

Supplementary Figure 6.

A. Lung Adenocarcinoma

GENIE Cohort v16.1

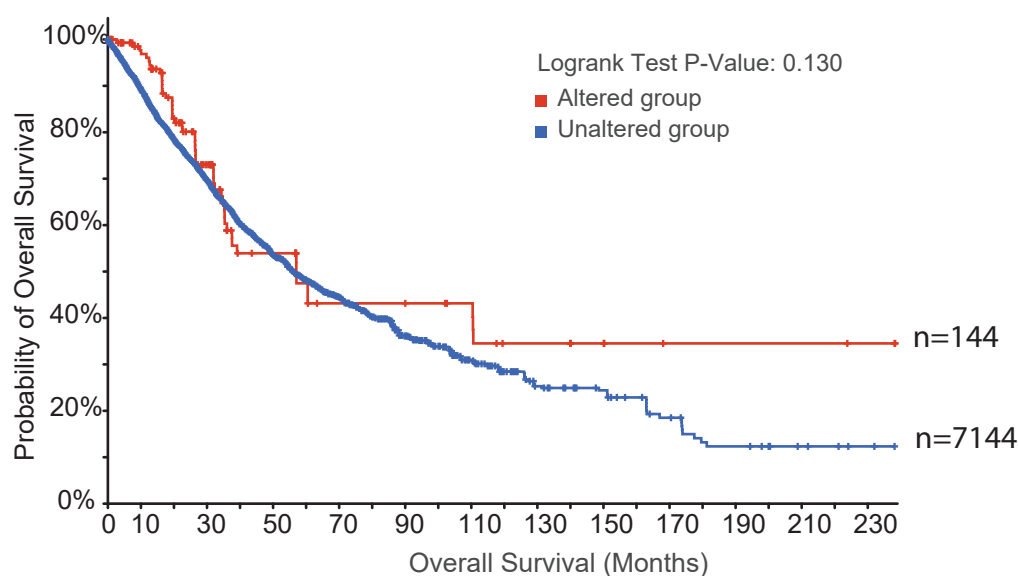
Patient ID	EpCAM	Driver gene mutations
GENIE-PROV-fb936feb0e-40fc12cacf	P178L	TP53(R273H),PIK3CG(A557V),KEAP1(G332C)
GENIE-MSK-P-0075350-T02-IM7	R126K	KRAS(G12L), CTNNB1(E101D),MTOR(E1601K)
GENIE-DFCI-200207-2455865*	E304K	EGFR(L858R), TP53(R273S), MET(Amp)
GENIE-MSK-P-0011142-T01-IM5	L264P	KRAS(G13D), TP53(A159P), NOTCH2(G1521W)
GENIE-DFCI-148911-1525863*	Q204E	BRAF(G464V),STK11(G242V),TP53(G245F)
GENIE-MSK-P-0054440-T02-IM6	Y174C	PRKCI(X473_splice), PTPRT(V609A)
GENIE-MSK-P-0016180-T01-IM6	A278V	PIK3CA(H1047R), NF1(K2428*), TP53(E339*)
GENIE-MSK-P-0093789-T01-IM7	N91S	TP53(G266L),WT1(R380Q),NOTCH3(S1817L)
GENIE-DFCI-037706-158964	Y215N	ATM(R2832H), TP53(E285*), CDKN2A(E33*)
GENIE-DFCI-052981-218952	D100G	KRAS(G12C), PTEN(H93R), TP53(M237I)
GENIE-MSK-P-0018327-T01-IM6	Q167L	BRAF(V600E),TP53(Q104*), DNMT3A(R688L)
GENIE-MSK-P-0019323-T01-IM6	Q246H	BRAF(G469V), TP53(X225_splice), ERBB4(V87M)
GENIE-MSK-P-0020779-T01-IM6	L148V	KRAS(G12D),NOTCH2(A2178S),PIK3C2G(S526R)
GENIE-MSK-P-0021387-T01-IM6*	L195V	TP53(H179Tfs*7), AURKA(R137L), TERT(R428L)
GENIE-PROV-b6f3feb09c-7b90e15fc3	I272T	KRAS(G12S),TP53(D208V),ABL2(R974*)
GENIE-UCSF-11973-8812T*	E223Q	NRAS(Q61K),RAC1(D11E),ERBB2(Q398P),AXL(K567N)
GENIE-MSK-P-0026043-T01-IM6	Q246E	MET(H1094Y), TP53(R248L), DNMT3A(D702N)
GENIE-MSK-P-0046101-T01-IM6	M73I	TP53(E326*), ERBB4(L1259V), FOXL2(M149I)
GENIE-MSK-P-0032035-T01-IM6	V212M	KRAS(G12D),TP53(Y205F),KEAP1(N414I),NOTCH1(V47F)
GENIE-DFCI-001194-3170353	R173H	TP53(E68*), PRKDC(K1333*),CBL(P684S),RHOH(S152L)
GENIE-MSK-P-0057134-T01-IM6	E301K	TP53(E51*),DUSP4(R96C),SMARCA4(D1188Y)
GENIE-DFCI-106474-805367	I288V	KRAS(G12C),DNMT3A(R882H),TP53(R249S)
GENIE-PROV-c262c39a54-b95d471997	R163L	MAP2K2(R313W),ATM(E1139D),CSF1R(A123G)
GENIE-MSK-P-0033766-T01-IM6	A87S	TP53(H179Y), ATM(A244T),FOXP1(S322*)
