Supplemental Information

Supplemental Materials and Methods

RNA extraction and quantitative real-time PCR (qRT-PCR)

Total RNA was extracted from cultured cells using TRIzol reagent (Invitrogen, Cat No. 15596018) according to the manufacturer's instructions. cDNA was synthesized from total RNA using random primers. The qRT-PCR primers for *TFAP4*, *CD133*, *NANOG*, *SOX2*, *ABCG2*, *BMI1*, *CD44*, *EpCAM*, *ALDH1A1*, *JUN*, *CCND1*, *TCF1*, *TCF4*, *DVL1*, *LEF1* and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) were designed using Primer Express version 5.0 software (Applied Biosystems). *TFAP4* forward primer: 5'-GCTGAGTCTCG GGGGTTAGT-3';*TFAP4* reverse primer: 5'- GTGCCCTCTTTGCAACATTT-3'; *CD133* forward primer: 5'-TTTTGGATTCATA TGCCTTCTGT-3'; *CD133* reverse primer: 5'-ACCC ATTGGCATTCTCTTTG-3'; *NANOG* forward primer: 5'- AT GGAGGAGGGAAGAGGAGA -3'; *NANOG* reverse primer: 5'- GATTTGTGGGCCTGAAGAAA-3'; *SOX2* forward primer: 5'- GCTTAGCCTCGTCGATGAAC-3'; *SOX2* reverse primer: 5'- AACCCCAAGATGCACA ACTC-3'; *ABCG2* forward primer: 5'-AAGCCATTGGTGTTTCCT TG -3'; *ABCG2* reverse primer: 5'-CTGGATCCTGAGCCTTTGG-3; *BMI1* forward primer: 5'- CAGGTGGGGATTA GCTCAG-3'; *BMI1* reverse primer: 5'-CTTTCATTGTCTTTTCCGCC-3'; *CD44* forward primer: 5'-CGTGGAATACACCTGCAAAG -3'; *CD44* reverse primer: 5'-CGGACA CC ATGGACAAGTTT-3'; *EpCAM* forward primer: 5'-GCTGGTGTGTGAACACTGCT-3'; *EpCAM* reverse primer: 5'-ACGCGTTGTGATCTCCTTCT-3'; *ALDH1A1* forward primer: 5'-CCACTCACTGAATCATGCA -3'; *ALDH1A1* reverse primer : 5'-TGAGCCAGTCACCTGT GTTC-3';*c-Jun* forward primer: 5'-GTCCTTCTTCTCTTGCGTGG-3'; *c-Jun* reverse primer: 5'-GGAGACAAGTGGCAGAGTCC-3'; *CCND1* forward primer: 5'-AGTTGTTGGGGCTC CTCAG-3'; *CCND1* reverse primer: 5'-AG ACCTTCGTTGCCCTCTGT-3'; *TCF1* forward primer: 5'-GTGCTGCTGCAGGTAGGACT-3'; *TCF1* reverse primer: 5'-CCATCCTCAAAG

AGCTGGAG-3’; *TCF4* forward primer: 5′-GCCTCTCATCAC TACAGCA-3′; and *TCF4* reverse primer, 5′-GGATGGGGGATTTGTCCTAC 3′ *DVL1* forward primer: 5’-TGCCCCAC CCATCAGTACCCTG-3’; *DVL1* reverse primer: 5’-CCCACTGCTCTTGCTCCCTTCA CT-3’; *LEF1* forward primer: 5’-GCAGCTATCAACCAGATCC-3’; *LEF1* reverse primer: 5’-GATGTAGGCAGCTGTCATTC-3’; *GAPDH* forward prime: 5'-ACATCCCCTCACCAAT A ACAAC-3'; *GAPDH* reverse primer: 5'-TAGCCAAATCATACTGCTCGTC-3'. Expression data were normalized to the geometric mean of housekeeping gene *GAPDH* to control the variability in expression levels and calculated as $2^{-[(\text{Ct of gene})-(\text{Ct of } GAPDH)]}$, where Ct represents the threshold cycle for each transcript.

