Data Supplement

DNA methylome and transcriptome landscapes of cancer-associated fibroblasts reveal a smoking-associated malignancy index

Sheng-Fang Su, Hao Ho, Jia-Hua Li, Ming-Fang Wu, Hsu-Chieh Wang, Hsiang-Yuan Yeh, Shuenn-Wen Kuo, Huei-Wen Chen, Chao-Chi Ho* and Ker-Chau Li*

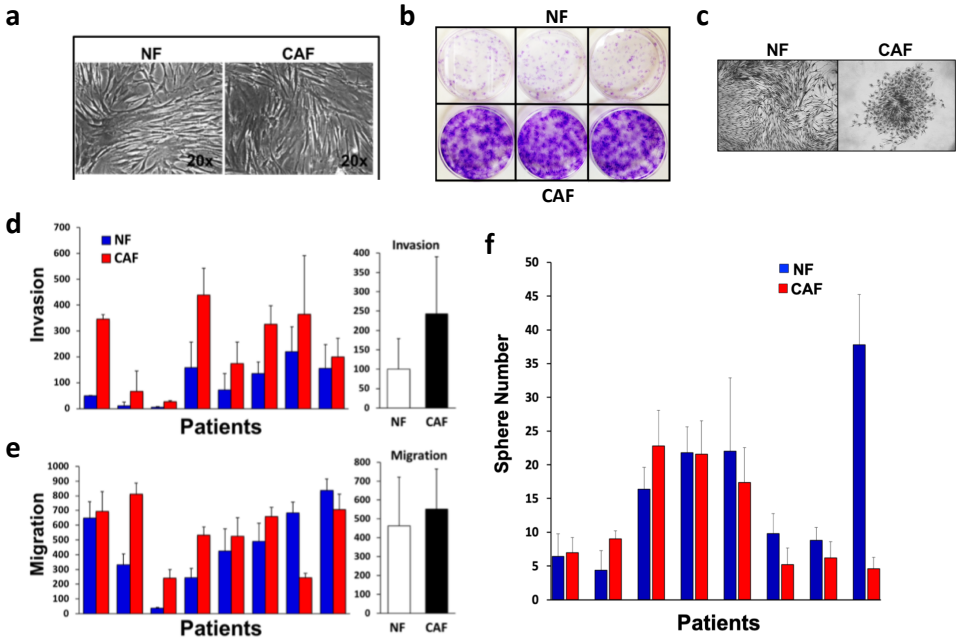
Correspondence to: kcli@stat.sinica.edu.tw; kcli@stat.ucla.edu

Contents:

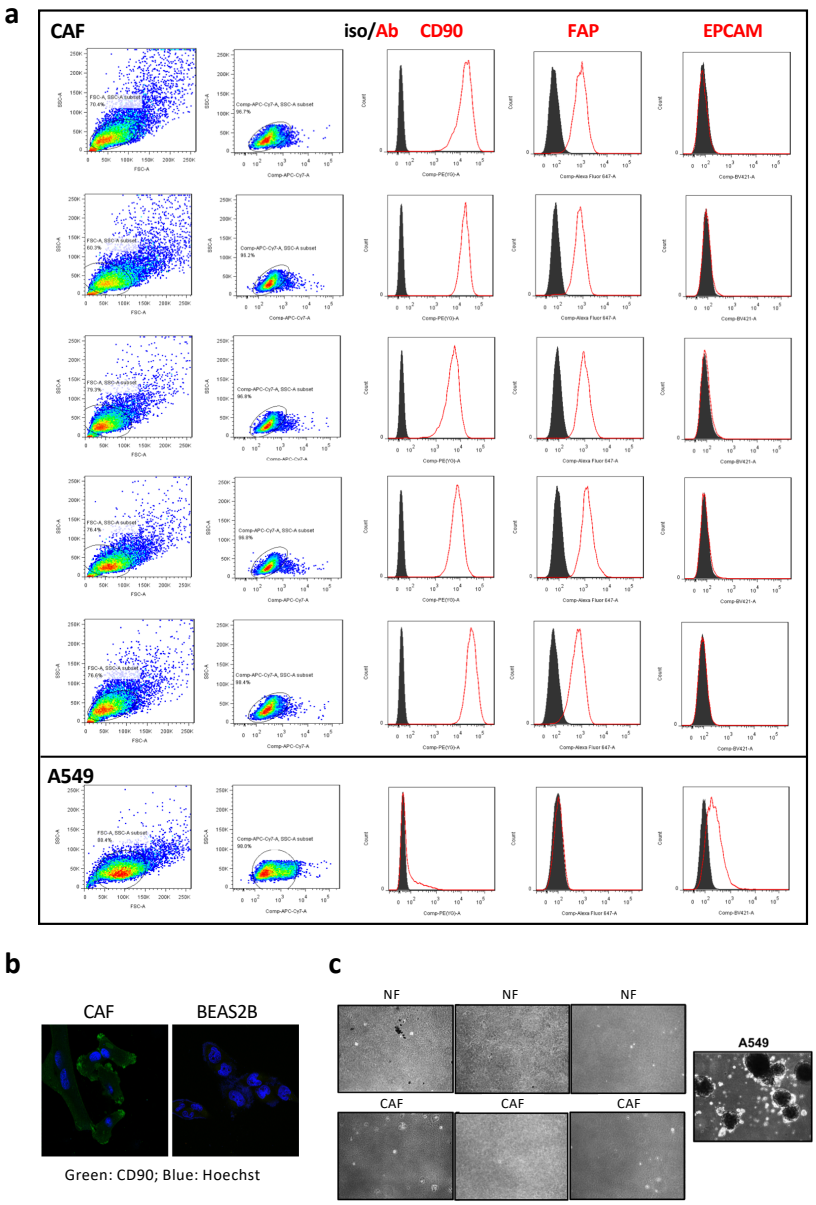
Additional supplements for this manuscript include the following:

Supplemental Data 1 to 5 as a separate Excel file

Supplemental Methods as a separate Word file



Supplemental Figure 1. Phenotype characterization of paired NFs and CAFs from NSCLC patients. (**a**) Cell morphology of the fibroblasts (NF, CAF) (200x). (**b**, **c**) Colony formation assay. 100 cells were seeded in the 35mm dishes and colonies were fixed with methanol and stained with 0.05% crystal violent on Day14. Before staining colonies were shown under the microscope (right). (**d**, **e**, **f**) Lung cancer cells were treated with conditioned medium (CM) from paired NF/CAF for 4 days, and then processed to conduct invasion (**d**), migration (**e**) and sphere forming assays (**f**). Numbers of cells migrated and invaded through the membrane were counted respectively. CM of NF/CAF from individual NSCLC patients was shown on the left and the average value was shown on the right panel (**d**, **e**). The number of spheres was counted coordinately (**f**). Data represent mean ± SD from three replicate experiments.



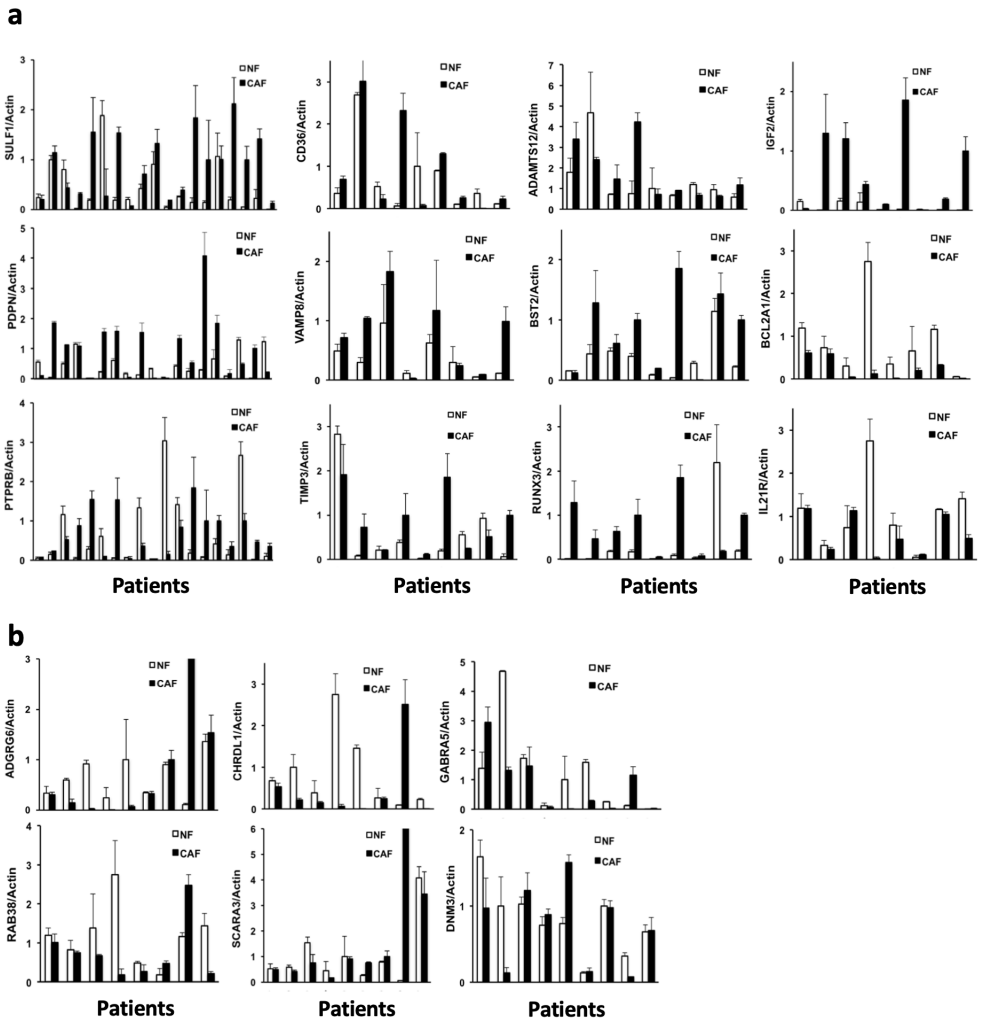
d

Chr7 EGFR L858R c.2573 T > G

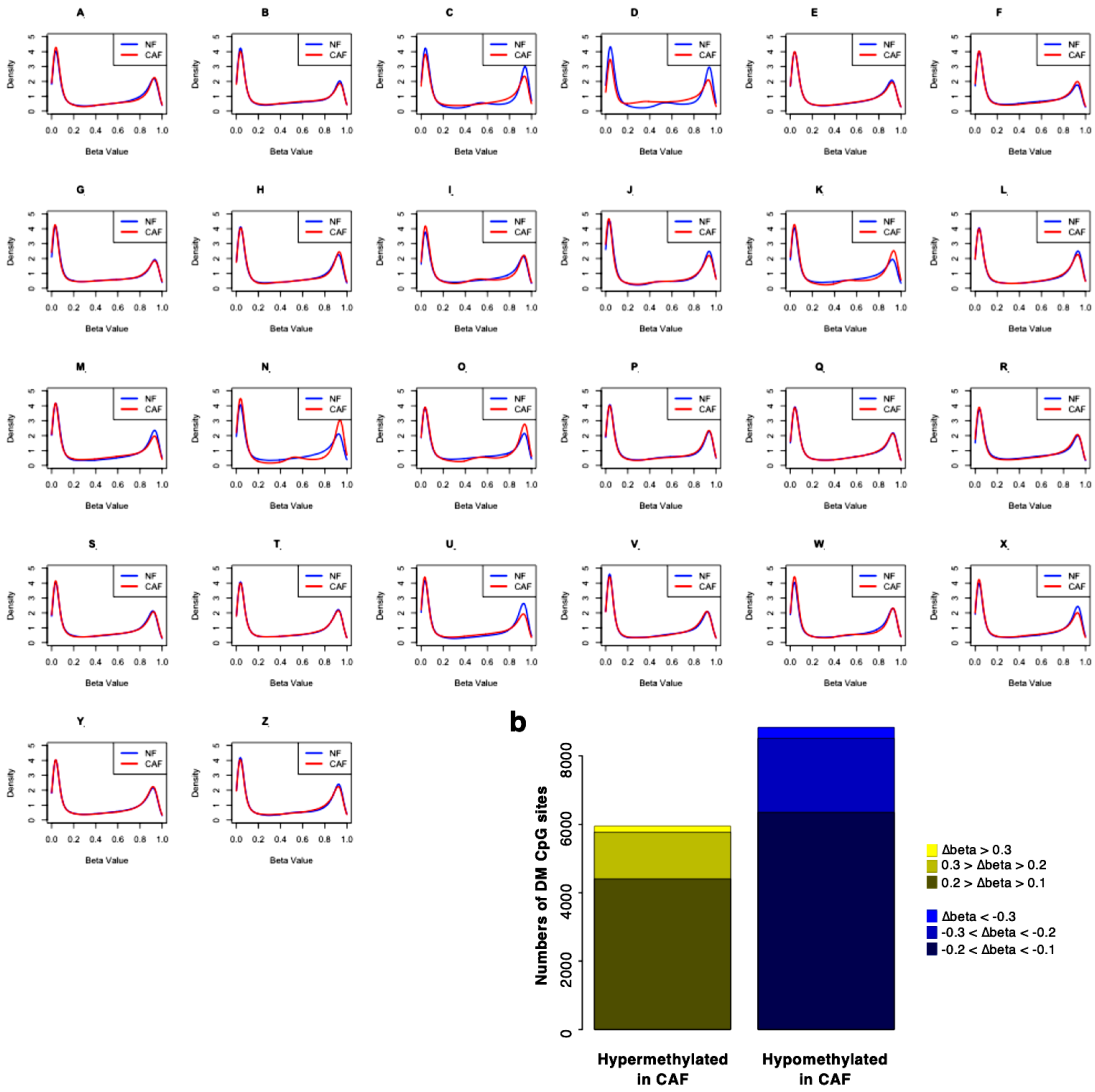
Sample	A	C	G	T	Total Reads
Tumor	1	1	77	542	621
Normal	0	0	1	673	674
CAF	0	1	0	990	991
NF	0	2	0	1008	1010

