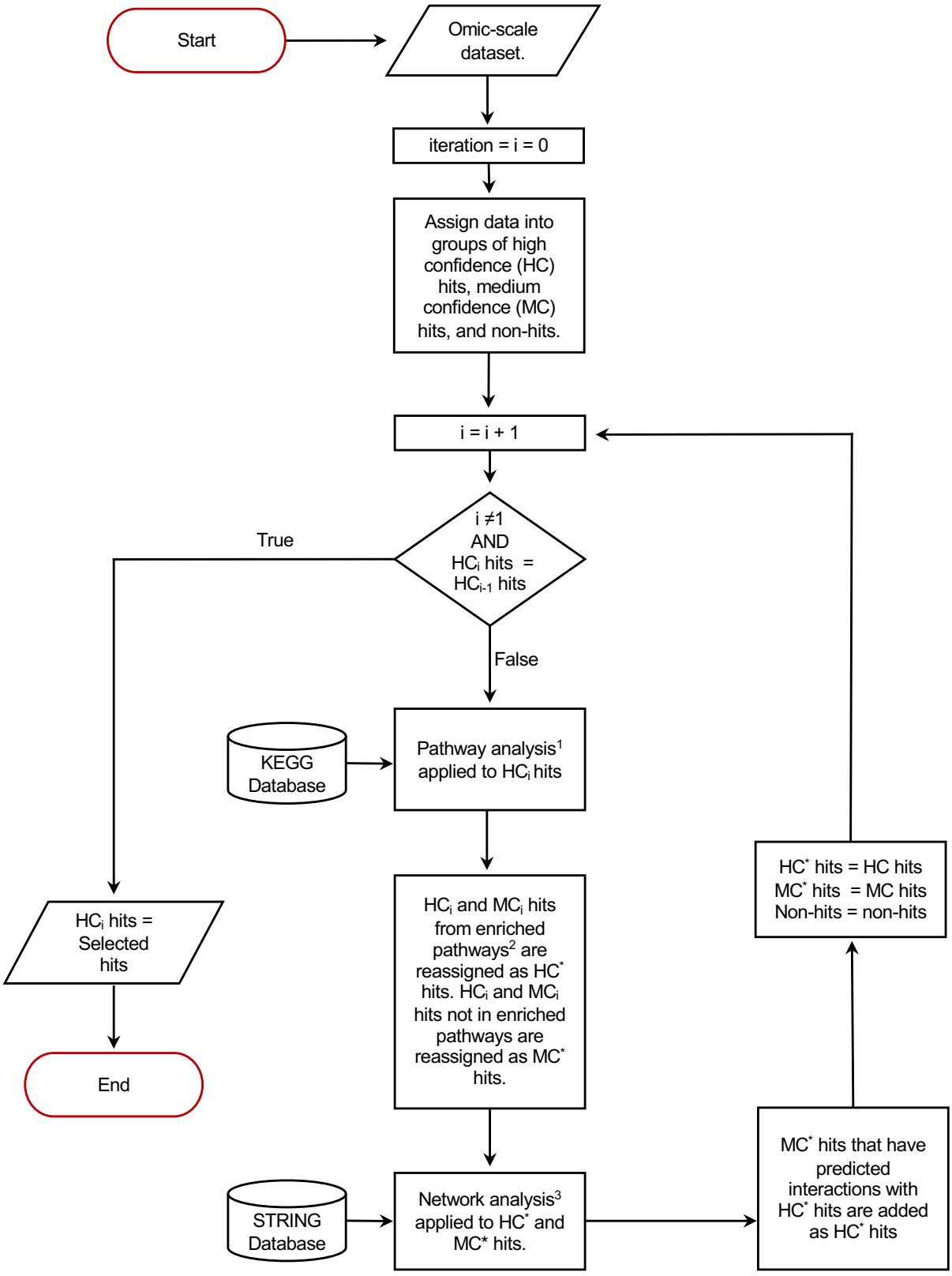
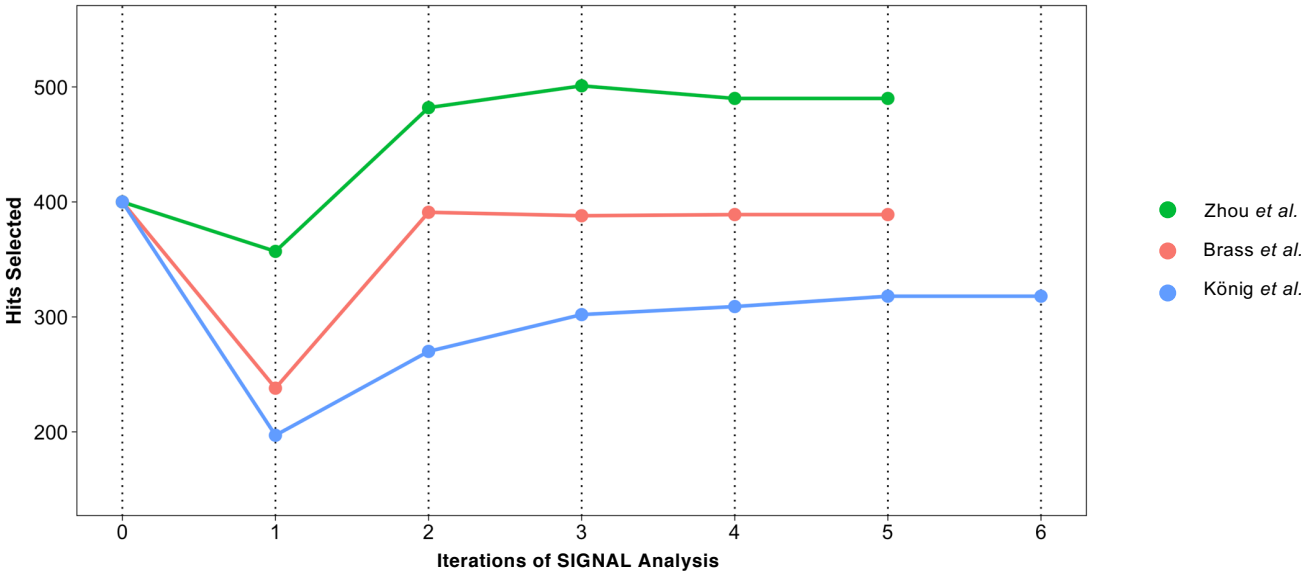
Supplementary Table 1: Design and hit selection methods for the three siRNA studies of early HIV dependency factors by Zhou *et al*., Brass *et al.*, and König *et al.*

