

Table S1. GSEA report of HCMDB gene-set enrichment in KPK v. KP cells, Related to Figure 1.

NAME	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX	LEADING EDGE
HCMDB	1826	0.35	1.73	0.00	0.08	0.00	2515.00	tags=21%, list=15%, signal=22%

Table S2. GSEA report of C3:TFT gene-set (MSigDB V6.2) enrichment in KPK v. KP cells, Related to Figure 1.

NAME	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX	LEADING EDGE
BACH1_01	243	0.37	1.74	0.00	0.11	0.10	2075	tags=18%, list=12%, signal=20%
ELK1_01	244	0.28	1.73	0.00	0.08	0.14	2412	tags=13%, list=14%, signal=15%
TNCATNTCCYR_UNKNOWN	120	0.30	1.71	0.00	0.07	0.19	3356	tags=22%, list=20%, signal=27%
AP1_Q4	250	0.38	1.65	0.00	0.17	0.29	2790	tags=21%, list=16%, signal=25%
SF1_Q6	231	0.34	1.64	0.00	0.15	0.29	1986	tags=15%, list=12%, signal=17%
HNF4_Q6	242	0.39	1.58	0.00	0.26	0.35	3095	tags=27%, list=18%, signal=33%
NRF2_Q4	229	0.37	1.57	0.00	0.27	0.35	3369	tags=24%, list=20%, signal=30%
AP1_Q6	236	0.36	1.55	0.00	0.30	0.41	2854	tags=22%, list=17%, signal=27%
RGTTAMWNATT_HNF1_01	67	0.51	1.54	0.00	0.28	0.41	1546	tags=28%, list=9%, signal=31%
PEA3_Q6	237	0.41	1.53	0.00	0.30	0.45	3326	tags=30%, list=20%, signal=37%
TGASTMAGC_NFE2_01	173	0.36	1.53	0.00	0.28	0.45	2422	tags=18%, list=14%, signal=21%
NFE2_01	245	0.38	1.50	0.00	0.35	0.56	3904	tags=29%, list=23%, signal=37%
TGANNYRGCA_TCF11MAFG_01	270	0.33	1.50	0.00	0.33	0.56	2263	tags=17%, list=13%, signal=19%
BACH2_01	247	0.32	1.50	0.00	0.32	0.56	1743	tags=14%, list=10%, signal=15%
TTCNRGNNTTC_HSF_Q6	138	0.33	1.50	0.00	0.30	0.56	1813	tags=14%, list=11%, signal=15%
AP1_Q6_01	246	0.36	1.49	0.00	0.29	0.56	1794	tags=15%, list=11%, signal=17%
NFKAPPAB_01	238	0.38	1.48	0.00	0.29	0.60	2107	tags=21%, list=12%, signal=24%
AP1FJ_Q2	248	0.37	1.48	0.00	0.28	0.60	2790	tags=21%, list=16%, signal=24%
AP1_01	238	0.34	1.47	0.00	0.29	0.60	2637	tags=21%, list=16%, signal=24%
TEL2_Q6	218	0.23	1.46	0.00	0.29	0.60	2931	tags=17%, list=17%, signal=20%
WCAANNYCAG_UNKNOWN	229	0.26	1.46	0.00	0.28	0.60	2837	tags=19%, list=17%, signal=22%
CCAWYNGGAAR_UNKNOWN	126	0.29	1.45	0.00	0.29	0.60	2850	tags=20%, list=17%, signal=24%
PPAR_DR1_Q2	240	0.28	1.43	0.00	0.32	0.60	2412	tags=15%, list=14%, signal=18%
NFKB_C	250	0.33	1.42	0.00	0.36	0.66	2707	tags=22%, list=16%, signal=26%
CEBPB_02	240	0.35	1.42	0.09	0.35	0.66	3208	tags=28%, list=19%, signal=34%
USF2_Q6	234	0.28	1.41	0.00	0.35	0.66	4351	tags=29%, list=26%, signal=38%
RGAGGAARY_PU1_Q6	466	0.31	1.41	0.09	0.36	0.71	3144	tags=26%, list=19%, signal=30%
