**File S1**  
**Supplement S1**

Table S1  Single Marker Analysis (SMA) of 200 replicates of chromosome 21 data with two isolated QTLs, and with sample size $N = 201$. GWT, CWT, BH, BY and LFDR are different SNP selection criteria for SMA. GWT is Genome-Wide Threshold which represents the $p$-value threshold $5.5 \times 10^{-8}$; CWT is Chromosome-Wide Threshold which represents the $p$-value threshold $5.5 \times 10^{-8} \times 3849034/50165$; BH denotes Benjamini-Hochberg FDR control; BY represents Benjamini-Yekutieli FDR control; LFDR1 and LFDR2 represent the approximate oracle procedure of Sun and Cai (2007) based on Efron’s Local FDR with two different sets of parameter estimates (see main text). For the FDR methods, a cut-off value of 0.05 was used. Comparison criteria are FWER, tFDR, TPR1, and TPR2. FWER is the empirical family-wise error rate (proportion of replicates with $\geq 1$ false positives); tFDR is an empirical thresholded false discovery rate mean(nFP/nSig) with nFP (nSig) being the number of false positives (significances), where a false positive is a SNP which is not in LD above a threshold $T$ with any QTL (causal SNP); $TPR1 = nCTP/Q$ is the first true positive rate with nCTP denoting the number of true QTLs identified and $Q$ denoting the number of true QTL (2*200) over all replicates; $TPR2 = nCLTP/Q$ is the second true positive rate with nCLTP denoting the number of true QTLs identified with the causal SNP or a linked SNP according to threshold $T$; $T$ is a threshold, on the absolute correlation between the allelic doses of the causal and a linked SNP, above which a linked SNP is counted as a true positive.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>$T$</th>
<th>GWT</th>
<th>BH</th>
<th>BY</th>
<th>LFDR1</th>
<th>LFDR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>tFWER</td>
<td>0.25</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0</td>
<td>0.036 $\pm$ 0.008</td>
<td>0</td>
<td>0.054 $\pm$ 0.010</td>
<td>0.059 $\pm$ 0.012</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0</td>
<td>0.040 $\pm$ 0.009</td>
<td>0</td>
<td>0.061 $\pm$ 0.010</td>
<td>0.064 $\pm$ 0.012</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0</td>
<td>0.055 $\pm$ 0.010</td>
<td>0.004 $\pm$ 0.003</td>
<td>0.086 $\pm$ 0.012</td>
<td>0.092 $\pm$ 0.014</td>
</tr>
<tr>
<td>TPR1</td>
<td>0.192 $\pm$ 0.020</td>
<td>0.510 $\pm$ 0.026</td>
<td>0.312 $\pm$ 0.025</td>
<td>0.535 $\pm$ 0.027</td>
<td>0.530 $\pm$ 0.027</td>
<td></td>
</tr>
<tr>
<td>TPR2</td>
<td>0.5 $\pm$ 0.021</td>
<td>0.520 $\pm$ 0.026</td>
<td>0.325 $\pm$ 0.025</td>
<td>0.550 $\pm$ 0.026</td>
<td>0.540 $\pm$ 0.027</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.7 $\pm$ 0.020</td>
<td>0.512 $\pm$ 0.026</td>
<td>0.320 $\pm$ 0.025</td>
<td>0.540 $\pm$ 0.027</td>
<td>0.532 $\pm$ 0.027</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.9 $\pm$ 0.020</td>
<td>0.510 $\pm$ 0.026</td>
<td>0.318 $\pm$ 0.025</td>
<td>0.538 $\pm$ 0.027</td>
<td>0.530 $\pm$ 0.027</td>
<td></td>
</tr>
</tbody>
</table>
Table S2  SMA of 200 replicates of chromosome 21 data with eight QTLs, and with sample size $N = 201$. See Table S1 for abbreviations and details.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>$T$</th>
<th>GWT</th>
<th>BH</th>
<th>BY</th>
<th>LFDR1</th>
<th>LFDR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>tFWER</td>
<td>0.25</td>
<td>0</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.135 + 0.024</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
<td>n.a</td>
</tr>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0</td>
<td>0.054 ± 0.006</td>
<td>0.010 ± 0.005</td>
<td>0.048 ± 0.006</td>
<td>0.048 ± 0.006</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0</td>
<td>0.061 ± 0.006</td>
<td>0.012 ± 0.005</td>
<td>0.054 ± 0.006</td>
<td>0.054 ± 0.006</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.016 ± 0.004</td>
<td>0.157 ± 0.008</td>
<td>0.065 ± 0.008</td>
<td>0.149 ± 0.008</td>
<td>0.147 ± 0.009</td>
</tr>
<tr>
<td>TPR1</td>
<td>0.25</td>
<td>0.296 ± 0.012</td>
<td>0.696 ± 0.016</td>
<td>0.528 ± 0.018</td>
<td>0.689 ± 0.016</td>
<td>0.678 ± 0.017</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.408 ± 0.015</td>
<td>0.792 ± 0.016</td>
<td>0.628 ± 0.020</td>
<td>0.788 ± 0.016</td>
<td>0.778 ± 0.017</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.358 ± 0.015</td>
<td>0.746 ± 0.016</td>
<td>0.584 ± 0.019</td>
<td>0.741 ± 0.016</td>
<td>0.729 ± 0.017</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.306 ± 0.013</td>
<td>0.699 ± 0.016</td>
<td>0.537 ± 0.018</td>
<td>0.695 ± 0.016</td>
<td>0.682 ± 0.017</td>
</tr>
</tbody>
</table>
Table S3  Penalized regression analysis using Lasso, Adaptive Lasso and Elastic Net, with FDR based selection of the penalty parameter values, of 200 replicates of chromosome 21 data with two isolated QTLs, and with sample size N = 201. FDR control was at the 0.05 level, and all methods used the analytic FDR except Lasso perm which used the permutation FDR. EN represents the Elastic Net with lasso weight ($\lambda_2$) set to 0.3, 0.5 or 0.9. AdaLasso represents the Adaptive Lasso using weights obtained by Lasso with CV (CV), Ridge Regression with CV (RR), and SMA (results for these three AdaLasso varieties were identical here). See Table S1 for other definitions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>T</th>
<th>Lasso perm</th>
<th>Lasso</th>
<th>AdaLasso (CV, RR, SMA)</th>
<th>EN ($\lambda_2 = 0.3$)</th>
<th>EN ($\lambda_2 = 0.5$)</th>
<th>EN ($\lambda_2 = 0.9$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0.022 ± 0.01</td>
<td>0.015 ± 0.008</td>
<td>0.015 ± 0.008</td>
<td>0.021 ± 0.010</td>
<td>0.019 ± 0.009</td>
<td>0.016 ± 0.008</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.025 ± 0.01</td>
<td>0.018 ± 0.008</td>
<td>0.018 ± 0.008</td>
<td>0.024 ± 0.010</td>
<td>0.021 ± 0.010</td>
<td>0.018 ± 0.008</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.027 ± 0.01</td>
<td>0.018 ± 0.008</td>
<td>0.018 ± 0.008</td>
<td>0.027 ± 0.010</td>
<td>0.023 ± 0.010</td>
<td>0.018 ± 0.008</td>
</tr>
<tr>
<td>TPR1</td>
<td></td>
<td>0.442 ± 0.024</td>
<td>0.382 ± 0.025</td>
<td>0.382 ± 0.025</td>
<td>0.400 ± 0.025</td>
<td>0.412 ± 0.026</td>
<td>0.392 ± 0.026</td>
</tr>
<tr>
<td>TPR2</td>
<td>0.5</td>
<td>0.512 ± 0.024</td>
<td>0.432 ± 0.027</td>
<td>0.432 ± 0.027</td>
<td>0.420 ± 0.026</td>
<td>0.432 ± 0.026</td>
