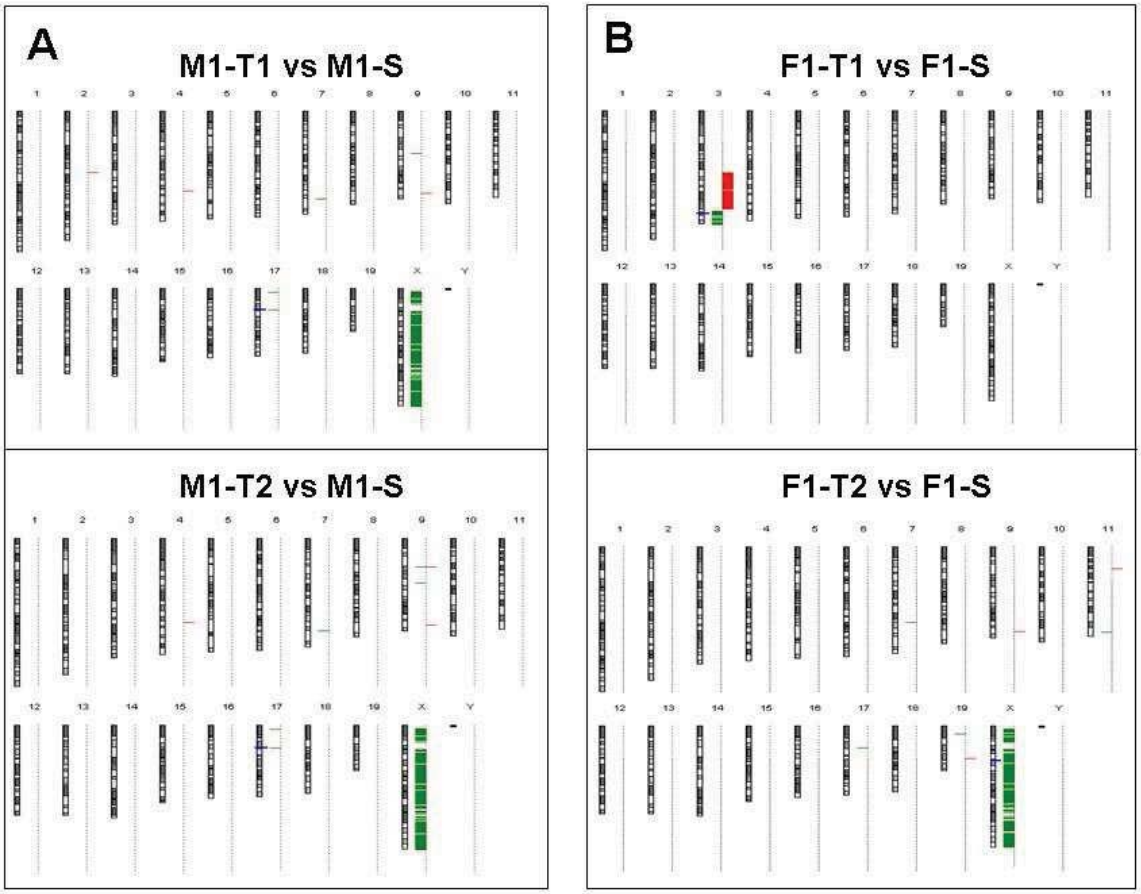
Supplementary Figures and Tables



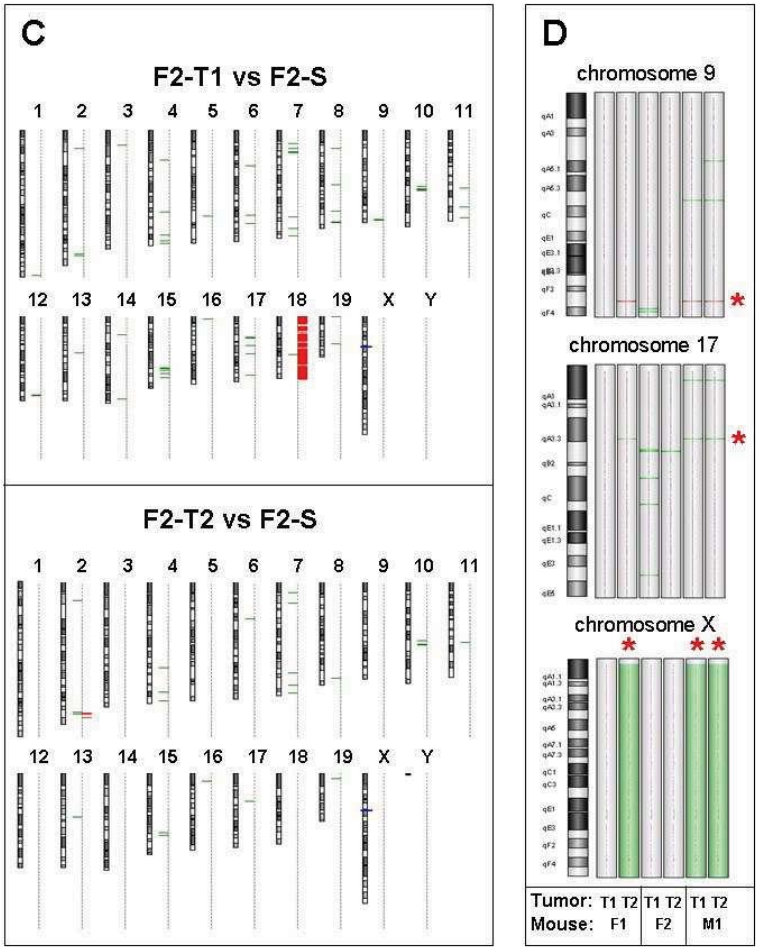


Figure S1. Distribution of chromosomal aberrations in the FIV-induced liver tumors. Comparative genomic hybridization (CGH) was performed on genomic DNA extracted from tumors (T1, T2) obtained from male mouse, M1 and female mice, F1 and F2, and compared to DNA extracted from the surrounding, non-tumorous, tissue (S). Amplifications (red) and deletions (green). (**a-c**) Distribution of aberrations according to chromosome. (**d**) Small aberrations in chromosomes 9 and 