

SigMod: an exact and efficient method to identify a strongly interconnected disease-associated module in a gene network

Yuanlong Liu, Myriam Brossard, Damian Roqueiro, Patricia Margaritte-Jeannin, Chloé Sarnowski, Emmanuelle Bouzigon, Florence Demenais

1 Supplementary Figures

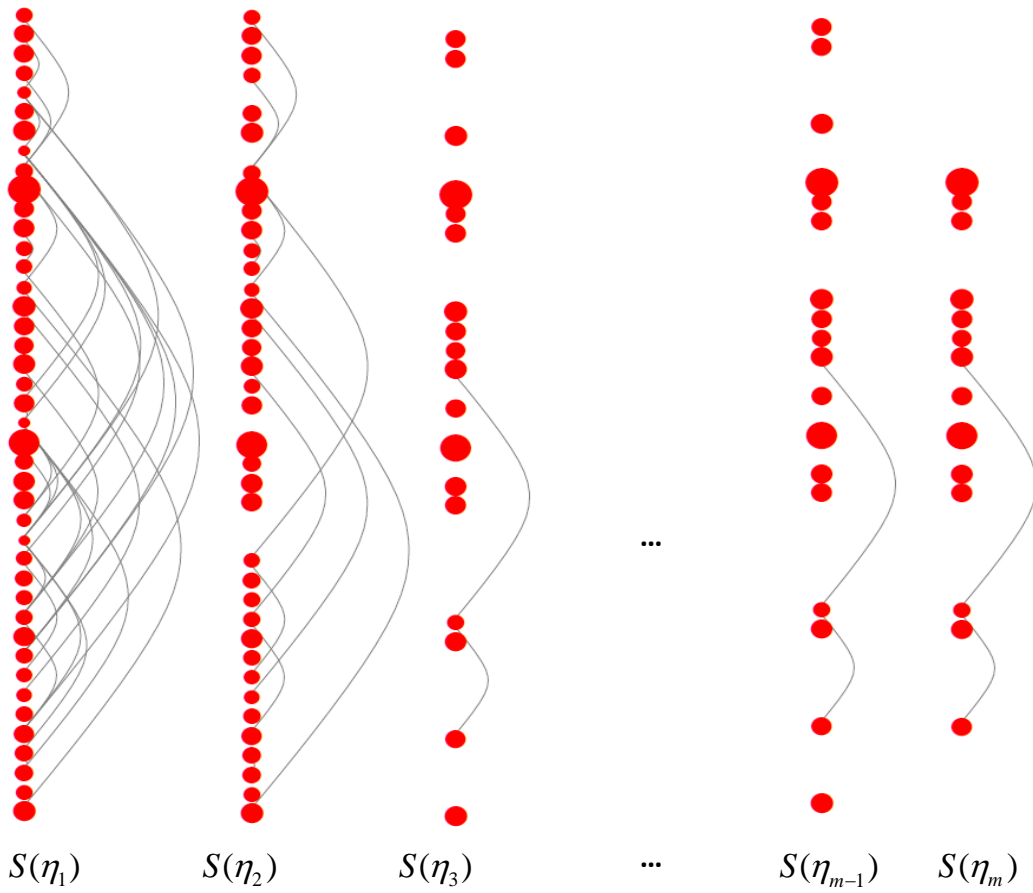


Figure S1: The selection path at a given λ value over the sparsity range $[\eta_{\min}, \eta_{\max}]$, defined as $\mathcal{P} = \langle S(\eta_1), \dots, S(\eta_m) \rangle$. Here $S(\eta_1), \dots, S(\eta_m)$ are the sequence of distinct modules selected by moving η from η_{\min} to η_{\max} , and η_1, \dots, η_m are some sparsity values leading to these distinct selections ($\eta_{\min} \leq \eta_1 < \dots < \eta_m \leq \eta_{\max}$). Red nodes represent genes and the curved lines are their connections. The *nesting property* is reflected by $S(\eta_1) \supseteq \dots \supseteq S(\eta_m)$.

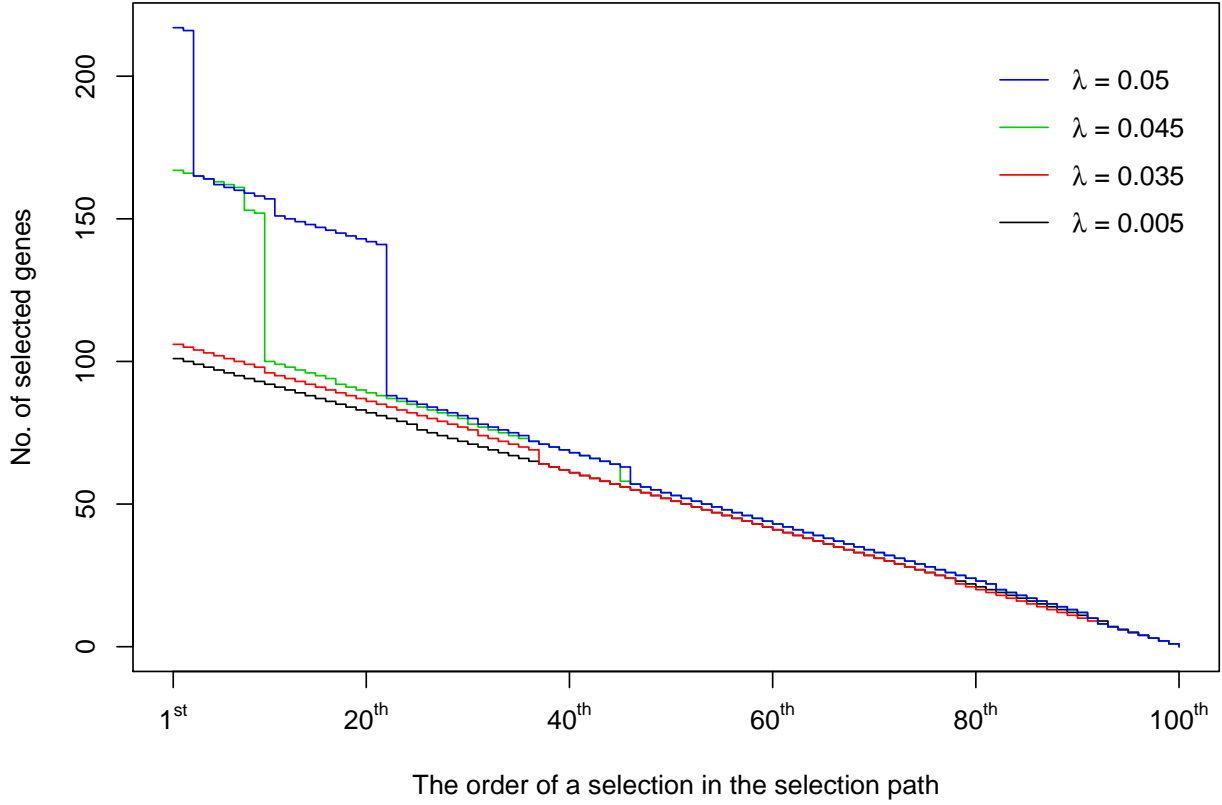


Figure S2: The number of genes in each module (S) from the selection path (\mathcal{P}). Four paths corresponding to four λ values are shown. The x-axis represents the order of each module in the path. The y-axis represents the number of genes in a module. Each path contains 100 distinct modules which end with an empty module (no gene selected). In a given path, a big size jump between two consecutive modules indicates some strongly interconnected genes are non-separable and are selected together. For larger λ values, the jumps are usually bigger, such as for $\lambda = 0.05$ (blue line).