GENIE-PROV-fba14a2091-f0188cb93f	Q201E	STK11(Q100*), CTNNB1(S37C), TP53(T150Hfs*20)
GENIE-DFCI-124025-1251941	I209V	TP53(N268I),RASA1(E763*),RICTOR(G71S),GATA3(A268S)
GENIE-DFCI-129774-1200816	K83E	KRAS(T50I),TP53(R249M),ALK(S691R),ERBB3(E925Q)
GENIE-DFCI-129774-1200816	D92V	KRAS(G12C),STK11(X199_splice),IKBKE(L269M)
GENIE-MSK-P-0041703-T01-IM6*	I181V	KRAS(G12D),ATR(G1948V),SMAD4(D103N)
GENIE-DFCI-084941-298944	I193V	PIK3CA(E726K),MSH6(Q4*),NSD2(G67S)
GENIE-DFCI-183356-3502543	Q262E	KRAS(G12V),FAT1(R3233C),TFE3(R134Q)
GENIE-DFCI-216708-3103272	V268F	TP53(V173L),PTPN11(L30F),NF1(X775_splice)
GENIE-UCSF-230021-3307T	G112D	EGFR(L858R), RET(R721W),ALK(T1087I)
GENIE-MSK-P-0048098-T02-IM6	D158Y	STK11(D53Gfs*12),KEAP1(G480W),JAK2(R215*)
GENIE-UCSF-10655-3862T	Q204H	TP53(E286Q),PTPRB(R874*),POLE(A2239V)
GENIE-UCSF-10736-6951T*	L286Q	KRAS(G12V),ATM(Y2437S),PTPRB(E1109D)
GENIE-UCSF-10977-2282T	E298V	TP53(N247I),SMAD3(R80W),ATR(Q1877E)
GENIE-DFCI-269017-3531071*	D96N	KRAS(G12D),TP53(Y163N),BRAF(G315*)
GENIE-MSK-P-0060172-T02-IM7*	V220A	EGFR(E746_A750del),PIK3CA(H1047R)
GENIE-DFCI-387986-4648988	S159N	ABCB11(M805I),KIF1B(L700del)
GENIE-DFCI-386840-4384366	M261T	BRAF(V600E),TP53(Q167*),DNMT3A(W330*)
GENIE-MSK-P-0064854-T01-IM7	S136F	KRAS(G12V),CTNNB1(S33F),PIK3CG(E740D)
GENIE-UCSF-14632-9666T	G237W	EGFR(L858R),PIK3CA(H1047R),PDGFRA(G132V)
GENIE-PROV-ae3f1d16fe-66eae2b26	S224C S228F	ATR(L2076V),BCR(I1124V),PTCH1(R1391W)
GENIE-PROV-5408feb7fb-c763276b13	N189Y	TP53(I251M),PIK3C2B(R1269H),DNMT3A(X803_splice)
GENIE-PROV-77be8dc9d9-e2f89c4af8	G112S	KEAP1(Y491*),KIF5B(Q794E),PAK3(D518Y)
GENIE-MSK-P-0068704-T01-IM7*	G245V	EGFR(T790M, L858R), ERBB2(S310F),TGFB1(M441I)
GENIE-PROV-f0235d7917-8cf0eda65c	V276L	KRAS(G12C),SMAD4(C363W),ERBB2(V308L),EGFR(K713T)
GENIE-PROV-08026afdad-9e829531c7	Y95H	TP53(R248L),KEAP1(G186S),TGFB2(N144S)
GENIE-UHN-OCT888181-ARC1	T172M	KRAS(G12C),KEAP1(S275*),SMAD4(D493H)
GENIE-MSK-P-0079689-T01-IM7	V275A	TP53(X225_splice),STK11(E223*),MSH6(R495*)

