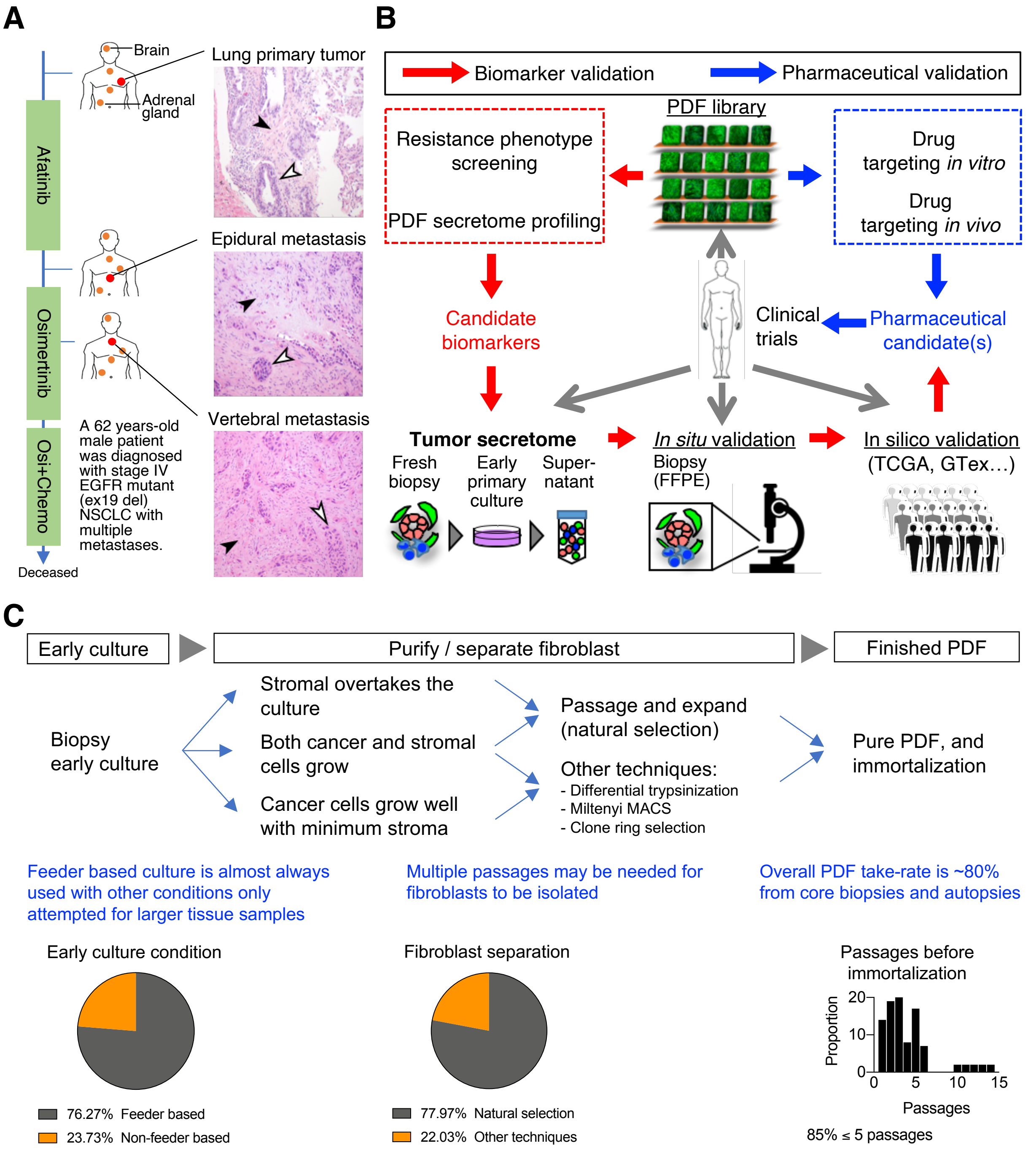


Figure S1. Establishment of patient-derived fibroblasts from tumor biopsies. Related to Figure 1. **A.** Biopsies from an *EGFR*+ lung cancer patient shows abundant tumor stroma is present at different timepoints and in multiple metastasis sites. Desmoplastic areas (black arrows) constitute the microenvironment surrounding cancer cells (white arrows) in the primary lung lesion and metastatic lesions. **B.** Schematic of the PDF library establishment, experimental use and data integration. **C.** The workflow of PDF model establishment process and validation.