17, as well as a complete deletion of the X chromosome, were common to the three tumors, F1-T2, M1-T1 and M1-T2, (marked by a red asterix).

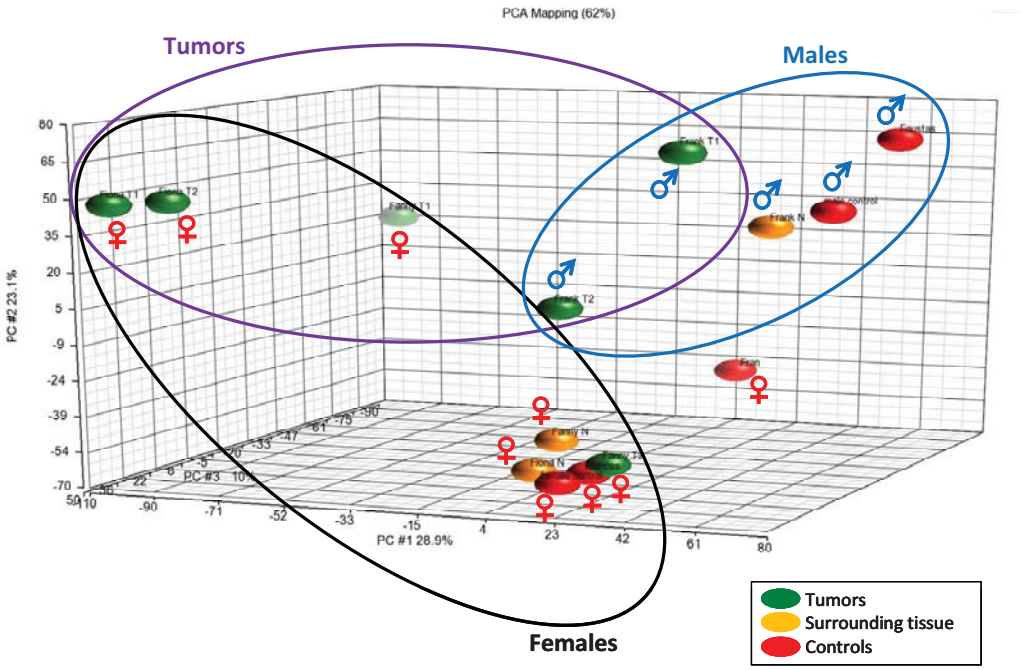


Figure S2. A graphical representation of the Principal Component Analysis (PCA) of genome scale gene expression profiling data obtained from FIV-transduced and control mice. Tumor tissue (T, green); matched non-tumorous liver tissue (N, yellow); liver tissue from healthy untransduced, or from tumor-free FIV-transduced mice (H, red). Male clusters (blue line); female clusters (black line); tumors (purple line).

