

# Figure S1:

(A) Density plot of a distance matrix of the average expression levels of all genes detected in epithelial and non-epithelial cells for a representative patient sample (P2). (B) Heatmap showing large-scale CNVs for chromosomal regions (x-axis) and individual cells (y-axis) from a representative tumor (P2), inferred based on the average expression of 100 genes surrounding each chromosomal position (columns). Red: amplification; Blue: deletion; (C) Scatter plot for identification of malignant cells based on InferCNV analysis. ‘Hallmarks of Cancer’ was used to describe the specific characteristics that distinguish cancer cells from normal cells. We, therefore, evaluate the expressions of DNA repair and apoptosis (Hallmarks of Cancer) gene sets from the Molecular Signatures Database (MSigDB) in the cells classified as malignant and non-malignant by the two, methods -- Distance Profiling (DP) and InferCNV. Boxplots showing the average expression levels of genes involved in (D) Apoptosis and (E) DNA repair in distinct cell populations classified as malignant and non-malignant by InferCNV and DP methods.