Plasmids, retroviral infection, and transfection

The primers used to amplify the cDNA are: TFAP4 primer up : 5’-GCCGAATTCACCATG GACTAC AAGGACGACGATGACAAGGAGTATTTCATGGTGCCCAC-3’; TFAP4 primer down : 5’- GCCGCTAGCTCAGGGAAGCTCCCCGTCCCCCG-3’. The amplified PCR products was further digested by BamH1 and EcoR1 restriction enzymes and subcloned into pSin-EF2 lentiviral vector. The sequences of RNAi#1 and RNAi#2 are GCCTTGCCAACAT TCCACTAA and TTAGTGGAATGTTGGCAAGGC, respectively. The reporter plasmids and siRNA were transfected into HCC cells using Lipofectamine 2000 (Life Technologies), according to the manufacturer’s instructions. The sense strand sequences of siRNAs, which were designed to target human cells, were as follows: si*DVL1*, 5’- GCAGAGUGAAGGGAGC AAATT-3’; si*LEF1*, 5’-UCAGAUGUCAACUCCAAACA A-3’. The reporter plasmids containing wild-type (CCTTTGATC; TOP flash) or mutated (CCTTTGGCC; FOP flash) TCF/LEF DNA binding sites were purchased from Upstate Biotechnology Company.

Primer for Chromatin immunoprecipitation-qPCR (ChIP-qPCR)

LEF1 forward 1: 5’- AAACCTTAACAGATGCGTTCAGCAG -3’; LEF1 reverse 1: 5’- CCA CTAATTGGCTAATAAAAACAGA -3’; LEF1 forward 2: 5’-ACAGTGCAA AAGGATCGTT

TTTTTT -3’; LEF1 reverse 2: 5’- TGTCATTATTTCAAACTTTTCCAGA -3’; LEF1 forward 3: 5’- TTCGTCTACTGCAAGAGCCAAGTTC -3’; LEF1 reverse 3: 5’- CTGGGAAGTGCA CGCAGATATG -3’; LEF1 forward 4: 5’- ATAGGGCAATTCATTTTAGCCATCC -3’; LEF1 reverse 4: 5’- AAGGAACAAATGAAATGCTATTAAT -3’; LEF1 forward 5: 5’- CTCTGGG CTGTGAGTGTGAGCGGAC -3’; LEF1 reverse 5: 5’- TTTGGTGTTTTTGGTCCACTGTG GG - 3’; TCF4 forward 1: 5’- ACGTGTGTCAACTATCATGGAAGGA-3’; TCF4 reverse 1: 5’-CTTTACATTTCTACAGTGCAATGTT-3’; TCF4 forward 2: 5’-AGGATTAACTAATTGA AGGA ATATA-3’; TCF4 reverse 2: 5’-GGTGCAAAAGGTTAAGGCTGACTTG-3’; TCF4 forward 3: 5’- ACAAACCCAGCTTGGTCAAGGCCAT-3’; TCF4 reverse 3: 5’-GCAAAAT GAAATGAA ACAAGTGTCA-3’; TCF4 forward 4: 5’- AGTTTGTTACTTTATTTTACTATTT -3’; TCF4 reverse 4: 5’- AATTTCCCCGCTGCGACCTACAACT -3’; TCF4 forward 5: 5’-AA TGGTGAAAATATATGAACTGTGG -3’; TCF4 reverse 5: 5’-CTTTATAACAAAAAGCCAG GCACTG - 3’; DVL1 forward 1: 5’- TCTGAAAGTACGTGGAGGACGGGAC -3’; DVL1 reverse 1: 5’- CCCATCTCCCCAAGACCTCCCTCCC -3’; DVL1 forward 2: 5’- CACCTGT GTCTGAGCAGCCGTGTTG -3’; DVL1 reverse 2: 5’- AGGAGGTGGGGGTCAGCCGAG AGCC -3’; DVL1 forward 3: 5’- GGCTCCCCCCGCCCCACCCCACGAC -3’; DVL1 reverse 3: 5’- CTGGAAGGACTGGCGGCTGCCTGTC -3’; DVL1 forward 4: 5’-CCGGCCCTGTCC TCGCGCCTGCATG -3’; DVL1 reverse 4: 5’- GATGGGGAAGGAGCCTGTCAGCACC -3’; DVL1 forward 5: 5’- GACCTTGGGCCGGTAAGCCAGGGTC -3’; DVL1 reverse 5: 5’- CAG GACCCTGGCCGACGGATGACTC -3’. The primers used to generate DVL1 and LEF1 luciferase reporters are: DVL1 primer up: 5’- TCTGAAAGTACGTGGAGGACGGGAC-3’; DVL1 primer down: 5’- CCCATCTCCCCAAGACCTCCCTCCC-3’; LEF1 primer up: 5’- AA ACCTTAACAGATGCGTTCAGCAG -3’; LEF1 primer down: 5’- CCACTAATTGGCTAAT AAAAACAGA -3’.

