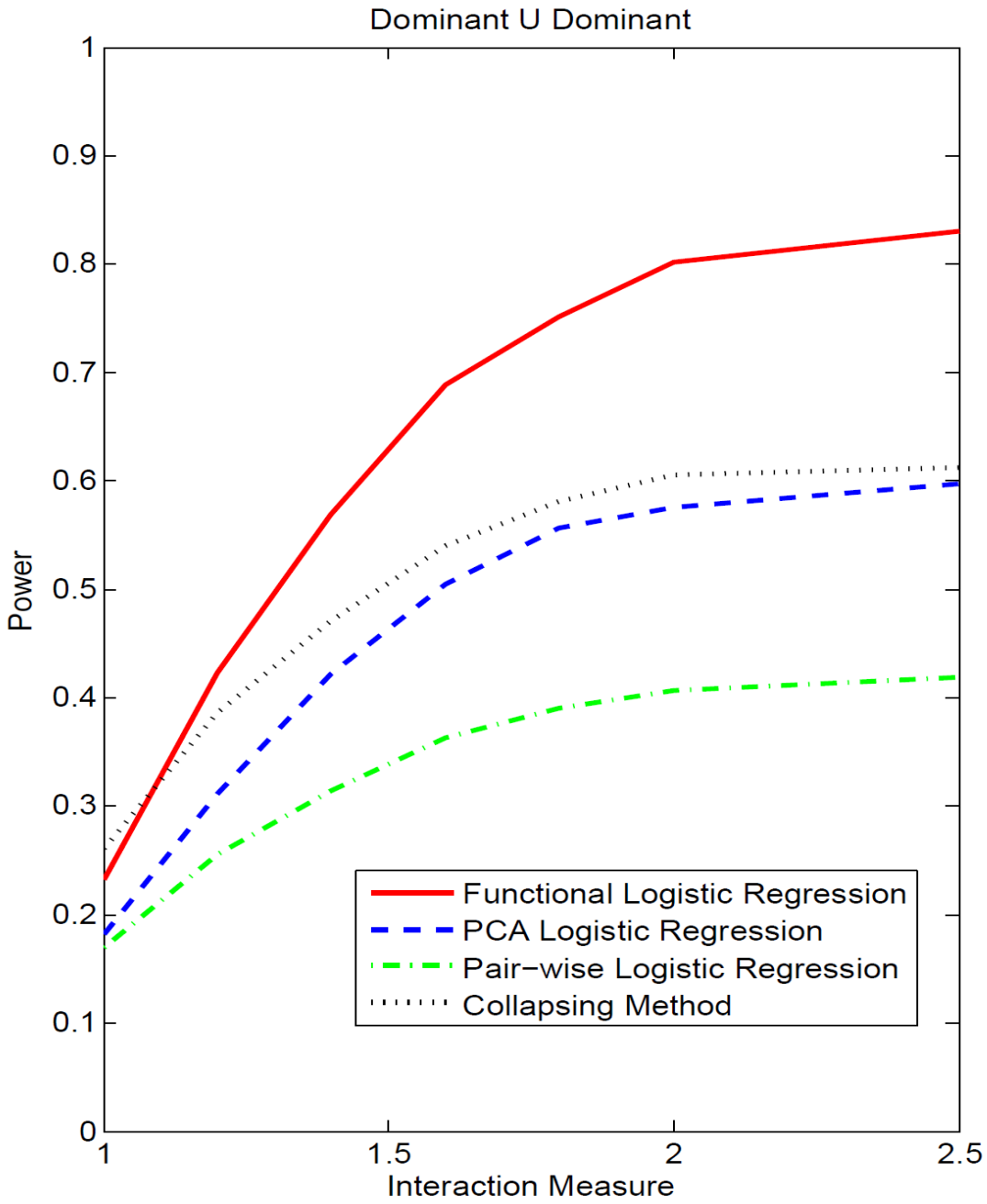


Figure S1. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of rare variants as a function of an interaction measure at the significance level $\alpha = 0.05$ under the dominant $\cup$ dominant model, assuming 2,000 cases and 2,000 controls, and 10% of risk variants.

