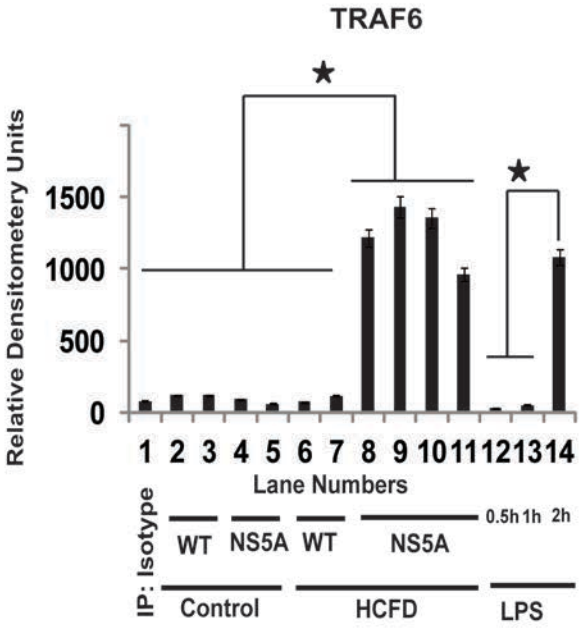
Supplementary Information List



Supplementary Figure 1. Fold change in Twist1 gene expression in TICs. Quantification of TRAF6 expressoin from Figure 1E using Image-J.

Suppl. Figure 1

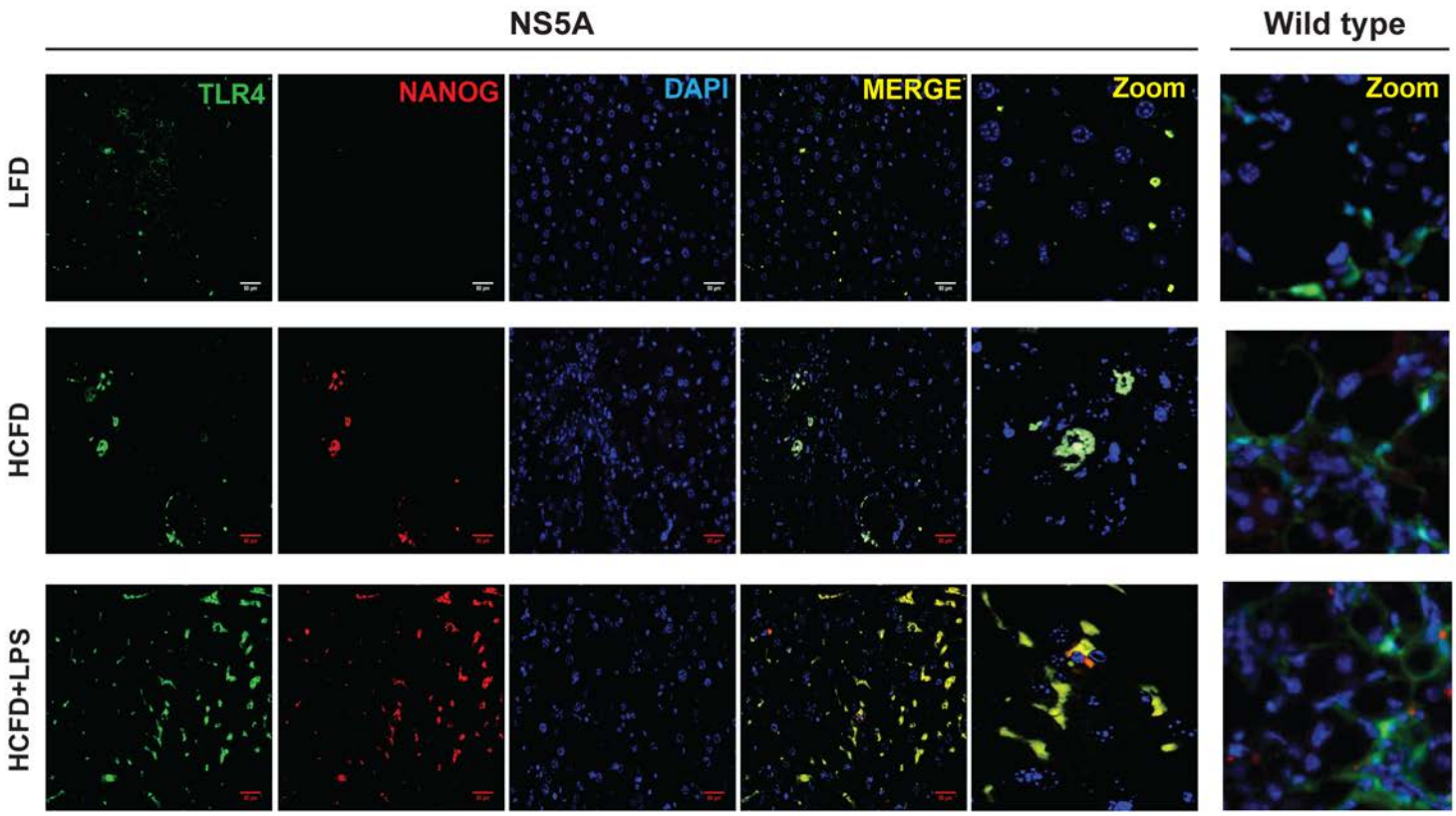
Uthaya Kumar *et al.*

NS5A+HCFD VS Non-Tg+HCFD Microarray gene expression analysis

Official Gene Name	Official gene symbol	Fold change in gene expression
osteomodulin	Omd	188.6
cytochrome P450	family 2	58.0
ubiquitin D	Ubd	47.6
cell death-inducing DNA fragmentation factor	alpha subun	45.8
sulfotransferase family 1E	member 1	31.7
neurotrophic tyrosine kinase	receptor	27.6
lymphocyte antigen 6 complex	locus D	26.4
neurotrophic tyrosine kinase	receptor	25.2
asporin	Aspn	25.1
rabphilin 3A	Rph3a	20.6
Transcribed locus		19.8
Nanog homeobox	Nanog	19.2
small proline-rich protein 1A	Sprr1a	16.0
chitinase 3-like 3	Chi3l3	15.1
camello-like 5	Cml5	14.5
G protein-coupled receptor 98	Gpr98	14.0
DnaJ (Hsp40) homolog	subfamily C	13.5
cholecystokinin A receptor	Cckar	13.4
fatty acid binding protein 4	adipocyte	13.4
fatty acid binding protein 4	adipocyte	12.9
lin-28 homolog (C. elegans)	Lin28	12.7
camello-like 4	Cml4	11.9
twist gene homolog 1 (Drosophila)	Twist1	11.9
plasma membrane associated protein	S3-12	11.5
CEA-related cell adhesion molecule 11	Ceacam11	11.5
C-type lectin domain family 7	member a	11.3
serine threonine kinase 31	Stk31	11.3
matrix metallopeptidase 13	Mmp13	11.0
pancreatic lipase-related protein 2	Pnliprp2	10.8
matrix metallopeptidase 16	Mmp16	10.3
cadherin 4	Cdh4	10.3
killer cell lectin-like receptor subfamily C	member 3	10.2
prolactin family 3	subfamily c	10.1
DNA cytosine methyltransferase mRNA		10.0
CEA-related cell adhesion molecule 11	Ceacam11	9.8
leucine rich repeat transmembrane neuronal 2	Lrrtm2	9.7
calcitonin/calcitonin-related polypeptide	alpha	9.6
solute carrier family 10	member 2	9.5
Crx opposite strand transcript 1	Crxos1	9.5
RIKEN cDNA D930048N14 gene	D930048N	9.4
uncoupling protein 1 (mitochondrial	proton carrie	9.4
mitogen activated protein kinase 10	Mapk10	9.3
histocompatibility 2	class II antig	9.3
synovial sarcoma	X breakpoin	9.2
similar to H-2 class I histocompatibility antigen	D-37 alpha c	9.1
chemokine (C-X-C motif) ligand 14	Cxcl14	8.8
epithelial membrane protein 1	Emp1	8.6
insulin-like growth factor 2 mRNA binding protein 3	Igf2bp3	8.6
C-type lectin domain family 4	member n	8.5
vacuolar protein sorting 11 (yeast)	Vps11	8.4
matrix metallopeptidase 12	Mmp12	8.3
glycoprotein (transmembrane) nmb	Gpnmb	8.2
eosinophil-associated	ribonuclease	8.2
keratin 14	Krt14	8.0
major urinary protein 3	Mup3	7.9
on-SMC condensin I complex	subunit G	7.9
serum amyloid A 2	Saa2	7.7
pyruvate dehydrogenase kinase	isoenzyme 4	7.6
thyrotropin releasing hormone receptor	Trhr	7.5
secreted frizzled-related protein 2	Sfrp2	7.5
ankyrin repeat domain 22	Ankrd22	7.3
fatty acid binding protein 7	brain	7.3
serine (or cysteine) peptidase inhibitor	clade B	7.3
neurogenin 1	Neurog1	7.3
coagulation factor XIII	A1 subunit	7.2
procollagen	type IX	7.2
phosphodiesterase 6A	cGMP-speci	7.1
coiled-coil domain containing 120	Ccdc120	7.1
hemoglobin Y	beta-like e	7.1
neuropeptide Y receptor Y5	Npy5r	7.0
folate hydrolase	Folh1	6.8
RIKEN cDNA 1700029I01 gene	OTTMUSG0	6.8
cytochrome b-245	beta polypep	6.8
ATP-binding cassette	sub-family D	6.7
expressed sequence C77370	C77370	6.7
phosphatidylinositol-4-phosphate 5-kinase	type II	6.6
ferritin mitochondrial	Ftmt	6.6
DBF4 homolog (S. cerevisiae)	Dbf4	6.6
cholecystokinin B receptor	Cckbr	6.5
vacuolar protein sorting 29 (S. pombe)	Vps29	6.5
DNA segment	Chr 1	6.5
expressed sequence AA517858	AA517858	6.5
C-type lectin domain family 4	member e	6.5
visual system homeobox 1 homolog (zebrafish)	Vsx1	6.4
C-type lectin domain family 2	member h	6.4
G-protein-coupled receptor 50	Gpr50	6.4
serine peptidase inhibitor	Kazal type 3	6.3
chitinase 3-like 1	Chi3l1	6.3
formyl peptide receptor	related sequ	6.3
solute carrier family 39 (zinc transporter)	member 9	6.3
phosphatase	orphan 1	6.2
serine (or cysteine) peptidase inhibitor	clade B	6.2
guanine nucleotide binding protein (G protein)	gamma 4 su	6.0
immunoglobulin superfamily	member 6	6.0
orosomucoid 2	Orm2	6.0
ATP-binding cassette	sub-family	5.9
hydroxysteroid (17-beta) dehydrogenase 13	Hsd17b13	5.9
zinc finger protein 286	Zfp286	5.8
presenilin 1	Psen1	5.8
Glucosaminyl (N-acetyl) transferase 2	I-branching e	5.8
caspase 4	apoptosis-re	5.7
minichromosome maintenance deficient 6 (MIS5 homolog	S. pombe) (S	5.7
retinal G protein coupled receptor	Rgr	5.7
zinc finger protein 1	Zfp1	5.7
T-box brain gene 1	Tbr1	5.6
brain abundant	membrane a	5.6
NLR family	pyrin domain	5.6
Amyloid beta (A4) precursor protein	App	5.6
sperm associated antigen 5	Spag5	5.5
Mus musculus	clone IMAGE	5.5
dynamin 1-like	Dnm1l	5.5
capping protein (actin filament)	gelsolin-like	5.5
leucine rich repeat transmembrane neuronal 1	Lrrtm1	5.5
lipocalin 2	Lcn2	5.5
armadillo repeat containing	X-linked 4	5.4
sarcoglycan	gamma	5.3
immunoglobulin mu binding protein 2	Ighmbp2	5.3
phosphatidylinositol 3-kinase	C2 domain c	5.3
eukaryotic translation initiation factor 3	subunit H	5.2
chondroadherin	Chad	5.2
arylsulfatase A	Arsa	5.2
kallikrein 1-related peptidase b16	Klk1b16	5.2
Nik related kinase	Nrk	5.1
betacellulin	epidermal	5.1
transient receptor potential cation channel	subfamily M	5.1
pyrin and HIN domain family	member 1	5.1
lymphocyte antigen 6 complex	locus A	5.1
tubulin	beta 2b	5.1
RIKEN cDNA A530058N18 gene	A530058N18	5.1
solute carrier family 15	member 3	5.0
BCL2/adenovirus E1B interacting protein 1	NIP2	5.0