<td>0.435 ± 0.027</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.495 ± 0.024</td>
<td>0.425 ± 0.027</td>
<td>0.425 ± 0.027</td>
<td>0.415 ± 0.026</td>
<td>0.428 ± 0.027</td>
<td>0.425 ± 0.027</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.482 ± 0.025</td>
<td>0.412 ± 0.027</td>
<td>0.412 ± 0.027</td>
<td>0.410 ± 0.026</td>
<td>0.422 ± 0.026</td>
<td>0.415 ± 0.026</td>
</tr>
</tbody>
</table>
Table S4  Penalized regression analysis using MCP with several fixed values of the second tuning parameter ($\lambda_2$) and with FDR based selection of the value for the first tuning parameter, of 200 replicates of chromosome 21 data with two isolated QTLs, and with sample size $N = 201$. FDR control was at the 0.05 level, and all methods used the analytic FDR. MCP represents Minimax Concave Penalty. See Table S1 for other definitions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>T</th>
<th>$\text{MCP} (\lambda_2 = 3)$</th>
<th>$\text{MCP} (\lambda_2 = 10)$</th>
<th>$\text{MCP} (\lambda_2 = 30)$</th>
<th>$\text{MCP} (\lambda_2 = 100)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0.022 ± 0.010</td>
<td>0.018 ± 0.008</td>
<td>0.020 ± 0.009</td>
<td>0.018 ± 0.008</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.025 ± 0.010</td>
<td>0.020 ± 0.009</td>
<td>0.022 ± 0.009</td>
<td>0.020 ± 0.009</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.025 ± 0.010</td>
<td>0.020 ± 0.009</td>
<td>0.022 ± 0.009</td>
<td>0.022 ± 0.009</td>
</tr>
<tr>
<td>TPR1</td>
<td></td>
<td>0.375 ± 0.025</td>
<td>0.380 ± 0.025</td>
<td>0.378 ± 0.025</td>
<td>0.380 ± 0.025</td>
</tr>
<tr>
<td>TPR2</td>
<td>0.5</td>
<td>0.432 ± 0.027</td>
<td>0.435 ± 0.027</td>
<td>0.432 ± 0.027</td>
<td>0.435 ± 0.027</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.422 ± 0.027</td>
<td>0.428 ± 0.027</td>
<td>0.425 ± 0.027</td>
<td>0.425 ± 0.027</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.410 ± 0.027</td>
<td>0.415 ± 0.027</td>
<td>0.412 ± 0.027</td>
<td>0.412 ± 0.027</td>
</tr>
</tbody>
</table>
Table S5  Penalized regression analysis using Lasso, Adaptive Lasso and Elastic Net, with FDR based selection of the penalty parameter values, of 200 replicates of chromosome 21 data with eight QTLs, and with sample size \( N = 201 \). FDR control was at the 0.05 level, and all methods used the analytic FDR except Lasso perm which used the permutation FDR. EN represents the Elastic Net with Lasso weight \((\lambda_2)\) set to 0.3, 0.5 or 0.9. AdaLasso represents the Adaptive Lasso using weights obtained by Lasso with CV (CV), Ridge Regression with CV (RR), and SMA. See Table S1 for other definitions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>T</th>
<th>Lasso perm</th>
<th>Lasso (CV)</th>
<th>AdaLasso (CV)</th>
<th>AdaLasso (RR, SMA)</th>
<th>EN ((\lambda_2 = 0.3))</th>
<th>EN ((\lambda_2 = 0.5))</th>
<th>EN ((\lambda_2 = 0.9))</th>
</tr>
</thead>
<tbody>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0.006 + 0.003</td>
<td>0.020 + 0.005</td>
<td>0.019 + 0.005</td>
<td>0.020 + 0.005</td>
<td>0.026 ± 0.005</td>
<td>0.023 ± 0.005</td>
<td>0.019 ± 0.005</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.008 + 0.003</td>
<td>0.023 + 0.005</td>
<td>0.022 + 0.005</td>
<td>0.023 + 0.005</td>
<td>0.028 ± 0.005</td>
<td>0.026 ± 0.005</td>
<td>0.022 ± 0.005</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.024 ± 0.006</td>
<td>0.042 ± 0.007</td>
<td>0.041 ± 0.006</td>
<td>0.042 ± 0.007</td>
<td>0.053 ± 0.006</td>
<td>0.048 ± 0.006</td>
<td>0.040 ± 0.006</td>
</tr>
<tr>
<td>TPR1</td>
<td>0.5</td>
<td>0.315 ± 0.011</td>
<td>0.378 ± 0.014</td>
<td>0.372 ± 0.014</td>
<td>0.378 ± 0.014</td>
<td>0.548 ± 0.017</td>
<td>0.466 ± 0.016</td>
<td>0.394 ± 0.014</td>
</tr>
<tr>
<td>TPR2</td>
<td>0.5</td>
<td>0.613 ± 0.016</td>
<td>0.692 ± 0.018</td>
<td>0.690 ± 0.018</td>
<td>0.692 ± 0.018</td>
<td>0.739 ± 0.018</td>
<td>0.718 ± 0.018</td>
<td>0.704 ± 0.018</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.514 ± 0.015</td>
<td>0.597 ± 0.018</td>
<td>0.594 ± 0.018</td>
<td>0.597 ± 0.018</td>
<td>0.661 ± 0.018</td>
<td>0.632 ± 0.018</td>
<td>0.609 ± 0.018</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.366 ± 0.012</td>
<td>0.441 ± 0.015</td>
<td>0.438 ± 0.015</td>
<td>0.441 ± 0.015</td>
<td>0.570 ± 0.017</td>
<td>0.509 ± 0.016</td>
<td>0.456 ± 0.015</td>
</tr>
</tbody>
</table>
Table S6  Penalized regression analysis using MCP with different fixed values of the second tuning parameter ($\lambda_2$) and with FDR based selection of the value of the first tuning parameter, of 200 replicates of chromosome 21 data with eight QTLs, and with sample size $N = 201$. FDR control was at the 0.05 level, and all methods used the analytic FDR. MCP represents Minimax Concave Penalty. See Table S1 for other definitions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>$T$</th>
<th>MCP ($\lambda_2 = 3$)</th>
<th>MCP ($\lambda_2 = 10$)</th>
<th>MCP ($\lambda_2 = 30$)</th>
<th>MCP ($\lambda_2 = 100$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0.018 ± 0.006</td>
<td>0.018 ± 0.005</td>
<td>0.018 ± 0.004</td>
<td>0.019 ± 0.005</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.020 ± 0.007</td>
<td>0.021 ± 0.005</td>
<td>0.021 ± 0.005</td>
<td>0.021 ± 0.005</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.029 ± 0.007</td>
<td>0.037 ± 0.006</td>
<td>0.037 ± 0.006</td>
<td>0.040 ± 0.006</td>
</tr>
<tr>
<td>TPR1</td>
<td></td>
<td>0.242 ± 0.010</td>
<td>0.325 ± 0.013</td>
<td>0.363 ± 0.014</td>
<td>0.378 ± 0.014</td>
</tr>
<tr>
<td>TPR2</td>
<td>0.5</td>
<td>0.643 ± 0.018</td>
<td>0.681 ± 0.018</td>
<td>0.694 ± 0.018</td>
<td>0.700 ± 0.018</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.476 ± 0.016</td>
<td>0.566 ± 0.017</td>
<td>0.593 ± 0.018</td>
<td>0.599 ± 0.018</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.289 ± 0.011</td>
<td>0.389 ± 0.014</td>
<td>0.428 ± 0.015</td>
<td>0.441 ± 0.015</td>
</tr>
</tbody>
</table>
Table S7  LD2lasso analyses of 20 randomly selected replicates of chromosome 21 data with two isolated QTLs, and with sample size N = 201. Lasso and EN ($\ell_2 = 0.5$) results are presented for comparison. For LD2lasso, $\phi$ is the relative weight on the Lasso penalty, and the LD function $h(r)$ was set to zero when the absolute correlation between two SNPs was $|r| < 0.85$ or 0.50. All methods used the analytic FDR method with control at the 0.05 level. See Table S1 for other definitions.