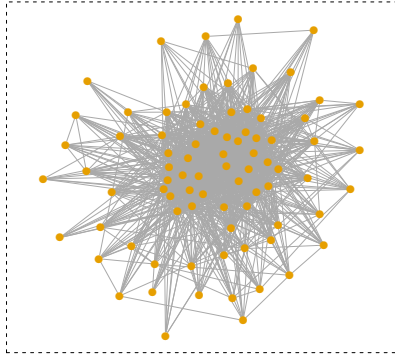
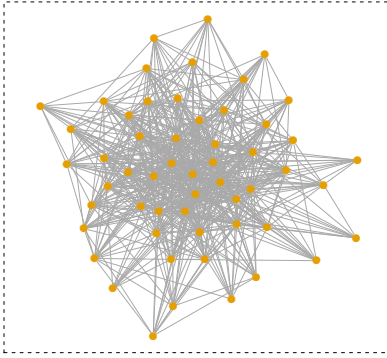
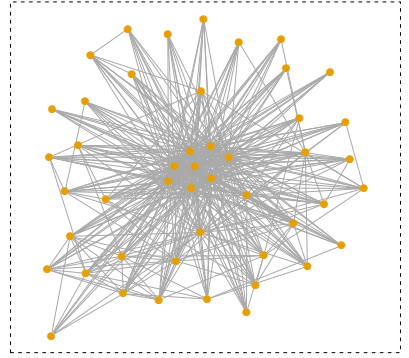
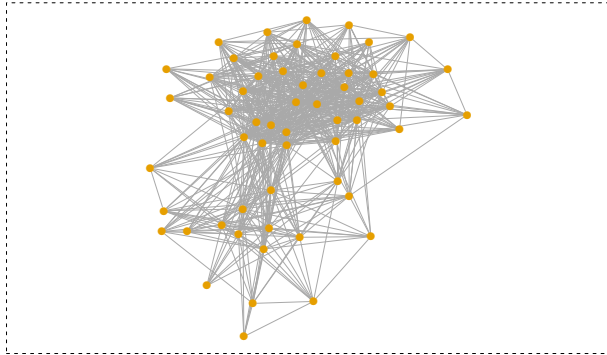
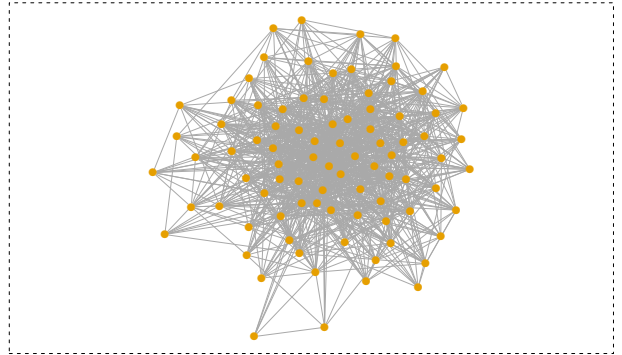
Module 1. Number of genes: 81**Module 2. Number of genes: 55****Module 3. Number of genes: 47****Module 4. Number of genes: 57****Module 5. Number of genes: 87**

Figure S3: Five strongly interconnected gene modules identified by CFinder were chosen as causal module for simulation study.

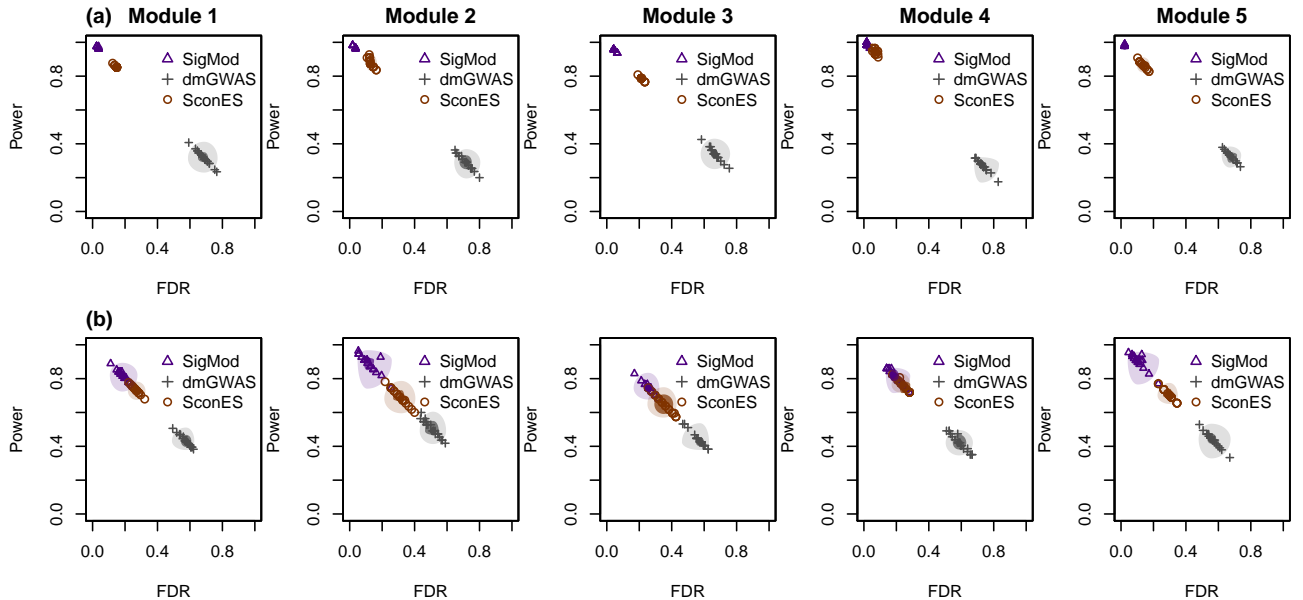


Figure S4: False discovery rate (FDR) versus power of three network analysis methods (SigMod, dmGWAS and SConES) applied to simulated data. In each experiment, one of the five modules identified by CFinder was set as the causal module. Plot (a) shows the results without adding noise to the GWAS data or GeneNet. Plot (b) shows the results with noise added to both gene-level P -values and gene network data.

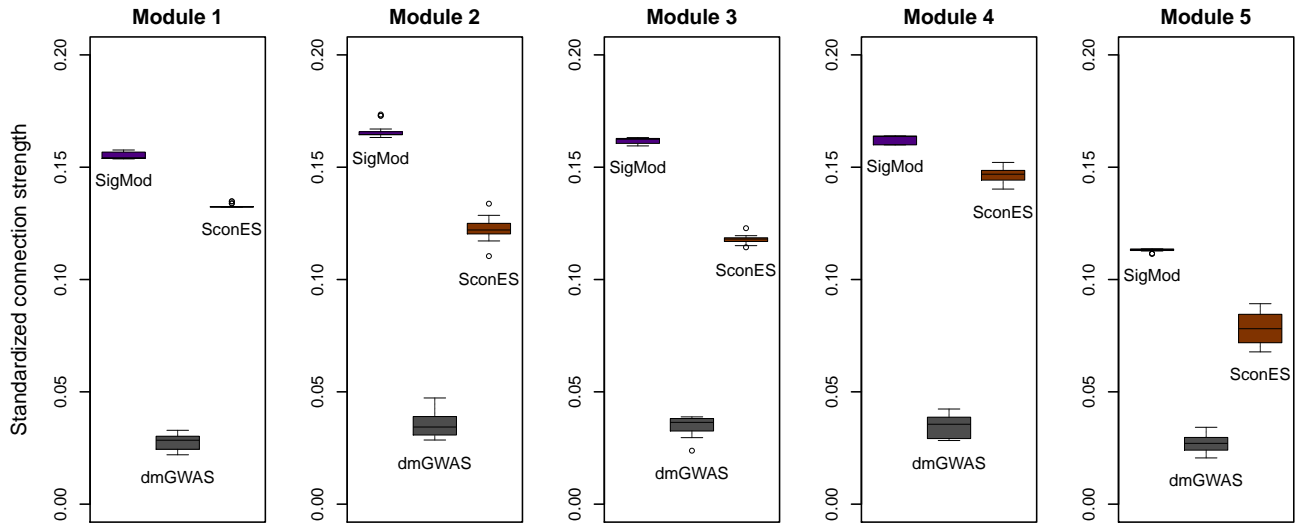


Figure S5: Box plots of the standardized connection strength (ρ) of gene modules identified by three network analysis methods (SigMod, SConES and dmGWAS). In each experiment, one of the five modules identified by CFinder was set as the causal module.