Supplemental Figure 2. No indication of cancer cell contamination in the primary cultured CAF cell lines. (a) Flow cytometry staining of CD90 and FAP (CAF markers), and EPCAM (epithelial marker) for the lung CAF cell lines and A549 (an epithelial lung cancer cell line). (b)

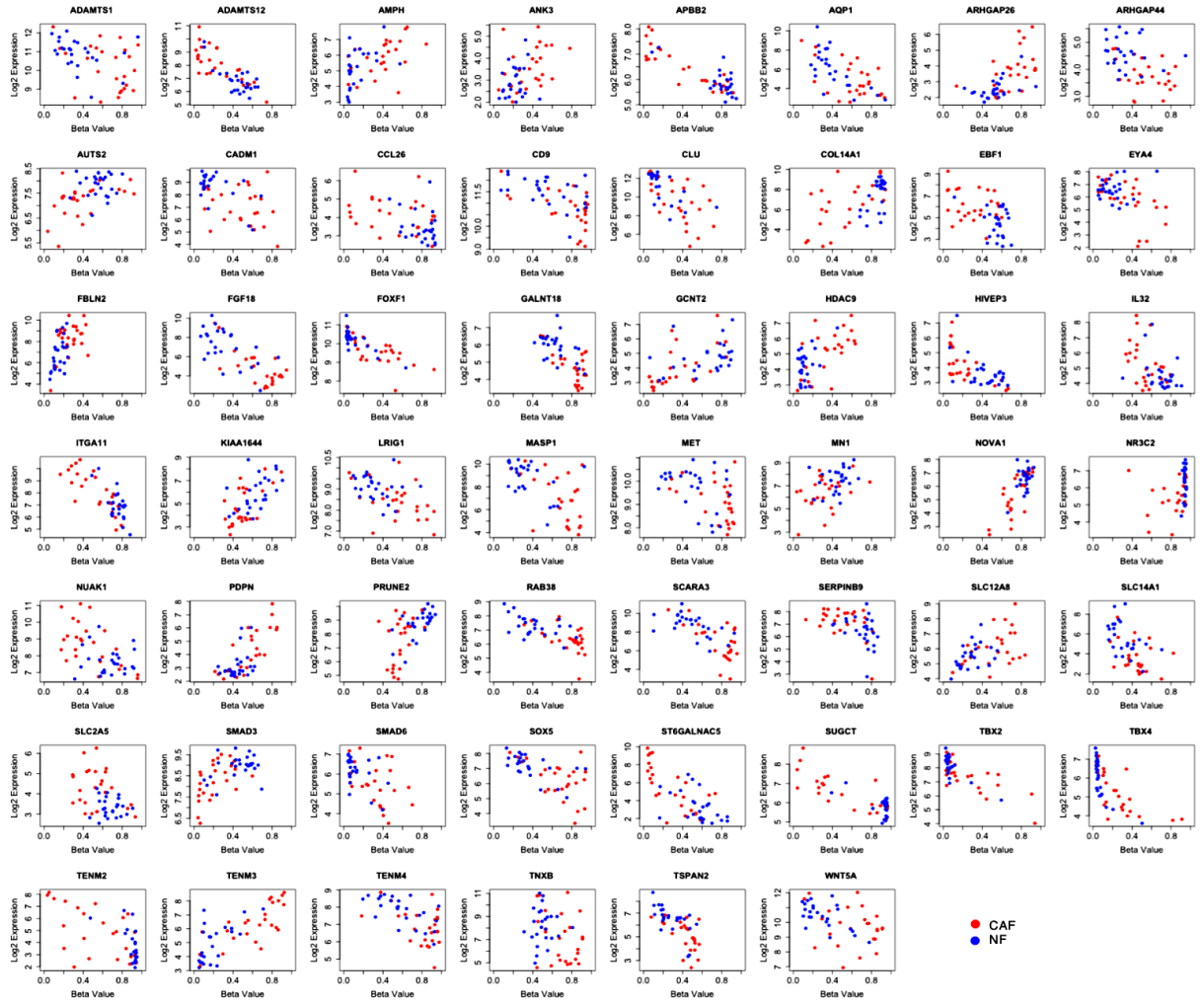
Immunofluorescence staining of CD90 (Green) for CAF and BEAS2B (a normal epithelial cell line). (c) Soft-agar assay on the NF and CAF cell lines and A549 to characterize tumorigenic feature. (d) TruSight® Tumor 15 (TST15) targeted NGS to detect EGFR mutation in the DNA samples from a NSCLC patient with EGFR L858R mutation.



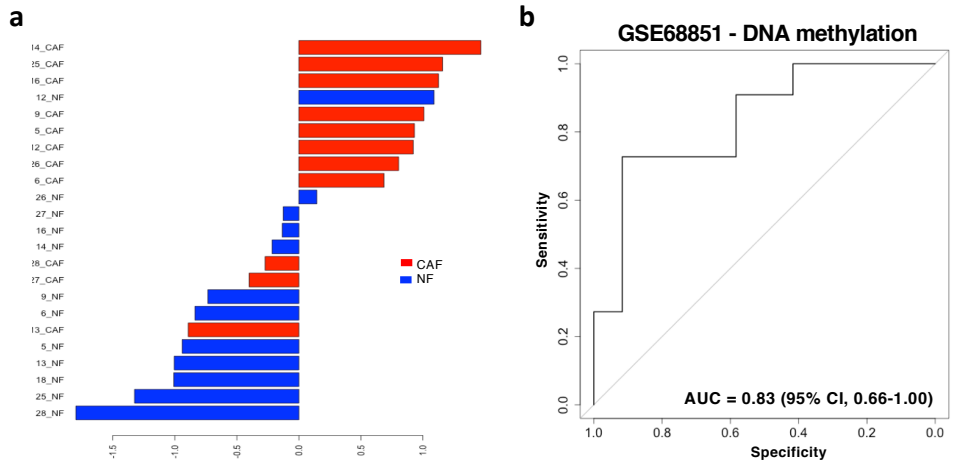
Supplemental Figure 3. Validation of differentially expressed genes between NF and CAF samples from NSCLC patients using real-time qPCR. (**a**, **b**) RNA samples of primary cultured NF and CAF pairs were reverse transcribed to cDNA by iScript RT transferase. mRNA levels of selected CAF up- (**a**) and down (**b**)-regulated genes were measured in paired NF/CAF samples. Actin was used as the internal control. Data represent mean ± SD from three replicate experiments.



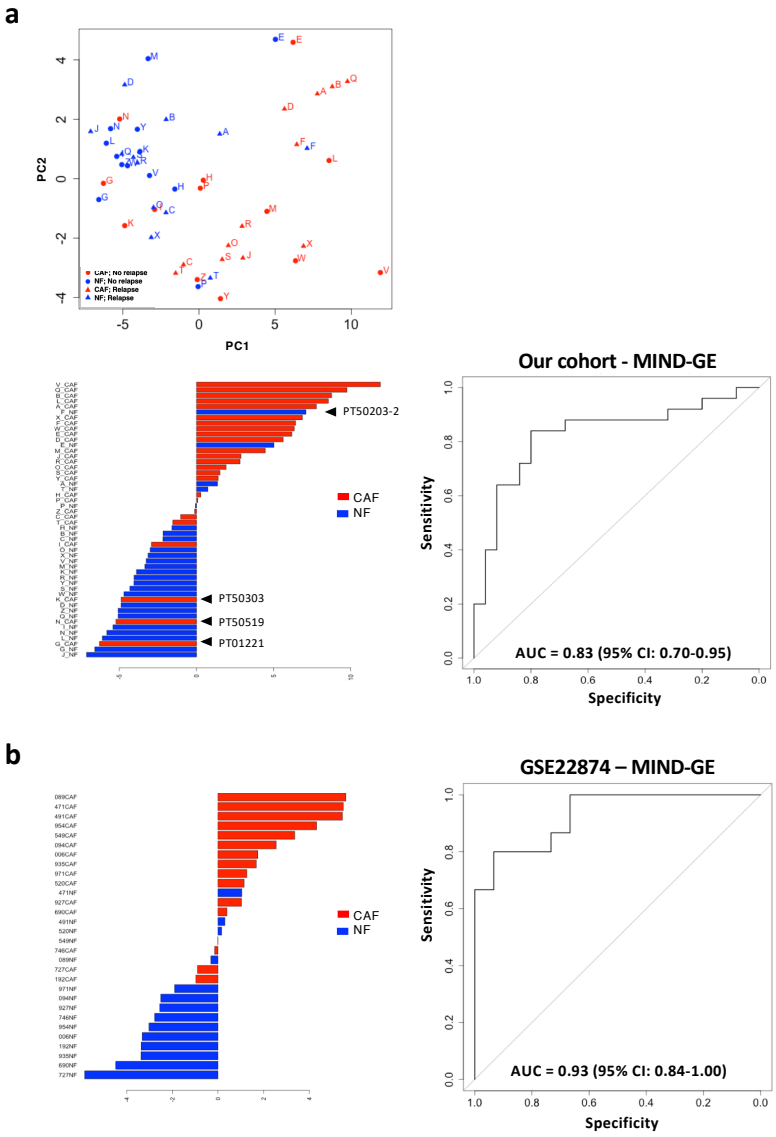
Supplemental Figure 4. The global methylation pattern in paired NF and CAF samples from 26 NSCLC patients. (**a**) The density estimation of the DNA methylation level in the beta-values of 26 NSCLC patients individually. Red: CAF. Blue: NF. (**b**) The number of hyper- and hypo-methylated DM CpGs in CAF was analyzed based on the changes in beta-values.



Supplemental Figure 5. Scatter plots of DNA methylation versus gene expression in NF/CAF pairs for 54 genes of MIND. Red: CAF. Blue: NF.

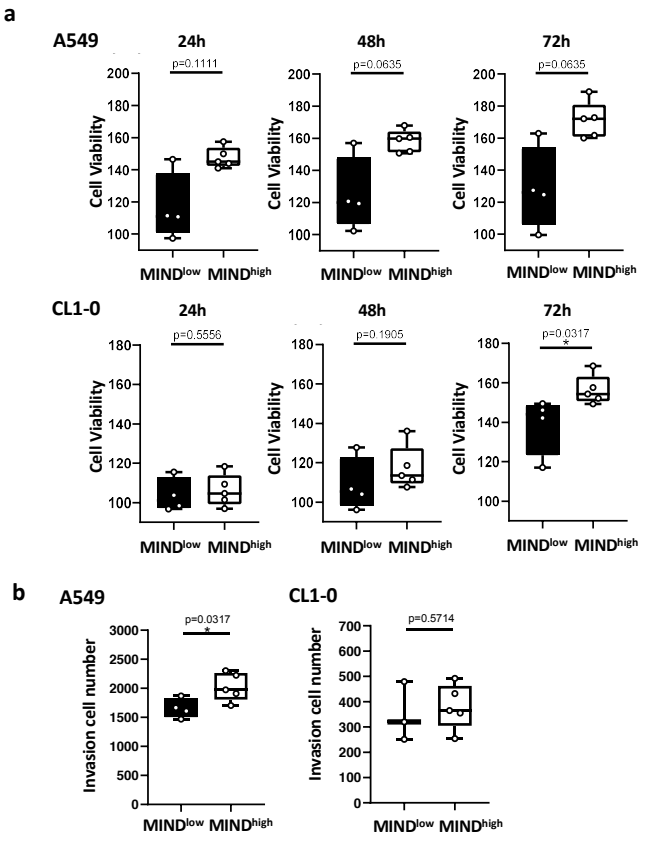


Supplemental Figure 6. Examination of MIND on the discrimination ability in an independent cohort. (**a**) MIND was applied in an independent cohort of NF/CAF samples from 12 NSCLC patients (GSE68851). The distribution of MIND, ordered from the largest value to the smallest, showed a clear classification between CAFs (Red) and NFs (Blue). (Sample 18_CAF lacked methylation value of cg21116457 in gene SLC14A1). (**b**) ROC curve showed the performance of MIND in NF/CAF classification with AUC of 0.83 (95% CI = 0.66-1.0), sensitivity of 73% and specificity of 92% (Youden's Index) in the GSE68851 dataset.

