Supporting Information

Clinicopathological indices to predict hepatocellular carcinoma molecular classification

Poh Seng Tan, Shigeki Nakagawa, Nicolas Goossens, Anu Venkatesh, Tiangui Huang, Stephen C. Ward, Xiaochen Sun, Won-Min Song, Anna Koh, Claudia Canasto-Chibuque, Manjeet Deshmukh, Venugopalan Nair, Milind Mahajan, Bin Zhang, Maria Isabel Fiel, Masahiro Kobayashi, Hiromitsu Kumada, Yujin Hoshida

Contents

regulatory network module (no.6) identified by PFNA.

Figure S7. Serum AFP levels according to HCC molecular subclasses in the training set.

Figure S8. Correlation of clinicopathological features with HCC molecular subclasses (univariable logistic regression analysis). (A: Training set, B: Validation set)

Figure S9. Receiver operating characteristic (ROC) curves for the S1, S2, and S3 subclass predictive indices based on clinicopathological features.

Figure S10. Required sample size in clinical trial and NNT in HCC molecular subclass-targeting therapy.

Table S1
The 30-gene HCC molecular classification signature.

Subclass	GeneID	Symbol	Description
S1	8826	IQGAP1	IQ motif containing GTPase activating protein 1
S1	6282	S100A11	S100 calcium binding protein A11 (calgizzarin)
S1	11031	RAB31	RAB31, member RAS oncogene family
S1	951	CD37	CD37 molecule
S1	10631	POSTN	periostin, osteoblast specific factor
S1	397	ARHGDIB	Rho GDP dissociation inhibitor (GDI) beta
S1	241	ALOX5AP	arachidonate 5-lipoxygenase-activating protein
S1	7805	LAPTM5	lysosomal associated multispanning membrane protein 5
S1	1462	CSPG2	chondroitin sulfate proteoglycan 2 (versican)
S1	10109	ARPC2	actin related protein 2/3 complex, subunit 2, 34kDa
S2	1280	COL2A1	collagen, type II, alpha 1 (primary osteoarthritis, spondyloepiphyseal dysplasia, congenital)
S2	2719	GPC3	glypican 3
S2	174	AFP	alpha-fetoprotein
S2	191	AHCY	S-adenosylhomocysteine hydrolase
S2	6894	TARBP1	Tar (HIV-1) RNA binding protein 1
S2	1820	ARID3A	AT rich interactive domain 3A (BRIGHT- like)
S2	2261	FGFR3	fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism)
S2	6599	SMARCC1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1
S2	6193	RPS5	ribosomal protein S5
S2	1974	EIF4A2	eukaryotic translation initiation factor 4A, isoform 2
S3	10249	GLYAT	glycine-N-acyltransferase
S3	462	SERPINC1	serpin peptidase inhibitor, clade C (antithrombin), member 1
S3	346	APOC4	apolipoprotein C-IV
S3	4522	MTHFD1	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase
S3	2875	GPT	glutamic-pyruvate transaminase (alanine aminotransferase)
S3	3242	HPD	4-hydroxyphenylpyruvate dioxygenase
S3	710	SERPING1	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)
S3	1798	DPAGT1	dolichyl-phosphate (UDP-N-acetylglucosamine) N-acetylglucosaminephosphotransferase 1 (GlcNAc-1-P transferase)
S3	5105	PCK1	phosphoenolpyruvate carboxykinase 1 (soluble)
S3	3081	HGD	homogentisate 1,2-dioxygenase (homogentisate oxidase

Table S2

Summary of clinical demographics in 9 HCC datasets used for evaluation of HCC subclass predictive performace of the 30-gene signature compared to the 619-gene signature.

	HCC-A	HCC-B	HCC-C	HCC-D	HCC-E	HCC-F	HCC-G	HCC-H	HCC-I
No. of samples	90	82	60	43	33	31	57	89	118
Sex (male %)	79%	na	73%	na	97%	68%	82%	58%	81%
Age	53 (12)	na	63	na	51 (10)	64 (8)	62 (14)	65 (9)	58 (9)
Etiology									
HBV	59%	83%	18%	14%	97%	32%	30%	-	23%
HCV	9%	5%	67%	26%	-	68%	19%	100%	69%
HBV+HCV	3%	-	-	-	-	-	-	-	1%
Alcohol	6%	-	-	12%	-	-	32%	-	3%
Others	4%	11%	15%	2%	3%	-	5%	-	-
Cirrhosis	50%	na	60	na	82%	na	63%	na	70%
Tumor stage	na	na	I / II / IIIA[*] 22/28/10	na	na	na	na	0/A/B/C[†] 9/57/5/8	I /II / III / IV* 35/55/12/4
Histology	G2/3/4[‡] 8/31/1	na	G1/2/3[‡] 4/45/10	G1/2/3/4[‡] 10/22/10/1	M/M+P/P[†] 11/19/2	W/M/P[§] 10/13/8	G1/2/3/4[‡] 7/51/35/7	W/M/P[§] 20/41/17	W/M/P[§] 24/65/16
Tumor size (cm)	na	na	4.0	na	7.0 [3.0.5,9]	3.1 [2.2,4.3]	7.9 (5.1)	3.5 [2.5,5.6]	2.2 [1.8,3.0]
α-fetoprotein (ng/ml)	>300, 51%	na	na	na	113 [12,4051]	na	≥100, 33%	18 [7,182]	38 [10,171]
Data source	GSE1898	a)	b)	c)	GSE364	d)	E-TABM-36	GSE9843	GSE10186
Reference	Nat Genet 36;1306 (2004)	Mol Biol Cell 13;1929 (2002)	Lancet 361;923 (2003)	Cancer Res 64;6058 (2004)	Nat Med 9;416 (2003)	Cancer Res 64;7263 (2004)	Hepatology 45;42 (2007)	Cancer Res 68;6779 (2008)	Cancer Res 69;7385 (2009)

HCC: hepatocellular carcinoma, HBV: hepatitis B virus, HCV: hepatitis C virus, na: not available

[*]: Liver Cancer Study Group of Japan, [†]: Bercelona Clinic Liver Cancer staging system,

[‡]: Edmondson-Steiner grade, [§]: UICC tumor differentiation (Well / Moderate /Poor),

Mean (SD, if available), Median [25%, 75%]

a) http://smd.stanford.edu/cgi-bin/publication/viewPublication.pl?pub_no=107

b) http://surgery2.med.yamaguchi-u.ac.jp/research/DNAchip/hcc-recurrence/index.html

c) http://www.medizin.uni-koeln.de/kliniken/pathologie/supplementaldata.html

d) http://www2.genome.rcast.u-tokyo.ac.jp/HCC/

Table S3

RT-qPCR primers.

Gene	Forward	Reverse
KRT19	CGACTACAGCCACTACTA	CAGACGGGCATTGTCGATC
EPCAM	ATAACCTGCTCTGAGCGAGTG	TGAAGTGCAGTCCGCAAACT
GLUL	AAGTGTGTGGAAGAGTTGCC	TGCTCACCATGTCCATTATC
PRL13A (housekeeping)	GTACGCTGTGAAGGCATCAA	GTTGGTGTTCATCCGCTTG

Table S4

Liver cancer-related gene signatures in literature.

Molecular signature	Type of correlate	Reference
Hoshida09_subclass	Molecular feature (classification)	Cancer Res 69;7385 (2009)
Boyault07_subclass	Molecular feature (classification)	Hepatology 45;42 (2007)
Lee04_subclass	Molecular feature (classification)	Hepatology 40;667 (2004)
Chiang08_subclass	Molecular feature (classification)	Cancer Res 68;6779 (2008)
Andersen10_KRT19	Molecular feature	Hepatology 51;1401 (2010)
Coulouarn08_TGFbeta	Molecular feature	Hepatology 47;2059 (2008)
KaposiNovak06_MET	Molecular feature	J Clin Invest 116;1582 (2006)
Villanueva11_KRT19	Molecular feature	Gastroenterology 140;1501 (2011)
Yamashita08_EpCAM	Molecular feature	Cancer Res 68;1451 (2008)
Andersen12_cholangio_subclass	Biological phenotype	Gastroenterology 142;1021 (2012)
Cairo08_hepatoblast_like	Biological phenotype	Cancer cell 14;471 (2008)
Lee06_progenitor_cell	Biological phenotype	Nat Med 12;410 (2006)
Oishi12_cholangio_stem_cell	Biological phenotype	Hepatology 56;1792 (2012)
Woo10_cholangio_like	Biological phenotype	Cancer Res 70;3034 (2010)
Yamashita09_cancer_stem_cell	Biological phenotype	Gastroenterology 136;1012 (2009)
Minguez11_vascular_invasion	Clinical phenotype	J Hepatol 55;1325 (2011)
Ye03_metastasis	Clinical phenotype	Nat Med 9;416 (2003)
Izuka03_early_recurrence	Clinical outcome	Lancet 361;923 (2003)
Kim12_survival	Clinical outcome	Hepatology 55;1443 (2012)
Kurokawa04_early_recurrence	Clinical outcome	J Hepatol 41;284 (2004)
Roessler10_metastasis	Clinical outcome	Cancer Res 70;10202 (2010)
Wang07_recurrence	Clinical outcome	Clin Cancer Res 13;6275 (2007)
Woo08_recurrence	Clinical outcome	Clin Cancer Res 14;2056 (2008)

Table S5

Concordance of histopathololological determination between evaluators.

		Training set (96 HCC tumors)	Validation set (99 HCC tumors)
Architectual pattern			
	Microtrabicular	73 (76%)	81 (82%)
	Macrotrabicular	78 (81%)	79 (80%)
	Compact	80 (83%)	77 (78%)
	Pseudoglandular	74 (77%)	79 (80%)
Cytological variant			
	SH-HCC	79 (82%)	86 (87%)
	Clear cell	77 (80%)	91 (92%)
	Fatty change	84 (88%)	92 (93%)
	Edmondson-Steiner grade I/II	79 (82%)	79 (80%)
	Osteoclastic cell	92 (96%)	96 (97%)
Immune cell infiltrate, score 0/1*		74 (77%)	72 (73%)

*Immune cell infiltrate score determned as 0: 0 t o5 cells, 1: 5 to 20 cells, 2: greater than 20 cells.

HCC: hepatocellular carcinoma, SH-HCC: steatohepatitic HCC

Table S6

HCC molecular subclasses and tumor characteristics of SH-HCC and non-SH-HCC variants.