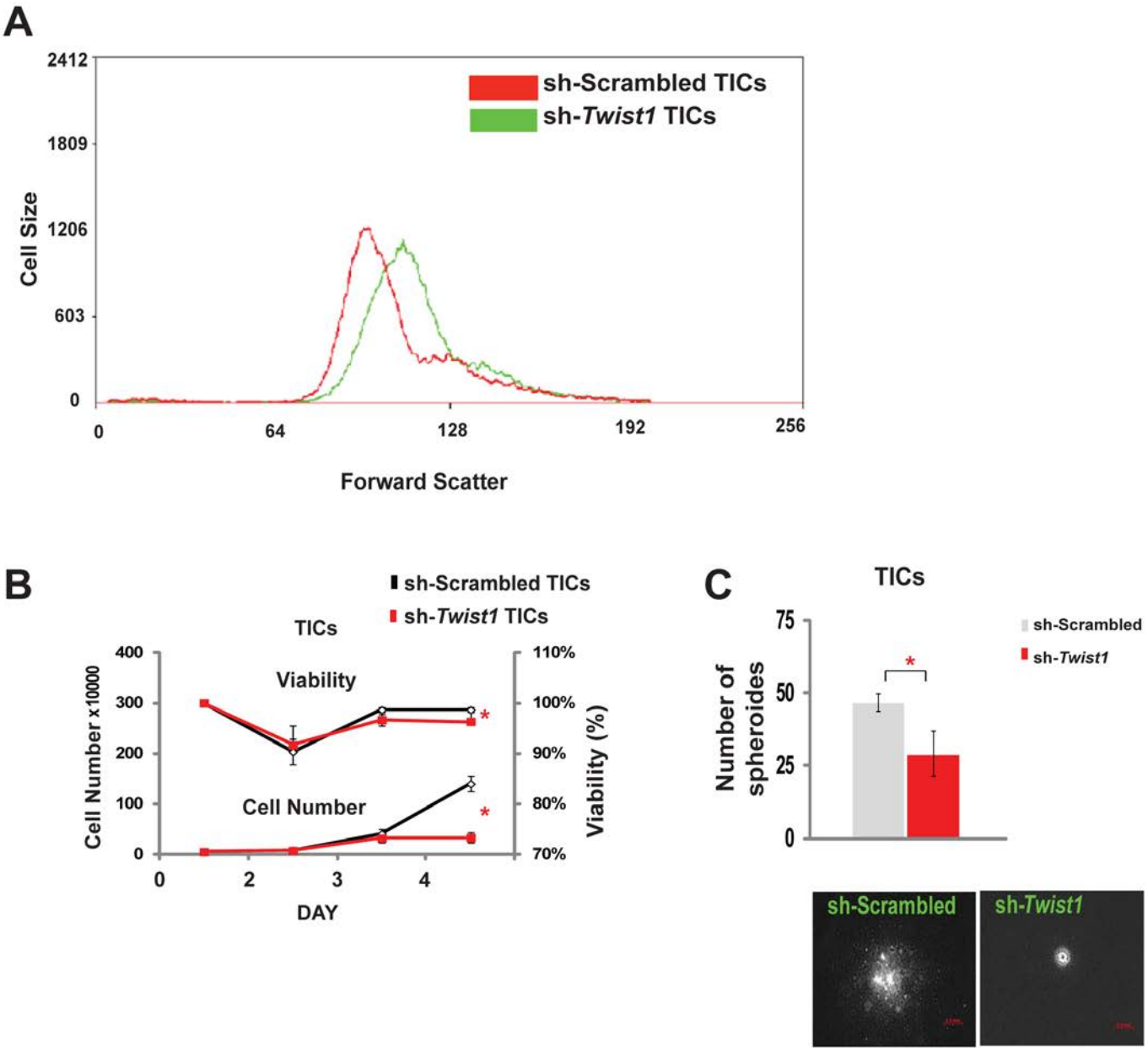
Official Gene Name	Official gene symbol	Fold change in gene expression
ATP/GTP binding protein-like 3	Agbl3	0.195
DPH5 homolog (S. cerevisiae)	Dph5	0.195
lipase	endothelial	0.195
ATP-binding cassette	sub-family B	0.194
leukemia inhibitory factor receptor	Lifr	0.193
bromodomain containing 4	Brd4	0.193
pancreatic lipase-related protein 2	Pnliprp2	0.192
RIKEN cDNA 4632419K20 gene	4632419K:	0.190
KH domain containing	RNA bindi	0.189
sortilin-related VPS10 domain containing receptor 2	Sorcs2	0.189
UDP-Gal:betaGlcNAc beta 1	3-galactosyltr	0.186
cell cycle exit and neuronal differentiation 1	Cend1	0.185
complement factor H-related 1	Cfhr1	0.184
FBJ osteosarcoma oncogene	Fos	0.184
transcription factor AP-2 beta	Tcfap2b	0.182
Metastasis associated lung adenocarcinoma transcript 1	Malat1	0.180
Microtubule-associated protein 1 light chain 3 beta	Map1lc3b	0.179
expressed sequence AA414993	AA414993	0.176
oxysterol binding protein-like 6	Osbpl6	0.174
Hedgehog-interacting protein	Hhip	0.173
keratin associated protein 16-1	Krtap16-1	0.172
integrin binding sialoprotein	Ibsp	0.171
lipocalin 4	Lcn4	0.170
B-cell leukemia/lymphoma 2 related protein A1a	Bcl2a1a	0.169
quiescin Q6 sulfhydryl oxidase 1	Qsox1	0.169
prokineticin receptor 1	Prokr1	0.169
ribosomal protein S19	Rps19	0.166
coatomer protein complex	subunit gam	0.166
annexin A2	Anxa2	0.163
immunoglobulin kappa chain variable 28 (V28)	butructural Ba	0.161
cholinergic receptor	nicotinic	0.157
Transcribed locus		0.157
vomeronasal 1 receptor	C8	0.156
RAD51 associated protein 1	Rad51ap1	0.154
protocadherin beta 14	Pcdhb14	0.152
myelin-associated oligodendrocytic basic protein	Mobp	0.152
nanos homolog 1 (Drosophila)	Nanos1	0.150
integrin alpha 4	Itga4	0.147
solute carrier family 22 (organic anion transporter)	member 7	0.142
chromatin licensing and DNA replication factor 1	Cdt1	0.140
Phosphoenolpyruvate carboxykinase 1	cytosolic	0.139
hepcidin antimicrobial peptide 1	Hamp1	0.139
androgen binding protein beta	Abpb	0.137
Phospholipase A2	group IB	0.137
procollagen	type V	0.136
recoverin	Rcvrn	0.135
cyclin B1	related sequ	0.135
doublecortin-like kinase 1	Dclk1	0.135
inactive X specific transcripts	Xist	0.130
tetratricopeptide repeat domain 16	Ttc16	0.129
RIKEN cDNA 1700016D06 gene	1700016D06	0.128
T-cell lymphoma invasion and metastasis 2	Tiam2	0.125
interferon alpha 4	Ifna4	0.124
defensin beta 10	Defb10	0.124
inactive X specific transcripts	Xist	0.116
vomeronasal 2	receptor	0.115
Predicted gene	EG629860	0.115
cytochrome P450	family 51	0.114
tripartite motif protein 16	Trim16	0.113
leptin receptor	Lepr	0.112
dolichyl-di-phosphooligosaccharide-protein glycotransferase	Ddost	0.111
fibroblast growth factor 4	Fgf4	0.110
defensin beta 5	Defb5	0.103
cDNA sequence BC014805	BC014805	0.097
serine/arginine-rich protein specific kinase 3	Srpk3	0.096
olfactory receptor 69	Olfr69	0.093
myosin	heavy polype	0.092
cDNA sequence BC014805	BC014805	0.090
paternally expressed 3	Peg3	0.088
hydroxy-delta-5-steroid dehydrogenase	3 beta- and s	0.088
oogenesin 1 /// similar to Oog1 protein	LOC1000390	0.087
RIKEN cDNA 1810022C23 gene	1810022C23	0.087
paternally expressed 3	Peg3	0.086
dopamine receptor 4	Drd4	0.083
Transcribed locus		0.082
cytochrome P450	family 51	0.081
solute carrier family 1	Scf1	0.076
Hedgehog-interacting protein	Hhip	0.075
cytochrome P450	family 51	0.071
regenerating islet-derived 3 gamma	Reg3g	0.071
cut-like 2 (Drosophila)	Cutl2	0.071
a disintegrin and metallopeptidase domain 3 (cyritestin)	Adam3	0.069
hepcidin antimicrobial peptide 1	Hamp1	0.067
Galactokinase 2	Galk2	0.067
vomeronasal 1 receptor	V1r	0.067
hepcidin antimicrobial peptide 2	Hamp2	0.061
paternally expressed 3	Peg3	0.058
myosin	heavy polype	0.049
hypothetical protein LOC100047923	LOC10004	0.048
killer cell lectin-like receptor	subfamily A	0.047
flavin containing monooxygenase 3	Fmo3	0.047
myosin	heavy polype	0.041
eukaryotic translation initiation factor 2	subunit 3	0.002

Supplementary Figure 2. RNA microarray analysis.



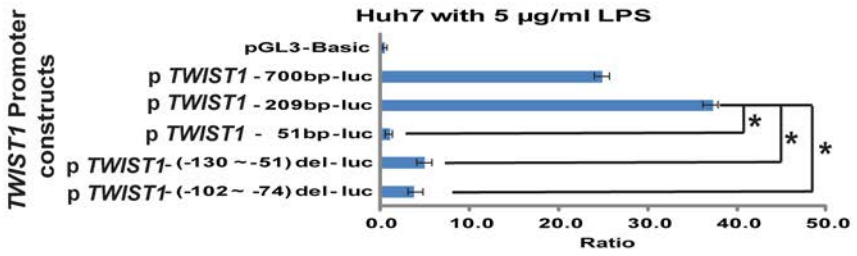
Supplementary Figure 3. Immunofluorescence staining on mouse liver sections for TLR4 and NANOG. Note: The TLR4 staining in the low fat diet (LFD) represents non-parenchymal cell (presumably such as Kupffer cells), whereas in liver sections from mice fed HCFD and in mice fed HCFD+LPS, the TLR4 is co-stained with NANOG representing that the origin of TLR4 in these livers is from both TICs and non-parenchymal cells. The scale bar indicates 50μm.