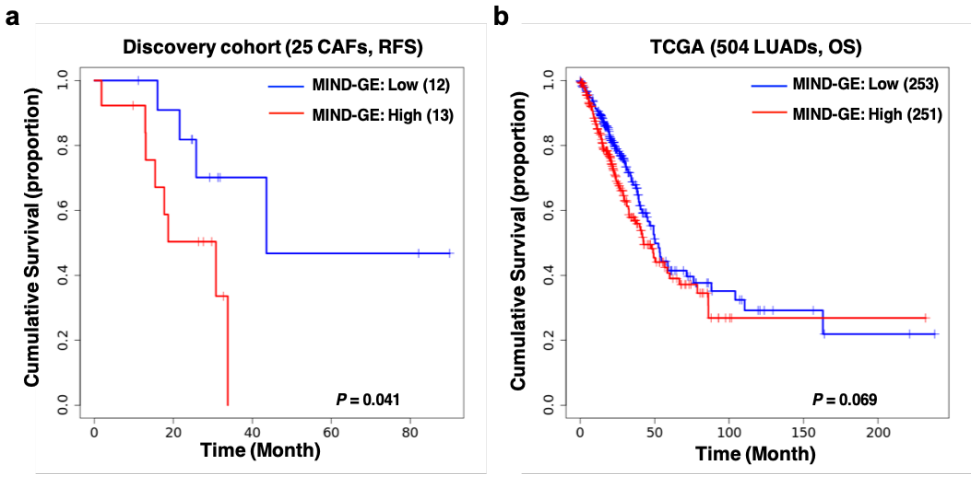


Supplemental Figure 7. Performance of MIND-GE. (**a**) Scatter plot of the first and second PCs showing gene expression profiles of 52 NF/CAF samples on the 54 genes. CAFs are shown in red and NFs in blue. Triangles indicate samples from patients with relapse while those with no relapse shown in circles. The gene expression signature (MIND-GE) was constructed as the weighted sum of the standardized expression weighted by the loadings of PC1. The distribution of MIND-GE, ordered from the largest value to the smallest, showed a clear partition between CAFs (Red) and NFs (Blue). ROC curve was shown with AUC 0.83 (95% CI = 0.70-0.95), sensitivity 84% and

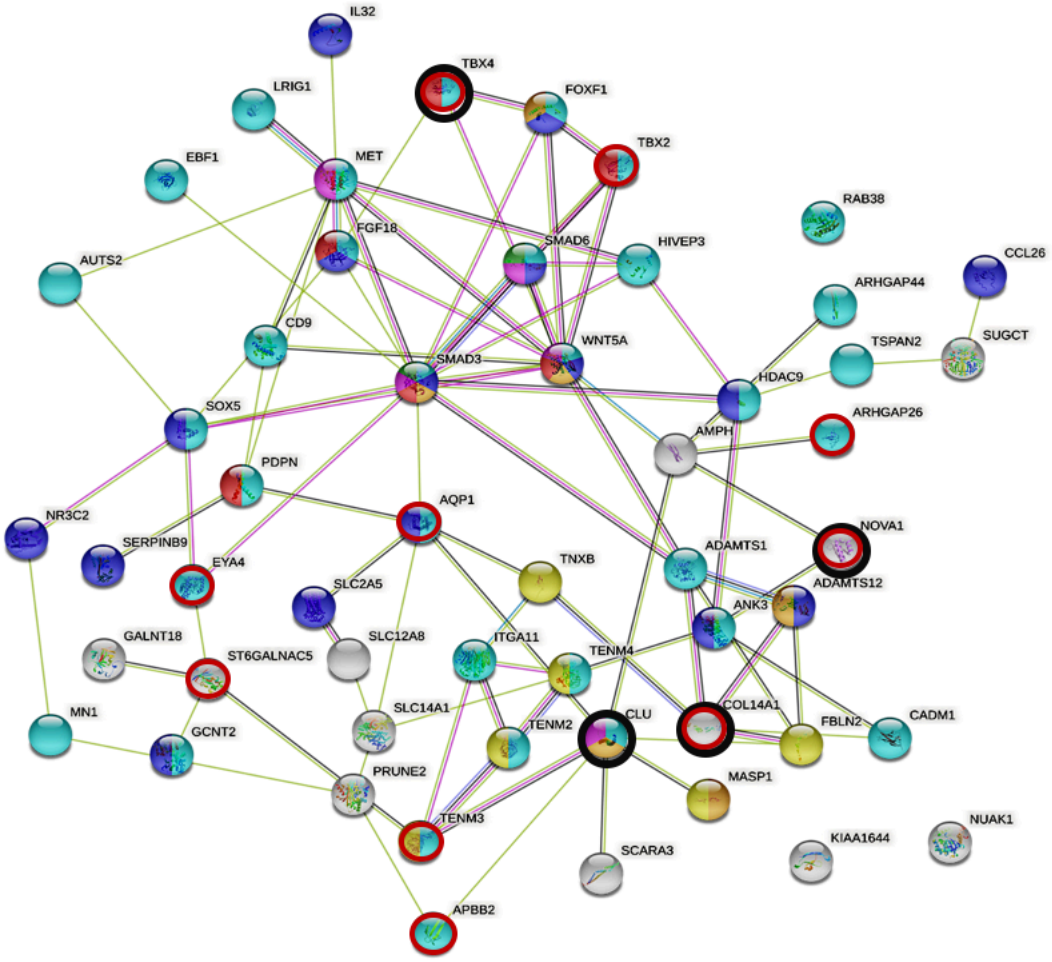
specificity 80% (Youden's Index). (**b**) MIND-GE was validated in an independent cohort of NF/CAF samples from 15 NSCLC patients (GSE22874). ROC curve was plotted with AUC of 0.93 (95% CI = 0.84-1.00), sensitivity of 80% and specificity of 93% (Youden's Index).



Supplemental Figure 8. Impact of the malignancy level of tumor microenvironment, as indexed by MIND, on cancer cell viability and invasion. Conditioned medium from the CAFs with high ($MIND^{high}$, n = 5) or low ($MIND^{low}$, n = 4) score of MIND were applied to the cancer cell lines A549 and CL1-0. (a) MTT assay for cell viability was performed at 24, 48 and 72 hours. (b) Cell invasion assay was performed at 18h. Data represent mean ± SD. Mann-Whitney U test was used for significance. $*P < 0.05$.



Supplemental Figure 9. Prognostic value of MIND-GE. (**a**) Kaplan-Meier analysis for relapse-free survival (RFS) prediction by MIND-GE in the CAF samples from our discovery cohort. Patients were stratified into the high and low groups using the median cutoff ($P = 0.041$, Log rank test). (**b**) Applying MIND-GE to the gene expression profiling of tumor samples from TCGA LUAD dataset. Patients were stratified into the high and low groups using the median cutoff (OS, $P = 0.069$).



Network Stats

number of nodes: 54
number of edges: 88
average node degree: 3.26
avg. local clustering coefficient: 0.383

expected number of edges: 41
PPI enrichment p-value: 1.47e-10
your network has significantly more interactions than expected (what does that mean?)

GO-term	description	count in network	strength	false discovery rate
GO:0032502	developmental process	33 of 5401	0.35	0.00020
GO:0070887	cellular response to chemical stimulus	16 of 2672	0.34	0.0251
GO:0050727	regulation of inflammatory response	6 of 338	0.81	0.0099
PMID:20691848	(2010) Lung organogenesis.	6 of 131	1.22	0.0154
PMID:27367735	(2016) Mechanisms of TGFBeta-Induced Epithelial-Mesenc...	5 of 188	0.98	0.0422
IPR013032	EGF-like, conserved site	6 of 193	1.05	0.00069
IPR036578	SMAD MH1 domain superfamily	2 of 8	1.96	0.0073

Navab R, et al. *PNAS*. (2011)
Vizoso M, et al. *Carcinogenesis*. (2015)

Supplemental Figure 10. The PPI analysis of the 54 genes constituent of MIND (STRING v11.0).

Supplemental Table 1. Clinical characteristics of NSCLC patients whose NF/CAF were primary cultured for genome-wide profiling

Discovery cohort (n=26)						Validation cohort (n=14)					
Age	**Gender**	**Histology**	**Stage**	**Smoking**	**Recurrence**	**Age**	**Gender**	**Histology**	**Stage**	**Smoking**	**Recurrence**
48	M	Squamous cell carcinoma	IIB	Y	Y	83	M	Squamous cell carcinoma	IIIB	Y	Y
86	F	Adenocarcinoma	IB	N	Y	75	M	Squamous cell carcinoma	IIB	Y	N
72	M	Adenocarcinoma	IIA	N	Y	69	M	Squamous cell carcinoma	IIA	Y	Y
68	M	Pleomorphic carcinoma	IIB	Y	Y	47	M	Squamous cell carcinoma	IIIA	Y	Y
59	M	Adenocarcinoma	IIA	N	N	70	M	Adenocarcinoma	IB	N	N
77	F	Adenocarcinoma	IIIA	N	Y	64	F	Adenocarcinoma	IIA	N	Y
71	F	Metastatic lung cancer	NA	N	N	68	M	Adenocarcinoma	IB	Y	N
54	M	Adenocarcinoma	IB	N	N	64	F	Adenocarcinoma	IA	N	N
57	F	Adenocarcinoma	IA	N	N	66	F	Adenocarcinoma	IB	N	N
85	M	Adenocarcinoma	IA	Y	Y	61	F	Adenocarcinoma	IV	N	N
58	F	Lymphoepithelioma-like	IIIA	N	N	67	M	Adenocarcinoma	IIIA	Y	N
76	M	Squamous cell carcinoma	IB	Y	N	59	F	Mucinous adenocarcinoma	IIA	N	N
76	M	Mucinous adenocarcinoma	IIIA	Y	N	81	F	Adenocarcinoma	IA	N	N
65	M	Adenocarcinoma	IB	N	N	60	F	Adenocarcinoma	IIIA	N	Y
77	F	Adenosquamous	IIIA	N	Y						
76	F	Squamous cell carcinoma	IIIA	NA	N						
54	M	Adenocarcinoma	IV	Y	Y						
81	F	Adenocarcinoma	IB	N	Y						
76	F	Adenocarcinoma	IIA	N	Y						
44	F	Adenosquamous	IIB	N	Y						
*63	M	Adenocarcinoma	IIB	Y	N						
67	M	Adenocarcinoma	IIIA	Y	N						
59	F	Adenocarcinoma	IB	N	N						
61	F	Adenocarcinoma	IIIA	N	Y						
61	F	Adenocarcinoma	IIIA	N	N						
64	F	Adenocarcinoma	IA	N	N						

* RNA not pass the QC Affymetrix HU133 PLUS2.0

Supplemental Table 2. Summary of biological mechanisms of 54 genes incorporated in MIND on tumor progression and their potential correlation with smoking.

Probeset	GeneSym	Reported	CTD[#]	Biological Mechanisms	Refs
cg22110896	TBX4	Ref1 Ref2	**P**	Tumor suppressor and epigenetic regulator in lung carcinogenesis; Super-enhancer-driving transcriptional factor in lung fibroblast; Putative candidates for epigenetic therapy in LUAD.	3-5
cg08594681	CLU	Ref1	**S, B, P**	Tumor suppressor in lung cancer and response to chemo- and radiotherapy; Regulation for cell cycle, DNA repair and immune response.	6-9
cg26179069	COL14A1	Ref1 Ref2	**S, B, P**	Fibril-associated collagens involved in the regulation of fibrillogenesis; DNA methylation marker in cancers (e.g., sarcoma).	10-12
cg19534149	NOVA1	Ref1 Ref2	**B, P**	Neuro-oncological ventral antigen 1; RNA binding protein; RNA splicing regulation (e.g., hTERT); Oncogene or tumor suppressor (cancer type-dependent); DNA methylation-driven.	13-16
cg26874542	FGF18		**B**	Involved in cell growth, migration, and epithelial-mesenchymal transition (EMT) via ERK, p38 and MMP26 /Akt/GSK3β/β-catenin.	17-21
cg21929472	SCARA3		**S, B, P**	Tumor suppressor gene; Cell apoptosis regulation via CPSF3, XIAP-caspases; Tumor growth and metastasis suppression; Scavenging receptor recognizing and contributing to oxidative stress, immune and cell death.	22-24
cg12547531	AQP1	Ref2	**S, B, P**	Down-regulated in colorectal carcinoma (CRC) via promoter hypermethylation; Tumor suppressor; Cell proliferation, migration, and senescence regulation.	25-27
cg00979704	TNXB		**B, P**	Involved in epithelial-mesenchymal transition (EMT) via PI3K/Akt pathway; Tumor suppressor; Down-regulated during tumor progression.	28,29
cg00730441	TBX2	Ref2	**B**	Downregulated and hypermethylated in lung CAFs; Acted in *trans* by promoting canonical Wnt (WNT3A); Tumor suppressor; Super-enhancer-driving transcriptional factor in lung fibroblast.	30-34
cg01957732	EYA4	Ref2	**S, B, P**	Tumor suppressor; hypermethylated and under-expressed in multiple independent lung tumor; Tumor metastasis inhibition via suppression of NF-κB.	35-39
cg15799109	SMAD3		**S, N, B, P**	Promoter hypermethylation and silencing, associated with hyperresponsiveness to TGF-β1 in lung CAFs; Tumor suppressor; suppression of NK cell-mediated immune surveillance via Smad3-E4BP4 axis.	2,40-42
cg07760722	NR3C2		**N, B, P**	The mineralocorticoid receptor and tumor suppressor.	43-45
cg09583036	RAB38		**S, B, P**	A member of the Ras small G protein family regulating intracellular vesicular trafficking; Alveolar type II epithelial cells development.	46-48
cg19092167	MASP1		**B, P**	A component of the lectin pathway of complement activation and the central enzyme in the innate immune system.	49-51
cg18010752	WNT5A		**S, B, P**	Tumor suppressor; β-catenin activation by FZD4; the canonical Wnt pathway inhibition via promoting β-	52-56

