SUPPLEMENTARY INFORMATION

EGFR inhibition attenuates liver fibrosis and development of hepatocellular carcinoma
Bryan C. Fuchs, Yujin Hoshida, Tsutomu Fujii, Suguru Yamada, Gregory Y. Lauwers, Christopher M. McGinn, Lan Wei, Toshihiko Kuroda, Michael Lanuti, Anthony D. Schmitt, Supriya Gupta, Andrew Crenshaw, Robert Onofrio, Bradley Taylor, Wendy Winckler, Todd R. Golub and Kenneth K. Tanabe

Experimental Procedures

Chemicals. Stock solutions of erlotinib (Tarceva™; Genentech, South San Francisco, CA) were prepared in DMSO and diluted in water before intraperitoneal (IP) injection.

Tissue Specimens. Tissue from human cirrhosis patients was obtained under protocols approved by the Dana-Farber Harvard Cancer Center Office for Protection of Human Subjects and the Partners Human Research Committee.

Cells and Culture Conditions. The human HSC cell line, TWNT-4 (1), was kindly provided by Dr. Sangeeta Bhatia (Massachusetts Institute of Technology, Cambridge, MA). Cells were cultured in DMEM (4.5 mg/ml glucose, 2 mM L-glutamine) with 10% fetal bovine serum (FBS) that was supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin (all from MediaTech CellGro, Herndon, VA). Cells were maintained at 37°C in a humidified incubator with 5% CO_2 in air.

Diethylnitrosamine (DEN) Rat Model. Male Wistar rats received weekly IP injections of DEN (50 mg/kg) or PBS control for 18 weeks. A subset of rats received either daily (5X a week) IP injections of 2 mg/kg erlotinib (N = 8) or vehicle (N = 8) during weeks 13 - 18. In a separate study, the erlotinib dose was lowered to 0.5 mg/kg (N = 8). Using an mg/m^2 conversion, a human chemotherapeutic dose of erlotinib equates to approximately 7 mg/kg in rats. The vehicle groups from the two studies were not

significantly different so they were combined together for analysis. Rats were sacrificed at 9, 13 and 19 weeks after a one-week washout period to eliminate acute effects of DEN. At the time of sacrifice, tumors were counted and measured, and the liver was then sectioned and fixed in 10% formalin. The remaining portions of the liver were snap-frozen and stored at -80°C until use.

Hepatectomy Study. Male Wistar rats received weekly IP injections of DEN (50 mg/kg) for 18 weeks. After 12 weeks, the rats underwent a survival hepatectomy and a liver wedge biopsy was removed for histology. Rats then received either 2 mg/kg erlotinib (N = 8) or vehicle (N = 8) daily (5 days/week) during weeks 13 - 18. Rats were sacrificed at 19 weeks after a one-week washout of DEN and the liver was harvested for histology.

Carbon tetrachloride (CCl_4) Mouse Model. Strain A/J male mice (Jackson Laboratory, Bar Harbor, ME) were treated three times a week for 18 weeks with either 0.1cc of a 40 percent solution of CCl_4 (Sigma) in olive oil or olive oil alone by oral gavage. A subset of mice received daily (5 days/week) IP injections of either 2 or 5 mg/kg erlotinib or vehicle (N = 8 for all three groups) during weeks 13 - 18. Using a mg/m^2 conversion, a human chemotherapeutic dose of erlotinib equates to approximately 30 mg/kg in mice. Mice were sacrificed at 19 weeks after a one-week washout to eliminate acute effects of $\underline{CCl_4}$, and the liver was harvested as described above.

Bile duct ligation (BDL) Rat Model. Male Wistar rats that had undergone a BDL (Charles River Labs) received daily (5 days/week) IP injections of 2 mg/kg erlotinib or vehicle (N = 8 for both groups) beginning 4 days after the BDL. Rats were sacrificed on Day 21, and the liver was harvested as described above.

Primary rat hepatic stellate cell (HSC) isolation. HSC were isolated according to an established protocol (2). Briefly, rat livers were perfused with a buffer containing collagenase A and pronase (both from Sigma). The digested liver was filtered through a 100 µm nylon membrane and hepatocytes were removed by centrifugation. The purity of isolated HSCs was assessed by their content of fluorescent retinoid droplets under UV excitation.

Histology, immunohistochemistry, immunofluorescence. Formalin-fixed samples were embedded in paraffin, cut into 5 µm-thick sections and stained with hematoxylin-eosin (H-E), Masson's trichrome and Sirius red according to standard procedures. Trichrome stained sections were analyzed to score the amount of liver disease according to the method of Ishak (3) as described in Supplementary Table 1. Collagen was morphometrically quantified on Sirius red stained sections with image processing software (ImageJ, NIH). Additional sections were stained with an antibodies specific for Ki67 (Abcam, Cambridge, MA), p-EGFR (Tyr1068) (Cell Signaling Technology, Beverly, MA) or alpha-smooth muscle actin (α-SMA; DakoCytomation, Denmark). Ki67 positive cells were morphometrically quantified with image processing software (ImageJ). For dual immunofluorescence, sections were co-stained with antibodies specific for phospho-EGFR (Tyr1068) (Cell Signaling Technology) and either α-SMA (Abcam), glial fibrillary acidic protein (GFAP; DakoCytomation) or desmin (Abcam) with detection by appropriate secondary antibodies labeled with either Cy3 or Alexa 488 according to the manufacturer's instructions. All slides were reviewed blindly by the same liver pathologist.

Liver Function Tests. A cardiac terminal blood withdrawal was performed at the time of sacrifice. Blood was allowed to clot for 2 h at room temperature before centrifugation at

2,000 rpm for 10 min at 4^{o} C. Serum was isolated and stored at -80^{o} C prior to use. Liver injury and liver failure were assessed by measuring the serum levels of several biochemical markers including alkaline phosphatase (ALP), alanine transaminase (ALT), aspartate transaminase (AST), total bilirubin (TBIL), albumin (Alb) and glucose (Glu) (DRI-CHEM 4000 Analyzer, Heska, Switzerland).

Hydroxyproline analysis and western blotting. Hydroxyproline in tissue was quantified by HPLC analysis as previously described (4). Western blot analysis was performed as previously described (5). Liver lysates from every animal from each group at all the time points were analyzed and representative samples are shown. For HSC experiments, either rat primary HSC cultures or TWNT-4 cells were serum starved overnight in DMEM with 0.1% FBS before the addition 100 ng/ml EGF with or without 2 μM erlotinib for 30 minutes. The following antibodies were used: EGFR, phospho-EGFR (Tyr1068), p44/42 mitogen-activated protein kinase (ERK), phospho-ERK (Thr202/Tyr204), proliferating cell nuclear antigen (PCNA), platelet-derived growth factor receptor-beta (PDGFRβ) (all from Cell Signaling Technology), α-SMA (Abcam) and β-actin (Sigma). Band intensities were quantified by image analysis software (Labworks 4.5; UVP Inc., Upland, CA). Each western blot was repeated to ensure reproducible results.

Real-Time PCR. Total RNA was isolated from animal liver tissue using TRIzol (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions and subsequently treated with DNase I (Promega, Madison, WI). 1 μg of total RNA from each sample was used to create cDNA by single strand reverse transcription (SuperScript III First-Strand Synthesis SuperMix for qRT-PCR; Invitrogen). Quantitative real-time PCR was performed using the 7900HT Fast Real-Time PCR System with commercial

TaqMan primers for cytochrome P450 2E1 (CYP2E1), amphiregulin (AREG), betacellulin (BTC), epiregulin (EREG), heparin-binding EGF-like growth factor (HB-EGF) and transforming growth factor-alpha (TGF-α) (Life Technologies, Grand Island, NY). Primers and conditions are available upon request. TWNT-4 cells were plated at 1 x 10^5 cells/ml in 10 ml media in a 100 mm plate and treated with or without 2 μM erlotinib the next day. Total RNA was extracted 72 hours later and cDNA synthesized as described above. Expression of α-SMA and α1(I) procollagen mRNA was analyzed by quantitative real-time PCR (LightCycler; Roche Diagnostics Corporation, Indianapolis, IN). mRNA expression was normalized to the expression of β-actin. Primer sequences are as follows: α-SMA forward TTCAATGTCCCAGCCATGTA and reverse GAAGGAATAGCCACGCTCAG, α1(I) procollagen forward AACATGACCAAAAACCAAAAGTG and reverse CATTGTTTCCTGTGTCTTCTGG, and Actin forward CCTGGACTTCGAGCAAGAGAT and reverse GCCGATCCACACGGAGTACT. All reactions were performed in duplicate and the experiment was repeated to ensure reproducible results.

Microarray Analysis. Total RNA was extracted from either non-tumor liver tissue or HCC tissue using TRIzol (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions and subsequently treated with DNase I (Promega, Madison, WI). Absence and presence of tumor was confirmed by H-E staining of the tissue sections. Genome-wide gene expression profiling for the rats and mice was performed using RatRef-12 and Mouse Ref-8 Expression BeadChip microarrays, respectively (Illumina, San Diego, CA). Scanned data were normalized using cubic spline algorithm (6), summarized into official gene symbol, and mapped to human orthologous genes using the mapping table provided by the Jackson laboratory (www.informatics.jax.org). The dataset is available at

NCBI Gene Expression Omnibus (GEO) with accession number GSE27641 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=xvelpumoegoykba&acc=GSE27641).

The 186-gene signature predictive of survival of patients with cirrhosis (7) and HCC (8) has been described elsewhere. EGF target gene signatures were previously generated in primary human foreskin microvascular endothelial cells (9), a normal human breast epithelial cell line (MCF10A) and a human cervical cancer cell line (HeLa) (10). Induction of the gene expression signatures was evaluated using Gene Set Enrichment Analysis (11). Briefly, genes on the microarray were rank-ordered according to differential expression between the experimental conditions. Over- or under-representation of each gene signature on the rank-ordered gene list was evaluated based on random permutation test. False discovery rate (FDR) <0.25 was regarded as statistically significant. All microarray data analysis was performed using GenePattern analysis toolkit (12) (www.genepattern.org)) and R statistical computing language (www.r-project.org).

Survival analysis. The association between high EGF expression and survival was analyzed by Cox score. Gene-expression profiling was previously performed according to the complementary DNA–mediated annealing, selection, extension, and ligation (DASL) assay (Illumina) on formalin-fixed, paraffin-embedded blocks of tumor and surrounding tissue from patients who were consecutively treated with surgery for primary HCC between 1990 and 2001 at Toranomon Hospital in Tokyo and for whom data on clinical outcomes (over a median follow-up period of 7.8 years) were available (8). Surrounding and tumor tissue samples from the same patients were analyzed to identify the samples with the highest EGF expression (90^{th} percentile).

