

iScience, Volume 27

Supplemental information

**AutoCancer as an automated multimodal
framework for early cancer detection**

Linjing Liu, Ying Xiong, Zetian Zheng, Lei Huang, Jiangning Song, Qiuzhen Lin, Buzhou Tang, and Ka-Chun Wong

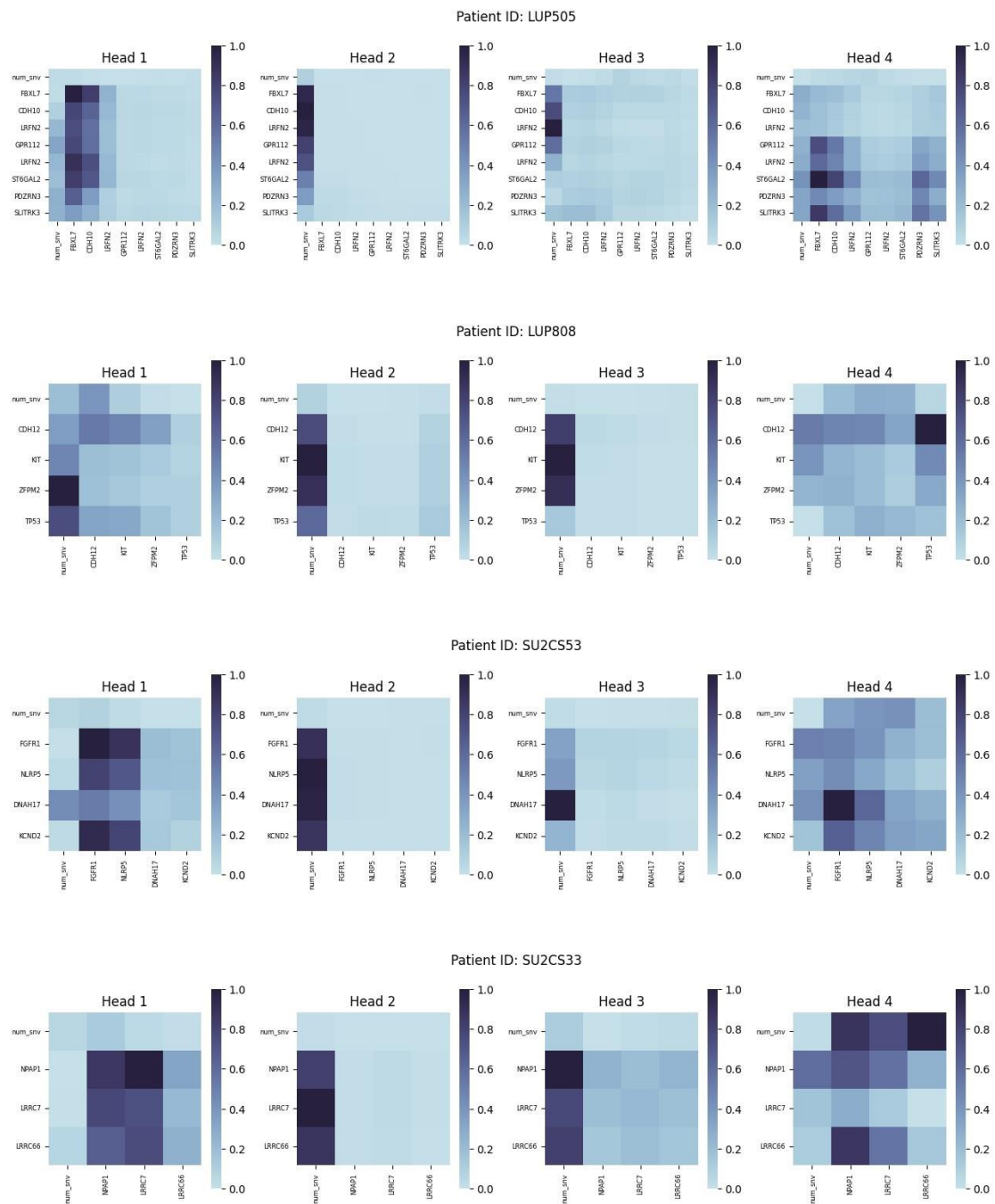


Figure S1. Visualization of attention matrix in each head, related to Figure 5.