				catenin degradation by ROR2; A cigarette smoke-responded gene.	
cg14124415	ADAMTS1		**S, B, P**	A dis-integrin and metalloproteinase with thrombospondin motif protein family; Tumor suppressor and antiangiogenic factor; Immune cell repertoire and inflammatory response regulation.	57-61
cg01766743	SOX5		**S, B, P**	Tumor growth, invasion and metastasis regulation; YAP1 interaction; EMT regulation and response to TGF-β via Twist1.	62-65
cg26262197	CADM1		**S, B, P**	Twist1-regulated tumor suppressor; Hippo pathway LATS2 interaction; YAP1 regulation; Tumor progression inhibition via c-Src, PI3K/Akt/mTOR pathway, and STAT.	66-69
cg00551679	FOXF1		**S, B, P**	Endothelial barrier enhancement through S1P/S1PR1 signaling; Lung regeneration stimulation and tumor growth inhibition.	70-73
cg25557280	MET		**B, P**	Tumor growth and senescence promotion in the tumorous microenvironment; TGF-β negative regulation of CXCL1 in CAFs through Smad2/3 binding to the promoter, and suppression of HGF/c-Met autocrine signaling	74-77
cg21621104	LRIG1		**S, B, P**	Tumor suppressor; A feedback negative regulator of receptor tyrosine kinases; Oncogenic signaling restriction from AR, Myc, ERBB oncogenic drivers.	78-80
cg00219210	ARHGAP44		**B**	GAP activity; Catalyzes Rac-GTP hydrolysis; Reduction of actin polymerization required for filopodia formation; A new target of mutant p53, which suppresses ARHGAP44 transcription.	81,82
cg17395738	TENM4		**S, B, P**	Involved in neuron development; Connection of H3K27me3 epigenetic markers for cancer prognosis.	83,84
cg07778290	CD9		**S, N, B, P**	Suppression of cancer cell motility and metastasis depending on cancer types; Adhesion, infection, and inflammatory regulation.	85-88
cg18778196	SMAD6		**P**	A mediator of TGF-β and BMP anti-inflammatory activity; Suppression of IL1R-TLR signaling through inhibition of NF-κB activation; Support of lung cancer cell growth and survival.	89-92
cg20967819	GALNT18		**B**	Response to mucin-type-Golgi stress; Catalyzes the initial reaction in O-linked oligosaccharide biosynthesis.	93,94
cg09811183	TSPAN2		**B, P**	Mutant P53-related; Cell motility and invasiveness enhancement via assisting CD44; An important paralog of CD9.	95,96
cg21116457	SLC14A1		**B**	Transmembrane urea transport facilitation; Tumor suppressor.	97,98
cg00560724	GCNT2		**S, N, B, P**	A I-branching N-acetylglucosaminyltransferase converting linear i-antigen to I-branching glycan; Negative regulation of tumor growth and survival.	99,100
cg14380519	PRUNE2		**S, B, P**	Tumor suppressor in prostate cancer; Cell apoptosis regulation.	101,102
cg21557473	AUTS2		**S, B, P**	Correlation with TGF-β, Hedgehog and Wnt signaling pathway.	103,104

cg00308159	MN1		**B, P**	Oncogene or tumor suppressor (cancer type-dependent); Regulation of the Tbx22 function on mammalian palate development.	105-107
cg00283524	KIAA1644		**B**	A Shisa-like protein.	108
cg05865493	HDAC9		**B**	Tumor progression and immune response regulation (cancer type-dependent).	109,110
cg05185926	ARHGAP26	Ref2	**S, B, P**	A CAF marker; Connection to the Rho kinase signaling pathway.	111
cg24332570	ANK3		**S, N, B, P**	Connection with Wnt and β-catenin signaling; Regulation of lung cancer and smoking-related.	112-114
cg08670693	APBB2	Ref2	**S, B, P**	Amyloid-beta A4 precursor protein-binding family B; Cell cycle regulation.	115,116
cg21604970	ST6GALNAC5	Ref2	**S, B, P**	EMT regulation via Met autocrine; Brain metastasis mediation.	117,118
cg21211213	EBF1		**S, B, P**	Modulation of TERT expression and PNO1/P53 pathway in cancer; Smoking status-associated.	119-121
cg11857452	HIVEP3		**S, B, P**	ERK/Wnt signaling regulation in osteoblasts and BMP-Runx2/VEGF; Pro-inflammatory regulator via binding to NF-κB.	122-124
cg11658060	NUAK1		**S, B, P**	AMPK-related kinase, involved in cell polarity, cell proliferation, cell adhesion, senescence, and tumor progression; Response to DNA damage-response and repair via interacting with TGF-β, SMAD3, P53, CDKN1A.	125-129
cg25602242	SLC2A5		**B**	A fructose transporter; Enhancement of cell proliferation, migration, invasion, and tumorigenesis via enhancing fructose utilization, stimulating fatty acid synthesis and AMPK/mTORC1 signaling; Adjacent lung adenocarcinoma cytoplasmic pro-B cell development.	130-132
cg25311162	TENM3	Ref2	**B**	Regulation of neuronal development and fibroblasts-like cell differentiation; A tumor marker with epigenetic regulation; Correlation with Wnt signaling	133-136
cg18015301	AMPH		**B**	UP-regulated in cancer; Tumor suppressor in breast and lung cancer via inhibiting PI3K/AKT and Ras-Raf-MEK-ERK signal pathways.	137-139
cg24838136	SLC12A8		**B**	A nicotinamide mononucleotide transporter; Tumor growth and metastasis promotion in colorectal, ovarian carcinoma, and bladder cancer via EMT regulation.	140,141
cg03466986	FBLN2		**S, B, P**	An extracellular matrix glycoprotein on tissue development/remodeling; Epithelial basement membrane stability; Correlation with TGF–β signaling; A driver of malignant progression in lung adenocarcinoma.	142-144
cg21158795	PDPN		**B, P**	A CAF malignant marker, promoting tumor growth and metastasis; A new inhibitory molecule on T cells; Enrichment in CD204 M2 mφ (TAM) and negative regulation of CD4+ effector T cells; Regulation of tumor metastasis via TGF-β/STATs, Wnt/β–catenin, and ROCK-LIMK-Cofilin signaling.	145-156
cg08646805	SUGCT		**B**	A mitochondrial enzyme synthesizing glutaryl-CoA from glutarate in tryptophan and lysine catabolism.	157
cg17405646	TENM2		**B, P**	Synapse formation in neuron development.	136,158, 159

cg11303839	CCL26		**S, B, P**	Ligand of CCR3, promoting tumor growth, invasion/migration, angiogenesis; Regulation of TAM and eosinophil recruitment.	160-163
cg23721083	ITGA11		**B, P**	Regulator of tumor niche cytokines (IGF-II, PDGF), promoting tumor progression; TGF-β and FGF2-regulated.	164-169
cg23963071	SERPINB9		**S, B, P**	Regulator of tumor immune escape, blockade of the granzyme B/perforin pathway; Promotion of the resistance to T cell-mediated killing during immune-checkpoint blockade therapy.	170-174
cg07700393	ADAMTS12		**B, P**	Tumor growth promotion via enhancing the transcriptional activity of β-catenin in the Wnt/β-catenin signaling.	175
cg00471190	IL32		**B, P**	A proinflammatory cytokine (cancer type-dependent), promoting tumor growth, invasion and metastasis (integrin β3/p38 MAPK, NF-κB-cytokines/metalloproteinase, and immunosuppressive effects); Tumor cell apoptosis induction; Enhancement of NK and cytotoxic T cell sensitivity; Smoking and COPD-related gene.	176-180

[#]CTD: The Comparative Toxicogenomics Database https://ctdbase.org [181]*updated Dec. 12, 2020*.
S: Tobacco Smoke Pollution-; **N:** nicotine-; **B:** Benzo(a)pyrene, B[a]P-; **P:** PMs-related genes
References in the supplemental materials.

References

1. Navab R, et al. Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer. *Proc Natl Acad Sci U S A*. 2011;108(17):7160-7165.
2. Vizoso M, et al. Aberrant DNA methylation in non-small cell lung cancer-associated fibroblasts. *Carcinogenesis.* 2015;36(12):1453-1463.
3. Xie T, et al. Transcription factor TBX4 regulates myofibroblast accumulation and lung fibrosis. *J Clin Invest.* 2016;126(8):3063-3079.
4. Khalil A, et al. Transcriptomic alterations in lung adenocarcinoma unveil new mechanisms targeted by the *TBX2* subfamily of tumor suppressor genes. *Front Oncol.* 2018;8:482.
5. Horie M, et al. TBX4 is involved in the super-enhancer-driven transcriptional programs underlying features specific to lung fibroblasts. *Am J Physiol Lung Cell Mol Physiol.* 2018;314(1):L177-L191.
6. Chen Z, et al. Inactivation of tumor suppressor gene Clusterin leads to hyperactivation of TAK1-NF-κB signaling axis in lung cancer cells and denotes a therapeutic opportunity. *Theranostics.* 2020;10(25):11520-11534.
7. Panico F, et al. Clusterin (CLU) and lung cancer. *Adv Cancer Res.* 2009;105:63-76.
8. Panico F, et al. Prognostic role of clusterin in resected adenocarcinomas of the lung. *Lung Cancer.* 2013;79(3):294-299.
9. Afanasyeva MA, et al. Clusterin is a potential lymphotoxin beta receptor target that is upregulated and accumulates in germinal centers of mouse spleen during immune response. *PLoS One.* 2014;9(5):e98349.
10. Renner M, et al. Integrative DNA methylation and gene expression analysis in high-grade soft tissue sarcomas. *Genome Biol.* 2013;14(12):r137.

11. Morris MR, et al. Identification of candidate tumour suppressor genes frequently methylated in renal cell carcinoma. *Oncogene.* 2010;29(14):2104-2117.
12. Li X, et al. Identification of a DNA methylome profile of esophageal squamous cell carcinoma and potential plasma epigenetic biomarkers for early diagnosis. *PLoS One.* 2014;9(7):e103162.
13. Ludlow AT, et al. NOVA1 regulates hTERT splicing and cell growth in non-small cell lung cancer. *Nat Commun.* 2018;9(1):3112.
14. Xin Y, et al. Neuro-oncological ventral antigen 1 (NOVA1): Implications in neurological diseases and cancers. *Cell Prolif.* 2017;50(4):e12348.
15. Kim EK, et al. Implications of NOVA1 suppression within the microenvironment of gastric cancer: association with immune cell dysregulation. *Gastric Cancer.* 2017;20(3):438-447.
16. Kim EK, et al. NOVA1 induction by inflammation and NOVA1 suppression by epigenetic regulation in head and neck squamous cell carcinoma. *Sci Rep.* 2019;9(1):11231.
17. Chen T, et al. Fibroblast growth factor 18 promotes proliferation and migration of H460 cells via the ERK and p38 signaling pathways. *Oncol Rep.* 2017;37(2):1235-1242.
18. Song N, et al. FGF18 enhances migration and the epithelial-mesenchymal transition in breast cancer by regulating AKT/GSK3β/β-catenin signaling. *Cell Physiol Biochem.* 2018;49(3):1019-1032.
19. Sonvilla G, et al. FGF18 in colorectal tumour cells: autocrine and paracrine effects. *Carcinogenesis.* 2008;29(1):15-24.
20. Ruiz-Camp J, Morty RE. Divergent fibroblast growth factor signaling pathways in lung fibroblast subsets: where do we go from here? *Am J Physiol Lung Cell Mol Physiol.* 2015;309(8):L751-L755.
21. Katoh M. Multi-layered prevention and treatment of chronic inflammation, organ fibrosis and cancer associated with canonical WNT/β-catenin signaling activation. *Int J Mol Med*. 2018;42(2):713-725.
22. Cheng C, et al. The scavenger receptor SCARA1 (CD204) recognizes dead cells through spectrin. *J Biol Chem.* 2019;294(49):18881-18897.
23. Jiang L, et al. CSR1 suppresses tumor growth and metastasis of human hepatocellular carcinoma via inhibition of HPIP. *Eur Rev Med Pharmacol Sci*. 2017;21(17):3813-3820.
24. Zhu ZH, et al. CSR1 induces cell death through inactivation of CPSF3. *Oncogene.* 2009;28(1):41-51.
25. Wei X, Dong J. Aquaporin 1 promotes the proliferation and migration of lung cancer cell in vitro. *Oncol Rep.* 2015;34(3):1440-1448.
26. Smith E, et al. Reduced aquaporin-1 transcript expression in colorectal carcinoma is associated with promoter hypermethylation. *Epigenetics*. 2019;14(2):158-170.
27. Chen M, et al. AQP1 modulates tendon stem/progenitor cells senescence during tendon aging. *Cell Death Dis*. 2020;11(3):193.
28. Yang N, et al. A functional variant in TNXB promoter associates with the risk of esophageal squamous-cell carcinoma. *Mol Carcinog*. 2020;59(4):439-446.
29. Lévy P, et al. Microarray-based identification of tenascin C and tenascin XB, genes possibly involved in tumorigenesis associated with neurofibromatosis type 1. *Clin Cancer Res*. 2007;13(2 Pt 1):398-407.
30. Khalil AA, et al. TBX2 subfamily suppression in lung cancer pathogenesis: a high-potential marker for early detection. *Oncotarget*. 2017;8(40):68230-68241.
31. Crawford NT, et al. TBX2 interacts with heterochromatin protein 1 to recruit a novel