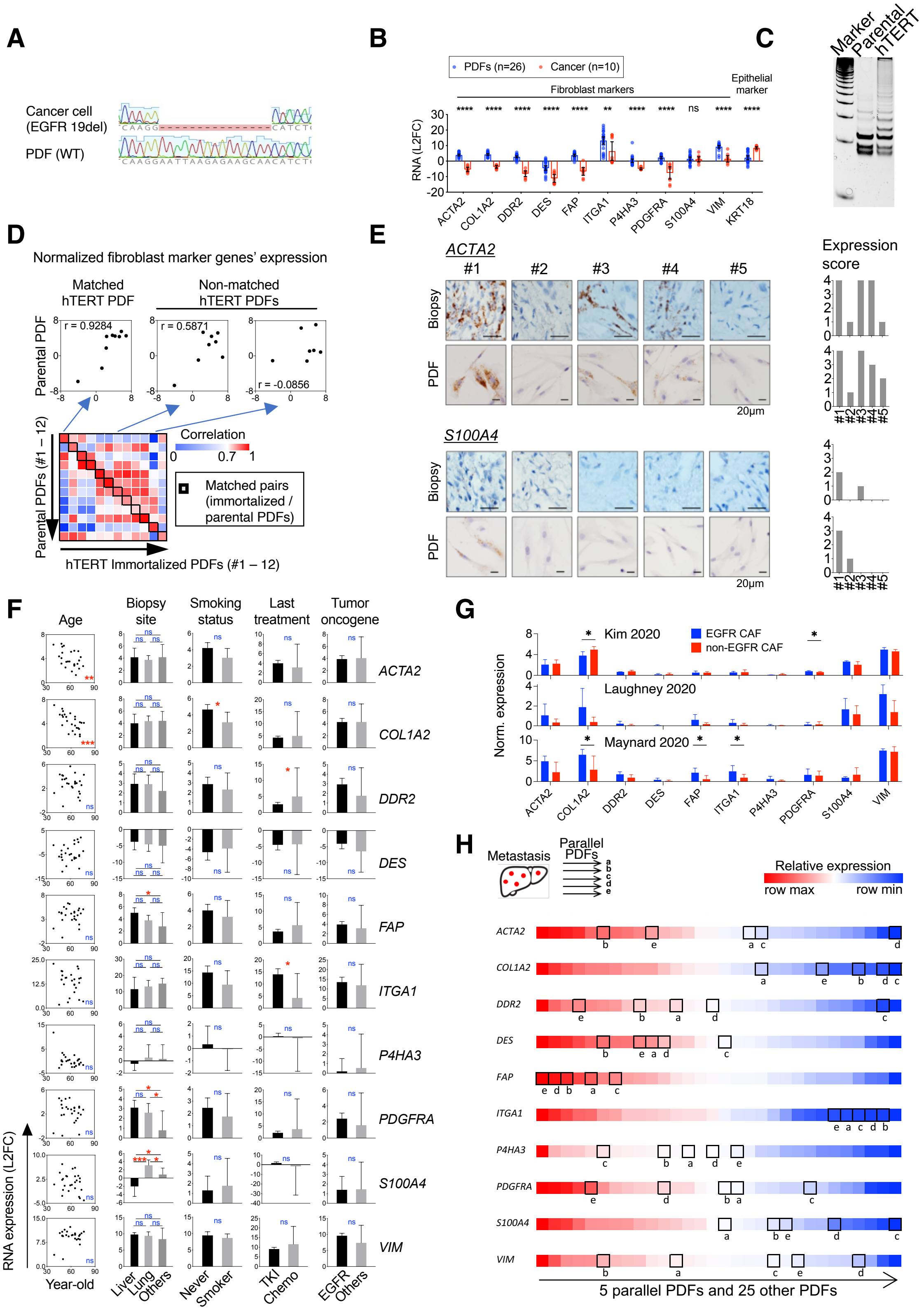
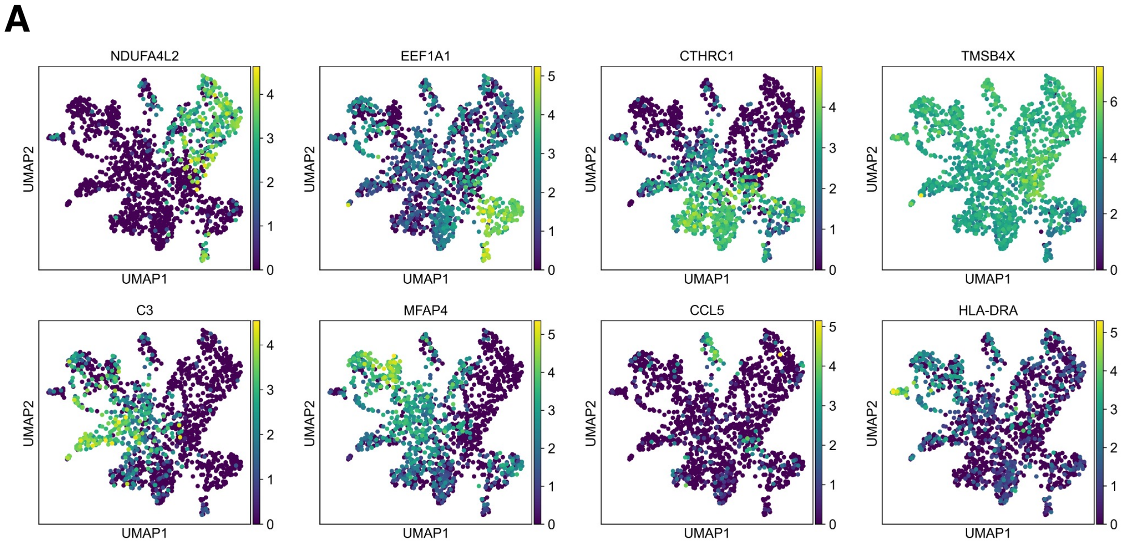


Figure S2. Validation of PDFs and their expression of CAF markers. Related to Figure 1.
A. An example of *EGFR* genotyping analysis comparing PDF and the corresponding cancer cells. **B.** The mRNA level of 10 CAF marker genes and the epithelial marker gene KRT18 assessed by q-PCR in PDFs