Statistical Analysis. An unpaired two-tailed t-test was used to compare differences in body weights, liver weights, quantification of Sirius red stained sections, liver function tests and RNA expression.

Supplementary Table 1. Ishak scoring.

Score	Description
0	No fibrosis
1	Fibrous expansion of some portal areas, with or without short fibrous septa
2	Fibrous expansion of most portal areas, with or without short fibrous septa
3	Fibrous expansion of most portal areas, with occasional portal to portal bridging
4	Fibrous expansion of most portal areas, with marked portal to portal bridging as well as portal areas to central bridging
5	Marked bridging with occasional nodules
6	Probable to definite cirrhosis

Supplementary Table 2. Enrichment of drug metabolizing enzymes in erlotinib-treated DEN rats and CCl_4 mice.

DEN Rats			
Gene Set	**NES**	**p-value**	**FDR**
KEGG_DRUG_METABOLISM_CYTOCHROME_P450	1.64	0.006	0.016
KEGG_DRUG_METABOLISM_OTHER_ENZYMES	1.11	0.326	0.302
CCl_4 Mice			
Gene Set	**NES**	**p-value**	**FDR**
KEGG_DRUG_METABOLISM_CYTOCHROME_P450	2.07	<0.001	<0.001
KEGG_DRUG_METABOLISM_OTHER_ENZYMES	1.22	0.212	0.18

NES: normalized enrichment score, FDR: False discovery rate.

Supplementary Table 3. The 186-gene signature.

GeneID	Gene symbol	Description	Cox score
Genes Correlated with Poor Survival			
2488	FSHB	follicle stimulating hormone, beta polypeptide	4.80
6456	SH3GL2	SH3-domain GRB2-like 2	4.21
23029	RBM34	RNA binding motif protein 34	4.19
23397	NCAPH	non-SMC condensin I complex, subunit H	4.02
1950	EGF	epidermal growth factor (beta-urogastrone)	3.97
7204	TRIO	triple functional domain (PTPRF interacting)	3.90
1293	COL6A3	collagen, type VI, alpha 3	3.87
3983	ABLIM1	actin binding LIM protein 1	3.86
3680	ITGA9	integrin, alpha 9	3.81
4922	NTS	Neurotensin	3.78
5055	SERPINB2	serpin peptidase inhibitor, clade B (ovalbumin), member 2	3.69
4316	MMP7	matrix metallopeptidase 7 (matrilysin, uterine)	3.59
5593	PRKG2	protein kinase, cGMP-dependent, type II	3.44
9170	EDG4	endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 4	3.40
4843	NOS2A	nitric oxide synthase 2A (inducible, hepatocytes)	3.33
2043	EPHA4	EPH receptor A4	3.25
6672	SP100	SP100 nuclear antigen	3.19
2326	FMO1	flavin containing monooxygenase 1	3.04
2877	GPX2	glutathione peroxidase 2 (gastrointestinal)	3.02
496	ATP4B	ATPase, H+/K+ exchanging, beta polypeptide	2.99
8870	IER3	immediate early response 3	2.98
7456	WIPF1	WAS/WASL interacting protein family, member 1	2.98
3489	IGFBP6	insulin-like growth factor binding protein 6	2.93
1501	CTNND2	catenin (cadherin-associated protein), delta 2 (neural plakophilin-related arm-repeat protein)	2.92
2200	FBN1	fibrillin 1	2.91
2629	GBA	glucosidase, beta; acid (includes glucosylceramidase)	2.85
22858	ICK	intestinal cell (MAK-like) kinase	2.85
10523	CHERP	calcium homeostasis endoplasmic reticulum protein	2.81
9734	HDAC9	histone deacetylase 9	2.81
51406	NOL7	nucleolar protein 7, 27kDa	2.80
8826	IQGAP1	IQ motif containing GTPase activating protein 1	2.79
120	ADD3	adducin 3 (gamma)	2.79
306	ANXA3	annexin A3	2.78
10362	HMG20B	high-mobility group 20B	2.76
6558	SLC12A2	solute carrier family 12 (sodium/potassium/chloride transporters), member 2	2.75
1282	COL4A1	collagen, type IV, alpha 1	2.75
1359	CPA3	carboxypeptidase A3 (mast cell)	2.74

3855	KRT7	keratin 7	2.74
5271	SERPINB8	serpin peptidase inhibitor, clade B (ovalbumin), member 8	2.69
4791	NFKB2	nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	2.67
165	AEBP1	AE binding protein 1	2.67
7041	TGFB1I1	transforming growth factor beta 1 induced transcript 1	2.66
2013	EMP2	epithelial membrane protein 2	2.63
596	BCL2	B-cell CLL/lymphoma 2	2.63
5698	PSMB9	proteasome (prosome, macropain) subunit, beta type, 9 (large multifunctional peptidase 2)	2.59
10097	ACTR2	ARP2 actin-related protein 2 homolog (yeast)	2.59
780	DDR1	discoidin domain receptor family, member 1	2.58
6541	SLC7A1	solute carrier family 7 (cationic amino acid transporter, y+ system), member 1	2.56
5420	PODXL	podocalyxin-like	2.56
1307	COL16A1	collagen, type XVI, alpha 1	2.55
10437	IFI30	interferon, gamma-inducible protein 30	2.55
9852	EPM2AIP1	EPM2A (laforin) interacting protein 1	2.55
301	ANXA1	annexin A1	2.53
6366	CCL21	chemokine (C-C motif) ligand 21	2.47
22856	CHSY1	carbohydrate (chondroitin) synthase 1	2.45
162	AP1B1	adaptor-related protein complex 1, beta 1 subunit	2.45
7004	TEAD4	TEA domain family member 4	2.39
54898	ELOVL2	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 2	2.39
6925	TCF4	transcription factor 4	2.38
9819	TSC22D2	TSC22 domain family, member 2	2.38
1847	DUSP5	dual specificity phosphatase 5	2.36
8030	CCDC6	coiled-coil domain containing 6	2.36
962	CD48	CD48 molecule	2.35
10188	TNK2	tyrosine kinase, non-receptor, 2	2.35
1601	DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	2.35
4017	LOXL2	lysyl oxidase-like 2	2.34
6035	RNASE1	ribonuclease, RNase A family, 1 (pancreatic)	2.34
4026	LPP	LIM domain containing preferred translocation partner in lipoma	2.33
7852	CXCR4	chemokine (C-X-C motif) receptor 4	2.33
6586	SLIT3	slit homolog 3 (Drosophila)	2.31
11259	FILIP1L	filamin A interacting protein 1-like	2.25
6363	CCL19	chemokine (C-C motif) ligand 19	2.23
11214	AKAP13	A kinase (PRKA) anchor protein 13	2.23
Genes Correlated with Good Survival			
223	ALDH9A1	aldehyde dehydrogenase 9 family, member A1	-3.34
7276	TTR	transthyretin (prealbumin, amyloidosis type I)	-3.27

6018	RLF	rearranged L-myc fusion	-3.23
3612	IMPA1	inositol(myo)-1(or 4)-monophosphatase 1	-3.22
5207	PFKFB1	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 1	-3.22
6296	ACSM3	acyl-CoA synthetase medium-chain family member 3	-3.21
151	ADRA2B	adrenergic, alpha-2B-, receptor	-3.19
5771	PTPN2	protein tyrosine phosphatase, non-receptor type 2	-3.12
5691	PSMB3	proteasome (prosome, macropain) subunit, beta type, 3	-3.09
5502	PPP1R1A	protein phosphatase 1, regulatory (inhibitor) subunit 1A	-3.07
27346	TMEM97	transmembrane protein 97	-3.06
5313	PKLR	pyruvate kinase, liver and RBC	-3.01
9252	RPS6KA5	ribosomal protein S6 kinase, 90kDa, polypeptide 5	-3.00
1528	CYB5A	cytochrome b5 type A (microsomal)	-2.96
6447	SCG5	secretogranin V (7B2 protein)	-2.93
25828	TXN2	thioredoxin 2	-2.90
5340	PLG	plasminogen	-2.88
6309	SC5DL	sterol-C5-desaturase (ERG3 delta-5-desaturase homolog, S. cerevisiae)-like	-2.87
367	AR	androgen receptor (dihydrotestosterone receptor; testicular feminization; spinal and bulbar muscular atrophy; Kennedy disease)	-2.84
3479	IGF1	insulin-like growth factor 1 (somatomedin C)	-2.84
8802	SUCLG1	succinate-CoA ligase, GDP-forming, alpha subunit	-2.84
23498	HAAO	3-hydroxyanthranilate 3,4-dioxygenase	-2.83
735	C9	complement component 9	-2.83
9013	TAF1C	TATA box binding protein (TBP)-associated factor, RNA polymerase I, C, 110kDa	-2.82
1371	CPOX	coproporphyrinogen oxidase	-2.82
7507	XPA	xeroderma pigmentosum, complementation group A	-2.82
3026	HABP2	hyaluronan binding protein 2	-2.81
2690	GHR	growth hormone receptor	-2.77
5105	PCK1	phosphoenolpyruvate carboxykinase 1 (soluble)	-2.76
6718	AKR1D1	aldo-keto reductase family 1, member D1 (delta 4-3-ketosteroid-5-beta-reductase)	-2.76
128	ADH5	alcohol dehydrogenase 5 (class III), chi polypeptide	-2.75
16	AARS	alanyl-tRNA synthetase	-2.70
732	C8B	complement component 8, beta polypeptide	-2.69
51237	MGC29506	NA	-2.67
10159	ATP6AP2	ATPase, H+ transporting, lysosomal accessory protein 2	-2.67
9732	DOCK4	dedicator of cytokinesis 4	-2.66
5627	PROS1	protein S (alpha)	-2.66
7709	ZBTB17	zinc finger and BTB domain containing 17	-2.65
1603	DAD1	defender against cell death 1	-2.65
1678	TIMM8A	translocase of inner mitochondrial membrane 8 homolog A (yeast)	-2.65
3155	HMGCL	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase (hydroxymethylglutaricaciduria)	-2.65