	Training and validation sets			Training set			Validation set		
	SH-HCC (36 tumors)	Non-SH-HCC (159 tumors)	P value	SH-HCC (19 tumors)	Non-SH-HCC (77 tumors)	P value	SH-HCC (17 tumors)	Non-SH-HCC (82 tumors)	P value
Molecular subclass, n (%)			<0.001			0.01			0.008
S1	21 (58%)	39 (25%)		11 (58%)	19 (25%)		10 (59%)	20 (24%)	
S2/S3	15 (42%)	120 (75%)		8 (42%)	58 (75%)		7 (41%)	62 (76%)	
AFP (ng/mL), median (IQR)	20 (6-92)	24 (6-142)	0.63	8 (5-62)	16 (4-75)	0.92	26 (9-118)	43 (10-197)	0.63
> 400 ng/mL	3 (8.3%)	20 (12.6%)	0.58	1 (5%)	7 (9%)	1.00	2 (12%)	13 (16%)	1.00
Tumor size (cm), median (IQR)	2.0 (1.8-3.0)	2.3 (2.0-3.0)	0.46	2.3 (2.0-3.1)	2.4 (2.0-2.9)	0.96	1.9 (1.7-2.9)	2.3 (1.8-3.0)	0.30
Tumor number, n (%)									
Single	32 (89%)	123 (77%)	0.92	18 (95%)	64 (83%)	0.56	14 (82%)	59 (72%)	0.84
Multiple	4 (11%)	24 15%)		1 (5%)	13 (17%)		3 (18%)	11 (13%)	
Edmondson-Steiner grade, n (%)									
I	4 (11%)	29 (18%)	0.17	1 (5%)	11 (14%)	0.06	3 (18%)	18 (22%)	0.80
II	23 (64%)	105 (66%)		11 (58%)	55 (71%)		12 (71%)	50 (61%)	
III	8 (22%)	25 (16%)		6 (32%)	11 (14%)		2 (12%)	14 (17%)	
IV	1 (3%)	0 (0%)		1 (5%)	0 (0%)		0 (0%)	0 (0%)	
Steatosis in non-tumor liver*, n (%)									
Absent/minimal	11 (31%)	101 (64%)	0.001	6 (32%)	46 (60%)	0.01	5 (29%)	55 (67%)	0.02
Mild	16 (44%)	41 (26%)		8 (42%)	24 (31%)		8 (47%)	17 (21%)	
Moderate	6 (17%)	14 (9%)		3 (16%)	7 (9%)		3 (18%)	7 (9%)	
Severe	3 (8%)	3 (2%)		2 (11%)	0 (0%)		1 (6%)	3 (4%)	
Steatohepatitis in non-tumor liver, n (%)	14 (39%)	11 (7%)	<0.001	6 (32%)	5 (6%)	0.007	8 (47%)	6 (7%)	<0.001

HCC: hepatocellular carcinoma, SH-HCC: steatohepatitic HCC, AFP: alpha-fetoprotein.

*Absent/minimal <5%; Mild = 5%-33%; moderate = 34%-66%; severe > 66%.

Table S7
Gene sets associated with histopathological features in the validation set (Gene Set Enrichment Analysis).

Histopathological features	Gene set database	Gene set	NES	p-value	FDR
Microtrabecular	Kyoto Encyclopedia of Genes and Genomes (KEGG) (Curated pathway database)	KEGG BUTANOATE METABOLISM	2.42	0.000	0.000
		KEGG TRYPTOPHAN METABOLISM	2.42	0.000	0.000
		KEGG TYROSINE METABOLISM	2.35	0.000	0.000
		KEGG FATTY ACID METABOLISM	2.31	0.000	0.000
		KEGG VALINE LEUCINE AND ISOLEUCINE DEGRADATION	2.17	0.000	0.000
		KEGG LYSINE DEGRADATION	2.06	0.000	0.001
		KEGG PROPANOATE METABOLISM	2.05	0.000	0.001
		KEGG STEROID HORMONE BIOSYNTHESIS	2.00	0.000	0.001
		KEGG DRUG METABOLISM CYTOCHROME P450	1.98	0.000	0.002
		KEGG METABOLISM OF XENOBIOTICS BY CYTOCHROME P450	1.94	0.000	0.003
		KEGG RETINOL METABOLISM	1.92	0.002	0.003
		KEGG PYRUVATE METABOLISM	1.89	0.000	0.005
		KEGG RIG I LIKE RECEPTOR SIGNALING PATHWAY	1.87	0.002	0.006
		KEGG GLYCINE SERINE AND THREONINE METABOLISM	1.86	0.004	0.006
		KEGG BETA ALANINE METABOLISM	1.79	0.005	0.012
		KEGG HISTIDINE METABOLISM	1.78	0.002	0.013
		KEGG PHENYLALANINE METABOLISM	1.78	0.009	0.013
		KEGG GLYCOLYSIS GLUCONEOGENESIS	1.78	0.000	0.013
		KEGG CITRATE CYCLE TCA CYCLE	1.77	0.004	0.012
		KEGG PPAR SIGNALING PATHWAY	1.76	0.005	0.013
		KEGG GLYCEROLIPID METABOLISM	1.75	0.007	0.014
		KEGG ALLOGRAFT REJECTION	1.73	0.005	0.015
		KEGG PEROXISOME	1.73	0.003	0.015
		KEGG ABC TRANSPORTERS	1.68	0.013	0.021
		KEGG TYPE I DIABETES MELLITUS	1.65	0.009	0.029
		KEGG GRAFT VERSUS HOST DISEASE	1.64	0.013	0.030
		KEGG ADIPOCYTOKINE SIGNALING PATHWAY	1.62	0.015	0.035
		KEGG LINOLEIC ACID METABOLISM	1.60	0.024	0.041
		KEGG JAK STAT SIGNALING PATHWAY	1.58	0.003	0.047
		KEGG AUTOIMMUNE THYROID DISEASE	1.57	0.010	0.047
	Reactome (Curated pathway database)	REACTOME INTERFERON ALPHA BETA SIGNALING	2.28	0.000	0.000
		REACTOME PHASE1 FUNCTIONALIZATION OF COMPOUNDS	2.07	0.000	0.005
		REACTOME TRANSPORT OF GLUCOSE AND OTHER SUGARS BILE SALTS AND ORGANIC ACIDS METAL IONS AND AMINE COMPOUNDS	1.99	0.000	0.011
		REACTOME BILE ACID AND BILE SALT METABOLISM	1.97	0.000	0.012
		REACTOME BIOLOGICAL OXIDATIONS	1.96	0.000	0.010
		REACTOME CYTOCHROME P450 ARRANGED BY SUBSTRATE TYPE	1.96	0.002	0.008
		REACTOME RIG I MDA5 MEDIATED INDUCTION OF IFN ALPHA BETA PATHWAYS	1.86	0.002	0.024
	Gene Ontology - Biological process	NITROGEN COMPOUND CATABOLIC PROCESS	2.00	0.000	0.030
		AROMATIC COMPOUND METABOLIC PROCESS	2.00	0.002	0.015
		AMINE CATABOLIC PROCESS	1.93	0.000	0.028
		ORGANIC ACID METABOLIC PROCESS	1.89	0.000	0.036
		CARBOXYLIC ACID METABOLIC PROCESS	1.88	0.000	0.037
	- Molecular function	OXIDOREDUCTASE ACTIVITY	2.35	0.000	0.000
		OXIDOREDUCTASE ACTIVITY ACTING ON THE ALDEHYDE OR OXO GROUP OF DONORS	2.24	0.000	0.000
		ELECTRON CARRIER ACTIVITY	2.03	0.000	0.005
		MONOOXYGENASE ACTIVITY	1.95	0.000	0.011
	- Cellular component	MITOCHONDRION	2.07	0.000	0.001
	Chemical and genetic perturbations (Experimentally defined target gene signature in literature)	CHIANG LIVER CANCER SUBCLASS PROLIFERATION DN	3.50	0.000	0.000
		HOSHIDA LIVER CANCER SUBCLASS S3	3.18	0.000	0.000
		LEE LIVER CANCER SURVIVAL UP	3.07	0.000	0.000
		CHIANG LIVER CANCER SUBCLASS CTNNB1 UP	3.05	0.000	0.000
		WOO LIVER CANCER RECURRENCE DN	3.00	0.000	0.000
		HSIAO LIVER SPECIFIC GENES	2.78	0.000	0.000
		BOYAULT LIVER CANCER SUBCLASS G123 DN	2.76	0.000	0.000
		CAIRO HEPATOBLASTOMA CLASSES DN	2.76	0.000	0.000
		CAIRO HEPATOBLASTOMA DN	2.66	0.000	0.000
		BROWNE INTERFERON RESPONSIVE GENES	2.57	0.000	0.000
		BOYAULT LIVER CANCER SUBCLASS G3 DN	2.52	0.000	0.000
		OHGUCHI LIVER HNF4A TARGETS DN	2.51	0.000	0.000
		HOSHIDA LIVER CANCER LATE RECURRENCE DN	2.48	0.000	0.000
		DAUER STAT3 TARGETS DN	2.45	0.000	0.000
		MOSERLE IFNA RESPONSE	2.45	0.000	0.000
		FARMER BREAST CANCER CLUSTER 1	2.42	0.000	0.000
		BOWIE RESPONSE TO TAMOXIFEN	2.41	0.000	0.000
		BOWIE RESPONSE TO EXTRACELLULAR MATRIX	2.39	0.000	0.000
		RADAEVA RESPONSE TO IFNA1 UP	2.32	0.000	0.000
		BOSCO INTERFERON INDUCED ANTIVIRAL MODULE	2.31	0.000	0.000
		SEITZ NEOPLASTIC TRANSFORMATION BY 8P DELETION UP	2.29	0.000	0.000
		HOSHIDA LIVER CANCER SURVIVAL DN	2.28	0.000	0.000
		UROSEVIC RESPONSE TO IMIQUIMOD	2.28	0.000	0.000
		HECKER IFNB1 TARGETS	2.24	0.000	0.000
		ZHU CMV 8 HR UP	2.20	0.000	0.000
		ZHANG INTERFERON RESPONSE	2.20	0.000	0.000
		EINAV INTERFERON SIGNATURE IN CANCER	2.19	0.000	0.000
		PIONTEK PKD1 TARGETS UP	2.14	0.000	0.000
		CHIANG LIVER CANCER SUBCLASS POLYSOMY7 UP	2.14	0.000	0.000
		LEE LIVER CANCER DENA DN	2.12	0.000	0.000
		YAMASHITA LIVER CANCER STEM CELL DN	2.12	0.000	0.001
		IIZUKA LIVER CANCER PROGRESSION G2 G3 UP	2.12	0.000	0.001
		SU LIVER	2.11	0.000	0.001
		LEE LIVER CANCER	2.09	0.000	0.001
		TAKEDA TARGETS OF NUP98 HOXA9 FUSION 8D UP	2.09	0.000	0.001
		LEE LIVER CANCER CIPROFIBRATE DN	2.08	0.000	0.001
		TAKEDA TARGETS OF NUP98 HOXA9 FUSION 3D UP	2.05	0.000	0.002
		BENNETT SYSTEMIC LUPUS ERYTHEMATOSUS	2.05	0.000	0.002
		SANA RESPONSE TO IFNG UP	2.04	0.000	0.002
		BOYAULT LIVER CANCER SUBCLASS G6 UP	2.02	0.000	0.002
		LEE LIVER CANCER ACOX1 DN	2.02	0.000	0.003
		TAKEDA TARGETS OF NUP98 HOXA9 FUSION 10D UP	2.00	0.000	0.003
		CAIRO LIVER DEVELOPMENT DN	2.00	0.000	0.003
		KRASNOSELSKAYA ILF3 TARGETS UP	2.00	0.000	0.003
		MAHADEVAN RESPONSE TO MP470 UP	1.99	0.002	0.004
		LEE LIVER CANCER MYC E2F1 DN	1.95	0.000	0.005
		GRAESSMANN RESPONSE TO MC AND SERUM DEPRIVATION UP	1.93	0.000	0.007
		COULOUARN TEMPORAL TGFB1 SIGNATURE DN	1.92	0.000	0.008