		Zhou *et al.*	Brass *et al.*	König *et al.*
Experimental Conditions And Design	Cell Type	HeLa P4/R5 Cells	HeLa-derived TZM-bl Cells	293T Cells
	Treatment	HXB2 HIV- 1	HIV IIIB	VSV-G pseudotyped HIV-1 reporter virus encoding luciferase
	Readout 1	Tat activation of expression of the β-Gal reporter	p24 (product of gag gene)	HIV-1 Vector encoded luciferase
	Time point: Readout 1	48h	48h	24h
	Readout 2	Tat activation of expression of the β-Gal reporter	β-Gal (Tat dependent)	MuLV and AAV
	Time point: Readout 2	96h	72h	24h
Hit Selection, Bioinformatics, and Secondary Screening	Cell Viability Correction	Decrease of cell viability by 2 SDs or more	Decrease of cell viability by 2 SDs or more	Cell toxicity screen
	Z score cutoff	2 SSMD relative to the negative control	2 SDs greater than the plate mean	2 siRNAs with ≥45% reduction in HIV infectivity
	Bioinformatics Used in Hit Selection	In silico screening for expression in activated T cells and Macrophages	None	"evidence score" based on functional, biochemical, and transcriptional data. Yeast to hybrid protein interaction database, NCBI HIV-1 Protein Interaction Database, MCODE, Ontogeny-based pattern identification algorithm
	Secondary Screening	Rescreening by independent siRNAS	Rescreening of pooled siRNAs in single siRNA assay	Rescreening of pooled siRNAs in single siRNA assay