| Criterion | T  | Lasso ($\ell_2 = 0.5$) | LD2lasso $\phi = 0.9$ $h(r) = r^2$ $|r| < 0.85$ | LD2lasso $\phi = 0.5$ $h(r) = r^2$ $|r| < 0.85$ | LD2lasso $\phi = 0.9$ $h(r) = |r|$ $|r| < 0.85$ | LD2lasso $\phi = 0.5$ $h(r) = |r|$ $|r| < 0.85$ |
|-----------|----|-----------------------|-----------------------------------------------|-----------------------------------------------|------------------------------------------------|------------------------------------------------|
|           |    |                       |                                               |                                               |                                                |                                                |
|           |    | tFDR                  |                                               |                                               |                                                |                                                |
|           | 0.25 | 0                     | 0                                             | 0                                             | 0                                              | 0                                              |
|           | 0.3  | 0                     | 0                                             | 0                                             | 0                                              | 0                                              |
|           | 0.5  | 0                     | 0                                             | 0                                             | 0                                              | 0                                              |
|           | TPR1  | 0.450 ± 0.089         | 0.425 ± 0.092                               | 0.475 ± 0.098                               | 0.500 ± 0.101                                 | 0.475 ± 0.098                                 | 0.450 ± 0.101                                 |
|           | TPR2  | 0.450 ± 0.095         | 0.475 ± 0.098                               | 0.500 ± 0.101                               | 0.525 ± 0.098                                 | 0.500 ± 0.101                                 | 0.475 ± 0.098                                 |
|           |       |                        |                                               |                                               |                                                |                                                |                                                |
|           | 0.5  | 0.450 ± 0.095         | 0.475 ± 0.098                               | 0.500 ± 0.101                               | 0.525 ± 0.098                                 | 0.500 ± 0.101                                 | 0.475 ± 0.098                                 |
|           | TPR2  | 0.450 ± 0.095         | 0.475 ± 0.098                               | 0.500 ± 0.101                               | 0.525 ± 0.098                                 | 0.500 ± 0.101                                 | 0.475 ± 0.098                                 |
|           |       |                        |                                               |                                               |                                                |                                                |                                                |
|           | 0.9  | 0.425 ± 0.098         | 0.450 ± 0.095                               | 0.475 ± 0.098                               | 0.525 ± 0.098                                 | 0.475 ± 0.098                                 | 0.450 ± 0.101                                 |