		STAMBOLSKY TARGETS OF MUTATED TP53 DN	1.92	0.000	0.008
		LEE LIVER CANCER E2F1 DN	1.91	0.000	0.008
	Liver cancer gene signatures in literature	WOO10_CHOLANGIO_LIKE_DOWN	3.41	0.000	0.000
		HOSHIDA09_SUBCLASS_S3	3.17	0.000	0.000
		CHIANG08_SUBCLASS_CTNNB1	3.06	0.000	0.000
		LEE04_SUBCLASS_B	3.05	0.000	0.000
		WOO08_RECURRENCE_GOOD	2.99	0.000	0.000
		VILLANUEVA11_KRT19_NEGATIVE	2.47	0.000	0.000
		ANDERSEN10_KRT19_NEGATIVE_CLSB	2.46	0.000	0.000
		KIM12_SURVIVAL_GOOD	2.42	0.000	0.000
		CHIANG08_SUBCLASS_POLY7	2.14	0.000	0.000
		YAMASHITA09_CANCER_STEM_CELL_DOWN	2.10	0.000	0.000
		BOYAULT07_SUBCLASS_G6	2.02	0.000	0.000
		COULOUARN08_TGFB1_DOWN	1.95	0.000	0.001
		CAIRO08_HEPATOBLAST_LIKE_GOOD_CLS1	1.86	0.000	0.002
		BOYAULT07_SUBCLASS_G5G6	1.70	0.004	0.008
		MINGUEZ11_VASCULAR_INVASION_NEGATIVE	1.65	0.008	0.012
		KAPOSINOVAK06_MET_DOWN	1.64	0.013	0.012
		YAMASHITA08_EPCAM_DOWN	1.58	0.034	0.022
		WANG07_RECURRENCE_GOOD	1.56	0.042	0.026
Macrotrabecular/compact	Kyoto Encyclopedia of Genes and Genomes (KEGG)	KEGG DNA REPLICATION	1.88	0.002	0.040
	(Curated pathway database)	KEGG SPLICEOSOME	1.82	0.003	0.040
		KEGG CELL CYCLE	1.81	0.000	0.037
		KEGG VIBRIO CHOLERAE INFECTION	1.74	0.005	0.048
	Reactome	REACTOME CELL CYCLE	2.40	0.000	0.000
	(Curated pathway database)	REACTOME CELL CYCLE MITOTIC	2.39	0.000	0.000
		REACTOME MITOTIC PROMETAPHASE	2.26	0.000	0.000
		REACTOME MITOTIC M M G1 PHASES	2.24	0.000	0.001
		REACTOME MEIOTIC RECOMBINATION	2.22	0.000	0.001
		REACTOME CLEAVAGE OF GROWING TRANSCRIPT IN THE TERMINATION REGION	2.22	0.000	0.001
		REACTOME RNA POL I PROMOTER OPENING	2.20	0.000	0.001
		REACTOME DNA REPLICATION	2.19	0.000	0.001
		REACTOME DEPOSITION OF NEW CENPA CONTAINING NUCLEOSOMES AT THE CENTROMERE	2.16	0.000	0.001
		REACTOME CHROMOSOME MAINTENANCE	2.11	0.000	0.002
		REACTOME MEIOSIS	2.11	0.000	0.002
		REACTOME RNA POL I TRANSCRIPTION	2.10	0.000	0.002
		REACTOME GLYCOSPHINGOLIPID METABOLISM	2.07	0.000	0.003
		REACTOME METABOLISM OF NUCLEOTIDES	2.01	0.003	0.005
		REACTOME RECRUITMENT OF MITOTIC CENTROSOME PROTEINS AND COMPLEXES	2.00	0.000	0.005
		REACTOME PROTEIN FOLDING	1.99	0.000	0.006
		REACTOME TRANS GOLGI NETWORK VESICLE BUDDING	1.97	0.000	0.006
		REACTOME PREFOLDIN MEDIATED TRANSFER OF SUBSTRATE TO CCT TRIC	1.97	0.002	0.006
		REACTOME MEIOTIC SYNAPSIS	1.95	0.000	0.007
		REACTOME S PHASE	1.95	0.000	0.006
		REACTOME G2 M CHECKPOINTS	1.95	0.000	0.006
		REACTOME MITOTIC G2 G2 M PHASES	1.95	0.000	0.006
		REACTOME PACKAGING OF TELOMERE ENDS	1.94	0.005	0.006
		REACTOME TRANSCRIPTION	1.93	0.000	0.007
		REACTOME MEMBRANE TRAFFICKING	1.90	0.000	0.009
		REACTOME TELOMERE MAINTENANCE	1.90	0.000	0.009
		REACTOME ACTIVATION OF ATR IN RESPONSE TO REPLICATION STRESS	1.90	0.005	0.009
		REACTOME METABOLISM OF RNA	1.88	0.000	0.010
		REACTOME MITOTIC G1 G1 S PHASES	1.88	0.000	0.009
		REACTOME G0 AND EARLY G1	1.88	0.000	0.009
		REACTOME LOSS OF NLP FROM MITOTIC CENTROSOMES	1.87	0.005	0.010
		REACTOME SPHINGOLIPID METABOLISM	1.87	0.000	0.009
		REACTOME G1 S TRANSITION	1.87	0.000	0.010
		REACTOME RNA POL I RNA POL III AND MITOCHONDRIAL TRANSCRIPTION	1.84	0.003	0.013
		REACTOME E2F MEDIATED REGULATION OF DNA REPLICATION	1.84	0.000	0.013
		REACTOME APOPTOTIC EXECUTION PHASE	1.82	0.003	0.015
		REACTOME METABOLISM OF NON CODING RNA	1.81	0.005	0.016
		REACTOME HIV LIFE CYCLE	1.79	0.000	0.019
		REACTOME SIGNALING BY NOTCH1	1.79	0.000	0.019
		REACTOME CELL CYCLE CHECKPOINTS	1.79	0.003	0.018
		REACTOME DOUBLE STRAND BREAK REPAIR	1.78	0.007	0.020
		REACTOME RNA POL II TRANSCRIPTION	1.75	0.003	0.024
		REACTOME DNA STRAND ELONGATION	1.75	0.005	0.024
		REACTOME ASSEMBLY OF THE PRE REPLICATIVE COMPLEX	1.74	0.000	0.024
		REACTOME PROCESSING OF CAPPED INTRON CONTAINING PRE MRNA	1.74	0.000	0.024
		REACTOME LATE PHASE OF HIV LIFE CYCLE	1.74	0.003	0.024
		REACTOME GOLGI ASSOCIATED VESICLE BIOGENESIS	1.73	0.008	0.025
		REACTOME APOPTOTIC CLEAVAGE OF CELLULAR PROTEINS	1.73	0.007	0.025
		REACTOME AMYLOIDS	1.73	0.003	0.025
		REACTOME APC C CDC20 MEDIATED DEGRADATION OF MITOTIC PROTEINS	1.73	0.003	0.024
	Gene Ontology - Biological process	CHROMOSOME SEGREGATION	2.38	0.000	0.001
		RNA PROCESSING	2.35	0.000	0.001
		MICROTUBULE CYTOSKELETON ORGANIZATION AND BIOGENESIS	2.30	0.000	0.000
		RNA SPLICING	2.17	0.000	0.003
		MICROTUBULE BASED PROCESS	2.11	0.000	0.005
		M PHASE OF MITOTIC CELL CYCLE	2.10	0.000	0.004
		CHROMOSOME ORGANIZATION AND BIOGENESIS	2.10	0.000	0.004
		MITOSIS	2.10	0.000	0.003
		MITOTIC CELL CYCLE	2.09	0.000	0.003
		M PHASE	2.06	0.000	0.004
		CELL CYCLE PROCESS	2.04	0.000	0.004
		CELL CYCLE PHASE	1.92	0.000	0.015
		CELL CYCLE GO 0007049	1.88	0.000	0.022
		CELL CYCLE CHECKPOINT GO 0000075	1.83	0.008	0.033
	- Cellular component	CENTROSOME	2.18	0.000	0.000
		CHROMOSOMEPERICENTRIC REGION	2.13	0.000	0.002
		SPINDLE	2.11	0.000	0.002
		MICROTUBULE ORGANIZING CENTER	2.09	0.000	0.004
		KINETOCHORE	2.05	0.000	0.004
		CHROMOSOMAL PART	2.01	0.000	0.005
		RIBONUCLEOPROTEIN COMPLEX	1.99	0.000	0.006

