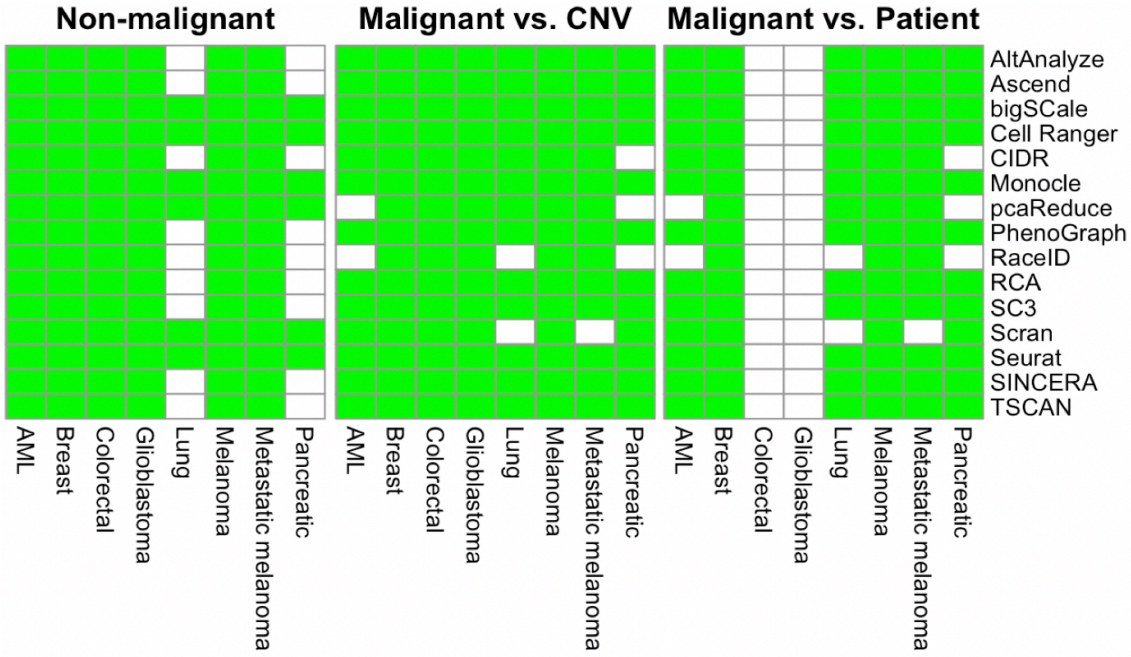
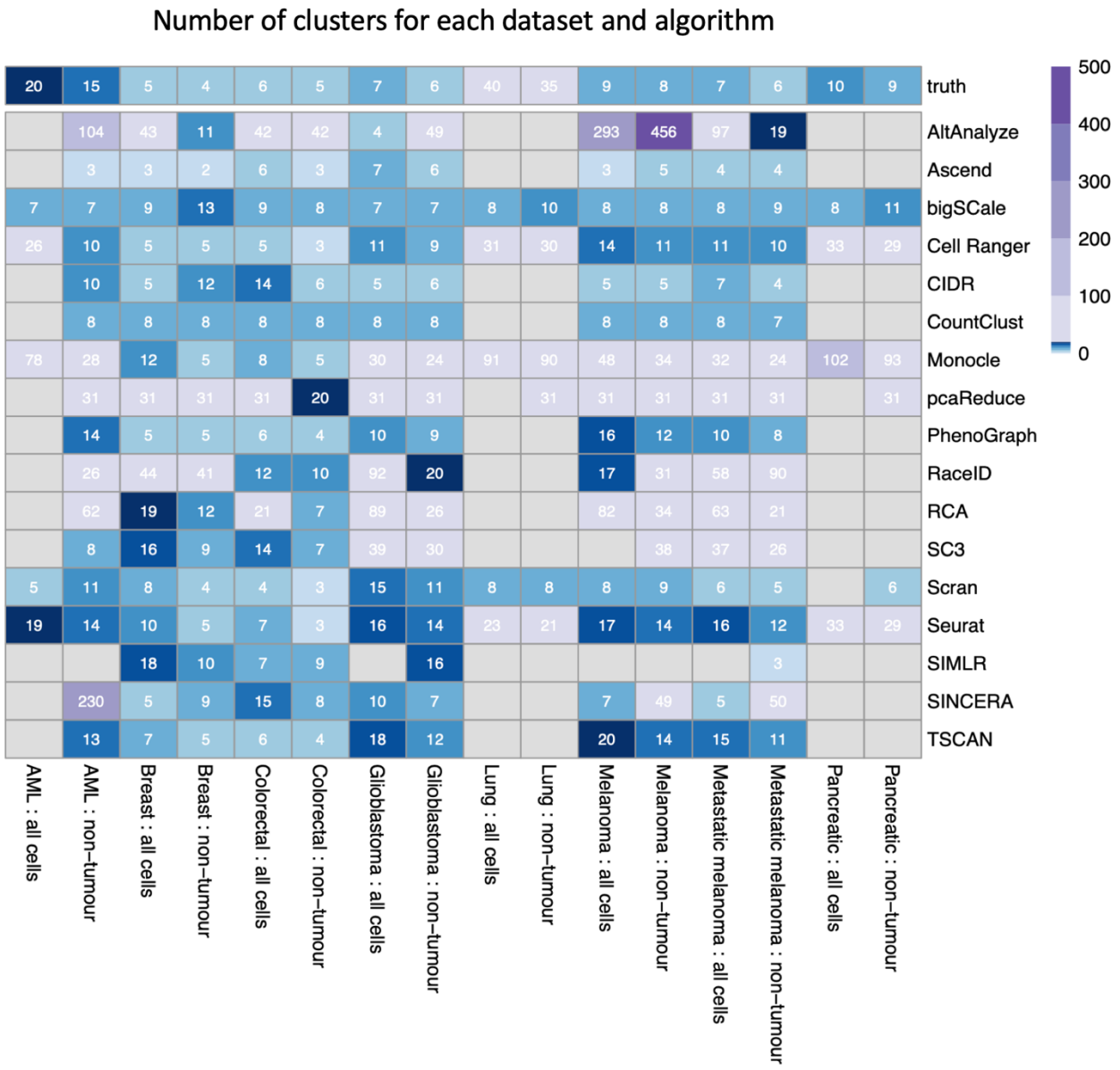
Evaluation of Single-cell RNA-seq Clustering Algorithms on Cancer Tumor Datasets

Alaina Mahalanabis, Andrei L. Turinsky, Mia Husic, Erik Christensen, Ping Luo, Alaine Naidas, Michael Brudno, Trevor Pugh, Arun K. Ramani, Parisa Shooshtari

Supplementary Figures



Supplementary Figure S1: Clustering quality was assessed for each of the 15 algorithms (table rows), using eight different scRNA datasets. For each dataset, the assessment was done separately on malignant or the non-malignant cells, thereby creating different data versions (table columns). The table represents clustering partitions for which the results were available (green), and those for which we are unable to obtain results, primarily due to the large dataset size (white).



Supplementary Figure S2: The top row shows the number of cell types in the original data annotations. The rest of the heatmap represents the number of clusters detected by each algorithm (rows). Columns represent each dataset when using either all cells or only non-malignant cells. In the majority of cases, the algorithms detected more clusters than the number of annotated cell types. Grey cells represent partitions that were unavailable.