| Criterion | T  | LD2lasso $\phi = 0.9$, $h(r) = r^2$, $|r| < 0.50$ | LD2lasso $\phi = 0.5$, $h(r) = r^2$, $|r| < 0.50$ |
|-----------|----|-------------------------------------------------|-------------------------------------------------|
|           |    |                                                 |                                                 |
|           |    | tFDR                                            |                                                 |
|           | 0.25 | 0                                                | 0                                                |
|           | 0.3  | 0                                                | 0                                                |
|           | 0.5  | 0                                                | 0                                                |
|           | TPR1  | 0.475 ± 0.098                                   | 0.450 ± 0.101                                   |
|           | TPR2  | 0.500 ± 0.101                                   | 0.475 ± 0.098                                   |
|           |       | 0.500 ± 0.101                                   | 0.475 ± 0.098                                   |
|           |       | 0.425 ± 0.098                                   | 0.450 ± 0.101                                   |
Table S8  LD2lasso analyses of 20 randomly selected replicates of chromosome 21 data with eight QTLs, and with sample size \( N = 201 \). Lasso and EN (\( \lambda_2 = 0.5 \)) results are presented for comparison. For LD2lasso, \( \varphi \) is the relative weight on the Lasso penalty, and the LD function \( h(r) \) was set to zero when the absolute correlation between two SNPs was \(|r| < 0.85\). All methods used the analytic FDR method with control at the 0.05 level. See Table S1 for other definitions.

| Criterion | T   | Lasso \((\lambda_2 = 0.5)\) | LD2lasso \(\varphi = 0.9\) \(h(r) = r^2\) \(|r| = 0.85\) | LD2lasso \(\varphi = 0.5\) \(h(r) = r^2\) \(|r| = 0.85\) | LD2lasso \(\varphi = 0.9\) \(h(r) = |r|\) \(|r| = 0.85\) | LD2lasso \(\varphi = 0.5\) \(h(r) = |r|\) \(|r| = 0.85\) |
|-----------|-----|-----------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| tFDR      | 0.25| 0.005 ± 0.005               | 0.191 ± 0.041                   | 0.201 ± 0.041                   | 0.195 ± 0.041                   | 0.204 ± 0.041                   |
|           | 0.3  | 0.005 ± 0.005               | 0.194 ± 0.041                   | 0.204 ± 0.041                   | 0.198 ± 0.041                   | 0.207 ± 0.041                   |
|           | 0.5  | 0.024 ± 0.014               | 0.252 ± 0.048                   | 0.259 ± 0.048                   | 0.256 ± 0.048                   | 0.269 ± 0.051                   |
| TPR1      | 0.269 ± 0.041 | 0.469 ± 0.044 | 0.500 ± 0.051 | 0.562 ± 0.054 | 0.500 ± 0.051 | 0.544 ± 0.06 |
| TPR2      | 0.631 ± 0.057 | 0.688 ± 0.060 | 0.769 ± 0.057 | 0.769 ± 0.057 | 0.769 ± 0.057 | 0.744 ± 0.066 |
|           | 0.569 ± 0.057 | 0.625 ± 0.057 | 0.650 ± 0.057 | 0.656 ± 0.057 | 0.65 ± 0.057 | 0.625 ± 0.067 |
|           | 0.9   | 0.512 ± 0.048 | 0.569 ± 0.054 | 0.600 ± 0.054 | 0.569 ± 0.054 | 0.581 ± 0.06   |
| Criterion | T   | LD2lasso \(\varphi = 0.9\) \(h(r) = r^2\) \(|r| = 0.50\) | LD2lasso \(\varphi = 0.5\) \(h(r) = r^2\) \(|r| = 0.50\) |
| tFDR      | 0.25 | 0.194 ± 0.041 \((0.00058)\) | 0.166 ± 0.041 \((0.0037)\) |
|           | 0.3  | 0.199 ± 0.041 \((0.00041)\) | 0.169 ± 0.041 \((0.0031)\) |
|           | 0.5  | 0.255 ± 0.048 \((6.271e-05)\) | 0.223 ± 0.048 \((0.0045)\) |
| TPR1      | 0.556 ± 0.054 | 0.650 ± 0.057 |
| TPR2      | 0.5  | 0.756 ± 0.060 | 0.775 ± 0.060 |
|           | 0.7  | 0.650 ± 0.057 | 0.694 ± 0.060 |
|           | 0.9  | 0.594 ± 0.057 | 0.662 ± 0.057 |
Table S9  Multi-split analysis using the method of Meinshausen et al. (2009) of 20 randomly chosen replicates of chromosome 21 data with two isolated QTLs or eight QTLs, and with sample size N=201. The analysis used Benjamini-Hochberg (BH) or Benjamini-Yekutieli (BY) FDR control. See Table S1 for other definitions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>T</th>
<th>Two QTLs</th>
<th>Eight QTLs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lasso BH</td>
<td>BY</td>
</tr>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0 0 0</td>
<td>0.005 ± 0.005</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0 0 0</td>
<td>0.005 ± 0.005</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0 0 0</td>
<td>0.024 ± 0.014</td>
</tr>
<tr>
<td>TPR1</td>
<td>0.5</td>
<td>0.350 ± 0.089 0.150 ± 0.073 0.050 ± 0.035</td>
<td>0.269 ± 0.041 0.094 ± 0.025 0.038 ± 0.016</td>
</tr>
<tr>
<td>TPR2</td>
<td>0.7</td>
<td>0.450 ± 0.095 0.175 ± 0.076 0.050 ± 0.035</td>
<td>0.631 ± 0.057 0.206 ± 0.051 0.069 ± 0.032</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.425 ± 0.098 0.175 ± 0.076 0.050 ± 0.035</td>
<td>0.419 ± 0.047 0.100 ± 0.025 0.044 ± 0.019</td>
</tr>
</tbody>
</table>
Table S10  Adaptive Lasso with Local FDR estimation with bootstrap size 100 and different cut-off values for the local FDR (locFDR), of 200 replicates of chromosome 21 data with two isolated QTLs, and with sample size N=201. For comparison, results for the Lasso with analytic FDR control are also shown. See Table S1 for other definitions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Two QTLs</th>
<th>Eight QTLs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lasso</td>
<td>AdaLasso locFDR=0.1</td>
</tr>
<tr>
<td>tFDR 0.25</td>
<td>0.015 ± 0.008</td>
<td>0.020 ± 0.006</td>
</tr>
<tr>
<td>0.3</td>
<td>0.018 ± 0.008</td>
<td>0.020 ± 0.006</td>
</tr>
<tr>
<td>0.5</td>
<td>0.018 ± 0.008</td>
<td>0.020 ± 0.006</td>
</tr>
<tr>
<td>TPR1</td>
<td>0.382 ± 0.025</td>
<td>0.150 ± 0.020</td>
</tr>
<tr>
<td>TPR2 0.5</td>
<td>0.432 ± 0.027</td>
<td>0.175 ± 0.021</td>
</tr>
<tr>
<td>0.7</td>
<td>0.425 ± 0.027</td>
<td>0.175 ± 0.021</td>
</tr>
<tr>
<td>0.9</td>
<td>0.412 ± 0.027</td>
<td>0.175 ± 0.021</td>
</tr>
</tbody>
</table>