HNF4_01	246	0.35	1.40	0.00	0.35	0.71	2319	tags=18%, list=14%, signal=20%
TGACCTTG_SF1_Q6	222	0.32	1.40	0.10	0.35	0.71	2901	tags=21%, list=17%, signal=25%
STTCRNTT_IRF_Q6	178	0.32	1.39	0.00	0.37	0.75	2624	tags=22%, list=16%, signal=26%
HNF4_DR1_Q3	238	0.29	1.38	0.00	0.36	0.75	3354	tags=21%, list=20%, signal=26%
NFKAPPAB65_01	227	0.33	1.38	0.10	0.36	0.75	2027	tags=18%, list=12%, signal=20%
ETS_Q4	234	0.25	1.37	0.00	0.38	0.75	3249	tags=22%, list=19%, signal=27%
HNF1_01	221	0.41	1.37	0.00	0.37	0.75	2708	tags=29%, list=16%, signal=33%
YGACNNYACAR_UNKNOWN	86	0.39	1.36	0.00	0.38	0.79	3157	tags=28%, list=19%, signal=34%
AP1_Q4_01	243	0.31	1.36	0.10	0.37	0.79	2500	tags=21%, list=15%, signal=24%
HNF1_Q6	231	0.40	1.36	0.00	0.36	0.79	2141	tags=23%, list=13%, signal=26%
ARNT_01	238	0.23	1.36	0.00	0.36	0.79	3459	tags=18%, list=20%, signal=22%
PPARA_01	34	0.44	1.35	0.00	0.36	0.79	1630	tags=18%, list=10%, signal=19%
TTGCWCAAY_CEBPB_02	59	0.41	1.35	0.19	0.36	0.79	3210	tags=32%, list=19%, signal=40%
NFKB_Q6_01	219	0.36	1.34	0.00	0.36	0.79	1992	tags=18%, list=12%, signal=20%
AP1_C	255	0.33	1.34	0.00	0.36	0.79	2075	tags=17%, list=12%, signal=19%
LMO2COM_01	237	0.33	1.34	0.10	0.35	0.79	2368	tags=19%, list=14%, signal=21%
PPARA_02	118	0.33	1.34	0.20	0.35	0.79	4325	tags=32%, list=26%, signal=43%
MEIS1_01	228	0.33	1.33	0.09	0.36	0.79	2730	tags=20%, list=16%, signal=23%
TGGNNNNNKCCAR_UNKNOWN	386	0.33	1.33	0.10	0.36	0.79	3308	tags=28%, list=20%, signal=34%
CREL_01	247	0.31	1.33	0.09	0.35	0.79	1882	tags=16%, list=11%, signal=18%
AREB6_03	237	0.32	1.33	0.10	0.36	0.79	2019	tags=18%, list=12%, signal=20%
STAT1_03	223	0.27	1.32	0.09	0.36	0.85	2568	tags=16%, list=15%, signal=18%
AREB6_02	232	0.36	1.32	0.09	0.37	0.85	2792	tags=25%, list=16%, signal=29%
WGGAATGY_TEF1_Q6	337	0.31	1.31	0.00	0.37	0.85	2856	tags=23%, list=17%, signal=27%
RTTTNNNYTGGM_UNKNOWN	144	0.33	1.30	0.09	0.37	0.85	2100	tags=24%, list=12%, signal=27%
NF1_Q6	243	0.29	1.30	0.09	0.37	0.85	2195	tags=18%, list=13%, signal=21%
PAX2_02	236	0.34	1.30	0.09	0.36	0.85	3308	tags=26%, list=20%, signal=32%
TCF11MAFG_01	189	0.34	1.30	0.09	0.36	0.85	2188	tags=16%, list=13%, signal=18%
STAT5A_02	131	0.33	1.30	0.09	0.35	0.85	2801	tags=23%, list=17%, signal=27%
P300_01	229	0.28	1.30	0.09	0.35	0.85	2144	tags=16%, list=13%, signal=18%
AP1_Q2	242	0.34	1.30	0.00	0.35	0.85	2790	tags=21%, list=16%, signal=25%
YKACATTT_UNKNOWN	253	0.31	1.30	0.09	0.35	0.85	3492	tags=29%, list=21%, signal=36%

Table S4. GSEA report of Bach1 signature gene-set enrichment in KP siFbxo22 v. KP siCtrl cells, Related to Figure 4.