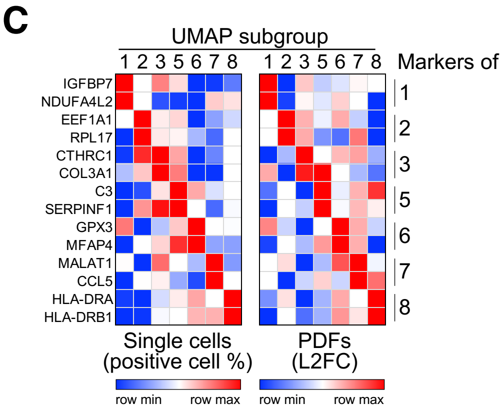
(n=26) and cancer cell lines (n=10). Means with 95%CI are shown. Two-tailed t test is used. **C.** Example of the results of TRAPeze assay showing the telomerase activity in a hTERT transduced PDF and its parental PDF. **D.** Fibroblast marker genes (n=10) signature assessed by qRT-PCR in 12 parental PDFs and compared to their hTERT immortalized derivatives. **E.** Expressions of fibroblast markers ACTA2 (αSMA) and S100A4 (FSP1) in *EGFR+* lung cancer sample and PDFs developed from the corresponding biopsies. In situ staining is performed by RNAscope, and the result is scored based on manufacture's semi-quantitative criteria. **F.** Breakdown of PDFs' expression of CAF markers (n=10) based on patients' clinical features. Patients' clinical features are shown in x-axis and RNA expression (qRT-PCR, L2FC) is shown in y-axis. Mean value with 95%CI is shown (two-tailed Spearman's r for age correlation analysis and t-test for group comparison). **G.** Expression of CAF markers in lung CAF single cells based on the oncogene background of NSCLC. Mean value of CAFs in each tumor sample is first calculated, and bars show the average of all samples in each group. Two-sided t-test is used. **H.** Five liver metastasis samples obtained from a single patient were analyzed for the expression of CAF marker genes using qRT-PCR. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.



B

1	2	3	4*	5	6	7	8
IGFBP7	EEF1A1	CTHRC1	TMSB4X	C3	GPX3	MALAT1	HLA-DRA
NDUFA4L2	FTH1	COL3A1	RPLP1	CCDC80	MFAP4	CCL5	TYROBP
RGS5	RPL17	VCAN	RPS14	SERPINF1	A2M	CXCR4	HLA-DRB1
COX4I2	GABARAP	COL1A1	RPS12	DCN	CFD	BTG1	FCER1G
PPP1R14A	GAPDH	BGN	SERF2	JUN	RGCC	TMSB4X	HLA-DPA1
CALD1	NBEAL1	COL1A2	LGALS1	NNMT	TIMP3	B2M	CD74
PDGFRB	MIF	LUM	RPS18	PTGDS	SEPP1	CD52	LYZ
HIGD1B	RPL3	POSTN	RPL34	C7	CYR61	ARHGDIB	SRGN
COL4A1	LDHA	AEBP1	RPLP2	RARRES1	ZFP36	CCL4	LAPTM5
TINAGL1	HNRNPA1	SULF1	RPS19	IGF1	PTGDS	HLA-E	HLA-DPB1
LHFP	RPS2	COL5A2	TMSB10	C1S	DUSP1	PTPRC	AIF1

* UMAP-4 is corresponding to poor quality single cells, and is excluded from further analysis



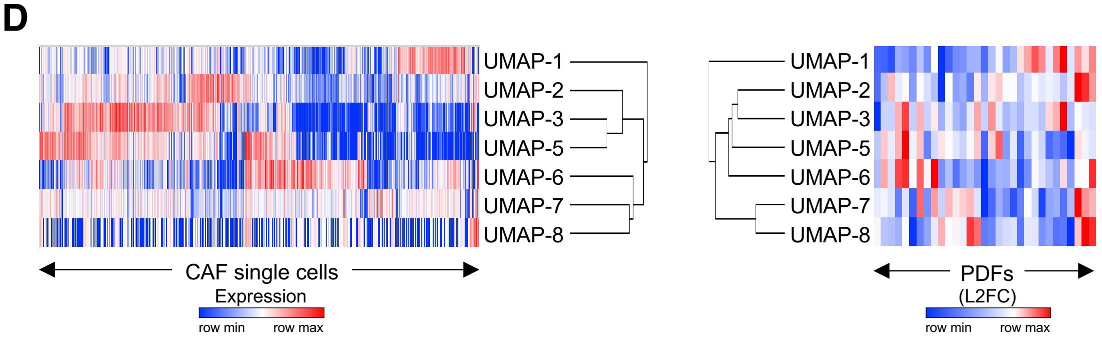


Figure S3: PDFs captured the overall molecular repertoire of NSCLC CAFs. Related to Figure 1. **A.** Uniform Manifold Approximation and Projection (UMAP) analysis of 1,465 individual fibroblasts in NSCLC resection samples from Lambrechts et al. (2018) reveals 8 molecular classes. The expression of representative marker genes is shown by each UMAP class. Color scale is based on the raw counts. **B.** The top 10 markers per UMAP class. UMAP-4 corresponds to poor quality single cells and is excluded from further analysis. **C.** PDFs mapped to each UMAP class consistently expressed the corresponding marker genes defined by CAF single cells. **D.** Unsupervised clustering shows the signature proximity between PDF molecular classes (right) and that between CAF single cells (left). Each row is the average of top marker genes per UMAP class, each column is a unique PDF or CAF single cell.

