

## Supplementary Materials

### Supplementary Table legends:

**Table S1. Patient tissue information:** Clinical and culture properties of the 87 patient tumor tissues used in this study. Highlighted green and yellow samples are paired samples from two different patients.

**Table S2. CNA segments:** Final copy number segment file from SMASH data for all samples

**Table S3. RNA-seq sample information:** Sample details for samples used in transcriptomic profiling

**Table S4. RNA-seq signatures:** Gene expression signatures analysis for the various PDOs

**Table S5. Ki67 proliferation:** Percent Ki67 per organoid analysis for various normal and tumor PDOs

**Table S6. *In-vivo* transplant:** Overview of the PDO xenotransplant experiments N1 and N2, along with information on whether tumor formation and/or metastasis was observed.

**Table S7. Flow cytometry data and single cell experiment design:** Summary of EPCAM/CD49f flow cytometry data for the various lines. N= multiple passages for the tumor PDOs, and multiple different patient-derived normal organoid lines for the normal (NM0s) PDOs. Bottom panel: Experiment design for single cell experiment.

**Table S8. Differential gene expression of tumor vs normal LP cells:** Table of differential gene expression for comparing tumor LPs vs normal LPs

**Table S9. Cluster-specific gene markers:** Cluster specific markers for each of the clusters identified in TNBC PDOs

## Supplemental Figure Legends

**Figure S1.** related to Figure 1,2 **A. & B.)** Variant allele frequencies comparisons on a patient-wise manner for the various SNVs identified in **A.)** tumor and/or PDOs **B.)** different passage PDOs. R represents Pearson correlation. **C.)** Oncoplot for pathogenic variants in tumor vs scraping derived organoids. **D.)** Oncoplot for tumor tissue (p0 organoids) for samples that did not result in successful organoid cultures. **E.)** Loss of copy number alterations in one PDO model overtime in culture.

**Figure S2.** related to Figure 3 **A.)** Heatmap of sample-by-sample distance matrix showing Euclidean distances between the different PDOs. Samples were clustered using hierarchical clustering with Ward linkage. type: Normal= reductive mammoplasty derived normal organoids, Paired Normal= Adjacent or Distal to the tumor paired normal, Normal outgrowth= no pathogenic mutations were found, Luminal= ER/PR+ organoids; driver mutation: Other= trace mutations (see Fig 1E), None= no pathogenic mutations were found, NA= not assessed **B.)** Copy number profiles of NH48N and NH48T **C.)** Gene expression signatures for selected modules. Pink box highlights the samples defined as LT-TNBCs **D.)** Gene expression signatures for luminal-progenitor gene sets and c-MYC amplification. Each dot represents a different PDO; Luminal N=12, Normal N= 7, TNBC N= 19, TNBC met= 4. **E.)** Copy number data for selected MYC genes

**Figure S3.** related to Figure 3 **A.)** Selected culture images for organoids that did not result in successful cultures. **B.)** Ki67 IHC and H&E images for PDOs derived from normal and tumor pairs of the same patient, scale bars= 100µm. **C.)** Distribution of percent Ki67

positive cells per organoid for the various normal and tumor-derived organoids. The numbers on top of violin points indicate the number of organoids counted. The black line in the violin represents median. Data is plotted using GraphPad Prism. **D.)** Growth rate of organoid formation as measure by 3D Cell Titre Glo assay d6 post seeding, bars represent average with standard error and individual data points are measurements at different passages per PDO. Data is plotted using GraphPad Prism. **E.)** Dose response curves for three different drugs on 5 different PDO lines, data-points are presented as mean $\pm$ SEM and curve is the non-linear fit.

**Figure S4.** related to Figure 4 **A.)** H&E staining of PDO-X tumors from NH84T and NH93T that were not palpable and were only identified upon histological analysis. **B.)** Visualization of micro-metastasis and single cell metastasis using IHC for human mitochondria labeling in lymph nodes (upper panels) of the NOD/SCID mice transplanted with various PDOs. Bottom panels show human mitochondria labeling in lungs, liver and lymph-node from HCM-CSHL-0655-C50. Black arrows point to human mitochondria labeled micro- or single cell-metastasis. **C.)** CD117 and CK5/6 labeling of sections from an AdCC-like patient tumor (upper panels) and the PDO-X derived tumor (lower panels) from the same patient, scale bar on bottom panel=100 $\mu$ M.