725	C4BPB	complement component 4 binding protein, beta	-2.62
7189	TRAF6	TNF receptor-associated factor 6	-2.62
1967	EIF2B1	eukaryotic translation initiation factor 2B, subunit 1 alpha, 26kDa	-2.61
3990	LIPC	lipase, hepatic	-2.60
10026	PIGK	phosphatidylinositol glycan anchor biosynthesis, class K	-2.60
80344	WDR23	WD repeat domain 23	-2.59
5982	RFC2	replication factor C (activator 1) 2, 40kDa	-2.58
2915	GRM5	glutamate receptor, metabotropic 5	-2.56
6391	SDHC	succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa	-2.55
2073	ERCC5	excision repair cross-complementing rodent repair deficiency, complementation group 5 (xeroderma pigmentosum, complementation group G (Cockayne syndrome))	-2.54
2158	F9	coagulation factor IX (plasma thromboplastic component, Christmas disease, hemophilia B)	-2.54
157567	ANKRD46	ankyrin repeat domain 46	-2.54
417	ART1	ADP-ribosyltransferase 1	-2.54
1486	CTBS	chitobiase, di-N-acetyl-	-2.54
2542	SLC37A4	solute carrier family 37 (glucose-6-phosphate transporter), member 4	-2.53
211	ALAS1	aminolevulinate, delta-, synthase 1	-2.53
27072	VPS41	vacuolar protein sorting 41 homolog (S. cerevisiae)	-2.51
2642	GCGR	glucagon receptor	-2.51
10694	CCT8	chaperonin containing TCP1, subunit 8 (theta)	-2.51
25874	BRP44	brain protein 44	-2.50
2868	GRK4	G protein-coupled receptor kinase 4	-2.50
3336	HSPE1	heat shock 10kDa protein 1 (chaperonin 10)	-2.50
79731	NARS2	asparaginyl-tRNA synthetase 2, mitochondrial (putative)	-2.49
667	DST	dystonin	-2.49
27032	ATP2C1	ATPase, Ca++ transporting, type 2C, member 1	-2.48
10327	AKR1A1	aldo-keto reductase family 1, member A1 (aldehyde reductase)	-2.48
2010	EMD	emerin (Emery-Dreifuss muscular dystrophy)	-2.47
799	CALCR	calcitonin receptor	-2.45
22839	DLGAP4	discs, large (Drosophila) homolog-associated protein 4	-2.45
6240	RRM1	ribonucleotide reductase M1 polypeptide	-2.44
29937	NENF	neuron derived neurotrophic factor	-2.44
29887	SNX10	sorting nexin 10	-2.44
5372	PMM1	phosphomannomutase 1	-2.44
6999	TDO2	tryptophan 2,3-dioxygenase	-2.43
2944	GSTM1	glutathione S-transferase M1	-2.43
6721	SREBF2	sterol regulatory element binding transcription factor 2	-2.42
26469	PTPN18	protein tyrosine phosphatase, non-receptor type 18 (brain-derived)	-2.42
27163	ASAHL	N-acylsphingosine amidohydrolase (acid ceramidase)-like	-2.41
5336	PLCG2	phospholipase C, gamma 2 (phosphatidylinositol-specific)	-2.41

3760	KCNJ3	potassium inwardly-rectifying channel, subfamily J, member 3	-2.40
5833	PCYT2	phosphate cytidylyltransferase 2, ethanolamine	-2.40
2705	GJB1	gap junction protein, beta 1, 32kDa	-2.39
7108	TM7SF2	transmembrane 7 superfamily member 2	-2.39
8991	SELENBP1	selenium binding protein 1	-2.38
316	AOX1	aldehyde oxidase 1	-2.37
10444	ZER1	zer-1 homolog (C. elegans)	-2.37
130	ADH6	alcohol dehydrogenase 6 (class V)	-2.36
2956	MSH6	mutS homolog 6 (E. coli)	-2.36
8671	SLC4A4	solute carrier family 4, sodium bicarbonate cotransporter, member 4	-2.34
9097	USP14	ubiquitin specific peptidase 14 (tRNA-guanine transglycosylase)	-2.34
727	C5	complement component 5	-2.32
5893	RAD52	RAD52 homolog (S. cerevisiae)	-2.32
116496	FAM129A	family with sequence similarity 129, member A	-2.31
10458	BAIAP2	BAI1-associated protein 2	-2.31
6744	SSFA2	sperm specific antigen 2	-2.30
5446	PON3	paraoxonase 3	-2.30
2646	GCKR	glucokinase (hexokinase 4) regulator	-2.30
1385	CREB1	cAMP responsive element binding protein 1	-2.30
23316	CUTL2	cut-like 2 (Drosophila)	-2.29
6427	SFRS2	splicing factor, arginine/serine-rich 2	-2.28
3156	HMGCR	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	-2.28
2677	GGCX	gamma-glutamyl carboxylase	-2.27
1555	CYP2B6	cytochrome P450, family 2, subfamily B, polypeptide 6	-2.26
7739	ZNF185	zinc finger protein 185 (LIM domain)	-2.26
378	ARF4	ADP-ribosylation factor 4	-2.23
10965	ACOT2	acyl-CoA thioesterase 2	-2.22
513	ATP5D	ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit	-2.22
1369	CPN1	carboxypeptidase N, polypeptide 1	-2.20
5331	PLCB3	phospholipase C, beta 3 (phosphatidylinositol-specific)	-2.20
3642	INSM1	insulinoma-associated 1	-2.18
5442	POLRMT	polymerase (RNA) mitochondrial (DNA directed)	-2.14
11145	HRASLS3	HRAS-like suppressor 3	-2.13

Cox scores were calculated in the training set as previously reported (5).

Supplementary Table 4. Enrichment of the 186-gene signature in rodent models of chronic liver disease.

Model	Enrichment of Poor-Prognosis Signature	Enrichment of Good-Prognosis Signature	Reference
KLF6 +/- mouse	0.05	0.24	(13)
Mst1/2 -/- mouse	0.81	0.003	(14)
Notch1 tg mouse	0.45	0.01	(15)
DEN rat	0.001	< 0.001	this study
CCL4 mouse	0.19	0.004	this study
Bile duct ligated rat	0.001	0.11	(16)

False discovery rates (FDR) in Gene Set Enrichment Analysis is presented.

Supplementary Figure Legends

Supplementary Figure 1. DEN induces cirrhosis and hepatocellular carcinoma (HCC) in a rat model. Male Wistar rats received weekly intraperitoneal injections of PBS or DEN (50 mg/kg diluted in PBS) for 8, 12 or 18 weeks. **(A)** Representative rat livers at the time of sacrifice. **(B)** Representative H&E staining of formalin-fixed paraffin embedded (FFPE) liver tissue (Magnification 100X). **(C)** Representative Masson's trichrome staining of FFPE liver tissue (Magnification 100X). **(D)** Masson's trichrome stains of liver sections were scored by the method of Ishak (n = 6 for all groups). **(E)** Collagen levels were morphometrically quantified from Sirius red stained sections. **(F)** Rat weights. **(G)** Liver weight expressed as percent body weight at week 18. **(G)** Representative H&E staining of a well-differentiated HCC in a liver after 18 weeks of DEN administration. ** $p < 0.01$ compared to PBS.

Supplementary Figure 2. DEN induces progressive liver injury and liver failure in a rat model. Male Wistar rats received weekly intraperitoneal injections of PBS or DEN for 8, 12 or 18 weeks. Serum levels (n = 4 for all groups) of **(A)** alkaline phosphatase (ALP), **(B)** alanine transaminase (ALT), **(C)** aspartate transaminase (AST) **(D)** total bilirubin (TBIL) **(E)** albumin (Alb) and **(F)** glucose (Glu). * $p < 0.05$ and ** $p < 0.01$ compared to PBS.

Supplementary Figure 3. Expression of cytochrome P450 2E1 (CYP2E1) in DEN-injured rats and CCl_4-injured mice. (A) Male Wistar rats received PBS (-) or DEN (+) for 18 weeks. DEN-injured rats received vehicle control (-) or erlotinib (0.5 (+) or 2 mg/kg (++)) during weeks 13 -18. Expression of CYP2E1 was determined by quantitative real-time PCR. ## $p < 0.01$ compared to PBS. **(B)** Male A/J mice were administered either

0.1cc of a 40% solution of CCl_4 in olive oil (+) or olive oil alone (-) for 18 weeks. CCl_4-injured mice received vehicle control (-) or erlotinib (2 (+) or 5 mg/kg (++)) during weeks 13 - 18. Expression of CYP2E1 was determined by quantitative real-time PCR. ## $p < 0.01$ compared to olive oil alone and ** $p < 0.01$ compared to CCl_4-injured.

Supplementary Figure 4. Deregulation of cirrhosis-related molecular pathways in human and rodent transcriptome datasets. (A) Enrichment of cirrhosis-related molecular pathways in human cirrhosis and rodent models of liver injury, fibrosis, and/or carcinogenesis. We analyzed previously reported transcriptome datasets from human (NCBI Gene Expression Omnibus, GSE6764, 10 healthy livers vs. 13 cirrhotic livers) (17), Bile duct ligation (BDL) rat (GSE13747, 6 BDL rats vs. 6 controls) (16), Notch1 transgenic mouse (GSE33560, 6 transgenic vs. 5 controls) (15), Klf6 knockout mouse (GSE23375, 4 knockouts vs. 4 controls) (13), and Mst1/2 knockout mouse (14) in addition to the current datasets for DEN rat and CCl_4 mouse. The molecular pathway gene sets representing inflammatory response (GeneOntology, www.geneontology.org), blood coagulation and global metabolism (KEGG, www.genome.jp/kegg), bile acid synthesis and platelet-related biology (REACTOME, www.reactome.org), and glucose metabolism (SABiosciences, www.sabiosciences.com) were obtained from Molecular Signature Database (MSigDB, www.broadinstitute.org/msigdb). Induction or suppression of each pathway gene set was quantitatively evaluated by GSEA (www.broadinstitute.org/gsea) (11) as normalized gene set enrichment score. **(B)** Similarity of cirrhosis-related molecular pathway activation status between human cirrhosis and the rodent models of liver injury, fibrosis, and/or carcinogenesis. Similarity was assessed for each group of each molecular pathway. For the groups with multiple pathway gene sets, similarity of the profiles of the normalized gene set enrichment score was measured by Pearson correlation coefficient and the samples were clustered by

using hierarchical clustering with average linkage method. For the groups with only one pathway gene set, clustering was performed based on similarity of the normalized gene set enrichment score. In summary, DEN rat, CCl_4 mouse, BDL rat, and Mst1/2ko mouse showed globally similar pathway deregulation patterns and modest correlation to human ("All Pathways"). Examining each pathway separately, we found that "Bile Acid Synthesis" and "Platelet-Related Pathways" were generally very similar between human and all the rodent models, while induction of "Inflammatory Response" and suppression of "Coagulation" pathways in human were closest to the DEN rat. For "Global Metabolism", all rodent models (with Klf6-ko mouse being a little outstanding) formed a cluster and were distinct from human, but some models (including DEN rat and BDL rat) did show a weak positive correlation with human. Lastly, "Glucose Metabolism" showed the most distinct difference between human and the rodent models where a negative correlation was observed, and this may explain the modest recovery of serum glucose levels in our models after treatment with erlotinib.