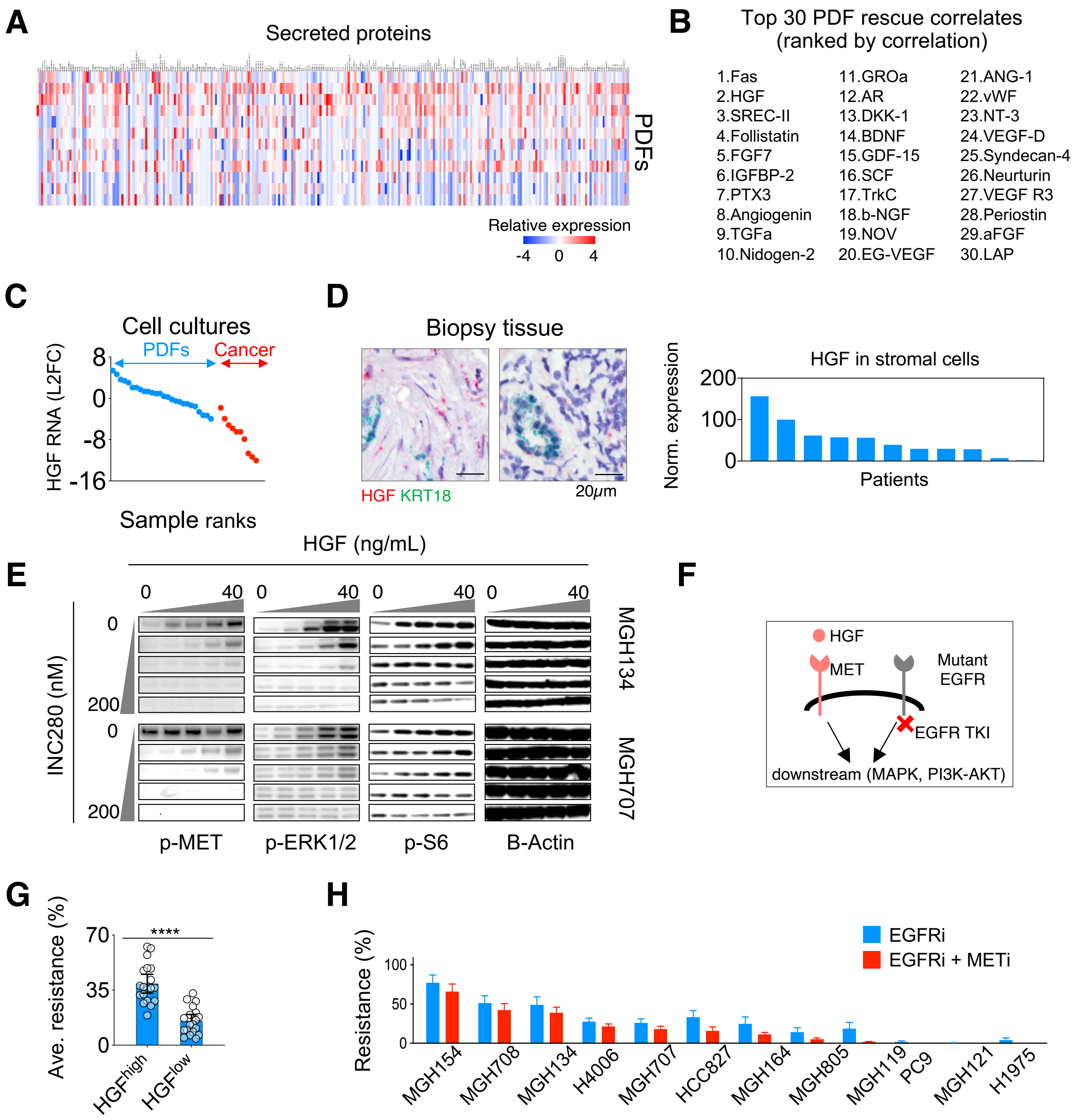


Figure S4: HGF partially defines PDFs functional heterogeneity in conferring EGFRi resistance. Related to Figure 3.

A. Comprehensive secretome profiling was carried out in 12 PDFs, with the heatmap showing the profile of secretion of 273 proteins detectable in at least one PDF. **B.** The list of secreted proteins strongly correlated with PDF rescue. **C.** The heterogeneous HGF mRNA expression in PDFs. Cancer cells' expression is shown for comparison. **D.** RNAscope staining of HGF in stroma cells in *EGFR*+ lung cancer (n=11). HGF expression is measured by staining quantification (dots per 100 stromal cells, see Methods). **E.** Signaling changes induced by recombinant HGF (0, 0.625, 2.5, 10, and 40ng/mL) added to EGFR-driven NSCLC models in the absence of presence of the MET inhibitor INC280 (200nM). MET phosphorylation and downstream pathway activation measured by western blotting with the indicated antibodies. **F.** Schematics illustrating the bypass signaling mediated by alternative receptor kinase activation such as HGF-MET. **G.** Comparing PDFs with higher than medium HGF secretion with PDFs with lower than medium HGF secretion, assessed by ELISA, among 38 *EGFR+* PDFs in rescuing cancer cells from EGFRi effect**. H.** The effect of METi in countering PDF-driven resistance to EGFRi in EGFR+ cancer models. Each bar is the average effect of 38 EGFR+ PDFs, with error bar showing 95%CI.

