Supporting Information

for

Postchronic Single-Walled Carbon Nanotube Exposure Causes Irreversible Malignant Transformation of Human Bronchial Epithelial Cells through DNA Methylation Changes

Jin Wang,[†,#] *Xin Tian,*[‡,#] *Jie Zhang,*[†] *Lirong Tan,*[†] *Nan OuYang,*[†] *Beibei Jia,*[†] *Chunying Chen,*[§] *Cuicui Ge,*[*,‡] *and Jianxiang Li*[*,†]

[†]Department of Toxicology, School of Public Health, Medicine College, Soochow University, Suzhou, 215123, China

[‡]State Key Laboratory of Radiation Medicine and Protection, School of Radiation Medicine and Protection, School for Radiological and Interdisciplinary Sciences (RAD-X), Collaborative Innovation Center of Radiation Medicine of Jiangsu Higher Education Institutions, Soochow University, Suzhou 215123, China

[§]Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety & CAS Center for Excellence in Nanoscience, National Center for Nanoscience and Technology of China, Chinese Academy of Sciences, Beijing 100190, China

[#]J.W. and X.T. contributed equally to this work.

[*]Corresponding authors:

Cuicui Ge, E-mail: ccge@suda.edu.cn; Jianxiang Li, E-mail: aljxcr@suda.edu.cn

Other Methods

DNA Methylation Microarray Profiling. The level of epigenome-wide DNA methylation of BEAS-2B cells was analyzed by Infinium Methylation EPIC 850K Beadchip (Illumina Inc., USA). DNA methylation is determined by measuring the fluorescent signals from the unmethylated (U) and methylated (M) probes specific for each site included in the array. The probe signal intensities were normalized by quantile algorithm in 'methylumi' and 'minif' packages from R Bioconductor and the normalized signal intensities were then used to calculate the β values. After calculating the arithmetic average of the signal values of all the probes in each region, a pooled *t*-test was used to screen the differentially methylated genes (DMGs) of the sample group, as defined by a threshold of $P < 0.05$ and $|\Delta\beta| > 0.14$. To verify the reliability of the methylation microarray result, the methylation status of the top 3 DMGs were measured by methylation-specific PCR (MSP). The total DNA was extracted using the DNA extraction kit (Tiangen, China) and modified using the EpiTect Bisulfite Kit (Qiagen, Germany). MSP primers were designed using MethPrimer (http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi) (primers and conditions available in **Table S9**).

mRNA Sequencing. Total cellular RNA was isolated from BEAS-2B cells by using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. mRNA sequencing was performed on the Illumina HiSeq 4000 platform, using 150 bp paired-end sequencing with a fragment size of 200-400 bp. The Illumina TruSeq RNA kit was used for preparing the mRNA sequencing

library. The mRNAs with expression profiles that differed between the samples were normalized as fragments per kilobase of transcript per million mapped reads (FPKM). Differential expression analysis of two groups was performed using the "DEGseq" package. $P < 0.05$ and $|\log2 \text{ (fold change)}| \geq 1$ were set as the threshold for differentially expressed genes (DEGs).

Real-Time Quantitative PCR. A total of 1.5 μg RNA isolated from each sample was reversely transcribed into complementary DNA (cDNA) using Revert Aid First Strand Complementary DNA Synthesis kit (ThermoFisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions. RT-qPCR was performed using NovoScript® SYBR Two-Step qRT-PCR kit (novoprotein, China) with a QuantStudioTM 6 Flex qRT-PCR system (ABI, Rockford, IL, USA). The internal control for the analysis of mRNA was GAPDH. The primer used for qPCR were listed in **Table S10**.

Enrichment Analyses. To investigate the biological function of DEGs or DMGs, the online enrichment tool DAVID (https://david.ncifcrf.gov/) was used to undergo the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. The results were visualized by R package "GOplot".

Methylation Pattern of DMGs in Lung Cancer. The MEXPRESS (https://mexpress.be) is a user-friendly online tool for analyzing and visualization of TCGA data, providing a simple way for researchers to analyze the multi-omics data in the TCGA database. The correlation between the mRNA expression levels and DNA

methylation values of DMGs in lung cancer were assessed using the MEXPRESS tool.