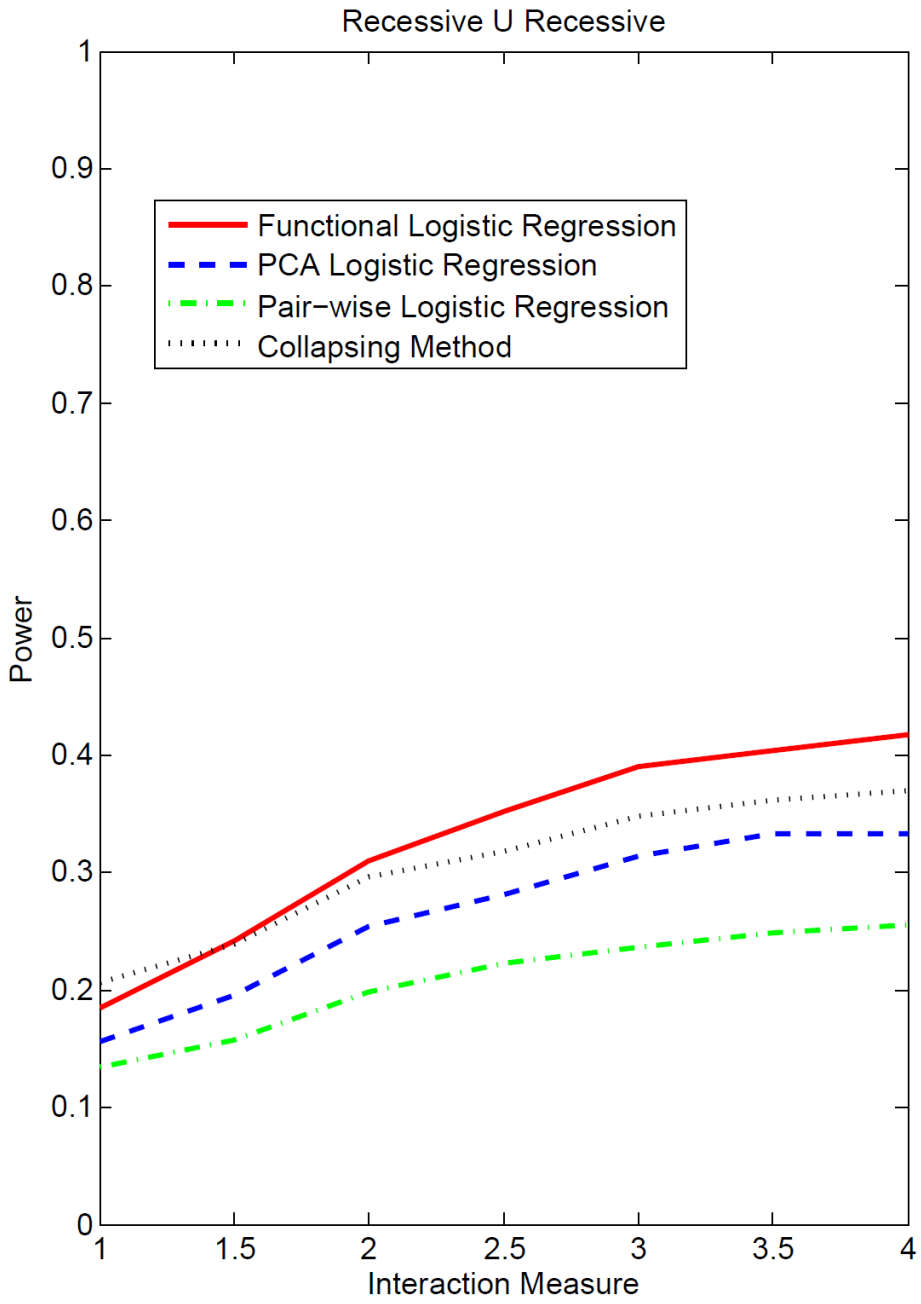


Figure S2. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of rare variants as a function of an interaction measure at the significance level $\alpha = 0.05$ under the recessive $\cup$ recessive model, assuming 2,000 cases and 2,000 controls, and 10% of risk variants.

