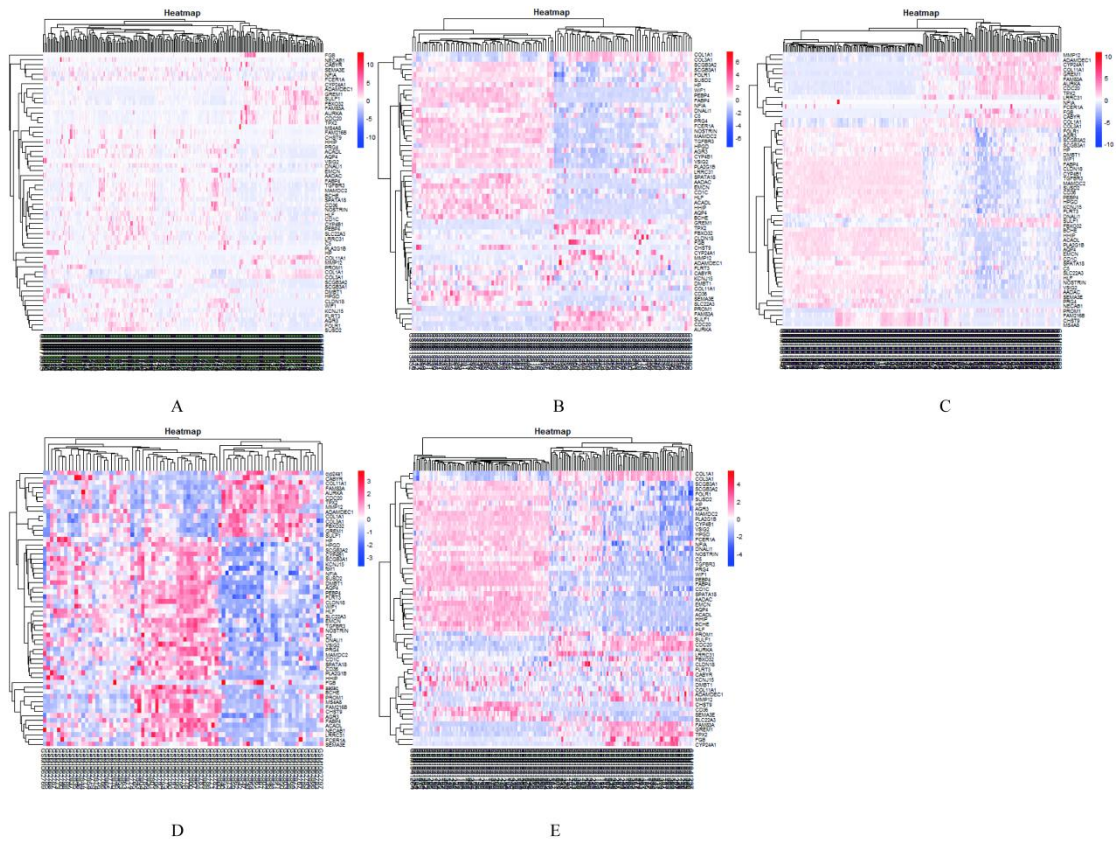


Supplementary files



Supplementary Figure S1. Two-dimensional hierarchical clustering of the significant and 58 DEGs in all samples. Genes are in rows; samples are in columns. Both up-regulated and down-regulated mRNAs can be seen in tumors compared with normal tissues in 3 lung adenocarcinoma datasets(B, C, E); Both up-regulated and down-regulated mRNAs can be seen in smokers and non-smokers in 2 smoking related lung adenocarcinoma datasets(A, D). **A:** GSE31210. **B:** GSE32863. **C:** GSE40791. **D:** GSE43458. **E:** GSE75037.

Supplementary Table S1. 58 DEGs in smoking related lung adenocarcinoma.

Gene symbol	Gene name
AADAC	arylacetamide deacetylase
ACADL	acyl-CoA dehydrogenase, long chain
ADAMDEC1	ADAM like decysin 1
AGR3	anterior gradient 3
AQP4	aquaporin 4
AURKA	aurora kinase A
BCHE	butyrylcholinesterase
C5	complement component 5
CABYR	calcium binding tyrosine phosphorylation regulated
CD1C	CD1c molecule
CD36	CD36 molecule (thrombospondin receptor)
CDC20	cell division cycle 20
CHST9	carbohydrate sulfotransferase 9
CLDN18	claudin 18
COL11A1	collagen type XI alpha 1 chain
COL1A1	collagen type I alpha 1 chain
COL3A1	collagen type III alpha 1 chain
CYP24A1	cytochrome P450, family 24, subfamily A, polypeptide 1
CYP4B1	cytochrome P450, family 4, subfamily B, polypeptide 1
DMBT1	deleted in malignant brain tumors 1
DNALI1	dynein axonemal light intermediate chain 1
EMCN	endomucin
FABP4	fatty acid binding protein 4, adipocyte
FAM216B	family with sequence similarity 216, member B
FAM83A	family with sequence similarity 83, member A
FBXO32	F-box protein 32
FCER1A	Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide
FGB	fibrinogen beta chain
FLRT3	fibronectin leucine rich transmembrane protein 3
FOLR1	folate receptor 1
GREM1	gremlin 1
HHIP	hedgehog interacting protein
HLF	hepatic leukemia factor
HP	haptoglobin
HPGD	hematopoietic prostaglandin D synthase