Supplemental Figure legends

Supplemental Figure 1. Elevated TFAP4 expression is associated with overall survival in human HCC. Kaplan-Meier analysis of the association between TFAP4 expression and overall survival in the TCGA dataset **(A)** and 197 primary HCC specimens **(B).** HCC: hepatocellular carcinoma; TCGA: the cancer genome atlas; TFAP4: transcription factor AP-4.

Supplemental Figure 2. TFAP4 overexpression activates the Wnt/β-catenin signaling. (A) GSEA analysis indicating a significant correlation between TFAP4 mRNA expression and Wnt-suppressed gene signatures (LABBE_WNT3A_TARGETS _DN, LABBE_TARGETS_ OF_TGFB1_AND_WNT3A_DN). **(B)** Prediction and validation of the interaction between TFAP4 and the *TCF4* promoter. **(C)** Real-time PCR analysis of the mRNA expression of TCF4, LEF1 and DVL1 genes. Bars in B and C indicate the mean ± SD of three independent experiments; * $P < 0.05$. DVL1: dishevelled segment polarity protein 1; GSEA: gene set enrichment analysis; LEF1: lymphoid enhancer binding factor 1; TFAP4: transcription factor AP-4; TCF4: transcription factor 4.

Supplemental Figure 3. TFAP4 overexpression enhances the TIC-like phenotype through activation of Wnt/β-catenin pathway in human HCC. (A) Real-time PCR analysis of stemness-related markers in the indicated cells. **(B)** Real-time PCR analysis of the expression of Wnt/β-catenin signaling pathway downstream genes in the indicated cells. Bars represent the mean ± SD of three independent experiments. *$P < 0.05$. HCC: hepatocellular carcinoma; TIC: tumor-initiating cell; TFAP4: transcription factor AP-4.

Supplemental Tables

Supplemental Table 1. Clinicopathological Characteristics of Clinical Samples and Expression of TFAP4 in Liver Cancer

Characteristics	No. patients	(%)
Age (years)		
≤ 60	104	(52.8)
> 60	93	(47.2)
Gender		
Male	109	(55.3)
Female	88	(44.7)
Clinical stage		
I-II	63	(32.0)
III-IV	134	(68.0)
Histological differentiation		
Well/Moderate	127	(64.5)
Poor	70	(35.5)
HBsAg		
positive	175	(88.8)
negative	22	(11.2)
AFP		
> 400 **ng/ml**	171	(86.8)
≤ 400 **ng/ml**	26	(13.2)
Tumor size		
> 5 cm	121	(61.4)

≤ 5 cm	76	(38.6)
Vital status (at follow-up)		
Alive	131	(66.5)
Death due to liver cancer cause	66	(33.5)
Expression of TFAP4		
Low expression	87	(44.2)
High expression	110	(55.8)
Relapse		
Yes	63	(32.0)
No	134	(68.0)

Supplemental Table 2. Correlation between TFAP4 Expression and Clinicopathologic Charateristics of Liver Cancer Patient

