

Supplementary material

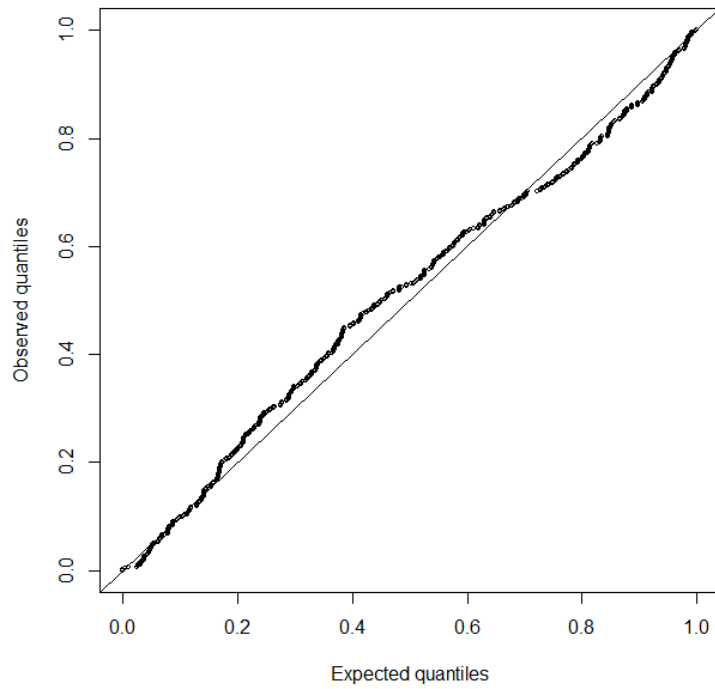


Figure S1: Observed quantiles of the p-value distribution ($n = 500$) under the null hypothesis plotted against the expected quantiles of a uniform distribution. The solid line is of equation $y = x$.

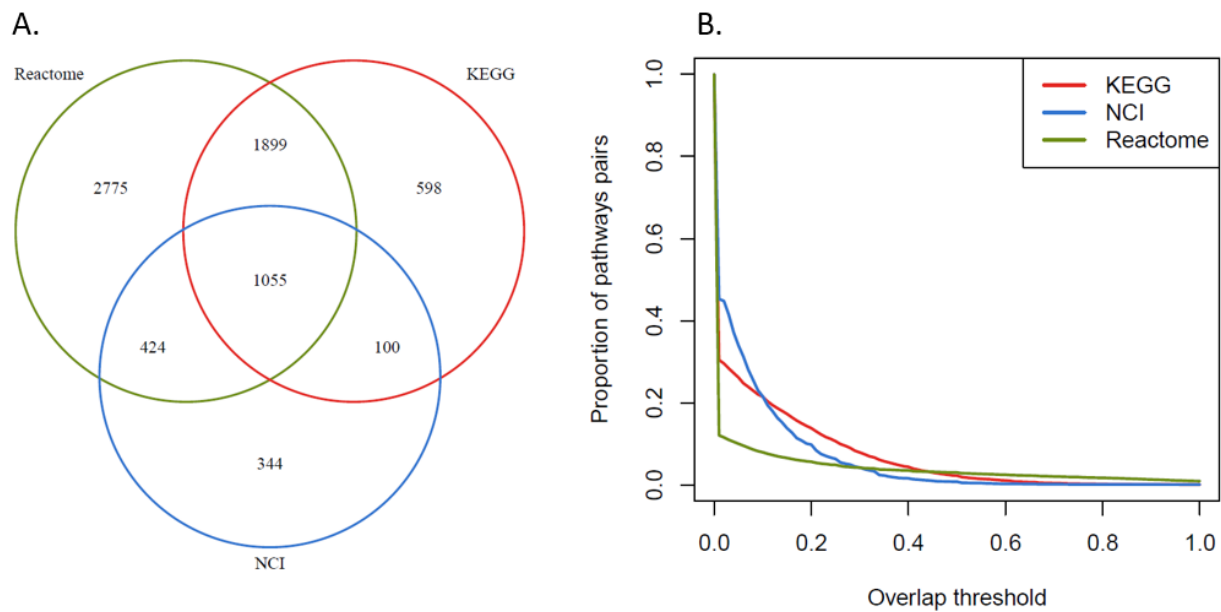


Figure S2: Overlap in gene content between pathways and pathway databases. The Venn diagram (A) represents the number of genes shared or not between the three different databases. A gene is considered part of a database if it occurs in at least one pathway in this database. Overlap within pathways is summarized by computing the redundancy measure (B) in each database. Redundancy is characterized by the proportion of pathway pairs with a higher overlap than a given threshold as a function of this threshold.

