

# Papalexi positive controls

2023-03-13

We look at SCEPTR (with skew-normal) and Seurat p-values for the positive control pairs on the Papalexi (gene) data. We indicate, for each gene targeted, whether it survived Mixscape and whether it is a transcription factor. Transcription factors are displayed first; they are top candidates for our discovery analysis.

Gene	Seurat p-value	SCEPTR p-value	Survived Mixscape	TF
STAT1	7.5e-172	2.8e-156	TRUE	TRUE
STAT2	1.7e-34	5.2e-34	TRUE	TRUE
STAT3	1.2e-14	1.2e-17	FALSE	TRUE
NFKBIA	2.1e-14	1.3e-13	FALSE	TRUE
IRF1	6.5e-14	6.1e-11	TRUE	TRUE
SMAD4	1.1e-05	1.7e-09	TRUE	TRUE
ETV7	1.5e-07	1.4e-06	FALSE	TRUE
SPI1	0.00065	6.7e-05	TRUE	TRUE
STAT5A	0.016	0.036	FALSE	TRUE
IRF7	0.00084	0.048	FALSE	TRUE
MYC	0.3	0.092	TRUE	TRUE
ATF2	0.59	0.45	FALSE	TRUE
POU2F2	0.72	0.61	FALSE	TRUE
JAK2	1.7e-192	8.2e-157	TRUE	FALSE
IFNGR2	3.1e-37	2.9e-38	TRUE	FALSE
CMTM6	3.8e-30	2.4e-30	FALSE	FALSE
UBE2L6	2.8e-23	2.1e-23	FALSE	FALSE
TNFRSF14	4.1e-14	7.6e-14	FALSE	FALSE
BRD4	8.5e-16	3.7e-10	TRUE	FALSE
CUL3	0.0016	0.0013	TRUE	FALSE
MARCH8	0.22	0.092	FALSE	FALSE
CD86	0.061	0.26	FALSE	FALSE
PDCD1LG2	0.55	0.38	FALSE	FALSE
CAV1	0.36	0.42	FALSE	FALSE
IFNGR1	0.1	0.87	TRUE	FALSE

- SCEPTR with skew-normal fit is quite capable of producing very small p-values!
- The SCEPTR and Seurat p-values are generally concordant. It's not atypical, however, for them to differ by an order of magnitude or two.
- Some perturbations pass Mixscape even though their Seurat positive control p-values are not small. This suggests that it is possible for a perturbation to have significant associations with expressions of genes that are not its direct target even if the association with its direct target is not significant.
- Some perturbations do not pass Mixscape even though their Seurat / SCEPTR positive control p-values are small. These are examples of Mixscape being too aggressive. In Papalexi et al, they say that they do not detect any differentially expressed genes for perturbations that did not pass Mixscape. This appears at odds with the fact that there are cases not passing Mixscape where the Seurat positive p-values are small. This discrepancy can be reconciled if the latter small p-values did not pass Seurat's other filters, e.g. log fold change.
- Initial perturbations to explore for our Papalexi discovery analysis are STAT3, NFKBIA, and ETV7. These TFs, for which SCEPTR's positive control p-values are small, did not pass Mixscape.