Supplementary Figure 5. DEN administration reproduces a gene expression signature associated with poor survival in human cirrhosis patients. Male Wistar rats received weekly intraperitoneal injections of DEN or PBS for 8, 12 or 18 weeks. GSEA analysis of the **(A)** 73-gene poor-prognosis signature and **(B)** 113-gene good-prognosis signature in response to DEN administration for 12 weeks or 18 weeks.

Supplementary Figure 6. Erlotinib reverses a gene expression signature associated with poor survival in human cirrhosis patients and inhibits HSC activation in CCl_4-treated mice. Male A/J mice were administered either 0.1cc of a 40% solution of CCl_4 in olive oil (OO) (+) or olive oil alone (-) by oral gavage three times a week for 18 weeks. A subset of mice received vehicle control (-) or erlotinib (2 (+) or 5

mg/kg (++)) daily (5 days/week) during weeks 13 - 18. Mice were sacrificed at 19 weeks after a one-week washout to eliminate acute effects of CCl_4. GSEA analysis of the **(A)** 73-gene poor-prognosis signature and **(B)** 113-gene good-prognosis signature in response to CCl_4 administration for 18 weeks. **(C)** Erlotinib reverses the 186-gene signature that is predictive of prognosis in human cirrhosis and HCC patients. **(D)** Representative photomicrographs of liver sections that were stained for α-SMA as a marker for activated HSC (Magnification 40X).

Supplementary Figure 7. Expression of EGFR ligands in DEN-injured rats, CCl_4-injured mice and BDL rats. (A) Male Wistar rats received PBS (-) or DEN (+) for 18 weeks. DEN-injured rats received vehicle control (-) or erlotinib (0.5 (+) or 2 mg/kg (++)) during weeks 13 -18. The expression of amphiregulin (AREG), betacellulin (BTC), epiregulin (EREG), heparin-binding EGF (HB-EGF) and transforming growth factor-alpha (TGF-α) was determined by quantitative real-time PCR. # $p < 0.05$ compared to PBS. **(B)** Male A/J mice were administered either 0.1cc of a 40% solution of CCl_4 in olive oil (+) or olive oil alone (-) for 18 weeks. CCl_4-injured mice received vehicle control (-) or erlotinib (2 (+) or 5 mg/kg (++)) during weeks 13 - 18. The expression of amphiregulin (AREG), betacellulin (BTC), epiregulin (EREG), heparin-binding EGF (HB-EGF) and transforming growth factor-alpha (TGF-α) was determined by quantitative real-time PCR. # $p < 0.05$ compared to olive oil alone and * $p < 0.05$ or ** $p < 0.01$ compared to CCl_4-injured. **(C)** Male Wistar rats that had undergone a sham operation (-) or BDL (+) received vehicle control (-) or erlotinib (2 mg/kg (+)) beginning 4 days after the BDL and ending on Day 21. The expression of amphiregulin (AREG), betacellulin (BTC), epiregulin (EREG), heparin-binding EGF (HB-EGF) and transforming growth factor-alpha

(TGF-α) was determined by quantitative real-time PCR. # $p < 0.05$ or ## $p < 0.01$ compared to sham operation and * $p < 0.05$ compared to BDL.

Supplementary Figure 8. EGF-target gene signatures are induced in the surrounding non-tumoral liver tissue of DEN-treated animals and reversed after treatment with erlotinib. GSEA evaluating induction of previously reported EGF target gene signatures in **(A,B,C)** the surrounding non-tumoral liver tissue of rats treated with PBS or DEN for 18 weeks, or **(D,E,F)** the surrounding non-tumoral liver tissue or **(G,H,I)** tumor tissue of rats treated with DEN for 18 weeks plus either vehicle or erlotinib (2 mg/kg) during weeks 13 - 18. EGF target gene signatures were generated in **(A,D,G)** primary human foreskin microvascular endothelial cells (9), **(B, E,H)** a normal human breast epithelial cell line MCF10A, and **(C,F,I)** a human cervical cancer cell line HeLa (10).

Supplementary Figure 9. EGFR signaling is present in HSC in human liver cirrhosis and erlotinib inhibits HSC activation. (A) Liver sections from human cirrhosis patients were co-stained for p-EGFR (Y1068) and either GFAP, desmin or α-SMA (Magnification 400X). **(B)** Representative western blot analysis of TWNT-4 cells treated with 100 ng/ml EGF in the absence or presence of 2 μM erlotinib for 30 minutes. **(C)** Expression of α-SMA and α1(I) procollagen in TWNT-4 cells after treatment with or without 2 μM erlotinib for 72 hours was determined by quantitative real-time PCR.

Supplementary Figure 10. Erlotinib inhibits HSC activation in BDL rats. Male Wistar rats that had undergone a sham operation (-) or BDL (+) received IP injections of vehicle control (-) or erlotinib (2 mg/kg (+)) beginning 4 days after the BDL and ending on Day

21. Representative photomicrographs of liver sections that were stained for α-SMA as a marker for activated HSC (Magnification 100X).

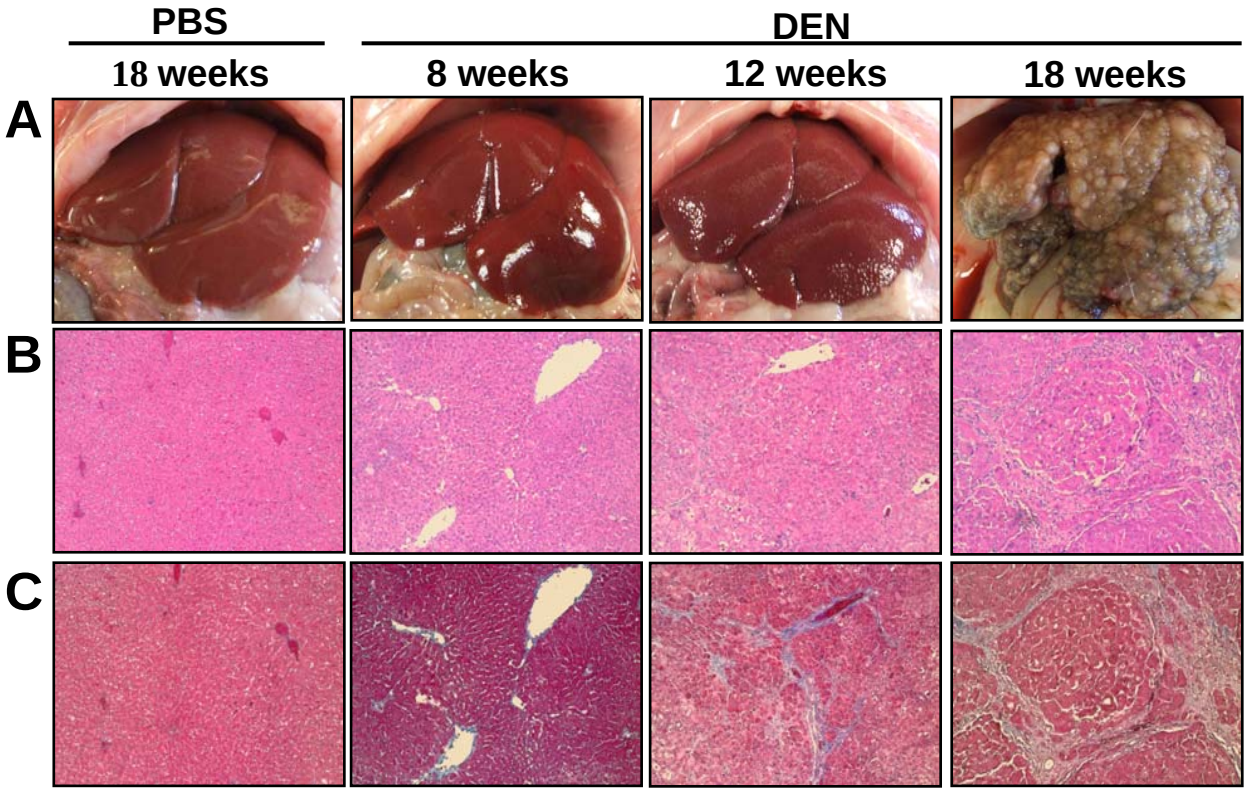
Supplementary Figure 11. EGF expression in the surrounding non-tumoral liver tissue but not the tumor is associated with survival in HCC patients. Gene-expression profiling was previously performed on tumor and surrounding tissue from patients who were consecutively treated with surgery for primary HCC and for whom data on clinical outcomes (over a median follow-up period of 7.8 years) were available (8). **(A)** Surrounding non-tumor and **(B)** HCC tissue samples from the same patients were analyzed to identify the samples with the highest EGF expression (90^{th} percentile). The association between survival and high EGF expression in either **(C)** surrounding non-tumor or **(D)** HCC tissue samples was analyzed by Cox score.