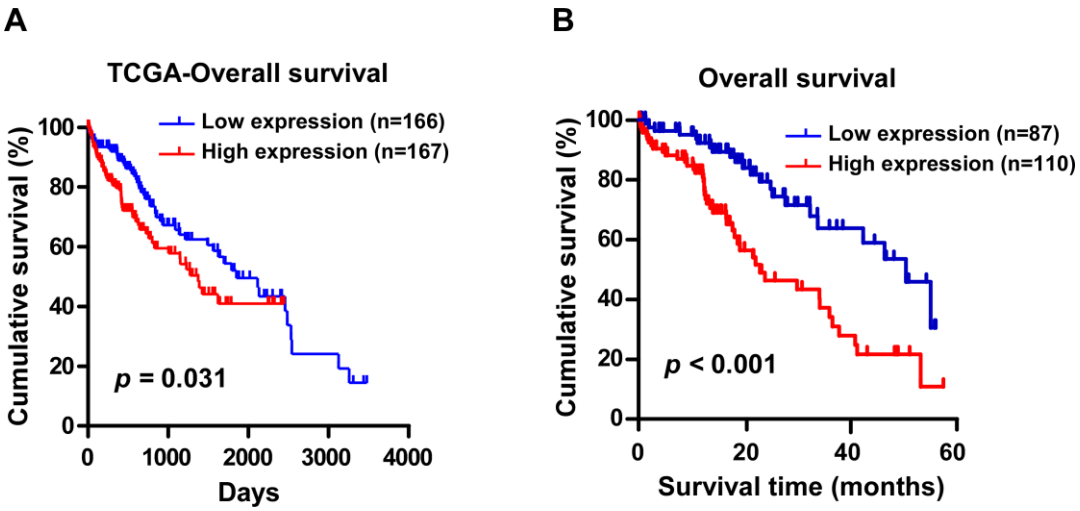
Characteristic		TFAP4 Protein level		Chi-square test
		Low(87)	**High(110)**	**P value**
Age (years)	≤ 60	46	58	*0.984*
	> 60	41	52	
Gender	Male	53	56	*0.160*
	Female	34	54	
Clinical stage	I-II	49	14	*<0.001*
	III-IV	38	96	
Histological differentiation	Well/Moderate	63	64	*0.038*
	Poor	24	46	
Tumor size (cm)	> 5	42	79	*0.001*
	≤ 5	45	31	
AFP (ng/ml)	> 400	71	100	*0.055*
	≤ 400	16	10	
Vital status	Alive	65	66	*0.030*
	Death	22	44	
HBsAg	Positive	77	98	*0.897*
	Negative	10	12	
Relapse	Yes	19	44	*0.007*
	No	68	66	

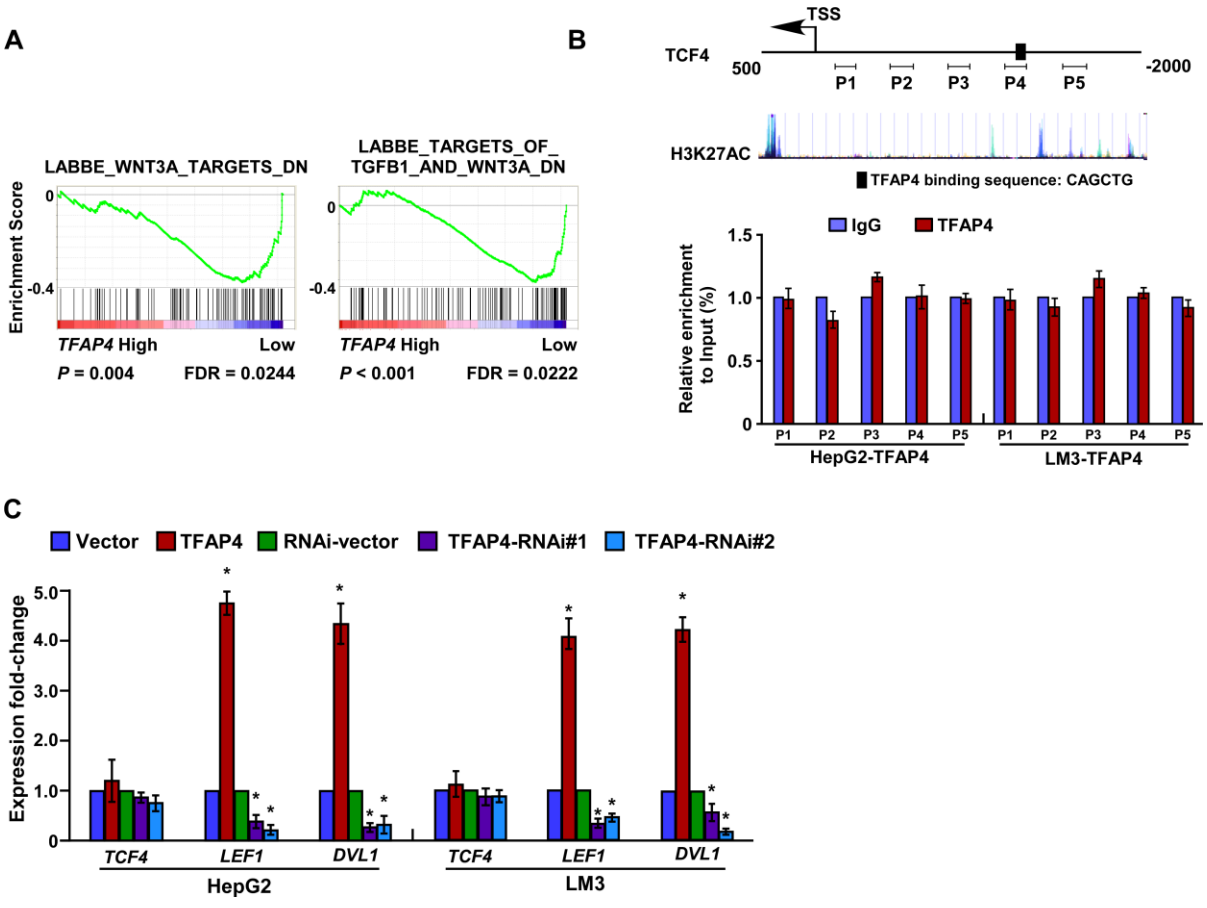
Supplemental Table 3. Spearman Analysis of Correlation between TFAP4 and Clinicopathological Characteristics