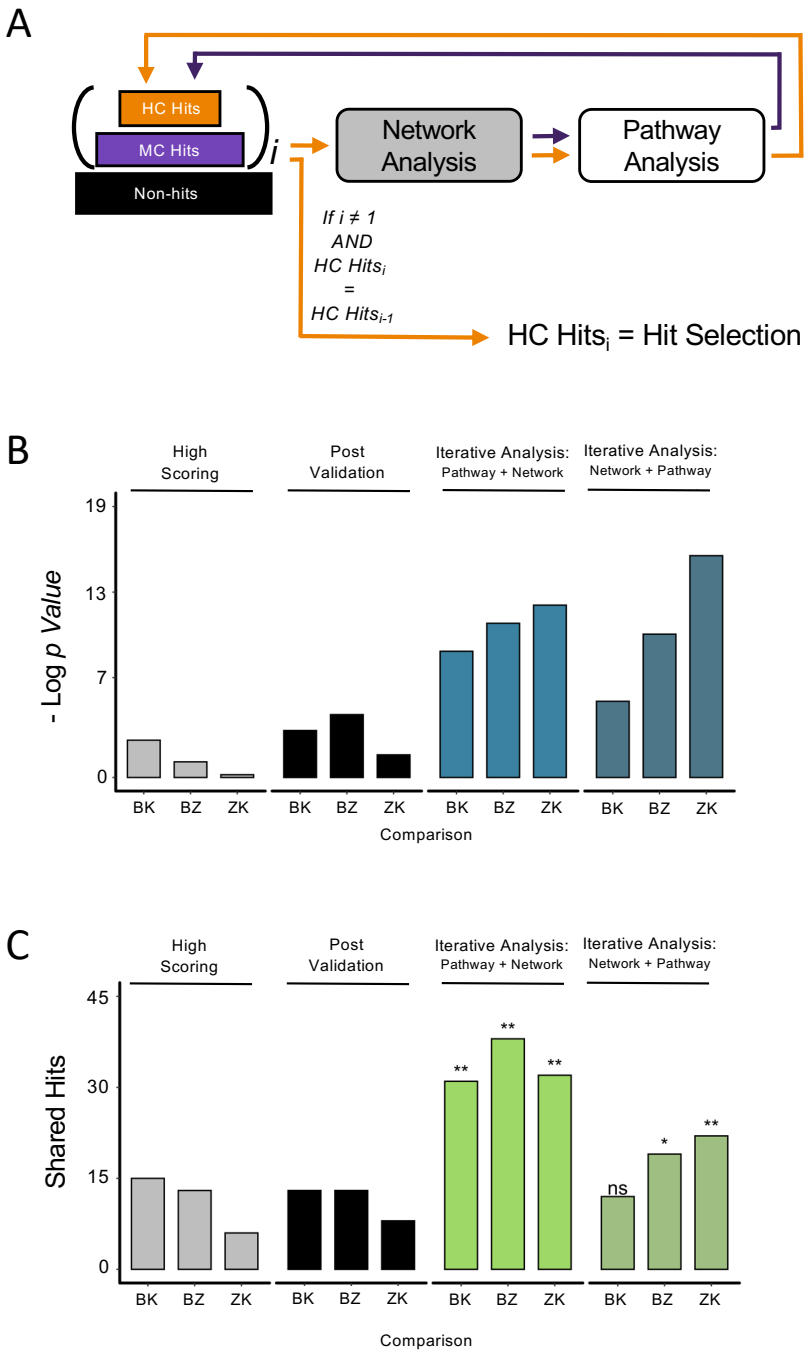
1. Hypergeometric test, alternative hypothesis = "greater than", null hypothesis = no shared enrichment.
2. $p < 0.05$
3. Direct neighbor functional approach, evidence source: Experimental & Database, Confidence ≥ 0.4

Supplementary Figure 1 (related to figure 3D): Hit selection by iterative application of pathway and network analysis.
Flowchart of the Selection by Iterative pathway Group and Network Analysis Looping (SIGNAL) hit selection pipeline.