	CHROMOSOME	1.95	0.000	0.008
	MICROTUBULE CYTOSKELETON	1.94	0.000	0.008
	NUCLEAR PART	1.87	0.000	0.016
	RUFFLE	1.85	0.005	0.017
	MICROTUBULE	1.85	0.007	0.017
	SPLICEOSOME	1.82	0.005	0.018
	INTRACELLULAR NON MEMBRANE BOUND ORGANELLE	1.78	0.000	0.024
	CONDENSED CHROMOSOME	1.78	0.009	0.023
	NON MEMBRANE BOUND ORGANELLE	1.77	0.000	0.023
	ER GOLGI INTERMEDIATE COMPARTMENT	1.75	0.014	0.026
	CYTOSKELETAL PART	1.74	0.000	0.025
	GOLGI APPARATUS PART	1.67	0.008	0.039
	PERINUCLEAR REGION OF CYTOPLASM	1.67	0.010	0.037
Oncogenic pathway targets	CORDENONSI YAP CONSERVED SIGNATURE	1.92	0.000	0.009
	VEGF A UP.V1 DN	1.84	0.000	0.014
	RB P107 DN.V1 UP	1.80	0.000	0.014
Chemical and genetic perturbations (Experimentally defined target gene signature in literature)	CAIRO HEPATOBLASTOMA CLASSES UP	3.48	0.000	0.000
	CROONQUIST NRAS SIGNALING DN	3.19	0.000	0.000
	CHIANG LIVER CANCER SUBCLASS PROLIFERATION UP	3.18	0.000	0.000
	WINNEPENNINCKX MELANOMA METASTASIS UP	3.13	0.000	0.000
	ROSTY CERVICAL CANCER PROLIFERATION CLUSTER	3.12	0.000	0.000
	PATIL LIVER CANCER	3.07	0.000	0.000
	CROONQUIST IL6 DEPRIVATION DN	3.06	0.000	0.000
	BENPORATH PROLIFERATION	3.03	0.000	0.000
	SOTIRIOU BREAST CANCER GRADE 1 VS 3 UP	3.03	0.000	0.000
	RODRIGUES THYROID CARCINOMA POORLY DIFFERENTIATED UP	2.99	0.000	0.000
	LEE LIVER CANCER SURVIVAL DN	2.96	0.000	0.000
	LI WILMS TUMOR VS FETAL KIDNEY 1 DN	2.93	0.000	0.000
	RHODES UNDIFFERENTIATED CANCER	2.91	0.000	0.000
	BOYAULT LIVER CANCER SUBCLASS G123 UP	2.91	0.000	0.000
	WHITEFORD PEDIATRIC CANCER MARKERS	2.91	0.000	0.000
	ZHOU CELL CYCLE GENES IN IR RESPONSE 24HR	2.90	0.000	0.000
	BIDUS METASTASIS UP	2.90	0.000	0.000
	CHANG CYCLING GENES	2.89	0.000	0.000
	KOBAYASHI EGFR SIGNALING 24HR DN	2.89	0.000	0.000
	HOSHIDA LIVER CANCER SUBCLASS S2	2.87	0.000	0.000
	BORCZUK MALIGNANT MESOTHELIOMA UP	2.87	0.000	0.000
	BOYAULT LIVER CANCER SUBCLASS G3 UP	2.84	0.000	0.000
	SHEDDEN LUNG CANCER POOR SURVIVAL A6	2.82	0.000	0.000
	MITSIADES RESPONSE TO APLIDIN DN	2.81	0.000	0.000
	GRAHAM NORMAL QUIESCENT VS NORMAL DIVIDING DN	2.80	0.000	0.000
	MISSIAGLIA REGULATED BY METHYLATION DN	2.80	0.000	0.000
	BURTON ADIPOGENESIS 3	2.77	0.000	0.000
	KANG DOXORUBICIN RESISTANCE UP	2.75	0.000	0.000
	KONG E2F3 TARGETS	2.75	0.000	0.000
	ZHAN MULTIPLE MYELOMA PR UP	2.74	0.000	0.000
	DUTERTRE ESTRADIOL RESPONSE 24HR UP	2.73	0.000	0.000
	PUJANA BRCA2 PCC NETWORK	2.71	0.000	0.000
	WHITFIELD CELL CYCLE LITERATURE	2.68	0.000	0.000
	KAUFFMANN MELANOMA RELAPSE UP	2.66	0.000	0.000
	PUJANA BRCA CENTERED NETWORK	2.64	0.000	0.000
	YU MYC TARGETS UP	2.63	0.000	0.000
	ZHOU CELL CYCLE GENES IN IR RESPONSE 6HR	2.61	0.000	0.000
	SARRIO EPITHELIAL MESENCHYMAL TRANSITION UP	2.60	0.000	0.000
	MORI IMMATURE B LYMPHOCYTE DN	2.60	0.000	0.000
	FEVR CTNNB1 TARGETS DN	2.59	0.000	0.000
	KAMMINGA EZH2 TARGETS	2.59	0.000	0.000
	BLUM RESPONSE TO SALIRASIB DN	2.58	0.000	0.000
	FOURNIER ACINAR DEVELOPMENT LATE 2	2.57	0.000	0.000
	HORIUCHI WTAP TARGETS DN	2.57	0.000	0.000
	JOHNSTONE PARVB TARGETS 3 DN	2.57	0.000	0.000
	FERREIRA EWINGS SARCOMA UNSTABLE VS STABLE UP	2.56	0.000	0.000
	WONG EMBRYONIC STEM CELL CORE	2.56	0.000	0.000
	MARTINEZ RESPONSE TO TRABECTEDIN DN	2.56	0.000	0.000
	MORI LARGE PRE BII LYMPHOCYTE UP	2.55	0.000	0.000
	BURTON ADIPOGENESIS PEAK AT 24HR	2.55	0.000	0.000
Transcription factor targets	KRCTCNNNNMANAGC UNKNOWN	2.17	0.000	0.005
	SGCGSSAAA V$E2F1DP2 01	2.13	0.000	0.003
	V$E2F1 Q6	2.09	0.000	0.003
	GGCNNMSMYNTTG UNKNOWN	2.01	0.000	0.007
	GKCGCNNNNNNNTGAYG UNKNOWN	1.92	0.000	0.013
	RRCCGTTA UNKNOWN	1.87	0.003	0.018
	V$E2F4DP1 01	1.86	0.000	0.019
	V$E2F Q4 01	1.82	0.000	0.026
	V$E2F1 Q4 01	1.78	0.000	0.032
	V$E2F Q4	1.77	0.000	0.030
	V$E2F1 Q6 01	1.77	0.000	0.028
	GCCATNTTG V$YY1 Q6	1.77	0.000	0.027
	V$E2F 02	1.76	0.000	0.026
	V$E2F Q6	1.76	0.000	0.024
	V$E2F1 Q3	1.76	0.000	0.023
	V$E2F Q3	1.75	0.000	0.023
	V$E2F1 Q4	1.75	0.000	0.022
	V$E2F4DP2 01	1.73	0.000	0.026
	V$E2F1DP1 01	1.73	0.000	0.025
	V$E2F Q3 01	1.72	0.000	0.025
	V$E2F1DP2 01	1.72	0.000	0.024
	V$DR4 Q2	1.69	0.003	0.031
	V$E2F1DP1RB 01	1.67	0.006	0.035
	V$NFAT Q6	1.64	0.003	0.044
	ATCMNTCCGY UNKNOWN	1.63	0.018	0.044
	V$NFMUE1 Q6	1.63	0.003	0.044
Liver cancer gene signatures in literature	CHIANG08_SUBCLASS_PROLIFERATION	3.19	0.000	0.000
	LEE04_SUBCLASS_A	3.02	0.000	0.000
	BOYAULT07_SUBCLASS_G1G2G3	2.96	0.000	0.000
	HOSHIDA09_SUBCLASS_S2	2.84	0.000	0.000
	BOYAULT07_SUBCLASS_G3	2.84	0.000	0.000
	VILLANUEVA11_KRT19_POSITIVE	2.82	0.000	0.000
	WOO10_CHOLANGIO_LIKE_UP	2.60	0.000	0.000
	BOYAULT07_SUBCLASS_G2G3	2.37	0.000	0.000
	KIM12_SURVIVAL_POOR	2.27	0.000	0.000
	BOYAULT07_SUBCLASS_G1G2	2.23	0.000	0.000
	ANDERSEN10_KRT19_POSITIVE_CLSA	2.21	0.000	0.000
	BOYAULT07_SUBCLASS_G1	2.19	0.000	0.000
	MINGUEZ11_VASCULAR_INVASION_POSITIVE	2.14	0.000	0.000

		YAMASHITA08_EPCAM_UP	2.12	0.000	0.000
		WOO08_RECURRENCE_POOR	2.04	0.000	0.000
		CAIRO08_HEPATOBLAST_LIKE_POOR_CLS2	1.87	0.000	0.002
		KAPOSINOVAK06_MET_UP	1.78	0.005	0.006
		HOSHIDA09_SUBCLASS_S1	1.76	0.000	0.007
		COULOUARN08_TGFB1_UP	1.64	0.006	0.017
Pseudoglandular	Biocarta (Curated pathway database)	BIOCARTA CHEMICAL PATHWAY	2.00	0.000	0.028
	Chemical and genetic perturbations (Experimentally defined target gene signature in literature)	CHIANG LIVER CANCER SUBCLASS CTNNB1 UP	2.60	0.000	0.000
	Liver cancer gene signatures in literature	CHIANG_CTNNB1_UP	2.55	0.000	0.000
		BOYAULT_G6_UP	1.82	0.000	0.022
SH-HCC	Reactome (Curated pathway database)	REACTOME COLLAGEN FORMATION	2.16	0.000	0.009
		REACTOME EXTRACELLULAR MATRIX ORGANIZATION	2.14	0.000	0.005
	Gene Ontology - Cellular component	COLLAGEN	2.04	0.000	0.018
		EXTRACELLULAR MATRIX PART	1.96	0.000	0.021
		EXTRACELLULAR MATRIX	1.96	0.000	0.015
		PROTEINACEOUS EXTRACELLULAR MATRIX	1.95	0.000	0.012
	Oncogenic pathway targets	CAHOY ASTROGLIAL	2.41	0.000	0.000
		ESC V6.5 UP EARLY.V1 DN	2.34	0.000	0.000
		MEL18 DN.V1 UP	2.20	0.000	0.001
		BMI1 DN MEL18 DN.V1 UP	2.14	0.000	0.000
		RPS14 DN.V1 UP	2.10	0.000	0.001
		BMI1 DN.V1 UP	2.00	0.000	0.001
		CORDENONSI YAP CONSERVED SIGNATURE	1.82	0.003	0.010
		BCAT BILD ET AL UP	1.79	0.006	0.011
		STK33 SKM UP	1.72	0.000	0.016
		RB DN.V1 DN	1.70	0.000	0.019
		RB P107 DN.V1 UP	1.70	0.000	0.017
		HOXA9 DN.V1 UP	1.68	0.000	0.017
		STK33 UP	1.62	0.000	0.028
		SIRNA EIF4GI DN	1.60	0.007	0.031
		STK33 NOMO UP	1.58	0.000	0.033
		KRAS.DF.V1 UP	1.58	0.005	0.032
		RAF UP.V1 UP	1.55	0.005	0.040
	Chemical and genetic perturbations (Experimentally defined target gene signature in literature)	WOO LIVER CANCER RECURRENCE UP	3.20	0.000	0.000
		HOSHIDA_LIVER_CANCER_SUBCLASS_S1	3.18	0.000	0.000
		TURASHVILI_BREAST_LOBULAR_CARCINOMA_VS_LOBULAR_NORMAL_DN	2.82	0.000	0.000
		ANASTASSIOU CANCER MESENCHYMAL TRANSITION SIGNATURE	2.82	0.000	0.000
		CHIANG LIVER CANCER SUBCLASS CTNNB1 DN	2.79	0.000	0.000
		TURASHVILI BREAST LOBULAR CARCINOMA VS DUCTAL NORMAL U	2.71	0.000	0.000
		HOSHIDA LIVER CANCER SURVIVAL UP	2.65	0.000	0.000
		SCHUETZ BREAST CANCER DUCTAL INVASIVE UP	2.62	0.000	0.000
		CHARAFE BREAST CANCER LUMINAL VS MESENCHYMAL DN	2.55	0.000	0.000
		BURTON ADIPOGENESIS 8	2.53	0.000	0.000
		KIM GLIS2 TARGETS UP	2.52	0.000	0.000
		IGLESIAS E2F TARGETS UP	2.50	0.000	0.000
		BOYAULT LIVER CANCER SUBCLASS G5 DN	2.47	0.000	0.000
		CROONQUIST STROMAL STIMULATION UP	2.47	0.000	0.000
		LINDGREN BLADDER CANCER CLUSTER 2B	2.44	0.000	0.000
		CERVERA SDHB TARGETS 1 UP	2.43	0.000	0.000
		RODWELL AGING KIDNEY NO BLOOD UP	2.40	0.000	0.000
		KHETCHOUMIAN TRIM24 TARGETS UP	2.40	0.000	0.000
		CLASPER LYMPHATIC VESSELS DURING METASTASIS DN	2.37	0.000	0.000
		VECCHI GASTRIC CANCER ADVANCED VS EARLY UP	2.36	0.000	0.000
		RODWELL AGING KIDNEY UP	2.33	0.000	0.000
		BOQUEST STEM CELL UP	2.31	0.000	0.000
		MCLACHLAN DENTAL CARIES UP	2.31	0.000	0.000
		PETRETTO CARDIAC HYPERTROPHY	2.31	0.000	0.000
		PICCALUGA ANGIOIMMUNOBLASTIC LYMPHOMA UP	2.28	0.000	0.000
		PASINI SUZ12 TARGETS DN	2.28	0.000	0.000
		CHICAS RB1 TARGETS CONFLUENT	2.26	0.000	0.000
		VERHAAK GLIOBLASTOMA NEURAL	2.25	0.000	0.000
		JEON SMAD6 TARGETS UP	2.24	0.000	0.000
		KEEN RESPONSE TO ROSIGLITAZONE DN	2.23	0.000	0.000
		MIKKELSEN MEF LCP WITH H3K4ME3	2.22	0.000	0.000
		BASSO CD40 SIGNALING UP	2.22	0.000	0.000
		TURASHVILI BREAST DUCTAL CARCINOMA VS DUCTAL NORMAL UP	2.22	0.000	0.000
		MCLACHLAN DENTAL CARIES DN	2.21	0.000	0.001
		NADLER OBESITY UP	2.21	0.000	0.001
		REN ALVEOLAR RHABDOMYOSARCOMA DN	2.20	0.000	0.001
		GERHOLD ADIPOGENESIS DN	2.20	0.000	0.001
		LIU SMARCA4 TARGETS	2.20	0.000	0.001
		LINDSTEDT DENDRITIC CELL MATURATION D	2.18	0.000	0.001
		GU PDEF TARGETS UP	2.17	0.000	0.001
		NAKAMURA ADIPOGENESIS LATE DN	2.15	0.003	0.002
		HELLEBREKERS SILENCED DURING TUMOR ANGIOGENESIS	2.15	0.000	0.002
		PETROVA ENDOTHELIUM LYMPHATIC VS BLOOD DN	2.15	0.000	0.002
		NAKAMURA ADIPOGENESIS EARLY DN	2.14	0.000	0.002
		LIM MAMMARY LUMINAL MATURE DN	2.13	0.000	0.002
		STEGER ADIPOGENESIS DN	2.12	0.000	0.002
		WIELAND UP BY HBV INFECTION	2.11	0.000	0.002
		ONDER_CDH1_SIGNALING_VIA_CTNNB1	2.10	0.000	0.003
		PARK APL PATHOGENESIS DN	2.09	0.000	0.003
		FARMER BREAST CANCER CLUSTER 5	2.09	0.000	0.003
	microRNA targets	CCCNNNNNNAAGWT UNKNOWN	1.95	0.000	0.033
	Liver cancer gene signatures in literature	WOO08_RECURRENCE_UP	3.20	0.000	0.000
		HOSHIDA09_SUBCLASS_S1	3.13	0.000	0.000
		WOO10_CHOLANGIO_LIKE_UP	2.65	0.000	0.000
		KIM12_SURVIVAL_POOR	2.48	0.000	0.000
		OISHI12_CHOLANGIO_STEM_CELL_DOWN	1.90	0.000	0.003
		CHIANG08_SUBCLASS_PROLIFERATION	1.77	0.000	0.012
Clear cell	Kyoto Encyclopedia of Genes and Genomes (KEGG) (Curated pathway database)	KEGG_MISMATCH_REPAIR	1.96	0.000	0.046
	Chemical and genetic perturbations	HOSHIDA LIVER CANCER SUBCLASS S2	2.82	0.000	0.000

