iScience, Volume *25*

Supplemental information

Transcriptional regulatory networks of circulating immune cells in type 1 diabetes: A community knowledgebase

Scott A. Ochsner, Rudolf T. Pillich, Deepali Rawool, Jeffrey S. Grethe, and Neil J. McKenna

Primary Supplemental Material

This section relates to:

Supplemental text

Supplemental Figs. S1 to S6

Other Supplemental Materials for this manuscript include the following:

Tables S1-21

Supplemental Text

Hypergeometric test gene set cut-offs for figure plots

Figure 1. A. k = Xhonneux IAAsig; s = Top ranked T1D consensome genes (GMFC > 1.3; p < 5E-03); N = universe of T1D consensome genes (n = 22549). **B.** k = Xhonneux IAAsig; s = Top ranked T1D consensome genes (GMFC > 1.3; p < 5E-03) & top ET1DREs (ENR > 7.7 & p<5E-04); N = universe of T1D consensome genes (n = 22549). **C.** k = Top ET1DRE genes (ENR>6.6; p < 3E-05); s = top teplizumab R v C-induced genes, LFC>0.34 & p < 0.05; N: universe of probesets (n = 35722). **D.** k = T1DKP T1D GWAS p < 1E-150; s = Top ranked ET1DRE genes (ENR>6.3; p < 5E-07); N = universe of T1D consensome genes (n = 22549).

Figure 2. A. k = T1DKP T1D GWAS SNPs (p < 3E-05); s = top ET1DRE p < 0.05 nodes (ENR>3.85, p < 3E-04); N = all nodes included in HCT intersection analysis (n = 961). **B.** k = PAGGs; s = ET1DRE nodes OR>2.45, p < 4E-05); N = all nodes included in HCT intersection analysis (n = 961). **C.** k = ET1DRE nodes ENR>1.7; p < 0.05; s = top negatively ranked Shifrut nodes (num = 4, score ≤3E-08, sgRNA = 4, rank >15, lfc ≤-0.6); N = universe of genes in Shifrut screen (n = 19106). **D.** k = ET1DIE nodes ENR>1.6; p < 0.05); s = top negatively ranked Shifrut nodes (num = 4, FDR ≤3E-09, sgRNA = 4, rank ≥ 20, lfc ≤-0.9); N = universe of genes in Shifrut screen (n = 19106). **E.** k = genes mapped to “autoimmunity” (HP:0002960); s = ET1DRE nodes (OR>2.5 p < 4E-08); N = all nodes included in HCT intersection analysis (n = 961). **F.** k = genes mapped to “autoimmunity” (HP:0002960); s = ET1DIE nodes (OR>2.5 p < 2E-05); N = all nodes included in HCT intersection analysis (n = 961).

Figure 4. A. k = Travaglini NK-specific p < 1E-20; s = Top ET1DREs: (T1D consensome GMFC>1.35 & p < 5E-05; ENR > 6 and p<2E-03); N = all genes in consensome (n = 22,549). **C.** k = Hesslein NK transcription factors; s = ET1DRE∩NK nodes OR>5.5 & p < 6E-04; N = all nodes included in the HCT intersection analysis (n = 961). **D.** k = Top ET1DRE∩NK genes: rank > 13 & ET1DRE ENR > 6, p < 2E-03; s = $CD57^{+}$-specific and $CD56^{dim}$ v $CD56^{bright}$ LFC>1.5 & p < 2E-03; N = all probesets on array (n = 16,495). **E.** k = $CD56^{dim}$-induced (LFC>1.2 & p < 3E-02); $CD57^{+}$-specific; ET1DRE∩NK ENR > 6.6 & p < 3E-05, rank >15; s = teplizumab responder v control LFC>0.34 & p < 0.05; N = all probesets on array (n = 35,722). **F.** k = Beckmann et al. MIS-C signature; s = ET1DRE∩NK genes (ENR > 3.8; p < 1.2 E-02) & rank > 35; N = universe of T1D consensome genes (n = 22,549). **G.** k = Beckmann et al. MIS-C signature; s = $CD57^{+}$-specific ET1DRE∩NK genes (ENR > 3.8; p < 1.2 E-02) & rank > 34; N = universe of T1D consensome genes (n = 22549). **H.** k = Beckmann et al. MIS-C signature; s = teplizumab induced (LFC ≥ 0.3, p < 0.05) ET1DRE∩NK (ENR>7.7, p < 3E-03 & rank > 17); N = universe of T1D consensome genes (n = 35722).