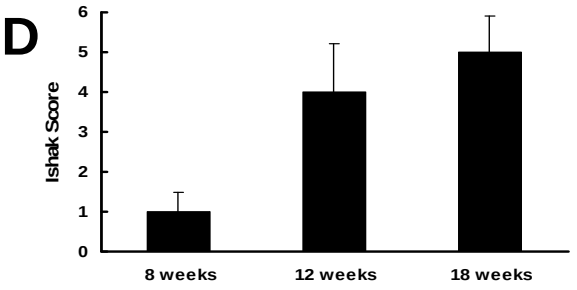
Supplementary References

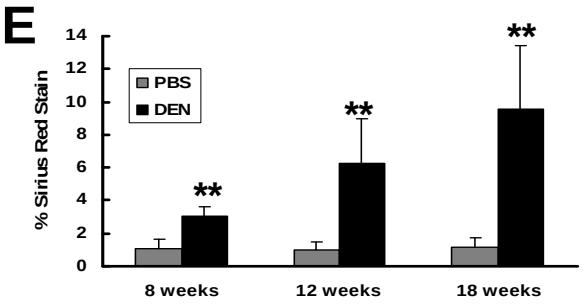
1. Shibata N, Watanabe T, Okitsu T, Sakaguchi M, Takesue M, Kunieda T, Omoto K, et al. Establishment of an immortalized human hepatic stellate cell line to develop antifibrotic therapies. Cell Transplant 2003;12:499-507.
2. Friedman SL, Roll FJ. Isolation and culture of hepatic lipocytes, Kupffer cells, and sinusoidal endothelial cells by density gradient centrifugation with Stractan. Anal Biochem 1987;161:207-218.
3. Ishak KG. Chronic hepatitis: morphology and nomenclature. Mod Pathol 1994;7: 690-713.
4. Hutson PR, Crawford ME, Sorkness RL. Liquid chromatographic determination of hydroxyproline in tissue samples. J Chromatogr B Analyt Technol Biomed Life Sci 2003;791:427-430.
5. Fuchs BC, Fujii T, Dorfman JD, Goodwin JM, Zhu AX, Lanuti M, Tanabe KK. Epithelial-to-mesenchymal transition and integrin-linked kinase mediate sensitivity to epidermal growth factor receptor inhibition in human hepatoma cells. Cancer Res 2008;68:2391-2399.
6. Workman C, Jensen LJ, Jarmer H, Berka R, Gautier L, Nielser HB, Saxild HH, et al. A new non-linear normalization method for reducing variability in DNA microarray experiments. Genome Biol 2002;3:research0048.
7. Hoshida Y, Villanueva A, Sangiovanni A, Sole M, Hur C, Andersson KL, Chung RT, et al. Prognostic gene-expression signature for patients with hepatitis C-related early-stage cirrhosis. Gastroenterology 2013;144:1024-1030.
8. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, Gupta S, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med 2008;359:1995-2004.
9. Zhang HT, Gorn M, Smith K, Graham AP, Lau KK, Bicknell R. Transcriptional profiling of human microvascular endothelial cells in the proliferative and quiescent state using cDNA arrays. Angiogenesis 1999;3:211-219.
10. Amit I, Citri A, Shay T, Lu Y, Katz M, Zhang F, Tarcic G, et al. A module of negative feedback regulators defines growth factor signaling. Nat Genet 2007;39:503-512.
11. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545-15550.
12. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. Nat Genet 2006;38:500-501.
13. Tarocchi M, Hannivoort R, Hoshida Y, Lee UE, Vetter D, Narla G, Villanueva A, et al. Carcinogen-induced hepatic tumors in KLF6+/- mice recapitulate aggressive human hepatocellular carcinoma associated with p53 pathway deregulation. Hepatology 2011;54:522-531.
14. Zhou D, Conrad C, Xia F, Park JS, Payer B, Yin Y, Lauwers GY, et al. Mst1 and Mst2 maintain hepatocyte quiescence and suppress hepatocellular carcinoma development through inactivation of the Yap1 oncogene. Cancer Cell 2009;16:425-438.
15. Villanueva A, Alsinet C, Yanger K, Hoshida Y, Zong Y, Toffanin S, Rodriquez-Carunchio L, et al. Notch signaling is activated in human hepatocellular carcinoma and induces tumor formation in mice. Gastroenterology 2012;143:1660-1669.
16. Moreno M, Chaves JF, Sancho-Bru P, Ramalho F, Ramalho LN, Mansego ML, Ivorra C, et al. Ghrelin attenuates hepatocellular injury and liver fibrogenesis in rodents and influences fibrosis progression in humans. Hepatology 2010;51:974-985.

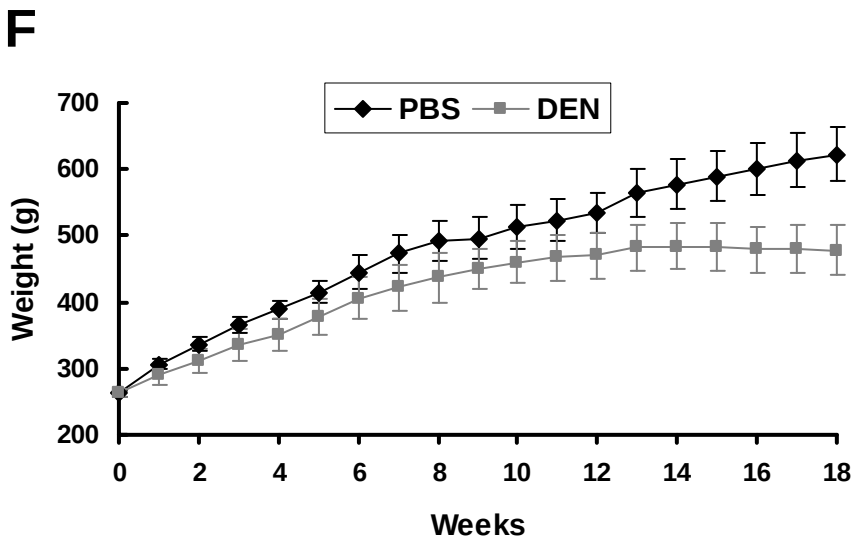
17. Wurmbach E, Chen YB, Khitrov G, Zhang W, Roayaie S, Schwartz M, Fiel I, et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology 2007;45:938-947.