(Experimentally defined target gene signature in literature)	BOYAULT LIVER CANCER SUBCLASS G123 UP	2.39	0.000	0.001
	SHEN SMARCA2 TARGETS UP	2.24	0.000	0.007
	BIDUS METASTASIS UP	2.24	0.000	0.006
	PUJANA XPRSS INT NETWORK	2.13	0.000	0.020
	LY AGING MIDDLE DN	2.12	0.000	0.018
	YAMASHITA LIVER CANCER WITH EPCAM UP	2.12	0.000	0.017
	KAUFFMANN MELANOMA RELAPSE UP	2.11	0.000	0.015
	ZHOU CELL CYCLE GENES IN IR RESPONSE 24HR	2.10	0.000	0.016
	SONG TARGETS OF IE86 CMV PROTEIN	2.10	0.000	0.015
	KAMMINGA EZH2 TARGETS	2.07	0.000	0.018
	CHANG CYCLING GENES	2.02	0.000	0.032
	PATIL LIVER CANCER	2.00	0.000	0.036
	ISHIDA E2F TARGETS	1.97	0.000	0.040
	BOYAULT LIVER CANCER SUBCLASS G1 UP	1.97	0.000	0.038
	WINNEPENNINCKX MELANOMA METASTASIS UP	1.97	0.000	0.036
	CHOI ATL STAGE PREDICTOR	1.96	0.000	0.036
	GABRIELY MIR21 TARGETS	1.96	0.000	0.036
	ZHOU CELL CYCLE GENES IN IR RESPONSE 6HR	1.95	0.000	0.036
	PUJANA BRCA2 PCC NETWORK	1.95	0.000	0.034
	MITSIADES RESPONSE TO APLIDIN DN	1.93	0.000	0.037
	WHITFIELD CELL CYCLE LITERATURE	1.90	0.006	0.046
	PUJANA BRCA CENTERED NETWORK	1.89	0.000	0.049
	CROONQUIST IL6 DEPRIVATION DN	1.89	0.000	0.047
	GRAHAM NORMAL QUIESCENT VS NORMAL DIVIDING DN	1.88	0.000	0.046
	FUJII YBX1 TARGETS DN	1.88	0.000	0.045
	MISSIAGLIA REGULATED BY METHYLATION DN	1.88	0.000	0.045
	MORI IMMATURE B LYMPHOCYTE DN	1.88	0.000	0.044
	CAIRO HEPATOBLASTOMA CLASSES UP	1.87	0.000	0.046
	MILI PSEUDOPODIA HAPTOTAXIS UP	1.86	0.000	0.045
microRNA targets	ATAACCT,MIR-154	2.21	0.000	0.004
Liver cancer gene signatures in literature	HOSHIDA09_SUBCLASS_S2	2.79	0.000	0.000
	BOYAULT07_SUBCLASS_G1G2G3	2.39	0.000	0.000
	YAMASHITA08_EPCAM_UP	2.20	0.000	0.001
	BOYAULT07_SUBCLASS_G1G2	1.97	0.004	0.005
	BOYAULT07_SUBCLASS_G1	1.96	0.000	0.004
	LEE04 SUBCLASS A	1.76	0.000	0.026
	VILLANUEVA11_KRT19_POSITIVE	1.70	0.000	0.030
	OISHI12_CHOLANGIO_STEM_CELL_UP	1.70	0.000	0.026
	ANDERSEN10_KRT19_POSITIVE_CLSA	1.63	0.036	0.043

Gene sets with FDR <0.05 or top 50 gene sets are shown.

NES: normalized enrichment score, FDR: false discovery rate, SH-HCC: Steatohepatitic-hepatocellular carcinoma. Click each gene set name for detailed information.

p-value/FDR of "0.000" indicates <0.001.

Table S8
Key driver genes in gene regulatory network modules (Planar Filtered Network Analysis, Key Driver Analysis).

Module	Size (no. genes)	Key driver genes
PFN-module_no1	1436	LCN1,SEC11A,RPL13,ST13,SF3A1,SURF1,IFNA5,RPS3A,CTBP1,DMBT1,CLIC4,SMS,PPP2CB,CA6,ILF3,DHCR24,HIP2,STX4,MAP2K5,THBS2,MOXD1,VIL2,CEACAM8,RQCD1,COL6A3,SDC2,MMP19,COPS8,RGS10,GALNT10,FRG1,PDAP1,TFAP2C,RAD23B,TNFRSF8,PRAF2,NDUFB1,RAB2A,ATP6V1C1,HNRPA2B1,IER3,ARL2BP,BPGM,DAG1,PHF21A,SMARCD2,ABCF3,C21orf91,PMM2,PPP3CA,HLA-F,PPP2R5C,SAT1,ATP1B1,CSTF2,SNAPC5,TRIM14,AES,DDX3Y,SDC4,BCAT2,PPARD,PSMD6,ADRM1,RPS6KB2,NIPSNAP1,AKR1B1,MAPKAPK3,CREB1,EIF2S1,PLK2,TRIM26,TSN,DNAJA1,PC,PIK4CB,APP,FN1,RPS6KB1,ANXA7,TRAK2,HNRPC,SLC22A18,EPAS1,CITED2,EIF3S10,PPIH,CD151,PARK7,GUK1,LSM3,PRKACA,SERPINA1,REL,TMEM97,BMI1,GIPC1,SSR4,DDX17,NGFRAP1,CD46,HLA-DQB1,TCTA,UBXD6,USP34,XRCC5,CLTB,FOXA1,PAFAH1B2,PCBP2,SRRM2,ZEB1,CNIH,LAMB2,PTTG1,TJP1,MRPL12,PENK,RARRES1,RPL36AL,C4BPB,LUM,APOB,GNB1,SAFB2,POLR2J,CRK,PCK2,CYB5A,FCGR1A,MGST2,PGK1,NPC2,ECHS1,HADHB,GTF2A1,PPIB,ZNF423,ATP5O,COPS2,ELK1,PSEN1,UCK2,PSCD1,COL3A1,HOXB5,IFRD2,PIB5PA,TAF11,ZNF3,LY96,RAC2,TUBG1,ACAT1,FABP4,MRPS31,RYK,UCN,ZNF22,DDX1,NAPA,PAPSS2,SLC9A3R1,SREBF2,WWP1,TOP2A,EVI2B,PTPRC,COL1A2,COL1A1,ECM2,SNRPB2,ANP32B,REEP5,CXCL12,KNG1,TNIP1
PFN-module_no10	166	FOXF2,GHRH,MMP8,PRR5,NOX4,MC4R,AMELX,ANGPTL4,SPINK2,SRPK3,SLC2A3,TGFB2,TCF15,CHRNA2,SPRR3,CSDC2,LCT,ART3,KIR3DL2,MYOG,RPL11,SEMA4D,PLAUR
PFN-module_no11	214	ARHGDIB,ATP2A3,CD3E,LTB,CD52,EVI2B,IGJ,BTG2,PTPRC,AIF1,IL16,SH2D2A,LCP1,CD27,LAPTM5,GMFG,SPOCK2,C6orf32,CD3D,ARHGAP25,LCP2,CD48,PRKCB1,STK17B,GPSM3,FGL2,CXCR4,PSCDBP,LSP1,MS4A6A,TMSB10
PFN-module_no13	266	ACOX2,ADH1B,SEC14L2,MTHFD1,HSD11B1,DCXR,AQP9,CPB2,ALDH6A1,HPD,APOC4,CDO1,ADH6,DBT,PKM2,SERPINC1,RGN,NR1I3,SCP2,PCK2,PAH,GCDH,FLJ20699,SLC10A1,PON3,TAT,C6,FAAH,PPARA,C4BPA,CAT
PFN-module_no14	139	TCEAL4,YY1,DYNLL1,HINT1,ACTN1,SDCBP,ZC3H15,PSME1,SRP19,EIF4B,SNRPC,RAB4A,CAV1,EIF3S2,REST,UVRAG,MBP,SF3B2,AP2S1,SUB1,REEP5,ATP5G3,RAP1B,RPS6KA3
PFN-module_no15	87	TOP2A,NCAPH,DLG7,PRC1,TTK,LMNB1,SMC4,UBE2C,NEK2,NUSAP1,HIST1H2AE,KIAA0101,PTTG1,CCNA2,GINS1,MCM6,ZWINT,BRCA1,CDK2,RRM2,HIST1H2BD,ACAA1,C22orf5
PFN-module_no18	577	LBX1,IFNA1,SOX11,LALBA,BMPR1B,CACNG1,LRP2,MYT1,GPR3,EPYC,FABP7,STAR,CYP2A7,DRD5,WNT1,NEUROD2,PSG6,NEUROG1,CSRP3,OR2H1,PRM1,COX7A1,PPM1E,SCD5,CSF3,GABRB1,CHRNB2,MCF2,NDP,ACRV1,CLUL1,CRISP1,KIF2C,BMP4,MPL,CD1A,KRT81,ATP12A,MOG,AMPH,BRAF,PTK2,CEP350,NUAK1,CUL4B,RIT2,BLK,CYP24A1,KRT76,NPPA,PYY,DNTT,KRT75,MAGEA9,CCL1,CDH8,CST1,HTR6,SUMO1,ZP2,MYOC,SCG2,VIP,ADORA2B,DLG3,HOXA1,KLK6,KRT13,LEFTY2,ATP1A2,DLX2,HTR2A,NHLH2,SLC2A5,CCR5,GPR143,GYPE,OPRM1,DUSP2,TBR1,CRYBB1
PFN-module_no2	859	ST13,SURF1,SEC11A,CTBP1,SMS,PPP2CB,HIP2,SF3A1,RQCD1,RPL13,FRG1,PDAP1,RAD23B,NDUFB1,CSTF2,AES,LCN1,AKR1B1,EIF2S1,IFNA5,HNRPC,PPIH,PARK7,GUK1,HNRPA2B1,LSM3,PRKACA,TMEM97,RPL36AL,SSR4,NGFRAP1,CD46,HLA-DQB1,CLTB,CNIH,PTTG1,C4BPB,APOB,GNB1,POLR2J,REL,VIL2,PCK2,CYB5A,FCGR1A,MGST2,NPC2,ECHS1,HADHB,SAFB2,SERPINA1,PPIB,COPS2,PSEN1,UCK2,IFRD2,BRE,PLS3,BRD8,LY96,PSMA7,RAC2,TUBG1,ACAT1,CRK,FABP4,MRPS31,RYK,STX4,UCN,ZNF22,DDX1,NUP155,NAPA,PAPSS2,SLC9A3R1,SREBF2,WWP1,BRP44,COL18A1,TOP2A,CHPF,EVI2B,PTPRC,SMAD5,SRI,GNG10,HAX1,PDK2,PTPRM,SIPA1,TBCB,LAMB2,LSM1,PSMA3,TIMM17A,UQCRFS1,USP1,IDI1,ITPR2,CENPB,FBL,ITGA6,MGST3,CCT5,HOXC4,UBD,ANP32B,DHX9,REEP5,COPE,ELAVL1,SHC1,FAM3C,RAB5A,KNG1,AGXT
PFN-module_no20	961	DIRAS1,SH3GL3,MYCBP,AOC2,CNR1,TNNI3,BAPX1,APOBEC1,GABRQ,PCSK1,CKS2,TNRC4,ADARB1,BRCA2,CD6,UNC5C,TLE2,EPHA5,KLK4,ERAF,F2RL1,GIPR,GATA1,IL3,VTN,SSPN,XK,MYLK2,PAPPA,PNLIP,ERCC4,HIST1H4B,CUL4A,CBX3,RHOA,CALD1,SPRY2,TNFRSF14,QPRT,CPA1,RXRG,NOL14,SFTPB,SLC25A6,WT1,IFNGR2,RNF2,ANGPT2,REGL,CNTN1,FOXI1,CLPS,SYT17,TCFL5,CSF3R,SFRS1,SMARCC1,GP9,PFDN5,REV3L,ZFP161,ADNP,LMNB2,MALL,INPP5E,CDKN2B,CHST2,SYNCRIP,TAZ,ZIC1,ATP2A1,GPRC5B,PTHR1,MDC1,MYBPC1,POU1F1,RB1CC1,TP73,Septin 4,CTNNB1,NELL1,MYOM2,SLC15A2,DTYMK,HNRPU,RNPEP,SULF1,EXDL2,NAV3,NMT1,NPR1,TFAP2A,TPR,CDS2,CEBPZ,HMMR,ST18,COMMD3,MCM7,XRCC6,IVNS1ABP,NR2F2,PAX2,PPARG,STAT3,HTR7,NOV,RORA,FGFR4,GNE,CASP7,CD86,DNTTIP2,FUT2,LAMC2,RAC3,RBM9,UBE3C,ABCE1,JARID1C,RBM4,ACOT7,ATP6V1E2,C5orf15,EXOSC10,IFIT2,NME1,RPN2,SNTB1,DNASE1L2,MS4A4A,PHKG1,SNX17,CDH12,EEF1G
PFN-module_no22	71	HIST1H2BF,PCNA,YBX1,HIST1H1E,TRIM28,HIST1H4E,ILF2,DLG7,HIST1H3H,TROAP,RELA,PRC1,BIRC5
PFN-module_no26	121	TSHB,ELAVL3,SERPINB2,DMN,FLJ33790,SNCB,NINJ2,KIAA0274,TBL3,LEF1,HOXA5,KL,SLC17A7,CNGB1,SCUBE2,XRCC2,GUCY1A3,NKX2-5
PFN-module_no3	120	RNF185,AMELY,GRM1,ALK,PRKCG,MYH6,GDPD5,SFTPA2,KCTD2,SEMA3B,EPHA4,CCNG1,FOLR1,ARHGDIG,GML,LLGL2,SLC6A12,TP53I11
PFN-module_no30	85	MAG,MMP13,IFNA21,NEFH,MMP15,PSG3,STC2,OLFM1,SOX3,EDN2,DEFA4,ALPPL2
PFN-module_no32	133	POU4F2,PGC,CTRC,AQP6,C1orf61,GUCY2D,CKM,SCNN1G,TSPYL5,FUT3,MME,IVL,PSG1,SLC5A6,RBMX,DNASE1L3,DPEP1,RHD,SCEL,LAMA3,AMACR,EDN2,IRX5