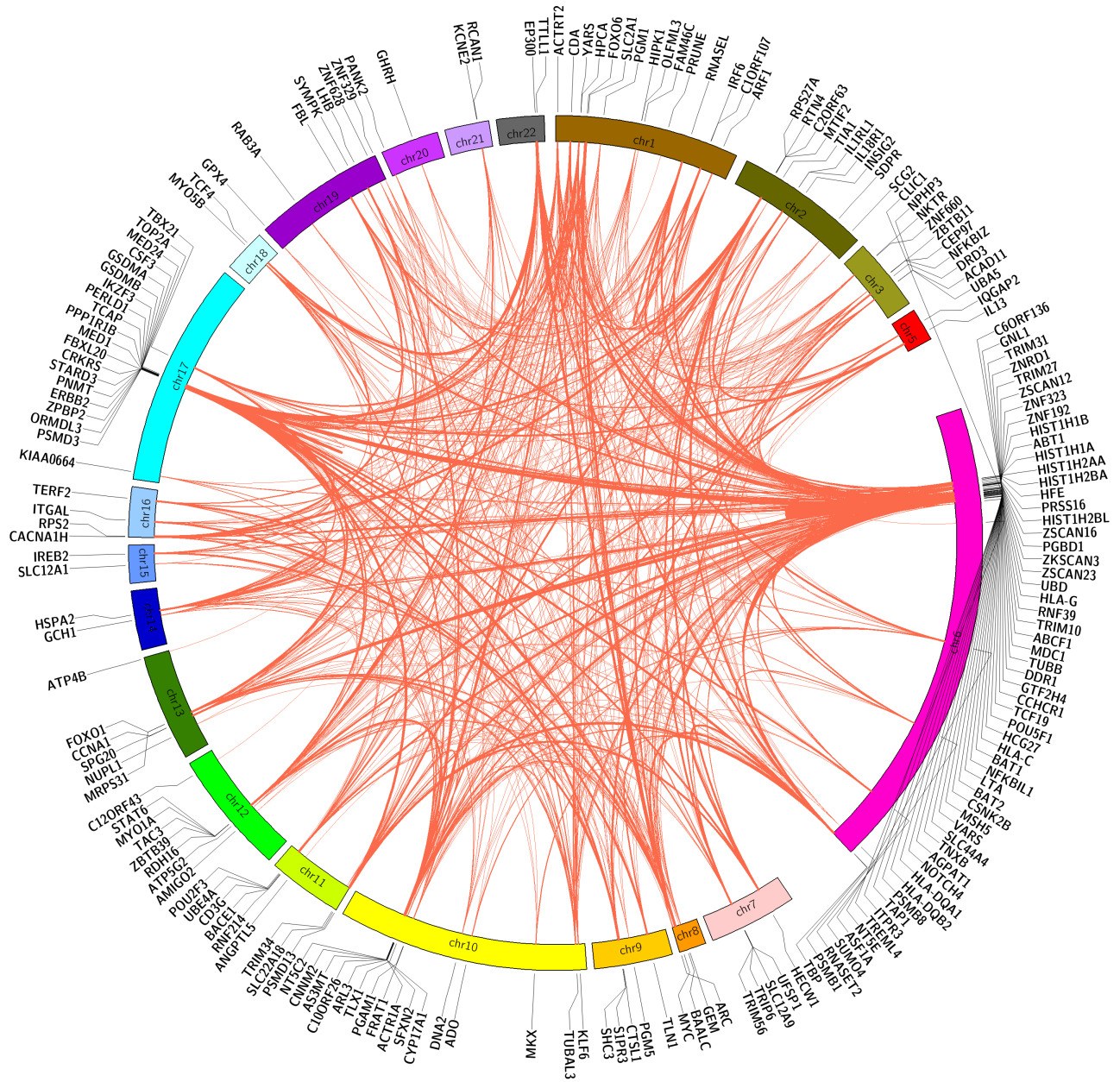


Figure S6: The gene module identified by applying SigMod to asthma META1 dataset. This module contains 190 genes and 1295 connections. In the figure the edge widths are proportional to edge weights, with smallest weight of 0.151 and largest weight of 0.999. Chromosome lengths have been rescaled for better layout.

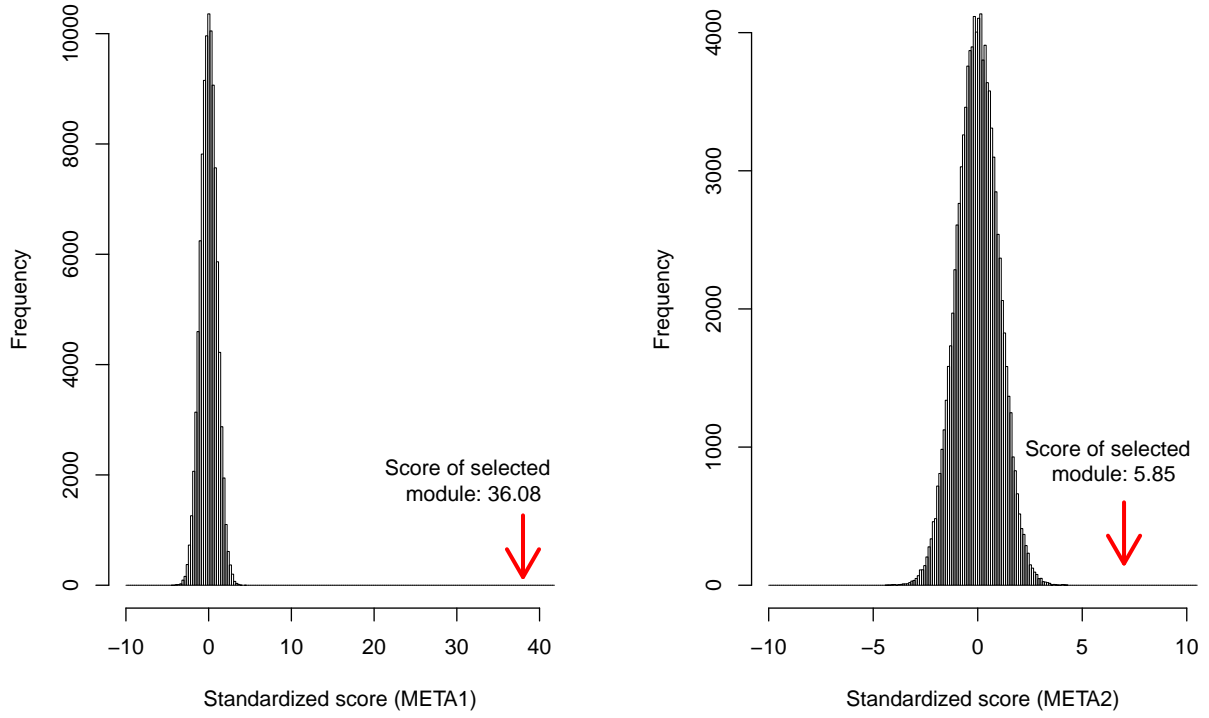


Figure S7: Histogram of the standardized module scores z^* of 100,000 random modules sampled from GeneNet (each module has 190 genes). The selected gene module has a significantly higher score than the random modules ($P < 10^{-5}$), evaluated using META1 and META2 dataset respectively.

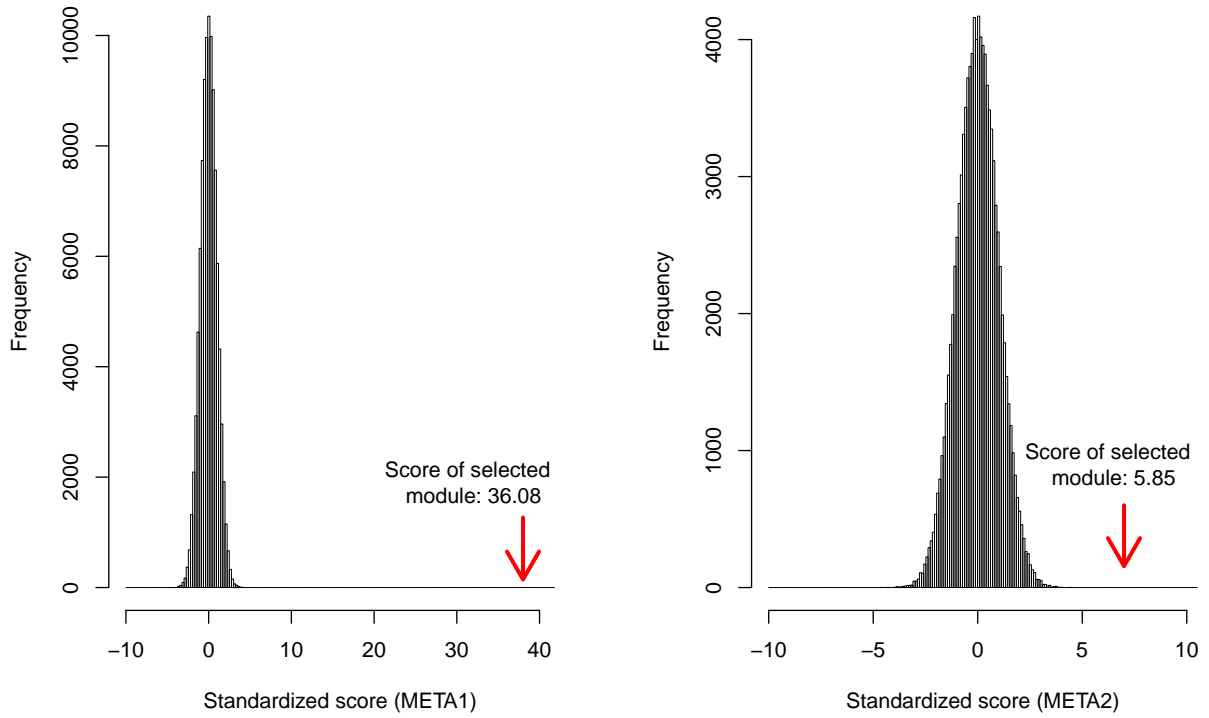


Figure S8: Histogram of the scores of selected gene module based on 100,000 random permutations of SNP-level statistics using Circular Genomic Permutation (CGP). The observed module score is significantly higher than those obtained from the permutation samples ($P < 10^{-5}$), evaluated using META1 and META2 dataset respectively.