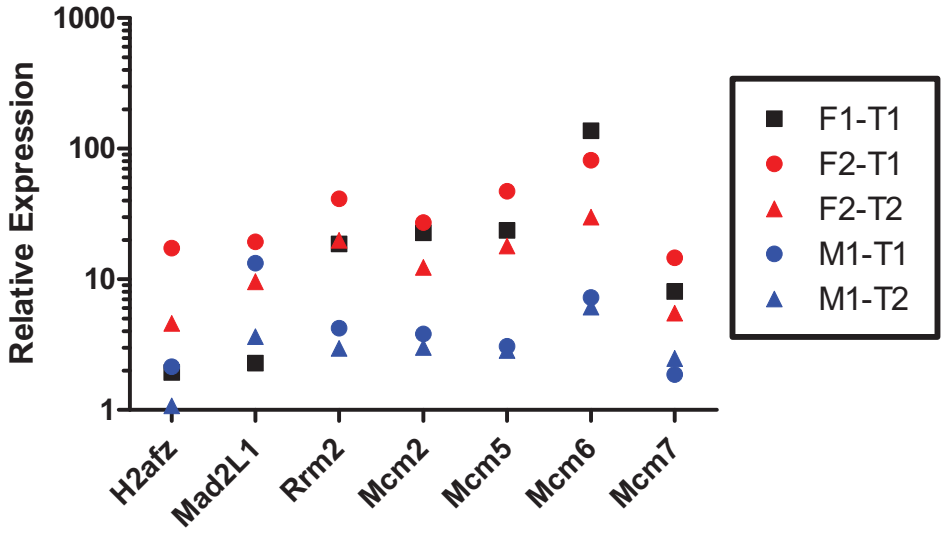


Figure S3. Validation of genome-wide microarray expression data by qRT-PCR. Total RNA was extracted from the five FIV-induced HCCs that shared gene expression patterns. Seven E2F target genes that were significantly upregulated in the genome-wide expression arrays were validated by qRT-PCR. Relative quantitation (RQ) of the tumor tissue was calculated using non-tumorous tissue as the calibrator (RQ=1). HPRT was used as the endogenous control.