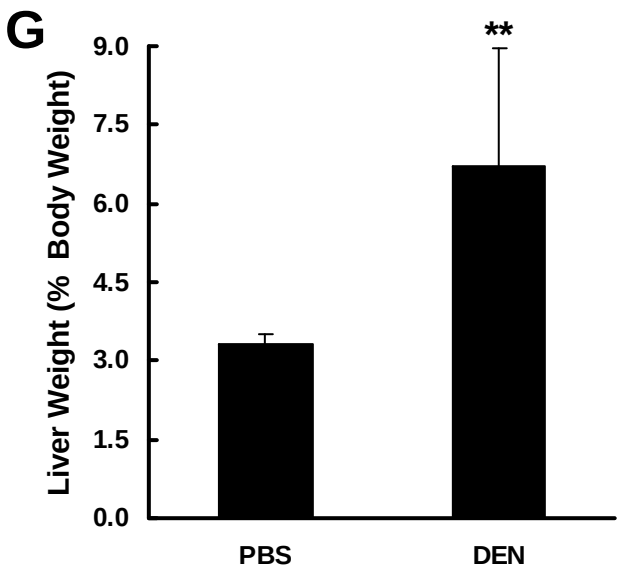
Supplementary Figure 1

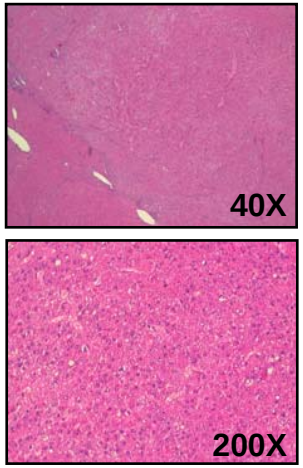




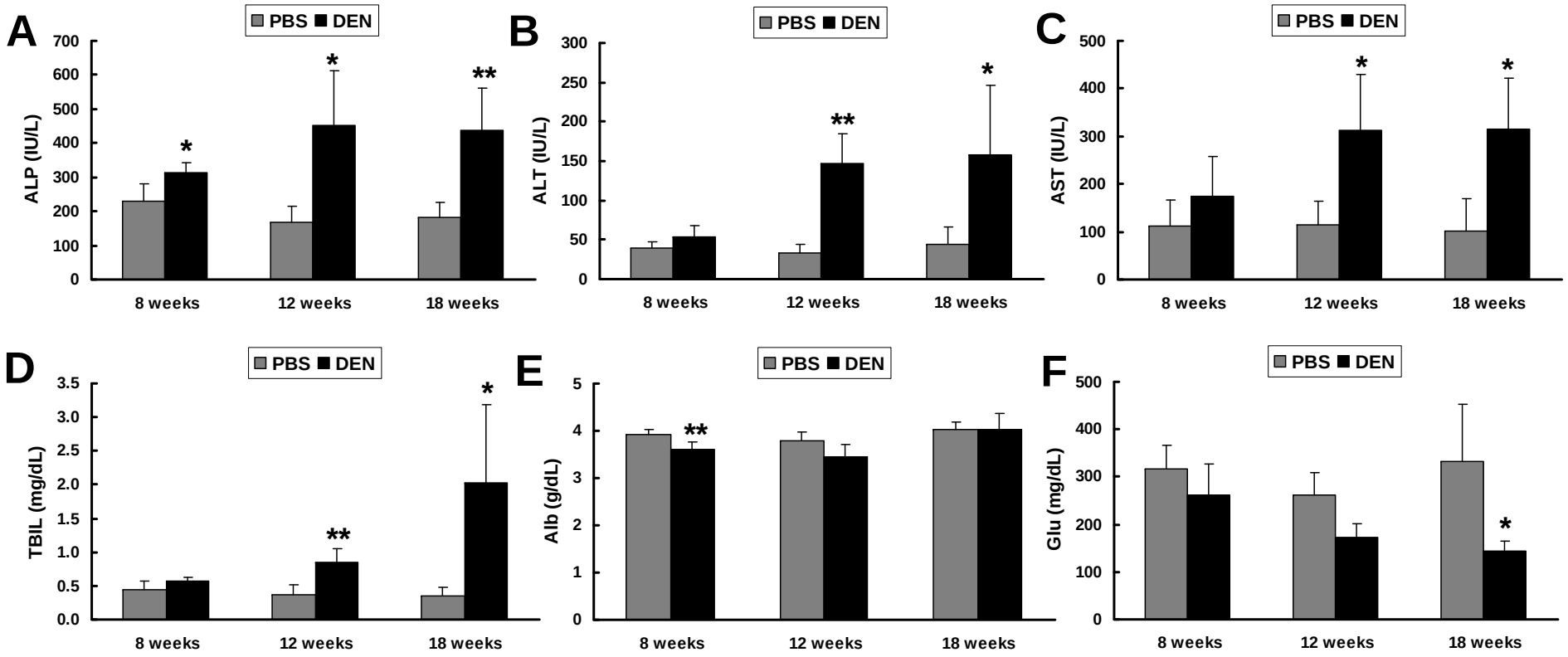




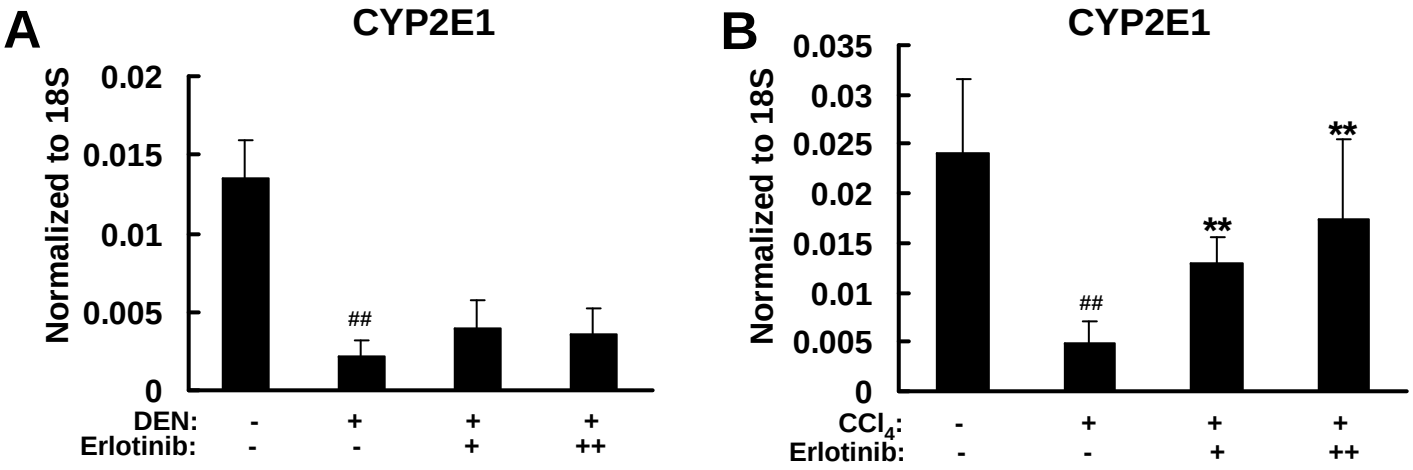




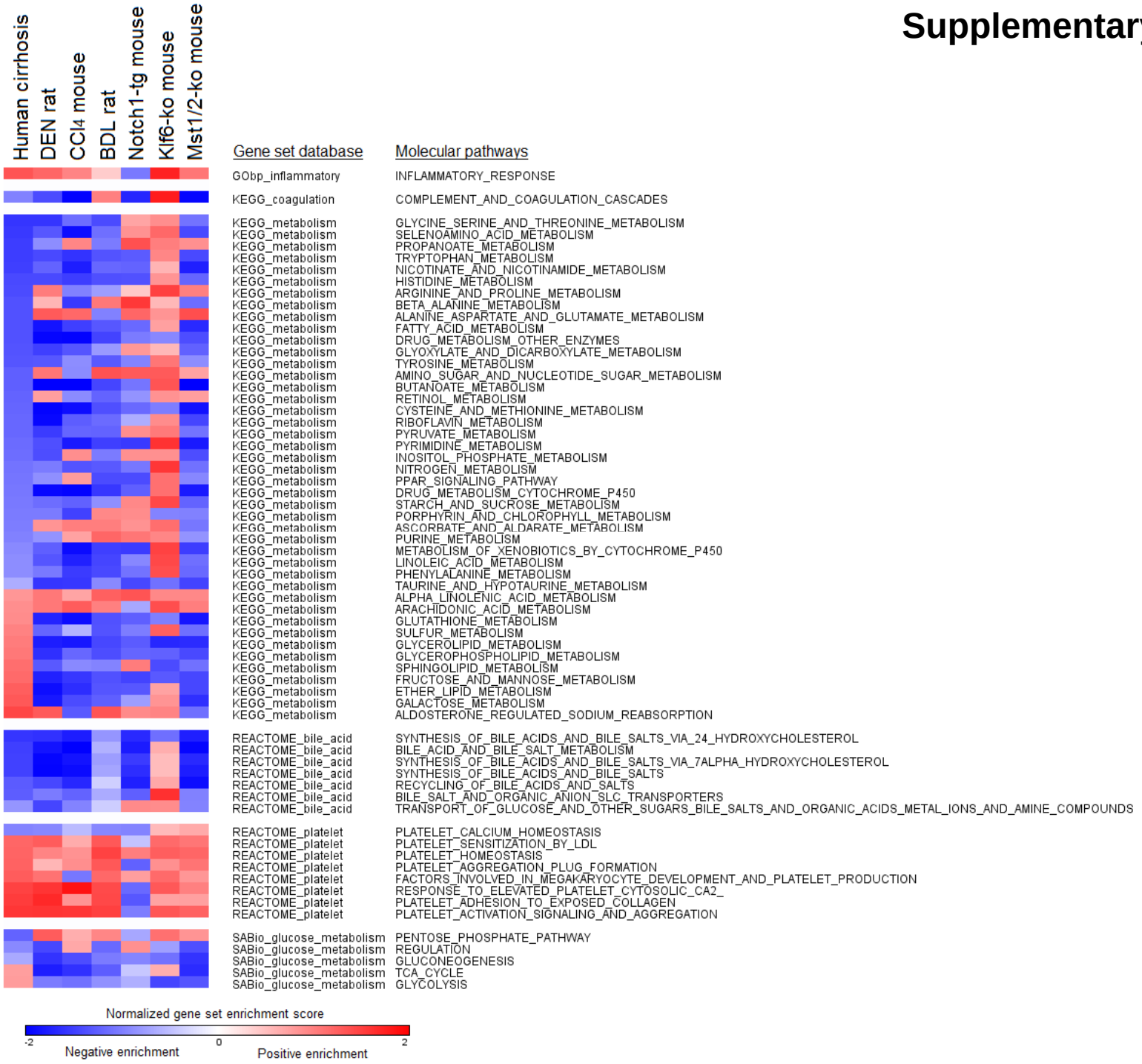
Supplementary Figure 2



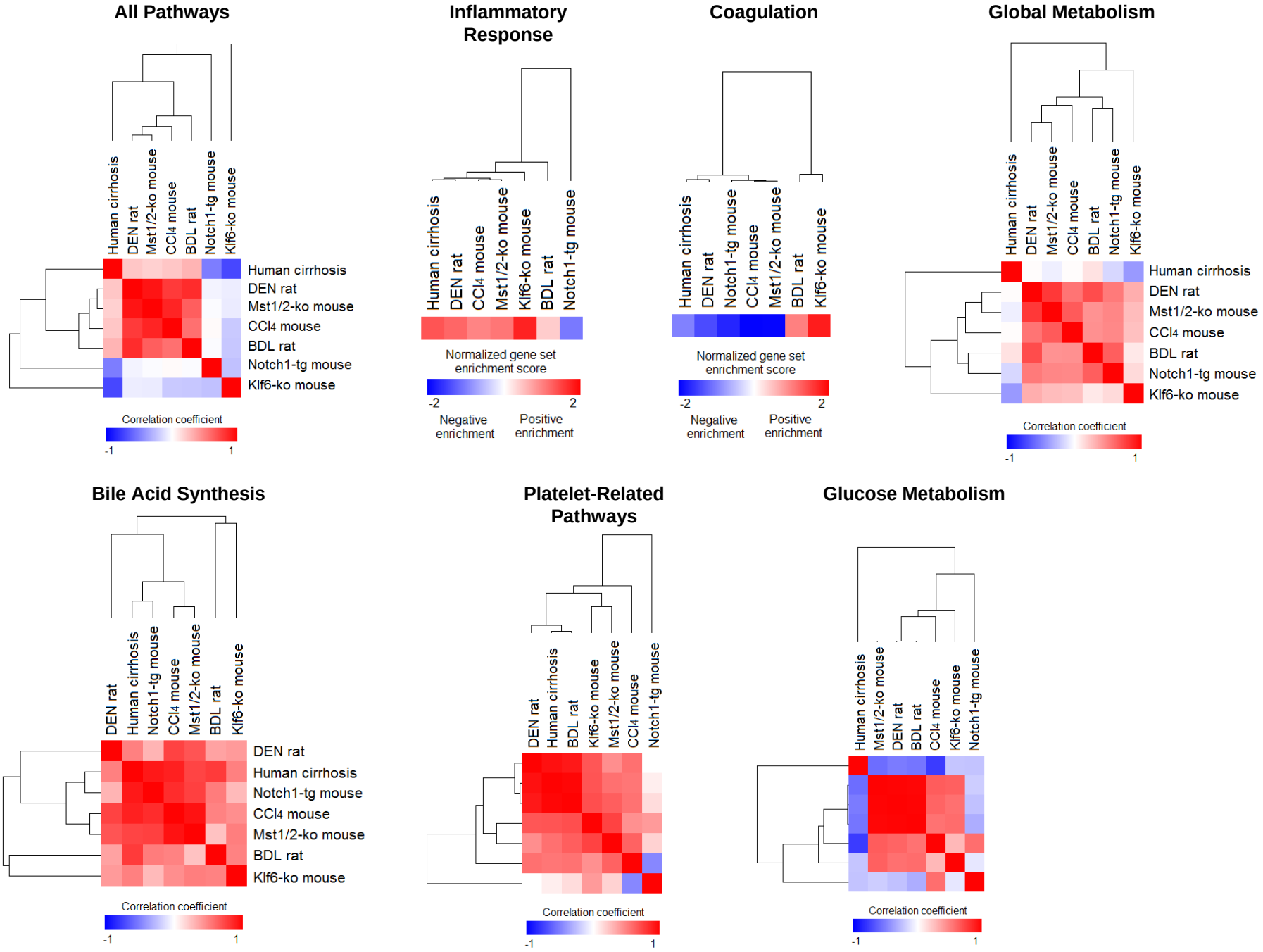
Supplementary Figure 3

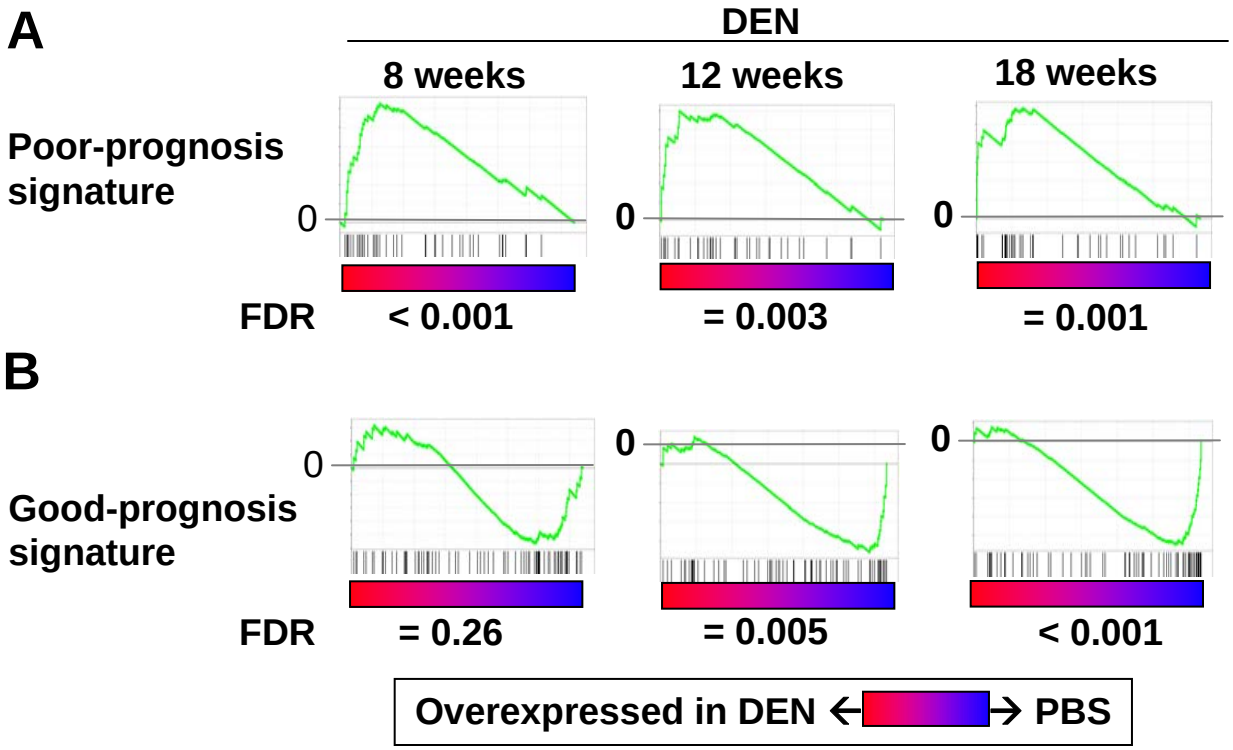