**Figure S5.** related to Figure 5 **A.)** Flow-cytometry contour plots for EPCAM and CD49f for selected tumor PDOs (in red contours) overlaid onto normal-derived organoids (in blue contours). **B.)** Expression of ITGA6 and EPCAM from RNA-seq data for the different PDOs. Dark pink dots: true normal PDOs, light pink dots: paired normal PDOs and black

dots: tumors. Circled region represents long term propagating organoids. **C.)** Different passage flow-cytometry plots for NH95T, NH85TSc and HCM-CSHL-0366-C50 showing stability of the cell populations over time in culture. **D.)** Flow-cytometry plots for some more TNBC and lymph-met derived organoids.

**Figure S6.** related to Figure 6 **A. & B.)** Cell cycle related **A.)** UMAP-plot and **B.)** gene expression in the cycling clusters **C.)** Cluster specific markers associated with each G1 cluster identified **D.)** Expression patterns of the various luminal/basal markers in normal and TNBC PDOs. HC-0366 is PDO HCM-CSHL-0366-C50 and HC-0655 is PDO HCM-CSHL-0655-C50 **E.)** Mammary lineage scores for each of the PDOs in specific clusters **F.)** Gene-set scores for each of the specified gene-set between the tumor and normal samples. **G.)** Motif enrichment for differentially expressed genes that were significantly upregulated in tumor LPs compared to normal LPs (clusters 2, 7 and 11). Score=  $-10 \cdot NES \cdot \text{padj.value}$  **H.)** Representative culture images for three different PDOs treated with DAPT and MYCi **I.)** Normal organoid formation from sorted EPCAM+/CD49f+ and CD49f+ve populations from three different normal PDO lines in DAPT or MYCi containing medium. Significance was assessed by multiple-comparisons by two-way ANOVA ns= not significant, \*\*\* pvalue<0.005.

**Figure S7.** related to Figure 7 **A.)** Cluster distribution for each of the TNBC PDOs tested. **B.)** EPCAM and ITGA6 expression for all the TNBC samples **C.)** Feature plots showing mammary lineage scores for each of the PDOs in specific clusters **D.)** Boxplots showing scores for the different gene-sets tested **E.)** Feature plots showing the gene-set score

distribution among the different TNBC PDOs. The first panel is the combined score distribution for all samples and provides the legend for the whole row. **F.)** Marker expression for one gene per cluster for the combined tumor organoids (first panel) and for individual TNBC organoids.