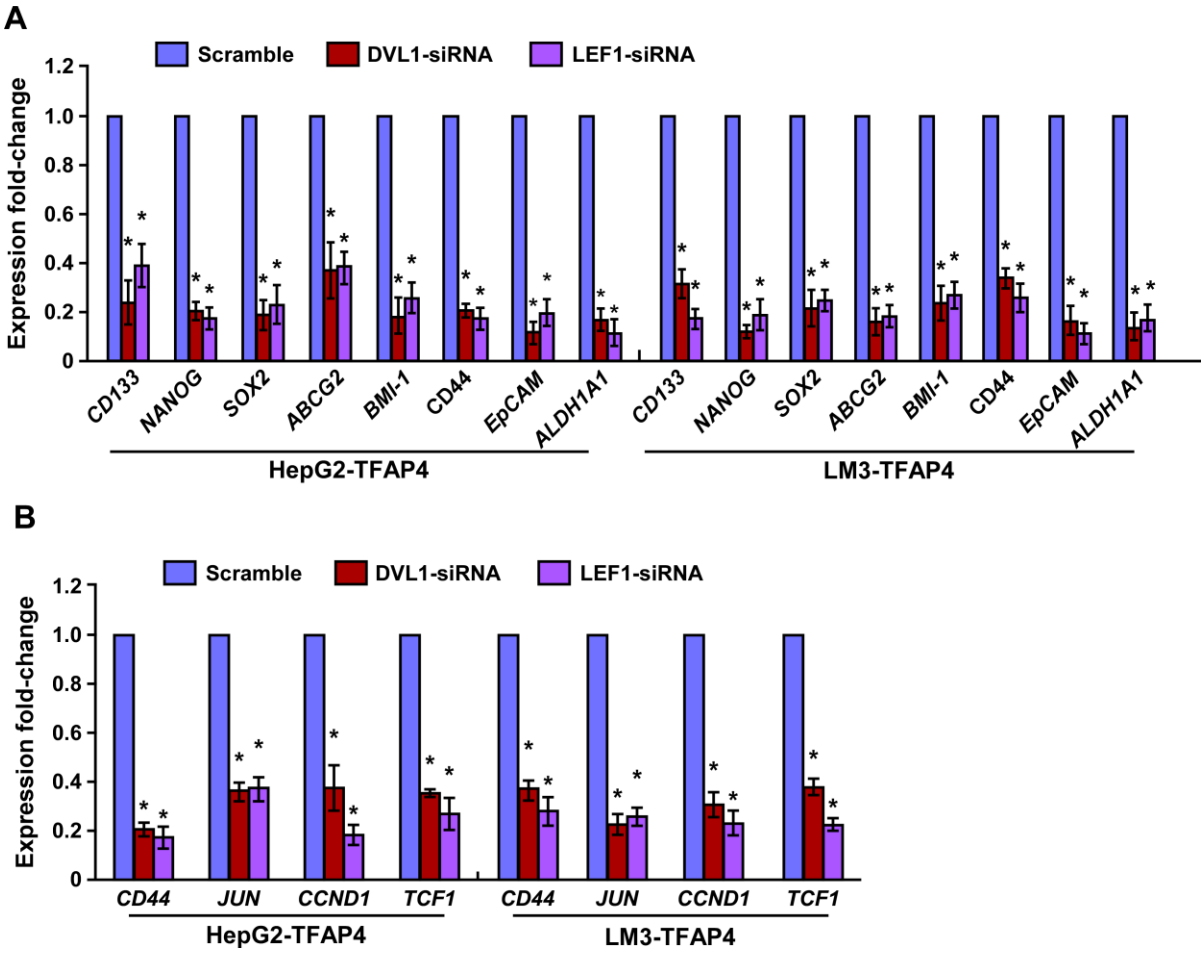
Variables	TFAP4 expression level	
	Spearman Correlation	***P* Value**
Survival time	-0.305	*<0.001*
Vital status	0.155	*0.030*
HBsAg	0.009	*0.898*
Age	0.001	*0.984*
Clinical stage	0.403	*<0.001*
Histological differentiation	0.148	*0.038*
AFP	0.136	*0.056*
gender	0.100	*0.162*
Relapse	0.193	*0.006*
Tumor size	0.240	*0.001*