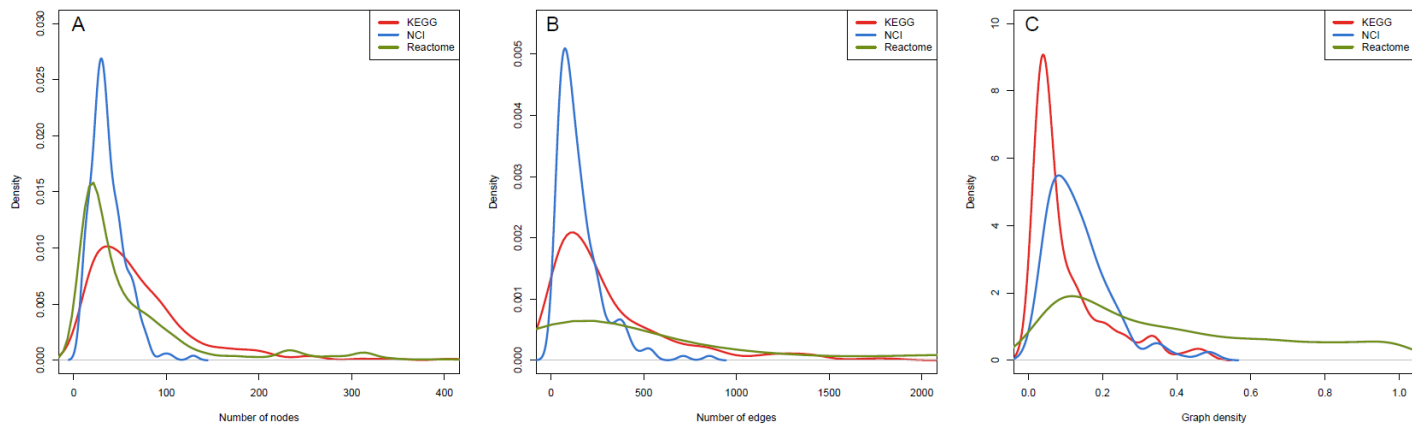


Figure S3: Pathway databases characteristics. The number of nodes (A), edges (B) and the graph density (C) distributions are represented for the three pathway databases: KEGG (red), NCI (blue) and Reactome (green). The graph density corresponds to the number of observed edges over the total number of possible edges, and it is a measure of the network connectivity.

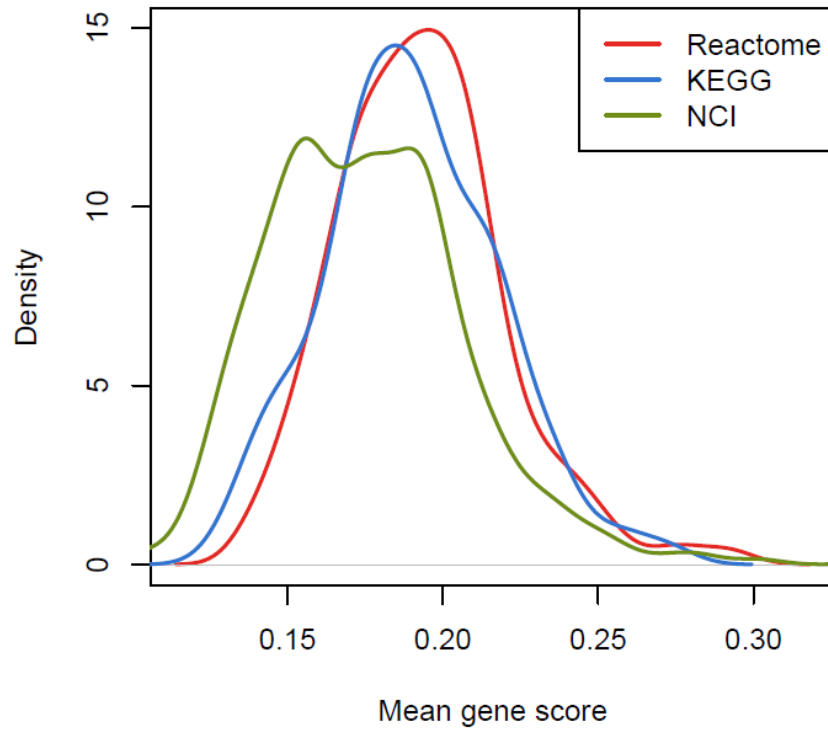


Figure S4: Mean gene scores (probability of convergent adaptation) distributions for the three databases. For each pathway, the average gene score has been computed and its distribution for each database is represented: KEGG (blue), NCI (green), Reactome (red).