Table S1. Characterization of SWCNTs

Property	Value	Unit	Characterization method
Outer Diameter	<2	nm	HRTEM, Raman
Purity	>95	wt%	TGA, TEM
Length	5-30	μm	TEM
Special Surface Area	>490	m^2/g	BET
Ash content	<5	wt%	HRTEM, TGA
Electric Conductivity	>100	s/cm	NA
Tap Density	0.14	g/cm^3	NA
Ignition point	610	^oC	NA
I_g/I_d	>20	NA	Raman

NA: not applicable; HRTEM: high resolution transmission electron microscope; TGA: thermogravimetric analyzer

Table S2. Physicochemical characterization of SWCNTs and carbon black

	unit	SWCNTs	Carbon black
Hydrodynamic size in H_2O	nm	158.3 ± 9	212.7 ± 0.6
Hydrodynamic size in DMEM	nm	194.5 ± 22	260.5 ± 0.2
ζ potential in H_2O	mV	-33.1 ± 0.6	-9.1 ± 0.1
ζ potential in DMEM	mV	-9.3 ± 0.4	-6.8 ± 0.5
Elemental analysis	wt.%	0.54% Al, 0.68% Fe	0.31% Al, 0.74% Fe

Table S3. Top 10 genes methylated in the promoter region

TSS1500	*P* value	β difference	TSS200	*P* value	β difference
AREG	4.29	-0.314	XAGE5	2.51	-0.338
ATIC	4.00	-0.312	CXCL10	3.10	-0.279
CD2AP	4.90	-0.304	FUCA1	4.55	-0.268
NDUFV3	3.91	-0.291	C9orf91	3.50	-0.265
EIF3I	2.54	-0.287	EMR3	3.14	-0.262
ZNF438	3.25	-0.285	RANBP17	3.75	-0.261
ECH1	4.50	-0.270	SCNN1G	3.31	-0.254
PRKAG1	2.33	-0.267	SIM2	3.81	-0.250
TTC26	2.22	-0.263	TLL1	2.89	-0.247
EMP2	2.14	-0.262	TMSB15B	2.84	-0.246
SREBF1	2.62	0.268	MIR628	4.44	0.269
FKSG83	2.95	0.285	TCEAL7	2.56	0.271
TAB3	3.12	0.286	RTN4RL1	3.39	0.272
GFAP	2.60	0.290	FOXP4-AS1	2.80	0.284
PIK3CG	2.01	0.301	C21orf128	2.16	0.284
AWAT2	2.92	0.329	KCNG3	3.51	0.289
SRPX2	3.19	0.336	MIR325	2.18	0.297
ZCCHC5	2.65	0.339	ITM2A	2.68	0.316
YES1	2.71	0.355	SNORA8	3.06	0.336
COX10-AS1	2.97	0.361	CNBD1	2.90	0.339

Table S4. Biological process enrichment of the differentially methylated genes

Term	Count	*P* value	Gene ratio (%)
Astrocyte development	4	0.010	25.00%
Macrophage derived foam cell differentiation	3	0.012	50.00%
Cilium morphogenesis	10	0.018	7.35%
Mitotic spindle assembly	5	0.020	13.89%
Transcription-coupled nucleotide-excision repair	7	0.021	9.46%
Nucleotide-excision repair, DNA incision	5	0.024	13.16%
Potassium ion transmembrane transport	9	0.025	7.44%
Integrin-mediated signaling pathway	8	0.026	8.08%
Hemopoiesis	6	0.028	10.17%
Positive regulation of transcription from RNA polymerase ii promoter	40	0.031	4.08%
Nucleotide-binding oligomerization domain containing signaling pathway	4	0.035	16.00%
Phosphatidylcholine biosynthetic process	4	0.035	16.00%
Sleep	3	0.039	27.27%
Blood circulation	5	0.041	11.11%
Phosphatidylinositol 3-kinase signaling	4	0.043	14.81%
Cardiac muscle cell differentiation	4	0.043	14.81%
Fatty acid transport	3	0.046	25.00%
Excitatory postsynaptic potential	4	0.047	14.29%
Negative regulation of interferon-gamma production	4	0.047	14.29%
Covalent chromatin modification	8	0.047	7.08%
Cellular response to starvation	5	0.047	10.64%
Organ regeneration	5	0.047	10.64%
Chemical synaptic transmission	13	0.049	5.42%