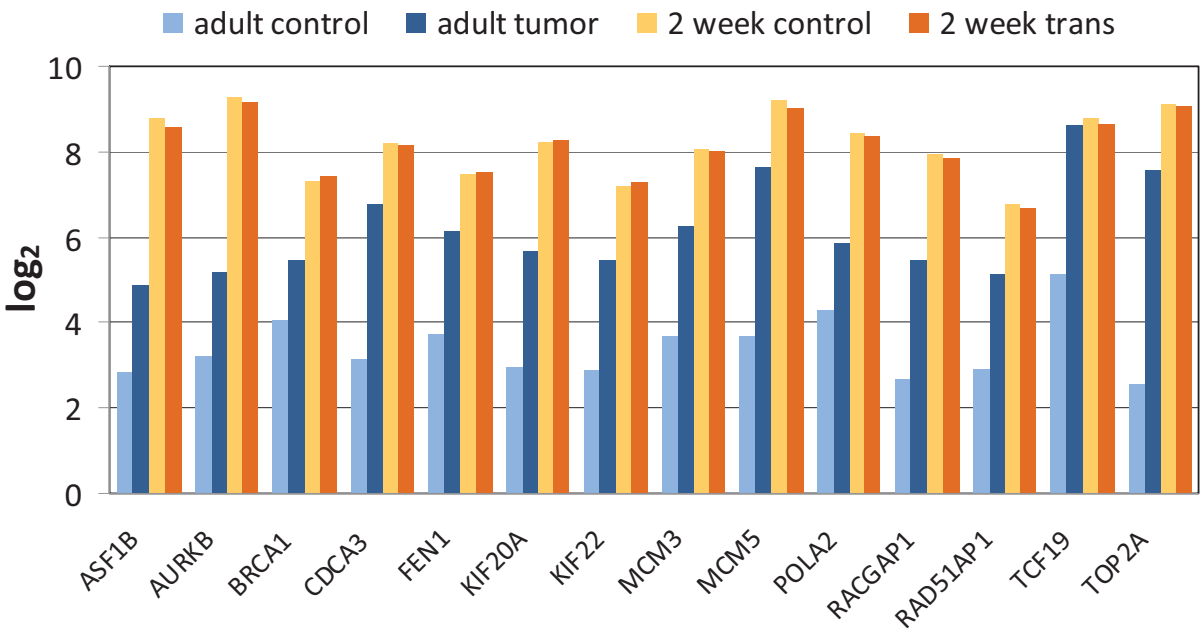


Figure S4. High level of E2F target gene expression in livers of young mice. Expression levels of E2F target genes common to humans and rats that were upregulated in the FIV-induced tumors, as determined by genome-wide array analysis. Microarray analysis was performed on total RNA extracted from FIV-induced liver tumors that developed in adult mice (dark blue) and matched surrounding tissue (light blue), as well as from livers of 10 day old mice untransduced (yellow) and FIV-transduced (orange). The following housekeeping genes were expressed at the same level in all datasets: Arl6ip1, Copb2, Eif2s3x, Man1a, Srp9, Tfg, and Ywhae.

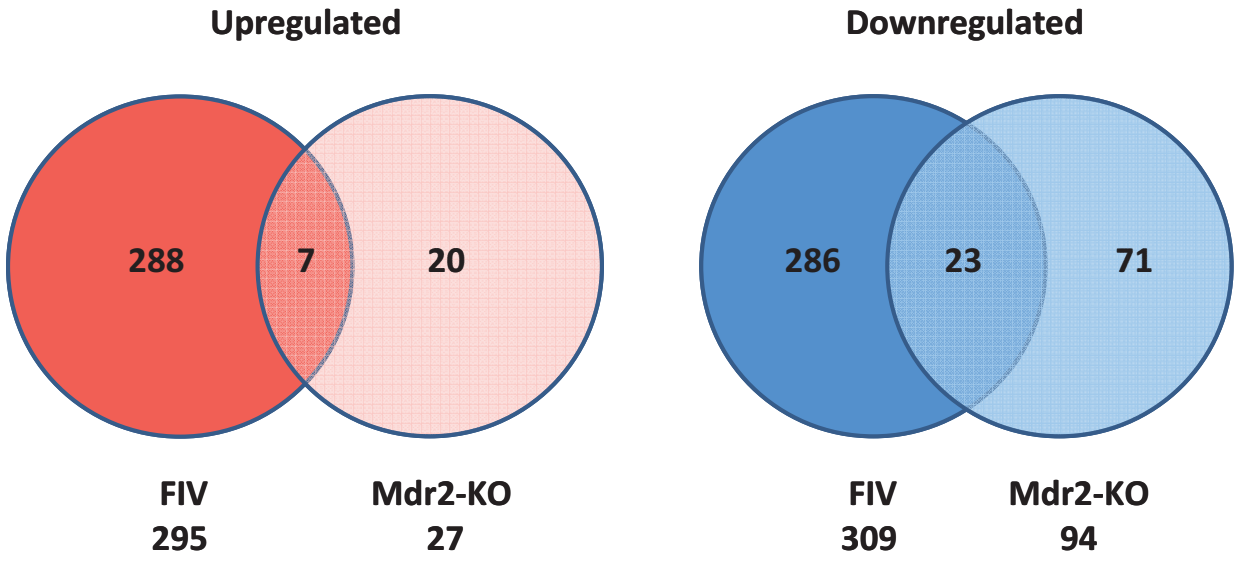


Figure S5. Comparison of deregulated genes between the two murine HCC models, namely, FIV-induced (current study) and inflammation-mediated Mdr2-KO mice[34]. The Venn diagram demonstrates numbers of gene probes that were either up- or down-regulated in at least four tumors in each dataset. Fold-change threshold for tumor versus non-tumor was $\geq$ 1.8 for up-regulated and $\leq$ -1.8 for down-regulated genes. Six tumors were analyzed in each HCC model and compared to matched non-tumor tissue (3 samples in FIV dataset and 4 in the Mdr2-KO dataset).