Supplementary Figure 2 (related to figure 3D): Iterations of integrated analysis of the three studies of HIV HDFs.

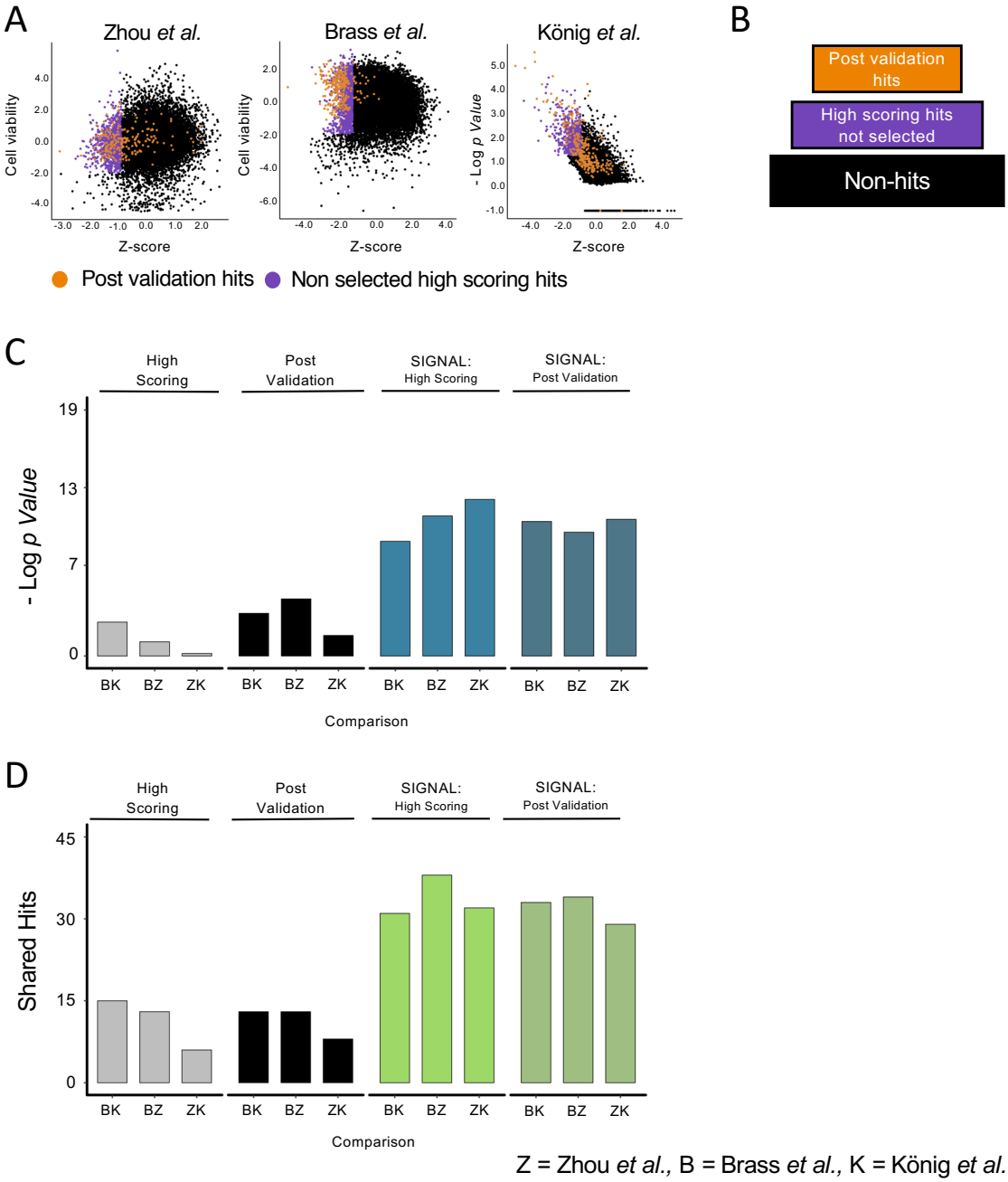
0 on the x-axis represents the high confidence set of hits at the analysis input stage. The high confidence hit sets are contracted and expanded through iterative analysis cycles. Analysis terminates when high confidence sets do not change between two consecutive iterations.



