Supplementary Methods

Calculating perturbation signatures for single cells

Let $X = \{x_1, \ldots, x_N\}$ represent a normalized single-cell dataset, with *N* cells.

For each cell x_i, we perform the following procedure:

1. Identify Y_i, a subset of *X* consisting of cells that receive a 'non-targeting' gRNA, and were present in the same biological replicate *r* as x_i

Identify the set $\{y_{i,1}, \ldots, y_{i,k}\}$ of nearest neighbors to x_i, based on the top 40 principal components described above. We set the hyperparameter *k* = 20 by default, and identified neighbors using the Randomized Approximate Nearest Neighbors (RANN) algorithm [1].

2. Compute the average expression profile of this local neighborhood $\overline{y_i} = \frac{\sum_{j=1}^{k} y_{i,j}}{k}$

3. Compute the 'local' perturbation signature: $p_i = x_i - \overline{y_i}$

This calculation is implemented in the CalcPerturbScore function in Seurat.

Clustering single cells based on their perturbation signature

Perturbation signatures were centered but not scaled using the ScaleData() function. We ran PCA using the perturbation signatures of the top 2000 most variable genes defined using the RNA assay. The first 40 components were used as input for UMAP visualization in two-dimensions [2].

Mixture-model based classification of KO and NP cells

The objective of this procedure is to identify cells that received a targeting guide but exhibited no detectable transcriptomic evidence of perturbation. We perform the following procedure independently, for each targeted gene *g.*

1. We perform differential expression testing between all cells that receive gRNA targeting gene *g,* and all cells that receive a NT gRNA. The gene set *DEG* represents the set of genes that pass a Bonferroni-adjusted p-value threshold of 0.05. If *DEG* consists of fewer than five genes, we stop the procedure, and label all cells as non-perturbed.

2. *Let* $P^g = \{p_1^g, \ldots, p_N^g\}$, represent a set of single-cell perturbation signatures for *N* cells, each of which receives gRNA targeting gene *g.* Similarly, let $P^{NT} = \{p_1^{NT}, \ldots, p_M^{NT}\}$, represent a set of single-cell perturbation signatures for all *M* cells that receive a non-targeting gRNA.

3. For each cell, the perturbation signature is a vector, with length equal to the size of the *DEG* gene set. We project this into a single dimension, representing a perturbation score *s* for each cell. We find that reducing the dimensionality of these data substantially improves the robustness to overfitting in downstream analyses. To calculate the score, we first calculate a vector representing the difference in the average perturbation signature of targeted and non-targeted cells. We then project each cell's perturbation signature onto this vector. Specifically:

Let $\overline{p}^g = \frac{\sum_{i=1}^{N} p_i^g}{N}$ and $\overline{p}^{NT} = \frac{\sum_{i=1}^{M} p_i^{NT}}{M}$

Then the perturbation score *s* for cell *i* is defined by:

$$s_i^g = (p_i^g) \cdot (\overline{p}^g - \overline{p}^{NT}) \text{ and}$$
$$s_i^{NT} = (p_i^{NT}) \cdot (\overline{p}^g - \overline{p}^{NT})$$

4. We model the perturbation scores of non-targeting cells with a Gaussian distribution:
$$s^{NT} \sim N(u^{NT}, \sigma^{NT}) \text{ where } u^{NT} = \frac{\sum_{i=1}^{M} s_i^{NT}}{M} \text{ and } \sigma^{NT} = \frac{\sum_{i=1}^{M} (s_i^{NT} - u^{NT})^2}{M-1}$$

5. We model the perturbation scores of targeted cells using a mixture of two Gaussian distributions. One mode represents cells that resemble NT cells due to a lack of a detectable perturbation, and therefore is parameterized by the previously measured u^{NT} and σ^{NT}

$$Y_{NP} \sim N(u^{NT}, \sigma^{NT})$$
$$Y_{KO} \sim N(u^{KO}, \sigma^{KO})$$

$$p(s^g) = (1-\theta)N(u^{NT}, \sigma^{NT}) + \theta N(u^{KO}, \sigma^{KO})$$

This requires estimating three parameters: the mean and standard deviation rate for the perturbation score of KO cells (u^{KO}, σ^{KO}) and the mixing rate (or 'perturbation rate') θ. We learn these parameters using the function *normalmixEM* from the *mixtools* package [3].

6. We calculate the probability that each cell *i* was successfully perturbed by a gRNA targeting gene *g:*

$$p(i)^g_{perturbed} = \frac{1}{1 + \left(\frac{1}{\sigma^{NT}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{s_i^g - u^{NT}}{\sigma^{NT}}\right)^2} \Big/ \frac{1}{\sigma^{KO}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{s_i^g - u^{KO}}{\sigma^{KO}}\right)^2}\right)}$$

7. All targeted cells with a perturbation probability > 0.5 are classified as KO cells, while the remainder of cells are classified as NT cells.

8. We repeat steps 1-7 until the classifications converge. In this manuscript, all analyses converged within 5 iterations.

At the conclusion of this procedure, each cell is assigned one of three identities:

- If the cell received a NT gRNA, it retains its assignment as non-targeting (NT)
- If the cell received a targeting gRNA, and is classified in step 8 as NT, it is assigned a non-perturbed (NP) label
- If the cell received a targeting gRNA, and is classified in step 8 as KO, it receives a perturbed/knock-out (KO) label.