Table S5. Top 20 differentially expressed genes in PE60 cells

Up	-Log10 (P value)	Log2 (Fold change)	Down	-Log10 (P value)	Log2 (Fold change)
FLG	<300	-4.814	PLAGL1	41.40	1.417
SCN9A	257.30	-4.050	ENPP1	38.60	1.451
TNS4	121.95	-3.666	C1QTNF1	19.71	1.474
AC079466.1	199.04	-3.582	TIE1	58.29	1.481
PSG1	<300	-3.205	THBD	22.89	1.488
MFAP5	97.36	-3.089	TBKBP1	38.23	1.494
CHAT	45.11	-2.927	COL4A1	39.46	1.496
PDE1C	<300	-2.552	COL4A2	78.13	1.526
FYB	141.30	-2.471	APCDD1L	18.84	1.537
PSG11	50.11	-2.466	C3	139.00	1.545
PSG5	144.22	-2.360	PRKD3	95.14	1.563
SLAMF7	41.80	-2.324	CAMK2N1	41.20	1.599
DSP	108.38	-2.158	CXCL1	82.34	1.618
PSG2	131.83	-2.140	CYP26B1	48.75	1.710
PSG6	67.75	-2.114	L1CAM	142.75	1.758
FLJ22447	41.57	-2.094	CHI3L1	34.55	1.818
GRIK5	42.02	-2.058	IGFN1	55.42	1.878
EREG	125.00	-2.055	MMP9	38.27	2.018
PLAC8	30.30	-2.035	ADCY1	87.14	2.100
KIRREL3	40.67	-2.011	CXCL8	187.31	2.130

Table S6. Biological process enrichment of the differentially expressed genes

Term	Count	*P* value	Gene ratio (%)
Female pregnancy	7	<0.001	7.87%
Angiogenesis	9	<0.001	4.04%
Immune response	12	<0.001	2.85%
Inflammatory response	11	0.001	2.90%
Peripheral nervous system development	4	0.001	16.67%
Collagen catabolic process	5	0.001	7.81%
Chemokine-mediated signaling pathway	5	0.002	7.04%
Response to molecule of bacterial origin	3	0.002	33.33%
Chemotaxis	6	0.002	4.92%
Wound healing	5	0.003	6.25%
MAPK cascade	8	0.003	3.05%
Signal transduction	18	0.005	1.55%
Peptidyl-tyrosine phosphorylation	6	0.006	3.92%
Skeletal muscle cell differentiation	4	0.006	8.16%
Cellular response to lipopolysaccharide	5	0.010	4.42%
Synapse assembly	4	0.010	6.56%
Positive regulation of neutrophil chemotaxis	3	0.011	13.64%
Regulation of glucose metabolic process	3	0.011	13.64%
Positive regulation of phosphorylation	3	0.015	12.00%
Extracellular matrix organization	6	0.015	3.06%
Cellular response to interleukin-1	4	0.016	5.63%
Patterning of blood vessels	3	0.018	10.71%
Ossification	4	0.021	5.00%
Peyer's patch morphogenesis	2	0.022	66.67%
Positive regulation of MAPK cascade	4	0.022	4.94%
Mesoderm development	3	0.023	9.38%
Positive regulation of protein kinase b signaling	4	0.024	4.76%
Skin development	3	0.026	8.82%
Protein localization to adherens junction	2	0.029	50.00%
Axon guidance	5	0.030	3.14%
Response to lipopolysaccharide	5	0.033	3.05%
Collagen fibril organization	3	0.034	7.69%
Transcription from RNA polymerase ii promoter	9	0.036	1.75%