Table S1. FIV integration sites within liver tumors

Tumor ID	Chromosomal Location[a]	Gene Symbol	Intragenic[b]	Intergenic (bps)[c]		Integration orientation[d]
F1(T1)	Chr4: 82315639	Nfib	intron 10			opposite
	Chr4: 106175353				>100kb	
	Chr5: 115025088	Spp13	intron 1			same
	Chr6: 4677827	Sgce	intron 10			same
	Chr13: 77132015	Ankrd32	intron 1			opposite
	Chr14: 22596501	1700112E06Rik	exon 4-intron 5			opposite
	Chr18: 9309351	ccny		upstream	4580	same
	Chr19: 29513373	A930007I19Rik	intron 2			same
F1(T2)	Chr1: 134747726	Syt2	3'UTR			same
	Chr2: 76984794	Ttn		upstream	2500	same
	Chr4: 64586468				>100kb	
	Chr10: 84917375	Ric8b		upstream	241	opposite
	Chr19: 6904891	Prdx5		downstream	4754	same
	Chr19 :61095340	Grk5		downstream	3038	opposite
F2(T1)	Chr1: 15845050	Terf1		downstream	1621	opposite
	Chr1: 13539165	Tram1		downstream	50699	same
	Chr5: 88759162	AK046278	exon1			
	Chr6: 122437822	Rimklb		downstream	9480	opposite
	Chr11: 12876885				>100kb	
	Chr18: 84327800	Zfp407	intron 7			same
	ChrX: 72959054	Nsdh1		downstream	526	opposite
F2(T2)	Chr3: 107325730	AK048987	3'UTR			opposite
	Chr5: 143645250	Cyth3	intron 1			same
M1(T1)	Chr3: 63958416	Slc33a1	intron 1			same
	Chr9: 108531707	Qrich1	intron 2			opposite
	Chr12: 85327916	Nek9	intron 16			same
	Chr16: 31262903	Ppp1r2	intron 2			same
M1(T2)	Chr16: 4626884	Coro7	intron 9			opposite

[a] Location of first nucleotide following 3' LTR according to the UCSC Genome Browser, Dec 2011 mouse assembly.

[b] Location of integration within gene.

[c] Position and distance of viral insertion relative to nearest gene within 100kb.

[d] Orientation of vector LTR with respect to the direction of transcription of the targeted gene.

Table S2. Syntenic regions of human common fragile sites in the mouse genome.

Fragile site name	Human genome[a]	Size (bp)	Mouse genome[a]	Size (bp)	Tumors bearing a coinciding chromosomal aberration[b]	Location of aberration
FRA1E	chr1: 68362127-156421365	88,059,239	chr3: 86799634-159599772	72,800,139	F1-T1	chr3: 87535405-159870502
FRA1H	chr1: 205641715-225711350	20,069,636	chr1: 181768214-197047501	15,279,288		
FRA2G	chr2: 139785273-188118461	48,333,189	chr2: 39363343-84313407	44,950,065		
FRA3B	chr3: 57917214-63984740	6,067,527	chr14: 8637918-14953921	6,316,004		
FRA4F	chr4: 93567701-93586956	19,256	chr6: 63311998-63322971	10,974		
FRA6E	chr6: 160023022-167303093	7,280,072	chr17: 7175001-13208514	6,033,514		
FRA6F	chr6: 100641521-155090183	54,448,663	chr10: 3011989-51231556	48,219,568		
FRA6H	chr6: 29430682-33405196	3,974,515	chr17: 34042414-37636846	3,594,433	F2-T1, F2-T2	chr17: 34622785-35297915
FRA7E	chr7: 76539913-92524425	15,984,513	chr5: 3144811-20984943	17,840,133		
FRA7G	chr7: 92596878-150189329	57,592,452	chr6: 3320884-48858948	45,538,065		
FRA7K	chr7: 107559442-111923382	4,363,941	chr12: 40896394-45712997	4,816,604		
FRA7H	chr7: 92596878-150189329	57,592,452	chr6: 3320884-48858948	45,538,065		
FRA7I	chr7: 92596878-150189329	57,592,452	chr6: 3320884-48858948	45,538,065		
FRA8C	chr8: 97515808-146131251	48,615,444	chr15: 32717008-76767049	44,050,042		
FRA9E	chr9: 27315071-122528763	95,213,693	chr4: 34896196-70204380	35,308,185		
FRA9G	chr9: 6745997-27290706	20,544,710	chr4: 73894394-94597321	20,702,928		
FRA11E	chr11: 26252973-57510525	31,257,553	chr2: 84322813-110832617	26,509,805	M1-T1	chr2: 85936918-85946800
FRA11F	chr11: 71304680-88990549	17,685,870	chr7: 94258787-109226497	14,967,711		
FRA11G	chr11: 106957827-134348749	27,390,923	chr9: 25880811-53831242	27,950,432	M1-T2	chr9: 37858515-37931443
FRA13A	chr13: 33378325-40146259	6,767,935	chr3: 52065711-56976945	4,911,235		
FRA13E	chr13: 52124034-101887709	49,763,676	chr14: 79986707-125114860	45,128,154		
FRA16D	chr16: 45250774-88637531	43,386,758	chr8: 87782679-126060687	38,278,009		