PFN-module_no33	223	GPR109B,PRL,NAP1L3,CNTN2,SART3,RBP3,OPRM1,TH,TEK,CRHR1,LIPF,SFTPC,ARF4,EIF4A1,TP73L,ATP5D,SEPHS1,ALAS2,MFAP5,COX6A1,EPM2AIP1,IL11RA,NPY5R,MAEA,SLC14A1,HAGH,POLR3D,SPINK5,TMEM187,H2AFZ,GSTM1,IL9R,USP14,GPSN2,COX5B,ITSN1,SLC6A3,ISG15,MYBPC3,CHI3L1,GRIA1,ABCB4,SYNGR1,MLN
PFN-module_no34	104	REN,CPA2,EFNA3,HSPB2,STAT4,IRX4,CRYGD,BLMH,CD93,CTAG1B,MN1,ARSA,ICAM4,TRAF5,ING3,ZNF165,GNG5,TAF6,ZNF593
PFN-module_no37	117	CTRC,PGC,AQP6,POU4F2,MME,C1orf61,FUT3,SLC5A6,PPY,CKM,GUCY2D,SCNN1G,TSPYL5,LAMA3,AMACR,AMPD1,IRX5,RHD,RBMX,PRKACG
PFN-module_no4	126	RAPGEF1,CNP,FDPS,CDKN2D,PPP2R5B,ELK1,FKBP1A,EHMT2,PIK3R2,TCF3,FOXO3,ST6GAL1,DAP,TAF7,ARHGEF12,PGK1,EXT1,TNIP1
PFN-module_no51	1760	LCN1,SEC11A,RPL13,ST13,SF3A1,IFNA5,SURF1,RPS3A,DMBT1,CTBP1,CLIC4,CA6,ILF3,SMS,DHCR24,PPP2CB,THBS2,STX4,VIL2,MAP2K5,UQCRFS1,MOXD1,MMP19,CEACAM8,HIP2,COL6A3,GALNT10,SDC2,RAB5A,COPS8,RGS10,PSMD7,KLHDC3,RQCD1,TFAP2C,RAB2A,ATP6V1C1,TNFRSF8,C21orf91,HNRPA2B1,PRAF2,NDUFB7,DDX3Y,IER3,ARL2BP,BPGM,DAG1,FRG1,PHF21A,SMARCD2,ABCF3,FARSA,PMM2,PPP1R7,PPP3CA,PDAP1,HLA-F,PPP2R5C,RAD23B,SAT1,ATP1B1,SNAPC5,TRIM14,SDC4,BCAT2,PPARD,PSMD6,ADRM1,RPS6KB2,NIPSNAP1,NDUFB1,MAPKAPK3,CREB1,PLK2,TRIM26,TSN,ARHGDIB,DNAJA1,PC,PIK4CB,APP,FN1,RPS6KB1,ANXA7,CSTF2,TRAK2,PDXK,AES,SLC22A18,UQCRC1,EPAS1,SNAI2,CITED2,EIF3S10,AKR1B1,CD151,HIST1H4C,EIF2S1,BMI1,GIPC1,DDX17,TCTA,UBXD6,USP34,FOXA1,PAFAH1B2,PCBP2,SRRM2,ZEB1,LAMB2,HNRPC,REL,TJP1,PTTG1,MRPL12,PENK,PPIH,SAFB2,LTB,PARK7,GUK1,LSM3,PRKACA,SERPINA1,ATP2A3,TMEM97,RPL36AL,SSR4,NGFRAP1,CD3E,CD46,HLA-DQB1,XRCC5,CLTB,CNIH,ELK1,CNP,LUM,RARRES1,CRK,CD52,GTF2A1,PGK1,APOB,EVI2B,TOP2A,C4BPB,PTPRC,GNB1,HOXB5,PIB5PA,RAPGEF1,POLR2J,TAF11,ZNF3,CYB5A,ELAVL1,PCK2,FCGR1A,MGST2,RPP38,HADHB,NPC2,AP3S1,ECHS1,IGJ,PPIB,ZNF423,ATP5O,COPS2,PIK3CA,PSEN1,RBM39,UCK2,PSCD1,RAC2,APOC2,COL3A1,IFRD2,BRE,PLS3,BRD8,LY96,PSMA7,TUBG1,ACAT1,FABP4,MRPS31,NEK2,RYK,UCN,ZNF22,DDX1,FDPS,COL1A2,COL1A1,ECM2,SNRPB2,ANP32B,MFAP4,REEP5,CXCL12,FKBP1A,TNIP1,KNG1,TAF7,TMEPAI
PFN-module_no52	160	POU4F2,CTRC,PGC,AQP6,CKM,C1orf61,RHD,IVL,FUT3,GUCY2D,SCNN1G,TSPYL5,MME,DNASE1L3,PSG1,SLC5A6,RBMX,DPEP1,PPY,AMPD1,SCEL,LAMA3,SERHL2,AMACR,PRKACG
PFN-module_no53	273	ACOX2,ADH1B,SEC14L2,MTHFD1,HSD11B1,DCXR,CPB2,AQP9,ALDH6A1,CDO1,HPD,APOC4,ADH6,DBT,PKM2,SERPINC1,RGN,NR1I3,SCP2,PCK2,PAH,GCDH,PON3,FLJ20699,SLC10A1,TAT,C6,FAAH,PPARA,C4BPA,HGD,SULT2A1,ACY1,CAT,PCSK6
PFN-module_no54	548	NOX4,AMELX,GHRH,CARTPT,ATBF1,NEUROD1,MMP8,FOXF2,GRK1,TPH1,PKMYT1,LGALS7,PDE4A,CCNU,MC4R,TGFB2,KIF14,OCA2,ART3,PRR5,PLAGL1,CBFA2T3,RPA3,SPINK2,FCGBP,GNAO1,GNRH1,AIM1,NEFM,DEGS1,FEV,COL4A3,SFRS11,NR2C1,KIAA1539,ANGPTL4,NUP214,ILVBL,EEF1A2,IL13RA1,KBTBD10,TMEM4,TNNI1,TDGF1,ADCY2,CHRNA2,PFTK1,CYP11A1,DUSP9,MYL4,SALL1,NEBL,SLC14A2,ECE2,PSMB4,UBXD2,RPL11,SPRR3,CCR1,GPR116,SERPINB7,TBX5,BCL2L1,MYOZ2,SYPL1,LAMA4,SLC34A1,TACR3,SRPK3,PLAUR,GABRA1
PFN-module_no56	696	LBX1,SOX11,IFNA1,LALBA,BMPR1B,CACNG1,LRP2,MYT1,GPR3,EPYC,ELAVL3,STAR,FABP7,CYP2A7,DRD5,WNT1,NEUROD2,PSG6,NEUROG1,CSRP3,OR2H1,PRM1,SCD5,COX7A1,PPM1E,TSHB,CSF3,SERPINB2,GABRB1,CHRNB2,MCF2,CLUL1,KIF2C,DMN,NDP,ACRV1,CRISP1,BMP4,MPL,CD1A,KRT81,MOG,ATP12A,AMPH,BRAF,FLJ33790,PTK2,SNCB,CEP350,NUAK1,CUL4B,NPPA,PYY,RIT2,BLK,CYP24A1,KRT76,NINJ2,PRSS1,DNTT,KRT75,MAGEA9,KIAA0274,SUMO1,CCL1,CDH8,CST1,HTR6,ZP2,MYOC,SCG2,VIP,ADORA2B,DLG3,HOXA1,KLK6,KRT13,LEFTY2,ATP1A2,DLX2,HTR2A,NHLH2,SLC2A5,CCR5,GPR143,GYPE,OPRM1,TBL3,PLP1,TRPM8,DUSP2,GUCY1A3,TBR1,CRYBB1
PFN-module_no57	129	RNF185,AMELY,GRM1,ALK,PRKCG,MYH6,GDPD5,SFTPA2,SEMA3B,KCTD2,EPHA4,CCNG1,MAPK8,FOLR1,ARHGDIG,GML,LLGL2,SLC6A12,TP53I11,NR5A1
PFN-module_no6	124	IGFBP7,LUM,MFAP4,CXCL12,GEM,COL6A2,TAGLN,AEBP1,COL3A1,PTGIS,EPHA3,C7,HGF,BGN,BMP5,COL1A2,FILIP1L,GFPT2,TMEPAI
PFN-module_no9	498	NOX4,AMELX,ATBF1,CARTPT,NEUROD1,GHRH,TPH1,FOXF2,GRK1,PKMYT1,MMP8,PDE4A,LGALS7,CCNU,OCA2,CBFA2T3,GNAO1,GNRH1,AIM1,FEV,TGFB2,ART3,SFRS11,KIF14,RPA3,FCGBP,KIAA1539,ILVBL,NEFM,DEGS1,TMEM4,COL4A3,NR2C1,PLAGL1,PRR5,TDGF1,ADCY2,NUP214,PFTK1,EEF1A2,TNNI1,DUSP9,MYL4,SALL1,SLC14A2,ECE2,PSMB4,UBXD2,ANGPTL4,CYP11A1,IL13RA1,CCR1,GPR116,SERPINB7,TBX5,BCL2L1,SYPL1,LAMA4,NEBL,SLC34A1,KRT32,KBTBD10,SRPK3,USP9X,ZNF711,ARID1A,INSR,ARNT2,PLAUR,GABRA1