H. Yi et al. 11 SI
Table S11  Two dimensional (2D) MCP implemented in PUMA, across 200 replicates of chromosome 21 data with eight QTLs, and with sample size N=201. Final SNP selection was based on p-values with different cut-offs.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>T</th>
<th>2D MCP with different p-value thresholds</th>
<th>2D MCP analytic FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1x10^{-07}</td>
<td>1x10^{-06}</td>
</tr>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0.025 ± 0.006</td>
<td>0.061 ± 0.009</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.025 ± 0.006</td>
<td>0.061 ± 0.009</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.025 ± 0.006</td>
<td>0.066 ± 0.01</td>
</tr>
<tr>
<td>TPR1</td>
<td></td>
<td>0.191 ± 0.009</td>
<td>0.217 ± 0.009</td>
</tr>
<tr>
<td>TPR2</td>
<td>0.5</td>
<td>0.516 ± 0.017</td>
<td>0.581 ± 0.016</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.386 ± 0.013</td>
<td>0.433 ± 0.013</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.231 ± 0.009</td>
<td>0.261 ± 0.009</td>
</tr>
</tbody>
</table>
**Table S12  Joint analysis of chromosomes 19, 21 and 22 by single marker analysis.** Compared are pooled and separate analyses using the local FDR (locFDR) based thresholding procedure or the BH procedure, and the locFDR based grouping procedure of Cai and Sun (2009) referred to as Group locFDR. See Table S1 for other definitions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>T</th>
<th>Pooled locFDR</th>
<th>Separate locFDR</th>
<th>Group locFDR</th>
<th>Pooled BH</th>
<th>Separate BH</th>
</tr>
</thead>
<tbody>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0.026 ± 0.004</td>
<td>0.036 ± 0.004</td>
<td>0.052 ± 0.005</td>
<td>0.050 ± 0.004</td>
<td>0.058 ± 0.005</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.029 ± 0.004</td>
<td>0.039 ± 0.004</td>
<td>0.058 ± 0.005</td>
<td>0.056 ± 0.005</td>
<td>0.065 ± 0.005</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.087 ± 0.006</td>
<td>0.117 ± 0.007</td>
<td>0.135 ± 0.007</td>
<td>0.127 ± 0.007</td>
<td>0.156 ± 0.007</td>
</tr>
<tr>
<td>TPR1</td>
<td>0.5</td>
<td>0.576 ± 0.014</td>
<td>0.610 ± 0.013</td>
<td>0.634 ± 0.013</td>
<td>0.636 ± 0.013</td>
<td>0.659 ± 0.012</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.624 ± 0.014</td>
<td>0.654 ± 0.013</td>
<td>0.682 ± 0.013</td>
<td>0.682 ± 0.014</td>
<td>0.699 ± 0.012</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.582 ± 0.014</td>
<td>0.616 ± 0.013</td>
<td>0.642 ± 0.013</td>
<td>0.644 ± 0.013</td>
<td>0.665 ± 0.012</td>
</tr>
</tbody>
</table>
Table S13  Analyses of chromosomes 19, 21 and 22 by separate and pooled Lasso PR with analytic FDR control at the 0.05 level. Also compared are the results from separate and pooled analyses of the three chromosomes for individual chromosomes 21 and 22. See Table S1 for other abbreviations / definitions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>T</th>
<th>Separate</th>
<th>Pooled</th>
<th>C21 Separate</th>
<th>C21 Pooled</th>
<th>C22 Separate</th>
<th>C22 Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td>tFDR</td>
<td>0.25</td>
<td>0.028 ± 0.005</td>
<td>0.023 ± 0.004</td>
<td>0.021 ± 0.004</td>
<td>0.010 ± 0.003</td>
<td>0.025 ± 0.009</td>
<td>0.024 ± 0.007</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.028 ± 0.005</td>
<td>0.024 ± 0.004</td>
<td>0.021 ± 0.004</td>
<td>0.010 ± 0.003</td>
<td>0.025 ± 0.009</td>
<td>0.024 ± 0.007</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.041 ± 0.005</td>
<td>0.035 ± 0.005</td>
<td>0.034 ± 0.005</td>
<td>0.024 ± 0.005</td>
<td>0.028 ± 0.009</td>
<td>0.028 ± 0.008</td>
</tr>
<tr>
<td>TPR1</td>
<td>0</td>
<td>0.399 ± 0.011</td>
<td>0.414 ± 0.013</td>
<td>0.422 ± 0.013</td>
<td>0.390 ± 0.014</td>
<td>0.305 ± 0.022</td>
<td>0.510 ± 0.024</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.656 ± 0.013</td>
<td>0.664 ± 0.016</td>
<td>0.733 ± 0.016</td>
<td>0.687 ± 0.017</td>
<td>0.348 ± 0.023</td>
<td>0.572 ± 0.025</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.577 ± 0.012</td>
<td>0.586 ± 0.015</td>
<td>0.636 ± 0.015</td>
<td>0.591 ± 0.016</td>
<td>0.340 ± 0.023</td>
<td>0.568 ± 0.025</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.456 ± 0.011</td>
<td>0.472 ± 0.014</td>
<td>0.486 ± 0.014</td>
<td>0.449 ± 0.015</td>
<td>0.340 ± 0.023</td>
<td>0.565 ± 0.025</td>
</tr>
</tbody>
</table>
Figure S1   Histogram of the z-values pertaining to 21,530 SNPs on chromosome 21. The blue solid curve represents the N(0,1) distribution, which fits the empirical z-value distribution well.
HABC Study

The Health ABC study is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in older adults. Health ABC enrolled well-functioning, community-dwelling black (n=1281) and white (n=1794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. The present study sample consists of 786 white participants with available genotyping and IL-6 SR data.