Suppl. Fig. 3
Uthaya Kumar et al.



Supplementary Figure 4. *Twist1* is required for the mesenchymal phenotypes, cell proliferation, and self-renewal abilities of TICs. (A) Flow cytometry analysis (forward scatter affected by cell size) indicating the change in cell size after *Twist1* knockdown in TICs (B) Analysis of cell number and viability post-infection with Lentivirus expressing sh-*Twist1* or sh-scrambled into TICS; significant decrease in both cell number and viability was seen with the infection of the former compared to the later. *P<0.05, n=5. (C) Sphere formation assay demonstrated the significant decrease in the number of spheroids formed whne the *Twist1* gene is silenced in TICs. *P<0.05, n=3.

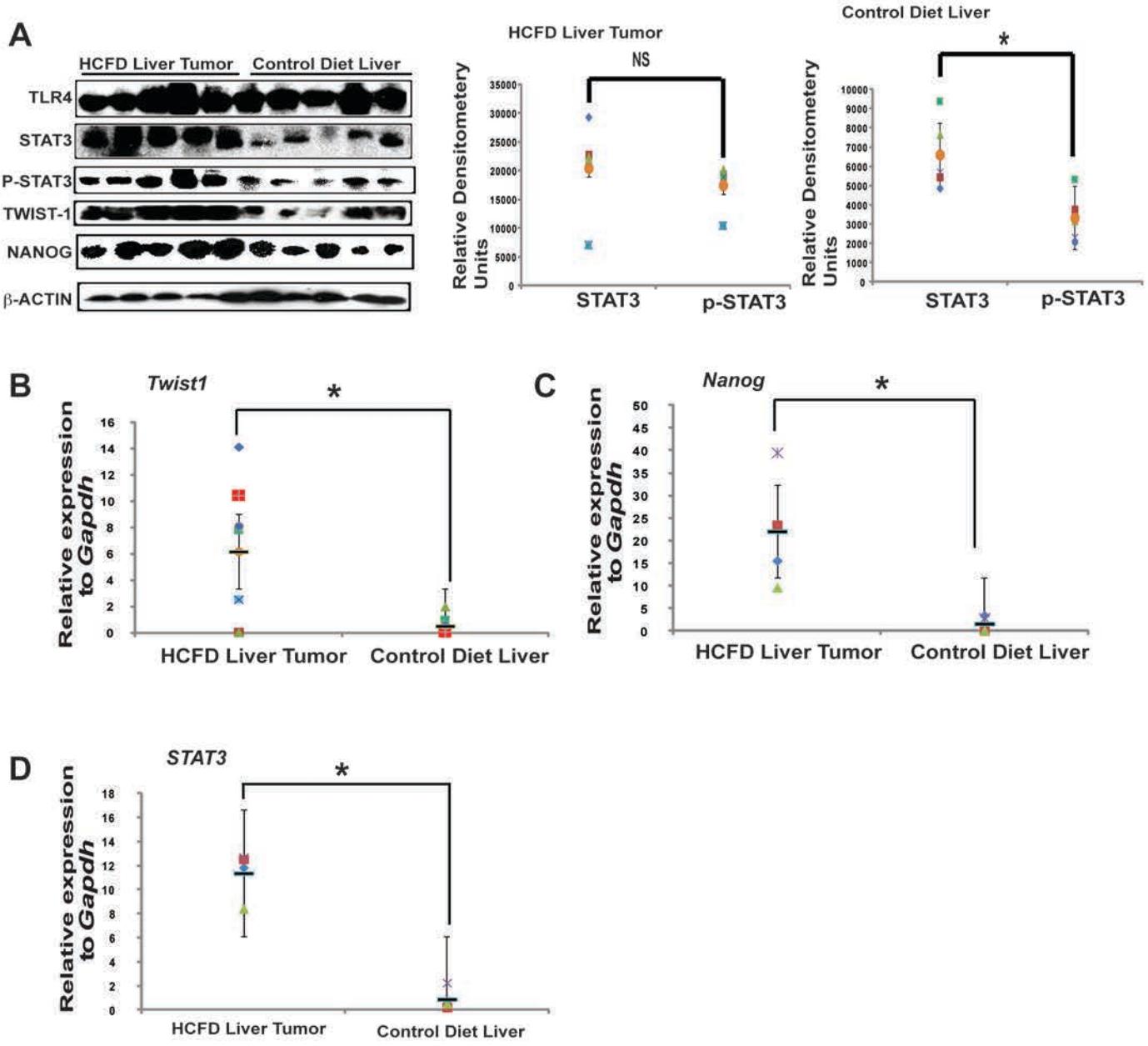
Suppl. Figure 4
Uthaya Kumar *et al.*



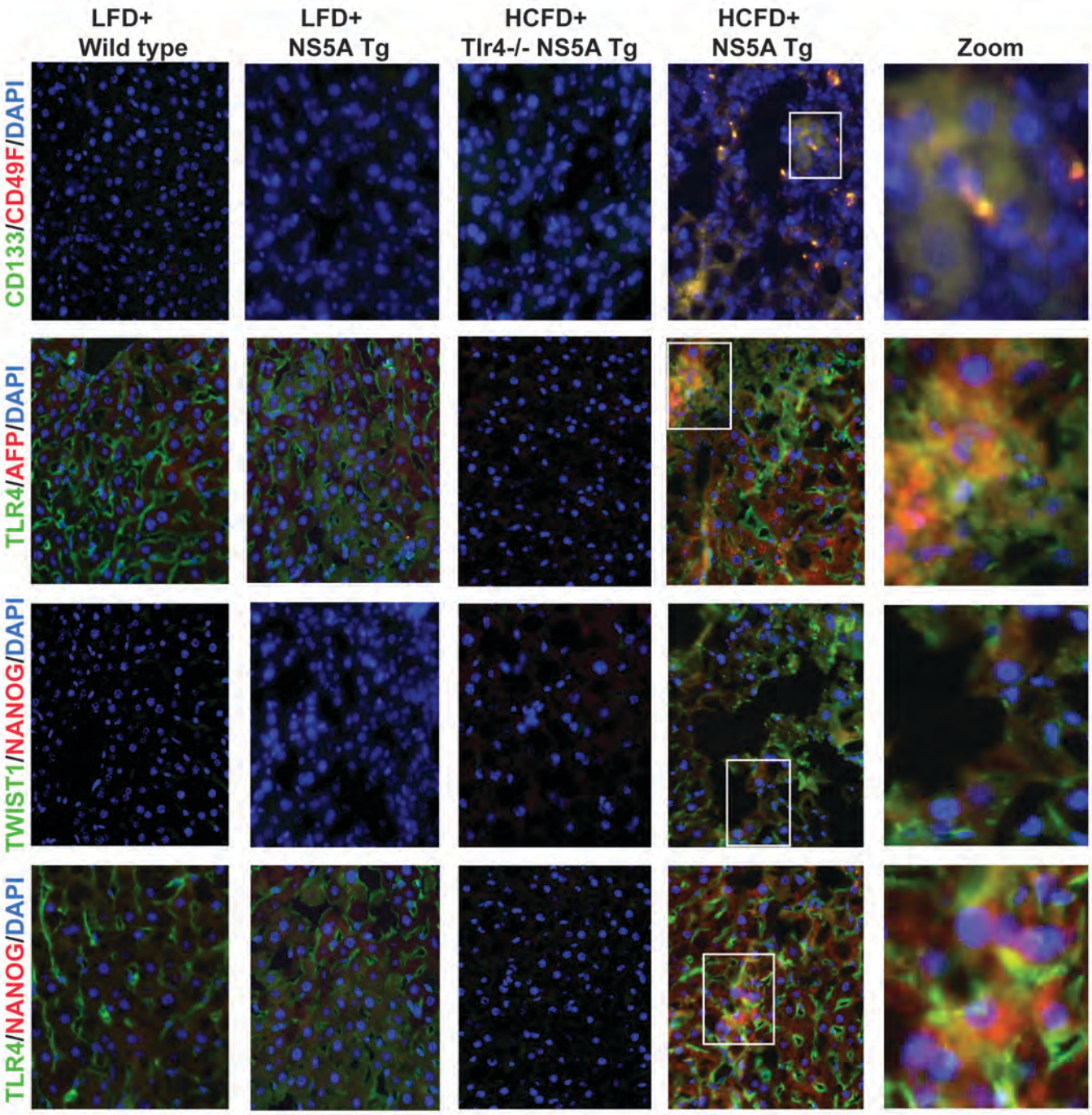
Supplementary Figure 5. *TWIST1* promoter activation in Huh7 cells. *TWIST1* promoter analysis with deletion constructs demonstrates the importance of the proximal segments (-209 to -51) in LPS-induced *TWIST1* promoter activity.

Suppl. Figure 5

Uthaya Kumar *et al*.

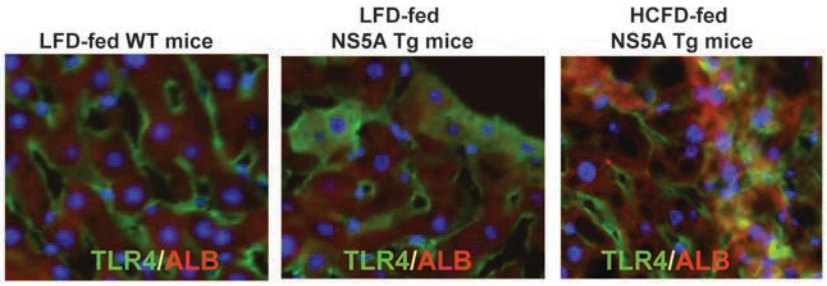


Supplementary Fig. 6. Induction of TLR4/NANOG/P-STAT3/TWIST1 pathway components in mouse HCC. (A) TLR4, STAT3, P-STAT3, NANOG and, TWIST1 protein levels are increased in HCC specimens from HCFD *NS5A* Tg mice, as compared with cirrhotic or healthy livers, N = 5 samples/cohort, n=3 (B) *Twist1* (C) *Nanog* (D) *Stat3* mRNA profiling showing significant increase in HCFD *NS5A* Tg mice Liver tumor, in contrast with the healthy livers, *p<.05, N=8 samples/cohort, n=3.