Nail development	2	0.044	33.33%
Collagen-activated tyrosine kinase receptor signaling pathway	2	0.044	33.33%
Positive regulation of receptor binding	2	0.044	33.33%
Outflow tract morphogenesis	3	0.045	6.52%
Cellular response to amino acid stimulus	3	0.047	6.38%
Cellular response to tumor necrosis factor	4	0.048	3.64%
Regulation of cell proliferation	5	0.048	2.70%

Table S7. KEGG pathway enrichment of the differentially expressed genes

Term	Count	*P* value	Gene ratio (%)
Cytokine-cytokine receptor interaction	9	0.001	3.70%
Pathways in cancer	11	0.002	2.80%
Amoebiasis	6	0.002	5.66%
TNF signaling pathway	6	0.002	5.61%
ECM-receptor interaction	5	0.007	5.75%
Hematopoietic cell lineage	5	0.007	5.75%
Glutathione metabolism	4	0.010	7.84%
Legionellosis	4	0.012	7.41%
NOD-like receptor signaling pathway	4	0.013	7.14%
ARVC	4	0.021	5.97%
Salmonella infection	4	0.037	4.82%
Protein digestion and absorption	4	0.043	4.55%

Table S8. Gene list of the cross-analysis of DMGs and DEGs

Gene symbol	log2 (Fold change)	β difference	LUAD	LUSC
PLEKHA7	-1.61924	-0.1824	up	down
GLIS1	-1.56449	-0.15586	up	up
SYBU	-0.85087	-0.15331	down	down
KANK4	-1.8115	-0.1493	down	down
CD7	-0.60415	-0.14431	no	down
PPL	-1.07512	0.143542	down	no
PLAC8	-2.03495	0.147265	down	down
MYEOV	-0.9121	0.147738	up	up
ZNF154	-1.14443	0.156159	no	down
CCR7	-1.45614	0.156251	no	down
SPINT2	-1.45623	0.173276	up	up
HAS3	-0.95282	0.182273	down	up
PPP1R3C	-0.93161	0.200882	down	down
PKP2	-0.72219	0.207832	up	up
CAMK2N1	-2.01098	0.211313	no	down
NLRP10	-1.34794	0.213733	no	up
ZNF674-AS1	-0.61396	0.22099	-	-
NNMT	-0.76466	0.224918	no	down
MYPN	-1.36403	0.227799	up	no
ITM2A	-1.02233	0.316059	down	down
PRSS12	0.589252	-0.14974	down	down
TMPRSS9	0.863492	-0.26303	up	up
PIM2	1.656682	-0.23817	up	no
ELMOD1	0.635065	-0.17515	down	down
ZMIZ2	0.653848	-0.16824	up	no
PAPPA	1.838318	-0.16107	no	no
SNAI3	0.681246	-0.15791	down	down
CACTIN-AS1	0.59136	-0.1545	-	-
SNX19	0.765857	-0.15034	no	down
IL32	0.913916	0.14385	no	down
RGS17	0.957131	0.146475	up	up
SPOCD1	1.093398	0.151233	down	down

RAB6B	0.94437	0.176995	up	up
MAPK8IP1	0.868597	0.177736	no	down
CCDC71L	0.683085	0.197512	-	-
KIRREL3	1.599182	0.220459	up	no
ABCC9	0.93518	0.256213	down	down

Table S9. Methylation-specific PCR primers used in this study

Primers	Sequence	Product size
PIM2_MF	GGAGTTTTGGATAGTTATAAGCGC	103
PIM2_MR	CCTTTCCTAAAAAATTAATTTACCG	
PIM2_UF	GAGTTTTGGATAGTTATAAGTGTGA	103
PIM2_UR	TCCTTTCCTAAAAAATTAATTTACCAC	
ABCA2_MF	TGAAATCGTTTATTTTTTAGGTTCG	137
ABCA2_MR	TACAAAAACTCCCTAACTACAACCG	
ABCA2_UF	GAAATTGTTTATTTTTTAGGTTTGT	135
ABCA2_UR	ACAAAAACTCCCTAACTACAACCAC	
CRYBG3_MF	TTTTTAAAAGTTCGAAGAAAATCGT	173
CRYBG3_MR	GACCTAAAAACCTAAAAAAAATCGC	
CRYBG3_UF	TTTTTAAAAGTTTGAAGAAAATTGT	174
CRYBG3_UR	CAACCTAAAAACCTAAAAAAAATCAC	

