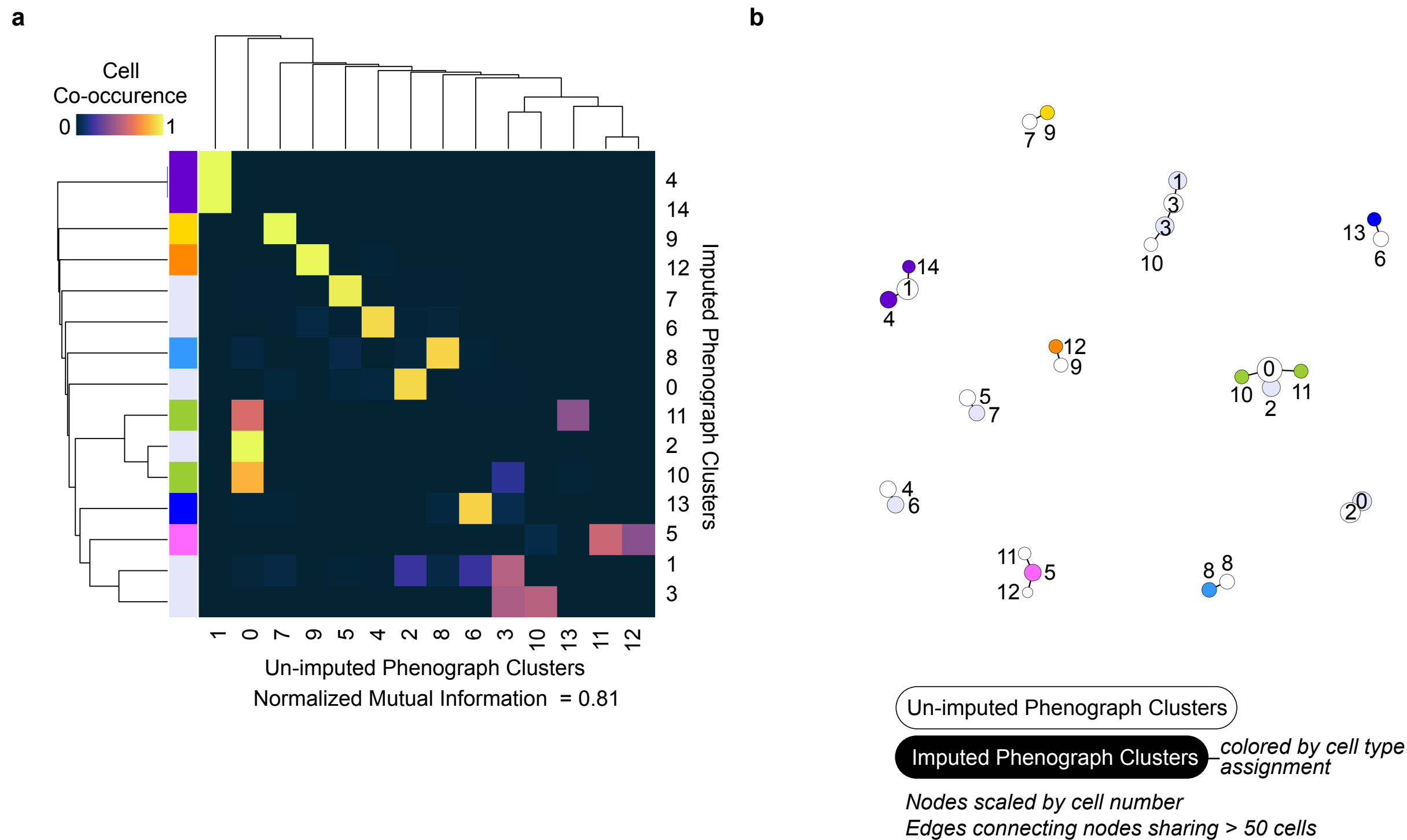


Supplementary Figures



Supplementary Figure 1. Robustness of cluster structure to imputation. **a**, Cell co-occurrence matrix within the epithelium derived from the normal lung and primary tumor when clustered independently using principal components applied to the normalized count matrix (Un-imputed Phenograph Clusters) or the imputed count matrix (Imputed Phenograph Clusters), as described in the Methods. The row of each Imputed Phenograph cluster is colored by its annotated cell lineage (as shown in [Fig. 2b](#)). **b**, A bipartite graph representation of the co-occurrence matrix. Each circular node represents an Un-Imputed or Imputed Phenograph Cluster; the nodes of Un-Imputed Phenograph Clusters are hollow and nodes of Imputed Phenograph Clusters are shaded by their annotated cell lineage. The diameter of each node is proportional to its cell number. Edges connect nodes sharing more than 50 cells.