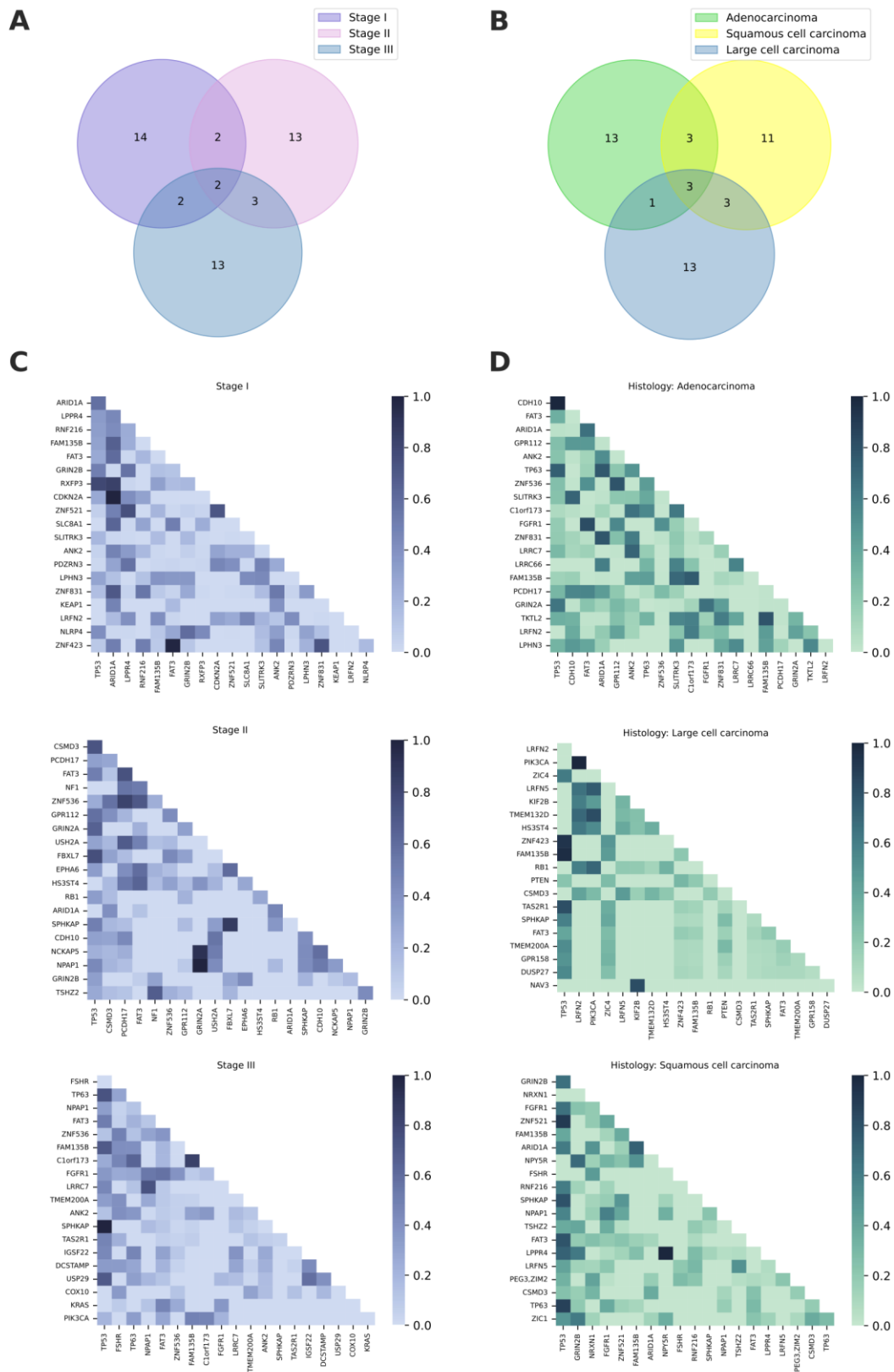


Figure S2. Top 20 focused genes vary across different stages and histology subtypes, related to Figure 5. (A) Overlap of top focused gene in different stages. (B) Overlap of top focused gene in different histology subtypes. (C) Attention matrix across top focused genes in different stages. (D) Attention matrix across top focused genes in different histology subtypes.

Table S2. Top genes unique to each stage, related to Figure 5.

Stage I	Stage II	Stage III
LPPR4	CSMD3	FSHR
RNF216	PCDH17	TP63
RXFP3	NF1	C1orf173
CDKN2A	GPR112	FGFR1
ZNF521	GRIN2A	LRRC7
SLC8A1	USH2A	TMEM200A
SLITRK3	FBXL7	TAS2R1
PDZRN3	EPHA6	IGSF22
LPHN3	HS3ST4	DCSTAMP
ZNF831	RB1	USP29
KEAP1	CDH10	COX10
LRFN2	NCKAP5	KRAS
NLRP4	TSHZ2	PIK3CA
ZNF423		

Table S3. Top genes unique to each subtype, related to Figure 5.