Figure 5. A. k = TRPC6M LFC< -2.5 & p < 5E-03; s = top NEOT1DI: T1D v C LFC>2 & p < 0.05; N = all genes in RNA-Seq experiment (n = 13,215). **B.** Hypergeometric test gene sets: k = IAV LFC>14 & p < 2E-03; s = top NEOT1DI: T1D v C LFC>2 & p < 0.05; N = all genes in RNA-

Seq experiment (*n* = 13,215). **C.** k = IAV LFC>14 & *p* < 2E-03; s = top NEOT1DI: T1D v C LFC>2 & *p* < 0.05 || TRPC6M v WT LFC< -2.7 & *p* < 5E-03; N = all genes in RNA-Seq experiment (*n* = 13,215). **D.** k = Reactome interferon alpha/beta signaling; s = TRPC6M↓∩NEOT1DI:NEOT1DI OR > 7 & TRPC6M↓∩NEOT1DI p>1E-13 IAV; N = all nodes in HCT intersection analysis (*n* = 961). **E.** k = CNB1KO v WT LFC< -0.2 & *p* < 0.05; s = top NEOT1DI: T1D v C LFC>2 & *p* < 0.05; N = all genes in RNA-Seq experiment (*n* = 13,215). **F.** k = CNB1KO v WT LFC< -0.2 & *p* < 0.05; s = top NEOT1DI (T1D v C LFC>2 & *p* < 0.05) and TRPC6M v WT LFC< -1.6 & *p* < 5E-03; N = all genes in RNA-Seq experiment (*n* = 13,215). **G.** k = CNB1KO v WT LFC< -0.2 & *p* < 0.05 and STAT2 ChIP-Seq consensome %ile ≥98; s = top NEOT1DI (T1D v C LFC>2 & *p* < 0.05) and TRPC6M v WT LFC< -1.6 & *p* < 5E-03; N = all genes in RNA-Seq experiment (*n* = 13,215). **H.** Interferon alpha beta signaling pathway targets are very strongly enriched among TCS-NEOT1DI genes relative to NEOT1DI genes. Refer to table S14 for the underlying data.

Figure 6. A. k = ST1D:MT1D ML10P Ratio > 9; s = screen top 175 genes; N = all genes in screen (*n* = 21379). **D.** k = clade 150c nodes; s = ST1DDs; N = all nodes in HCT intersection analysis (*n* = 961). **E.** k = clade 150c nodes; s = screen top 340 genes; N = all genes in screen (*n* = 21379). **F.** k = Green-IR; s = ST1D↑ (cohort 1 ST1D v MT1D LOG FC>1.3 & *p* < 5E-03; cohort 2 ST1D v MT1D LOG FC>0.8 & *p* < 5E-03); N = all probesets on array (*n* = 30633). **G.** k = Green-IR; s = ST1D↑ (cohort 1 ST1D v MT1D LOG FC>1.3 & *p* < 5E-03; cohort 2 ST1D v MT1D LOG FC>0.8 & *p* < 5E-03) & PCF11 ChIP-Seq consensome %ile > 75; N = all probesets on array (*n* = 30633). **H**. k = ST1D↑ predicted clade 150c targets (cohort 1 ST1D v MT1D LOG FC>1.6 & *p* < 4E-05; cohort 2 ST1D v MT1D LOG FC>0.9 & *p* < 1E-02) and in the consensome 75th ile for at least two of the three members of clade 150c.