Supplementary Figure 4A

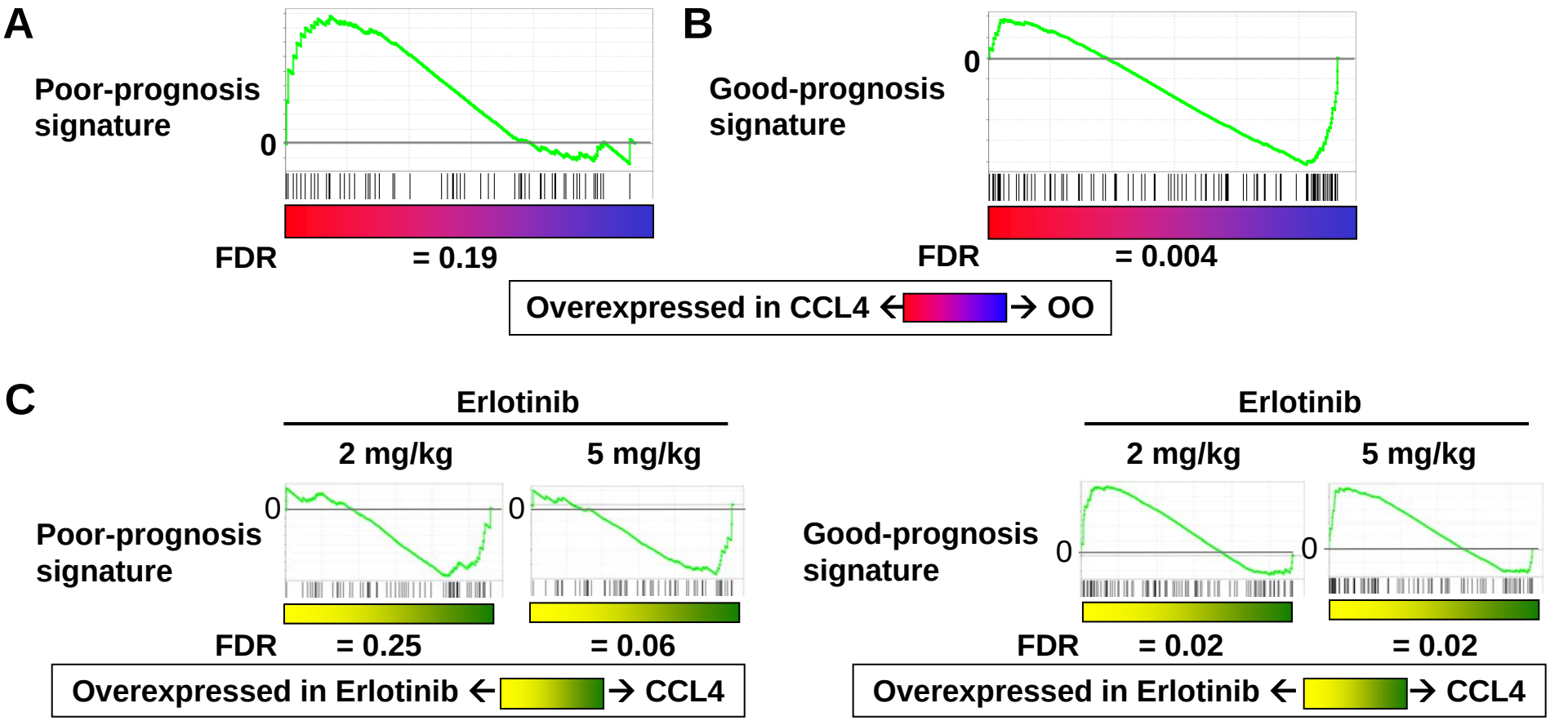


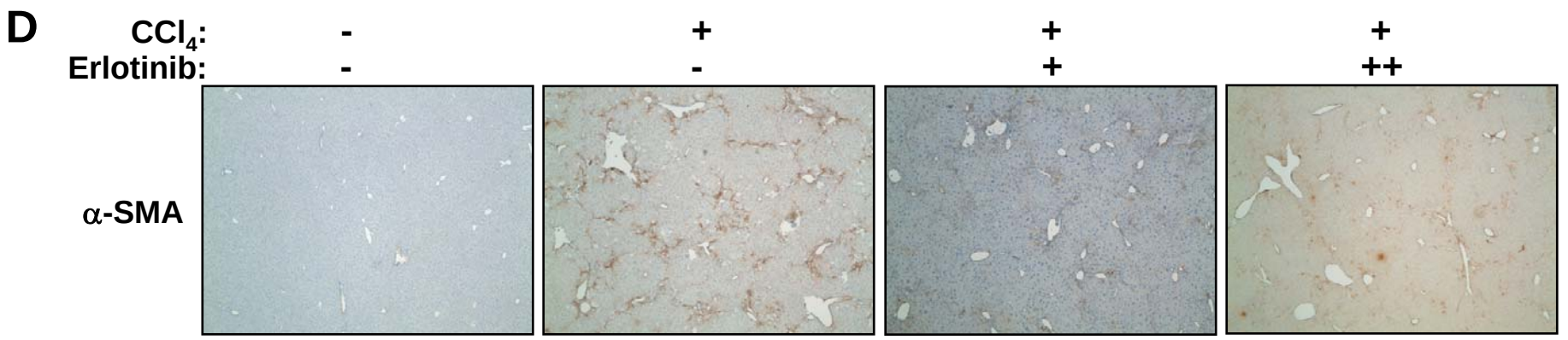
Supplementary Figure 4B



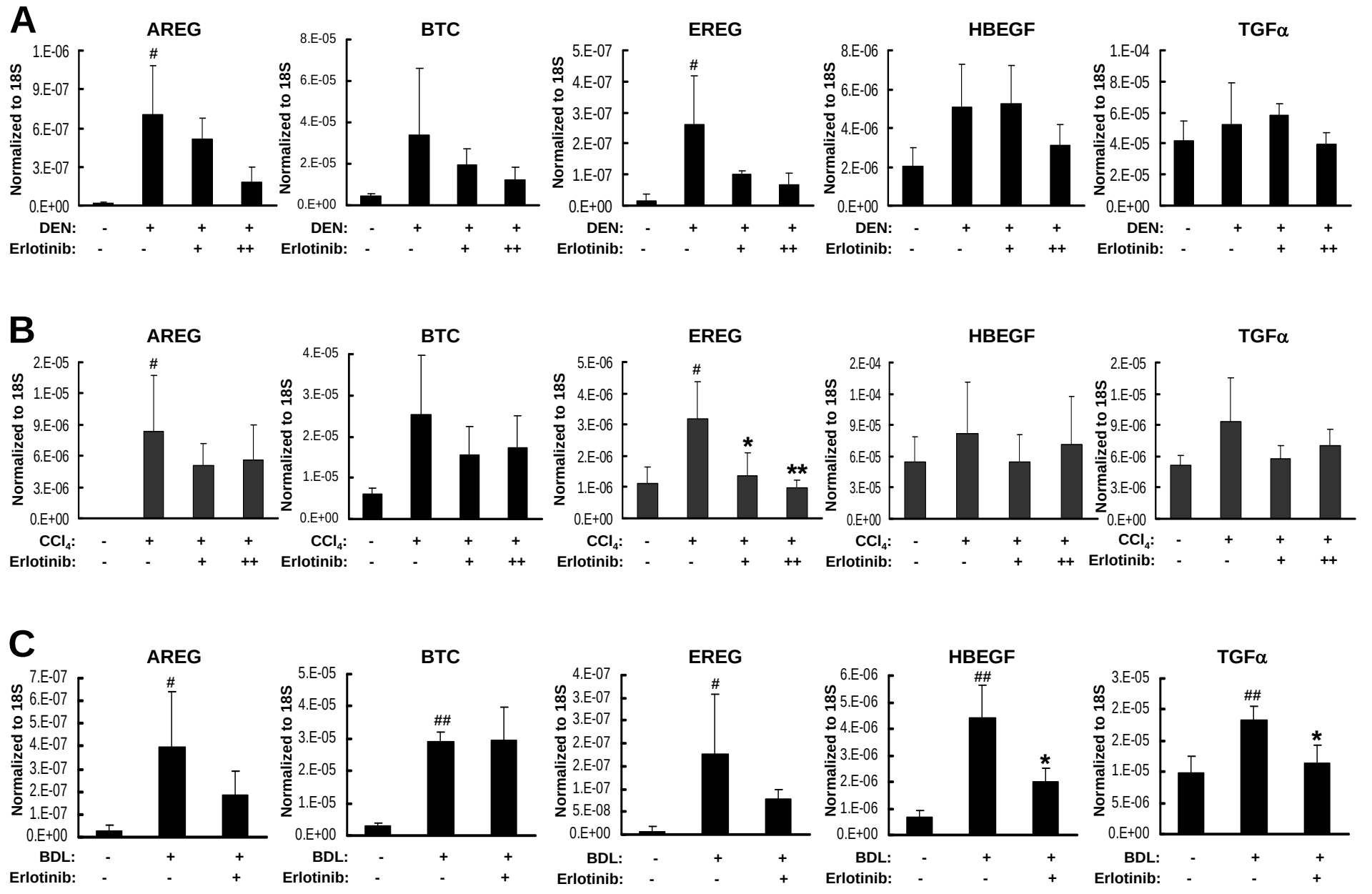

A
DEN
8 weeks
12 weeks
18 weeks
Poor-prognosis signature
0
0
0
FDR
< 0.001
= 0.003
= 0.001
B
0
0
0
Good-prognosis signature
FDR
= 0.26
= 0.005
< 0.001
Overexpressed in DEN
PBS