Fig S1

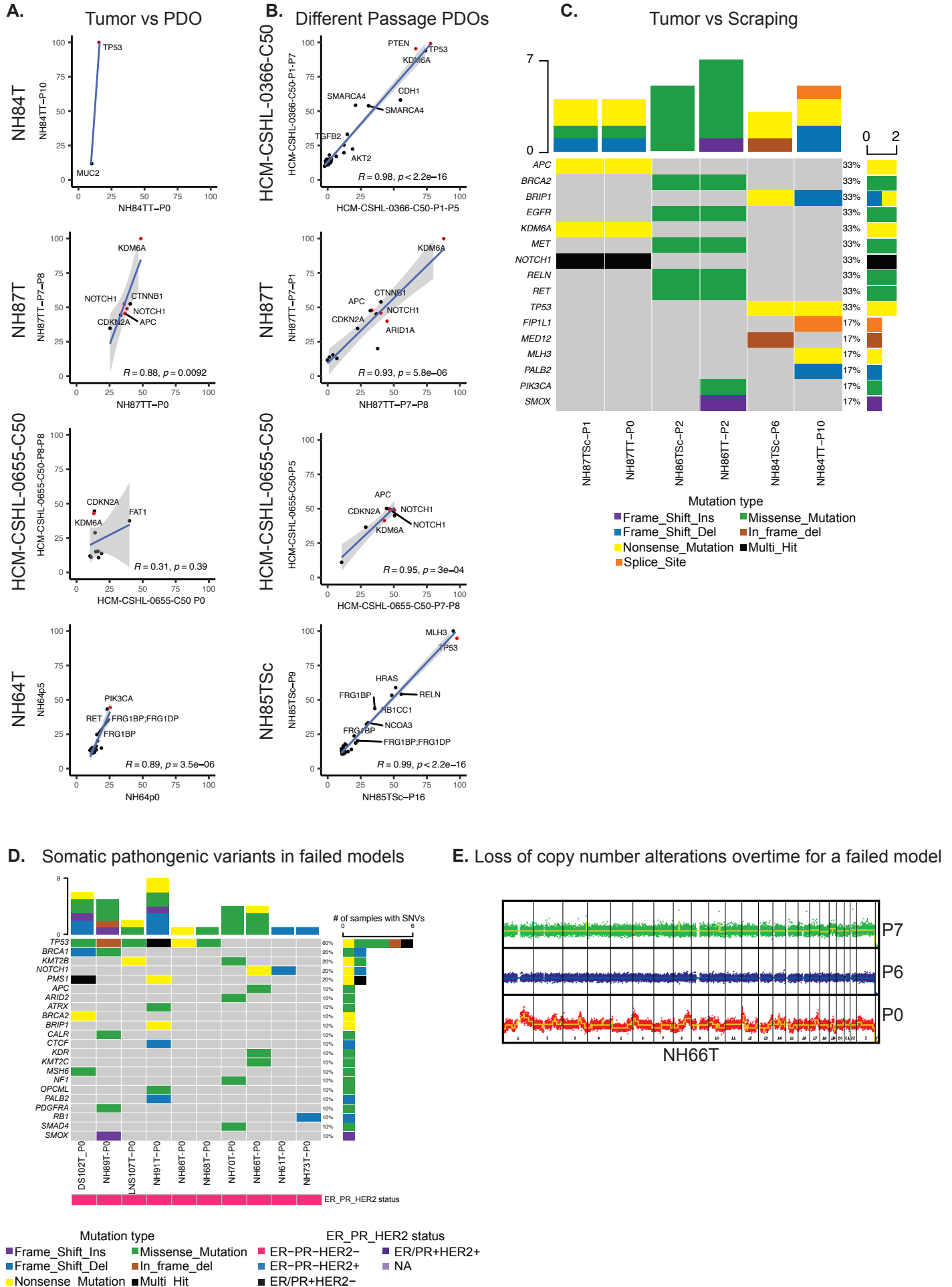


Fig S2