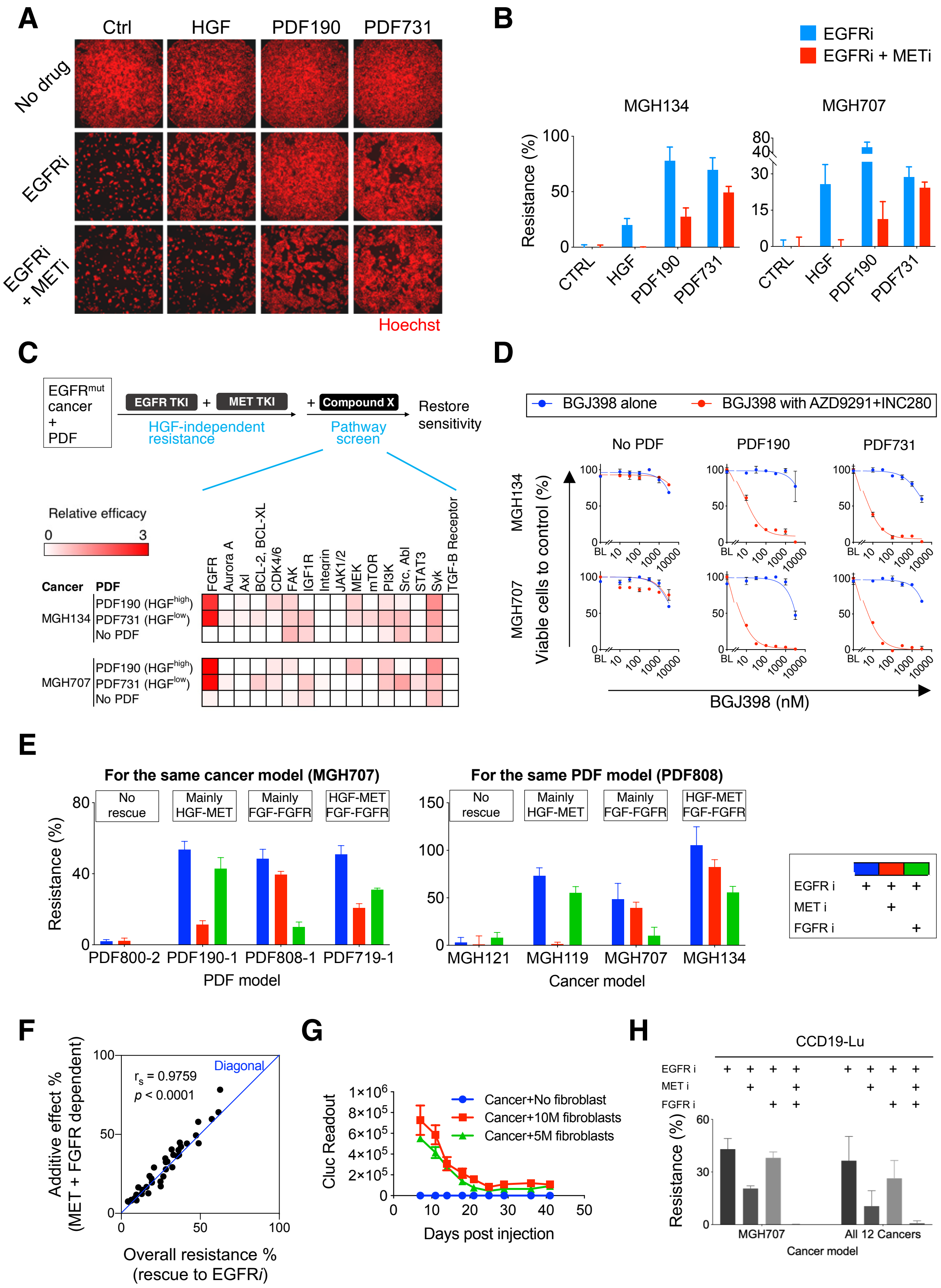


Figure S5: PDFs-driven activation of MET and FGFR pathways in cancer cells conferred resistance against EGFRi. Related to Figure 3 and Figure 4.

A-B. Representative images (**A**) and quantification (**B**) of HGF^{high} PDF (PDF190) and HGF^{low} PDF (PDF731) rescuing cancer cell lines MGH134 (A and B) and MGH707 (B) from EGFRi alone or EGFRi in combination with METi. Values are the averages of 4 replicates with error bars representing 95%CI. **C.** Efficacy of indicated compounds to negate HGF-independent resistance across 2 PDFs and 2 cancer cells. The relative efficacy (IC_{50} shift) of each compound is measured with the compound's IC_{50} in the presence of EGFR and MET dual inhibition subtracted from the IC_{50} when the testing compound is used alone. **D.** Dose response curves for the effect of BGJ398 alone or in the presence of EGFRi + METi on cancer cells' viability. **E.** Examples showing different modalities of PDF-mediated rescuing activities of different PDFs on the same cancer cell line (left) and the same PDF on different cancer cell lines (right). **F.** The PDFs' overall rescue activity against EGFRi compared with the additive effects of FGFR-mediated rescue against EGFRi plus METi and MET-mediated rescues against EGFRi plus FGFRi. Each dot represents a PDF's effect. **G.** Lack of persistence of human fibroblast in nude mice. Fibroblasts (CCD19-Lu) were engineered to secrete luciferase (Cypridina Luciferase, Cluc) and co-injected with MGH707 cancer cells subcutaneously in nude mice. Seven days after injection, tail vein blood was serially collected to measure the Cluc level. **H.** Conditioned media from CCD19-Lu prominently rescues 12 EGFR-driven cancer cell lines, including MGH707, through activating both MET and FGFR *in vitro*. Mean value and 95%CI are shown.