NAME	GS follow link to MSigDB	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX	LEADING EDGE
GSEA_BACH1_integrated_signature	GSEA_BACH1_integrated_signature	287	0.43	1.58	0.00	0.09	0.00	3104.00	tags=30%, list=17%, signal=35%

Table S6. Univariate and Multivariate Cox regression analyses for overall survival in the TCGA lung adenocarcinoma patient cohort, Related to Figure 7.

Description: Results of univariate and multivariable Cox proportional hazards model on overall survival in the TCGA LUAD cohort (all patients).

Characteristic	Univariate		Multivariable		P _{interaction}
	HR (95% CI)	p	HR (95% CI)	p	
Bach1 MSigDB signature (Strong vs weak correlation)	1.25(1.07-1.46)	0.00396	1.39(1.13-1.70)	0.00127	
Gender (Male vs Female)	1.08(0.80-1.46)	0.602	1.11(0.78-1.58)	0.53141	
Age (Years)	1.00(0.98-1.02)	0.576	1.01(0.99-1.03)	0.11725	
TNM Stage (Stage III/IV vs I/II)	2.67(1.95-3.67)	9.38E-10	1.18(0.75-1.86)	0.45328	
T score (T3/T4 vs T1/T2)	2.43(1.65-3.57)	5.68E-06	1.87(1.19-2.94)	0.0061	0.8934
N score (N1/N2 vs N0)	2.52(1.86-3.41)	1.96E-09	2.42(1.60-3.66)	2.56E-05	0.909
Smoking History (reformed > 15yrs vs non-smoker)	0.88(0.52-1.48)	0.647	0.89(0.50-1.60)	0.72014	
Smoking History (reformed <= 15yrs vs non-smoker)	1.09(0.68-1.76)	0.693	1.17(0.68-2.01)	0.56499	
Smoking History (current smoker vs non-smoker)	0.84(0.50-1.42)	0.526	0.87(0.48-1.58)	0.66912	
Tumor Purity (Frac. cancer nuclei)	1.16(1.00-1.34)	0.0376	1.37(1.129-1.66)	0.00141	0.7051

HR = Hazard ratio; CI = Confidence Interval; TNM Stage = Stage classification per Union for International Cancer Control (UICC);

T score = Primary tumor size/invasiveness; N score = Lymph node metastasis; Tumor Purity = From Campbell et al., Nat. Genet. 2016;48(6) using ABSOLUTE [Carter et al., Nat. Biotechnol. 2012;30(5)]

P_{interaction} = p-value of interaction between Bach1 signature and other significant covariates (model comparison; likelihood ratio test)

Table S7. Primer Sequences, Related to STAR methods section.

Description: oligonucleotides table

qPCR oligos		
Gene Name	Forward Sequence	Reverse Sequence
Bach1	GAGCCTTGCCCGTATGCTTG	TGGCAGCTTCACCTCACAGT
Hmox1	ACAGAGGAACACAAAGACCAG	GTGTCTGGGATGAGCTAGTG
Gapdh	CTTTGTCAAGCTCATTTCCTGG	TCTTGCTCAGTGCCTTG