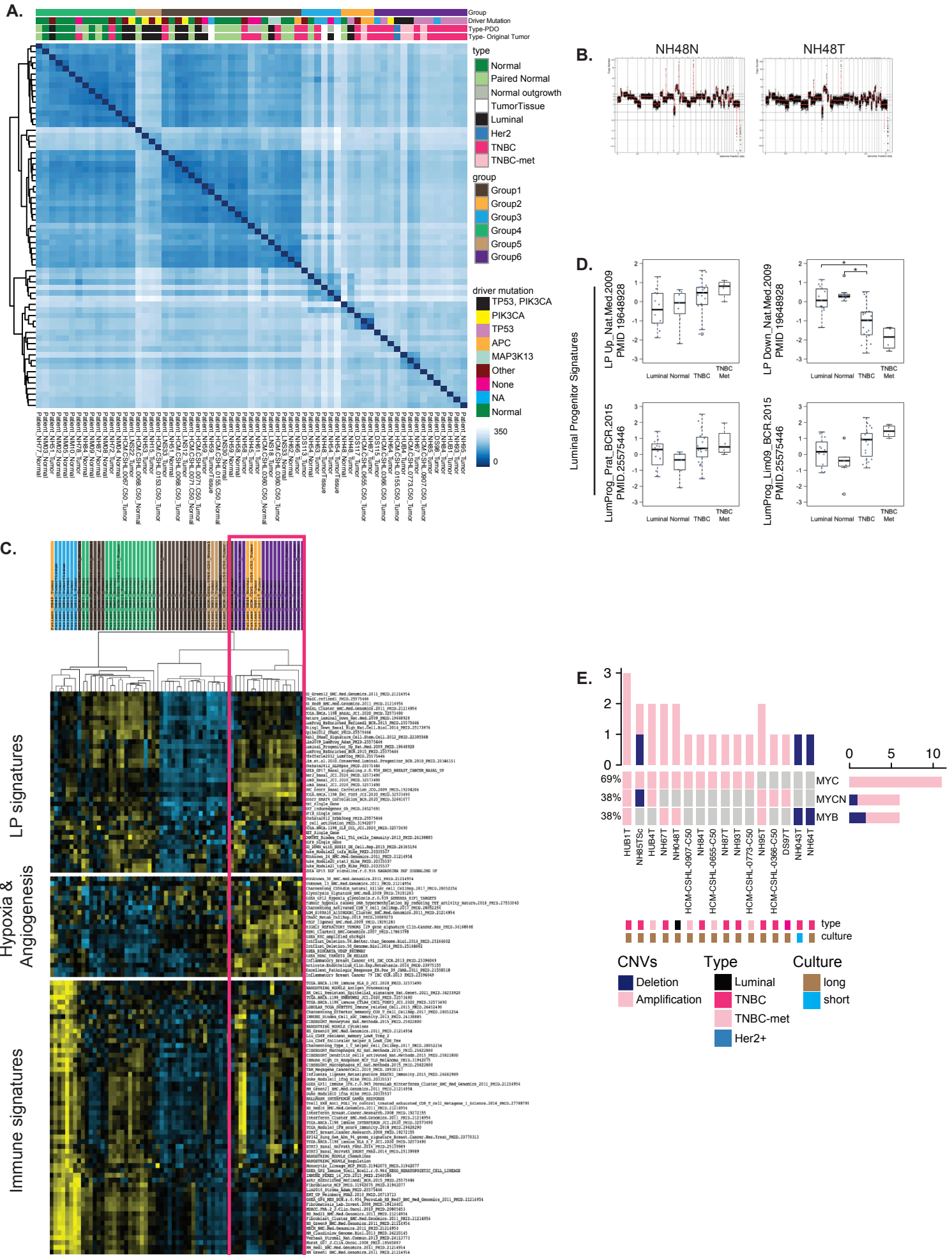


Fig S3