In addition to returning a KO or NP label, *mixscape* returns a perturbation probability (as defined in step 7) for each targeted cell.

This calculation is implemented in the RunMixscape function in Seurat.

Determining the optimal number of NT nearest neighbors for perturbation signature calculation

Our ability to calculate perturbation signatures is dependent on identifying a matched set of k-nearest neighbors within the control population. To assess our robustness to the parameter setting for k (k=20 by default), we varied k across a range of values (3, 10, 20, 30, 200), and ran the full *mixscape* workflow. To quantify performance, we compared the *mixscape* posterior probabilities and *mixscape* classifications with the values obtained with our default k=20 (Extended Data Figure 2A, B). We found extremely similar results spanning a reasonable range for k (10, 20, 30), but that very large deviations (i.e k=3, k200) did affect downstream results. We suggest that analyses with very small values of k (i.e. k=3) may be susceptible to technical variation due to the sparsity of single cell data. By contrast, setting a large value of k (i.e. k=200) may blend together cells in different biological states. We leave k as a user-defined parameter with a default of 20, and suggest setting values within the range of 10 to 30.

Finding k nearest neighbors using Seurat integration.

We also benchmarked an alternative workflow where control cells were identified after Seurat integration between non-targeting and targeted cells. Briefly, we split our cells into two objects one containing all non-targeting cells (NT) and the other all cells that expressed a gene targeting gRNA (T). After normalizing and finding variable features in each object separately, we used the FindIntegrationAnchors() and IntegrateData() functions to integrate the objects, specifying the NT object as a reference. We then calculated control neighbors based on the integrated datasets and proceeded with the full *mixscape* procedure. Based on the final *mixscape* classifications and posterior probabilities, we observed extremely similar results between the two workflows (Extended Data Figure 2C, D). We note that in cases where perturbations result in very strong transcriptomic shifts, performing integration may improve the ability to define matched control cells, although we did not observe meaningful differences here.

Benchmarking *mixscape* against MIMOSCA and MUSIC

MIMOSCA [4] and MUSIC [5] provide alternative computational frameworks for identifying non-perturbed cells in single cell pooled CRISPR screens. We ran MIMOSCA using the model-fitting procedure with default parameters, as specified in the 'Computational Workflow' section of the Github repository README (https://github.com/asncd/MIMOSCA). MIMOSCA requires a gene expression matrix and a file with all target gene classifications. For consistency, our gene expression matrix consisted of all genes used to build our *mixscape* classification model. To run MIMOSCA we used default parameters, which represented optimized values as described in the Perturb-seq publication (sklearn.linear_model.ElasticNet(l1_ratio = 0.5, alpha = 0.0005, max_iter = 10000).

Similarly to MIMOSCA, MUSIC requires the gene expression matrix and a file with all target gene classifications. For consistency, our gene expression matrix consisted of all genes used to build our *mixscape* classification model. MUSIC performs QC to remove low quality cells, and runs SAVER [6] on the gene expression matrix to impute mRNA expression values. The newly imputed matrix, together with the provided classifications, are used to classify cells. We ran MUSIC following the illustratrated example in the Github repository README (https://github.com/bm2-lab/MUSIC, with default parameters.

For benchmarking analyses, prior to running the three methods, we randomly sampled 1,000 cells expressing NT gRNA and re-labeled them as a new targeted gene class, representing a negative control (NEG CTRL). These cells should all be classified as NP (Extended Data Figures 5 and 6).

Linear Discriminant Analysis-based dimensionality reduction

After removing non-perturbed cells, we apply Linear Discriminant Analysis (LDA), followed by UMAP [2], to visualize the remaining cells in two dimensions. We apply LDA as an alternative linear reduction technique to PCA. While PCA aims to identify a low-dimensional subspace that maximally retains variation in a dataset, LDA aims to identify a low-dimensional subspace that maximally discriminates different groups of the data. In our case, the input to LDA is a single-cell data matrix and a set of group labels (the mixscape-derived classes).

In principle, we can use normalized gene expression as an input data matrix to LDA. However, this approach can lead to overfitting, as the total number of genes may be of a similar magnitude to the total number of cells. We therefore aimed to first reduce the dimensionality of our data in an unsupervised way, while retaining the sources of variation that distinguished each perturbation. We performed the following procedure for each targeted gene *g*:

1. From the previously computed set of perturbation signatures *P*, we extract all cells that are labeled by mixscape as KO for gene *g*, along with all non-targeted cells.
2. We perform unsupervised PCA. As input features to PCA, we use the gene set DEG, as previously calculated during mixscape classification.
3. We project this subspace onto all cells in the dataset.
4. We retain the first 10 projected components for all cell and expect that this subspace will retain differences between KO and NT cells.

At the conclusion of this procedure, all retained components are used as input to linear discriminant analysis using the lda function in the MASS R package [7]. This procedure is implemented in the MixscapeLDA function in Seurat.