Supplementary Figure 6.

B. Lung Squamous Cell Carcinoma

Patient ID	EpCAM	Driver gene mutations
GENIE-MSK-P-0041477-T01-IM6	E147*	TP53(Y163C), PIK3CA(R617W), JAK1(P209L)
GENIE-PROV-c48e0e3e23-bd35a7bc78	A165T	PIK3CA(E542K), TP53(S127F), CEBPA(T337S)
GENIE-PROV-276782d53f-cf5d341df1	Q246H	TP53(R175G), SMAD4(X436_splice), HGF(P180S)
GENIE-PROV-9d17b4b1e8-5825c59136	H150Q	TP53(M246V), PTEN(M35I), SMAD4(K50*)
GENIE-PROV-aacf5ba00e-fa980c2072	R173H	TP53(Q136*), FGFR2(P303R), SOX9(G305A)
GENIE-MSK-P-0041617-T01-IM6	A16E	KRAS(L19F), PIK3CA(E545K), TP53(L145Q)
GENIE-UCSF-13402-0998T	I193V	TP53(R342*), KIT(V654M), ERBB4(C189*)
GENIE-PROV-425b36f22f-86ee4cc590	T17M	TP53(R248P), ALK(P117R), INSR(D483N)
GENIE-UCSF-783433-75291T	Y32N	BRAF(V600E), TP53(R342*), FGFR2(L770V)
GENIE-PROV-ea08758a92-14894cdd37	A18P	TP53(R342*), KEAP1(G379V), ALK(Y1359C)
GENIE-DFCI-449876-4541134	G103W	TP53(E298*), ATM(L2519V), APC(P2392S)
GENIE-MSK-P-0015934-T03-IM7	E217K	TP53(R249M), PIK3CA(C378R), NF1(L2348F)
GENIE-PROV-7fcfb456a3-f0ad24f020	E25V	TP53(X224_splice), RET(K666N), PIK3CD(R641L)
GENIE-DFCI-577479-5191985	R173L	HRAS(G13V), PTEN(T319*), BRCA2(X2985_splice)
GENIE-UHN-OCT651159-ARC1	A2V	TP53(R273L), NFE2L2(E79K), TPR(L742F)
GENIE-PROV-80f21bed8e-971db6d694	M305I	TP53(P278S), PIK3CA(P471A), FAT1(K1801*)
GENIE-MSK-P-0082884-T01-IM7	E101V	TP53(G334V), RB1(S618*), ERBB4(C1106F)
GENIE-PROV-72ea2b3530-7b9ad68680	C27Y	TP53(X10_splice), CDKN2A(D108Y), PRKDC(L3450*)