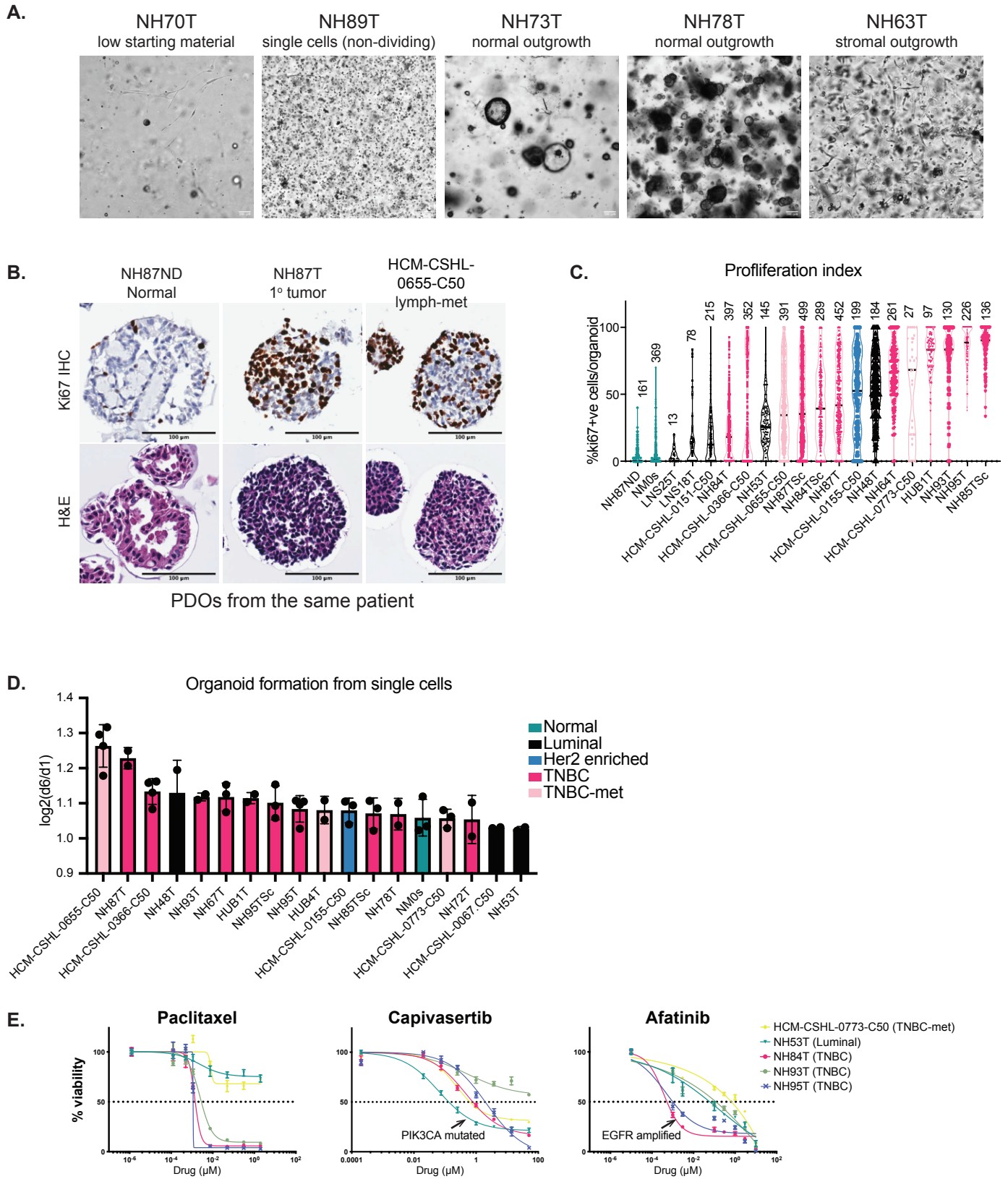




Fig S4

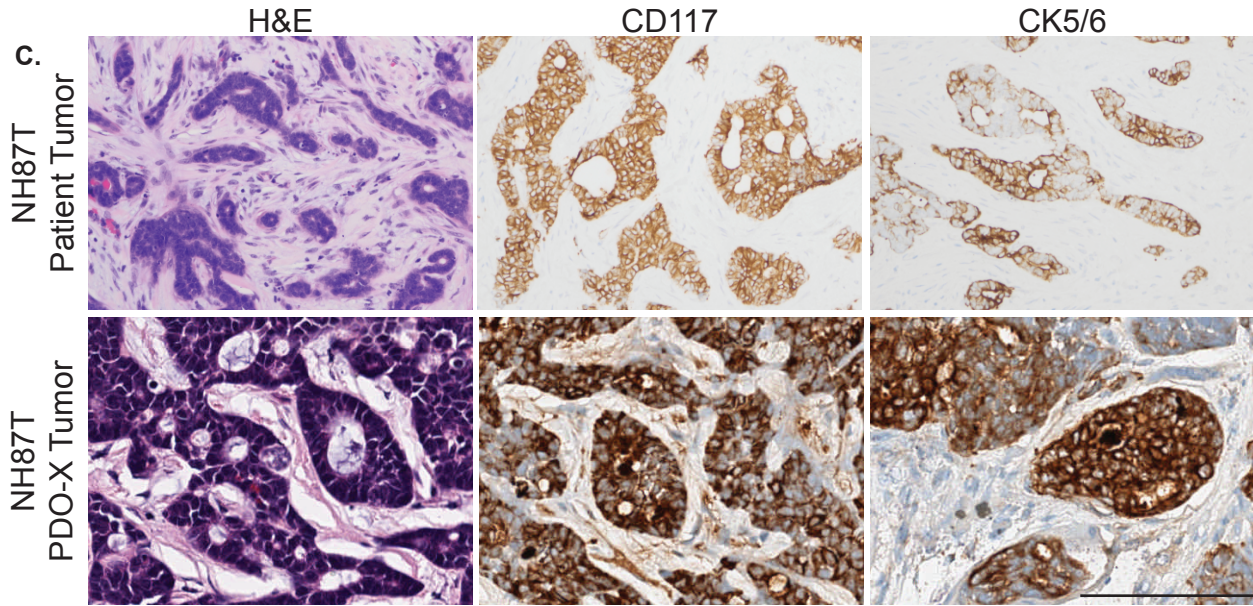