repression complex to EGR1-targeted promoters to drive the proliferation of breast cancer cells. *Oncogene.* 2019;38(31):5971-5986.
32. Nehme E, et al. Epigenetic suppression of the T-box subfamily 2 (TBX2) in human non-small cell lung cancer. *Int J Mol Sci*. 2019;20(5):1159.
33. Horie M, et al. TBX4 is involved in the super-enhancer-driven transcriptional programs underlying features specific to lung fibroblasts. *Am J Physiol Lung Cell Mol Physiol.* 2018;314(1):L177-L191.
34. Nandana S, et al. Bone metastasis of prostate cancer can be therapeutically targeted at the TBX2-WNT signaling axis. *Cancer Res.* 2017;77(6):1331-1344.
35. Wilson IM, et al. EYA4 is inactivated biallelically at a high frequency in sporadic lung cancer and is associated with familial lung cancer risk. *Oncogene.* 2014;33(36):4464-4473.
36. Mo SJ, et al. EYA4 functions as tumor suppressor gene and prognostic marker in pancreatic ductal adenocarcinoma through β-catenin/ID2 pathway. *Cancer Lett.* 2016;380(2):403-412.
37. Kim SJ, et al. EYA4 acts as a new tumor suppressor gene in colorectal cancer. *Mol Carcinog*. 2015;54(12):1748-1757.
38. Mo SJ, et al. EYA4 inhibits hepatocellular carcinoma growth and invasion by suppressing NF-κB-dependent RAP1 transactivation. *Cancer Commun (Lond).* 2018;38(1):9.
39. Selamat SA, et al. DNA methylation changes in atypical adenomatous hyperplasia, adenocarcinoma in situ, and lung adenocarcinoma. *PLoS One.* 2011;6(6):e21443.
40. Tang PM, et al. Smad3 promotes cancer progression by inhibiting E4BP4-mediated NK cell development. *Nat Commun.* 2017;8:14677.
41. Qian Z, et al. Investigating the mechanism by which SMAD3 induces PAX6 transcription to promote the development of non-small cell lung cancer. *Respir Res.* 2018;19(1):262.
42. Ikemori R, et al. Epigenetic SMAD3 repression in tumor-associated fibroblasts impairs fibrosis and response to the antifibrotic drug nintedanib in lung squamous cell carcinoma. *Cancer Res*. 2020;80(2):276-290.
43. Yang S, et al. A novel MIF signaling pathway drives the malignant character of pancreatic cancer by targeting NR3C2. *Cancer Res.* 2016;76(13):3838-3850.
44. Yang C, et al. MicroRNA-766 promotes cancer progression by targeting NR3C2 in hepatocellular carcinoma. *FASEB J.* 2019;33(1):1456-1467.
45. Zhang DL, et al. Genome-wide identification of transcription factors that are critical to non-small cell lung cancer. *Cancer Lett*. 2018;434:132-143.
46. Hsieh JJ, et al. RAB38 is a potential prognostic factor for tumor recurrence in non-small cell lung cancer. *Oncol Lett.* 2019;18(3):2598-2604.
47. Chang JW, et al. Comparison of genomic signatures of non-small cell lung cancer recurrence between two microarray platforms. *Anticancer Res.* 2012;32(4):1259-1265.
48. Zhang L, et al. Rab38 targets to lamellar bodies and normalizes their sizes in lung alveolar type II epithelial cells. *Am J Physiol Lung Cell Mol Physiol.* 2011;301(4):L461-L477.
49. Kang JU, et al. Identification of novel candidate target genes, including EPHB3, MASP1 and SST at 3q26.2-q29 in squamous cell carcinoma of the lung. *BMC Cancer.* 2009;9:237.
50. Gulla KC, et al. Activation of mannan-binding lectin-associated serine proteases leads to generation of a fibrin clot. *Immunology.* 2010;129(4):482-495.
51. Takahashi M, et al. Essential role of mannose-binding lectin-associated serine protease-1 in activation of the complement factor D. *J Exp Med.* 2010;207(1):29-37.

52. Fu HD, et al. Wnt5a mediated canonical Wnt signaling pathway activation in orthodontic tooth movement: possible role in the tension force-induced bone formation. *J Mol Histol.* 2016;47(5):455-466.
53. Ren D, et al. Wnt5a induces and maintains prostate cancer cells dormancy in bone. *J Exp Med.* 2019;216(2):428-449.
54. Hu B, et al. Epigenetic activation of Wnt5a drives glioblastoma stem cell differentiation and invasive growth. *Cell.* 2016;167(5):1281-1295.e18.
55. Serra R, et al. Wnt5a as an effector of TGFβ in mammary development and cancer. *J Mammary Gland Biol Neoplasia.* 2011;16(2):157-167.
56. Whang YM, et al. Wnt5a is associated with cigarette smoke-related lung carcinogenesis via protein kinase C. *PLoS One.* 2013;8(1):e53012.
57. Martino-Echarri E, et al. Contribution of ADAMTS1 as a tumor suppressor gene in human breast carcinoma. Linking its tumor inhibitory properties to its proteolytic activity on nidogen-1 and nidogen-2. *Int J Cancer.* 2013;133(10):2315-2324.
58. Obika M, et al. Tumor growth inhibitory effect of ADAMTS1 is accompanied by the inhibition of tumor angiogenesis. *Cancer Sci*. 2012;103(10):1889-1897.
59. Rodríguez-Baena FJ, et al. ADAMTS1 protease is required for a balanced immune cell repertoire and tumour inflammatory response. *Sci Rep.* 2018;8(1):13103.
60. Ricciardelli C, et al. The ADAMTS1 protease gene is required for mammary tumor growth and metastasis. *Am J Pathol*. 2011;179(6):3075-3085.
61. Tyan SW, et al. Breast cancer cells induce stromal fibroblasts to secrete ADAMTS1 for cancer invasion through an epigenetic change. *PLoS One*. 2012;7(4):e35128.
62. Kurtsdotter I, et al. SOX5/6/21 prevent oncogene-driven transformation of brain stem cells. *Cancer Res*. 2017;77(18):4985-4997.
63. Sun C, et al. SOX5 promotes breast cancer proliferation and invasion by transactivation of EZH2. *Oncol Lett.* 2019;17(3):2754-2762.
64. Hu J, et al. Sox5 contributes to prostate cancer metastasis and is a master regulator of TGF-β-induced epithelial mesenchymal transition through controlling Twist1 expression. *Br J Cancer*. 2018;118(1):88-97.
65. Zou H, et al. SOX5 interacts with YAP1 to drive malignant potential of non-small cell lung cancer cells. *Am J Cancer Res.* 2018;8(5):866-878.
66. Hartsough EJ, et al. CADM1 is a TWIST1-regulated suppressor of invasion and survival. *Cell Death Dis.* 2019;10(4):281.
67. Vallath S, et al. CADM1 inhibits squamous cell carcinoma progression by reducing STAT3 activity. *Sci Rep.* 2016;6:24006.
68. Ito T, et al. CADM1 associates with Hippo pathway core kinases; membranous co-expression of CADM1 and LATS2 in lung tumors predicts good prognosis. *Cancer Sci.* 2019;110(7):2284-2295.
69. van den Berg RM, et al. Comprehensive CADM1 promoter methylation analysis in NSCLC and normal lung specimens. *Lung Cancer.* 2011;72(3):316-321.
70. Cai Y, et al. FOXF1 maintains endothelial barrier function and prevents edema after lung injury. *Sci Signal.* 2016;9(424):ra40.
71. Bolte C, et al. FOXF1 transcription factor promotes lung regeneration after partial pneumonectomy. *Sci Rep.* 2017;7(1):10690.
72. Wu CY, et al. Highly expressed FOXF1 inhibit non-small-cell lung cancer growth via inducing tumor suppressor and g1-phase cell-cycle arrest. *Int J Mol Sci.* 2020;21(9):3227.

73. Melboucy-Belkhir S, et al. Forkhead Box F1 represses cell growth and inhibits COL1 and ARPC2 expression in lung fibroblasts in vitro. *Am J Physiol Lung Cell Mol Physiol.* 2014;307(11):L838-L847.
74. Fang WB, et al. TGF-β negatively regulates CXCL1 chemokine expression in mammary fibroblasts through enhancement of Smad2/3 and suppression of HGF/c-Met signaling mechanisms. *PLoS One*. 2015;10(8):e0135063.
75. Kanaji N, et al. Hepatocyte growth factor produced in lung fibroblasts enhances non-small cell lung cancer cell survival and tumor progression. *Respir Res.* 2017;18(1):118.
76. Boichuck M, et al. c-Met as a new marker of cellular senescence. *Aging (Albany NY)*. 2019;11(9):2889-2897.
77. Konstorum A, Lowengrub JS. Activation of the HGF/c-Met axis in the tumor microenvironment: a multispecies model. *J Theor Biol.* 2018;439:86-99.
78. Kou C, et al. LRIG1, a 3p tumor suppressor, represses EGFR signaling and is a novel epigenetic silenced gene in colorectal cancer. *Biochem Biophys Res Commun*. 2015;464(2):519-525.
79. Li Q, et al. LRIG1 is a pleiotropic androgen receptor-regulated feedback tumor suppressor in prostate cancer. *Nat Commun.* 2019;10(1):5494.
80. Yang B, et al. Lrig1 is a positive prognostic marker in hepatocellular carcinoma. *Onco Targets Ther*. 2016;9:7071-7079.
81. Galic M, et al. Dynamic recruitment of the curvature-sensitive protein ArhGAP44 to nanoscale membrane deformations limits exploratory filopodia initiation in neurons. *Elife*. 2014;3:e03116.
82. Xu J, et al. Mutant p53 promotes cell spreading and migration via ARHGAP44. *Sci China Life Sci.* 2017;60(9):1019-1029.
83. Tucker RP, Chiquet-Ehrismann R. Teneurins: a conserved family of transmembrane proteins involved in intercellular signaling during development. *Dev Biol.* 2006;290(2):237-245.
84. Ngollo M, et al. Global analysis of H3K27me3 as an epigenetic marker in prostate cancer progression. *BMC Cancer.* 2017;17(1):261.
85. Takeda T, et al. Adenoviral transduction of MRP-1/CD9 and KAI1/CD82 inhibits lymph node metastasis in orthotopic lung cancer model. *Cancer Res.* 2007;67(4):1744-1749.
86. Nakazawa Y, et al. Tetraspanin family member CD9 inhibits Aggrus/podoplanin-induced platelet aggregation and suppresses pulmonary metastasis. *Blood.* 2008;112(5):1730-1739.
87. Miyake M, et al. Motility-related protein-1 (MRP-1/CD9) reduction as a factor of poor prognosis in breast cancer. *Cancer Res.* 1996;56(6):1244-1249.
88. Reyes R, et al. Tetraspanin CD9: a key regulator of cell adhesion in the immune system. *Front Immunol.* 2018;9:863.
89. Kleeff J, et al. Smad6 suppresses TGF-β-induced growth inhibition in COLO-357 pancreatic cancer cells and is overexpressed in pancreatic cancer. *Biochem Biophys Res Commun.* 1999;255(2):268-273.
90. Zhang T, et al. Smad6 methylation represses NFκB activation and periodontal inflammation. *J Dent Res.* 2018;97(7):810-819.
91. Jeon HS, et al. SMAD6 contributes to patient survival in non-small cell lung cancer and its knockdown reestablishes TGF-beta homeostasis in lung cancer cells. *Cancer Res.* 2008;68(23):9686-9692.
92. Hata A, et al. Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumor suppressor. *Genes Dev*. 1998;12(2):186-197.

93. Jamaludin MI, et al. MGSE regulates crosstalk from the mucin pathway to the TFE3 pathway of the golgi stress response. *Cell Struct Funct.* 2019;44(2):137-151.
94. Shan A, et al. Polypeptide N-acetylgalactosaminyltransferase 18 non-catalytically regulates the ER homeostasis and O-glycosylation. *Biochim Biophys Acta Gen Subj.* 2019;1863(5):870-882.
95. Otsubo C, et al. TSPAN2 is involved in cell invasion and motility during lung cancer progression. *Cell Rep.* 2014;7(2):527-538.
96. Lafleur MA, et al. Tetraspanin proteins regulate membrane type-1 matrix metalloproteinase-dependent pericellular proteolysis. *Mol Biol Cell.* 2009;20(7):2030-2040.
97. Chan TC, et al. SLC14A1 prevents oncometabolite accumulation and recruits HDAC1 to transrepress oncometabolite genes in urothelial carcinoma. *Theranostics.* 2020;10(25):11775-11793.
98. Hou R, et al. SLC14A1: a novel target for human urothelial cancer. *Clin Transl Oncol.* 2017;19(12):1438-1446.
99. Nakamura K, et al. Aberrant methylation of GCNT2 is tightly related to lymph node metastasis of primary CRC. *Anticancer Res.* 2015;35(3):1411-1421.
100. Sweeney JG, et al. Loss of GCNT2/I-branched glycans enhances melanoma growth and survival. *Nat Commun.* 2018;9(1):3368.
101. Salameh A, et al. PRUNE2 is a human prostate cancer suppressor regulated by the intronic long noncoding RNA PCA3. *Proc Natl Acad Sci U S A.* 2015;112(27):8403-8408.
102. Tatsumi Y, et al. BMCC1, which is an interacting partner of BCL2, attenuates AKT activity, accompanied by apoptosis. *Cell Death Dis.* 2015;6(1):e1607.
103. Han Y, et al. AUTS2 is a potential therapeutic target for pancreatic cancer patients with liver metastases. *Med Hypotheses.* 2015;85(2):203-206.
104. Nagel S, et al. Deregulation of polycomb repressor complex 1 modifier AUTS2 in T-cell leukemia. *Oncotarget.* 2016;7(29):45398-45413.
105. Larmonie NSD, et al. MN1 overexpression is driven by loss of DNMT3B methylation activity in inv(16) pediatric AML. *Oncogene.* 2018;37(1):107-115.
106. Numata M, et al. High MN1 expression increases the in vitro clonogenic activity of primary mouse B-cells. *Leuk Res.* 2015;39(8):906-912.
107. Shu L, et al. MN1 gene loss-of-function mutation causes cleft palate in a pedigree. *Brain.* 2020;22:awaa431.
108. Pei J, Grishin NV. Unexpected diversity in Shisa-like proteins suggests the importance of their roles as transmembrane adaptors. *Cell Signal.* 2012;24(3):758-769.
109. Ning Y, et al. HDAC9 deficiency promotes tumor progression by decreasing the CD8(+) dendritic cell infiltration of the tumor microenvironment. *J Immunother Cancer.* 2020;8(1):e000529.
110. Ma Z, et al. Histone deacetylase 9 downregulation decreases tumor growth and promotes apoptosis in non-small cell lung cancer after melatonin treatment. *J Pineal Res.* 2019;67(2):e12587.
111. Bozóky B, et al. Novel signatures of cancer-associated fibroblasts. *Int J Cancer.* 2013;133(2):286-293.
112. Shriwash N, et al. Identification of differentially expressed genes in small and non-small cell lung cancer based on meta-analysis of mRNA. *Heliyon*. 2019;5(6):e01707.
113. Liu Y, et al. Identification of feature genes for smoking-related lung adenocarcinoma based on gene expression profile data. *Onco Targets Ther.* 2016;9:7397-7407.