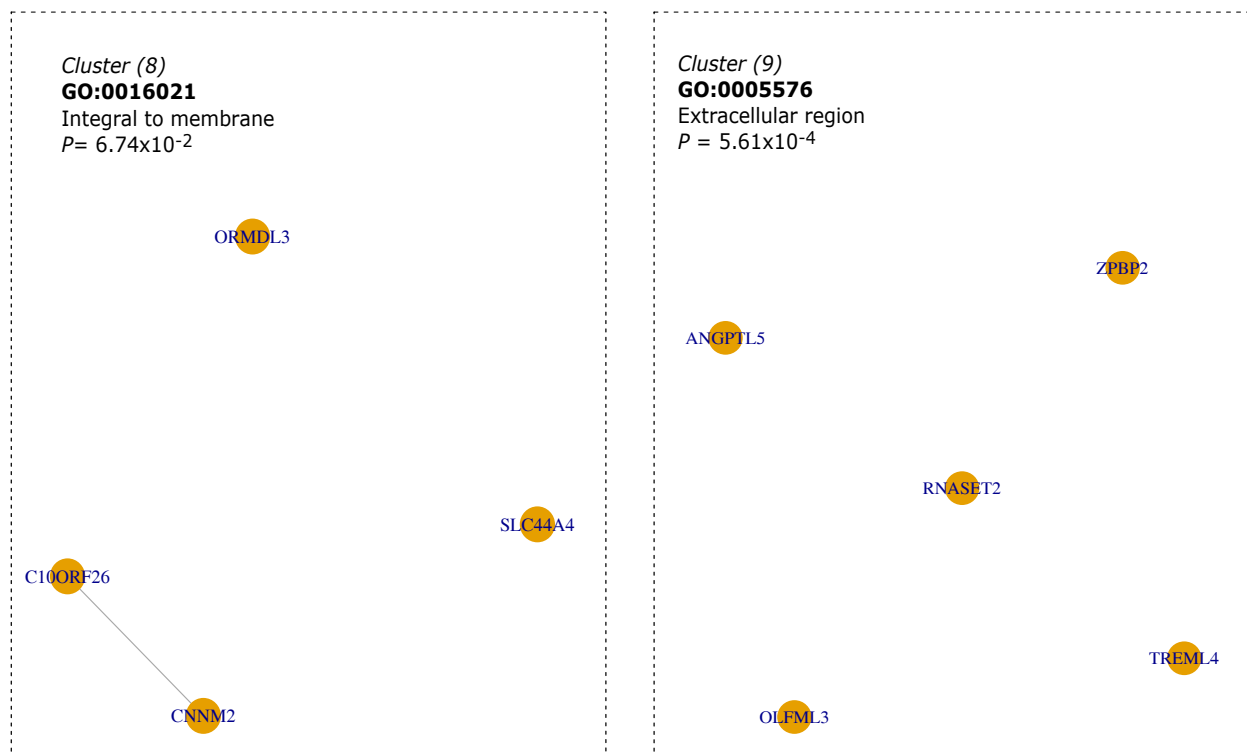


Figure S9: Two additional loosely connected or unconnected functional gene clusters out of nine gene clusters identified by DAVID in the selected gene module. The main function of each cluster is annotated using the gene ontology (GO) category with highest enrichment in these genes. P -values indicate the significance of enrichment of each GO category.

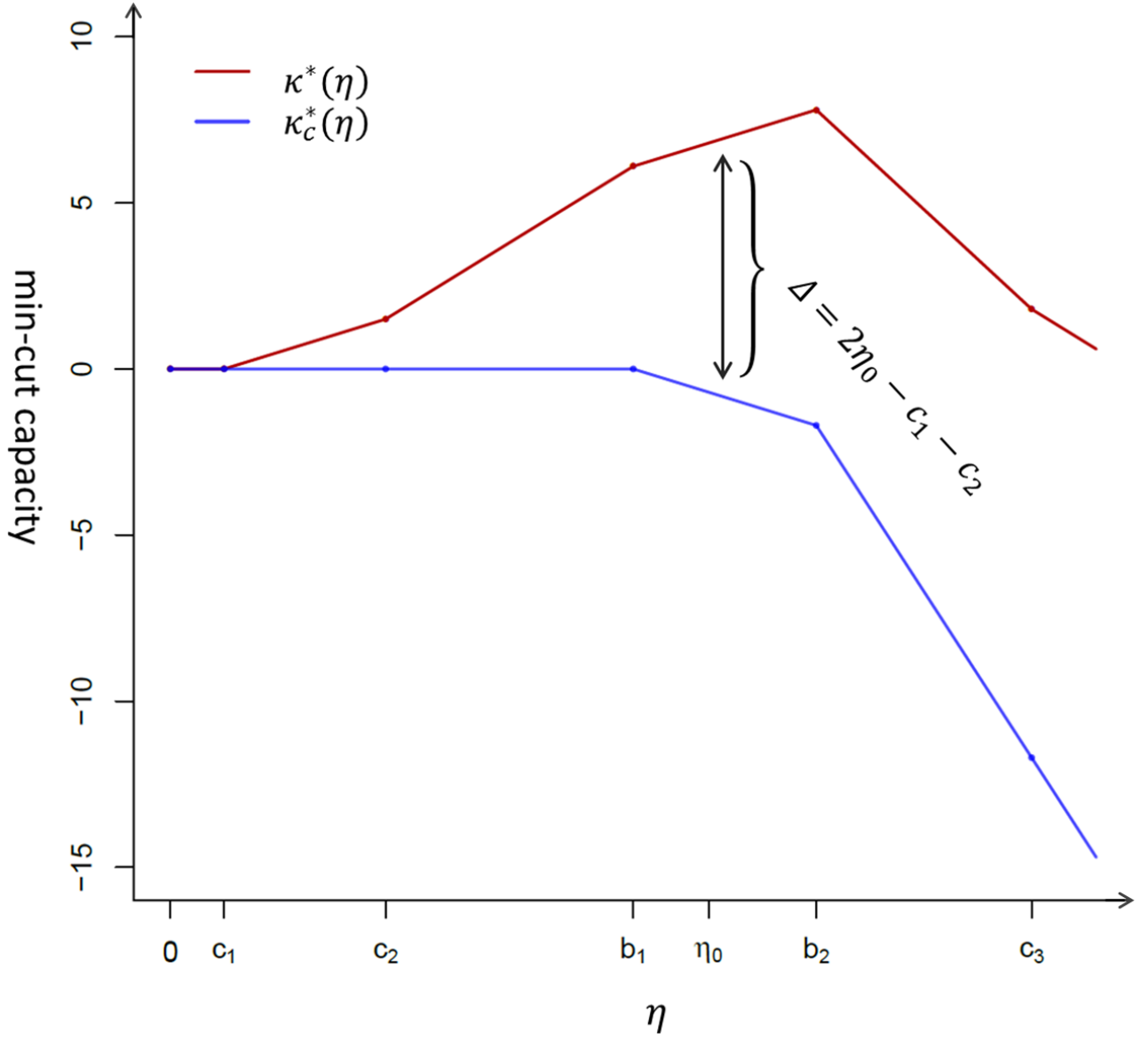


Figure S10: An example of *min-cut* capacity curve $\kappa^*(\eta)$ and its corrected curve $\kappa_c^*(\eta)$. $\kappa^*(\eta)$ has three change-points $\{c_1, c_2, c_3\}$ and two break-points $\{b_1, b_2\}$. It is piece-wise linear and concave between any consecutive change-points. The corrected curve $\kappa_c^*(\eta)$ is concave throughout its domain. The correction at η_0 is $\Delta = 2\eta_0 - c_1 - c_2$ according to Equation (S2).

2 Supplementary Tables

Table S1: Association statistics of 190 genes in the selected gene module for childhood asthma. $P_{\text{best-SNP}}$ is the best SNP P -value among all SNPs mapped to a gene. Gene P -value is the best SNP P -value corrected for gene length using CGP-permutation. “Rank” is the rank of a gene P -value among all genes mapped to the GeneNet evaluated using the META1 discovery dataset. The genes having significant P -values ($P \leq 0.05$) in both META1 and META2 datasets are shown in bold.