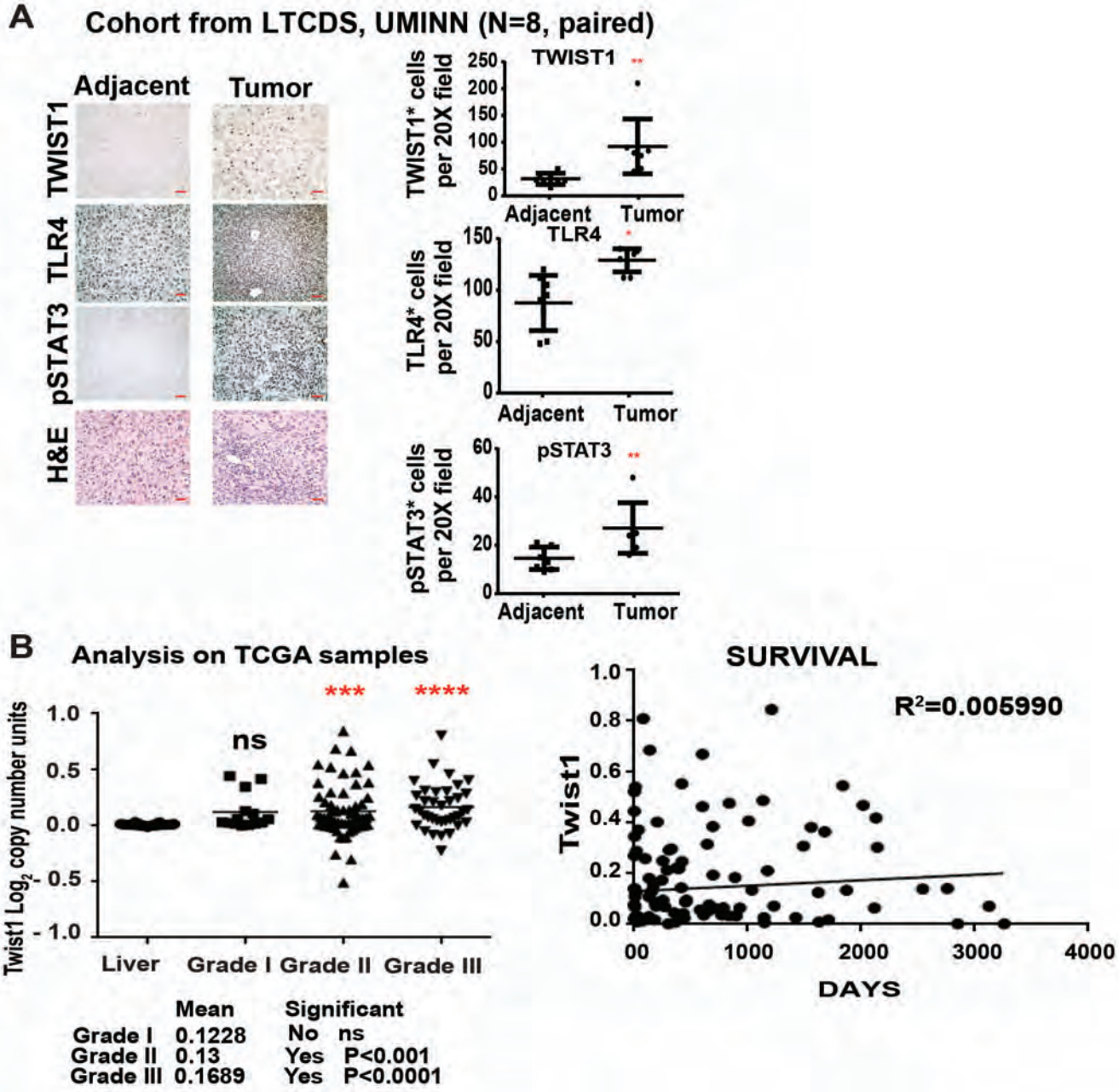
Supplementary Figure 7. Immunofluorescence staining of mouse liver sections for TLR4, NANOG, TWIST1, AFP, CD133 and CD49F. Note: Staining of TLR4, NANOG and TWIST1 is reduced in liver sections of *Tlr4* -/- NS5A Tg mice fed HCFD for 12 months in comparison to those of wild type and NS5A Tg mice or mice fed low fat diet (LFD). Parenchymal cells of NS5A Tg mice fed LFD have diffuse staining of TLR4 while hepatocytes of wild type mice fed LFD have TLR4 staining mainly in non-parenchymal areas. Note: Liver sections from NS5A Tg mice fed HCFD have co-expression (yellow) of TLR4 (green) and AFP (red) or NANOG (red) mainly in parenchymal cells (with larger nuclei) or TICs.

Suppl. Fig. 7
Uthaya Kumar et al.



Supplementary Figure 8. Immunofluorescence analysis of TLR4 and hapatocyte marker in NS5A Tg and wild type mice. The major source of TLR4 in the liver of wild type mice is from non-parenchymal cells, including the Kupffer cells and stellate cells. The low fat diet (LFD)-fed wild type mice have staining shows TLR4 positive cells, which are presumably Kupffer cells and stellate cells. Liver section of HCFD-fed NS5A Tg mice have TLR4 and NANOG-TLR4-double-positive cells, indicating that TLR4 origin is not only from Kupffer cells and stellate cells but rather from the TICs or hepatocytes. In liver of NS5A Tg mice, both parenchymal and non-parenchymal staining of TLR4 are positive while non-parenchymal area of wild type mice fed LFD mainly have positive staining of TLR4, indicating that hepatocytes and TICs of NS5A Tg mice have elevated levels of TLR4.

Suppl. Figure 8
Uthaya Kumar *et al*.



Supplementary Figure 9. Induction of TLR4, pSTAT3 and TWIST1 in human HCC. (**A**) Quantification of immunoperoxidase staining using Metamorph software showed increased staining intensity for Twist1, TLR4 and pSTAT3 in human liver tumors compared to the adjacent non-tumorous livers from Liver Tissue Cell Distribution System (LTCDS) of University of Minnesota (UMinn) (40X magnification; n=8 samples, paired). (**B**) Analysis of TCGA data of Twist 1 mRNA levedifferent stages of HCC grades and patient survival days with levels of TWIST1 expression.