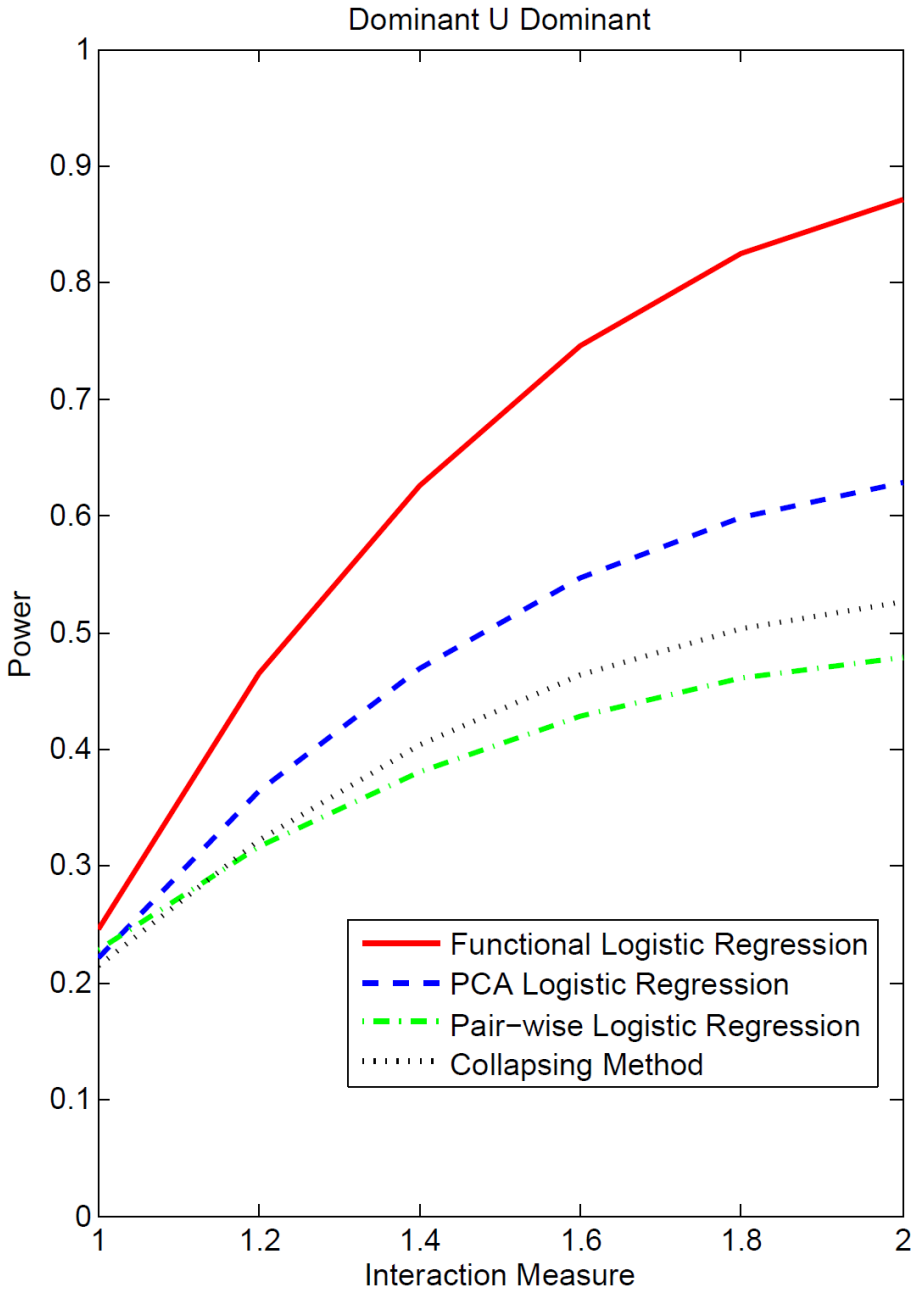


Figure S3. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of common variants as a function of an interaction measure at the significance level $\alpha = 0.05$ under the dominant $\cup$ dominant model, assuming 2,000 cases and 2,000 controls, and 10% of risk variants.

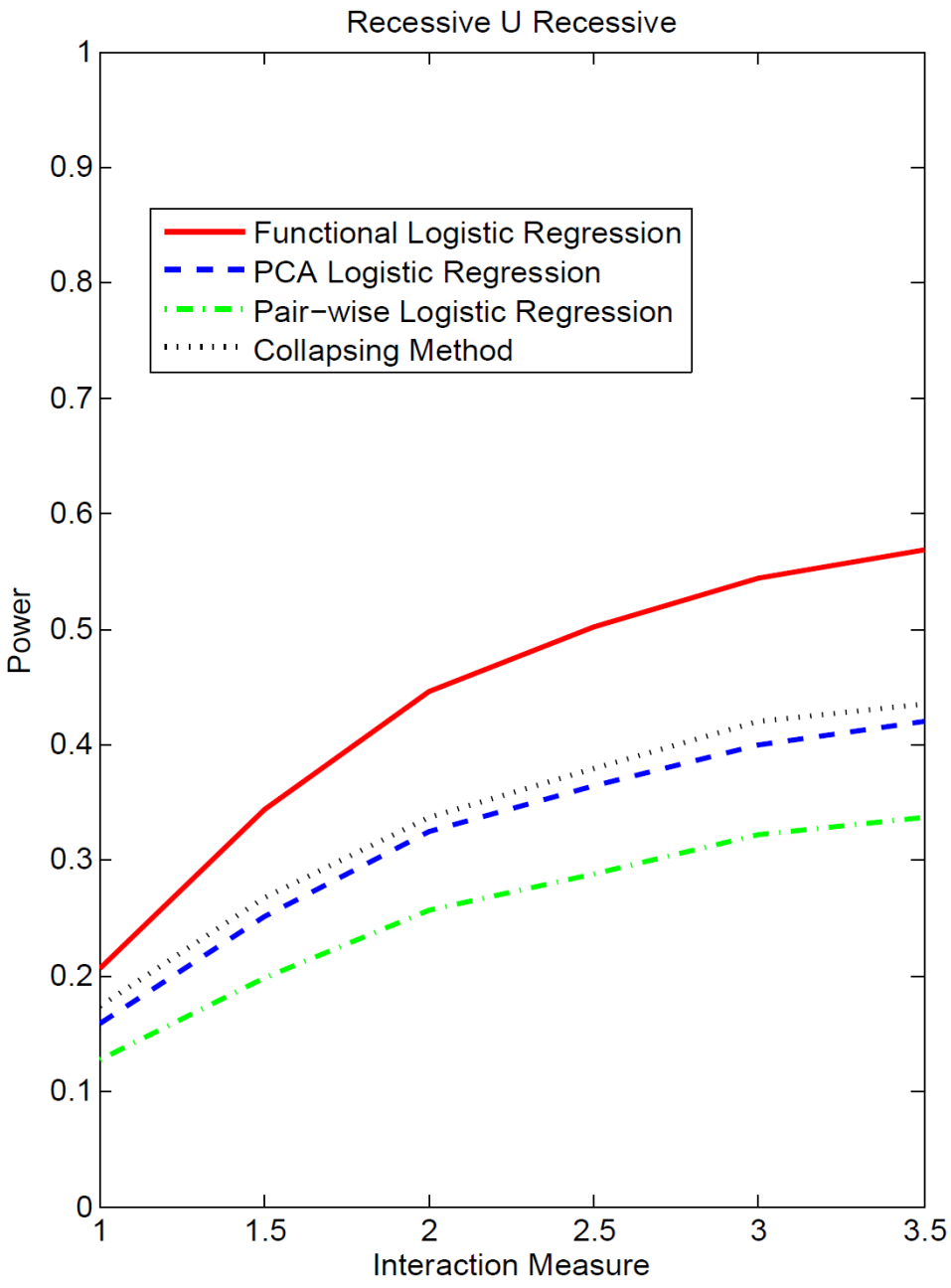


Figure S4. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of common variants as a function of an interaction measure at the significance level $\alpha = 0.05$ under the recessive $\cup$ recessive model, assuming 2,000 cases and 2,000 controls, and 10% of risk variants.