114. Durak O, et al. Ankyrin-G regulates neurogenesis and Wnt signaling by altering the subcellular localization of β-catenin. *Mol Psychiatry.* 2015;20(3):388-397.
115. Duilio A, et al. Fe65L2: a new member of the Fe65 protein family interacting with the intracellular domain of the Alzheimer's beta-amyloid precursor protein. *Biochem J.* 1998;330 (Pt 1)(Pt 1):513-519.
116. Penna I, et al. A novel snRNA-like transcript affects amyloidogenesis and cell cycle progression through perturbation of Fe65L1 (APBB2) alternative splicing. *Biochim Biophys Acta.* 2013;1833(6):1511-1526.
117. Chu C, et al. Stable ectopic expression of ST6GALNAC5 induces autocrine met activation and anchorage-independence in mdck cells. *PLoS One.* 2016;11(2):e0148075.
118. Bos PD, et al. Genes that mediate breast cancer metastasis to the brain. *Nature.* 2009;459(7249):1005-1009.
119. Xing M, et al. Genomic and epigenomic EBF1 alterations modulate TERT expression in gastric cancer. *J Clin Invest.* 2020;130(6):3005-3020.
120. Shen A, et al. EBF1-mediated upregulation of ribosome assembly factor PNO1 contributes to cancer progression by negatively regulating the p53 signaling pathway. *Cancer Res.* 2019;79(9):2257-2270.
121. Shui IM, et al. Prostate tumor DNA methylation is associated with cigarette smoking and adverse prostate cancer outcomes. *Cancer.* 2016;122(14):2168-2177.
122. Shim JH, et al. Schnurri-3 regulates ERK downstream of WNT signaling in osteoblasts. *J Clin Invest.* 2013;123(9):4010-4022.
123. Li Y, et al. Schnurri-3 regulates BMP9-induced osteogenic differentiation and angiogenesis of human amniotic mesenchymal stem cells through Runx2 and VEGF. *Cell Death Dis*. 2020;11(1):72.
124. Allen CE, et al. The κB transcriptional enhancer motif and signal sequences of V(D)J recombination are targets for the zinc finger protein HIVEP3/KRC: a site selection amplification binding study. *BMC Immunol.* 2002;3:10.
125. Kolliopoulos C, et al. Transforming growth factor β (TGFβ) induces NUAK kinase expression to fine-tune its signaling output. *J Biol Chem.* 2019;294(11):4119-4136.
126. Chen P, et al. High NUAK1 expression correlates with poor prognosis and involved in NSCLC cells migration and invasion. *Exp Lung Res.* 2013;39(1):9-17.
127. Hou X, et al. A new role of NUAK1: directly phosphorylating p53 and regulating cell proliferation. *Oncogene.* 2011;30(26):2933-2942.
128. Bernard D, Augert A. NUAK1 links genomic instability and senescence. *Aging* (Albany NY). 2010;2(6):317-319.
129. Sun X, et al. The regulation and function of the NUAK family. *J Mol Endocrinol.* 2013;51(2):R15-R22.
130. Weng Y, et al. SLC2A5 promotes lung adenocarcinoma cell growth and metastasis by enhancing fructose utilization. *Cell Death Discov.* 2018;4:38.
131. You J, et al. Low glucose transporter SLC2A5-inhibited human normal adjacent lung adenocarcinoma cytoplasmic pro-B cell development mechanism network. *Mol Cell Biochem.* 2015;399(1-2):71-76.
132. Chen WL, et al. GLUT5-mediated fructose utilization drives lung cancer growth by stimulating fatty acid synthesis and AMPK/mTORC1 signaling. *JCI Insight.* 2020;5(3):e131596.
133. Tsumura K, et al. Establishment of permutation for cancer risk estimation in the urothelium

based on genome-wide DNA methylation analysis. *Carcinogenesis.* 2019;40(11):1308-1319.
134. Luo SS, et al. Genome-wide analysis to identify a novel microRNA signature that predicts survival in patients with stomach adenocarcinoma. *J Cancer.* 2019;10(25):6298-6313.
135. Bastías-Candia S, et al. Wnt signaling upregulates Teneurin-3 expression via canonical and non-canonical Wnt pathway crosstalk. *Front Neurosci.* 2019;13:505.
136. Ziegler A, et al. Teneurin protein family: an emerging role in human tumorigenesis and drug resistance. *Cancer Lett.* 2012;326(1):1-7.
137. Yang H, et al. AMPH-1 is a tumor suppressor of lung cancer by inhibiting Ras-Raf-MEK-ERK signal pathway. *Lasers Med Sci.* 2019;34(3):473-478.
138. Chen Y, et al. AMPH1 functions as a tumour suppressor in ovarian cancer via the inactivation of PI3K/AKT pathway. *J Cell Mol Med.* 2020;24(13):7652-7659.
139. Jiang L, et al. miR-425 regulates cell proliferation, migration and apoptosis by targeting AMPH-1 in non-small-cell lung cancer. *Pathol Res Pract.* 2019;215(12):152705.
140. Grozio A, et al. Slc12a8 is a nicotinamide mononucleotide transporter. *Nat Metab.* 2019;1(1):47-57.
141. Li SL, et al. SLC12A8 plays a key role in bladder cancer progression and EMT. *Open Med* (Wars). 2020;16(1):58-67.
142. Ibrahim AM, et al. Fibulin-2 is required for basement membrane integrity of mammary epithelium. *Sci Rep.* 2018;8(1):14139.
143. Ono RN, et al. Latent transforming growth factor beta-binding proteins and fibulins compete for fibrillin-1 and exhibit exquisite specificities in binding sites. *J Biol Chem.* 2009;284(25):16872-16881.
144. Baird BN, et al. Fibulin-2 is a driver of malignant progression in lung adenocarcinoma. *PLoS One.* 2013;8(6):e67054.
145. Sakai T, et al. Link between tumor-promoting fibrous microenvironment and an immunosuppressive microenvironment in stage I lung adenocarcinoma. *Lung Cancer.* 2018;126:64-71.
146. Nakamura H, et al. Organoid culture containing cancer cells and stromal cells reveals that podoplanin-positive cancer-associated fibroblasts enhance proliferation of lungcancer cells. *Lung Cancer.* 2019;134:100-107.
147. Neri S, et al. Podoplanin-expressing cancer-associated fibroblasts lead and enhance the local invasion of cancer cells in lung adenocarcinoma. *Int J Cancer.* 2015;137(4):784-796.
148. Miyata K, et al. Podoplanin enhances lung cancer cell growth in vivo by inducing platelet aggregation. *Sci Rep.* 2017;7(1):4059.
149. Chen Y, et al. Podoplanin+ tumor lymphatics are rate limiting for breast cancer metastasis. *PLoS Biol.* 2018;16(12):e2005907.
150. Bresson L, et al. Podoplanin regulates mammary stem cell function and tumorigenesis by potentiating Wnt/beta-catenin signaling. *Development.* 2018;145(4):dev160382.
151. Takemoto A, et al. A critical role of platelet TGF-β release in podoplanin-mediated tumour invasion and metastasis. *Sci Rep.* 2017;7:42186.
152. Martín-Villar E, et al. Podoplanin mediates ECM degradation by squamous carcinoma cells through control of invadopodia stability. *Oncogene.* 2015;34(34):4531-4544.
153. Rayes J, et al. Functional significance of the platelet immune receptors GPVI and CLEC-2. *J Clin Invest.* 2019;129(1):12-23.
154. Saruwatari K, et al. Aggressive tumor microenvironment of solid predominant lung adenocarcinoma subtype harboring with epidermal growth factor receptor mutations. *Lung*

Cancer. 2016;91:7-14.
155. Peters A, et al. Podoplanin negatively regulates CD4+ effector T cell responses. *J Clin Invest.* 2015;125(1):129-140.
156. Chihara N, et al. Induction and transcriptional regulation of the co-inhibitory gene module in T cells. *Nature.* 2018;558(7710):454-459.
157. Niska-Blakie J, et al. Knockout of the non-essential gene SUGCT creates diet-linked, age-related microbiome disbalance with a diabetes-like metabolic syndrome phenotype. *Cell Mol Life Sci.* 2020;77(17):3423-3439.
158. Del Toro D, et al. Structural basis of teneurin-latrophilin interaction in repulsive guidance of migrating neurons. *Cell.* 2020;180(2):323-339.e19.
159. Sando R, et al. Latrophilin GPCRs direct synapse specificity by coincident binding of FLRTs and teneurins. *Science.* 2019;363(6429):eaav7969.
160. Lan Q, et al. CCL26 participates in the PRL-3-induced promotion of colorectal cancer invasion by stimulating tumor-associated macrophage infiltration. *Mol Cancer Ther.* 2018;17(1):276-289.
161. Zajkowska M, Mroczko B. Eotaxins and their receptor in colorectal cancer-a literature review. *Cancers* (Basel). 2020;12(6):1383.
162. Lin ZY, et al. Cancer-associated fibroblasts up-regulate CCL2, CCL26, IL6 and LOXL2 genes related to promotion of cancer progression in hepatocellular carcinoma cells. *Biomed Pharmacother.* 2012;66(7):525-529.
163. Korbecki J, et al. CC chemokines in a tumor: a review of pro-cancer and anti-cancer properties of the ligands of receptors CCR1, CCR2, CCR3, and CCR4. *Int J Mol Sci.* 2020;21(21):8412.
164. Ando T, et al. Integrin α11 in non-small cell lung cancer is associated with tumor progression and postoperative recurrence. *Cancer Sci.* 2020;111(1):200-208.
165. Navab R, et al. Integrin α11β1 regulates cancer stromal stiffness and promotes tumorigenicity and metastasis in non-small cell lung cancer. *Oncogene.* 2016;35(15):1899-1908.
166. Bansal R, et al. Integrin alpha 11 in the regulation of the myofibroblast phenotype: implications for fibrotic diseases. *Exp Mol Med.* 2017;49(11):e396.
167. Primac I, et al. Stromal integrin α11 regulates PDGFR-β signaling and promotes breast cancer progression. *J Clin Invest.* 2019;129(11):4609-4628.
168. Zhu CQ, et al. Integrin alpha 11 regulates IGF2 expression in fibroblasts to enhance tumorigenicity of human non-small-cell lung cancer cells. *Proc Natl Acad Sci U S A.* 2007;104(28):11754-11759.
169. Grella A, et al. FGF2 overrides TGFβ1-driven integrin ITGA11 expression in human dermal fibroblasts. *J Cell Biochem.* 2016;117(4):1000-1008.
170. Leask A. A centralized communication network: Recent insights into the role of the cancer associated fibroblast in the development of drug resistance in tumors. *Semin Cell Dev Biol.* 2020;101:111-114.
171. Rousalova I, et al. Expression of proteinase inhibitor-9/serpinB9 in non-small cell lung carcinoma cells and tissues. *Int J Oncol.* 2010;36(1):275-283.
172. Medema JP, et al. Blockade of the granzyme B/perforin pathway through overexpression of the serine protease inhibitor PI-9/SPI-6 constitutes a mechanism for immune escape by tumors. *Proc Natl Acad Sci U S A.* 2001;98(20):11515-11520.
173. Jiang P, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy

response. *Nat Med.* 2018;24(10):1550-1558.
174. Jiang L, et al. Direct tumor killing and immunotherapy through anti-SerpinB9 therapy. *Cell.* 2020;183(5):1219-1233.e18.
175. Li C, et al. ADAMTS12 acts as a cancer promoter in colorectal cancer via activating the Wnt/β-catenin signaling pathway in vitro. *Ann Transl Med.* 2020;8(6):301.
176. Wen S, et al. Cancer-associated fibroblast (CAF)-derived IL32 promotes breast cancer cell invasion and metastasis via integrin β3-p38 MAPK signalling. *Cancer Lett.* 2019;442:320-332.
177. Yan H, et al. Role of interleukin-32 in cancer biology. *Oncol Lett.* 2018;16(1):41-47.
178. Han S, Yang Y. Interleukin-32: Frenemy in cancer? *BMB Rep.* 2019;52(3):165-174.
179. Calabrese F, et al. IL-32, a novel proinflammatory cytokine in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2008;178(9):894-901.
180. Rong Y, et al. CIL-32 was involved in cigarette smoke-induced pulmonary inflammation in COPD. *Clin Respir J.* 2015;9(4):430-435.
181. The Comparative Toxicogenomics Database™ (CTD™) https://ctdbase.org updated Dec. 12, 2020.

Supplemental Table 3. The comparison of MIND with other CAF-associated signatures.