CRISPR gRNA oligos		
gRNA Target	Forward Sequence	Reverse Sequence
Fbxo22 gRNA#1	CACCGCGTGCTGAGTAACCTTGCGG	AAACCCGCAAGTTACTCAGCACGC
Fbxo22 gRNA#2	CACCGGTGCTTCGCGGATCTGTGG	AAACCCACAGATCCGCGAAGCACC
Bach1 gRNA#1	CACCGCCAGGACAGTCACATCACAC	AAACGTGTGATGTACTGTCTGGC
Bach1 gRNA#2	CACCGTCTGGCCTACGATTCTCGAG	AAACCTCGAGAATCGTAGGCCAGAC
Ho1 gRNA #1	CACCGTCAGGACCTGACCCCTGAG	AAACCTCAGGGGTCAGGTCCTGAC
Ho1 gRNA #2	CACCGCTGTGTGGCTGGCGTGCAA	AAACTTGACAGCCAGCCACACAGC

*G for U6 promoter transcription initiation
*Bpil (BbsI) overhang

Genotyping primers		
Name	Forward Sequence	Reverse Sequence
Bach1 CRISPR Genotype	CGGAGAGGAACGGAGCGC	CCCGAAAGGACGGAAGGACG
Fbxo22 CRISPR Genotype	GGCGAGGAATGGAGCCGCGAG	CCCGGCTACCGAAGCAGCGC

siRNA sequences		
mRNA Target	Cat.No.	Sequence
Human FBXO22	#1_DHARMACON, J-010812-06-0005	GCACCUUCGUGUUGAGUAA
	#2_DHARMACON, J-010812-07-0005	GGUGGGAGCCAGUAAUUAU
	#3_DHARMACON, J-010812-08-0005	GUUCGCAUCUUACCACAUUA
	#4_DHARMACON, J-010812-09-0005	GCUUAUGGAGGGAGUGUGU
Mouse FBXO22	#1_DHARMACON, J-055384-09-0005	GUCAGCACUUUCAGCGAUA
	#2_DHARMACON, J-055384-10-0005	GGGAUCAGGUAGCAAUCGA
	#3_DHARMACON, J-055384-11-0005	GAAUUUAACACUCGAAAGA
	#4_DHARMACON, J-055384-12-0005	GGAGAUGUAAUGAGGUAAA

Ultramers sequences	
mRNA Target	Sequence
Mouse Bach1 #2	TGCTGTTGACAGTGAGCGATCCACCTTGCTAGACTGTTATAGTGAAGCCACAGATGTATAACAGTCTAGACAAGGTGGAGTGCCTACTGCCTCGGA
Mouse Bach1 #4	TGCTGTTGACAGTGAGCGATCCTAGAGTTCAAACCTAAGTGAAGCCACAGATGTATAGTTAGTTTGAACCTCTAGGAGTGCCTACTGCCTCGGA
Mouse Fbxo22 #4	TGCTGTTGACAGTGAGCGCTCCCTCAAATGAAGGAATAATAGTGAAGCCACAGATGTATTTTCTTCAATTTGAGGGAATGCCTACTGCCTCGGA
Mouse Fbxo22 #5	TGCTGTTGACAGTGAGCGAACCAGGAAGGTGGAGTTAATAGTGAAGCCACAGATGTATTAACCTCCACCTTCTGTTGCTACTGCCTCGGA
Human BACH1 #3	TGCTGTTGACAGTGAGCGCCACAACAGTGTAAACAGAAAATAGTGAAGCCACAGATGTATTTTCTGTTAACACTGTTGTTGCCTACTGCCTCGGA
Human FBXO22 #1	TGCTGTTGACAGTGAGCGCTCCAAAGAATTTAACATTAAGTGAAGCCACAGATGTATCTAATGTTAAATTTCTGGATTGCCTACTGCCTCGGA
Human FBXO22 #5	TGCTGTTGACAGTGAGCGAAGCAAGTAGTCAGCACTTTCATAGTGAAGCCACAGATGTATGAAAGTGTGACTACTGCTGCTACTGCCTCGGA
shRNA control	TGCTGTTGACAGTGAGCGCAGATTATACTAGAAGTTGATAGTGAAGCCACAGATGTATACAACCTTCTAGTATAATCTGTTGCCTACTGCCTCGGA