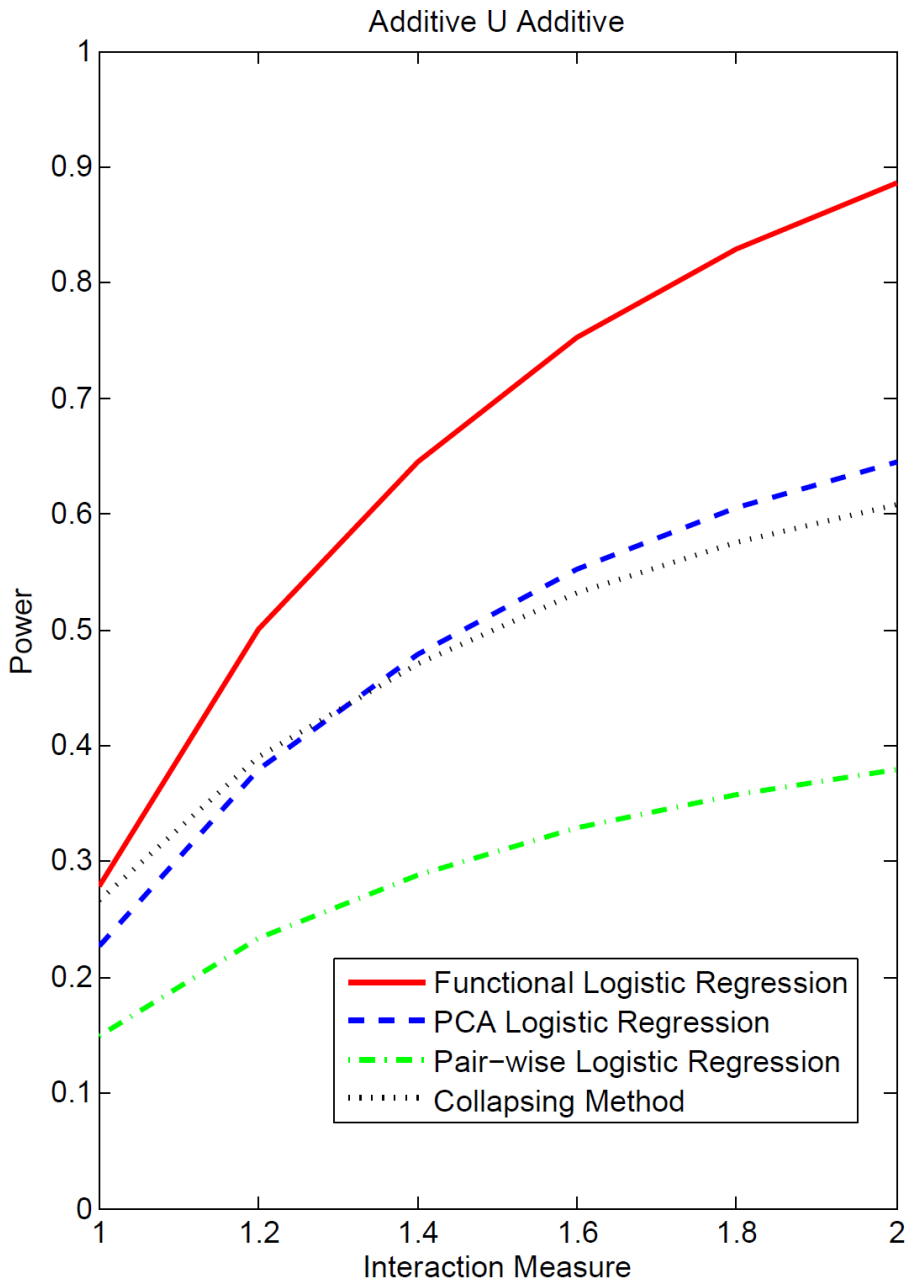


Figure S5. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of all variants as a function of an interaction measure at the significance level $\alpha = 0.05$ under the additive $\cup$ additive model, assuming 2,000 cases and 2,000 controls, and 10% of risk variants.