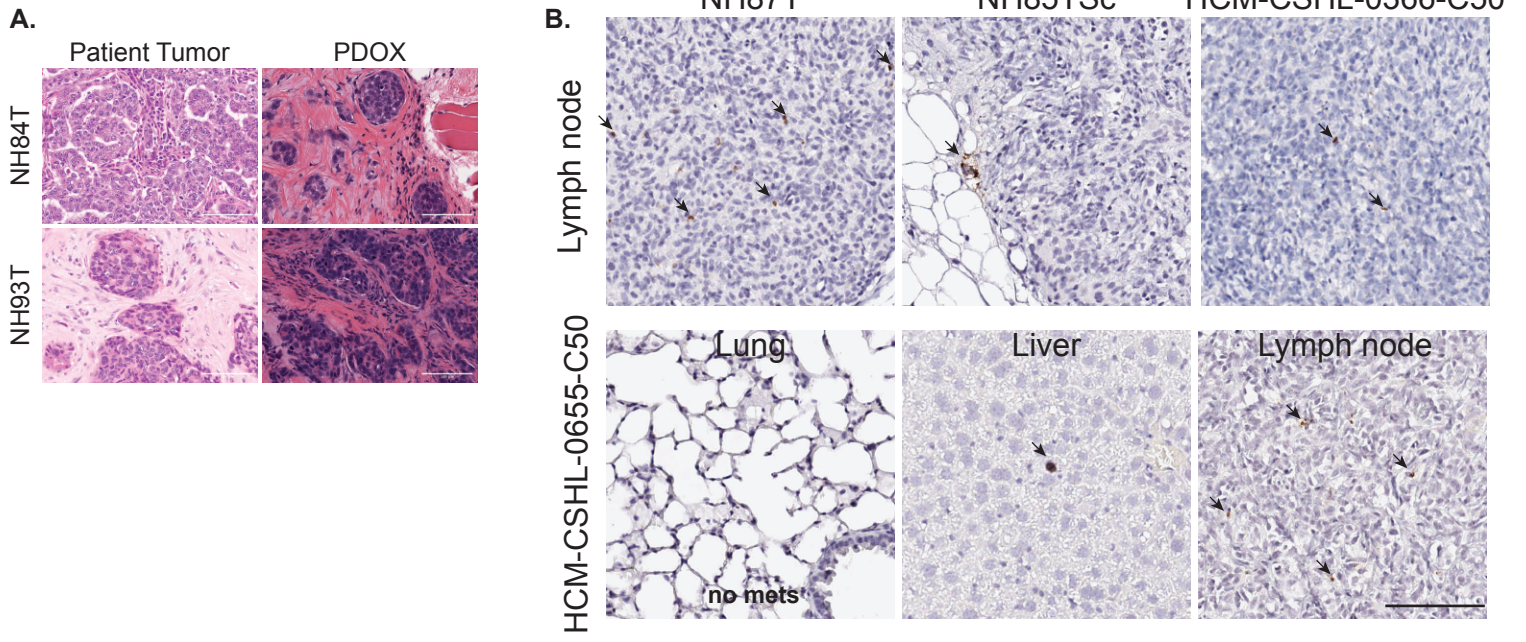