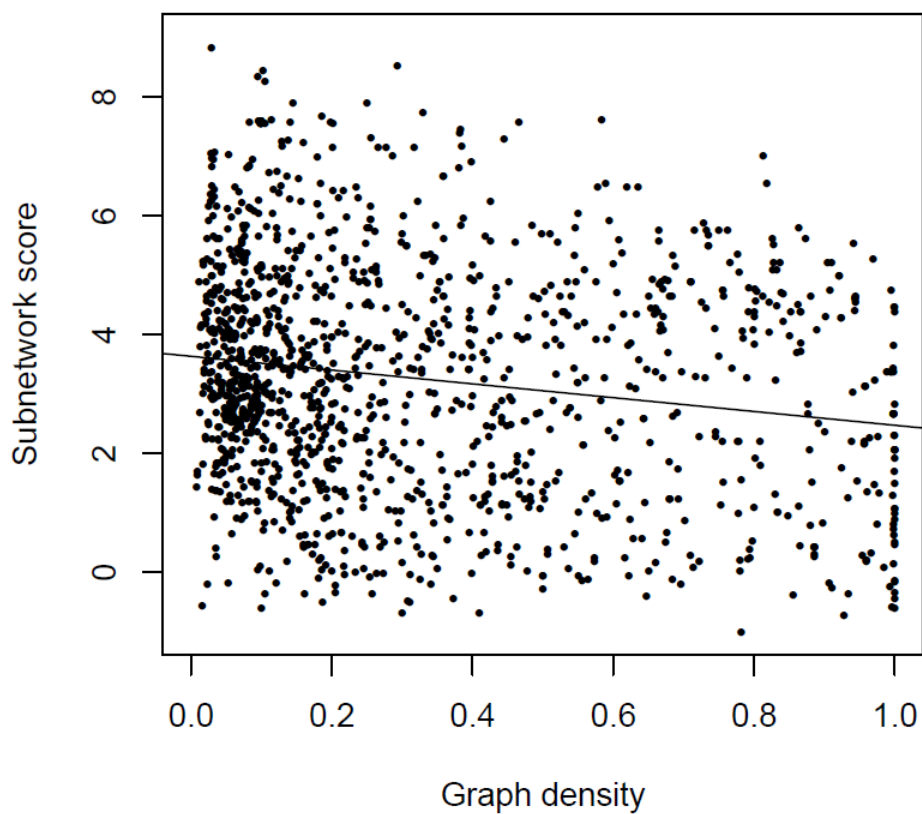


Figure S5: Correlation between graph density and estimated subnetwork score. The subnetwork scores obtained with the search algorithm for the three databases is plotted against the whole pathway density. The line represents the prediction of a linear regression of the scores against the density ($F = 42.11$, 1 and 1339 degrees of freedom, $p = 1.2 \cdot 10^{-10}$; $R = -0.17$).