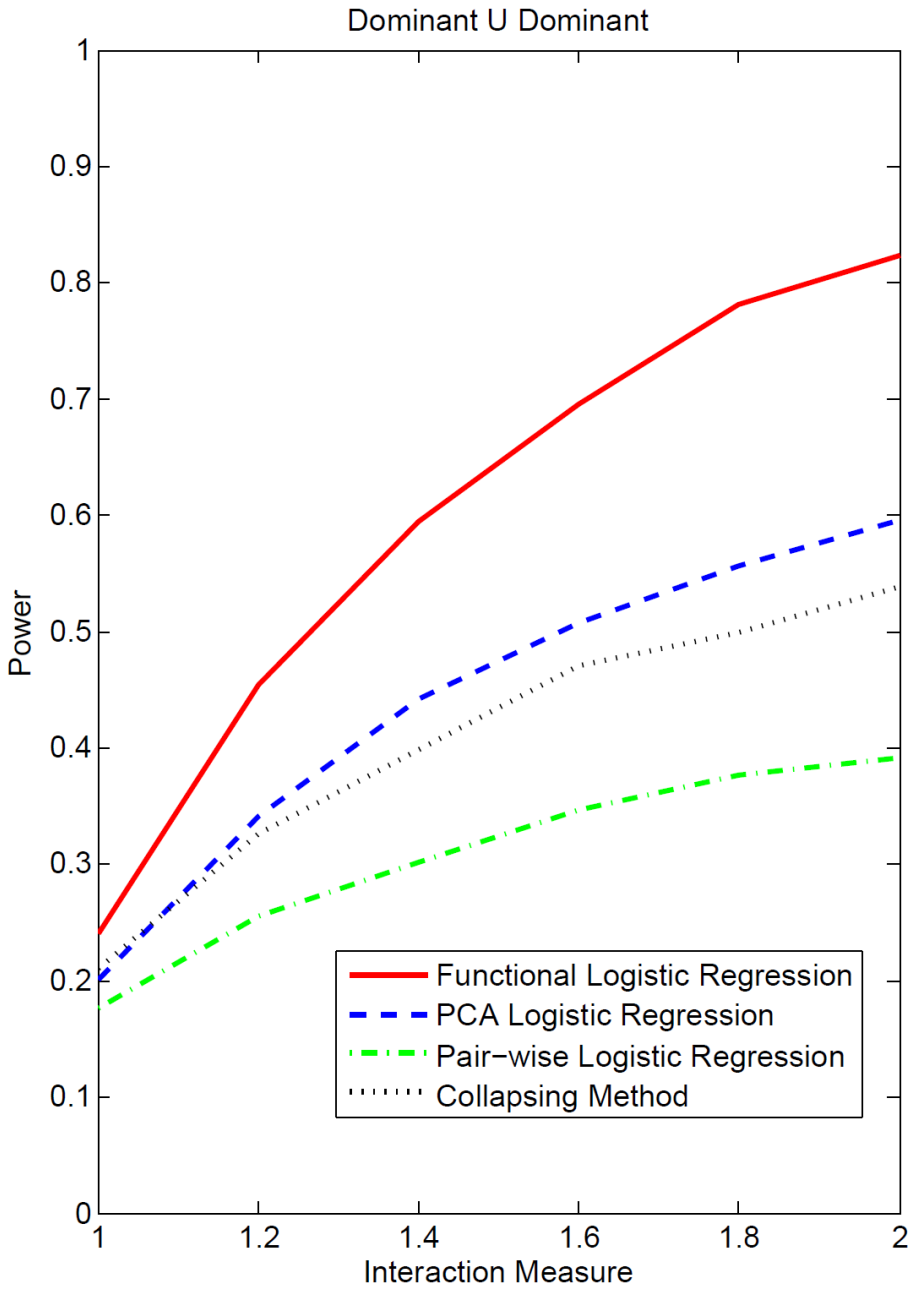


Figure S6. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of all variants as a function of an interaction measure at the significance level $\alpha = 0.05$ under the dominant $\cup$ dominant model, assuming 2,000 cases and 2,000 controls, and 10% of risk variants.

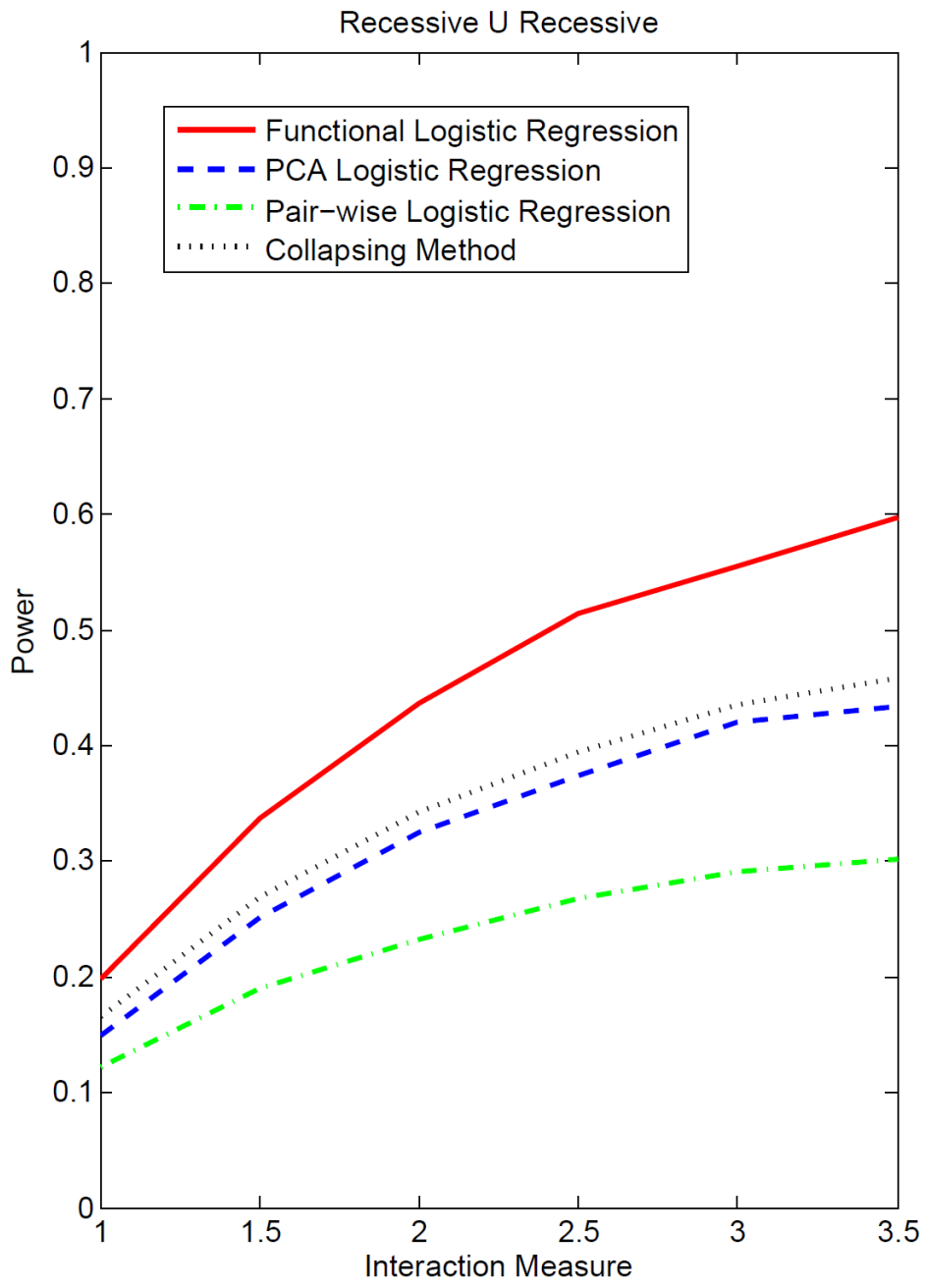


Figure S7. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of all variants as a function of an interaction measure at the significance level $\alpha = 0.05$ under the recessive $\cup$ recessive model, assuming 2,000 cases and 2,000 controls, and 10% of risk variants.