Figure S1. A. k = Top ET1DIE genes (ENR>10, *p* < 3E-07); s = top teplizumab R v C-repressed genes, LFC< -0.26 & *p* < 0.05; N = universe of probesets (*n* = 35722). **B.** k = T1DKP T1D GWAS *p* < 6E-108; s = Top ranked ET1DIE genes (ENR ≥ 6.25; *p* < 6E-03); N = universe of T1D consensome genes (*n* = 22549). **C.** k = top 3 PAGGs; s = Top ranked ET1DRE genes (ENR>6.3; *p* < 5E-07); N = universe of T1D consensome genes (*n* = 22549).

Figure S2. A. k = Ota-NK; s = ET1DRE ENR>7.7 & p<1E-05; N = all genes in consensome (n = 22,549). **C.** k = Brillantes NK transcription factors; s = ET1DRE∩NK nodes OR>5.5 & p< 6E-03; N = all nodes included in the HCT intersection analysis (n = 961). **D.** k = Top ET1DRE∩NK genes: rank >24 & ET1DRE ENR > 6, p < 2E-03; s = CD56^{dim} v CD56^{bright} LFC>1.5 & p<2E-03; N = all probesets on array (n = 16,495).

Figure S3. k = Reactome Interferon alpha/beta signaling (R-HSA-909733); s = NEOT1DI nodes OR>2.75 & *p* < 1E-12; N = all nodes included in HCT intersection analysis (*n* = 961).

Figure S4. A-C. k = BioGRID curated interacting proteins; s = nodes with $r \geq 0.95$; N = full list of 600 nodes that had at least two $p < 0.05$ intersections across the eight gene sets. **D.** k = siPCF11 v C DEGs $p <$ 3E-04; s = mean MACS2 score>650; N = all genes in PCF11 ChIP-Seq consensome (n = 3807). DEGs: differentially-expressed genes. **Figure S4. B.** k = FRIDMAN SENESCENCE; s = MEAN FC>6 & $p <$ 1E-15; N = All genes in ETOPO consensome (n = 21234).