Table S9

Gene regulatory network modules associated with histopathological features in the validation set. (Gene Set Enrichment Analysis)

Histopathological feature	Gene subnetwork	NES	p-value	FDR
Microtrabecular	PFN-MODULE_NO13	1.85	0.002	0.034
	PFN-MODULE_NO53	1.85	0.002	0.018
	PFN-MODULE_NO33	1.71	0.010	0.070
Macrotrabecular/Compact	PFN-MODULE_NO22	1.73	0.004	0.170
	PFN-MODULE_NO15	1.66	0.035	0.165
SH-HCC	PFN-MODULE_NO6	1.79	<0.001	0.098
Pseudoglandular	PFN-MODULE_NO56	1.86	0.003	0.055
	PFN-MODULE_NO18	1.78	0.016	0.080
Fatty change	PFN-MODULE_NO20	1.82	0.024	0.085

SH-HCC: steatohepatitic hepatocellular carcinoma, NES: normalized enrichment score, FDR: false discovery rate.

Table S10

Prognostic association of HCC molecular subclasses (univariable Cox regression modeling)

	Training set			Validation set		
	HR	95%CI	p-value	HR	95%CI	p-value
S1 subclass						
Death	0.76	0.39 - 1.47	0.41	0.89	0.49 - 1.64	0.72
Recurrence	1.03	0.64 - 1.65	0.91	0.93	0.54 - 1.61	0.81
S2 subclass						
Death	0.56	0.25 - 1.27	0.17	0.95	0.53 - 1.73	0.88
Recurrence	0.94	0.55 - 1.61	0.83	0.97	0.55 - 1.69	0.91
S3 subclass						
Death	1.78	0.99 - 3.20	0.05	1.14	0.67 - 1.96	0.63
Recurrence	1.02	0.66 - 1.57	0.94	1.09	0.66 - 1.81	0.73

HCC: hepatocellular carcinoma, HR: hazard ratio, CI: confidence interval.

Table S11

Performance of clinicopathological indices predictive of HCC molecular subclasses optimized for sensitivity/specificity.

	Training set					Validation set				
Molecular subclass	AUROC	Sensitivity	Specificity	PPV	NPV	AUROC	Sensitivity	Specificity	PPV	NPV
S1	0.69	70%	58%	43%	81%	0.71	83%	49%	42%	87%
S2	0.86	82%	81%	63%	92%	0.77	86%	69%	43%	95%
S3	0.83	69%	81%	71%	79%	0.73	65%	75%	71%	69%

Molecular subclass predictions using the clinico-histopathological indices were based on the following formulae and cut-off values optimized for sensitivity/specificity in the training set:

S1 = 1.45 x SH-HCC (0: no, 1: yes) + 1.18 x Immune cell infiltrate (0: 0-1, 1: 2); > 0.59 was predicted as S1 subclass.

S2 = 2.48 x macrotrabecular/compact (0: no, 1: yes) + 2.38 x serum alpha-fetoprotein (0: ≤ 400 ng/mL; 1: > 400 ng/mL) - 1.50 x pseudoglandular (0: no, 1: yes); > 1.68 was predicted as S2 subclass.

S3 = 1.37 x microtrabecular (0: no, 1: yes) - 2.97 x SH-HCC (0: no, 1: yes) - 1.59 x clear cell (0: no, 1: yes) - 1.13 x tumor grade (0: Edmondson-Steiner III or IV, 1: Edmondson-Steiner I or II); > 1.94 was predicted as S3 subclass.

HCC: hepatocellular carcinoma, AUROC: area under receiver operating characteristic curve, PPV: positive predictive value, NPV: negative predictive value, SH-HCC: steathepatitic HCC, HBV: hepatitis B virus.

Table S12

Required sample size in clinical trial and NNT in HCC molecular subclass-targeting therapy.

Molecular subclass		S1 (prevalence: 30%)			S2 (prevalence: 25%)			S3 (prevalence: 45%)		
ORR in targeted subclass		0.5	0.6	0.7	0.5	0.6	0.7	0.5	0.6	0.7
Sample size	All comers	184	121	88	293	184	129	76	53	39
	Subclass index-positive patients	43	30	22	62	43	32	31	22	16
NNT	All comers	6.7	5.6	4.8	8.0	6.7	5.7	4.4	3.7	3.2
	Subclass index-positive patients	3.3	2.8	2.4	4.0	3.3	2.9	2.8	2.3	2.0

Required sample sizes in one arm are shown (alpha=0.05, power=0.9). ORR: objective response rate, NNT: number needed to treat.

Table S13

Prognostic association of clear cell variant in literature.

	Yang, et al.		Emile, et al.*		Liu, et al.		Lai, et al.		Audisio, et al.**	
Characteristic	Clear cell (n=97)	Non-clear cell (n=118)	Clear cell (n=10)	Non-clear cell (n=108)	Clear cell (n=43)	Non-clear cell (n=597)	Clear cell (n=29)	Non-clear cell (n=50)	Clear cell (n=7)	Non-clear cell (n=130)
Male	69 (71%)	103 (87%)	8 (80%)	86 (80%)	32 (74%)	486 (81%)				
Hepatitis B infection	16 (16%)	12 (10%)	2 (20%)	25 (22%)	37 (86%)	443 (74%)				
Hepatitis C infection	47 (48%)	52 (44%)	4 (40%)	44 (41%)	6 (14%)	7 (1.2%)				
Hepatitis B & C infection	4 (4%)	1 (1%)	0 (0%)	2 (2%)						
Alcohol	NA	NA	2 (20%)	15 (14%)						
Cirrhosis	74 (76%)	70 (59%)	10 (100%)	108 (100%)	33 (77%)	514 (86%)				
Tumor stage	Mostly early-intermediate stage		Mostly early stage		Mostly intermediate-advanced stage		All advanced stage			
Treatment	Resection		Liver transplant		Resection		Best supportive care			
Prognosis	Trend of poorer prognosis and earlier recurrence in clear cell		No difference		Better prognosis in clear cell		Better prognosis in clear cell		Poorer prognosis in clear cell	
Reference	Pathol Int 46;503 (1996)		Histopathology 38;225 (2001)		Hepatol Res 38;291 (2008)		Cancer 44;1677 (1979)		Tumori 73;389 (1987)	

*Only tumors with > 90% clear cells were classified as clear cell variant positive.

**Only tumors with ≥ 50% clear cells were classified as clear cell variant positive.

Supplementary figure legend

Figure S1.
Study design.

Histopathological evaluation and determination of HCC molecular subclass were performed in the training set (left), and clinicopathological feature-based indices predictive of the molecular subclasses were developed. The indices were subsequently tested in an independent validation set (right).

Figure S2.
Concordance of HCC molecular subclass prediction between the full 619-gene signature and reduced 30-gene signature.

The subclass prediction was repeated with the 30-gene signature in the 9 datasets using nearest template prediction algorithm as previously described [Cancer Res 69;7385-92,2009] and concordance with the original prediction made by the 619-gene signature was evaluated. Median concordances for the prediction of overall subclasses, S1, S2, and S3 subclasses were 85% (IQR: 80%-90%), 78% (IQR: 74%-88%), 86% (IQR: 82%-93%), and 85% (IQR: 84%-89%), respectively.

Figure S3.
Induction of YAP target genes in S2 subclass (Gene Set Enrichment Analysis).

The signature of YAP activation (CORDENONSI_YAP_CONSERVED_SIGNATUE) was significantly enriched with S2 subclass. NES: normalized enrichment score, FDR: false discovery rate.

Figure S4.
The macrotrabecular/compact pattern-associated gene regulatory network modules (A: no.22, B: no.15) identified by Planar Filtered Network Analysis (PFNA).

Red nodes indicate key driver genes in the module determined by Key Driver Analysis (KDA). Green circles indicate top or bottom 50 genes deregulated in the HCC tumors with the macrotrabecular/compact pattern in the validation set.

Figure S5.
The pseudoglandular pattern-associated gene regulatory network module (no.56) identified by Planar Filtered Network Analysis (PFNA).

Red nodes indicate key driver genes in the module determined by Key Driver Analysis (KDA). Green circles indicate top or bottom 50 genes deregulated in the HCC tumors with the pseudoglandular pattern in the validation set.

Figure S6.

The steatohepatitic hepatocellular carcinoma (SH-HCC) variant-associated gene regulatory network module (no.6) identified by Planar Filtered Network Analysis (PFNA).