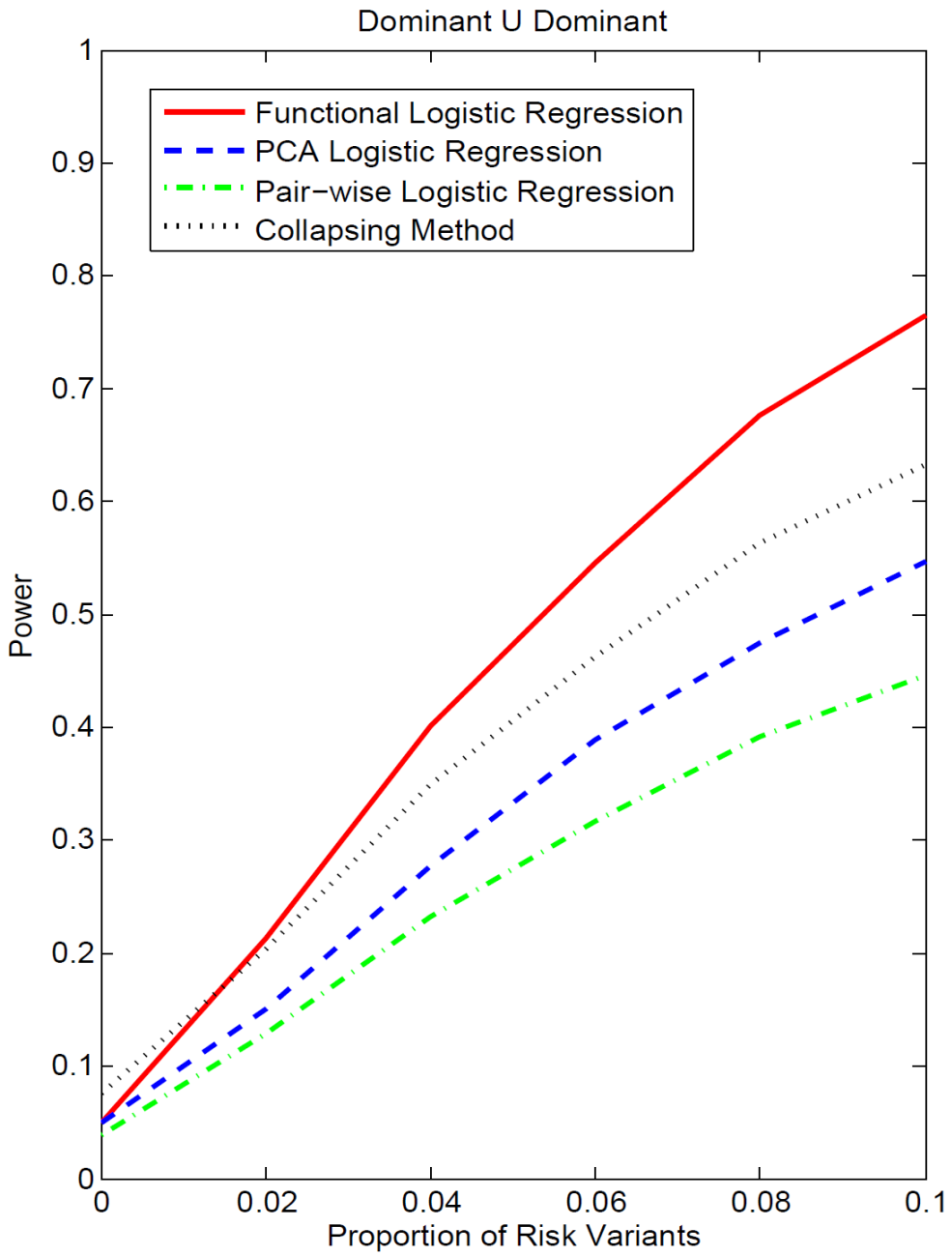


Figure S8. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of rare variants as a function of proportion of risk variants at the significance level $\alpha = 0.05$ under the dominant $\cup$ dominant model, assuming 2,000 cases and 2,000 controls and interaction measure of 2.5.