No.	Chr.	Gene	META1 (Discovery data)				META2 (Replication data)		
			best SNP	best SNP position	$P_{\text{best-SNP}}$	Gene P -value	$P_{\text{best-SNP}}$	Gene P -value	Rank
1	1	<i>FAM46C</i>	rs17037485	118162124	4.30E-04	1.45E-01	7.35E-03	6.80E-01	272
2	1	<i>HIPK1</i>	rs7545300	114503636	2.46E-04	5.32E-02	3.41E-03	2.64E-01	165
3	1	<i>OLFML3</i>	rs7364	114524661	3.66E-04	8.76E-01	9.63E-04	9.52E-01	58
4	1	<i>PGM1</i>	rs855311	64091739	3.87E-04	7.21E-03	8.28E-03	8.09E-02	294
5	1	<i>FOXO6</i>	rs7539614	41831047	1.22E-03	1.76E-04	7.89E-03	8.90E-04	284
6	1	<i>SLC2A1</i>	rs841847	43402708	8.64E-04	3.67E-02	1.30E-02	2.64E-01	422
7	1	<i>YARS</i>	rs2282294	33246177	2.59E-03	5.08E-01	1.67E-02	9.36E-01	524
8	1	<i>HPCA</i>	rs1284371	33352771	4.96E-03	7.86E-01	1.06E-02	9.39E-01	351
9	1	<i>CDA</i>	rs818202	20916791	3.63E-04	1.02E-01	3.91E-03	4.02E-01	178
10	1	<i>ACTRT2</i>	rs3795262	2938697	2.18E-03	3.06E-01	5.34E-03	4.76E-01	215
11	1	<i>PRUNE</i>	rs4970989	151003600	9.11E-04	1.24E-02	8.47E-03	6.48E-02	301
12	1	<i>RNASEL</i>	rs486907	182554557	2.15E-03	4.14E-02	1.19E-02	1.41E-01	387
13	1	<i>C1ORF107</i>	rs126280	210019824	3.31E-04	4.35E-01	5.72E-03	9.70E-01	227
14	1	<i>IRF6</i>	rs674433	209964875	1.14E-03	2.84E-01	1.07E-02	8.03E-01	353
15	1	<i>ARF1</i>	rs1188975	228277690	1.13E-03	1.49E-01	6.59E-03	4.17E-01	254
16	2	<i>TIA1</i>	rs2706769	70473453	2.93E-04	2.39E-01	3.05E-03	6.82E-01	157
17	2	<i>RPS27A</i>	rs2028139	55460017	9.54E-04	6.39E-01	3.26E-03	8.82E-01	161
18	2	<i>C2ORF63</i>	rs13032294	55404883	4.16E-04	5.01E-01	7.71E-03	9.90E-01	282
19	2	<i>MTIF2</i>	rs2043712	55480022	3.71E-04	3.29E-01	3.84E-03	8.24E-01	177
20	2	<i>RTN4</i>	rs11677099	55254165	4.09E-05	4.25E-02	2.12E-03	4.27E-01	124
21	2	<i>IL18R1</i>	rs3771166	102986222	1.87E-05	2.36E-05	8.20E-04	4.08E-04	52
22	2	<i>IL1RL1</i>	rs4988957	102968075	2.86E-05	2.17E-05	1.49E-03	5.63E-04	84
23	2	<i>INSIG2</i>	rs889904	118860471	9.40E-04	2.23E-01	8.66E-03	6.95E-01	309
24	2	<i>SDPR</i>	rs4853645	192704044	1.07E-03	2.50E-01	5.50E-03	5.73E-01	220
25	2	<i>SCG2</i>	rs2894511	224465827	9.31E-03	8.30E-01	8.56E-03	8.19E-01	305
26	3	<i>NFKBIZ</i>	rs604521	101550578	1.26E-03	1.17E-01	7.62E-03	3.55E-01	279
27	3	<i>ZNF660</i>	rs939649	44632212	3.27E-03	3.90E-02	7.55E-03	6.46E-02	277
28	3	<i>NKTR</i>	rs1062051	42672486	7.97E-04	1.01E-01	9.94E-03	4.88E-01	333
29	3	<i>CEP97</i>	rs2554962	101445118	7.15E-04	8.28E-02	4.25E-03	2.42E-01	187
30	3	<i>ZBTB11</i>	rs4683854	101385261	1.18E-03	7.80E-01	4.97E-03	9.77E-01	207
31	3	<i>DRD3</i>	rs324022	113887298	1.18E-04	5.06E-02	2.86E-03	3.57E-01	151
32	3	<i>NPHP3</i>	rs11708051	132410648	1.14E-03	5.87E-02	1.07E-02	2.74E-01	356
33	3	<i>UBA5</i>	rs1378807	132385190	1.27E-03	1.85E-01	6.34E-03	4.54E-01	247
34	3	<i>ACAD11</i>	rs2305627	132346992	2.84E-04	1.57E-01	4.12E-03	6.26E-01	183
35	5	<i>IQGAP2</i>	rs10514071	75994211	4.53E-05	1.54E-02	6.96E-03	5.35E-01	264
36	5	<i>IL13</i>	rs848	131996500	2.71E-04	1.02E-03	1.44E-03	2.44E-03	76
37	6	<i>PRSS16</i>	rs9393795	27217719	4.17E-04	3.34E-02	1.41E-03	5.54E-02	75
38	6	<i>ZNF192</i>	rs13205911	28124114	6.42E-05	7.37E-03	1.19E-03	4.14E-02	67
39	6	<i>ZNRD1</i>	rs8321	30032522	1.07E-03	1.89E-01	7.07E-03	5.34E-01	266
40	6	<i>ZSCAN12</i>	rs2232423	28366151	4.08E-04	9.53E-03	6.39E-03	7.38E-02	249
41	6	<i>TREML4</i>	rs7774363	41198145	2.36E-04	5.65E-02	4.30E-03	3.49E-01	188
42	6	<i>ABCF1</i>	rs3132610	30544401	7.59E-04	6.93E-01	5.32E-03	9.82E-01	213
43	6	<i>AGPAT1</i>	rs2269423	32145707	1.62E-04	1.28E-01	1.53E-03	3.82E-01	86
44	6	<i>BAT1</i>	rs2734583	31505480	3.56E-05	2.62E-01	9.48E-04	8.19E-01	55
45	6	<i>BAT2</i>	rs3132450	31596138	7.72E-04	6.16E-03	8.60E-03	4.01E-02	306
46	6	<i>C6ORF136</i>	rs9262135	30618906	4.69E-04	7.70E-01	6.79E-04	7.55E-01	44
47	6	<i>CCHCR1</i>	rs130073	31111180	3.05E-04	8.34E-02	6.46E-03	5.38E-01	252