Gender	Age at time of operation	Height	Weight	BMI	Primary Diagnosis Text	Total Cholesterol	AST	ALT	Necrosis	Bilobar	Tumor Stage "n" not stated	Vascular Invasion	Positive Margins	Intrahepatic metastasis	Tumor Grade 1=well 2=mod 3=poor 4=und n=can't assess b/c of necrosis	Multifocal (1=yes)	Microvascular Invasion	Macrovascular Invasion	steatohepatitis	Steatosis (>5%)	Cirrhosis	Fibrosis	Site of mets
M	68	1.76	89	28.73192	EtoH	142	471	457	1	1	T2	1	0	0	1	1	1	0	0	0	1	1	
F	66	1.52	73	31.59626	AIH	425	543	391	0	0	T2	1	0	0	1	0	1	0	1	1	1	1	
M	65	1.778	77.3	24.45209	Hep C	142	9812	3029	1	0	T2	0	0	0	2	1	0	0	0	0	1	1	
M	67	1.7	77	26.6436	Hep B	155	1548	815	0	1	T3a	0	0	0	1	1	0	0	0	0	1	1	
M	51	1.78	84	26.5118	Hep B	126	1467	489	0	0	T1	0	0	0	1	0	0	0	0	0	1	1	
F	59	1.65	81	29.75207	Hep C	210	824	889	0	1	T2	1	0	0	2	1	1	0	0	0	1	1	
M	67	1.78	71	22.40879	EtoH	n/a	2173	1407	0	0	T1	0	0	0	3	0	0	0	0	0	1	1	
M	64	1.63	67	25.21736	Hep C	99	413	200	1	0	T2	0	0	0	2	1	0	0	0	0	1	1	
M	52	1.68	53	18.77834	PrLiv Mal	193	2088	1019	0	0	T1	0	0	0	4	0	0	0	0	0	1	1	
M	71	1.75	83	27.10204	n/a	n/a	n/a	n/a	0	0	T2	1	0	0	n/a	0	1	1	0	0	0	0	
F	63	1.6	58	22.65625	PrLiv Mal	224	1880	1228	1	1	T2	0	0	0	2	1	0	0	0	0	1	1	
M	68	1.651	77.3	28.35864	Crypto	228	4992	1871	0	0	T1	0	0	0	3	0	0	0	0	1	1	1	
F	61	1.55	50	20.81165	Hep C	116	1269	683	0	0	T2	1	0	0	3	1	1	0	0	1	1	1	
M	61	1.83	89	26.57589	Hep C	162	1648	797	0	0	T1	0	0	0	3	0	0	0	0	0	1	1	
M	56	1.73	82	27.39818	Hep C	28	3452	878	0	0	T2	1	0	0	2	1	1	0	0	0	1	1	
M	59	1.65	62	22.77319	EtoH	137	1785	731	1	0	T3a	1	0	0	3	1	1	0	0	0	1	1	
M	69	1.83	105	31.35358	EtoH	138	1125	1057	0	1	T2	1	0	0	3	1	1	0	0	0	1	1	
M	62	1.7	93	32.17993	Hep C	173	490	219	1	1	T2	1	0	0	3	1	1	0	0	0	1	1	
M	65	1.8	81	25	Hep C	107	759	503	1	0	T1	1	0	0		0	1	0	0	0	1	1	
M	69	1.8	91	28.08642	Hep C	168	5142	1168	0	0	T3b	1	0	0	3	0	1	1	0	0	1	1	
M	66	1.73	87	29.0688	Hep C	122	5086	3287	0	1	T2	1	0	0	2	1	1	0	0	1	1	1	Lungs, ribs
M	65	1.7	72	24.91349	Hep B	201	649	488	0	0	T1	0	0	0	1	0	0	0	0	0	1	1	
M	60	1.7	72	26.84061	Hep C	117	2023	1722	1	1	T2	1	0	0	2	1	1	0	0	0	1	1	
M	66	1.85	79.5	23.22863	n/a	n/a	n/a	n/a	0	0	T1	0	0	2.8cm on S5	2	1	0	0	0	0	1	0	
M	68	1.79	86	26.84061	ALF	156	1009	400	0	0	T2	0	0	0	3	1	0	0	0	0	1	1	
M	68	1.83	80	23.88844	Hep B	180	765	444	1	1	T2	0	0	0	3	1	0	0	0	0	1	1	Abdominal nodes, liver
M	61	1.67	96	34.42217	Hep C	87	1418	1130	0	0	T2	1	0	0	3	0	1	0	0	0	1	1	
M	34	1.75	71	23.18367	Hep B	160	3635	911	1	1	T3a	1	0	0	3	1	1	0	0	0	1	1	
F	54	1.631	54.1	20.33711	Hep C	186	141	73	0	0	T3b	1	0	0	2	0	1	1	0	0	1	1	Lungs, r adrenal
M	48	1.699	69.1	23.93819	Hep C	184	8942	1014	0	0	T1	0	0	0	2	0	0	0	1	1	1	1	
M	64	1.78	86	27.14304	PrLiv Mal	97	7825	4690	1	0	T2	0	0	0	2	1	0	0	0	0	1	1	
M	74	n/a	75.29	n/a	n/a	n/a	n/a	n/a	0	0	T1	0	0	0	2	0	0	0	0	0	0	1	
M	46	1.73	81	27.06405	Hep C	178	1138	487	1	0	T1	0	0	0	2	0	0	0	0	0	1	1	
F	54	1.6	43	16.79688	PrLiv Mal	185	2279	2062	0	0	T3a	0	0	0	1	1	0	0	0	0	1	1	
F	46	1.52	70	30.29778	Hep C	144	2095	496	0	1	T2	0	0	0	4	1	0	0	0	0	1	1	
M	57	1.56	76	31.22945	Hep C	121	11564	1375	1	1	T3a	0	0	0	2	1	0	0	0	0	1	1	
M	60	1.7	67	23.18339	Hep C	129	4092	3838	1	1	T2	0	0	0	2	1	0	0	0	0	1	1	
M	73	1.8	68	20.98765	Hep C	130	2842	882	1	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	42	1.626	132.3	50.04017	Hep C	165	433	335	0	0	T3a	1	0	0	2	1	1	0	0	0	1	1	
F	71	1.52	81	35.05886	Hep C	151	4385	1047	1	1	T2	0	0	0	1	1	0	0	0	0	1	1	
M	53	1.78	102	32.1929	EtoH	164	21	29	1	0	T2	1	0	0	3	0	1	0	0	0	1	1	
F	66	1.52	77	33.32756	Hep C	89	4500	1288	1	0	T2	1	0	0	3	1	1	0	1	1	1	1	
M	49	1.95	65	17.09402	HCV	n/a	n/a	n/a	1	0	T3a	1	0		3	1	1	1	1	0	0	0	
M	67	1.88	102	28.85921	Hep B	48	402	413	0	1	T2	1	0	0	2	1	1	0	0	0	1	1	
M	62	n/a	57	n/a	Hep C	188	2792	320	1	0	T2	0	0	0	2	1	0	0	1	0	1	1	
M	75	1.77	93.4	29.81263	n/a	n/a	n/a	n/a	0	0	T3a	0	1	1	3	1	1	0	0	0	0	0	
M	65	1.7	80	27.68166	Hep B	82	1179	578	0	0	T2	0	0	0	2	1	0	0	1	1	1	1	
M	55	1.83	106	31.65218	Hep C	174	67	28	1	1	T3b	1	0	0	3	1	1	1	0	0	1	1	abdomen, lungs
M	49	1.78	76	23.98687	Hep C	95	284	596	0	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	58	1.8	93	28.7037	Hep C	270	262	386	0	0	T2	1	0	0	3	1	1	0	0	1	1	1	
M	60	1.77	79	25.21625	n/a	n/a	n/a	n/a	0	0	T1	0	0	0	3	0	1	0	0	0	0	0	
M	53	1.7	72	24.91349	Hep C	107	6735	2709	1	0	T2	1	0	0	3	0	1	0	0	1	1	1	
M	51	1.6	103	40.23438	Hep C	191	515	303	0	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	60	1.7	50	17.30104	PrLiv Mal	163	1573	810	1	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	56	1.72	73	24.6755	HBV	n/a	n/a	n/a	0	0	T1	0	0	0	2	0	0	0	1	0	0	1	lung
M	54	1.7	73	25.25952	Hep C	219	878	443	1	1	T2	0	0	0	2	1	0	0	0	1	1	1	
M	70	1.64	56	20.82094	HCV	n/a	n/a	n/a	0	0	T1	0	0	0	2	0	1	0	1	0	0	1	
M	52	1.68	59	20.9042	Hep B	244	2659	568	1	0	T3a	1	0	0	4	1	1	0	0	0	1	1	Liver, Lung, peritoneum
M	59	1.4	51	26.02041	Hep C	225	1771	450	1	0	T3a	1	0	0	2	1	1	0	0	0	1	1	
F	55	1.57	60.4	24.50404	HBV	n/a	n/a	n/a	0	0	T1	0	0	0	2	0	1	0	0	0	0	0	lung
M	64	1.85	90	26.29657	Hep C	118	306	93	0	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	56	1.73	91	30.40529	Hep C	101	831	254	0	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	48	1.68	62	21.96712	Hep C	122	372	305	1	0	T2	1	0	0	3	1	1	0	0	1	1	1	
M	55	1.75	80	26.12245	Hep B	213	1764	778	1	1	T3b	1	0	0	2	1	0	1	1	1	1	1	Lungs, liver, then with progression to L adrenal, pelvis
M	63	1.8	120	37.03704	Crypto	177	4207	1270	0	0	T2	0	0	0	3	1	0	0	1	1	1	1	
M	59	1.65	74	27.1809	Crypto	199	4620	2386	0	1	T4	1	0	0	3	1	1	1	1	1	1	1	Pelvic mass, lung lesions with innumerable liver lesions, carcinomatosis in 1 years time
M	53	1.72	80	27.04164	Hep B	184	1235	684	0	1	T3a	1	0	0	3	1	1	0	0	0	1	1	progresive spinal mets
M	67	n/a	n/a	n/a	Misc	n/a	8989	2085	0	0	T2	1	0	0	2	1	1	0	0	0	1	1	
M	61	1.78	70	22.09317	PrLiv Mal	166	925	726	0	1	T3b	1	0	0	2	1	1	1	0	0	1	1	Liver, lungs, iliac bone
M	55	1.72	94	31.77393	Hep C	166	3181	1707	1	0	T2	1	0	0	2	0	1	0	0	0	1	1	
M	67	1.7	93	32.17993	NASH	234	1217	994	0	0	T1	0	0	0	2	0	0	0	1	0	1	1	
M	57	1.83	88	26.27729	Hep C	150	337	200	1	1	T2	0	0	0	2	1	0	0	0	1	1	1	
M	47	1.54	52.8	22.26345	HBV	n/a	n/a	n/a	0	0	T1	0	0	0	3	0	1	0	1	0	0	0	
F	75	1.7	96	33.21799	Hep C	127	1099	276	1	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	53	1.78	145	45.76442	Hep C	n/a	2391	876	1	0	T2	0	0	0	2	1	0	0	0	1	1	1	
F	63	1.7	54	18.68512	Hep C	172	1002	746	1	0	T1	0	0	0	3	0	0	0	0	1	1	1	
F	72	1.75	79	25.79592	Hep C	185	494	670	1	1	T2	0	0	0	2	1	0	0	0	0	1	1	
M	51	1.63	65	24.4646	Hep C	226	5447	989	1	1	T3a	1	0	1	2	1	1	0	0	0	1	1	
M	57	1.88	137	38.76188	Hep C	144	740	528	1	1	T3b	1	0	0	2	1	0	1	0	0	1	1	Lungs, Brain
M	52	1.70	99.7	34.49827	HBV	n/a	n/a	n/a	0	0	T1	0	0	On Rt.lobe	3	On Rt.lobe	1	0	1	0	0	1	
M	68	1.6	64	25	Hep B	153	592	274	1	1	T2	0	0	0	2	1	0	0	0	0	1	1	
M	55	1.72	82	27.71769	Hep C	124	541	349	0	0	T2	0	0	0	3	1	0	0	0	0	1	1	
M	52	1.78	85	26.82742	Hep C	118	2559	1802	1	0	T3b	1	0	0	3	0	1	1	0	0	1	1	
F	54	1.65	55	20.20202	Hep B	141	1819	1018	1	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	61	1.536	62	26.27903	Hep C	143	2662	1228	1	0	T1	0	0	0	2	0	0	0	0	0	1	1	
F	36	1.67	72.5	25.99591	n/a	n/a	n/a	n/a	0	0	T1	0	0	0	2	0	0	0	0	0	0	0	
F	59	1.55	65	27.05515	PrLiv Mal	210	418	91	1	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	52	1.78	82	25.88057	Hep B	165	346	265	1	1	T2	0	0	0	3	1	0	0	0	0	1	1	
M	55	1.75	90	29.38776	Hep C	196	764	581	1	0	T1	0	0	0	3	0	0	0	0	0	1	1	
M	64	1.7	92	31.83391	Hep C	146	1169	480	0	1	T3a	0	0	0	2	1	0	0	0	0	1	1	Periesophageal, hilar nodes in lung, bone
M	50	1.64	61	22.67995	HBV	n/a	n/a	n/a	0	0	T3a	0	0	0	3	1	1	0	0	0	0	0	chest wall
M	45	1.80	95	29.32099	HBV,HCV	n/a	n/a	n/a	0	0	T1	0	0	0	2	0	1	0	0	0	0	0	n/a
M	50	1.68	96	34.01361	NASH	154	1311	342	0	0	T1	0	0	0	2	0	0	0	1	1	1	1	
M	50	1.82	89	26.86874	Hep C	116	196	72	0	0	T1	0	0	0	2	0	0	0	0	0	1	1	
F	72	1.52	69	29.86496	Crypto	113	1303	215	0	0	T2	0	0	0	2	1	0	0	0	0	1	1	
F	61	1.68	85	30.11621	NASH	149	1600	451	0	0	T1	0	0	0	2	0	0	0	0	0	1	1	
M	61	1.75	74	24.16327	Hep C	124	1062	872	0	0	T1	0	0	0	3	0	0	0	1	1	0	0	
M	55	1.75	75	24.4898	Hep C	174	329	86	1	0	T2	0	0	0	2	1	0	0	0	0	1	1	
M	77	1.67	58	20.79673	HCV	n/a	n/a	n/a	0	0	T2	0	0	0	2	1	0	0	0	0	0	0	
M	60	1.75	85	27.7551	NASH	222	3117	1592	0	1	T4	1	0	0	2	1	1	1	1	1	1	1	Liver, lungs, mediastinum
M	61	1.6	59	23.04688	EtoH	106	774	324	1	1	T2	1	0	0	3	1	1	0	0	1	1	1	
M	56	1.6	77	30.07813	Hep C	n/a	949	173	1	1	T2	0	0	0	2	1	0	0	0	0	1	1	
F	37	1.63	51	19.1953	Hep B	260	961	452	0	1	T3b	1	0	0	2	1	1	1	0	1	1	1	Peritoneum
M	53	1.92	116	31.46701	Hep C	70	1097	402	1	0	T2	0	0	0	2	1	0	0	0	0	1	1	
M	60	1.78	80	25.24934	Hep C	147	769	369	1	0	T3b	1	0	0	2	1	1	1	0	0	1	1	
M	58	1.65	69	21.67126	Hep C	102	375	134	1	0	T2	0	0	0	3	1	0	0	0	1	1	1	
M	69	1.8	105	32.40741	NASH	133	241	55	0	0	T1	0	0	0	1	0	0	0	1	1	1	1	
F	58	1.65	65	23.87511	Hep C	146	1963	1026	1	0	T3a	0	0	0	2	1	0	0	0	0	1	1	
M	56	1.68	140	49.60317	EtoH	104	11802	1200	0	0	T2	1	0	0	2	0	1	0	0	0	1	1	
M	54	1.68	88	31.17914	Hep C	213	575	269	1	0	T1	0	0	0	1	0	0	0	0	0	1	1	
M	73	1.57	56	22.71897	Hep C	216	1228	531	1	0	T1	0	0	0	3	0	0	0	0	0	1	1	
M	65	1.78	99	31.24605	EtoH	105	464	79	1	0	T1	0	0	0	2	0	0	0	0	1	1	1	
M	53	1.67	90	32.27079	HBV	n/a	n/a	n/a	0	0	T3a	0	0	0	n/a	1	0	0	0	0	0	0	