Table S10. qPCR primers used in this study

Primers	Sequence
GAPDH_F	GGAGCGAGATCCCTCCAAAAT
GAPDH _R	GGCTGTTGTCATACTTCTCATGG
FLG_F	TGAAGCCTATGACACCACTGA
FLG _R	TCCCCTACGCTTTCTTGTCCT
SCN9A_F	AGAGGGGTACACCTGTGTGAA
SCN9A _R	CCCAGGAAAATCACTACGACAAA
TNS4_F	AGCCAGGGGCTTTTGTCATAA
TNS4_R	AGACGACTCGATGAGGAAGTG
CXCL8_F	TTTTGCCAAGGAGTGCTAAAGA
CXCL8_R	AACCCTCTGCACCCAGTTTTC
ADCY1_F	AGGCACGACAATGTGAGCATC
ADCY1_R	TTCATCGAACTTGCCGAAGAG

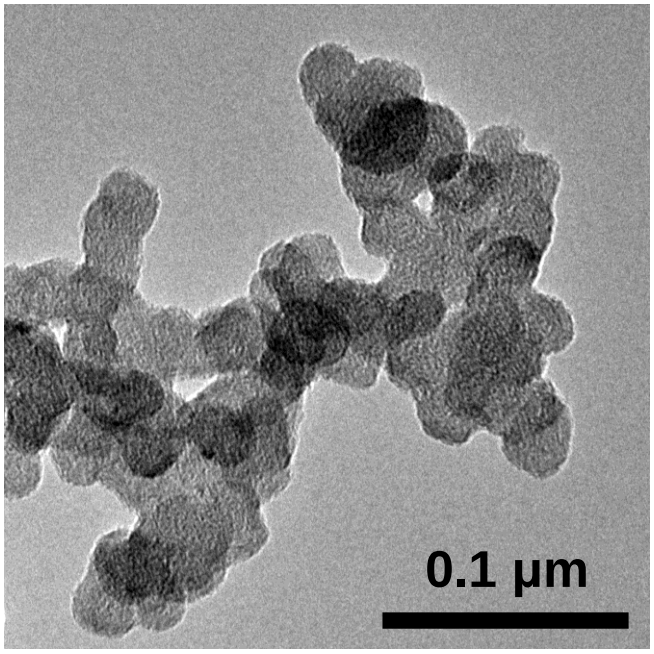

Figure S1. Representative TEM image of carbon black.

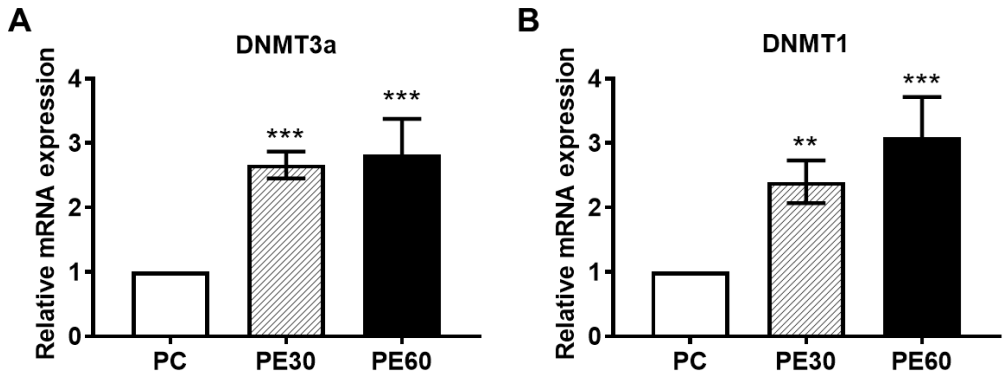

Figure S2. The expression of DNA methylase in PE cells. The mRNA expression levels of (A) DNMT3a and (B) DNMT1 in PE cells. $^{**}P < 0.01$ *vs*. PC cells; $^{***}P < 0.001$ *vs*. PC cells.