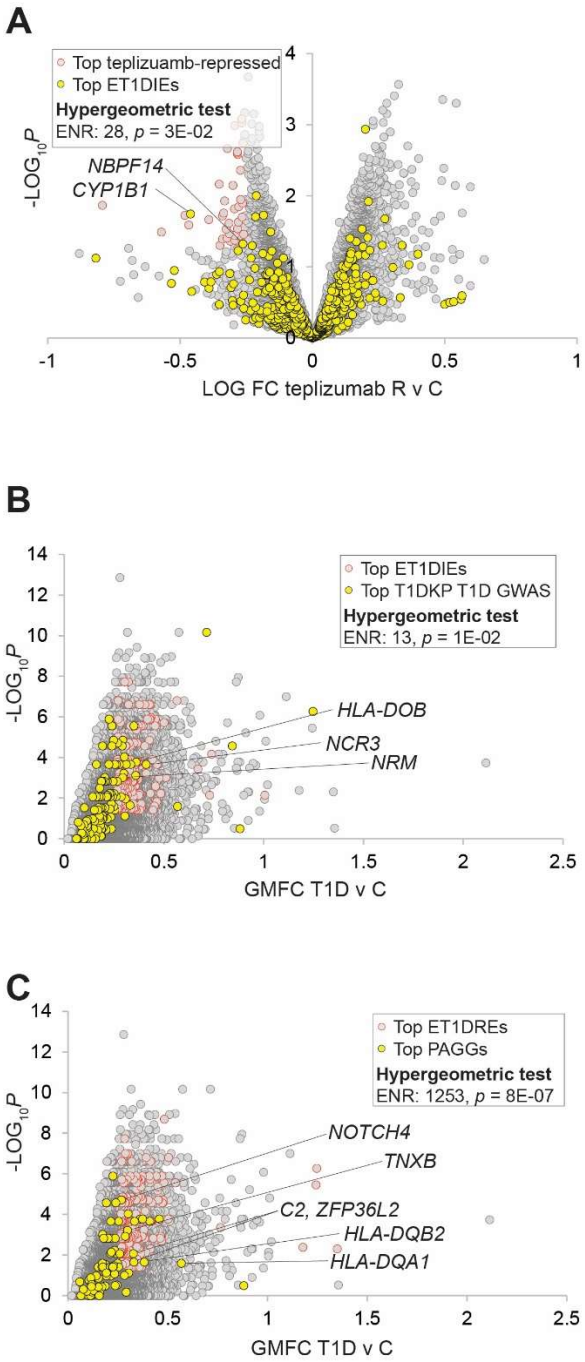

Figure S1. Validation of ET1DREs and ET1DIEs, related to Figure 1. Selected transcripts at the intersection of the indicated gene sets are labeled. Refer to supplementary text for cut-offs used for hypergeometric tests. **A.** Top-ranked ET1DIEs are enriched among top teplizumab-repressed genes in T1D subjects. **B.** Top ranked T1DKP GWAS T1D-associated genes are enriched among the top ranked ET1DIEs in the T1D consensome. **C.** Top ranked PAGGs are enriched among the top ranked ET1DREs in the T1D consensome.

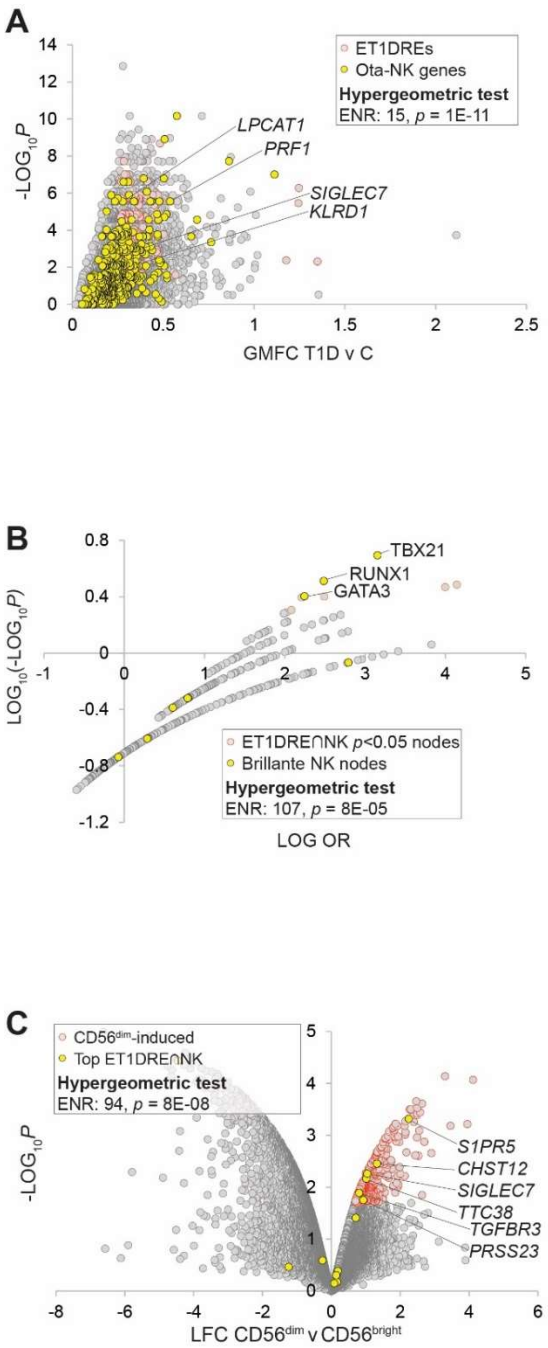

Figure S2. Supplementary data for use case 1, related to Figure 4. A. The highest ranked ET1DRE genes are enriched for NK cell-specific genes as described in Table S2 from the Ota study. **B.** Nodes that have significant transcriptional footprints within the ET1DREꓵNK gene set are enriched for canonical NK transcription factors per Brillantes and colleagues. **C.** ET1DREꓵNK genes are enriched among CD56dim-induced genes.