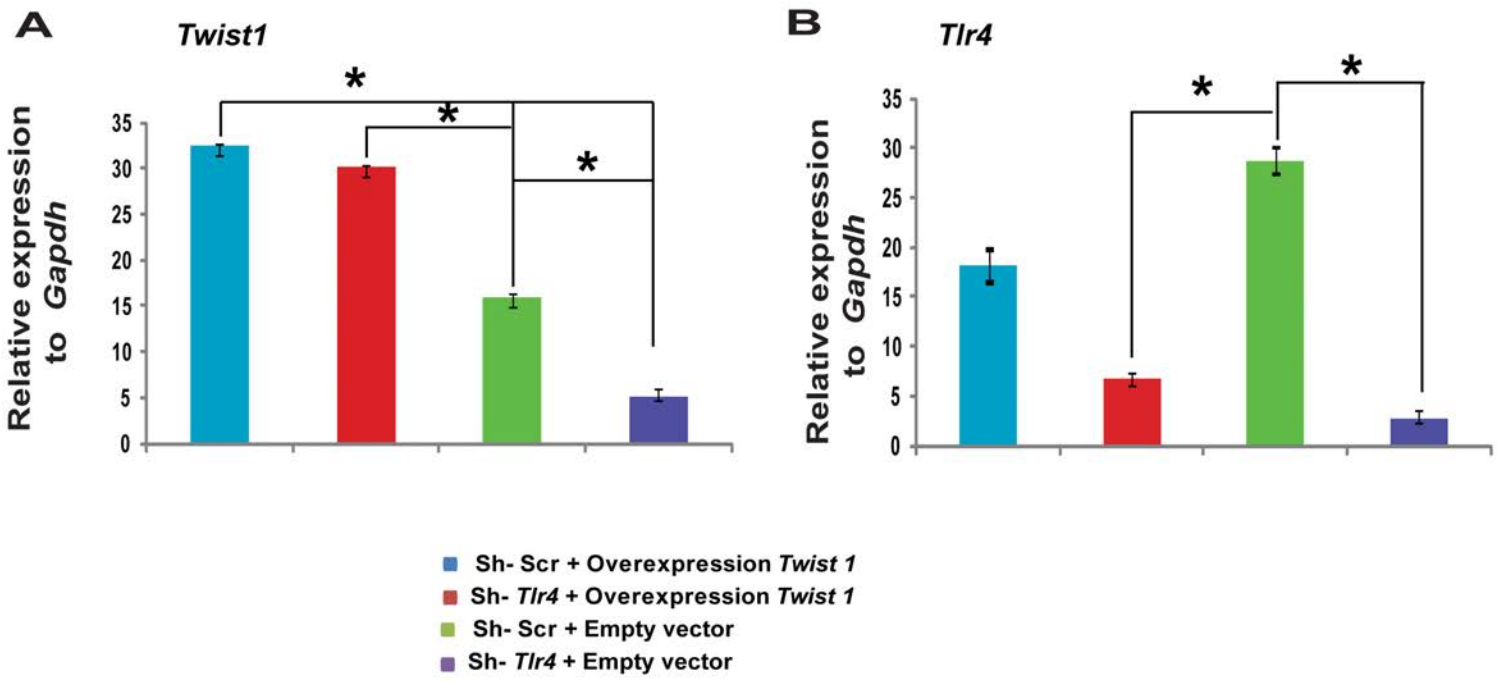
Supplementray Fig. 10. Clinical patient dataset summary used in this study.

Suppl. Fig 10
Uthaya kumar *et al*.

Supplemental Fig. 10. Clinical patient dataset summary used in this study.

Generally 1=present, 0=absent unless otherwise indicated.
Primary diagnosis text=clinical indication for transplantation
Necrosis=tumor necrosis
PrLivMal=primary liver malignancy
Crypto=cryptogenic cirrhosis

Suppl. Figure 10
Uthaya Kumar *et al*.



Supplementary Fig.11.*Twist1* by its very nature promotes tumor formation. (A) *Twist1* mRNA was analyzed using qRT-PCR in TICs post *Tlr4* silencing and *Twist1* overexpression, n=3. (B) *Tlr4* mRNA was analyzed using qRT-PCR in TICs post *Tlr4* silencing and *Twist1* overexpression, n=3.

Suppl. Fig. 11

Uthaya Kumar *et al.*

Supplementary Table 1: Clinicopathological Features of Patients with Hepatocellular Carcinoma

Feature	Value
Age (years)	59±8.51* (34-77)
Gender	
Male	93
Female	20
BMI	27±0.05* (19-50)
Total Cholesterol	157±0.48* (48-425)
AST	2124±21.74* (21-11802)
ALT	843±7.25* (55-3838)
Primary Diagnosis	
PrLive Mal	7
NASH	5
Crypto	4
HBV	21
HCV	60
Tumor stage	
T1	40
T2	46
T3a	16
T3b	9
T4	2
Tumor Grade	
1	19
2	64
3	35
4	3
Metastasis	16

* Value = Mean ± SEM

Supplementary Table 2. Liver histological grading of HCV NS5A Tg mice fed the low-fat diet, HCFD or HCFD+LPS for 12 months.

	Diet	Fatty liver (0-4+)	Spotty necrosis (0-2+)	Dysplasia (0-4+)	Inflammation (0-2+)
WT	Low fat diet	0.1 ± 0.1	0.1 ± 0.2	0.2 ± 0.1	0.2 ± 0.3
NS5A Tg	Low fat diet	0.1 ± 0.2	0.1 ± 0.1	0.1 ± 0.1^{NS}	0.2 ± 0.2^{NS}
Tlr4 –/–	Low fat diet	0	0.1 ± 0.3	0.1 ± 0.4^{NS}	0.1 ± 0.2^{NS}
Tlr4 –/– NS5A Tg	Low fat diet	0.1 ± 0.1	0.2 ± 0.1	0.2 ± 0.3^{NS}	0.2 ± 0.3^{NS}
WT	HCFD	2.3 ± 0.8	0.6 ± 0.2	0.2 ± 0.1^{NS}	0.6 ± 0.4^{NS}
NS5A Tg	HCFD	3.5 ± 0.6	1.1 ± 0.3*	1.4 ± 0.5*/*	1.2 ± 0.4*/*
Tlr4 –/–	HCFD	1.0 ± 0.7	0.4 ± 0.3	0.1 ± 0.2$^{NS/NS}$	0.4 ± 0.4$^{NS/NS}$
Tlr4 –/– NS5A Tg	HCFD	1.2 ± 0.8	0.5 ± 0.5	0.3 ± 0.5$^{NS/NS}$	0.5 ± 0.2$^{NS/NS}$
WT	HCFD + LPS	2.7 ± 0.7	0.7 ± 0.3	0.3 ± 0.3$^{NS/NS}$	0.7 ± 0.3$^{NS/NS}$
NS5A Tg	HCFD + LPS	3.7 ± 1.1	1.3 ± 0.5*	1.8 ± 0.6**/**	1.7 ± 0.6**/**
Tlr4 –/–	HCFD + LPS	1.2 ± 0.5	0.3 ± 0.3	0.2 ± 0.6$^{NS/NS}$	0.2 ± 0.4$^{NS/NS}$
Tlr4 –/– NS5A Tg	HCFD + LPS	1.5 ± 0.7	0.3 ± 0.6*	0.5 ± 0.4$^{NS/NS}$	0.4 ± 0.5$^{NS/NS}$

*P < 0.05, compared to respective HCFD diet-fed WT mice; **P < 0.05, compared to respective HCFD diet-fed NS5A Tg mice; $^{\#}P$ < 0.05, compared to HCFD plus LPS-fed WT mice. Fatty liver, 2+: 25%~50% heaptocytes with fat; 3+: 50%~75% with fat; 4+: >75% with fat. Inflammation, 1+: lesions encompassing less than 1/3 acinus; 2+: lesions larger than whole acini.

(WT-HCFD; *, P<0.05 **, P<0.01 ***, P<0.005, green scripts and symbols – statistical analysis in comparison to low fat diet (LFD), purple scripts and symbols - statistical analysis in comparison to HCFD)

Supplementary Table 3: Antibody list

Antibody	Manufacturer
TWIST1	SC-15393 (Santa Cruz Bio Technology)
NANOG	ab80892 (Abcam)
STAT3	#9139S (Cell Signalling)
P-STAT3	#9134S (Cell Signalling)
TLR4	SC-10741 (Santa Cruz Bio Technology)

TAK1	SC-7162 (Santa Cruz Bio Technology)
TRAF6	SC-7221 (Santa Cruz Bio Technology)
IKK-B	SC-8014 (Santa Cruz Bio Technology)
P-IKK-B	#2694 (Cell Signalling)
B-ACTIN	A5441 (SIGMA)

Supplementary Table 4: qRT-PCR primer sets

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
Twist-1	AGA TGT CAT TGT TTC CAG AGA	TTA GTT ATC CAG CTC CAG AGT
Nanog	AGG GTC TGC TA TGA GAT GCT	CAA CCA CTG GTT TTT CTG CCA
Stat3	GCC ACG TTG GTG TTT CAT AAT C	TTC GAA GGT TGT GCT GAT AGA G
Tlr4	ATG GCA TGG CTT ACA CCA CC	GAG GCC AAT TTT GTC TCC ACA
E-cad	CTG CTG CTC CTA CTG TTT CTA C	TCT TCT TCT CCA CCT CCT TCT
N-cad	CAG GGT GGA CGT CAT TGT AG	AGG GTC TCC ACC ACT GAT TC
Gapdh	TGG ATT TGG ACG CAT TGG TC	TTT GCA CTG GTA CGT GTT GAT

Supplementary Table 5: *In vitro* mutagenesis primer sets

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
Nanog- mut1 Proximal	GTT TGG GAG GAC GAA GGA GAC CCC GAG GAA GG	CCT TCC TCG GGG TCT CCT TCG TCC TCC CAA AC
Nanog- mut 2 Distal	AGG TCG TTT TTG CCT GGT TTG GGA GGA CG	CGT CCT CCC AAA CCA GGC AAA AAC GAC CT
Stat3- mut1 Proximal	TTT CCT ATA AAA CAT GAT TAC GTC CCT CCT CCT CAC G	CGT GAG GAG GAG GGA CGT AAT CAT GTT TTA TAG GAA A
Stat3- mut 2 Distal	CTG GAA AGC GGA AAC TAT GAT TAC GAA CTT CGA AAA GTC CC	GGG ACT TTT CGA AGT TCG TAA TCA TAG TTT CCG CTT TCC AG

Supplementary Table 6: ChIP-qPCR primer sets

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
Nanog- Proximal Binding Site	ATG GTT TGG GAG GAC GAG TTA	AAA GTT TCC GCT TTC CAG TCC
Stat3 - Distal Binding Site	GGA CTG GAA AGC GGA AAC T	GCA GAC TTG GAG GCT CTT ATA C
Stat3 – Proximal Binding Site	GCC AGG TCG TTT TTG AAT GG	CGT GCA GGC GGA AAG TTT GG

Specificity control -1	CCC AGC AAT CCC AAA TCG G	CAG CAA TGG CAA CAG CTT CTA
Specificity control -2	CTC ACG TCA GGC CAA TGA	GAG AGC TGC AGA CTT GGA G

SUPPLEMENTARY DISCUSSION

We demonstrated a synergistic interaction between alcohol consumption and HCFD, resulting in the highest observed tumor incidence in NS5A Tg mice (Fig. 2A). Additionally, a classical TLR4 activation was observed through canonical TAK-1, TRAF6 and pIKK-β signaling in both the HCFD- as well as HCFD - fed NS5A Tg mouse models (Figs. 1). We observed from RNA microarray analysis that *Twist1* was increased 11.9-fold in NS5A Tg mice fed HCFD (Fig. 2A). Long-term treatment of mice with HCFD activated *Tlr4-Nanog* signaling (Fig. 2D) and increased leptin and endotoxin levels in the plasma (Fig. 1B). A previous RNA microarray analysis of tissues from alcohol fed NS5A Tg mice[1] did not exhibit *Twist1* induction. These results led us to hypothesize that the adipose tissue-derived leptin-pSTAT3 axis and the TLR4-NANOG axis are needed for activation of *Twist1* in TICs. Consequently we analyzed the *Twist1* promoter for the functional importance of NANOG and pSTAT3 binding sites (Fig. 4). Our experiments showed that relative to the TSS, both NANOG proximal and STAT3 distal sites were required for maximum response to leptin and LPS stimulation, respectively. We postulate that this finding might be due to formation of a transcription complex comprised of these two DNA binding proteins on the *Twist1* promoter allowing contiguous stacking of these two trans-acting proteins.