Z = Zhou *et al.*, B = Brass *et al.*, K = König *et al.*

Supplementary Figure 3: Hit selection by iterative analysis with reverse pathway and network order.

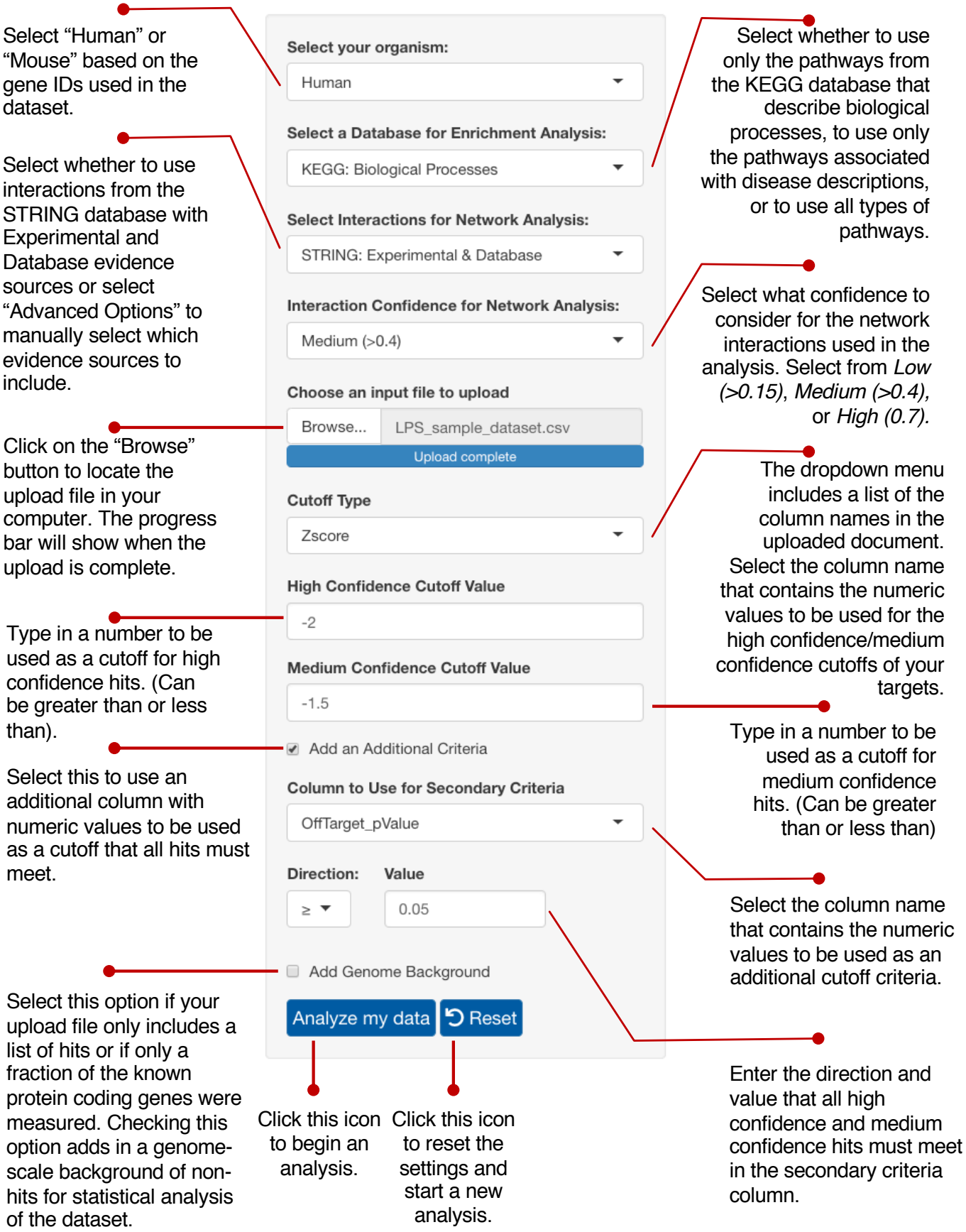
(A) Schematic of the iterative analysis as in Fig. 3D with the order of pathway and network analysis reversed. (B) statistical significance of the overlap across the three studies of HDFs when comparing hits selected by reverse iterative analysis versus highest scoring hits, post validation hits and hits selected by the alternative design of iterative analysis. (C) Number of shared hits between the hits selected by reverse iterative analysis from the three studies versus highest scoring hits, post validation hits and hits selected by the alternative design of iterative analysis.. Random permutation test scores: ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$