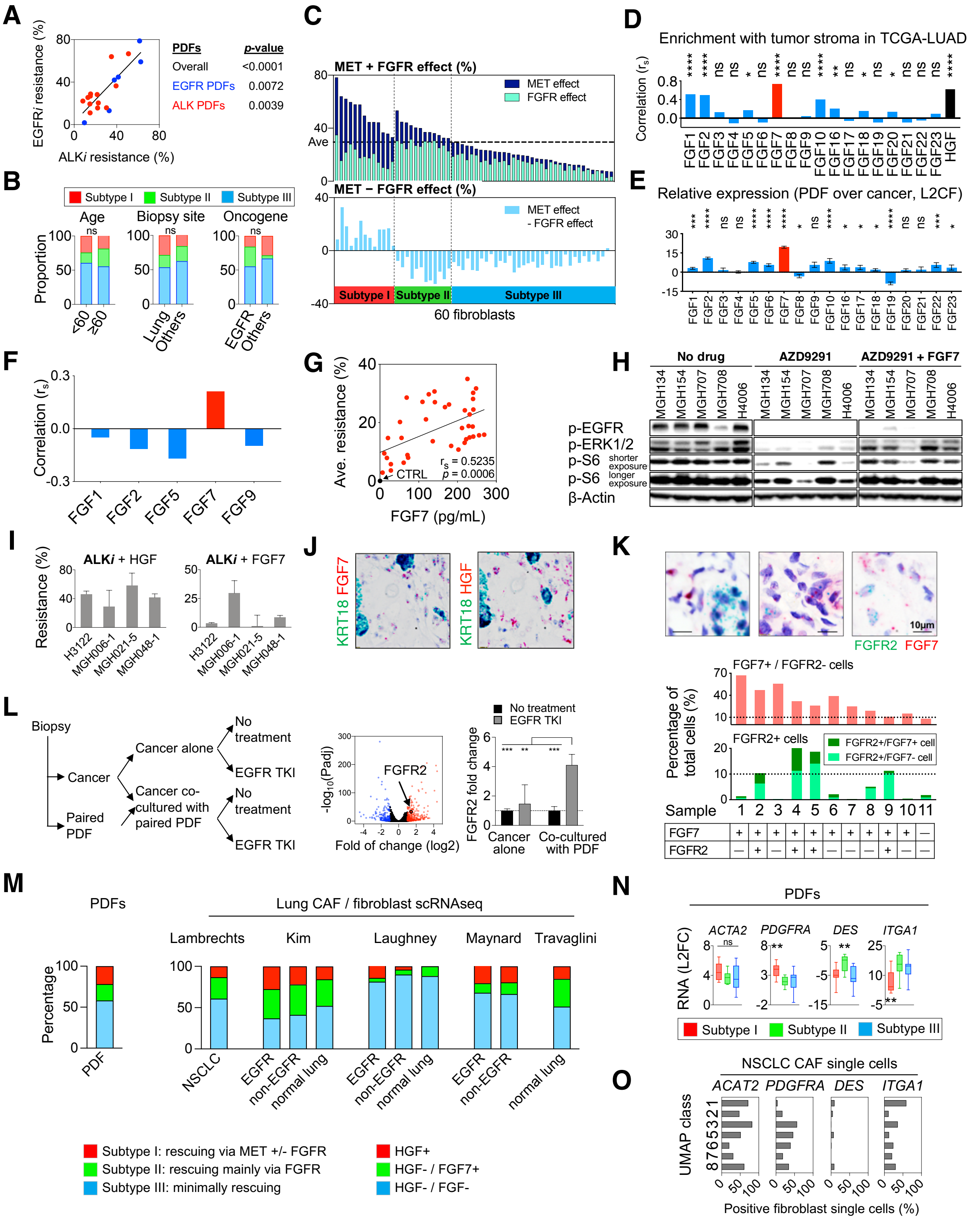


Figure S6: The specific FGF7-FGFR2 signaling axis between CAFs and cancer cells is a prevalent mechanism mediating EGFRi resistance in NSCLC. Related to Figure 5.