In support of such a functional model, Watt *et al*., showed that *Nanog* interacts with *Stat3* to regulate its own gene expression.[2] Building upon their research, we further established through sequential-ChIP-qPCR analysis (Fig. 4E) that these two transcription factors indeed interacted with one other to transactivate *Twist1*. The *in vitro* data were corroborated in mice and human tissue sections, where we demonstrated by IHC and IF that TWIST1 co-localized with TLR4, P-STAT3 and NANOG. Nevertheless, future experiments are warranted to understand how these transcription factors activate *Twist1.* Potential mechanisms could be histone modifications in the *Twist1* promoter or enhancer regions.[3] Master regulators involved in EMT during wound healing process have a robust expression of poised enhancer marks. This is to methodically shift the cells to the native state post remodeling. An understanding of such epigenetics marks in HCC associated TICs and specifically targeting the epigenetics marks is needed in both mouse and patient derived models.

Moreover, we observed that over-expression of *Twist1* in the absence of *Tlr4* can independently drive tumor formation and metastasis (Fig. 7) which underscores the importance of various TLR4 dependent oncogenic pathways. We speculate that this phenomenon might be due to basal level expression of *Tlr4* after shRNA treatment.

SUPPLEMENTARY MATERIALS AND METHODS

Isolation of mouse TICs using FACS

Tumor-initiating stem-like cells (TICs) were isolated from liver tumors in HCV-NS5A transgenic mice fed *ad lib* with an ethanol-containing liquid diet high in cholesterol and saturated fat (HCFD) (as previously described).[4] Briefly tumors were surgically resected and mechanically

dissociated by scissors. The tissue homogenate was digested with collagenase IV (BD Biosciences) and dispase (Sigma) mixture by incubation at 37°C for 2 hours. The resulting single cell suspensions were incubated with anti-CD133, anti-CD49f and anti-CD45 antibodies and separated using FACS sorting, according to the manufacturer's protocol as previously described[4]. Isolated TICs were maintained in Dulbecco's modified Eagle's medium nutrient mixture F-12 (DMEM/F12) containing 10% fetal bovine serum (FBS), 1% nucleosides, 1 μM dexamethasone, epidermal growth factor (EGF), 1 μg/ml penicillin, 1 μg/ml streptomycin and 1% nonessential amino acids (NEAA). $CD133^+$ TICs and $CD133^-$ control cells were cryopreserved in 60% FBS, 20% DMEM/F12, and 20% DMSO.

Plasmids, production and propagation of lentivirus and retrovirus vectors

The NS5A expression plasmid was constructed by inserting HCV-NS5A cDNA downstream of the CMV promoter into pcDNA3.1 (Invitrogen). All retroviruses were based on lentivirus (pPAX2: Addgene) or MMTV vectors (pVPack-GP: Stratagene). Lentivirus vectors were prepared by standard procedures using HEK293T cells. The packaging vector pPAX2 (Addgene), amphotropic envelope gene (VSV glycoprotein), packaging vector expressing GAG-POL: pMDV (Addgene), and shRNA expression cassette were co-transfected into HEK293T cells using BioT transfection reagent (Bioland Scientific LLC). Retroviruses expressing *Stat3C* and *Stat3D* were obtained from Prof. Daniel C. Link (Washington University of School of Medicine).[5] Retroviruses expressing *Stat3C* and *Stat3D* were produced using Phoenix cells/HEK293T.[6] 48 hours post transfection, the virus containing, cell supernatants were harvested, purified, mixed with polybrene (4 μg/ml), and used to infect cells (Huh7 / TICs). The lentivirus titers were determined using LentiX-gostick (Clontech). Human GIPZ lentiviral shRNAmir target gene set was used for human toll-like receptor 4 (TLR4) (RHS4531-NR_024169, RHS4430-98525129, RHS4430-98843572, and RHS4430-99137800) (Open Biosystems). To increase silencing effects and to reduce off-target effects, a combination of shRNA lentiviruses were used to knock down target genes. MOI was calculated on a case by case basis depending on empirical transduction efficiency. The TWIST1-pGL3 reporter constructs were obtained from Prof. Nakamura (Tokyo Medical and Dental University).[7]

Tumor collection and analysis

Tumor-bearing animals were sacrificed at day 30 or 35 (depending on the cell number injected) or whenever the tumor size exceeded the limit, and tumors were collected and measured for volume and weight. The tumor tissues were divided for (1) fixation with neutrally buffered 10% formalin for H&E staining and histological evaluation of the tumor; (2) fixation with 4% paraformaldehyde followed by sucrose treatment for subsequent immunostaining; and (3) snap-freezing in liquid N_2 for mRNA and protein analysis.

Endotoxin measurement

For endotoxin measurements, blood was collected from the inferior vena cava with pyrogen-free heparin as previously described.[8] Extreme precautions were taken to avoid or eliminate pyrogen and endotoxin contamination of all surgical instruments and laboratory supplies. Blood samples were transferred into appropriate glass tubes made pyrogen-free by heat-treatment at 180°C for 24 hours. Pyrogen-free water was supplied by the manufacturer (Kinetic-QCL, Santa Clara, CA; Biowhittaker). Just prior to assay, plasma samples were diluted and heated to 75°C for 10 minutes to denature endotoxin-binding proteins that can mask endotoxin detection. Levels of endotoxin were measured using the Limulus amebocyte lysate pyrogen test and a kinetic program (Kinetic test, Kinetic-QCL, Santa Clara, CA; Biowhittaker). The threshold of detection is

0.1 pg/ml.

Histology & immunohistochemistry

Tissue samples were either fixed in 10% neutral buffered formalin and cryopreserved (Cryomatrix™) or with 4% paraformaldehyde (PFA) and embedded in paraffin, followed by thin-sectioning and mounted on glass slides. Samples were stained with either hematoxylin & eosin (H&E) or processed for immunostaining as appropriate. For the latter, primary antibodies against NANOG (Rabbit ab80892, Abcam), pSTAT3 (Rabbit #9134, Cell Signaling technology), TLR4 (Mouse monoclonal antibody, SC293072, Santa Cruz), TLR4 (goat sc-8694, Santa Cruz Biotechnology), or TWIST1 (Rabbit polyclonal antibody, sc-15393, Santa Cruz Biotechnology) were used along with their respective secondary antibodies. Slides were mounted for microscopy using xylene based mounting media, which includes hematoxylin for nuclei counterstaining (Vector Laboratories), as per the manufacturer's recommendations. The stained samples were then subjected to morphometric analysis. To determine the specificity of IHC, serial sections were similarly processed except primary antibodies were omitted. The area of interest was quantified using Metamorph software. The data shown represent the means ± SD.

Quantitative real-time PCR (qRT-PCR)

Total RNA was extracted by using TRIzol Reagent (Invitrogen) and purified using the RNeasy mini kit (QIAGEN) according to the manufacturer's protocol. RNA concentration and purity were determined by A_{260} and A_{260}/A_{280} ratios, respectively. The RNA samples were treated with DNase I (Invitrogen) to remove residual traces of DNA. cDNA was obtained from 1 µg of total RNA, using SuperScript III reverse transcriptase (Invitrogen) and random primers in a final volume of 10 µl. cDNAs were amplified by PCR using the primer pairs listed in the Supplementary Table 4. Quantitative real-time PCR was performed on an ABI 7300 HT Real-Time PCR machine using 2X SYBR Green Master Mix (Applied Biosystems). Conditions for all reactions: 1 cycle at 50°C for 2 min, followed by 1 cycle at 95°C for 10 min, followed by 40 cycles at 95°C for 15 s and 60°C for 1 min. Specificity of the PCR products were tested by thermal dissociation curves. Gene expression was determined as relative ratio to β-Actin or GAPDH control via the ΔCt method. The data shown represents the means ± standard deviation (SD).

Gene array analysis of liver tumors

For identifying anti-apoptotic or proto-oncogenic proteins, we prepared serial cytosections of the mice liver tissues, stained them with H&E, and collected hepatocytes under homeostatic conditions, dysplastic, or transformed morphology by using laser-capture microscopy as described[9-11]. In order to identify changes associated with HCFD, comparative analysis were performed on the cells isolated from livers of mice fed HCFD. A gene microarray analysis requires a minimum of 100-200 cells and proteomic analysis requires approximately 50,000-100,000 cells for each cell phenotype[12]. The cells were lysed for RNA or protein extraction for gene chip analysis and 1D gel MS/MS analysis[9, 10, 12-14]. The cells collected from each group of three animals were isolated for RNA or protein individually and later combined to create a representative sample pool and provide sufficient amounts of material for analysis. For gene profiling, the Affymetrix mouse gene chip (GeneChip Mouse Genome 430A 2.0) was used, and analyses were performed in the Genome Core Facility at Los Angeles Children's Hospital. Five individually extracted, mouse liver RNA specimens were pooled for each experimental group for microarray analysis. Data analysis was performed by using Partek Pro 5.1 (Partek Inc.). The

normalization of the array data and statistical analysis were performed as described previously.[15-17]

Proliferation assay

TICs were initially seeded at $5x10^4$ cells per well in a 6-well plate. Cell number and viability were measured at day 0, 2, 3, and 4 by the Countess™ automated cell counter (Invitrogen) with trypan blue exclusion. All experiments were carried out using three biological replicates and were repeated three times. The data shown represent the means ± SD.

Wound healing (migration) assay

Cells were seeded in a 6-well plate and cultured until fully confluent. The confluent cell monolayer was slightly and quickly wounded with a linear scratch made with a sterile 200/100 µl pipette tip. The debris were removed, and the edges of the scratch were levelled with PBS washing. The open gap was inspected and photographed microscopically (10X object, Nikon) at 1 and 24 hours[18]. All experiments were carried out using three biological replicates and were repeated three times. The data shown represents the mean ± SD.

Soft-agar colony formation assay

Cells ($2.5x10^3$) were seeded in 0.35% agarose in TIC growth medium on a layer of 0.5% agar in the TIC growth medium. Cells were incubated for 10-14 days at 37℃ in a humidified atmosphere containing 5% CO_2 in air. The TIC growth medium (0.5 ml) was changed two or three times a week, as needed. At the end of the incubation period, colonies were stained with crystal violet (CV) followed by scanning for colony counts. The CV stain was also read at OD_{540}. All experiments were carried out using three biological replicates and were repeated three times. The data shown represent the means ± SD.

Site-directed mutagenesis

Site-directed mutagenesis was performed as per a PCR-based mutagenesis kit (Quikchange site-directed mutagenesis kit, Stratagene, USA) with Advantage polymerase (Clontech). Consensus NANOG and STAT3 binding sites AATGG and TTCCTATAA have been previously observed *in vitro.*[19, 20] The TWIST1 plasmid -209/+1, containing putative NANOG binding sites (5'-TAAT(G/T)(G/T)-3' or 5'-[CG][GA][CG]C[GC]ATTAN[GC]-3') and STAT3 binding sites (5-TTC(C/T)N(A/G)GAA-3), were mutated utilizing a forward mutagenic primer and a reverse primer as previously described.[20] The mutated sequences were confirmed by DNA sequencing. Primers used in this analysis are listed in Supplementary Table 5. The data shown represent the means ± SD.