[a]Chromosomal locations according to UCSC Genome Browser assemblies 2006

[b]Chromosomal aberrations as determined by CGH.

Table S3. Summary of FIV integration profiling following high throughput sequencing of transduced AML 12 cells

BIOINFORMATIC ANALYSIS	*DEFINITION*	*NO. OR % OF SEQUENCES*[a]
no. of raw sequences	Number of sequences that were generated by 454 high throughput sequencing.	93,115
output of sorter	The sorter removes the viral LTR, barcode sequences, and genomic sequences that do not immediately follow the LTR.	72,994
output of CD-HIT	Redundant sequences are removed from the output of sorter.	32,480
output of ISA	Sequences that are too small are removed from the output of CD-HIT; remaining sequences are matched to the mouse genome.	19,603
no hits found	No matches were found in the UCSC Genome Browser.	991
no gene within 150 kb	No gene was found within 150 kb of the integration site.	1,988
number of unique sequences	This equals the output of ISA minus (multiple hits + no hits).	17,576
RESULTS		
hits in gene	Number of integrations occurring within a gene.	6,437
% of integrations within genes	Number of integrations within genes/number of unique sequences.	37±2
number of integrations in repeat sequences	Number of sequences in LINEs, SINEs, LTRs.	1,036
% of integrations in repeat sequences	Percentage of sequences in LINEs, SINEs, LTRs.	9±4.8
number of integrations upstream of closest gene		5,818
% of integrations upstream of closest gene	Of integrations occurring between genes, the percentage that were upstream of the closest gene.	38±1
number of integrations downstream of closest gene		9,241
% of integrations downstream of closest gene	Of integrations occurring between genes, the percentage that were downstream of the closest gene.	62±1

[a]sum or average of four experiments±SD

Table S4: Genes deregulated in livers of tumor-free, transduced mice, compared to untransduced controls