Phenotypic Information

To measure the level of IL-6 SR, venipuncture was performed for each of the participants after an overnight fast of at least 8 h, and serum samples were then frozen at -70°C. IL-6 SR levels were measured by ultrasensitive ELISA (R&D Systems) and had CVs of 3.5%–5.2%.

Genotyping and Imputation

Genomic DNA was extracted from buffy coat collected using PUREGENE® DNA Purification Kit during the baseline exam. Genotyping was performed by the Center for Inherited Disease Research (CIDR) using the Illumina Human1M-Duo BeadChip system. Samples were excluded from the dataset for the reasons of sample failure, genotypic sex mismatch, and first-degree relative of an included individual based on genotype data. Genotyping was successful for 1,151,215 SNPs in 2,802 unrelated individuals (1663 Caucasians and 1139 African Americans). Imputation was done for the autosomes using the MACH software version 1.0.16. SNPs with minor allele frequency ≥ 1%, call rate ≥97% and HWE p≥10^{-6} were used for imputation. HapMap II phased haplotypes were used as reference panels. For Caucasians, genotypes were available on 914,263 high quality SNPs for imputation based on the HapMap CEPH reference panel (release 22, build 36).
Supplement S3

# Simulate 200 SNP genotype data sets using Hapgen2

# Change work directory to the directory containing "hapgen2" software
# and HapMap2 Ref#24 haplotype data sets
cd ~/.../Hapgen2

# Do the loop to generate 200 SNP genotype data sets.
# Note that the seed is the current time by default.
# So it's better to generate a few data sets in each loop,
# and do it for a few times with time gaps.
for i in {1..10}; do
    ./Hapgen2 -h hapmap_r24_b36_fwd.consensus.qc.poly.chr21_ceu.phased -l chr21_ceu.r24.legend -m genetic_map_chr21_CEU_b36.txt -dl 34238344 1 1 1 0 201 0 -o ./Results2/$i; done

---

## R Code to pre-process SNP genotype data

iteration <- 200  # Number of SNP data replicates
N <- 201          # Sample size
# Record number of SNPs remaining after filtering by
# MAF <= 0.01 and correlation >= 0.999.
nSNPMAF <- nSNPMAFVar <- nSNPVarCor <- numeric(iteration)
# Record SNP positions for each of the 200 SNP data.
SNP.pos <- matrix(0, iteration, 30000)

for (its in 1:iteration) {
    ### Read each SNP data set
    setwd(".../Hapgen2/Results2")
dataName <- paste(its,".controls.gen", sep="")
    Data <- read.table(dataName)
Genotype0 <- Data[, -c(1:5)]  # first 5 columns are info
    ### compute allelic dose from three-column SNP data
    Genotype1 <- matrix(0, nrow(Genotype0), N)  # P*N SNP matrix
    for (t in 1:N){
        index <- which(Genotype0[, (t-1)*3+1] ==1)
        Genotype1[index,t] <- 0
        index <- which(Genotype0[, (t-1)*3+2] ==1)
        Genotype1[index,t] <- 1
        index <- which(Genotype0[, t*3] ==1)
        Genotype1[index,t] <- 2
    }
row.names(Genotype1) <- Data[, 3] # SNP positions

### Remove SNPs with MAF <= 0.01
MAF <- rowSums(Genotype1) / (2 * ncol(Genotype1))
id1 <- c(which(MAF <= 0.01), which(MAF >= 0.99))
Genotype2 <- t(Genotype1[-id1,])
colnames(Genotype2) <- rownames(Genotype2)[-id1]
nSNPMAF[its] <- ncol(Genotype2) # no. SNPs after MAF filter

### Remove SNPs with variance 0
id2 <- c(which(colSums(Genotype2) == 0), which(colSums(Genotype2) == N), which(colSums(Genotype2) == (2*N)), which(colSums(Genotype2) == (3*N)))
Genotype3 <- Genotype2[-id2]
colnames(Genotype3) <- colnames(Genotype2)[-id2]
nSNPMAFVar[its] <- ncol(Genotype3) # Record number of SNPs after MAF <= 0.01 filtering
X <- Genotype3

### Remove SNPs with correlation > 0.999
for (t in 1:10) {
P <- ncol(X)
id2 <- c()
for (i in 1:(P-1)) {
  if ((i+500) <= P) {
    COR <- cor(X[,i], X[, (i+1):(i+500)])
    temp <- which(abs(COR) > 0.999)
    id2 <- c(id2, i+temp)
  }
  if ((i+500) > P) {
    COR <- cor(X[,i], X[, (i+1):P])
    temp <- which(abs(COR) > 0.999)
    id2 <- c(id2, i+temp)
  }
}
if (length(unique(id2)) > 0) {
tmp <- X[, -unique(id2)]
X <- tmp
}
if (length(unique(id2)) == 0) {
nSNPVarCor[its] <- ncol(X)
break
}
}

### Change data to minor alleles consistent for all SNPs
tmp <- colSums(X) / (2*nrow(X))
id <- which(tmp >= 0.5)
X[, id] <- 2 - X[, id]

### Record SNPs positions
SNP.pos[its, (1:ncol(X))] <- colnames(X)
### Save final, clean SNP data sets

setwd(".../Hapgen2/SNP_matrix")
dataName2 <- paste(its,".txt", sep="") # Define data name
write.table(X, dataName2, row.names=FALSE, col.names=FALSE)
}

### Save SNP positions for each SNP data replicate
write.table(SNP.pos, '.../Hapgen2/SNP_matrix/position.txt', row.names=FALSE, col.names=FALSE)

### Save number of SNPs after MAF and correlation scanning
write.table(nSNPCor, '.../Hapgen2/SNP_matrix/nSNPCor2.txt', row.names=FALSE, col.names=FALSE)

#### Get the common SNPs across 200 data replicates

### Get the common SNP positions across 200 data replicates
SNP.pos <- read.table('.../Hapgen2/SNP_matrix/position.txt')
pos1 <- SNP.pos[1, -which(SNP.pos[1,] == 0)]
pos2 <- SNP.pos[2, -which(SNP.pos[2,] == 0)]
comSNP <- pos1[ which(is.na(match(pos1, pos2)) == F) ]
for (i in 3:iteration) {
    tmp <- SNP.pos[i, -which(SNP.pos[i,] == 0)]
    comSNP <- tmp[ which(is.na(match(tmp, comSNP)) == F) ]
}
length(comSNP) # 9224