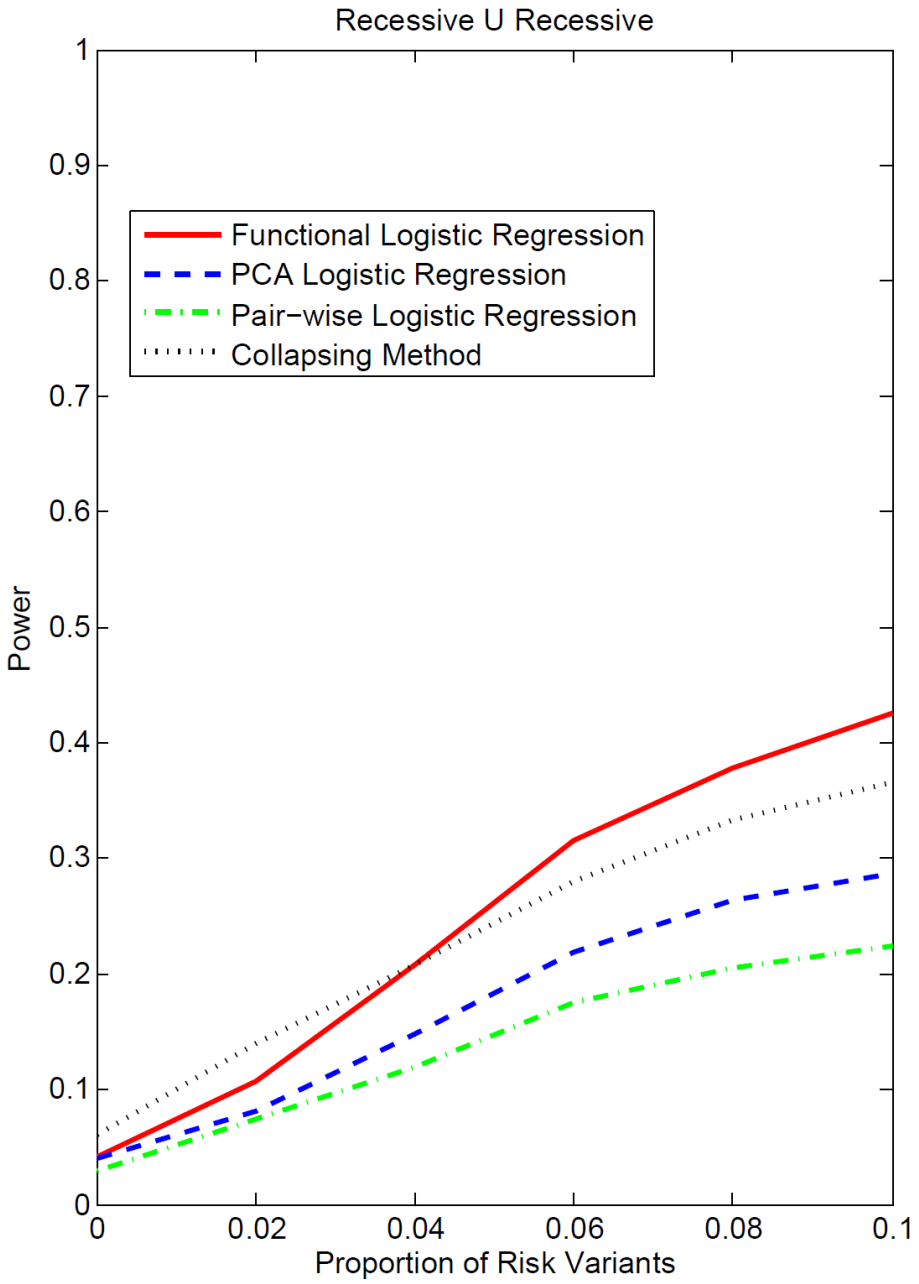


Figure S9. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of rare variants as a function of proportion of risk variants at the significance level $\alpha = 0.05$ under the Recessive $\cup$ Recessive model, assuming 2,000 cases and 2,000 controls and interaction measure of 2.5.