		†Validation cohort (CAF, RFS) Univariate Cox model		
Signature	**n**	**HR**	**95% CI**	**P**
MIND	14	16.15	1.8-2127.9	0.009*
Vizoso 2015-M	14	0.32	0.03-1.71	0.188
MIND-GE		Not applicable		
Navab 2011-GE		Not applicable		
Sandoval 2013-T	14	0.51	0-4.5	0.613
		GSE39279 (Tumor, RFS) Univariate Cox model		
Signature	**n**	**HR**	**95% CI**	**P**
MIND	190	1.83	1.22-2.75	0.004*
Vizoso 2015-M	198	1.16	0.78-1.74	0.457
MIND-GE		Not applicable		
Navab 2011-GE		Not applicable		
Sandoval 2013-T	197	2.59	1.58-4.23	0*
		TCGA LUAD (Tumor, OS) Univariate Cox model		
Signature	**n**	**HR**	**95% CI**	**P**
MIND	440	1.46	1.07-2.01	0.018*
Vizoso 2015-M	447	1.01	0.74-1.38	0.94
MIND-GE	504	1.31	0.98-1.76	0.07
Navab 2011-GE	504	1.46	1.09-1.97	0.012*
Sandoval 2013-T	445	1.43	0.84-2.44	0.188

Abbreviations: RFS, relapse free survival; OS, overall survival; n, number of patients with both the signature and the survival data. HR, hazard ratio; CI, confidence interval. Not applicable: the signatures are not applicable on these datasets due to the lack of gene expression data. † Firth's Penalized likelihood method was used for this small dataset. *$P < 0.05$.

Supplementary Data 1. (separate file)
614 differentially expressed (DE) probes

Supplementary Data 2. (separate file)
Significant KEGG pathways in DE genes

Supplementary Data 3. (separate file)
14,781 differentially methylated (DM) CpG sites

Supplementary Data 4. (separate file)
1193 DM-DE pairs

Supplementary Data 5. (separate file)
Methylation index for NF/CAF discrimination (MIND)

Supplementary Methods

Gene Expression Data Analysis

We performed Affymetrix Human U133 Plus2.0 array for gene expression profiling on 25 pairs of cancer-associated fibroblasts (CAFs) and normal fibroblasts (NFs) with adequate RNA quality (one sample not passing the QC was excluded) from non-small cell lung cancer (NSCLC) patients in our discovery cohort (N=26). This platform contains 54,675 probes covering 21,649 genes. RMA normalization was performed using the package affy[1] in *R*. Probes on the sex chromosomes and probes with chromosome or gene symbol annotated as "-" were removed after normalization, leaving 41,158 probes.

Differential Expression Analysis

To identify genes exhibiting differential expression (DE) between CAFs and NFs, we performed Significance Analysis in Microarray (SAM) analysis on the 25 NF/CAF pairs using the package samr[2] in *R*. A total of 614 DE probes were obtained (mean fold change > 1.5 and $Q < 0.1$), where 242 were up-regulated (189 genes), and 372 were down-regulated in CAFs (272 genes).

Pathway Enrichment Analysis

Significant KEGG[3] pathways for the DE genes were determined by the functional annotation tool of DAVID Bioinformatics Resource 6.8.[4] Seven KEGG pathways were identified at the false discovery level of 10%, including ECM-receptor interaction, PI3K-Akt signaling pathway, focal adhesion, TGF-β signaling pathway, pathways in cancer, hematopoietic cell lineage and cell cycle. Gene Set Enrichment Analysis (GSEA)[5] was also performed to identify KEGG pathways exhibiting significant differences in the global

expression profile between CAFs and NFs. One pathway, ECM-receptor interaction, was identified ($Q < 0.1$).

DNA Methylation Data Analysis

We used Illumina Infinium Methylation 450K array for DNA methylation profiling on the 26 NF/CAF pairs. This platform contains 485,460 CpG sites. Background correction and normalization were performed using the package methylumi[6] in *R*. After data preprocessing, sites on the sex chromosomes and sites with missing values in any sample were removed (472,676 sites remained).

Differential Methylation Analysis

To identify CpG sites exhibiting differential methylation (DM) between CAFs and NFs, we performed Wilcoxon signed-rank test on the 26 NF/CAF pairs. Benjamini-Hochberg method was used to control the false discovery rate. A total of 14,781 DM CpG sites were obtained (mean beta-value difference > 0.1 and $Q < 0.1$), where 5,951 (40%) were hypermethylated and 8,830 (60%) hypomethylated in CAF (Supplemental Figure 4b).

Correlation of clinical variables with CAF/NF Methylation Differences

We explored the heterogeneity of methylation differences for patients with varied clinical backgrounds. Each categorial clinical variable was coded as a binary variable; smoking: ever versus never, stage: early (I and II) versus late (III and IV), histology: adenocarcinomas versus non-adenocarcinomas. The delta beta-values ($\Delta\beta$) of the 14,781 DM CpG sites between patients grouped by each categorial clinical variable were compared by Welch's t-test. For the continuous clinical variable: age, Pearson's correlation coefficient was used. When controlling the false discovery rate at the level of 0.1 level, we obtained 3,707 (25%) smoking-associated DM CpG sites with greater differences in the

ever-smoker group than the never-smoker group. There was no other significant association found in the other clinical variables.

Cis-Correlation between Differentially Methylated CpG Sites and their Target Genes

Among the 18,837 genes with probes in both arrays, 419 were DE genes, whereas 220 (53%) had at least one DM CpG site. In contrast, among the 18,418 non-DE genes, only 3,999 (22%) genes had at least one DM CpG site. This significant enrichment (OR = 3.99, $P < 0.001$, Fisher exact test) suggests that the methylation changes significantly contributed to the differential expression of the target genes.

To identify the DM CpG sites that had cis-correlation with their target genes, we matched the 614 DE probes obtained earlier with 14,781 DM sites by gene symbol, leading to a total of 1,193 DM-DE pairs. For each DM-DE pair, Spearman's Rho was used to evaluate the cis-correlation between the methylation changes and the expression fold changes (FC). We identified 486 DM-DE pairs exhibiting significant correlations ($Q < 0.1$), and 482 of them had signs consistent with the ratios of the mean log2FC to the mean $\Delta\beta$.

Overall, 340 (71%) negatively correlated pairs were found among the 482 cis-correlated DM-DE pairs. Based on the location of the CpG sites to their target genes, DM-DE pairs can be assigned into three groups: promoter, within 1,500 bps to the transcription starting site (TSS); body, from the TSS to the end of the transcript; or multiple, in the promoter and body for different transcripts. Among the 57 pairs assigned to the promoter, 24 pairs (42%) were hypermethylated and down-regulated in CAFs while 23 pairs (40%) were hypomethylated and up-regulated. On the other hand, among the 395 pairs assigned to the body, 77 pairs (19%) were hypermethylated and down-regulated in CAFs while 187

(47%) were hypomethylated and up-regulated. And the rest 30 pairs are assigned to the "multiple" group. (Supplementary Methods Figure 1).

Candidate CpG Sites Selection

After intersecting the 3,707 smoking-associated DM CpG sites with the CpGs in the 482 cis-correlated DM-DE pairs, 93 CpG sites were obtained. To further select a smaller number of high-quality methylation probes, we imposed the following filtering steps. First, 10 CpG sites that have SNPs in the nearby region were excluded. Second, to be parsimony, for the remaining 83 CpG sites located in the 54 genes, we only kept the CpG sites with the largest beta-value difference between CAFs and NFs for each gene. The list of 54 selected CpG sites is given in Supplementary Data 5.

Derivation of the Methylation Index for NF/CAF Discrimination

Principal component (PC) analysis was conducted on the methylation profile of the 54 selected CpG sites. The first PC explained 67.0% total variance, and the second PC explained 7.5% total variance. A methylation index for NF/CAF discrimination (MIND) was defined as the weighted sum of the centered beta-values by the loadings of the first principal component.

The receiver operating characteristic (ROC) curve was used to illustrate the discrimination ability of MIND. The area under the ROC curve of our discovery cohort was 0.88 (95% CI: 0.80-0.97). Based on Youden's index, the optimal cutoff -0.59 had 88% sensitivity and 77% specificity. As a comparison, we studied the discrimination ability of 12 known CAF markers, including ten positive and two negative markers, in gene expression. There was no single known marker with discriminant ability as good as MIND (Supplementary Methods Figure 2). We noticed the sign inconsistent issue for some known

markers with our data. Five positive markers *TNC*, *PDFGRA*, *CSPG4*, *DES*, and *VIM* had lower expression in CAFs than NFs. We performed the linear discriminant analysis on the seven sign consistent markers. The area under the ROC curve of the best linear discriminant direction was 0.75 (95% CI: 0.60-0.89). The discrimination ability of MIND was significantly better ($P = 0.018$ DeLong's test). If we ignore the sign inconsistent issue and use the best linear discriminant direction of all 12 known CAF markers, the area under the ROC curve will be 0.89 (95% CI: 0.80-0.98), and the optimal cutoff has 88% sensitivity and 76% specificity. MIND showed the same discrimination ability to the best that the 12 known CAF markers can get.

We also examined the discrimination ability of MIND on an independent cohort consisting of 12 NF/CAF pairs from NSCLC patients (GSE68851).[7] The area under the ROC curve was 0.83 (95% CI: 0.66-1), and the optimal cutoff had 73% sensitivity and 92% specificity (Supplemental Figure 6).

The Prognostic Ability of MIND in CAFs

To examine the prognostics ability of the MIND index, we compared the relapse-free survival (RFS) curves of the two patient groups split by MIND using the log-rank test ($P = 0.013$, Figure 5a). In our discovery cohort, the cutoff was selected as the value separating CAFs from NFs by controlling the probability of misclassifying NFs at no more than 5% while maximizing the probability of correct classification of CAFs. This selecting procedure is objective and without using the survival data. Multivariate Cox regression with backward selection confirmed the significance of MIND adjusted by stage, age, gender and smoking (HR = 9.29, 95% CI: 1.14-75.44, $P = 0.037$; Table 1).

We also examined the prognostic power of MIND in an independent validation cohort consisting of 14 NSCLC patients. DNA methylation profiling of their CAF and NF samples was performed using Illumina Infinium Methylation 450K array. The same set of CpG sites and the weights derived from our discovery cohort were used when computing MIND. The median value of MIND was used to split patients into two groups. The results from the log-rank test (P = 0.007, Figure 5b) and the multivariate Cox regression with backward selection remain significant (HR = 29.17; 95% CI = 2.19-6520.53; P = 0.006). Specifically, among the 14 patients, none of the patients in the MIND-low group had RFS events. The package coxphf in *R* was used to fit the Cox regression with Firth's Penalized likelihood[8].

The Prognostic Ability of MIND in Tumor Samples

Although MIND was derived from the primary cultured CAFs/NFs, and although the signals from CAFs could be profiled together with the bulk of tumor tissues, we still explored the clinical value of MIND as a prognosis signature in two independent cohorts GSE39279 (444 NSCLCs)[9] and The Cancer Genome Atlas (TCGA)[10] lung adenocarcinomas (456 LUADs) that have tumor DNA methylation profiles. DNA methylation data and curated clinical data[11] were downloaded from Gene Expression Omnibus (GEO) and TCGA GDC data portal (https://gdc.cancer.gov/about-data/publications/pancanatlas).

For the GSE39279 cohort, after removing 13 patients with missing beta values of the 54 CpGs, the MIND scores of the remaining 431 patients were computed, using the weights already determined at the discovery stage earlier. Before looking at the survival data, patients were split evenly into two groups using the median cutoff (215 MIND-high

and 216 MIND-low). After that, we examined the patient survival data. There were 241 patients missing survival data, 126 from the MIND-high group and 115 from the MIND-low group. The deviation from the expected number of 120.5 in each group was statistically insignificant (P = 0.2863, Fisher exact test), an indication of missing at random. For patients with missing survival data, we simply dropped them from further analysis. This resulted in the survival analysis by the log-rank test with uneven size for MIND-high and MIND-low groups (Figure 5c). For the TCGA LUAD cohort, among the 54 CpG sites, 9 of them were unavailable in the methylation dataset. Thus, the available 45 CpG sites were used to compute the MIND scores for a total of 449 patients, using the weights already determined earlier at our discovery stage. Patients were split evenly into MIND-high and MIND-low groups, using the median cutoff. After removing patients with missing survival data, we performed the log-rank test on the remaining patients (223 MIND-high and 217 MIND-low; Figure 5d). Patients with high MIND scores have shorter RFS in the GSE39279 cohort (P = 0.003, Figure 5c) and shorter OS in TCGA LUAD (P = 0.018, Figure 5d). As for another way of patient splitting, the patients with survival data were evenly divided into high and low MIND score groups (Supplementary Methods Figure 3). The log-rank test results exhibited significant.

The Discrimination Ability and Prognostic Effect of the 54 Target Gene Expressions

Principal component (PC) analysis was performed using the standardized gene expression of the 54 target genes. The first PC explained 48.4% total variance, and the second PC explained 9.7% total variance. An expression index (MIND-GE) was defined as the weighted sum of the standardized expression by the loadings of the first principal component.

The area under the ROC curve for MIND-GE was 0.83 (95% CI: 0.70-0.95; Supplemental Figure 7a), and the optimal cutoff had 84% sensitivity and 80% specificity. Its discrimination ability was slightly lower than MIND in our cohort. We also examined the discrimination ability of MIND-GE on an independent cohort consisting of 15 NF/CAF pairs from NSCLC patients (GSE22874). The area under the ROC curve was 0.93 (95% CI: 0.84-1), and the optimal cutoff had 80% sensitivity and 93% specificity (Supplemental Figure 7b). MIND-GE also showed significant performance on recurrence prediction in our discovery cohort using the median split ($P = 0.041$; Supplemental Figure 9a). Applying MIND-GE to the gene expression profiling data of the TCGA LUAD cohort found marginally significant in OS prediction ($P = 0.069$; Supplemental Figure 9b). This result showed that MIND outperformed MIND-GE in the prognostic assessment.