A. Each PDF's average resistance effect to ALKi (lorlatinib) assessed in 4 *ALK*+ cancer models is plotted against their resistance effect to EGFRi (osimertinib) assessed in 12 *EGFR*+ cancer models. Each dot represents a PDF model, with 13 PDFs from *ALK*+ NSCLC and 6 PDFs from *EGFR*+ NSCLC included. **B.** Breakdown of the PDF functional subtypes based on patients' age, tumors' site-of-biopsy, and oncogene status. Two-tailed Fisher's exact test was used to compare the proportions of subtype I/II vs. subtype III PDFs depending on patient clinical features. **C.** PDFs are characterized based on their rescue via MET effect (rescue against EGFRi plus FGFRi) and FGFR effect (rescue against EGFRi plus METi) (top graph, MET effect plus FGFR effect are shown as stacked bars to estimate their rescue potential with both pathways) and their preference in using either of these pathways (bottom graph, MET effect minus FGFR effect, positive value indicates MET effect preference, negative value indicates FGFR effect preference). 60 PDFs (38 *EGFR+* and 22 non-*EGFR+* PDFs) were categorized into three distinctive functional subtypes (red, subtype I PDFs with significant MET-mediated rescue; green, subtype II PDFs with modest rescue through FGFR; blue, subtype III PDFs with minimum rescue effect). **D.** Correlation between expression levels of FGFs and HGF in lung adenocarcinoma with the stroma score in each tumor (relative quantification of stroma component abundance, previously published by Yoshihara et al., 2013) in the TCGA-LUAD dataset. Two-tailed spearman's r is used. **E.** Relative expression of FGF family members in PDFs (n=26) versus cancer cells (n=10) determined by qRT-PCR. The PDFs' FGF7 level over cancer cells are shown as -*ΔΔ*CT. Two-tailed t-test is used. **F.** Correlation between FGF ligands secretion in PDFs (n=12) and PDFs' FGFR-dependent rescue against EGFRi plus METi. **G.** The correlation between FGF7 concentration assessed by ELISA and PDFs' FGFR-dependent rescue against EGFRi plus METi. **H.** Western blotting analysis of EGFR, ERK and S6 phosphorylation status in the indicated conditions. FGF7 was used at 10ng/mL. **I.** The rescue effect of HGF and FGF7 on 4 ALK cancer models upon ALK inhibition (lorlatinib at 300nM). Mean value is shown with error bar showing 95% CI. **J.** Representative images of FGF7 and HGF detected by RNAscope in the stroma of xenograft tumor from co-injecting MGH707 cancer cells with CCD19-Lu human fibroblasts. **K.** Representative images of RNAscope staining of FGF7 and FGFR2 in *EGFR*+ lung cancer biopsies (n=11) A sample with more than 10% of cells showing FGF7 expression (without detectable FGFR2) was considered FGF7 positive and a sample with more than 10% of cells with detectable FGFR2 expression (regardless of FGF7 staining) was considered FGFR2 positive. **L.** In a cancer-PDF pair (MGH805-1 and PDF805-1) derived from the same patient biopsy (schematics on the left), only treating cancer cells with EGFRi in the presence of PDF leads to increased FGFR2 expression in cancer cells (RNAseq volcano plot and RNA expression on the right). **M.** The proportion of functionally classified PDFs (left, according to Figure 5A and Figure S6C) and the proportion of tumor and normal lung fibroblast single cells with high HGF, low HGF/high FGF7, and low HGF/low FGF7 levels (Right) in five previous tumor and normal lung scRNA-seq datasets (Lambrechts et al., 2018; Kim et al., 2020; Laughney et al., 2020; Maynard et al., 2020; and Travaglini et al., 2020). **N.** The mRNA level of four fibroblast markers *ACTA2* (αSMA), *PDGFRA*, *DES* and *ITGA1* across three subtypes of PDFs. Gene RNA expression is assessed by qRT-PCR (log2, normalized by house-keeping gene) (whiskers are maximum and minimum values). Two-tailed t test is used to compare the indicated single group to all other PDFs. **O.** The proportion of fibroblast single cells in the NSCLC scRNA-seq dataset from (Lambrechts et al., 2018) positive for key CAF markers per UMAP class. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