Fig S5

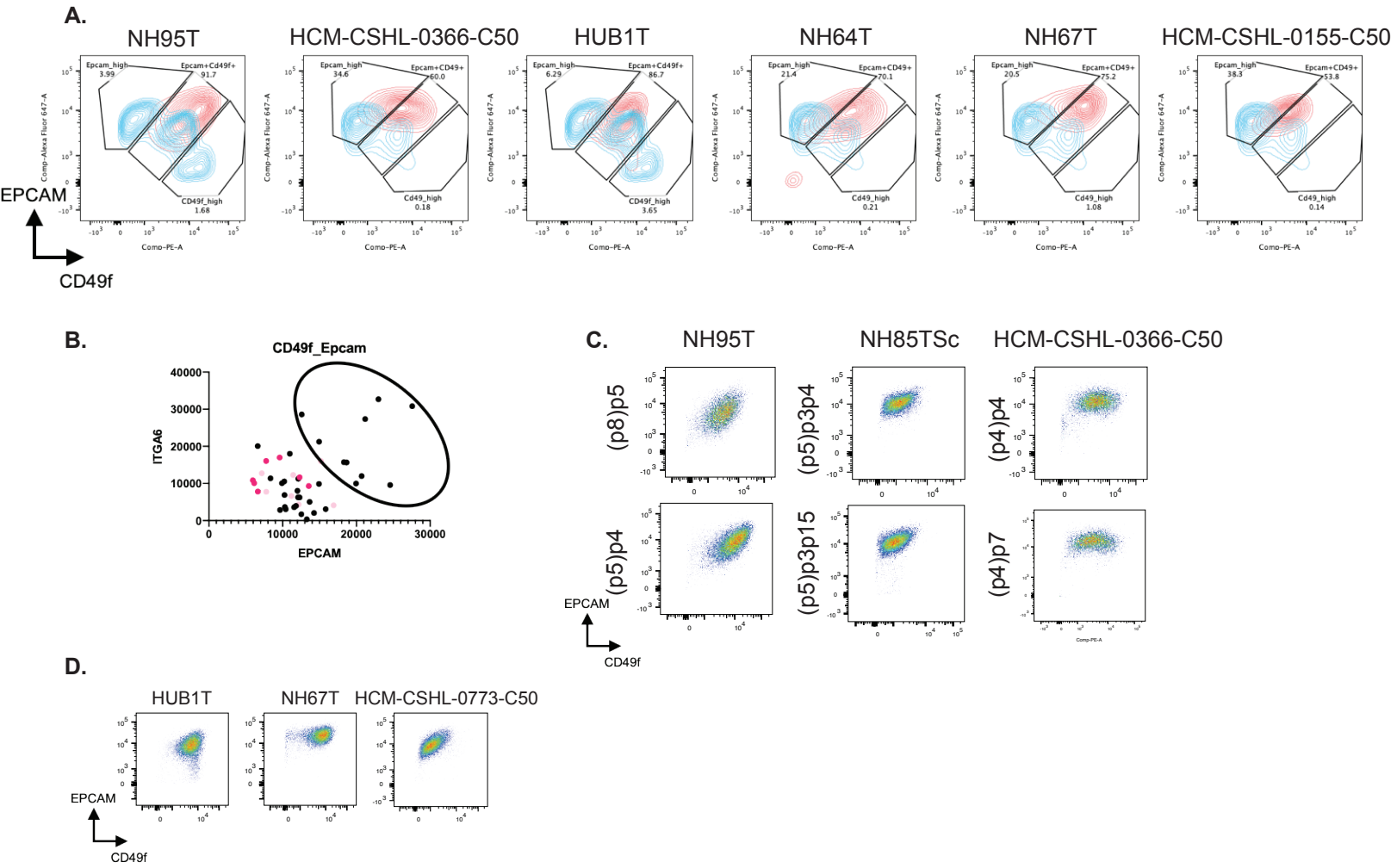