Supplementary Figure 6

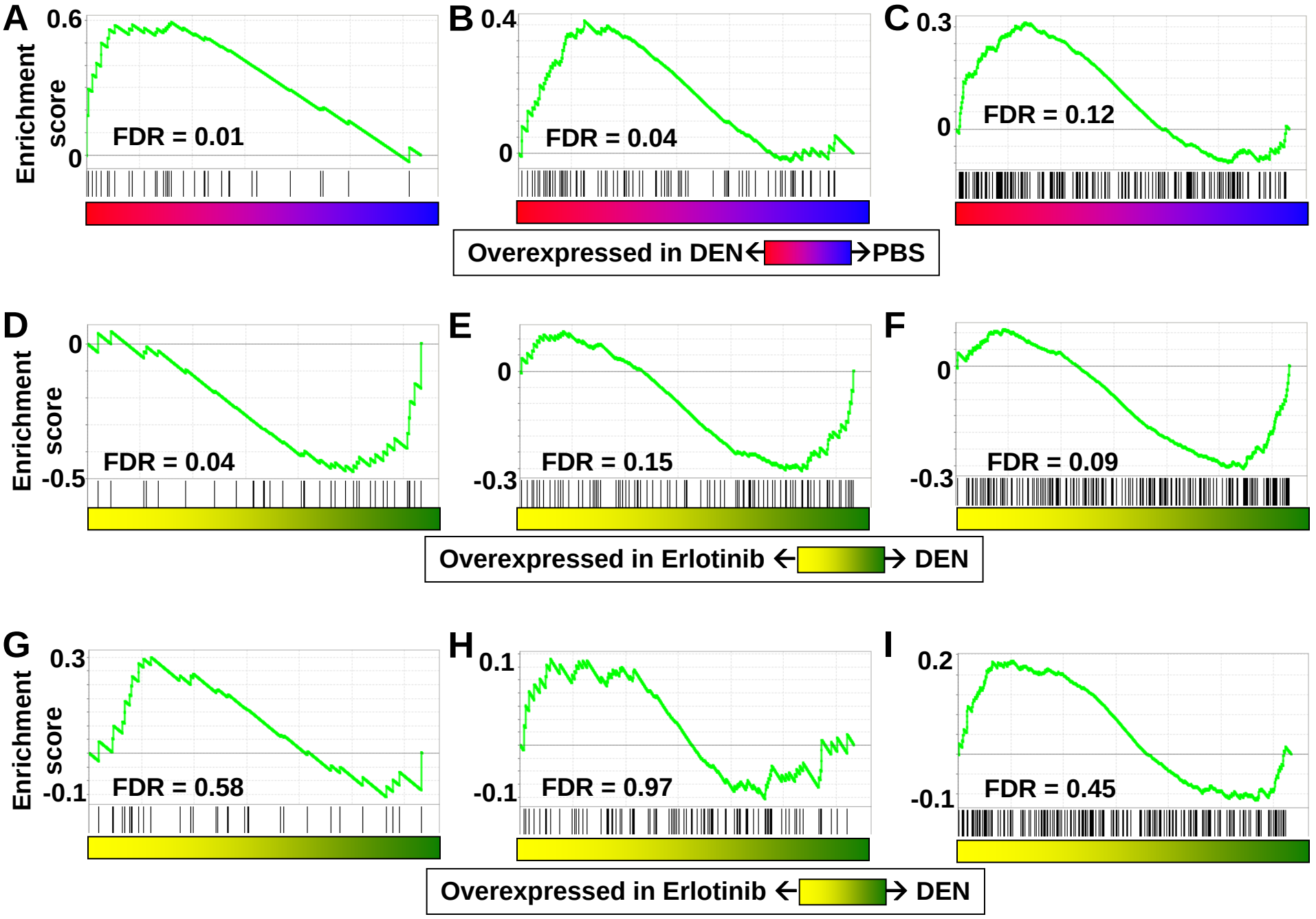




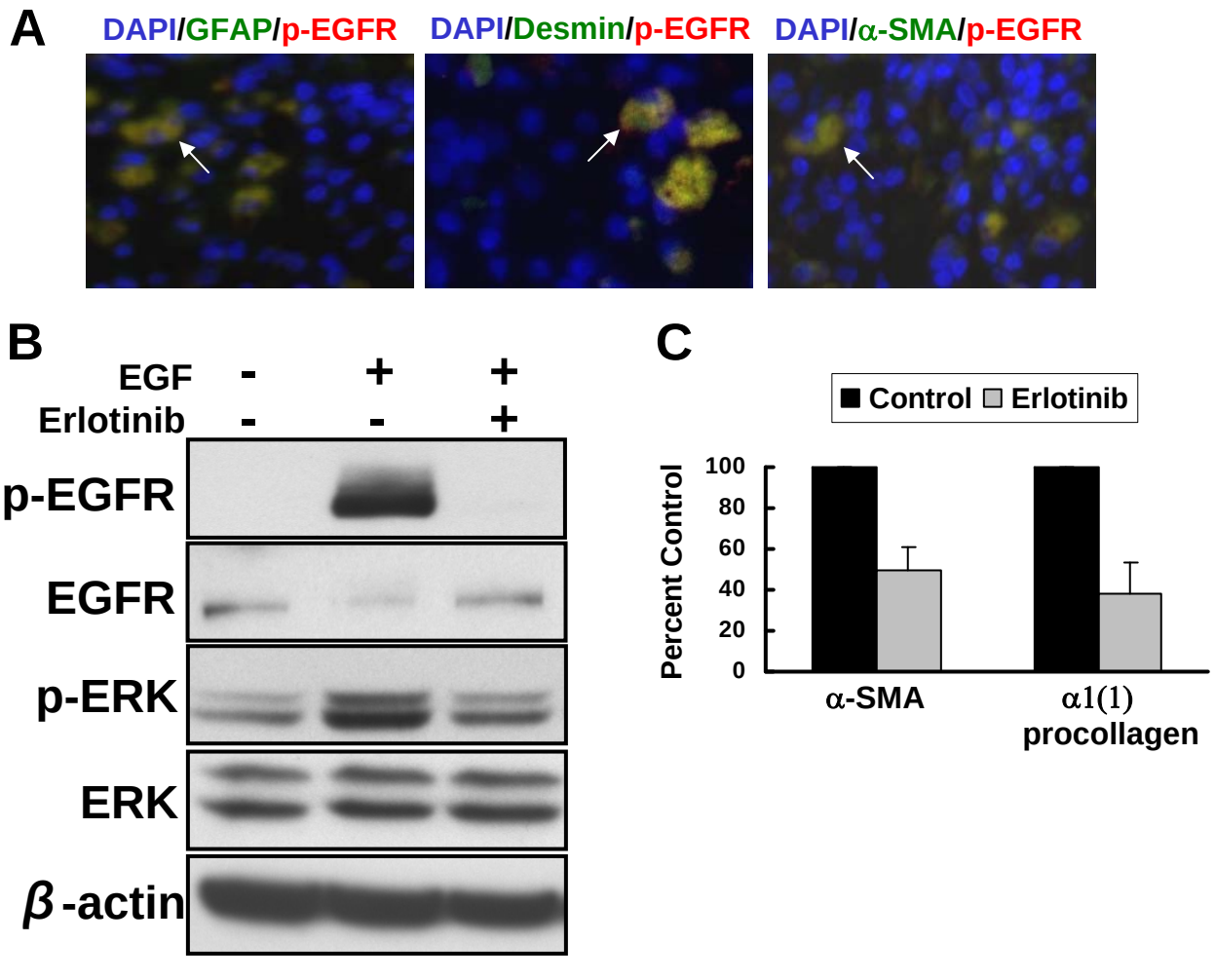
Supplementary Figure 7



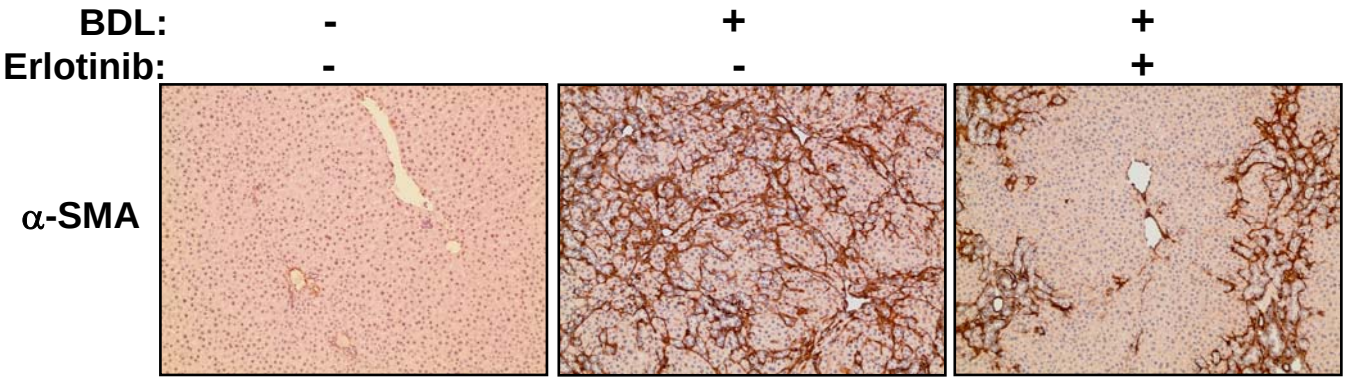
Supplementary Figure 8



Supplementary Figure 9



Supplementary Figure 10



Supplementary Figure 11