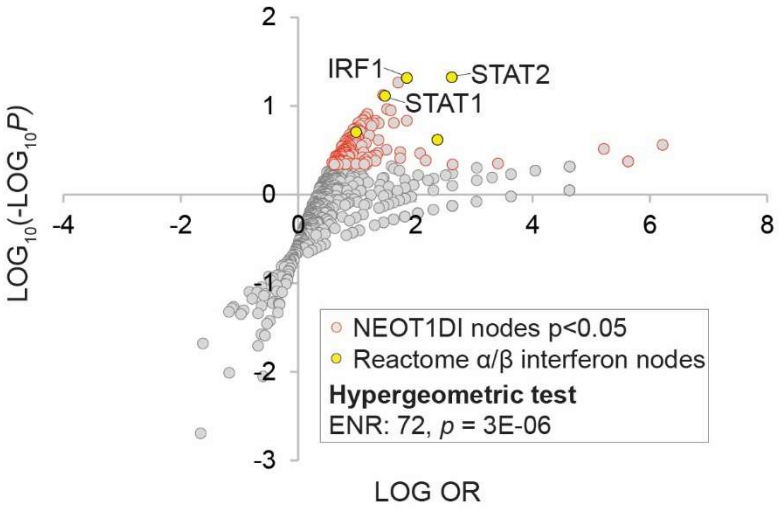

Figure S3. Supplementary data for use case 2, related to Figure 5. Canonical drivers of the type 1 interferon response are enriched among nodes with robust transcriptional footprints among NEOT1DI genes.