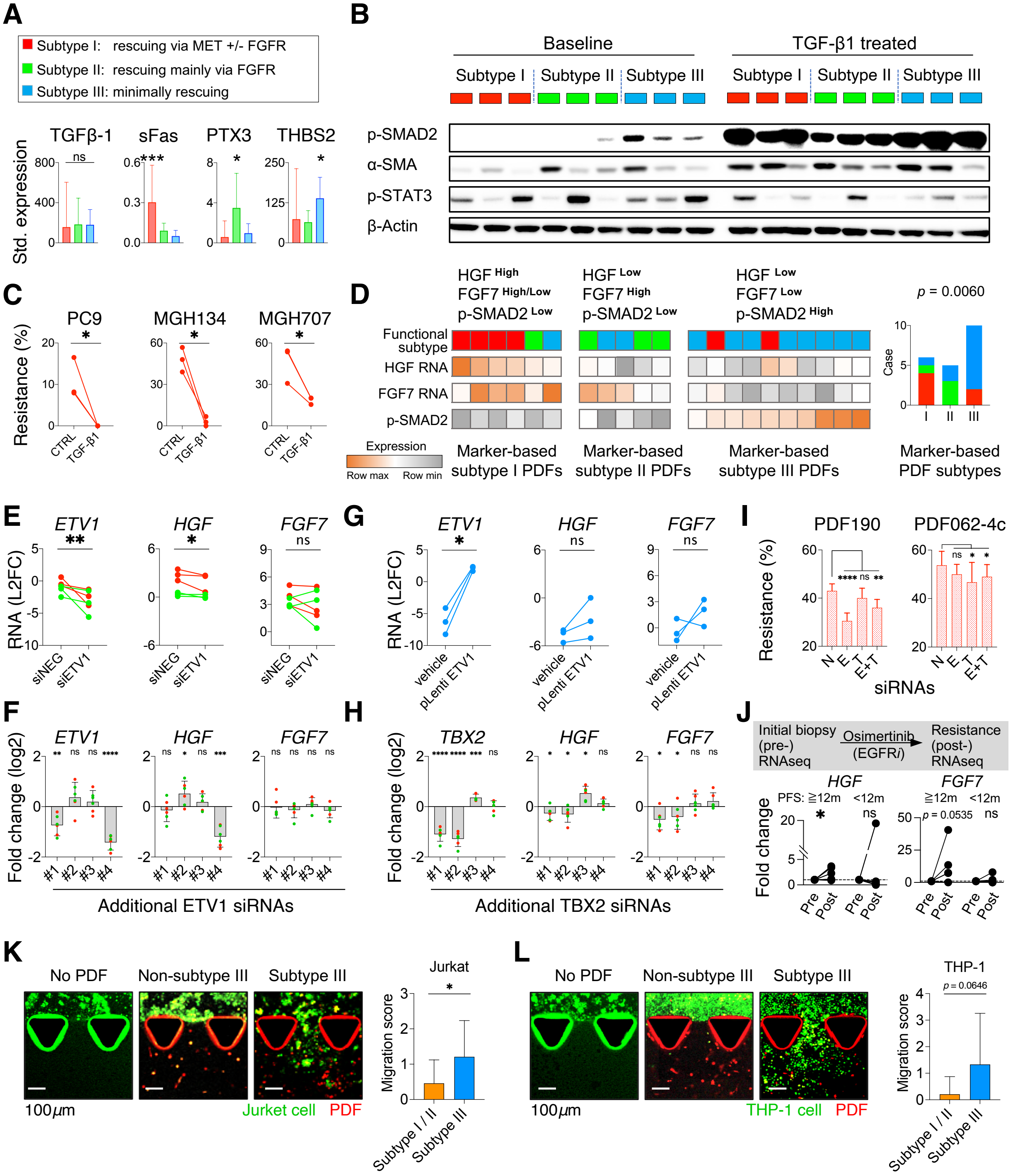


Figure S7: Mechanistic insight of the three subtypes of PDFs. Related to Figure 6, Figure 7 and Figure 8.

A. The protein level of TGF-β1 and three cytokines among three subtypes of PDFs (n=12) assessed by using a 448-multiplex ELISA array. Means with 95%CI are shown. Two-tailed Mann-Whitney U test is used to compare the indicated single group compared to all other PDFs. **B.** Western blotting showing phospho-SMAD2, αSMA, and phospho-STAT3 in three subtypes of PDFs at the baseline and after exposure to exogeneous TGF-β1 (10ng/mL for 24 hours). Lysates from this panel were also probed in Figure 6E. **C.**

Three subtype I PDFs with and without ectopically activating their intrinsic TGF-β signaling are assessed for their resistance effect on three *EGFR*+ NSCLC cancer models. PDFs are pre-incubated with TGF-β1 (20ng/mL) for 72 hours, PDF media was then changed and was banked for five days for cancer cell resistance assessment. **D.** 21 PDFs are classified based on proposed markers (HGF, FGF7, and phospho-SMAD2, see also Figure 6I) and are compared to their functional classification (see also Figure 5A). HGF and FGF7 RNA are assessed by qRT-PCR, and phospho-SMAD2 is assessed by western blotting and quantified by using ImageJ. Fisher's exact test is used. **E-F.** Gene expression change upon knocking down ETV1 in subtypes I (red) and II (green) PDFs. Both an siRNA pool (**E**) and a set of individual siRNAs (**F**) against ETV1 were used for validation. **G.** Gene expression change upon ectopically expressing ETV1 in subtype III (blue) PDFs. ETV1, HGF, and FGF7 RNA expression is assessed by qRT-PCR. **H.** Gene expression change upon knocking down TBX2 in subtypes I (red) and II (green) PDFs. A set of individual siRNAs against TBX2 were used for validation, in addition to an siRNA pool used in Figure 6M. (C, E-H), Mean and 95%CI are shown. Paired one-tail t-test is used. **I.** Subtype I PDFs treated with scramble siRNA (N), and siRNAs against ETV1 (E) and TBX2 (T) are assessed for their resistance effect on MGH707 (*EGFR*+ NSCLC cancer models). PDFs are pre-treated with an siRNA pool for 24 hours, then PDF media was changed and banked for five days for cancer cell resistance assessment. One-tail t-test is used with bars showing mean and 95%CI. **J.** RNAseq data of matched before (pre-) and after (post-) osimertinib treatment biopsies from 11 *EGFR*+ NSCLC patients (from Roper et al., 2020). The dynamic change of HGF and FGF7 RNA are shown according to patients' progression-free survival (PFS). One-tailed t-test was used. **K-L.** Example images and summary of non-subtype III PDFs (subtypes I and II, n=4, example of a subtype I PDF is shown) and subtype III PDFs (n=4) in chemoattracting Jurkat cells (**K**) or THP-1 cells (**L**). One of the representative interface areas is shown, migration score is assessed across all interfaces in the chip and averaged by three replicates. Each bar is the average level with error bar showing 95%CI, one-tailed t-test is used. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.