Gene Symbol	Gene Title	(O) Oncogene or (TS) Tumor	Signaling network	HCC associated genes
Genes up-regulated in FIV-transduced liver				
Arntl	aryl hydrocarbon receptor nuclear translocator-like		P53	
Cdkn1a	cyclin-dependent kinase inhibitor 1A (P21)	TS	P53	√
Chka	choline kinase alpha	O		
Ddit4	DNA-damage-inducible transcript 4	TS	P53	√
Dusp6	dual specificity phosphatase 6	O/TS		
Egr1	early growth response 1	O/TS	P53	√
G0s2	G0/G1 switch gene 2		PPAR	
Gadd45a	growth arrest and DNA-damage-inducible 45 alpha		P53	√
Hspb1	heat shock protein 1			
Igfbp1	insulin-like growth factor binding protein 1		P53	
Ndrg1	N-myc downstream regulated gene 1	TS	P53	√
Pdgfc	platelet-derived growth factor, C polypeptide	O		√
Pdk4	pyruvate dehydrogenase kinase, isoenzyme 4		Rb	
Pnrc1	proline-rich nuclear receptor coactivator 1			
Ppp1r3c	protein phosphatase 1, regulatory (inhibitor) subunit 3C		P53	
St5	suppression of tumorigenicity 5	TS		
Tbx3	T-box 3	O		
Trim34	Tripartite motif protein 34			
Genes down-regulated in FIV-transduced liver				
Armcx3	armadillo repeat containing, X-linked 3			
Cish	cytokine inducible SH2-containing protein			
Dact1	dapper homolog 1, antagonist of beta-catenin (xenopus)	TS	Wnt	√
Dbp	D site albumin promoter binding protein			
Foxq1	forkhead box Q1			
Mcm10	minichromosome maintenance deficient 10 (S.cerevisiae)			
Myc	myelocytomatosis oncogene	O	Myc	
Per3	period homolog 3 (Drosophila)	TS		√
Pfkfb3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3			
Phlda1	pleckstrin homology-like domain, family A, member 1	TS		
Plk3	polo-like kinase 3 (Drosophila)	TS		√
Rnd3	Rho family GTPase 3	TS	Rb	√
Tef	thyrotroph embryonic factor			
Thrsp	thyroid hormone responsive SPOT14 homolog (Rattus)			
Usp2	ubiquitin specific peptidase 2		P53	√

Table S5. HCC-associated oncogenes and tumor suppressors

HCC-associated oncogenes	HCC-associated tumor suppressors		
Akt	Ache	Hic1	Rprm
Birc2	Apc	Igfals	Runx3
Birc5	Arid2	Irf2	Scara5
Bmi1	Axin1	Kl	Sfn
Ccnd1	Blm	Klf4	Sfrp1
Cks1b	Brms1	Klf6	Sfrp2
Ctnnb1	Btbd9	Lats1	Sirt3
Ect2	Cacna1g	Lats2	Socs1
Egf	Cadm1	Men1	Sox1
Egfr	Casp8	Mxi1	Sparc
Eif5a2	Cdh1	Ndrg2	Stk11
Ezh2	Cdkn1a	Nf2	Sulf1
Hgf	Cdkn1b	Nr1h4	Trp53
Id1	Cdkn2a	Nrsn2	Tsc1
Jun	Cdkn2b	Pax5	Tsc2
Mad2l1	Creb3l3	Per3	Vac14
Met	Cyr61	Plagl1	Wrn
Mtdh	Dapk1	Plk4	Xpo4
Myc	Dcc	Pml	Zbbx
Pdgfb	Ddx20	Prdm2	
Pdgfra	Dlc1	Ptgs2	
Pdgfrb	Dleu2	Pycard	
Psmd10	Dpyd	Rassf1	
Rhoc	Fgf6	Rassf2	
Sirt1	Fgl1	Rassf5	
Smo	Fstl5	Rb1	
Sulf2	Gjd4	Rbl1	
Tgfa	Glo1	Rbl2	
Vegfa	Gorab	Recql	
Yap1	Gstp1	Recql5	

Table S6. Primer sequences of E2F target genes used for qRT-PCR

gene	orientation[a]	sequence
H2afz	F	GAAATCTAGGACAACCAGCCAC
	R	AAGTCTTTTGACGCATTTCCTG
Mad2L1	F	GTATCTCAATAATGTGGTGGAACAG
	R	ATATCAAACTGCCATCTTTCAAGG
Rrm2	F	GCTCTGAAACCCGATGAGAG
	R	GCCTCTGTAACTTGAACTTCTTGG
Mcm2	F	AAGAAGAGGATGGAGAGGAACTC
	R	CCTCCACATCTTCATCATCCAG
Mcm5	F	AGATGAGCTGCTCCAAGACATC
	R	CAGAGGCTGAAATGATGATGC
Mcm6	F	GAGAAACACGCTGGTTGTGAG
	R	AAGGTCTTCAAGGCTCGACAC
Mcm7	F	ATTGTCACTCGTGTGTCTGAAGTC
	R	ACATGATCAGAGGCATGAAAGTG

[a]F forward; R reverse