The results from this function are used as input for 2D visualization with UMAP (Figure 3F). We found that this procedure substantially improved the visualization and interpretability of ECCITE-seq data. We observed that cells characterized by different perturbations separated visually in the 2D embedding, but retained their global structure (for example, STAT1, JAK2, IFNGR1 and IFNGR2 are all upstream regulators of the IFNγ pathway, and these clusters are adjacent on the visualization). Moreover, as described above, we randomly sampled 1,000 cells expressing NT gRNA and re-labeled them as a new targeted gene class, representing a negative control (NEG CTRL). Despite receiving a different label in the LDA procedure, these cells were indistinguishable from NT controls in the resulting embedding, demonstrating that our procedure does not overfit the data (Supplementary Figure 5).

Supplementary References

1. Arya S, Mount DM, Netanyahu NS, Silverman R, Wu AY. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. J ACM. 1998;45: 891–923.

2. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv [stat.ML]. 2018. Available: http://arxiv.org/abs/1802.03426

3. Benaglia T, Chauveau D, Hunter D, Young D. mixtools: An R package for analyzing finite mixture models. 2009. Available: https://hal.archives-ouvertes.fr/hal-00384896/

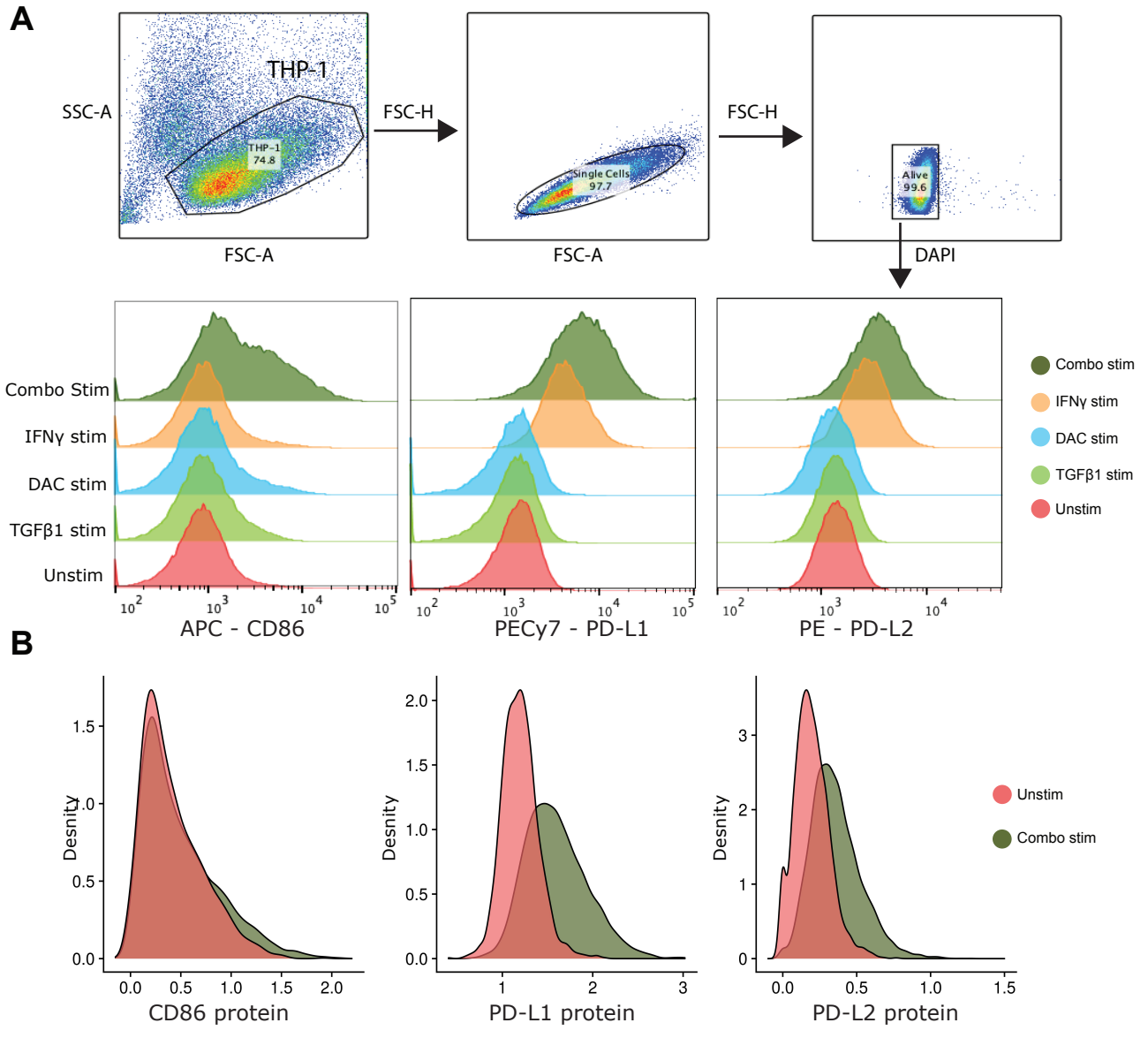
4. Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. Cell. 2016;167: 1853–1866.e17.

5. Duan B, Zhou C, Zhu C, Yu Y, Li G, Zhang S, et al. Model-based understanding of single-cell CRISPR screening. Nat Commun. 2019;10: 2233.

6. Huang M, Wang J, Torre E, Dueck H, Shaffer S, Bonasio R, et al. SAVER: gene expression recovery for single-cell RNA sequencing. Nat Methods. 2018;15: 539–542.