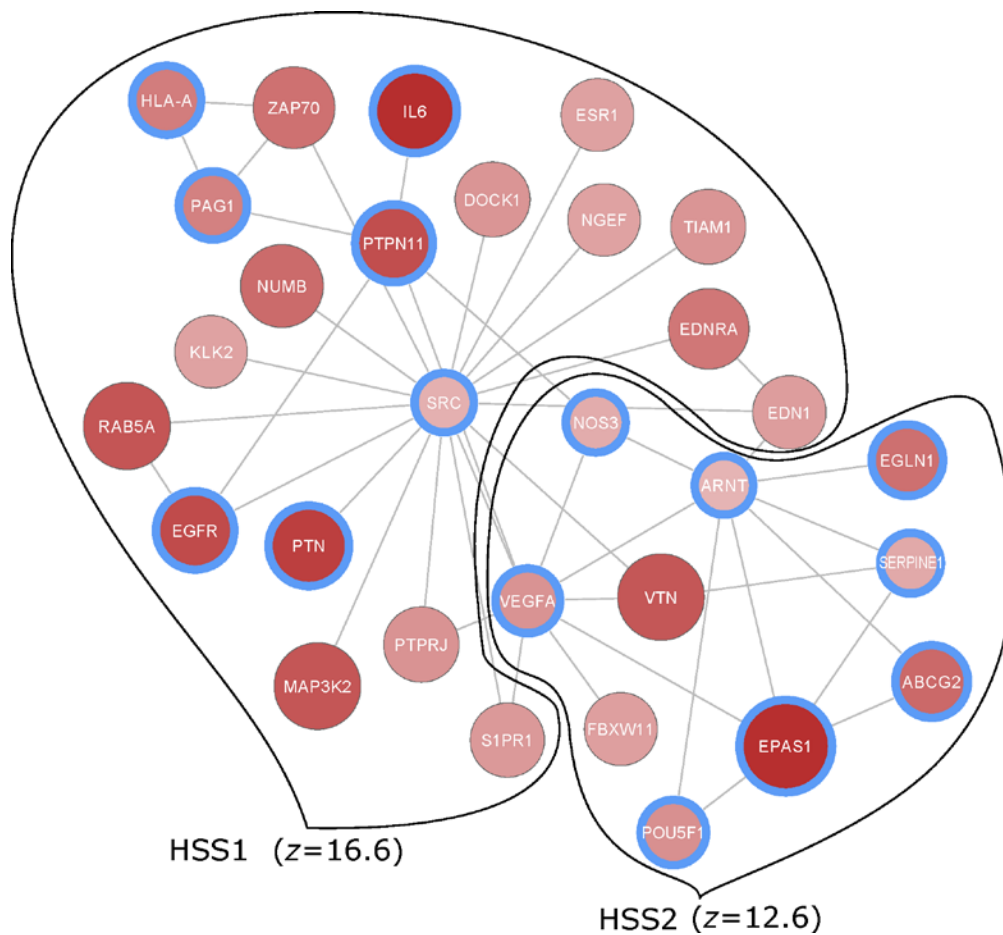


Figure S6: Significant subnetworks in the analysis of a large network of 2,354 genes and 21,537 interactions obtained by merging all pathways of the NCI database. The two HSS are represented with their z score. The color and node size are proportional to the individual gene score. A thick blue node border indicates that the gene has been identified with the individual pathway analysis.

Table S1: Computation runtime analysis. The values indicate the average *signet* execution time (in seconds) computed over 10 runs. We have examined the effect of three different parameters: number of simulated annealing iterations, network density (*d*), and network size.

		Iterations					
		1,000		5,000		10,000	
		d = 0.2	d = 0.8	d = 0.2	d = 0.8	d = 0.2	d = 0.8
Network size	100	2.27	2.21	12.8	13.12	25.21	24.38
	500	5.49	5.79	28.56	29.07	61.83	58.58
	1,000	15.99	16.07	75.53	78.48	172	162
	5,000	332	341	1628	1554	3157	3319

Table S2: Significant highest-scoring subnetworks (HSS). The database, pathway names and sizes are indicated for each significant HSS ($p < 0.01$). The HSS sizes, scores, p-values and q-values are shown, as well as the HSS composition (Entrez gene IDs). Significance threshold was determined on the basis of a q-value < 0.2 .

Database	Pathway	Pathway size	HSS size	HSS score	p-value	q-value	HSS genes (Entrez ID)
NCI	HIF-2-alpha transcription factor network	28	9	6.113	0.002	0.157	10370;135;2034;405;5054;5460;7422;8553;9429
NCI	Syndecan-3-mediated signaling events	18	5	5.221	0.006	0.194	1956;5764;6714;8573;9672
NCI	SHP2 signaling	48	25	5.206	0.006	0.194	10818;1950;1956;2770;284;2909;3479;3480;3558;3560;3569;3572; 3717;3791;4846;4908;4909;4915;4916;5295;55824;5781;627;7010; 7422
NCI	ErbB receptor signaling network	15	7	4.888	0.007	0.194	10718;1950;1956;2065;2066;3084;7039
KEGG	Chemical carcinogenesis	64	23	6.015	0.002	0.085	1139;124;125;127;128;130;131;1553;1646;218;54490;54575;54576 ;54577;54578;54579;54600;54657;54658;54659;7364;79799;873
KEGG	Cell adhesion molecules (CAMs)	87	3	5.923	0.002	0.092	22871;57502;9378
KEGG	Drug metabolism - cytochrome P450	65	12	5.832	0.002	0.092	124;125;126;127;128;130;131;218;221357;2941;2950;4258
KEGG	Antigen processing and presentation	47	7	5.177	0.004	0.101	3105;3106;3107;3133;3134;3135;3824
KEGG	Graft-versus-host disease	16	7	5.177	0.004	0.101	3105;3106;3107;3133;3134;3135;3824
KEGG	Tyrosine metabolism	32	18	5.213	0.004	0.101	124;125;126;127;128;130;131;1312;1644;218;259307;2805;2806;4 128;4129;4282;6898;7054