48	6	<i>CLIC1</i>	rs3131383	31704294	5.20E-04	3.68E-01	1.23E-03	4.76E-01	70
49	6	<i>CSNK2B</i>	rs9267531	31636742	5.29E-04	2.83E-01	2.31E-03	5.39E-01	129
50	6	<i>DDR1</i>	rs1049633	30867527	4.07E-05	1.71E-01	8.85E-04	6.01E-01	54
51	6	<i>GNL1</i>	rs3130247	30515043	9.78E-04	7.82E-01	3.83E-03	9.70E-01	175
52	6	<i>GTF2H4</i>	rs1264308	30879987	5.96E-05	1.55E-01	7.56E-04	4.31E-01	46
53	6	<i>HCG27</i>	rs3132511	31170020	3.92E-04	6.01E-02	4.52E-03	2.79E-01	195
54	6	HFE	rs1045537	26096748	7.39E-04	7.24E-03	4.92E-03	2.73E-02	206
55	6	<i>HLA-C</i>	rs2001181	31236998	5.32E-04	3.91E-01	4.73E-03	8.67E-01	205
56	6	<i>HLA-DQA1</i>	rs9272853	32610705	3.09E-05	2.31E-01	3.04E-04	4.57E-01	23
57	6	<i>HLA-DQB2</i>	rs9276584	32730835	6.99E-04	3.83E-02	8.24E-03	2.16E-01	293
58	6	<i>HLA-G</i>	rs1610696	29798803	1.25E-03	6.60E-01	8.39E-03	9.78E-01	297
59	6	<i>ITPR3</i>	rs3736893	33639760	2.51E-05	2.02E-01	1.44E-03	9.22E-01	78
60	6	<i>LTA</i>	rs909253	31540313	1.52E-03	1.40E-01	4.71E-03	2.64E-01	203
61	6	<i>MDC1</i>	rs3094093	30679628	2.67E-03	3.46E-01	9.87E-03	6.62E-01	330
62	6	<i>MSH5</i>	rs3117574	31725230	4.35E-04	7.34E-02	4.72E-03	3.16E-01	204
63	6	<i>NFKBIL1</i>	rs2071592	31515340	7.20E-05	4.85E-02	1.04E-03	1.85E-01	60
64	6	<i>NOTCH4</i>	rs436388	32186264	2.40E-04	1.07E-01	6.29E-03	6.75E-01	245
65	6	<i>POU5F1</i>	rs13409	31132140	6.15E-04	3.29E-02	6.05E-03	1.58E-01	237
66	6	<i>PSMB8</i>	rs9276810	32810443	4.43E-04	4.09E-01	2.55E-03	7.66E-01	142
67	6	<i>RNF39</i>	rs9261290	30038647	1.03E-03	3.14E-01	5.75E-03	6.92E-01	229
68	6	<i>SLC44A4</i>	rs9267659	31846234	1.69E-04	1.30E-01	2.00E-03	4.60E-01	117
69	6	<i>TAP1</i>	rs6457684	32814447	1.09E-04	3.08E-01	1.81E-03	8.37E-01	101
70	6	<i>TCF19</i>	rs7750641	31129310	2.14E-05	3.52E-02	4.99E-04	1.59E-01	31
71	6	<i>TNXB</i>	rs3130286	32042322	3.28E-04	2.53E-01	6.68E-03	8.91E-01	259
72	6	<i>TRIM10</i>	rs2523735	30122657	4.04E-04	3.24E-02	5.17E-03	1.79E-01	211
73	6	<i>TRIM31</i>	rs1116221	30071330	6.70E-04	3.52E-02	8.10E-03	2.05E-01	290
74	6	<i>TUBB</i>	rs8233	30692965	8.13E-04	3.59E-01	2.87E-03	6.03E-01	152
75	6	<i>UBD</i>	rs404240	29523957	4.22E-04	4.36E-01	2.68E-03	8.14E-01	148
76	6	<i>VAR5</i>	rs915652	31749142	3.83E-04	1.66E-01	2.46E-03	4.16E-01	133
77	6	<i>ABT1</i>	rs4634439	26598004	1.69E-03	9.26E-02	7.37E-03	2.36E-01	273
78	6	<i>ZNF323</i>	rs13217619	28306671	3.45E-04	1.22E-02	6.25E-03	1.03E-01	244
79	6	<i>ZSCAN16</i>	rs4713140	28097193	2.09E-03	6.35E-02	9.56E-03	1.79E-01	324
80	6	<i>ZSCAN23</i>	rs7766356	28400538	3.06E-04	3.15E-02	2.27E-03	9.84E-02	128
81	6	<i>TRIM27</i>	rs3135293	28877247	4.96E-05	1.60E-01	7.80E-04	4.87E-01	50
82	6	<i>HIST1H1A</i>	rs16891235	26017542	4.19E-04	2.46E-01	6.34E-04	2.17E-01	35
83	6	<i>HIST1H1B</i>	rs17763089	27835218	6.04E-04	6.24E-01	1.38E-03	7.58E-01	72
84	6	<i>HIST1H2BL</i>	rs200485	27775697	3.76E-03	7.01E-01	6.42E-03	8.28E-01	250
85	6	<i>PGBD1</i>	rs13211507	28257377	2.51E-04	1.66E-02	3.69E-03	1.01E-01	171
86	6	<i>ZKSCAN3</i>	rs6921919	28325201	5.77E-04	1.22E-02	6.43E-03	7.14E-02	251
87	6	<i>HIST1H2AA</i>	rs4711095	25726774	4.69E-03	4.21E-01	1.00E-02	6.25E-01	336
88	6	<i>HIST1H2BA</i>	rs9358872	25727517	4.18E-03	5.56E-01	1.11E-02	8.18E-01	366
89	6	<i>NT5E</i>	rs3812139	86196990	2.62E-04	5.67E-02	4.34E-03	3.26E-01	191
90	6	<i>ASF1A</i>	rs7772912	119218470	2.63E-03	1.73E-01	1.25E-02	4.52E-01	406
91	6	<i>SUMO4</i>	rs237025	149721690	5.64E-03	4.78E-01	5.44E-03	4.47E-01	219
92	6	<i>PSMB1</i>	rs12207633	170847181	2.53E-04	7.94E-02	4.16E-03	4.15E-01	185
93	6	<i>RNASET2</i>	rs1077453	167360024	4.28E-04	1.23E-01	5.33E-03	5.15E-01	214
94	6	<i>TBP</i>	rs12200657	170872108	1.34E-03	8.92E-02	9.76E-03	3.25E-01	325
95	7	<i>HECW1</i>	rs2330785	43276428	3.90E-05	2.26E-02	7.58E-03	7.27E-01	278
96	7	<i>TRIM56</i>	rs6948536	100731829	8.86E-04	6.55E-01	1.91E-03	7.87E-01	110
97	7	<i>SLC12A9</i>	rs314370	100453208	2.93E-04	2.25E-02	2.46E-03	7.99E-02	132
98	7	TRIP6	rs6706	100471044	3.32E-04	1.88E-02	1.20E-03	3.11E-02	68
99	7	<i>UFSP1</i>	rs12666989	100486754	2.02E-04	3.55E-01	6.35E-04	4.60E-01	36
100	8	<i>GEM</i>	rs1050616	95262253	9.92E-05	5.39E-01	1.40E-03	9.54E-01	74
101	8	<i>BAALC</i>	rs2454014	104154965	2.34E-04	1.97E-02	8.87E-03	2.86E-01	315
102	8	<i>MYC</i>	rs4645956	128750212	7.80E-03	5.32E-02	1.89E-02	1.04E-01	581
103	8	<i>ARC</i>	rs10097505	143694184	3.65E-03	6.81E-02	3.78E-03	5.54E-02	173
104	9	<i>TLN1</i>	rs10972567	35728019	1.42E-03	4.26E-02	1.10E-02	1.82E-01	363
105	9	<i>PGM5</i>	rs7874438	71114809	1.51E-04	7.30E-02	6.17E-03	6.58E-01	240
106	9	CTSL1	rs2378757	90343780	1.84E-03	6.89E-03	5.58E-03	1.37E-02	222