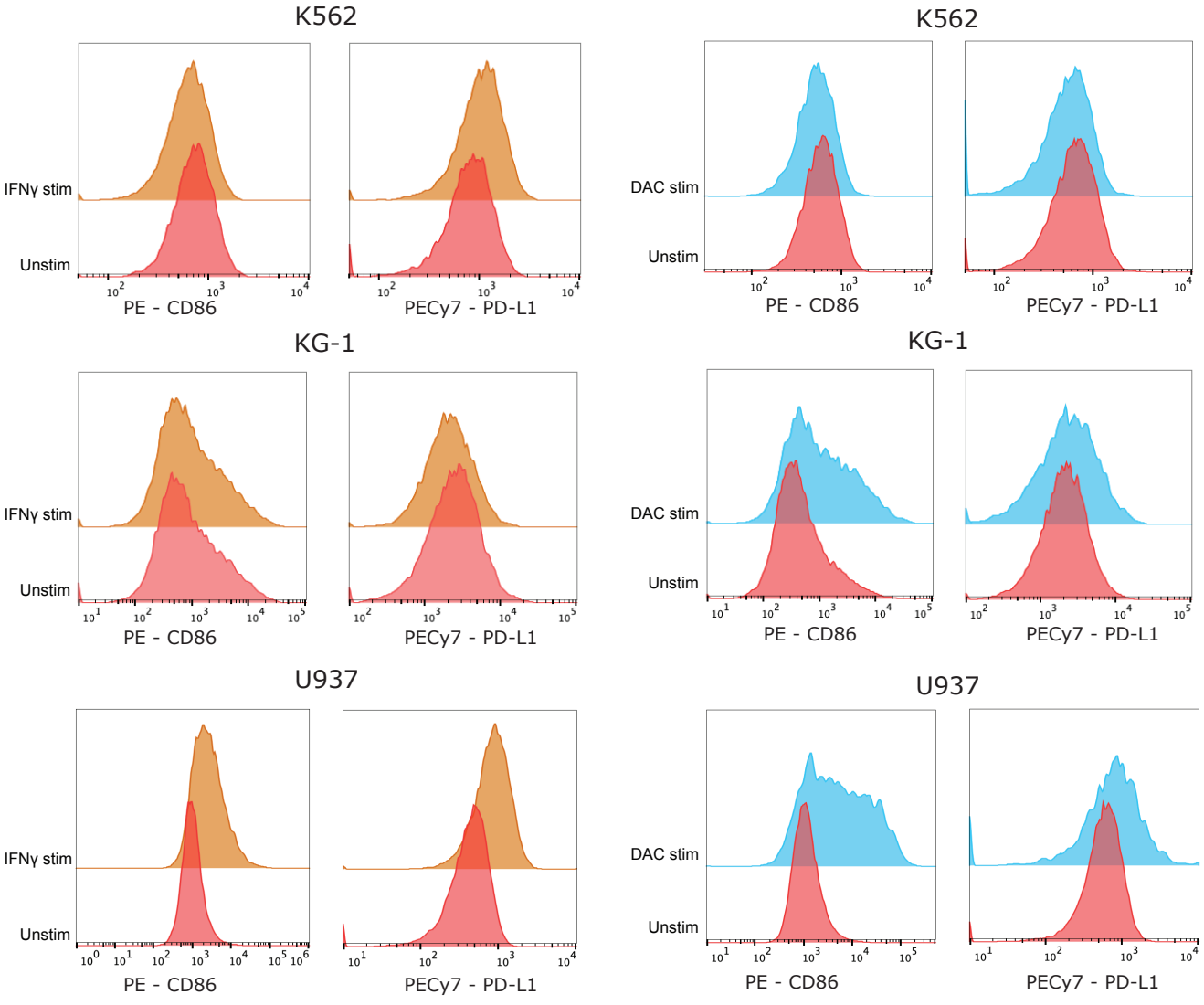
7. Venables WN, Ripley BD. Modern Applied Statistics with S. Springer Science & Business Media; 2013.