Supplementary Figure 4: Using post-validation hits for analysis by SIGNAL.
(A) Scores from three genome-wide studies of HDFs. Post-validation hits are in orange and the highest scoring 1000 genes not selected are in purple. (B) Schematic of three-tiered data using post validation hits as high confidence hits, and non-selected high scoring hits as medium confidence hits. (C) Statistical significance of the overlap across the three studies of HDF when comparing hits selected by SIGNAL analysis of post validation hits versus highest scoring hits, post validation hits, and hits selected by SIGNAL analysis of high scoring hits. (D) Number of shared hits between the hits selected by SIGNAL analysis of post validation hits versus highest scoring hits, post validation hits, and hits selected by SIGNAL analysis of high scoring hits. Random permutation test scores: ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$

GeneSymbol	EntrezID	PercInfected.Zscore	CellNumber.Zscore	assigned.value
CXCR4	7852	-4.96623446	0.755277194	1
C1orf52	148423	-3.637572822	1.925896515	1
MED14	9282	-3.435696475	0.976086257	1
ADAM10	102	-3.435673997	1.513583032	1
GCK	2645	-3.223640638	1.920937941	1
GPR21	2844	-3.201246647	-0.096137148	1
ZNF831	128611	-3.20098761	-0.243515999	1
CD4	920	-3.162525347	1.767687649	1
EGFR	1956	-3.162525347	1.487870985	1
WNT1	7471	-3.140163554	1.674727332	1
USP6	9098	-3.114415882	1.238823011	1
PLEKHA7	144100	-1.920126088	-1.705268716	0.5
DPH3	285381	-1.919642212	-0.473272261	0.5
NA	284861	-1.919642212	-0.200375892	0.5
PNMA6A	84968	-1.917641072	-1.574602053	0.5
EIF3G	8666	-1.917428352	-1.51997327	0.5
TFDP2	7029	-1.916264377	1.3078323	0.5
CLNS1A	1207	-1.915660931	0.543313511	0.5
MMP19	4327	-1.90908986	1.426094148	0.5
RECQL4	9401	-1.90908986	1.167972495	0.5
ZNF536	9745	-1.909010831	-0.111569716	0.5
NMUR2	56923	-3.690591512	-2.63385185	0
SMU1	55234	-3.686455305	-3.076050904	0
LSM8	51691	-3.62462392	-2.307921727	0
NAT10	55226	-3.460889184	-2.681841646	0
SGO1	151648	-3.435116303	-4.066983289	0
DHRS13	147015	-3.38711184	-2.495038806	0
XAB2	56949	-3.327710602	-3.064859872	0
HEG1	57493	-3.291433863	-2.728928693	0
COPB2	9276	-3.28475593	-4.284378793	0
PSMB6	5694	-3.261635121	-3.280551426	0

Supplementary Figure 5 (related to figure 6A): A sample input file for SIGNAL.
A sample dataset prepared for SIGNAL analyses using the data from the Brass *et al.* of study of essential factors for HIV infection. Gene column IDs are labeled as "EntrezID" and "GeneSymbol" (either one is sufficient for upload). The "PercInfected.Zscore" column includes the normalized Z scores and can be used to set cutoffs for the high confidence and medium confidence fields on the SIGNAL platform. To incorporate the "CellNumber.Zscore" in defining high confidence vs. medium confidence hits, a new column is created "assigned.value". Hits assigned as high confidence by both criteria are given a value of 1, hits assigned as medium confidence are given a value of 0.5. Hits that don't meet the two criteria are assigned a value of 0.