107	9	<i>S1PR3</i>	rs7858626	91612639	5.42E-06	1.50E-01	2.18E-04	4.03E-01	19
108	9	<i>SHC3</i>	rs1331180	91630548	1.98E-05	3.57E-02	1.63E-03	4.86E-01	91
109	10	<i>MKX</i>	rs2637277	27969032	3.54E-04	7.69E-03	8.82E-03	9.77E-02	314
110	10	<i>KLF6</i>	rs17731	3821561	1.24E-03	2.48E-01	4.64E-03	4.85E-01	200
111	10	<i>TUBAL3</i>	rs7097775	5435918	6.60E-04	3.12E-01	6.00E-03	8.04E-01	235
112	10	<i>ADO</i>	rs9990	64567938	3.35E-03	3.13E-01	9.24E-03	5.41E-01	317
113	10	<i>DNA2</i>	rs10998202	70224310	4.06E-04	3.75E-02	6.95E-03	2.64E-01	263
114	10	<i>ARL3</i>	rs2251772	104470918	1.52E-04	2.94E-01	1.74E-03	7.53E-01	98
115	10	<i>ACTR1A</i>	rs3781290	104244948	7.32E-04	3.68E-01	6.99E-03	8.76E-01	265
116	10	<i>FRAT1</i>	rs3781373	99080585	1.38E-03	6.45E-01	2.73E-03	7.77E-01	149
117	10	<i>CYP17A1</i>	rs10786712	104596396	4.21E-05	3.61E-01	7.25E-04	8.11E-01	45
118	10	<i>C10ORF26</i>	rs284858	104573936	6.52E-05	7.40E-02	2.59E-03	5.86E-01	145
119	10	<i>CNNM2</i>	rs943036	104836047	2.34E-05	6.52E-02	1.76E-03	6.72E-01	99
120	10	<i>NT5C2</i>	rs746293	104897254	1.78E-05	5.67E-02	9.50E-04	4.69E-01	56
121	10	<i>SFXN2</i>	rs1475644	104491164	2.21E-04	1.59E-01	2.26E-03	4.98E-01	127
122	10	<i>TLX1</i>	rs2235128	102896801	1.89E-03	5.71E-02	1.01E-02	1.81E-01	337
123	10	<i>AS3MT</i>	rs1591915	104646849	2.73E-05	3.42E-01	6.64E-04	8.71E-01	39
124	10	<i>PGAM1</i>	rs764223	99191935	9.68E-03	2.07E-01	8.87E-03	1.80E-01	316
125	11	<i>TRIM34</i>	rs3740998	5664831	4.94E-04	4.19E-02	7.55E-03	2.70E-01	276
126	11	<i>PSMD13</i>	rs577298	248002	9.13E-04	2.63E-03	1.35E-02	2.49E-02	430
127	11	<i>SLC22A18</i>	rs11024581	2939705	4.02E-04	1.48E-01	5.59E-03	6.17E-01	224
128	11	<i>ANGPTL5</i>	rs7109121	101766771	4.72E-04	1.48E-02	5.05E-03	7.64E-02	209
129	11	<i>CD3G</i>	rs3212264	118216234	3.83E-04	6.50E-01	1.83E-03	9.14E-01	104
130	11	<i>POU2F3</i>	rs7484249	120142118	7.26E-05	2.13E-02	2.38E-03	2.02E-01	131
131	11	<i>RNF214</i>	rs655023	117127620	1.09E-03	6.22E-02	9.43E-03	2.68E-01	322
132	11	<i>UBE4A</i>	rs12576486	118235490	2.42E-04	1.20E-01	3.12E-03	4.71E-01	159
133	11	<i>BACE1</i>	rs490460	117163765	1.01E-03	1.20E-01	8.31E-03	4.33E-01	295
134	12	<i>STAT6</i>	rs324015	57490100	8.73E-04	2.01E-02	5.35E-03	6.78E-02	216
135	12	<i>AMIGO2</i>	rs854889	47471037	4.42E-03	5.58E-01	7.31E-03	6.91E-01	270
136	12	<i>ATP5G2</i>	rs12422531	54067827	2.09E-03	4.83E-01	1.02E-02	8.70E-01	340
137	12	<i>RDH16</i>	rs901068	57346805	5.45E-04	2.77E-02	2.67E-03	6.94E-02	146
138	12	<i>TAC3</i>	rs733629	57406444	6.95E-04	7.30E-01	2.50E-03	9.36E-01	137
139	12	<i>ZBTB39</i>	rs4016338	57399015	3.46E-03	1.37E-03	9.48E-03	2.73E-03	323
140	12	<i>MYO1A</i>	rs17119344	57422934	7.69E-04	2.90E-01	6.63E-03	7.68E-01	257
141	12	<i>C12ORF43</i>	rs3751150	121442214	4.36E-04	1.25E-02	3.15E-03	4.30E-02	160
142	13	<i>NUPL1</i>	rs9551192	25904797	6.09E-04	7.89E-02	4.41E-03	2.65E-01	192
143	13	<i>CCNA1</i>	rs4245378	37007040	3.71E-03	2.62E-01	1.29E-02	5.40E-01	420
144	13	<i>SPG20</i>	rs9547247	36892264	2.99E-04	5.33E-01	6.17E-03	9.94E-01	241
145	13	<i>FOXO1</i>	rs7323267	41204015	5.74E-05	1.80E-01	2.55E-03	8.84E-01	143
146	13	<i>MRPS31</i>	rs9549281	41330171	1.73E-04	4.16E-01	1.59E-03	8.41E-01	90
147	13	<i>ATP4B</i>	rs11164142	114309226	5.49E-03	6.28E-01	8.79E-03	7.62E-01	313
148	14	<i>GCH1</i>	rs10498472	55354869	7.14E-04	1.21E-01	1.00E-02	5.86E-01	334
149	14	<i>HSPA2</i>	rs17101919	65007547	1.25E-02	3.02E-01	1.83E-02	3.94E-01	562
150	15	<i>SLC12A1</i>	rs11857986	48571605	2.74E-04	1.02E-01	6.60E-03	6.39E-01	255
151	15	<i>IREB2</i>	rs11630228	78736325	2.12E-04	4.52E-02	4.66E-03	3.39E-01	202
152	16	<i>ITGAL</i>	rs2230434	30518096	1.64E-04	2.28E-03	2.58E-03	1.58E-02	144
153	16	<i>CACNA1H</i>	rs4984637	1261282	1.02E-03	5.14E-01	7.44E-03	9.34E-01	275
154	16	<i>RPS2</i>	rs6366	2014031	1.38E-02	6.58E-01	1.23E-02	6.35E-01	397
155	16	<i>TERF2</i>	rs3785073	69401937	3.33E-03	9.11E-02	1.18E-02	2.16E-01	381
156	17	<i>KIAA0664</i>	rs11078312	2600186	9.92E-04	3.44E-02	5.59E-03	1.07E-01	223
157	17	<i>IKZF3</i>	rs907091	37921742	2.49E-15	1.55E-07	4.56E-05	7.09E-05	7
158	17	<i>CSF3</i>	rs25645	38173143	2.41E-06	1.19E-04	1.24E-04	4.50E-04	12
159	17	<i>ERBB2</i>	rs1058808	37884037	3.47E-06	1.62E-05	1.61E-04	8.19E-05	16
160	17	<i>STARD3</i>	rs11869286	37813856	4.57E-07	2.28E-04	9.08E-05	8.64E-04	9
161	17	<i>CRKRS</i>	rs11658678	37680096	1.91E-04	4.66E-02	2.54E-03	2.26E-01	141
162	17	<i>MED1</i>	rs10445306	37591422	4.30E-04	4.99E-02	2.50E-03	1.33E-01	136
163	17	<i>ORMDL3</i>	rs12603332	38082807	4.59E-15	7.31E-08	9.71E-06	7.18E-06	2
164	17	<i>PERLD1</i>	rs903502	37829604	5.32E-07	7.03E-05	1.55E-04	6.33E-04	15
165	17	<i>PNMT</i>	rs876493	37824545	1.34E-07	4.00E-05	3.08E-05	7.35E-05	6

166	17	PPP1R1B	rs2271309	37784990	2.20E-04	1.01E-03	1.07E-03	2.09E-03	62
167	17	TCAP	rs1053651	37822311	8.12E-06	3.65E-03	6.80E-05	2.68E-03	8
168	17	ZPBP2	rs12936231	38029120	1.03E-14	1.01E-07	2.24E-05	1.94E-05	4
169	17	FBXL20	rs8069451	37504933	4.12E-04	6.10E-02	7.84E-03	4.18E-01	283
170	17	GSDMA	rs7212938	38122680	2.42E-13	3.22E-07	2.66E-05	3.25E-05	5
171	17	GSDMB	rs9303281	38074046	2.59E-16	7.53E-08	5.49E-06	1.94E-05	1
172	17	MED24	rs12309	38175462	3.03E-06	1.20E-04	2.53E-04	9.25E-04	21
173	17	PSMD3	rs11655264	38138995	1.00E-07	6.93E-05	9.63E-05	5.80E-04	10
174	17	TOP2A	rs2586112	38572366	2.39E-04	7.33E-03	1.89E-03	2.44E-02	107
175	17	TBX21	rs12721470	45822579	1.91E-03	9.12E-03	1.14E-02	3.60E-02	372
176	18	MYO5B	rs7237973	47514403	6.74E-05	4.26E-04	1.21E-02	4.78E-02	389
177	18	TCF4	rs10515970	52980635	9.62E-05	2.95E-02	6.79E-03	5.24E-01	261
178	19	RAB3A	rs2271882	18309365	2.89E-03	1.38E-01	5.17E-03	1.79E-01	210
179	19	GPX4	rs713041	1106615	1.20E-03	4.01E-01	1.46E-03	3.70E-01	79
180	19	FBL	rs11083539	40325680	1.85E-03	3.51E-01	9.28E-03	7.41E-01	320
181	19	LHB	rs1800447	49519905	1.02E-02	5.45E-01	9.27E-03	5.17E-01	318
182	19	SYMPK	rs7258364	46356548	1.97E-04	8.16E-02	3.42E-03	4.42E-01	167
183	19	ZNF329	rs159667	58648359	2.79E-05	4.68E-01	4.01E-04	8.58E-01	28
184	19	ZNF628	rs4801677	55990975	3.54E-03	5.80E-01	8.04E-03	7.92E-01	289
185	20	PANK2	rs6052169	3896449	3.75E-04	2.88E-01	3.33E-03	7.33E-01	163
186	20	GHRH	rs4988492	35882698	6.86E-04	3.29E-01	1.54E-03	4.28E-01	88
187	21	KCNE2	rs1010668	35738158	4.47E-04	2.22E-01	1.48E-03	3.55E-01	83
188	21	RCAN1	rs1557270	35957254	2.52E-05	7.41E-02	1.96E-03	7.29E-01	113
189	22	TTLL1	rs5759126	43455665	3.11E-04	1.15E-01	6.81E-03	6.59E-01	262
190	22	EP300	rs1569857	41540934	8.82E-04	7.32E-02	1.29E-02	4.38E-01	419