Adenocarcinoma	Large cell carcinoma	Squamous cell carcinoma
PCDH17	GRIN2B	DUSP27
ZNF831	TSHZ2	GPR158
GRIN2A	PEG3	NAV3
LRRC66	ZIM2	RB1
ZNF536	RNF216	TAS2R1
SLITRK3	LPPR4	TMEM132D
CDH10	NPY5R	TMEM200A
C1orf173	NPAP1	KIF2B
TKTL2	NRXN1	ZNF423
LRRC7	ZIC1	PTEN
LPHN3	FSHR	ZIC4
GPR112	ZNF521	PIK3CA
ANK2		HS3ST4

Table S4. Top 50 gene pairs associated with non-small cell lung cancer, related to Figure 6.

Gene1	Gene2	Weight
TP53	SPHKAP	5.002291
TP63	TP53	4.281027
NF1	ITGB3	3.460196
TP53	FAM135B	3.3944192
PCDH17	FAT3	3.322363
ST6GAL2	FBXL7	3.2492893
TP53	DUSP27	3.2138388
ZNF536	FAT3	3.0865207
PPP1R3A	LRRC66	3.000533
CDKN2A	ARID1A	2.9695902
SLC8A1	C1orf173	2.8343005
NPAP1	FGFR1	2.765613
TP53	CDH10	2.6875825
ZNF423	TP53	2.6446936
RXFP3	PTCHD4	2.6365385
NPY5R	LPPR4	2.6264887
SLITRK4	FAT3	2.600594
ZNF521	LPPR4	2.5760121
ZNF521	TP53	2.56786
NF1	GPR112	2.5040803
RNF216	ARID1A	2.5034041
ZNF536	DUSP27	2.4904222
SLITRK4	KRAS	2.432415
NPAP1	GRIN2A	2.4271429
TP53	TAS2R1	2.3864439
RNF216	KCNJ12	2.3534539
TP53	LPPR4	2.3361993
LPPR4	HTR5A	2.315725
NCKAP5	GRIN2A	2.296194
LRRC4C	DARC	2.290184
TP53	HTR5A	2.2640498
TP53	NPAP1	2.2537193
RNF216	LRRC66	2.2354383
SLC8A1	NF1	2.2003665
RIMS2	LRRC7	2.1979532
KLHL6	CUL3	2.1748445
RB1	GRM8	2.172481

TP53	GPR158	2.1604207
RP1	KIT	2.1481342
ZNF521	CDKN2A	2.1349797
TKTL2	FAM135B	2.1266603
TBX22	LPPR4	2.1199358
PEG3, ZIM2	LRRC7	2.114758
SPHKAP	FBXL7	2.106174
TP53	CDH12	2.0939147
TMEM200A	TLR4	2.0883563
ZNF423	PTCHD4	2.0759187
USP29	IGSF22	2.072718
TP63	ARID1A	2.064836
ZNF831	ANK2	2.0594206

Table S5. Top 20 gene pairs across different stages, related to Figure 6.

Stage I		
Gene1	Gene2	Weight
NLRP5	FGFR1	3.2188225
PPP1R3A	LRRC66	3.000533
ZNF423	FAT3	2.93407
NF1	GRIN2B	2.6792438
RXFP3	PTCHD4	2.6365385
NLRP8	GLCC11	2.561884
PTEN	HTR5A	2.4076223
RNF216	KCNJ12	2.3534539
LRRC4C	DARC	2.290184
ZNF521	LPPR4	2.2273345
ZNF831	ZNF423	2.1767464
RB1	GRM8	2.172481
RP1	KIT	2.1481342
TBX22	LPPR4	2.1199358
ZNF80	USP29	2.0968084
ZNF423	PTCHD4	2.0759187
CHAT,SLC18A3	CDH7	2.067424
THSD7B	GLCC11	2.057376
ZNF423	EPHA10	2.0379512
NLRP5	KLHL6	2.0228753