Supplementary Figure 1



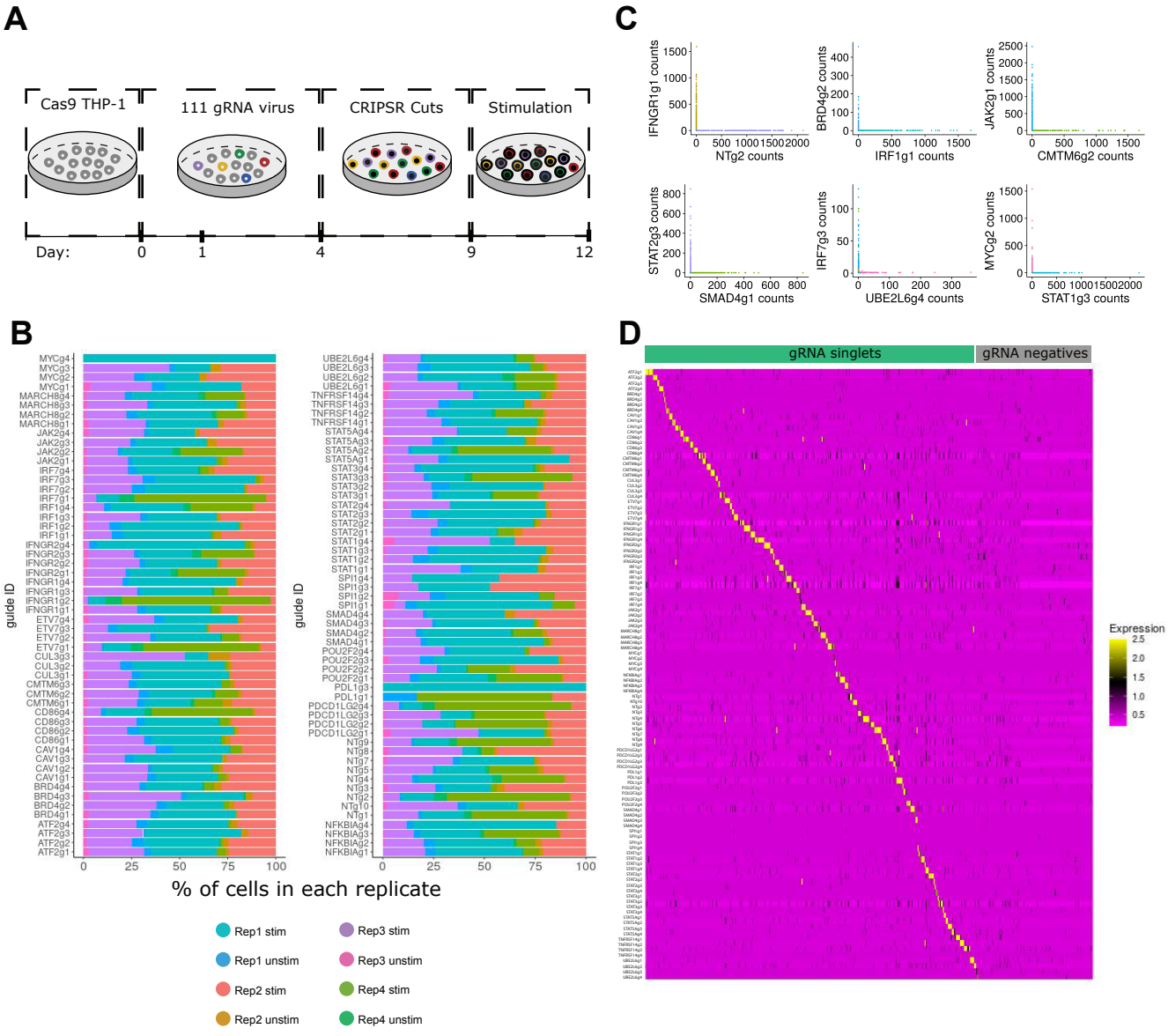
Supplementary Figure 1. THP-1 cell stimulation induces the expression of multiple immune checkpoint proteins (related to Figure 1).
(A) Top: Example gating strategy for flow cytometry to remove cell doublets and dead cells. The same gating strategy was used in Figures 1B, 3E, 4C, 4D, 5C and 5G. Bottom: Single-cell expression levels for CD86, PD-L1 and PD-L2 proteins across different stimulation conditions, as measured by flow cytometry. **(B)** Single-cell expression levels for CD86, PD-L1 and PD-L2 proteins across different stimulation conditions, as measured by CITE-seq. Relative levels are concordant with (A).