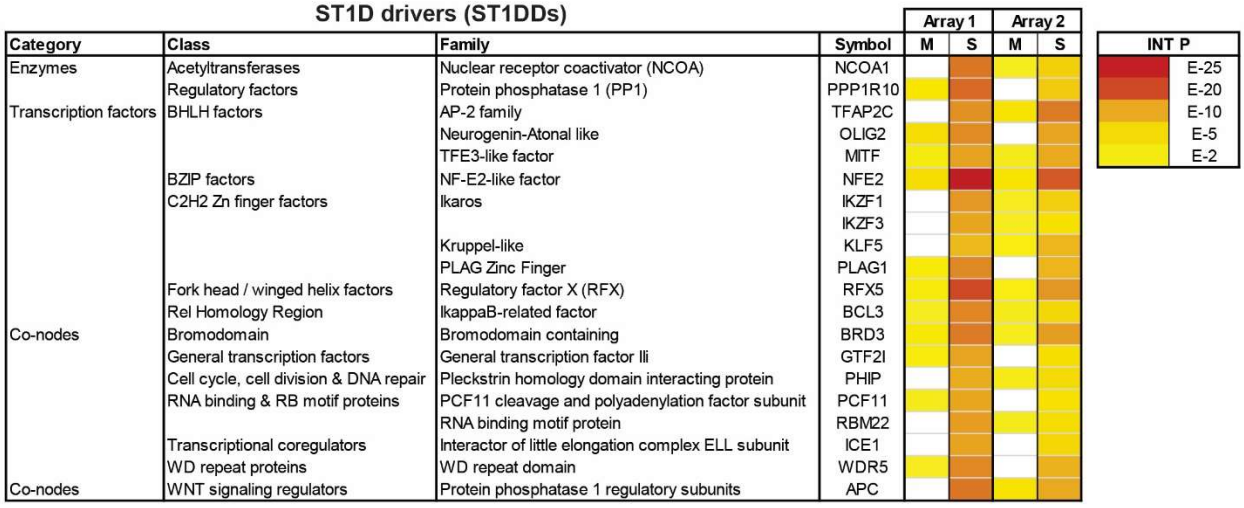

Figure S4. Candidate transcriptional regulatory drivers of severe type 1 diabetes (ST1DDs), related to Figure 6. T1D v control differential expression cut-offs: FC≥ ± 1.25 & p<0.05; ST1DD cut-off criteria: minimum Σ ST1D-induced OR = 3.8; minimum Σ ST1D -log10P = 14.5; minimum ST1D:MT1D -log10P ratio: 4.7. M, mild T1D v healthy control, S, severe T1D v healthy control.

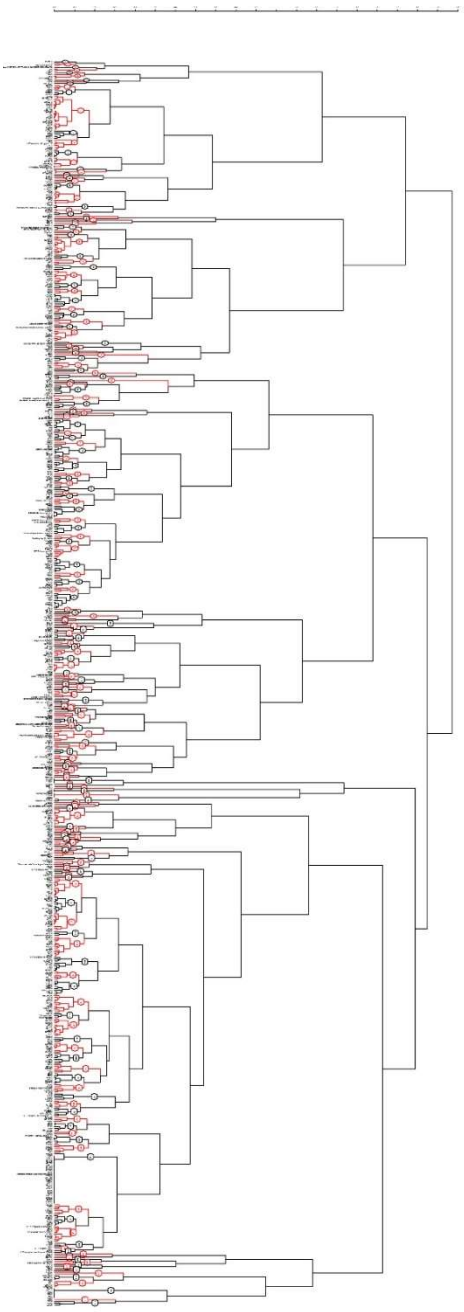

Figure S5. Use case 3: hierarchical clustering of nodes that had at least two significant (p<0.05) intersections across the eight independent gene sets in GSE33440, Related to Figure 6. Specific gene sets were MT1D v HC, FC>1.25 up & p<0.05, cohort 1 and cohort 2; MT1D v HC, FC<0.8 down & p<0.05, cohort 1 and cohort 2; ST1D v HC, FC>1.25 up & p<0.05, cohort 1 and cohort 2; and ST1D v HC, FC<0.8 down & p<0.05, cohort 1 and cohort 2. See Methods for details of the analyis settings in R.

