

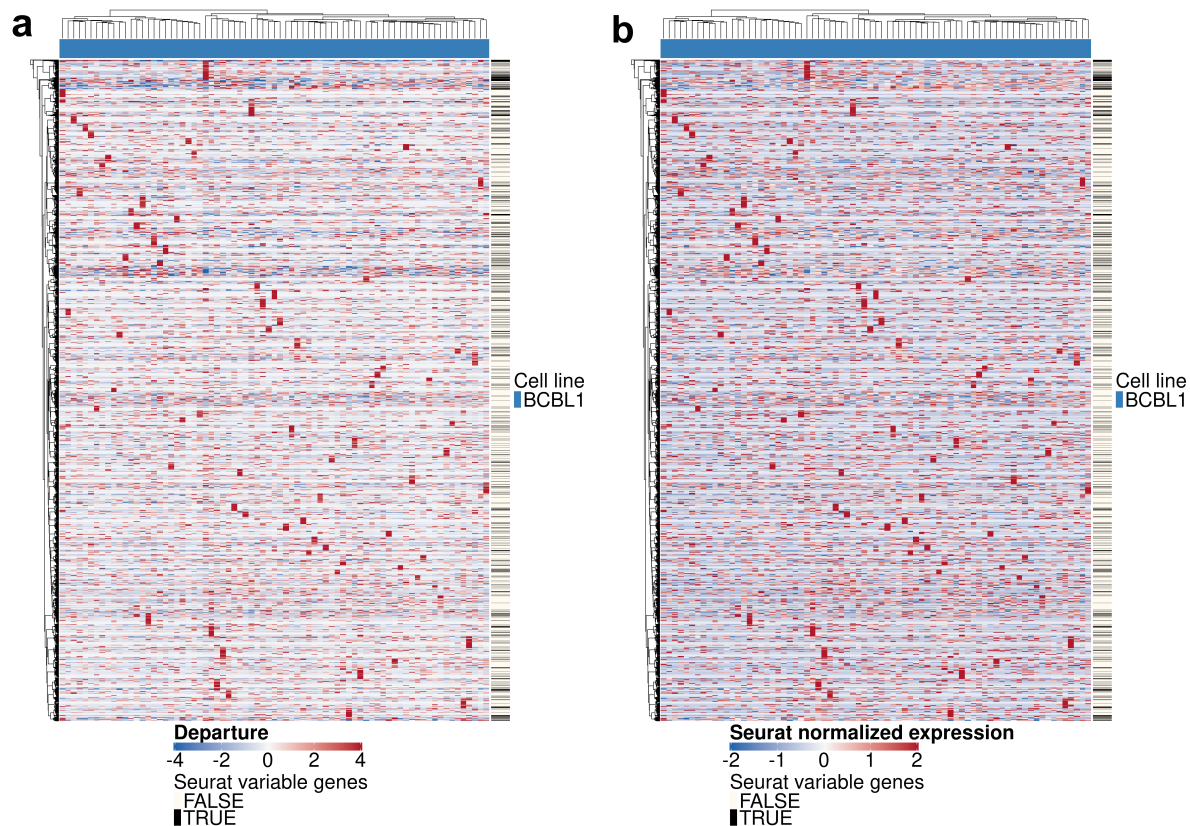


Figure S1: A heatmap view of the data representations of a single cell line data set (Plate 3). Data representations based on (a) DIPD and (b) Seurat normalized and scaled counts before feature selection. The black colored lines in the sidebars on the right represent the top 2,000 most variable genes kept by the Seurat pipeline. Visually, both data representations demonstrate this data set is homogeneous.