Supplementary Figure 2



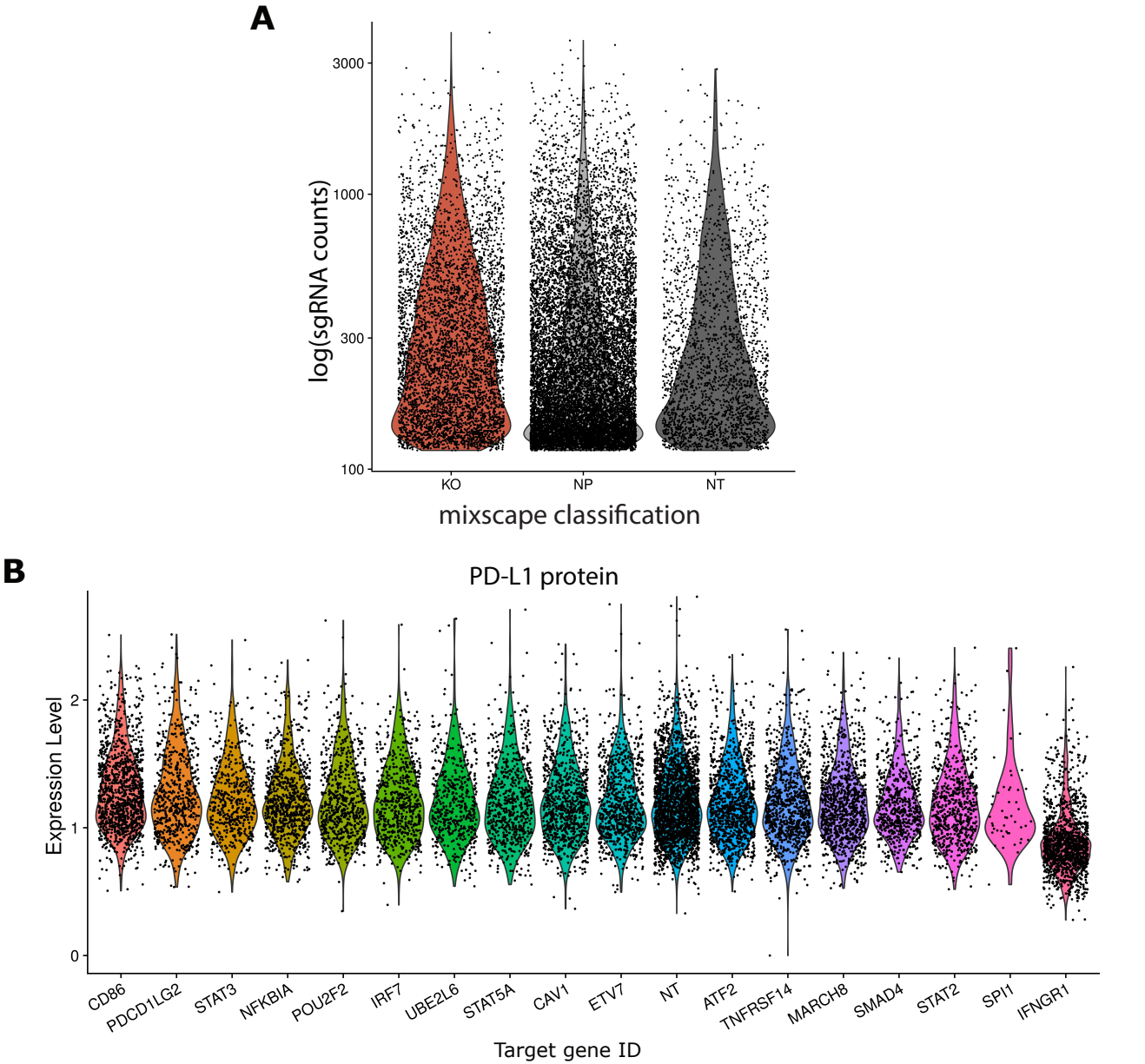
Supplementary Figure 2. Stimulatiom of K562, KG-1 and U937 cell lines induces the expression of various immune checkpoint proteins.
Single-cell expression levels for CD86, PD-L1 and PD-L2 proteins across different stimulation conditions and cell lines, as measured by flow cytometry.

Supplementary Figure 3



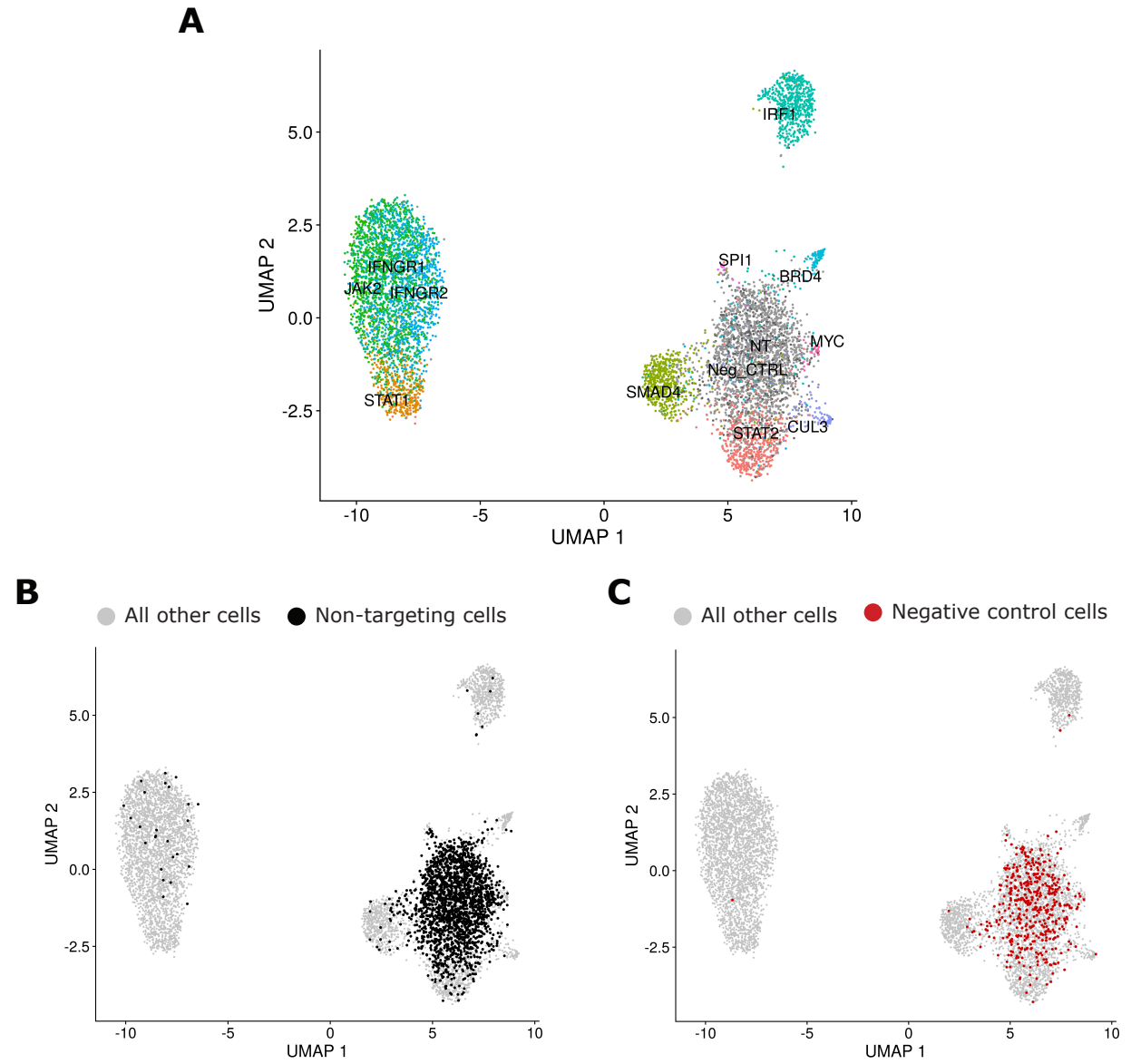
Supplementary Figure 3. Pre-processing of the 111-gRNA ECCITE-seq experiment (related to Figure 2). **(A)** Schematic representation of experimental setup. A Cas9-inducible THP-1 cell line was transduced with a virus containing 111 gRNAs (low MOI). Cells were treated with Blasticidin for three days to select for successfully transduced cells. Next, cells were treated with doxycycline to induce Cas9 expression and CRISPR-mediated gene KOs for 5 days. Lastly, cells were stimulated with TGFβ1, Decitabine and IFNγ for three days as previously described. ECCITE-seq was performed 12 days post viral transduction. **(B)** Barplots showing gRNA representation for each replicate (four biological replicates, split also by stimulation condition: stim and unstim). While replicates 1-3 had even representation of gRNAs, replicate 4 (oldest transduction) had skewed gRNA representation, likely as a result of long term cell culture. Replicate 4 was excluded from downstream analyses. **(C)** Example pairwise scatter plots showing mutually exclusive gRNA expression in single cells. number of sampled IFNGR1g2, PD-L1g1 and Non-targeting control cells. **(D)** gRNA expression heatmap showing the majority of cells express only one gRNA (singlets: green, negatives: grey). We also detected 30 gRNA doublet cells that are not shown here. gRNA counts are normalized (CLR normalization) and scaled (z-score).