### Save common SNP matrix across 200 data replicates
X.all <- matrix(0, N*iteration, length(comSNP))
for (its in 1:iteration) {
    dataName2 <- paste(its,".txt", sep="")
    X <- read.table(dataName2)
    X <- as.matrix(X)
id <- match(comSNP, SNP.pos[its, ])
    X.all[((its-1)*N +1):(its*N), ] <- X[, id]
}
write.table(X.all, '.../Hapgen2/SNP_matrix/commonSNPs.txt', row.names=FALSE, col.names=FALSE)

### Save common SNPs positions
write.table(comSNP, '.../Hapgen2/SNP_matrix/commSNPnames.txt', row.names=FALSE, col.names=FALSE)

#### Select two isolated QTL loci

### Select two QTL loci
loc <- c(174, 8566)
Q <- length(loc)   # Number of QTLs
MAF <- colSums(X.all) / (2*nrow(X.all))
MAF[loc]
locQTL <- comSNP[loc]  # locations: 14516982 44540242
X.2QTL <- X.all[, loc]

### Calculate variance and correlation of 2 QTLs
fi <- colSums(X.2QTL) / (2*nrow(X.2QTL))
2*fi*(1-fi)  # SNPs variance based on HWE
diag(var(X.2QTL))  # SNPs variance
cor(X.2QTL)  # correlation of two QTLs

### Test if the correlation of 2 QTLs is significant or not
require("psych")
r.test(N*iteration, cor(X.2QTL)[1,2])$p  # Not significant

### Save QTL data across 200 data replicates
write.table(X.2QTL, '.../Hapgen2/SNP_matrix/twoQTLs.txt',
row.names=FALSE, col.names=FALSE)

SNP.pos <- read.table(".../Hapgen2/SNP_matrix/position.txt")
N <- 201
Q <- length(loc)
locAll <- matrix(0, iteration, Q)
for (its in 1:iteration) {
  setwd(".../Hapgen2/SNP_matrix")
dataName2 <- paste(its,".txt", sep="")
X <- read.table(dataName2)
  # Get location of the two QTLs in each SNP data
loc <- which(is.na(match(SNP.pos[its, -which(SNP.pos[its, 0])], locQTL)) == F)
locAll2[its, ] <- loc
}
# Save two QTL locations across 200 SNP data replicates
write.table(locAll2, '.../Hapgen2/SNP_matrix/locAll2.txt',
row.names=FALSE, col.names=FALSE)

X.2QTL <- read.table(".../Hapgen2/SNP_matrix/twoQTLs.txt")
locAll <- read.table(".../Hapgen2/SNP_matrix/locAll2.txt")
h <- c(0.1, 0.1)  # Heritability of each QTL
beta <- sqrt(h/diag(var(X.2QTL)))  # var(Y) = 1
var_res <- 1 - sum(h)  # residual variance
for (its in 1:iteration) {
    setwd(".../Hapgen2/SNP_matrix")
dataName2 <- paste(its,".txt", sep="")
X <- read.table(dataName2)
Y <- as.matrix( X[, unlist(locAll[its, ])])%*%beta + rnorm(N, 0, sqrt(var_res))
setwd(".../Hapgen2/Response_2QTL")
Yname <- paste("Y", its, ".txt", sep="")
write.table(Y, Yname, row.names=FALSE, col.names=FALSE)
}

#################################################################
############# Select eight correlated QTL loci
#################################################################

### Select eight QTL loci: 2 groups of 2 SNPs with r^2=0.5, one
### group of 4 SNPs with 0.01<=r^2<=0.1.
loc <- c(195,198, 3337,3341,3343,3344, 8803,8814)
Q <- length(loc)
# Location of eight QTLs:
# 1st group: 14562752, 14569224
# 2nd group: 27048038, 27053105, 27068060, 27069599
# 3rd group: 45312259, 45360242
MAF <- colSums(X.all) / (2*nrow(X.all))
locQTL <- comSNP[loc]
X.8QTL <- X.all[, loc]

### variance and correlation of eight QTLs
fi <- colSums(X.8QTL) / (2*nrow(X.8QTL))
2*fi*(1-fi)
diag(var(X.8QTL))
cor(X.8QTL)

### Test if correlation between groups is significant or not
require("psych")
r.test(N*iteration, cor(X.8QTL)[1,3])$p     # not significant

### Save eight QTL data across 200 data replicates
write.table(X.8QTL, '.../Hapgen2/SNP_matrix/eightQTLs.txt',
row.names=FALSE, col.names=FALSE)

#################################################################
##### Get eight QTL locations for all SNP data replicates
#################################################################
SNP.pos <- read.table('.../Hapgen2/SNP_matrix/position.txt')
N <- 201
Q <- length(loc)
locAll <- matrix(0, iteration, Q)
for (its in 1:iteration) {

H. Yi et al. 21 SI
```r
setwd(".../Hapgen2/SNP_matrix")
dataName2 <- paste(its,".txt", sep="")
X <- read.table(dataName2)
loc <- which(is.na(match(SNP.pos[its, -which(SNP.pos[its, ] == 0)], locQTL)) == F)
locAll8[its, ] <- loc
write.table(locAll8, '.../Hapgen2/SNP_matrix/locAll8.txt',
row.names=FALSE, col.names=FALSE)

### Generate phenotype data Y with 8 QTLs

X.8QTL <- read.table(".../Hapgen2/SNP_matrix/eightQTLs.txt")
locAll <- read.table(".../Hapgen2/SNP_matrix/locAll8.txt")
h <- c(rep(0.04, 2), rep(0.05, 4), rep(0.04, 2)) # QTL h2
sum(h) # Total h2 = 0.68
E <- c(1, 1, 1, -1, 1, -1, -1, -1) # effect signs
beta <- sqrt(h/diag(var(X[, loc]))) * E # var(Y) = 1
var_res <- 1 - sum(h) # residual variance

for (its in 1:iteration) {
  setwd(".../Hapgen2/SNP_matrix")
dataName2 <- paste(its,".txt", sep="")
X1 <- read.table(dataName2)
Y <- as.matrix( X1[, locAll[its, ]])%*%beta + rnorm(N, 0, sqrt(var_res))
setwd(".../Hapgen2/Response_8QTL")
Yname <- paste("Y", its, ".txt", sep="")
write.table(Y, Yname, row.names=FALSE, col.names=FALSE)
}

### actual heritability explained by each QTL by considering LD

R <- cov(X[, loc])
H <- indH <- numeric(Q)
for (i in 1:Q) {
  temp <- 0
  for (j in 1:Q) {
    H[i] <- H[i] + R[i, j]*beta[i]*beta[j]
    if (j != i) temp <- temp + R[i, j]*beta[j]/R[i,i]
  }
  indH[i] <- R[i,i]*(beta[i] + temp)^2
}

sumH <- sum(H)
totH <- sumH / (sumH + var_res) # total heritability by taking into account LD
individualH <- indH / (sumH + var_res)
round(individualH, 3) # Actual individual heritability with LD
```
standardize <- function(X)
{
  n <- nrow(X)
  center <- colMeans(X)
  X.c <- sweep(X, 2, center)
  scale <- sqrt(apply(X.c,2,crossprod)/n)
  val <- sweep(X.c, 2, scale,"/")
  attr(val,"center") <- center
  attr(val,"scale") <- scale
  val
}