Confocal immunofluorescent microscopy

Immunofluoroscence staining of cryosections or paraffin sections was performed using primary antibodies against NANOG (Rabbit ab80892, Abcam), P-Stat3 (Rabbit #9134, Cell Signaling technology), TLR4 (Mouse monoclonal antibody, SC293072, Santa Cruz), TLR4 (goat sc-8694, Santa Cruz Biotechnology), or TWIST1 (Rabbit polyclonal antibody, sc-15393, Santa Cruz Biotechnology) (refer Supplementary Table 3). Specimens were mounted on glass slides according to the manufacturer's recommendations using mounting media which included DAPI

for nuclei counterstaining (Vector Laboratories). To determine the specificity of IF, serial sections were similarly processed except primary antibodies were omitted. Images were captured on a Zeiss LSM510 confocal microscope using sequential acquisition imaging. The degree of staining was categorized by the extent and the intensity of staining. Image analysis of nuclear translocation was performed using Metamorph or ImageJ v3.91 software (http://rsb.info.nih.gov/ij). A minimum of 10 high power fields were selected for image analysis. To avoid experimental bias for the staining colocalization of TLR4/NANOG/pSTAT3 with TWIST1, nuclear (DAPI) staining was used to identify fields with near-confluent cells for the purpose of maximizing the cell numbers used for analysis. The selected fields were then evaluated for the expression of TLR4, pSTAT3, NANOG, and TWIST1. Quantitative fluorescence data were exported from ImageJ generated histograms in Microsoft Excel software for further analysis and presentation. The data shown represent the means ± SD.

Tissue microarray analysis (TMA)

The HCC TMA was constructed as previously described.[21] Briefly, archived liver cases were reviewed, and areas containing HCC and benign hepatic parenchyma were marked for sampling. Three cores per HCC and matched benign from the same patient, measuring 0.6 mm in diameter, were obtained from selected regions in each donor paraffin block and transferred to a recipient paraffin block.

Spheroid assay

TICs (50 cells) were seeded onto Ultra low attachment 96-well plates (Corning Inc.), followed by incubation at 37℃ in a humidified atmosphere containing 5% CO_2 for 14 days. 100 µl/well of TIC growth medium was replaced twice a week. The number of colonies was counted under bright-field microscopy, and the proliferation was measured using counting numbers of spheroides and Luminescent Cell Viability Assay (Promega) followed by manufacturer's instructions. All experiments were carried out using 24 biological replicates and were repeated three times. The data shown represent the means ± SD.

Immunoblotting

Total cell lysates were prepared by lysing the cells in cold NP-40 buffer (150 mM NaCl, 1.0% NP-40, 10% Glycerol, and 50 mM Tris, pH 8.0) containing complete protease inhibitor mixture (Roche) for 1 h on ice, followed by centrifugation at 14,000 RPM for 15 min and collection of the clarified supernatant. Protein concentrations were determined using the DC protein assay Kit (Bio-Rad), and the supernatant was mixed with 6X Laemmli sample buffer. Proteins were separated on 10% SDS-PAGE and transferred to nitrocellulose membranes (Thermo). The membranes were blocked with 5% non-fat milk + 0.1% tween-20 for 1 h, followed by incubation with the primary antibodies: TWIST1 (Santa Cruz Biotechnology), E-CADHERIN (BD Biosciences), N-CADHERIN (Santa Cruz Biotechnology), TLR4 (Santa Cruz Biotechnology), NANOG (Abcam), pSTAT3 (Cell signaling Technology) and β-ACTIN (sigma) (all at 1:1,000 dilution) at 4°C overnight. Horseradish peroxidase–conjugated IgG (Santa Cruz Biotechnology; 1:2,000) was used to treat the membranes for 1 h at room temperature, and visualized with SuperSignal® West Pico Chemiluminescent substrate (Thermo). The immunoreactive bands were detected with Premium Clear Blue X-Ray films (Bioland Scientific LLC). Quantification of the bands was performed using ImageJ software. The data shown represent the means ± SD. Antibodies used for these studies are listed in Suppl. Table 3.

Promoter luciferase reporter assays

TICs obtained from NS5A transgenic mice (<10 passages in culture) were cultured in six-well plates and cotransfected using BioT (Bio land Scientific) with 1 μg Twist1 promoter-fused to *Firefly* luciferase reporter and 50 ng (SV40) *Renilla* luciferase expression vector to control for transfection efficiency. Forty-eight hours after transfection, cells were lysed in 1x passive lysis buffer, and luciferase activity was measured using the Dual-Glo Luciferase System (Promega) using a Lumat LB9501 luminometer (Berthold). At least three independent biological replicates were used for this experiment and were performed for at least total of three determination. Plasmids used in this assay are listed in Supplementary Table 5. The data shown represent the means ± SD.

Subcutaneous xenograft transplantation of the TICs into immunodeficient mice

NOG mice were purchased from Taconic and housed under pathogen-free conditions in accordance with approved Institutional Animal Care and Use Committee protocols. TICs (10^5) in 100 μl solution were mixed with 100 μl Matrigel (BD Biosciences) and were injected into the dorsal flanks of female NOG mice 8-9 weeks of age. Mice were anesthetized with ketamine (80 mg/kg) and xylazine (10 mg/kg) cocktail through I.P. during the procedure. The tumor volume was measured with a caliper and calculated according to the formula $V=[a \times (b^2)]/ 2$, where "V" represents tumor volume, "a" represents the largest, and "b" the smallest superficial diameter[22]. The data shown represents the mean ± SD.

Live animal imaging

The tumor bearing mice was monitored using noninvasive imaging by whole-body GFP imaging utilizing the bioluminescence imaging system (IVIS 200 Imaging Series, Xenogen) at day 21 and 35.

Chromatin immunoprecipitation (ChIP) and re-ChIP analysis

CD133+ liver TICs grown in 10-cm cell culture dishes following LPS and leptin treatment were fixed for 10 min at room temperature by addition of 1% paraformaldehyde to the growth medium. Cells were washed twice in cold PBS supplemented with complete protease inhibitor mixture and gently scraped from the plate. Cell lysis and chromatin immunoprecipitation (ChIP) were performed using the ChIP Assay Kit (Millipore). For chromatin fragmentation, cells were sonicated using a Branson Sonifier 450 on power setting 4 in 30-s bursts with 1 min cooling on ice for a total sonication time of 4 min. For immunoprecipitations, 8 μg of each antibody was used. Anti-Nanog (Abcam) and Anti Stat3 (Cell signaling technology) monoclonal antibody were used for immunoprecipitation. Preimmune IgG was used as the antibody specificity control. Immunoprecipitated DNA was quantified for *Twist1* promoters using q-PCR primers which are listed in Supplementary Table 6. The Re-ChIP or Sequential ChIP analysis was performed according to the manufacture's protocol (Active Motif Re-ChIP IT®), whereas all the initial sample preparation where the same as explained above. The data shown represent the means ± SD.

Statistical analysis

Statistical significance was estimated by un-paired, two-tailed Student's *t* test. P values are indicated in the figures. Bars represent the mean and error bars the SD. For most of the figures,

statistical significance is represented by asterisks above each column: *P<0.05, **P<0.005, ***P<0.001 and ****P<0.0001. Some figures have been represented with pound sign or ampersand, details of which are given in the respective figure legends. For Figure 7 B, statistical significance was calculated using two-way ANOVA method. In this specific analysis the time point used was where all mice were still alive, before any required euthanasia.

References

1. Machida K, Tsukamoto H, Mkrtchyan H, et al. Toll-like receptor 4 mediates synergism between alcohol and HCV in hepatic oncogenesis involving stem cell marker Nanog. Proc Natl Acad Sci U S A 2009;106:1548-53.
2. Torres J, Watt FM. Nanog maintains pluripotency of mouse embryonic stem cells by inhibiting NFkappaB and cooperating with Stat3. Nat Cell Biol 2008;10:194-201.
3. Yang MH, Hsu DS, Wang HW, et al. Bmi1 is essential in Twist1-induced epithelial-mesenchymal transition. Nat Cell Biol 2010;12:982-92.
4. Chen CL, Tsukamoto H, Liu JC, et al. Reciprocal regulation by TLR4 and TGF-beta in tumor-initiating stem-like cells. J Clin Invest 2013;123:2832-49.
5. McLemore ML, Grewal S, Liu F, et al. STAT-3 activation is required for normal G-CSF-dependent proliferation and granulocytic differentiation. Immunity 2001;14:193-204.
6. Saitoh S, Akashi S, Yamada T, et al. Lipid A antagonist, lipid IVa, is distinct from lipid A in interaction with Toll-like receptor 4 (TLR4)-MD-2 and ligand-induced TLR4 oligomerization. Int Immunol 2004;16:961-9.
7. Ohkuma M, Funato N, Higashihori N, et al. Unique CCT repeats mediate transcription of the TWIST1 gene in mesenchymal cell lines. Biochem Biophys Res Commun 2007;352:925-31.
8. Mathurin P, Deng QG, Keshavarzian A, et al. Exacerbation of alcoholic liver injury by enteral endotoxin in rats. Hepatology 2000;32:1008-17.
9. Marko-Varga G, Berglund M, Malmstrom J, et al. Targeting hepatocytes from liver tissue by laser capture microdissection and proteomics expression profiling. Electrophoresis 2003;24:3800-5.
10. Michel C, Desdouets C, Sacre-Salem B, et al. Liver gene expression profiles of rats treated with clofibric acid: comparison of whole liver and laser capture microdissected liver. Am J Pathol 2003;163:2191-9.
11. Espina V, Milia J, Wu G, et al. Laser capture microdissection. Methods Mol Biol 2006;319:213-29.
12. Wulfkuhle JD, Sgroi DC, Krutzsch H, et al. Proteomics of human breast ductal carcinoma in situ. Cancer Res 2002;62:6740-9.
13. Wulfkuhle JD, Liotta LA, Petricoin EF. Proteomic applications for the early detection of cancer. Nat Rev Cancer 2003;3:267-75.
14. Espina V, Mehta AI, Winters ME, et al. Protein microarrays: molecular profiling technologies for clinical specimens. Proteomics 2003;3:2091-100.
15. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440-5.
16. Sitkiewicz I, Musser JM. Expression microarray and mouse virulence analysis of four conserved two-component gene regulatory systems in group a streptococcus. Infect Immun 2006;74:1339-51.
17. Graham MR, Virtaneva K, Porcella SF, et al. Group A Streptococcus transcriptome dynamics during growth in human blood reveals bacterial adaptive and survival strategies. Am J Pathol 2005;166:455-65.

18. Rodriguez LG, Wu X, Guan JL. Wound-healing assay. Methods Mol Biol 2005;294:23-9.
19. Jauch R, Ng CK, Saikatendu KS, et al. Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog. J Mol Biol 2008;376:758-70.
20. Cheng GZ, Zhang WZ, Sun M, et al. Twist is transcriptionally induced by activation of STAT3 and mediates STAT3 oncogenic function. J Biol Chem 2008;283:14665-73.
21. Huang J, Yao JL, Zhang L, et al. Differential expression of interleukin-8 and its receptors in the neuroendocrine and non-neuroendocrine compartments of prostate cancer. Am J Pathol 2005;166:1807-15.
22. Carlsson G, Gullberg B, Hafstrom L. Estimation of liver tumor volume using different formulas - an experimental study in rats. J Cancer Res Clin Oncol 1983;105:20-3.