Supplementary Figure 4



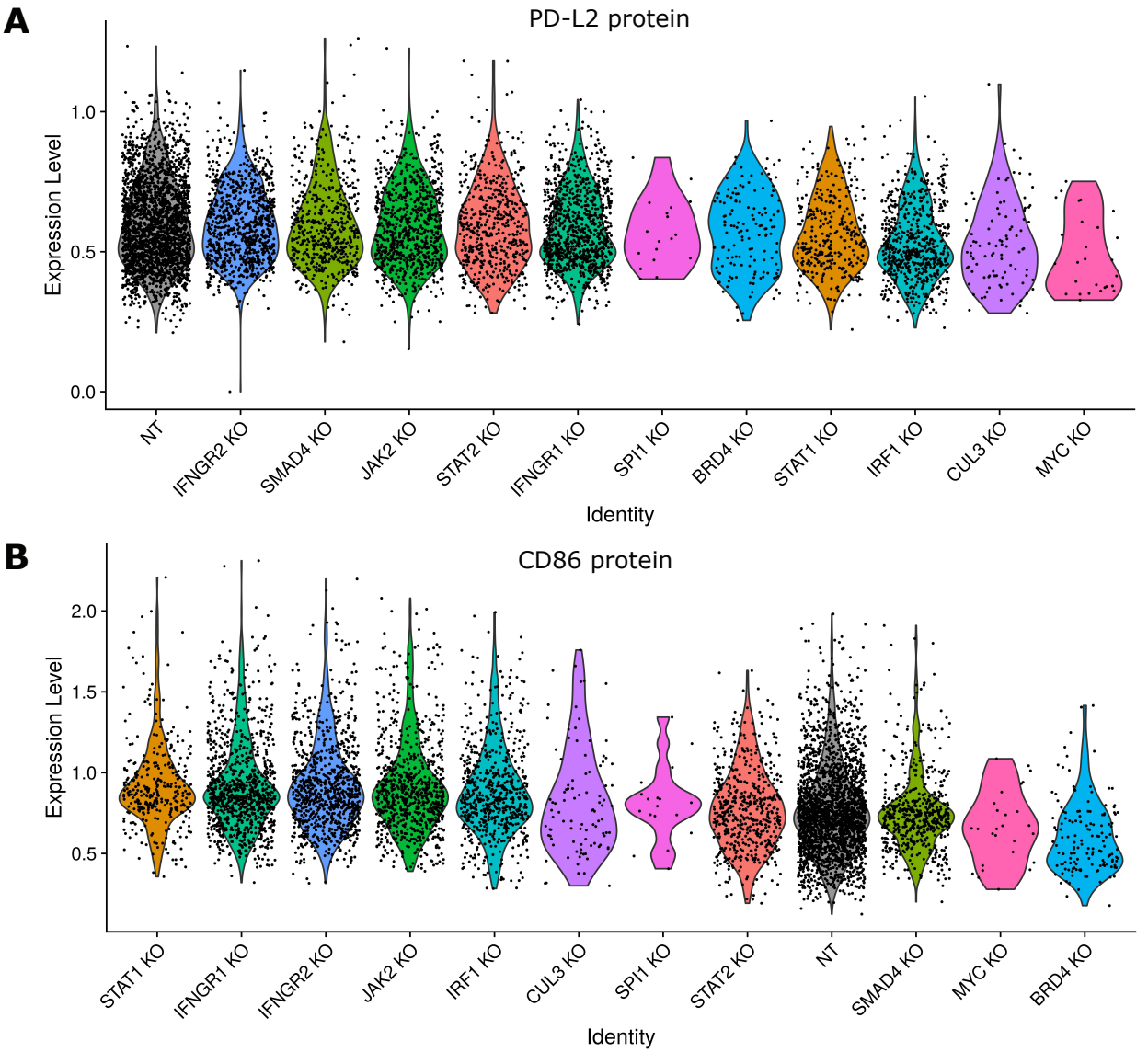
Supplementary Figure 4. *Mixscape* does not identify perturbed cells for all target genes.
(A) Violin plot showing the log-transformed gRNA counts in the dataset split by the global *mixscape* classification (NT, KO and NP). All three groups have similar count distributions suggesting that the NP cells are not a result of incorrect gRNA classification. **(B)** Violin plots showing PD-L1 protein levels in many target gene groups remain unchanged in response to their perturbation. For reference, IFNGR1 KO cells are included.

