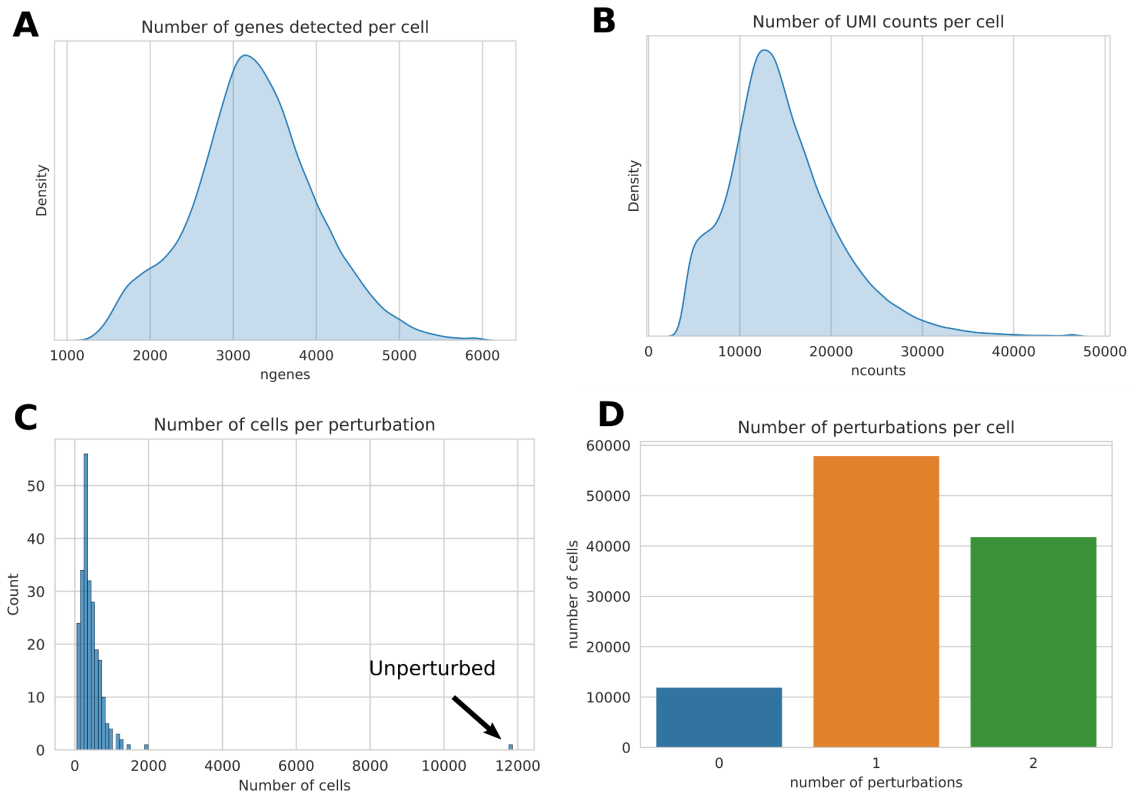


Supplement



Supp Fig 1: Exemplary information provided for each scPerturb dataset (here for (Norman et al., 2019)) (A) Number of genes that are detected with at least one count in a cell across all cells. (B) Total number of UMI counts per cell across all cells. Together with Supp Fig 1A this provides an overview over both sparsity and quality of the dataset. (C) Number of cells per perturbation. Depending on the application, perturbations with few cells can be filtered out before down-stream analysis. High imbalance in cell numbers per perturbation may also lead to biases in models. (D) Number of cells which received none, a single one, or two perturbations.