



Supplementary Fig. 1. (A) UMAP plot depicting 41 clusters identified within the NSCLC patient dataset comprising 64369 cells. (B) Stacked bar plots representing the frequency of clusters 0 to 40 across the mutations in the dataset. (C) Markers and references for cell type assignment by cluster. Cell type annotation was performed using the following markers: Myeloid (*CD14*, *LYZ*), T cells (*CD3E*, *CD3D*), NK cells (*NKG7*, *GNLY*), cancer cells (*NAPSA*, *NKX2-1*), epithelial cells (*CAPS*, *SNTN*), B cells (*MS4A1*, *LY9*), fibroblast (*COL1A2*, *DCN*), endothelial cells (*FLT1*, *PECAM1*), plasma cells (*MZB1*, *JCHAIN*), and mast cells (*KIT*, *CPA3*). (D) Dot plot and violin plot representing the expression of cell type subset markers by 41 clusters. (E) Feature plot confirming the expression of cell type subset markers. NSCLC, non-small cell lung cancer; UMAP, Uniform Manifold Approximation and Projection.