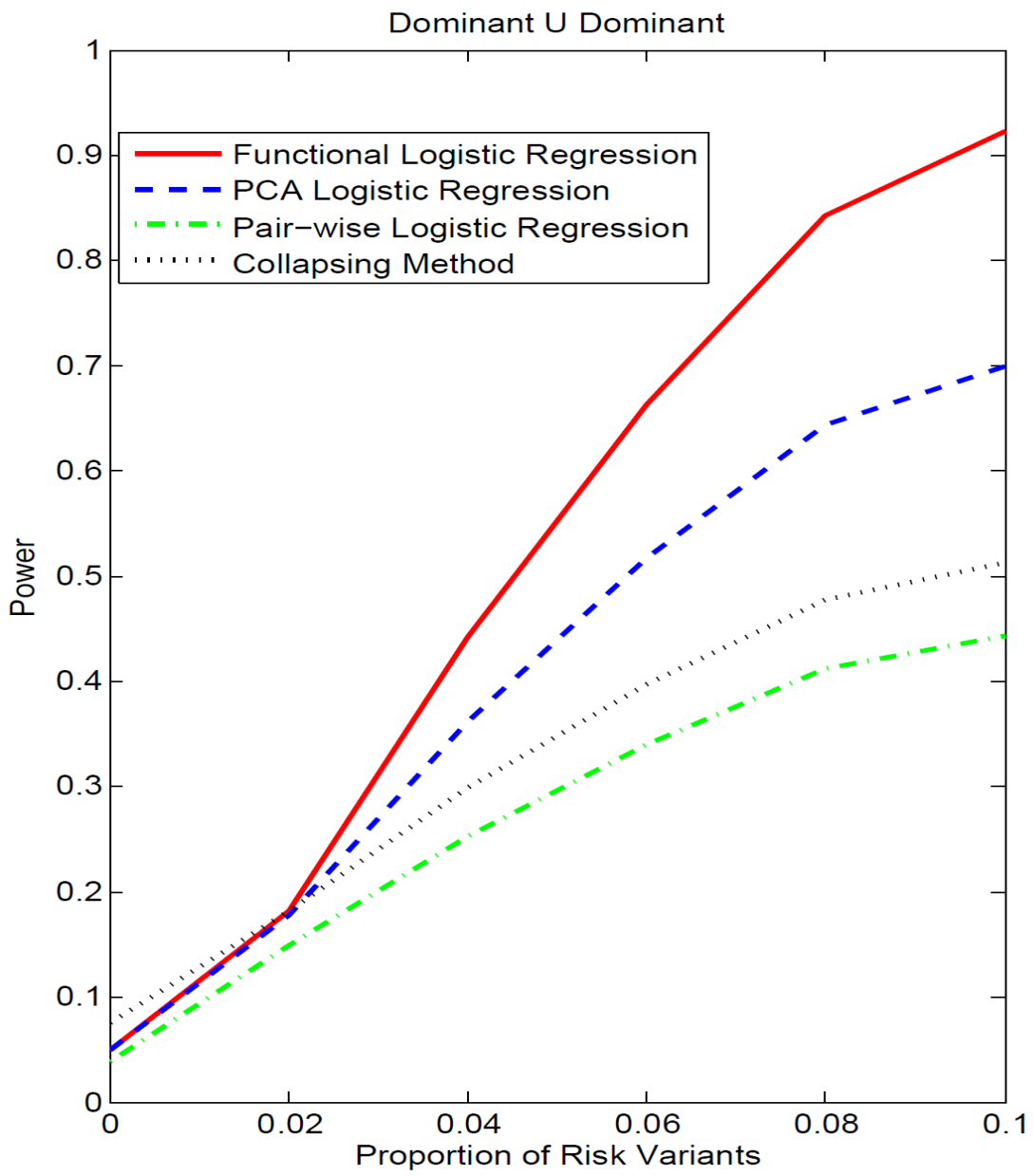


Figure S10. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of common variants as a function of proportion of risk variants at the significance level $\alpha = 0.05$ under the dominant $\cup$ dominant model, assuming 2,000 cases and 2,000 controls and the interaction measure of 2.5.

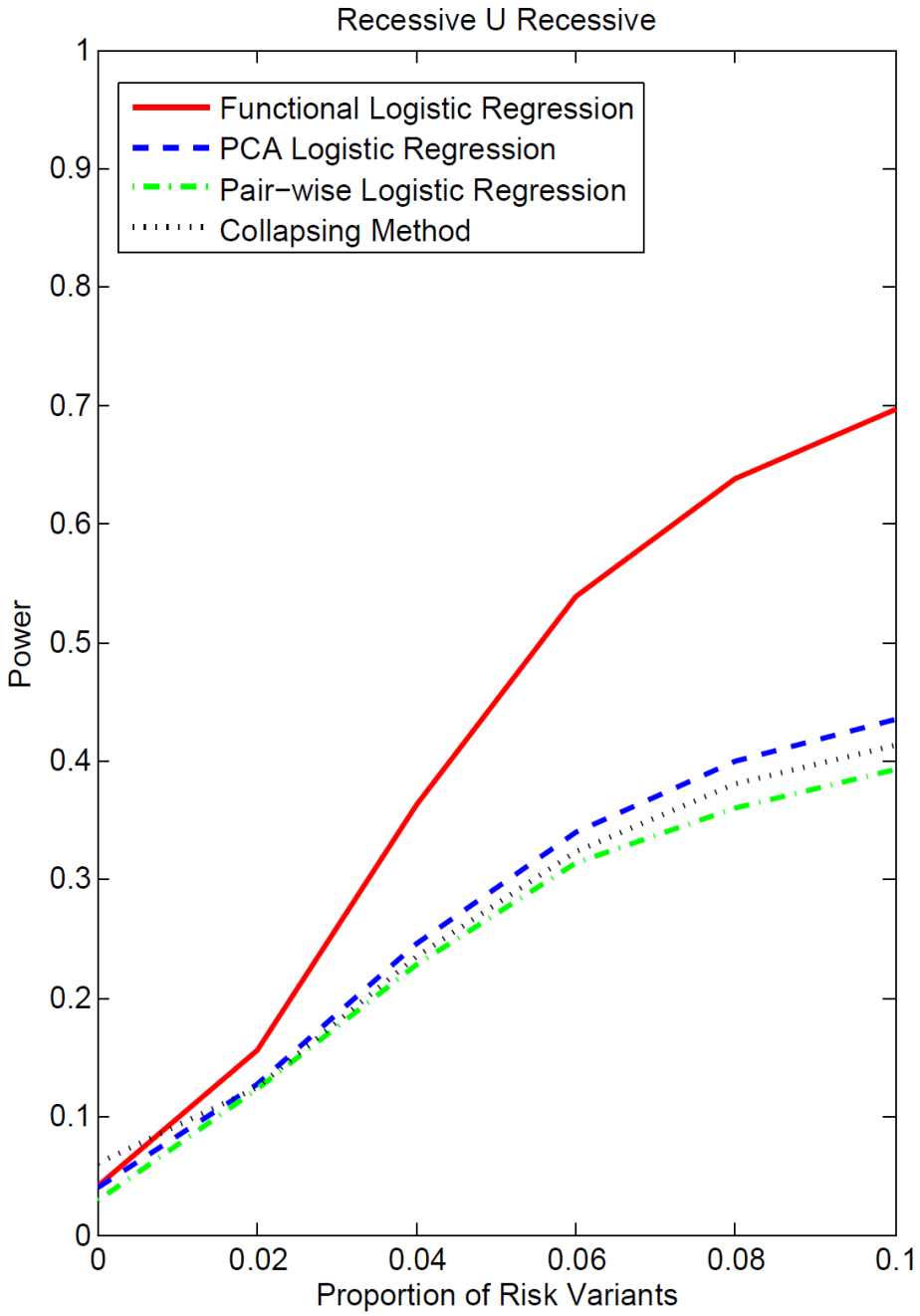


Figure S11. Power curves of four statistics: the function logistic regression, the PCA logistic regression, collapsing method, and the pair-wise logistic regression where permutations were used to adjust for multiple testing, for testing interaction between two genomic regions that consist of common variants as a function of proportion of risk variants at the significance level $\alpha = 0.05$ under the recessive $\cup$ recessive model, assuming 2,000 cases and 2,000 controls and the interaction measure of 2.5.