Comparisons of Molecular Prognostic Signatures

To demonstrate the prognostic ability of MIND compared to other prognostic signatures, we benchmarked against 5 molecular signatures, including 2 methylation-based CAF-associated signatures (**MIND** and **Vizoso 2015-M**[12]), 2 expression-based CAF-associated signatures (**MIND-GE** and **Navab 2011-GE**[13]), and 1 methylation-based tumor-associated signature (**Sandoval 2013-T**[14]), on the three datasets described in the manuscript (Supplemental Table 3).

MIND: A methylation-based CAF-associated signature derived from our discovery cohort, representing a linear combination of the centered beta value of 54 CpG sites.

Vizoso 2015-M: Vizoso et al. identified that abnormal DNA methylation on the *EDARADD* promoter can be used as a prognostic biomarker. In this comparison, the

average of the centered beta value of the 3 DM CpG sites they identified was used as a signature.

MIND-GE: An expression-based CAF-associated signature derived from our discovery cohort, representing a linear combination of the standardized gene expression of 54 genes.

Navab 2011-GE: An expression-based CAF-associated signature, representing a linear combination of the standardized gene expression of 11 genes (13 probesets) derived by Navab et al. The weights from the original paper were applied directly.

Sandoval 2013-T: A methylation-based tumor-associated signature derived by Sandoval et al. using the dataset from GSE39279. The signature is based on the number of hypermethylated events of the 5 genes they identified. In this comparison, patients with any of the 10 DM-CpGs hypermethylated (beta value > 0.4) in the 5 genes were assigned to the high-risk group, otherwise were assigned to the low-risk group.

The first four signatures were first dichotomized by the median value among all the patients with methylation/expression profiles. Prognostic ability of the dichotomized signatures was examined by univariate Cox proportional hazard regression. Notably, some signatures were slightly modified when applied to the TCGA LUAD cohort because some CpGs' methylation values and some genes' expression values are not available in their preprocessed dataset. MIND was defined as the linear combination of the centered beta value of the 45 available CpGs using the same weights as in the original version. MIND-GE was defined as the linear combination of the standardized gene expression of the 48 available genes using the same weights as in the original version. When computing Navab's 11-gene signature which was defined at the probeset level, the summation of the

weights corresponding to the probesets were used for the linear combination of the 11 genes.

The result showed that MIND is the only significant signature in the validation dataset when applying the signatures on the CAF samples, indicating that MIND can better reflect the malignancy level of CAFs than the other molecular signatures. Moreover, when applying the signatures on the tumor samples, MIND remains significant in both tumor datasets GSE39279 and the TCGA. Sandoval 2013-T is the most significant signature in the tumor dataset GSE39279 since it was derived from this dataset using the survival data. However, it's not significant in the TCGA dataset. On the other hand, Navab 2011-GE is significant in the TCGA dataset with the same hazard ratio as MIND. Together, the results illustrate the robust prognostic ability of MIND.

References

1. Gautier L, Cope L, Bolstad BM, et al: Affy - Analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20:307–315, 2004

2. Tibshirani R, Chu G, Narasimhan B, et al: samr: SAM: Significance Analysis of Microarrays. *R* package version 2.0, 2011

3. Kanehisa M, Goto S, Furumichi M, et al: KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res 38:355–360, 2009

4. Huang DW, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57, 2009

5. Subramanian A, Tamayo P, Mootha VK, et al: Gene set enrichment analysis : A knowledge-based approach for interpreting genome-wide. Proc Natl Acad Sci U S A 102:15545–15550, 2005

6. Davis S, Du P, Bilke S, et al: methylumi: Handle Illumina methylation data. *R* package version 2.22.0, 2017

7. Vizoso M, Puig M, Javier Carmona F, et al: Aberrant DNA methylation in Non Small Cell Lung Cancer associated fibroblasts. Carcinogenesis 36:1453–1463, 2015

8. Heinze G, Ploner M: coxphf: Cox Regression with Firth's Penalized Likelinhood. *R* package version 1.13.0, 2018

9. Sandoval J, Mendez-Gonzalez J, Nadal E, et al: A prognostic DNA methylation signature for stage I non-small-cell lung cancer. J Clin Oncol 31:4140–4147, 2013

10. Collisson E a., Campbell JD, Brooks AN, et al: Comprehensive molecular profiling of lung adenocarcinoma. Nature 511:543–50, 2014

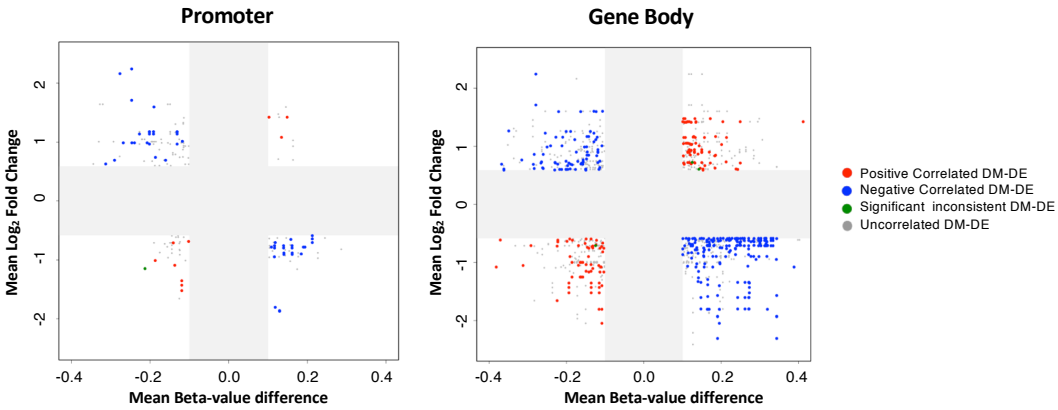
11. Liu J, Lichtenberg T, Hoadley KA, et al: An Integrated TCGA Pan-Cancer Clinical

Data Resource to Drive High-Quality Survival Outcome Analytics. Cell 173:400-416.e11, 2018

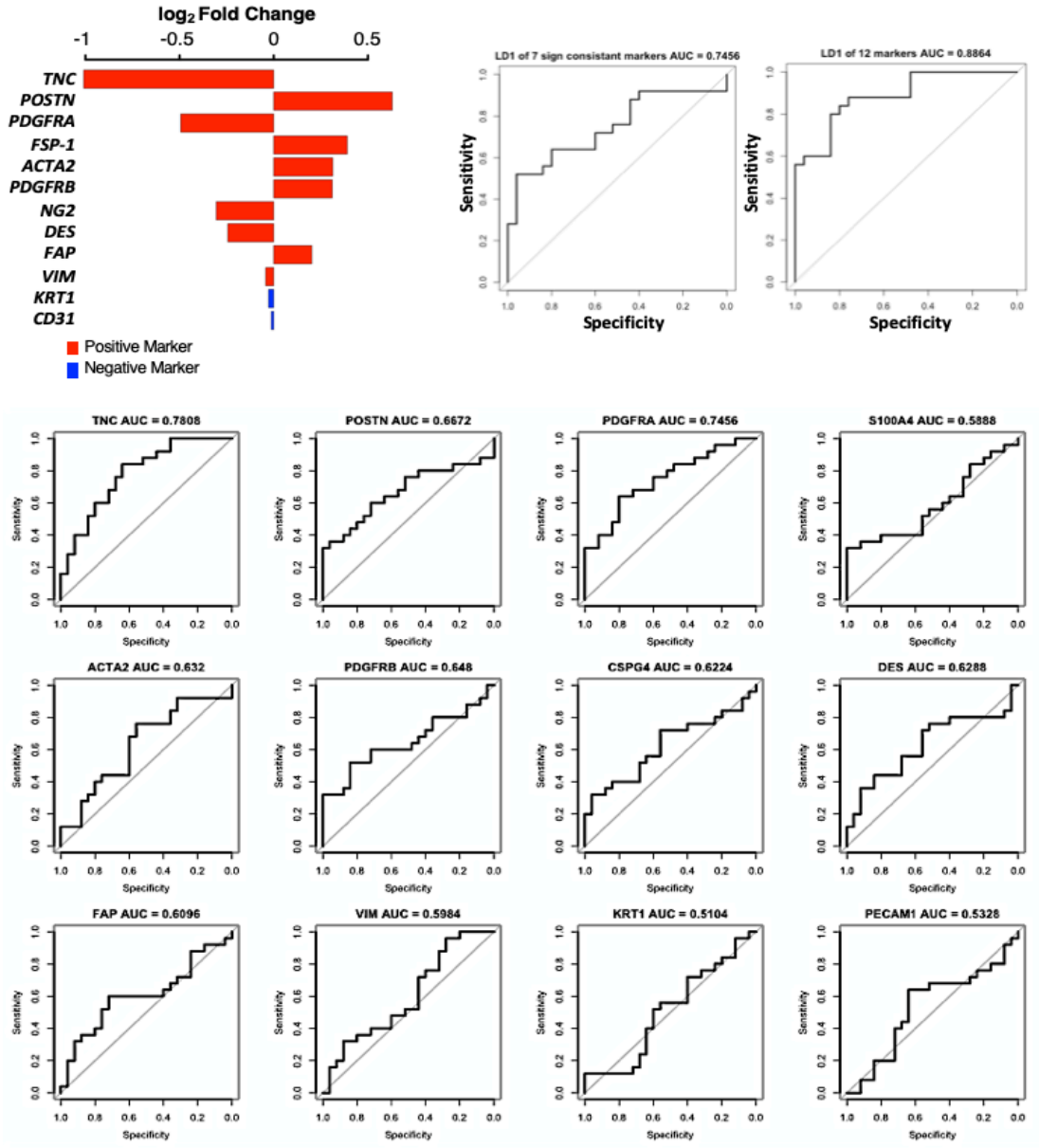
12. Vizoso M, Puig M, Carmona FJ, Maqueda M, Velasquez A, Gomez A, et al. Aberrant DNA methylation in non-small cell lung cancer-associated fibroblasts. *Carcinogenesis.* 2015;36(12):1453-63.

13. Navab R, Strumpf D, Bandarchi B, Zhu CQ, Pintilie M, Ramnarine VR, et al. Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer. *Proc Natl Acad Sci U S A.* 2011;108(17):7160-5.

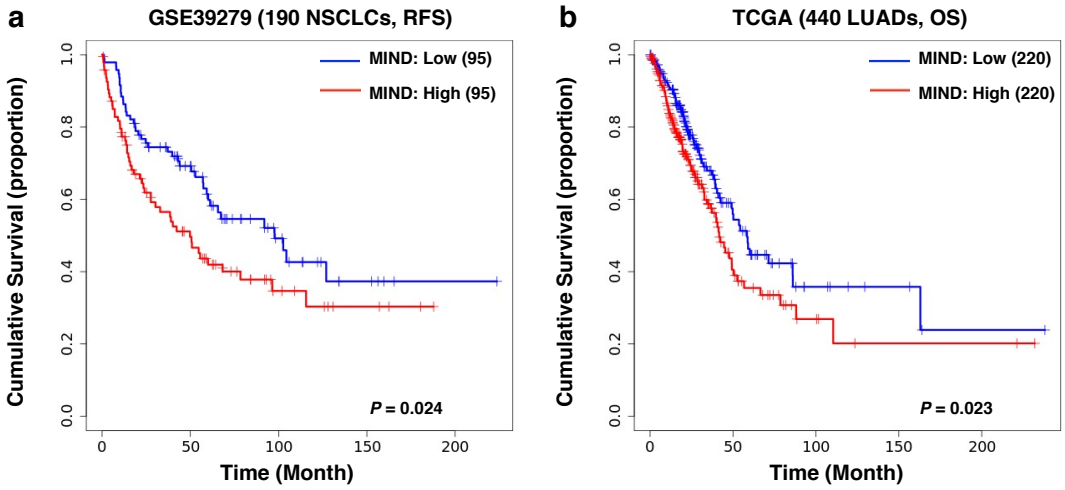
14. Sandoval J, Mendez-Gonzalez J, Nadal E, Chen G, Carmona FJ, Sayols S, et al. A prognostic DNA methylation signature for stage I non-small-cell lung cancer. *J Clin Oncol.* 2013;31(32):4140-7.



Supplementary Methods Figure 1. **Negatively correlated DM-DE probes were enriched in the promoter regions.** With regard to genomic location of CpG sites, significantly cis-correlated DM-DE pairs were plotted with the mean beta-value difference in DNA methylation versus mean log fold change in gene expression in the promoter region (left) and gene body (right), respectively. Grey zone reflected the selective threshold. In the promoter region, 47/57 (82.5%) were negatively correlated whereas 264/395 (66.8%) were in the body.



Supplementary Methods Figure 2. The discrimination ability of 12 known CAF markers in gene expression.



Supplementary Methods Figure 3. **Significance of MIND for its prognostic power.** Applying MIND to DNA methylation profiling of tumor samples from two independent cohorts with equal size groups. **(a)** 190 NCSLC Patients in GSE39279 with non-missing MIND score and survival data were stratified into the $MIND^{high}$ and $MIND^{low}$ groups using the median cutoff (P = 0.024). **(b)** 440 patients in the TCGA LUAD cohort with non-missing MIND score and survival data were stratified into the $MIND^{high}$ and $MIND^{low}$ groups using the median cutoff (P = 0.023). RFS: relapse-free survival; OS: overall survival.