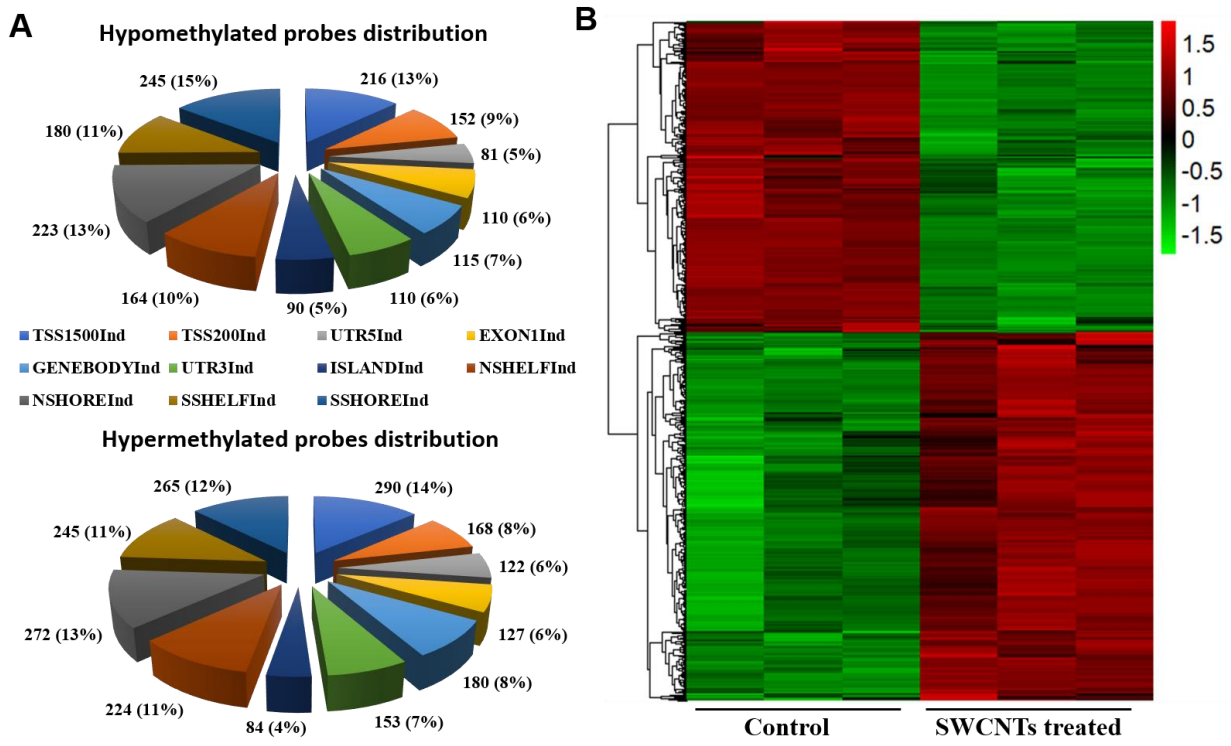

Figure S3. Differentially methylated probes and differentially expressed genes in PE60 cells. (A) Percentage of differentially methylated probes in PE60 cells *vs.* PC60 cells. (B) Heatmap of the 148 differentially expressed genes of PE60 cells. The colors in the heatmap represent the normalized expression values, with low expression values being colored in shades of green and high expression values in shades of red.