Supplementary Figure 6 (related to figure 6A): Setting up an analysis on SIGNAL.
Guide for the control panel for setting up an analysis session on the *signal.niaid.nih.gov* web interface.

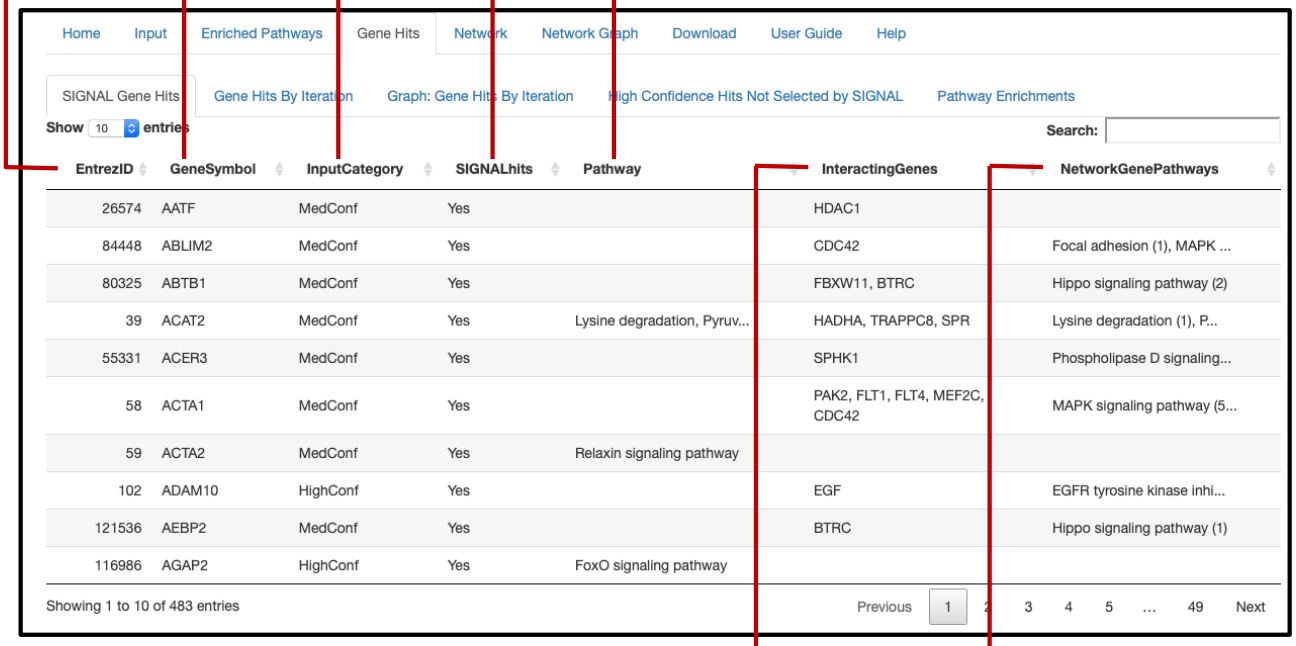
NCBI EntrezID.

HGNC Gene Symbol.

The confidence category this gene was assigned to at the start of the analysis.

If the gene was selected as a hit by SIGNAL.

Names of KEGG Pathways that were identified by SIGNAL as enriched for in the dataset that the gene is an annotated member of.