Table S2: 15 KEGG pathways enriched in the selected gene module for childhood asthma

Enriched KEGG	<i>P</i> -value	Ratio
Epstein-Barr virus infection	3.28E-04	12/202
Herpes simplex infection	1.57E-03	10/185
Thyroid hormone signaling pathway	1.57E-03	8/118
HTLV-I infection	3.34E-03	11/258
Antigen processing and presentation	3.34E-03	6/77
Inflammatory bowel disease (IBD)	9.14E-03	5/65
Type I diabetes mellitus	1.15E-02	4/43
Proteasome	1.15E-02	4/44
Phagosome	1.50E-02	7/154
Viral carcinogenesis	1.78E-02	8/205
Viral myocarditis	2.42E-02	4/59
Glucagon signaling pathway	3.07E-02	5/101
Central carbon metabolism in cancer	3.20E-02	4/67
Allograft rejection	3.55E-02	3/38
Graft-versus-host disease	4.08E-02	3/41

Note: *P*-values are FDR adjusted; Ratio is the number of genes from the 190 module genes that map to the pathway divided by the total number of genes that map to the canonical pathway.

3 Supplementary Notes

3.1 Computing the selection path at any given λ value

In our formulation, a module is selected by solving the optimization problem

$$\arg \max_{\mathbf{u}} \mathbf{z}^\top \mathbf{u} + \lambda \mathbf{u}^\top \mathbf{A} \mathbf{u} - \eta \|\mathbf{u}\|_0. \quad (\text{S1})$$

Therefore the selection can vary for different values of λ and η . To conveniently trace the selection change, we develop the path algorithm that enables computing all distinct module selections at any fixed λ value. Specifically, we define the selection path of a given λ value over a sparsity range $[\eta_{\min}, \eta_{\max}]$ as the sequence of distinct modules selected by moving η from η_{\min} to η_{\max} , denoted as $\mathcal{P} = \langle S(\eta_1), \dots, S(\eta_m) \rangle$ (as described in the main paper). Note that the entire selection path can be obtained by setting $\eta_{\min} = 0$ and $\eta_{\max} = +\infty$.

To compute \mathcal{P} , we define the capacity function $\kappa^*(\eta)$ as the capacity of the s - t min-cut on G_{st} (s - t min-cut, capacity and G_{st} are defined in the main paper). Note that $\kappa^*(\eta)$ can be expressed as $\kappa^*(\eta) = (k_1 - k_2) \times \eta + C$, where k_1 is the number of edges connecting the selected nodes and the sink node t in G_{st} ; k_2 is the number of edges connecting the unselected nodes and the source node s ; C is some constant independent of η . Vary the value η will not change the cut edges unless: (1) η is set as a value that causes the rewiring of an edge of G_{st} from s to t according to Equation (3) (given in the main paper), or (2) η is set a a value that leads to the change of a selection. Therefore $\kappa^*(\eta)$ is a continuous and piece-wise linear function of η . Its slope changes at either a break-point or a change-point, where a η value is called a break-point if it leads to the change of selection, and called a change-point if it causes the rewiring of an edge.

Thus, computing the selection path is equivalent to finding the break-points of $\kappa^*(\eta)$. As our formulation satisfies all conditions of *parametric maximum flow* algorithm (Gallo *et al.*, 1989), hence $\kappa^*(\eta)$ is concave between any consecutive change-points (Gallo *et al.*, 1989). Therefore we can transform $\kappa^*(\eta)$ into a concave function throughout its domain, by correcting at each of its change-points so that the slope does not change at these values (as shown in Figure 10). We achieve this by using the following correction function

$$\kappa_c^*(\eta) = \kappa^*(\eta) - \sum_{c_p \in \mathcal{C}; c_p \leq \eta} (\eta - c_p), \quad (\text{S2})$$

where \mathcal{C} represents the change-points that can be obtained from Equation (3) in the main paper, simply, $\mathcal{C} = \{c_p \mid c_p > 0; c_p = z_p + \lambda d_p; 1 \leq p \leq n\}$.

The corrected capacity function $\kappa_c^*(\eta)$ has no change-points but the same break-points as $\kappa^*(\eta)$. It is also piece-wise linear and strictly concave throughout its domain. Thereby all break-points can be calculated by applying the iterative contraction algorithm described in Gallo *et al.* (1989). When all break-points are obtained, the selection path can be computed by setting η at each of the break-points and solving the corresponding optimization problem defined by Equation (S1). The pseudocode is given in Algorithm 1.

3.2 The nesting property and memoryless property of our module selection method

The module selection by solving Equation (S1) has the nesting property that $S(\eta_1) \subseteq S(\eta_0)$ if $\eta_1 > \eta_0$. It also has the memoryless property, that if a gene is not selected by setting η at some value (e.g., $\eta = \eta_0$), then it can be removed from the GeneNet when computing the selection at a η value greater than η_0 , as demonstrated by the following proposition. Here $S(\eta)$ represents the module selected by setting the sparsity parameter as η .

Proposition 1. Denote the selected genes by setting the sparsity parameter at η_0 as $S(\eta_0)$. The subnetwork of G induced by $S(\eta_0)$ is denoted as G_{sub} . Then, for any $\eta_1 > \eta_0$,

$$S(\eta_1) = S_{sub}(\eta_1),$$

where $S_{sub}(\cdot)$ represents the selection by solving Equation (S1) defined on the subnetwork G_{sub} .

Proof. As the module selection via Equation (S1) satisfies all conditions of *parametric maximum flow* algorithm (Gallo *et al.*, 1989), therefore the selection has the nesting property (according to Lemma 2.4 in Gallo *et al.* (1989)):

$$S(\eta_1) \subseteq S(\eta_0) \text{ for any } \eta_1 > \eta_0.$$

Denote S_0 as the node indices of $S(\eta_0)$, then maximizing $f(\mathbf{u}, \eta_1)$ is equivalent to maximizing $f\left([\mathbf{u}_{S_0}^\top, \mathbf{0}^\top]^\top, \eta_1\right)$. This is because according to the nesting property, the nodes unselected by setting the sparsity parameter at η_0 will not be selected by setting at a larger value η_1 neither. Note $f\left([\mathbf{u}_{S_0}^\top, \mathbf{0}^\top]^\top, \eta_1\right) \equiv f(\mathbf{u}_{S_0}, \eta_1)$, where the latter is the objective function defined on the induced network G_{sub} . This leads to $S(\eta_1) = S_{\text{sub}}(\eta_1)$, hence Proposition 1 holds. \square

Algorithm 1 Compute the selection path \mathcal{P} over a sparsity range $[\eta_{\min}, \eta_{\max}]$

```

1: function SELECTION_PATH( $\eta_{\min}, \eta_{\max}$ )
2:   Compute  $S(\eta_{\min})$  and  $S(\eta_{\max})$  ▷ As described in the main paper
3:   if  $S(\eta_{\min}) == S(\eta_{\max})$  then ▷ No more selection between them
4:      $path \leftarrow S(\eta_{\min})$ 
5:   else ▷ Use the divide-and-conquer strategy
6:     Compute  $\kappa^*(\eta)$  at  $\eta_{\min}$  and  $\eta_{\max}$ 
7:     Compute  $\kappa_c^*(\eta)$  at  $\eta_{\min}$  and  $\eta_{\max}$  according to Equation (S2)
8:     Compute the tangent lines of  $\kappa_c^*(\eta)$  at  $\eta_{\min}$  and  $\eta_{\max}$ 
9:      $\eta_{\text{mid}} \leftarrow \eta \in [\eta_{\min}, \eta_{\max}]$  where two tangent lines intersect ▷ The middle point
10:     $path\_head \leftarrow \text{SELECTION\_PATH}(\eta_{\min}, \eta_{\text{mid}})$ 
11:     $path\_tail \leftarrow \text{SELECTION\_PATH}(\eta_{\text{mid}}, \eta_{\max})$ 
12:     $path \leftarrow path\_head \cup path\_tail$  ▷ Put two sub-paths together
13:   end if
14:    $path \leftarrow$  order the selections in  $path$  by the size (number of nodes) of each selection
15:   return  $path$ 
16: end function

```

References

Gallo, G. *et al.* (1989). A fast parametric maximum flow algorithm and applications. *SIAM Journal on Computing*, **18**(1), 30–55.