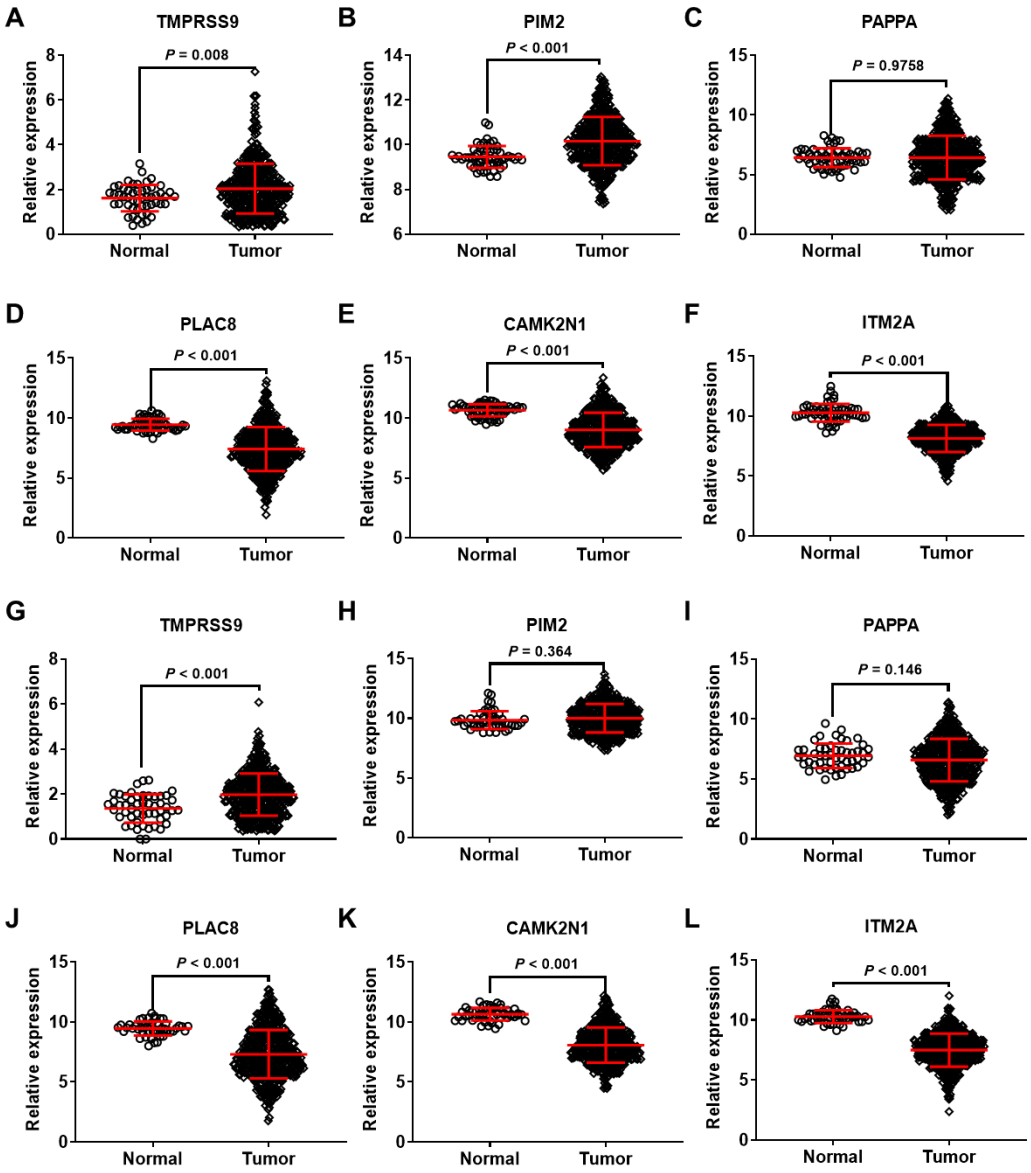

Figure S4. Expression of different expressed DMGs in lung cancer in TCGA database. Expression of TMPRSS9 (A), PIM2 (B), PAPPA (C), PLAC8 (D), and CAMK2N1 (E), as well as ITM2A (F) in TCGA LUAD dataset. Expression of TMPRSS9 (G), PIM2 (H), PAPPA (I), PLAC8 (J), and CAMK2N1 (K), as well as ITM2A (L) in TCGA LUSC dataset.