### this function performs Elastic Net Penalized Regression
### with analytic FDR control
### Arguments:
### X.D: N x p.D matrix of design covariates
### X.SNP: N x p.SNP matrix of SNP covariates (gene doses)
### N: sample size
### p.D: number of design covariates
### p.SNP: number of SNPs
### y: N x 1 vector of continuous phenotypes (assumed normal)
### alpha: weight on the lasso proportion of the penalty
### nlambda: number of lasso/L1 tuning parameter values (lambda)
### FDR.level: desired level at which to control the FDR
### Strong recommendations:
### Give meaningful names to the columns of X.D.
### Name the columns of X.SNP with the corresponding SNP names (rsxxx).
### p.SNP should be a large value (at least several thousand).
### alpha should be in the range of 0.5 to 0.7.
### nlambda should be at least 100, we recommend 1000
### (required for achieving FDR control near the desired level)
### (and small nlambda is not computationally advantageous).
### FDR control method is conservative, we therefore recommend
### setting FDR.level to 0.05.
### Additional information on matrix X.D:
### If your design includes factors in addition to covariates,
### please create matrix X.D by using the model.matrix function
### of the stats R package as shown below:
### X.D <- model.matrix(~chip+site+age+pc1)[,-1]
### where chip and site are defined as factors and age and pc1
### are covariates. E.g.,
### chip <- as.factor(chip), chip <- factor(chip,labels=temp)
### In the penalized regression, each column of X.D will be...
## EN.R

### treated as a covariate to which no shrinkage will be applied.
### Requires:
### current version of R package glmnet
### Output:  EN.results <- EN.FDR(X.D,X.SNP,y,alpha,nlambda,FDR.level)
### EN.results$alpha: alpha as provided in argument list
### EN.results$nlambda: lambda as provided in argument list
### EN.results$FDR.level: FDR.level as provided in argument list
### EN.results$n.SNP: number of SNPs selected (nonzero coefficients)
### EN.results$which.SNP: identifiers of selected SNPs
### EN.results$betas.SNP: coefficient estimates of selected SNPs
### EN.results$FDR.achieved: nominal level of FDR control (should
### be just below the desired level)
### EN.results$check.FDR: should be zero (if set to 1 then FDR
### control was not achieved - something went wrong!)

```r
EN.FDR <- function(X.D,X.SNP,y,alpha,nlambda,FDR.level) {
  library(glmnet)
  if(is.null(X.D)) {
    X.D.s <- NULL
    n.fix <- 0
  } else {
    X.D.s <- standardize(X.D)
    n.fix <- dim(X.D.s)[2]
  }
  X.SNP.s <- standardize(X.SNP)
  scale <- attr(X.SNP.s, "scale")
  nz <- which(scale > 1e-6)
  X.SNP.s <- X.SNP.s[,nz,drop=FALSE]
  n.Xsnp <- dim(X.SNP.s)[2]
  X <- cbind(X.D.s,X.SNP.s)
  penalty.factor <- c(rep(0,n.fix),rep(1,n.Xsnp))
  idx <- seq(1,nlambda)
  # perform EN penalized regression on grid of lambda values
  EN.fit <-
  glmnet(X,y,penalty.factor=penalty.factor,alpha=alpha,nlambda=nlambda)
  lseq <- EN.fit$lambda
  beta <- as.matrix(EN.fit$beta)
  if(is.null(X.D)) {
    beta.SNP <- beta
  } else {
    beta.SNP <- beta[-(1:n.fix),]
  }
  c.nzero <- colSums(beta.SNP!=0)
  R.vec <- rep(0,length(lseq))
  R.vec[which(c.nzero!=0)] <- 1/c.nzero[which(c.nzero!=0)]
  # analytic FDR calculation
  yy <- y - mean(y)
  sigk <- sqrt( colSums( (}
matrix(rep(yy,length(lseq)),byrow=F,length(y))-X%*%beta )^2 ) )
FDR <- 2*n.Xsnp*pnorm(q=(-length(y)*lseq*alpha/sigk),mean=0,sd=1,lower.tail=TRUE,log.p=FALSE)*R.vec
# find lambda value closest from below to the desired FDR level
diff <- FDR.level - FDR
### this way:
  lim <- min( (sum(diff>0)+10),nlambda )
diff <- abs(diff[1:lim])
lbest <- which.min(diff)[1]
### or this way (ensures the selected lambda has FDR <= FDR.level):
diff <- diff[diff>0]
  idx <- idx[diff>0]
  lbest <- idx[which.min(diff)[1]]
### end lbest

# output:
check.FDR <- 0
if(FDR[length(FDR)]<FDR.level) check.FDR <- 1 #should never happen
which.SNP <- which( beta.SNP[,lbest] !=0 )
betas.SNP <- beta.SNP[which.SNP,lbest]
which.SNP <- rownames(beta.SNP)[which.SNP]
n.SNP <- length(which.SNP)
FDR.achieved <- FDR.level - diff[lbest]
list(alpha=alpha,nlambda=nlambda,FDR.level=FDR.level,n.SNP=n.SNP,
     which.SNP=which.SNP,betas.SNP=betas.SNP,FDR.achieved=FDR.achieved,
     check.FDR=check.FDR)