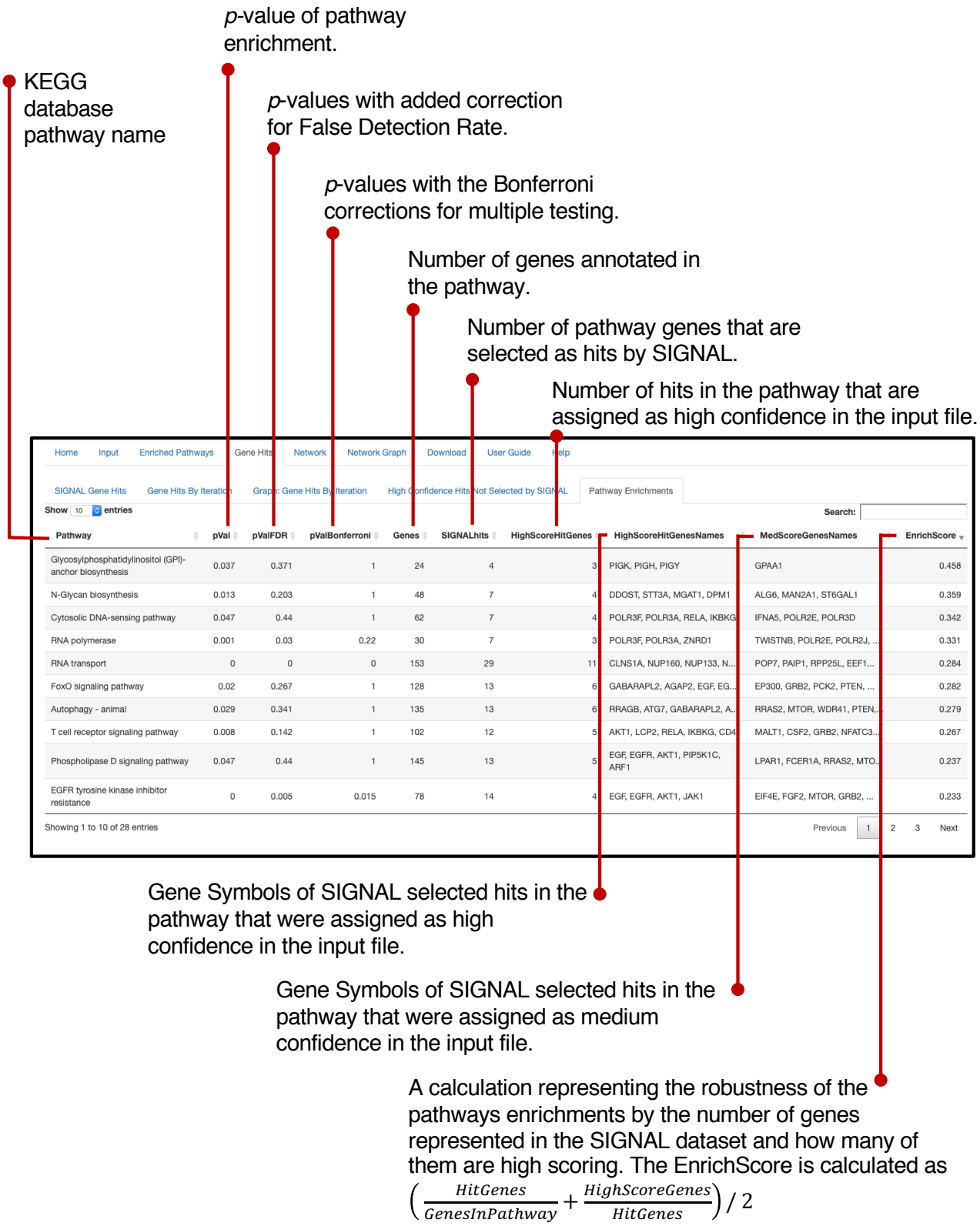


Names of SIGNAL gene hits from the analysis that have predicted interactions with the listed hit gene of the row. The interactions are determined based on the network settings set by the user at the start of the analysis.

Names of enriched pathways from the analysis which the interacting genes are members of.

Number in parenthesis indicates the number of Interacting genes that are members of the preceding pathway.

Supplementary Figure 7 (related to figure 6): SIGNAL Gene Hits table on SIGNAL. Guide for the "SIGNAL Gene Hits" table generated after an analysis session on *signal.niaid.nih.gov* is complete.



Supplementary Figure 8 (related to figure 6): Pathway Enrichments table in SIGNAL. Guide for the "Pathway Enrichments" table generated after an analysis session on *signal.niaid.nih.gov* is complete.