Stage II		
Gene1	Gene2	Weight
NF1	ITGB3	3.460196
RIMS2	FAM71B	2.3495607
ST6GALNAC5	PIK3CG	2.2404597
RIMS2	LRRC7	2.1979532
SPHKAP	FBXL7	2.106174
TMEM200A	TLR4	2.0588264
ZNF536	PCDH17	2.0050983
TSHZ2	SLC8A1	1.9921732
KLHL6	ANK2	1.9284734
PCDH17	AKT1	1.9082471
EPHA10	ASTN1	1.8331211
TP63	RB1	1.7755973
GRIN2A	FGFR1	1.7673779
ZNF521	RB1	1.734963
SLC8A1	KEAP1	1.7137923
ST6GAL2	RIT1	1.6830616
TMEM132D	PIK3CA	1.6815755
GRIN2A	CXCL16	1.6730747
TSHZ2	NF1	1.656126
ZNF831	ZIM3	1.6461878

Stage III		
Gene1	Gene2	Weight
TP53	SPHKAP	3.5790436
TAS2R1	NLRP12	2.9556944
TP63	TP53	2.6376
TP53	FAM135B	2.4896145
NF1	GPR112	2.4237237
PPP1R3A	NLRP8	2.167505
NPAP1	LRRRC66	2.1150541
TP53	GPR158	2.0772412
FGFR1	FAT3	2.0228293
ZNF423	TP53	1.913132
ZFPM2	TP53	1.8626779
GPR139	CHAT,SLC18A3	1.7379782
PIK3CA	FAM135B	1.7266072
SLITRK3	C1orf173	1.6878777
PIK3CA	C1orf173	1.6648972
TP63	IGSF22	1.6589303
PIK3CA	MMP16	1.647397
TP53	IGSF22	1.5960064
TP53	MMP16	1.5473424
LRRRC7	LRRRC66	1.5400691

Table S6. Top 20 gene pairs across different subtypes, related to Figure 6.

Adenocarcinoma		
Gene1	Gene2	Weight
TP53	NCKAP5	4.594971
NLRP5	FGFR1	3.2188225
PPP1R3A	LRRC66	3.000533
ZNF423	FAT3	2.93407
TP53	CDH10	2.6875825
RXFP3	MAGEL2	2.6447456
RXFP3	PTCHD4	2.6365385
NLRP8	GLCCI1	2.561884
RNF216	ARID1A	2.5034041
NPAP1	GRIN2A	2.4271429
ZNF536	DUSP27	2.361931
KEAP1	KCNA4	2.33498
RB1	ARID1A	2.2926912
PDGFRA	ERBB4	2.2722187
RNF216	LRRC66	2.2354383
RIMS2	LRRC7	2.1979532
ZNF831	ZNF423	2.1767464
RB1	GRM8	2.172481
SLITRK3	PPP1R3A	2.1534142
TKTL2	FAM135B	2.1266603

Large cell carcinoma		
Gene1	Gene2	Weight
PIK3CA	LRFN2	2.0425072
ZNF423	TP53	1.913132
NAV3	KIF2B	1.6688011
PIK3CA	LRFN5	1.5455643
TMEM132D	LRFN2	1.4265536
LRFN2	HS3ST4	1.2747283
LRFN2	KIF2B	1.2248298
TP53	SPHKAP	1.1959748
PIK3CA	HS3ST4	1.0939729
TP53	GPR158	1.0075513
ZNF423	ZIC4	0.9846345
LRFN5	CSMD3	0.9403834
PIK3CA	KIF2B	0.9347547
LRFN2	CSMD3	0.8521757
PIK3CA	TMEM132D	0.8232899
TAS2R1	TP53	0.8149611
LRFN5	HS3ST4	0.7868407
ZIC4	TAS2R1	0.7427185
PIK3CA	RB1	0.7098927
RB1	HS3ST4	0.6881671

Squamous cell carcinoma		
Gene1	Gene2	Weight
TP53	SLITRK4	3.3913813
RXFP3	ARID1A	2.4295762
RIMS2	FAM71B	2.3495607
TP53	RXFP3	2.2045038
RP1	KIT	2.1481342
TP53	SPHKAP	2.0849867
LRRC66	BRAT1	2.068101
FAM135B	ALK	2.064527
CDKN2A	ARID1A	1.9739954
GRIN2B	CTNND2	1.9530003
NPY5R	CUL3	1.9203162
RIT1	EPHA6	1.9117112
NPY5R	GRIN2B	1.8170743
TP53	GRIN2B	1.7383668
PIK3CA	FAM135B	1.7266072
ST6GAL2	RIT1	1.6830616
TSHZ2	NF1	1.656126
LPPR4	GRIN2B	1.6214553
TP53	ARID1A	1.6032948
TP53	RNF216	1.5940902