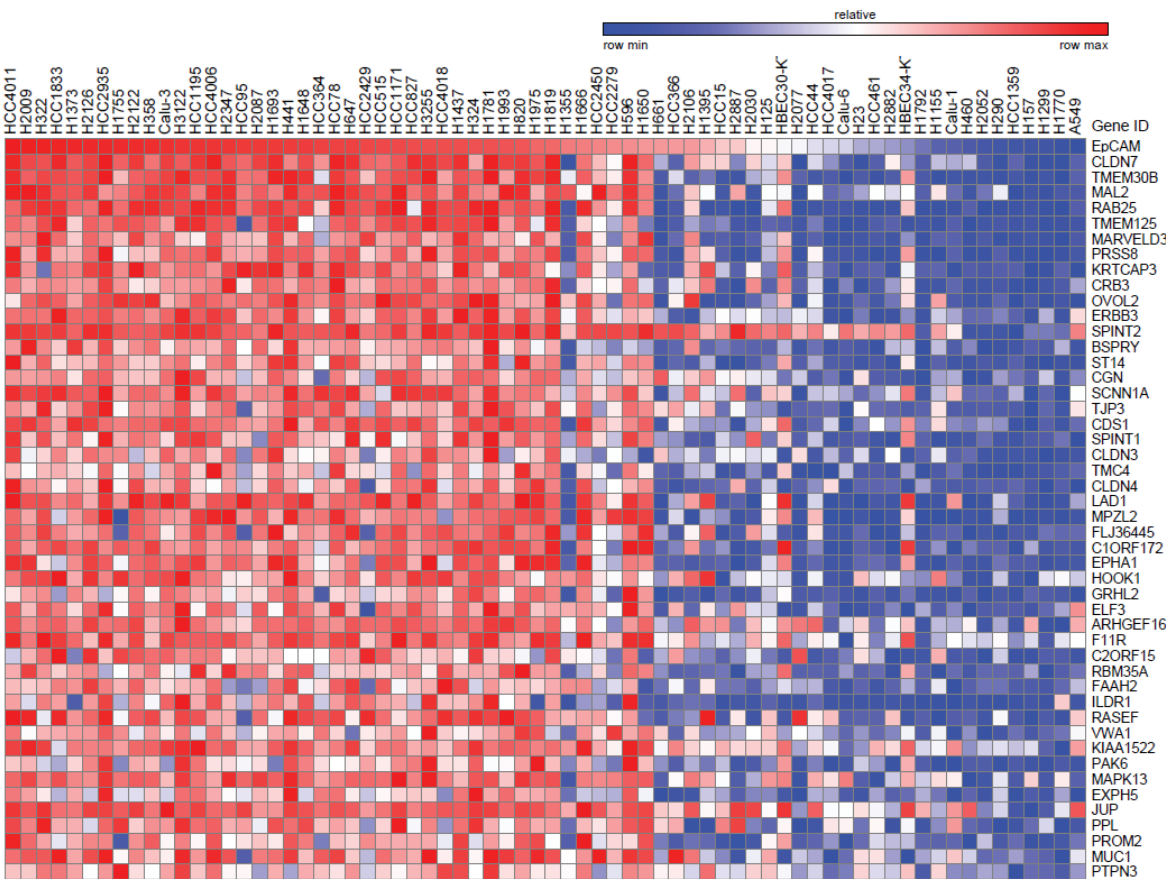
C.



Altered group: Median month overall survival: 57.03

Unaltered group: Median month overall survival: 56.21

D. 76 EMT gene signature (NSCLC)



Supplementary Figure 6. Somatic mutation of EpCAM in lung cancer, (A) Mutations in adenocarcinoma, **(B)** squamous cell carcinoma. **(C)** Kaplan-Meier survival plot. **(D) 1.** 76-EMT gene signature, which designates cell types epithelial or mesenchymal, was used on dataset GSE32989 of 70 NSCLC lines. Heat plot shows that EpCAM expression is restricted to epithelial cell.