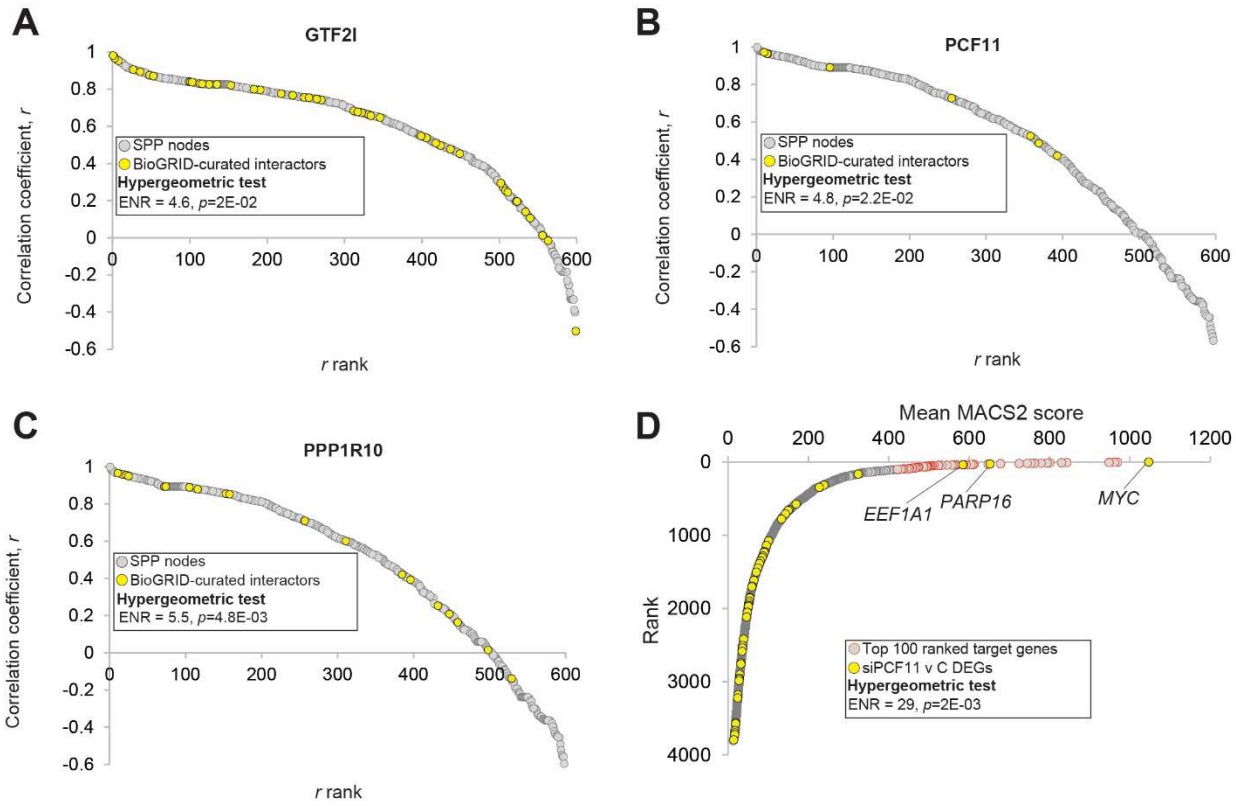

Figure S6. Use case 3: literature-curated interacting nodes for GTF2I, PCF11 and PPP1R10 are enriched among nodes whose HCT intersection profiles correlate most strongly with these three nodes. Related to Figure 6. A-C For each of the three nodes we retrieved the correlation coefficients (r) for all 599 other nodes from the results of the hierarchical clustering analysis. We then retrieved the list of unique curated interacting human proteins from BioGRID and performed hypergeometric distribution test for this set of interacting proteins. Hypergeometric test parameters: k = BioGRID curated interacting proteins; s = nodes with r ≥ 0.95; N = full list of 600 nodes that had at least two p<0.05 intersections across the eight gene sets. **D.** Validation of the human PCF11 ChIP-Seq consensome against a set of genes differentially expressed in response to PCF11 knockdown. Mean MACS2 score: mean peak height computed across publicly archived ChIP-Seq datasets in which human PCF11 is the IP antigen. Rank: rank of gene by mean MACS2 score where 1 is the highest average MACS2 score. Ref: K. Kamieniarz-Gdula et al, 2019.