Fig S6

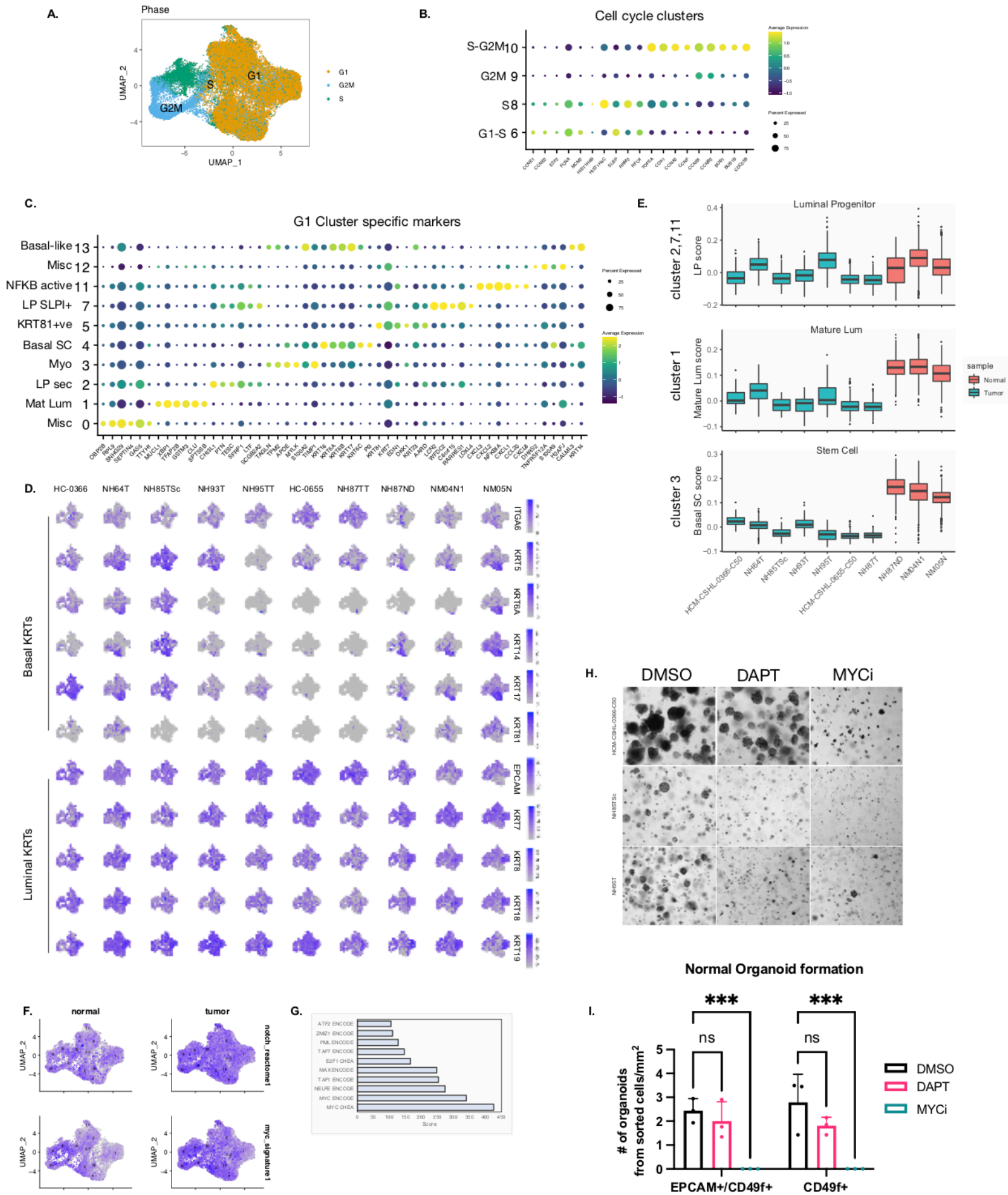


Fig S7