Supplementary Figure 5



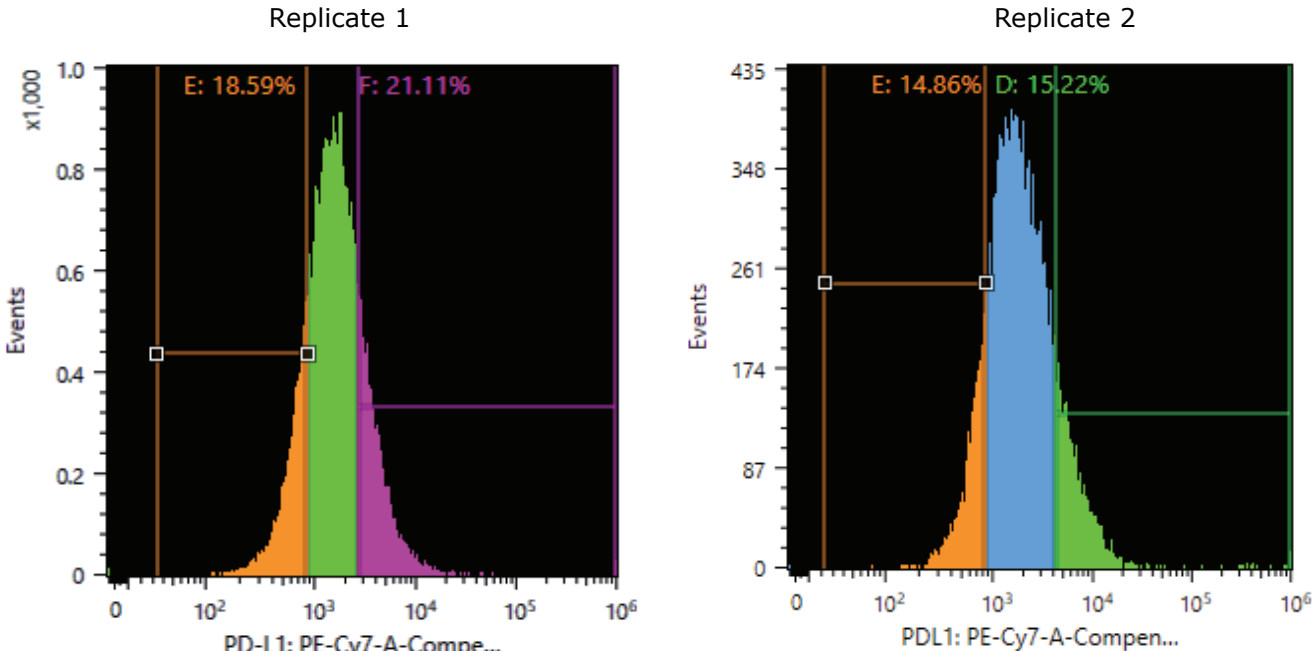
Supplementary Figure 5. Linear Discriminant Analysis for the clustering of ECCITE-seq data (related to Figure 3).
(A) UMAP visualization of all 7,421 NT, KO and negative control (Neg_CTRL) cells after running Linear Discriminant Analysis (Supplementary Methods). Cells from each targeted class separate visually on the 2D embedding. However, Neg_CTRL cells cluster with NT cells despite having a distinct label, demonstrating that our procedure does not overfit the data and induce separation when there are no transcriptomic differences. **(B)** Same as in (A), highlighting the non-targeting cells (black). **(C)** Same as in (A), highlighting the negative control cells (red).

Supplementary Figure 6



Supplementary Figure 6. CD86 and PD-L2 expression in KO cells.
(A) Violin plots showing PD-L2 protein levels across KO target gene classes. **(B)** Violin plots showing CD86 protein levels across KO target gene classes.

Supplementary Figure 7



Supplementary Figure 7. FACS gating strategy for validation pooled CRISPR screen (related to Figure 5). Flow cytometry plots showing gating for PD-L1 high and PD-L1 low expressing cells. Each gate represents 15-20% of the total number of cells in the data. Gating strategy is shown for both replicates and was used to obtain data for Figure 5B.