Supplemental Table 4. Univariate and Multivariate Analyses of Various Prognostic Parameters in Liver Cancer Patients by Cox-regression Analysis

	Univariate analysis			Multivariate analysis		
	Relative risk	95% confidence interval	*P*	Relative risk	95% confidence interval	*P*
Stage	0.125	1.127-1.838	*0.003*	1.360	1.033-1.792	*0.029*
Recurrence	0.250	1.260-3.360	*0.004*	1.703	1.022-2.839	*0.041*
TFAP4	0.263	1.585-4.451	*<0.001*	1.825	1.027-3.244	*0.040*
Histological differentiation	0.252	1.070-2.873	*0.026*	1.792	1.080-2.974	*0.024*


A
TCGA-Overall survival
Low expression (n=166)
High expression (n=167)
Cumulative survival (%)
100
80
60
40
20
0
p = 0.031
0
1000
2000
3000
4000
Days
B
Overall survival
Low expression (n=87)
High expression (n=110)
Cumulative survival (%)
100
80
60
40
20
0
p < 0.001
0
20
40
60
Survival time (months)


A
LABBE_WNT3A_TARGETS_DN
LABBE_TARGETS_OF_
TGFB1_AND_WNT3A_DN
Enrichment Score
0
-0.4
TFAP4 High
Low
P = 0.004
FDR = 0.0244
P < 0.001
FDR = 0.0222
B
TSS
TCF4
500
-2000
P1
P2
P3
P4
P5
H3K27AC
TFAP4 binding sequence: CAGCTG
IgG
TFAP4
Relative enrichment to Input (%)
1.5
1.0
0.5
0
HepG2-TFAP4
LM3-TFAP4
C
Vector
TFAP4
RNAi-vector
TFAP4-RNAi#1
TFAP4-RNAi#2
Expression fold-change
5.0
4.0
3.0
2.0
1.0
0
TCF4
LEF1
DVL1
HepG2
LM3


A
Scramble
DVL1-siRNA
LEF1-siRNA
Expression fold-change
1.2
1.0
0.8
0.6
0.4
0.2
0
CD133
NANOG
SOX2
ABCG2
BMI-1
CD44
EpCAM
ALDH1A1
CD133
NANOG
SOX2
ABCG2
BMI-1
CD44
EpCAM
ALDH1A1
HepG2-TFAP4
LM3-TFAP4
B
Scramble
DVL1-siRNA
LEF1-siRNA
Expression fold-change
1.2
1.0
0.8
0.6
0.4
0.2
0
CD44
JUN
CCND1
TCF1
CD44
JUN
CCND1
TCF1
HepG2-TFAP4
LM3-TFAP4