KCNJ15	potassium inwardly-rectifying channel, subfamily J, member 15
LRRC31	leucine rich repeat containing 31
MAMDC2	MAM domain containing 2
MMP12	matrix metalloproteinase 12
MS4A8	membrane-spanning 4-domains, subfamily A, member 8
NECAB1	N-terminal EF-hand calcium binding protein 1
NFIA	nuclear factor I/A
NOSTRIN	nitric oxide synthase trafficking
PEBP4	phosphatidylethanolamine-binding protein 4
PLA2G1B	phospholipase A2, group IB (pancreas)
PRG4	proteoglycan 4
PROM1	prominin 1
SCGB3A1	secretoglobin, family 3A, member 1
SCGB3A2	secretoglobin, family 3A, member 2
SEMA3E	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E
SLC22A3	solute carrier family 22 (organic cation transporter), member 3
SPATA18	spermatogenesis associated 18
SULF1	sulfatase 1
SUSD2	sushi domain containing 2
TGFBR3	transforming growth factor beta
TPX2	TPX2, microtubule-associated
VSIG2	V-set and immunoglobulin domain containing 2
WIF1	WNT inhibitory factor 1

Supplementary Table S2. Gene Ontology enrichment of 58 DEGs in smoking related lung adenocarcinoma.

Ontology	Term	P Value	Fold_Enrichment
Cellular component	GO:0005576~extracellular region	5.87E-07	3.840461402
Biological process	GO:0006898~receptor-mediated endocytosis	1.81E-05	12.39131351
Cellular component	GO:0009986~cell surface	3.20E-05	6.004217185
Cellular component	GO:0005615~extracellular space	3.33E-05	3.623926185
Biological process	GO:0030199~collagen fibril organization	2.11E-04	33.76973353
Cellular component	GO:0009897~external side of plasma membrane	4.50E-04	9.167002012
Biological process	GO:0030574~collagen catabolic process	9.11E-04	20.57843137
Cellular component	GO:0005581~collagen trimer	0.002701753	14.14906832
Cellular component	GO:0005887~integral component of plasma membrane	0.002956991	2.759818274
Cellular component	GO:0005578~proteinaceous extracellular matrix	0.008668498	6.071428571
Molecular function	GO:0005044~scavenger receptor activity	0.009237451	20.28966346
Biological process	GO:0043152~induction of bacterial agglutination	0.011858397	164.627451
Molecular function	GO:0005201~extracellular matrix structural constituent	0.017459197	14.5358783
Molecular function	GO:0003824~catalytic activity	0.019143607	6.907119476
Cellular component	GO:0005788~endoplasmic reticulum lumen	0.020229039	6.779761905
Biological process	GO:0030198~extracellular matrix organization	0.020540948	6.719487795

Cellular component	GO:0043203~axon hillock	0.02093908 8	92.97959184
Molecular function	GO:0030674~protein binding, bridging	0.02266811	12.6481019
Biological process	GO:0040036~regulation of fibroblast growth factor receptor signaling pathway	0.02357895 2	82.31372549
Biological process	GO:0007179~transforming growth factor beta receptor signaling pathway	0.03068162 2	10.73657289
Molecular function	GO:0048407~platelet-derived growth factor binding	0.03274470 1	59.02447552
Biological process	GO:0006955~immune response	0.03628767 5	3.91039076
Molecular function	GO:0030247~polysaccharide binding	0.04149110 6	46.37637363
Biological process	GO:0002063~chondrocyte development	0.04376295 3	43.90065359
Biological process	GO:0030168~platelet activation	0.04602387	8.589258312
Cellular component	GO:0071682~endocytic vesicle lumen	0.04722922 1	40.67857143
Molecular function	GO:0050431~transforming growth factor beta binding	0.04727891 5	40.57932692

Supplementary Table S3. KEGG pathway of 58 DEGs in smoking related lung adenocarcinoma.

Term	P Value	Fold Enrichment
ECM-receptor interaction	0.004065856	11.76670924
Platelet activation	0.012332612	7.874643875
PPAR signaling pathway	0.025914484	11.45936982
Protein digestion and absorption	0.042764316	8.724747475