Red nodes indicate key driver genes in the module determined by Key Driver Analysis (KDA). Yellow circles indicate top or bottom 50 genes deregulated in the HCC tumors with SH-HCC variant in the validation set.

Figure S7.
Serum alpha-fetoprotein (AFP) levels according to HCC molecular subclasses in the training set.

AFP levels were significantly higher in the S2 subclass compared to the rest ($p<0.001$, Wilcoxon rank-sum test).

Figure S8.
Correlation of clinicopathological features with HCC molecular subclasses (univariable logistic regression analysis). (A: training set, B: validation set)

SH-HCC: steatohepatitic hepatocellular carcinoma, DCP: des-gamma-carboxy prothrombin, AFP: alpha-fetoprotein, L3: lens culinaris agglutinin-reactive fraction 3, HCV: hepatitis C virus, HBV: hepatitis B virus.

Figure S9.
Receiver operating characteristic (ROC) curves for the S1, S2, and S3 subclass-predictive indices based on clinicopathological features.

Receiver operating characteristic (ROC) curves for the S1, S2, and S3 subclass-predictive indices based on tumor-related clinicopathological features in the training set (upper panels) and the validation set (lower panels). Red and black arrows indicate cut-off values optimized for PPV and sensitivity/specificity, respectively which were determined in the training set and applied to the validation set. AUROC: area under ROC curve.

Figure S10.
Required sample size in clinical trial and NNT in HCC molecular subclass-targeting therapy.

Sample size required to detect therapeutic response in clinical trial of HCC drugs (upper panel) and NNT (lower panel) based on the HCC molecular subclass-predictive indices were determined. Positive predictive value (PPV) for each subclass was derived from the validation set (Table 4). Sample sizes were calculated at alpha error of 0.05, power of 0.90, and treatment/control ratio of 1. Based on previously published biomarker-enriched clinical trials in lung cancer [39], ORRs of 50% , 60%, and 70% were tested. NNT: number needed to treat, ORR: objective response rate.

Figure S1

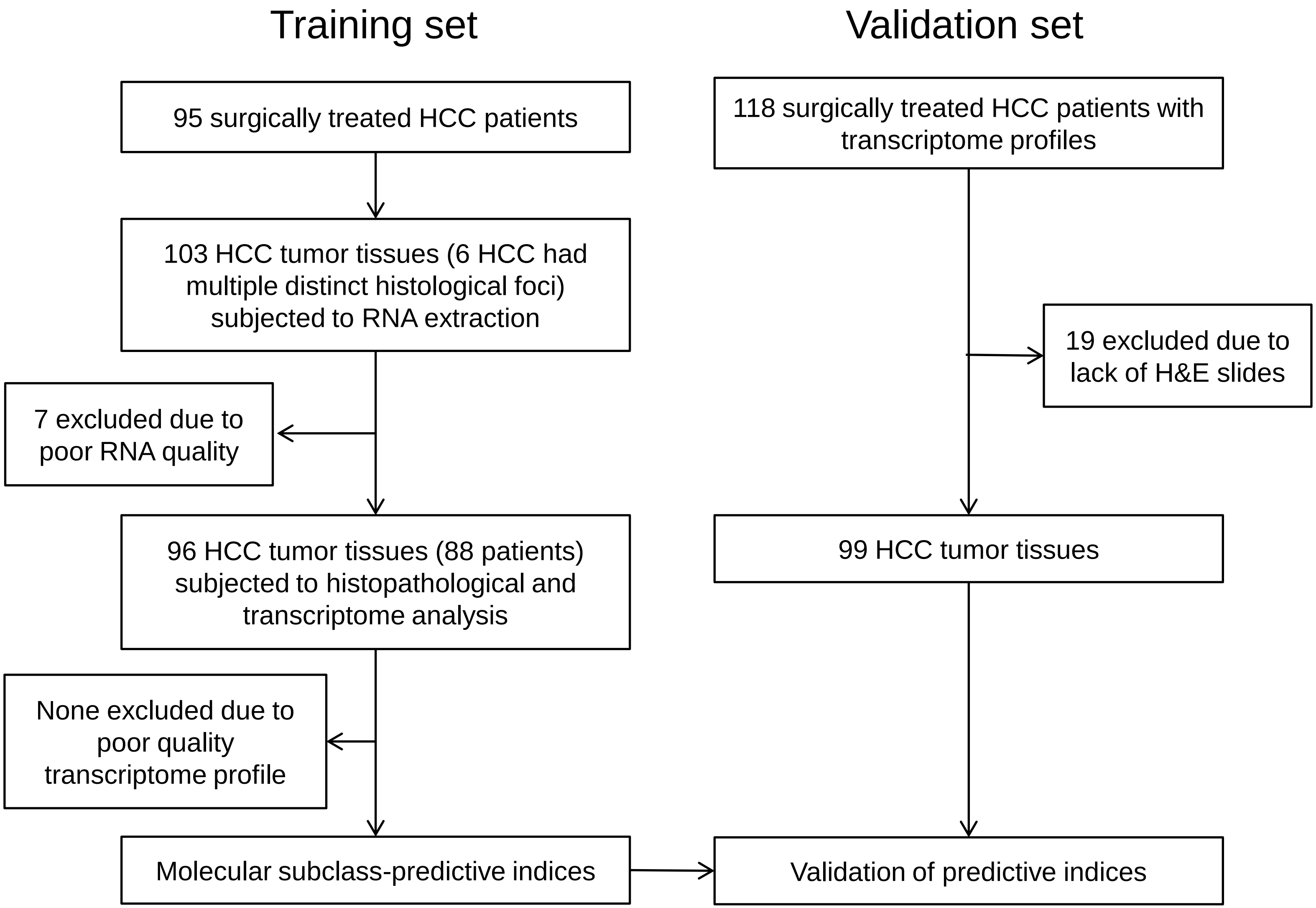

Training set
Validation set
95 surgically treated HCC patients
103 HCC tumor tissues (6 HCC had multiple distinct histological foci) subjected to RNA extraction
7 excluded due to poor RNA quality
96 HCC tumor tissues (88 patients) subjected to histopathological and transcriptome analysis
None excluded due to poor quality transcriptome profile
Molecular subclass-predictive indices
118 surgically treated HCC patients with transcriptome profiles
19 excluded due to lack of H&E slides
99 HCC tumor tissues
Validation of predictive indices

Figure S2

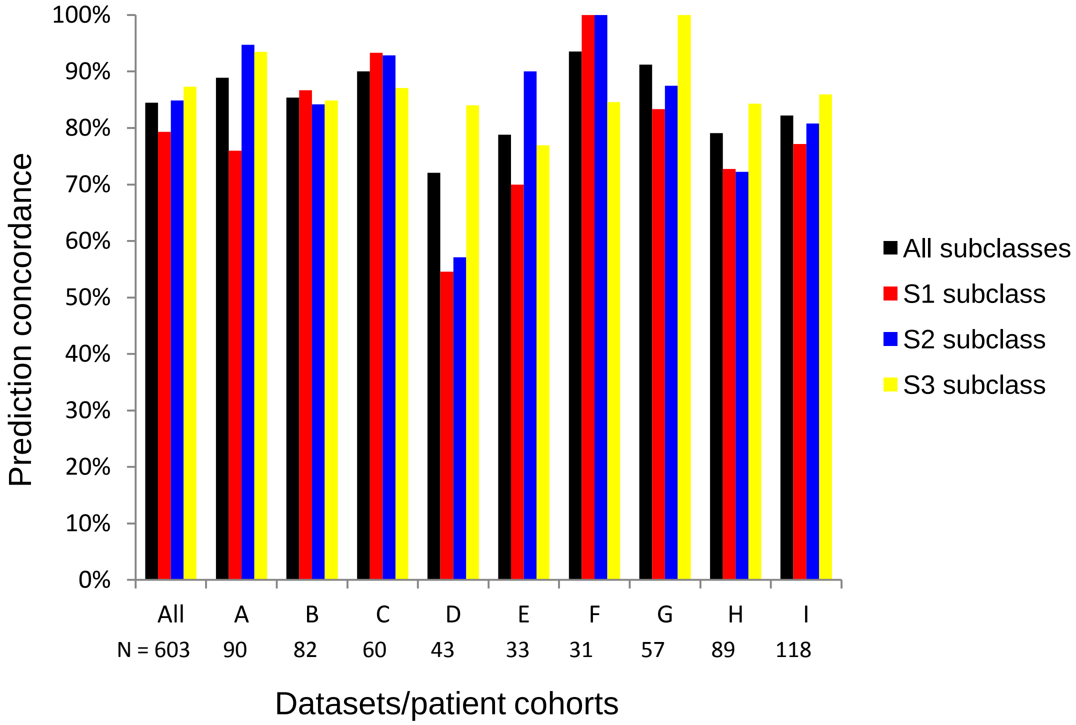

Prediction concordance
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
All
A
B
C
D
E
F
G
H
I
N = 603
90
82
60
43
33
31
57
89
118
Datasets/patient cohorts
All subclasses
S1 subclass
S2 subclass
S3 subclass

Figure S3

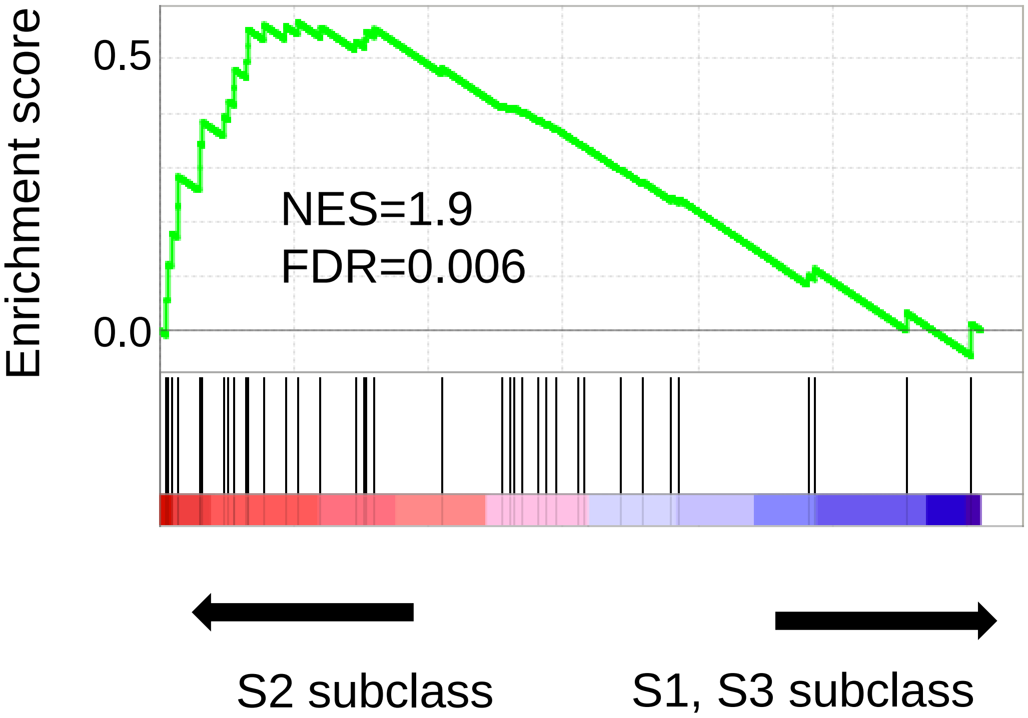

Enrichment score
0.5
0.0
NES=1.9
FDR=0.006
S2 subclass
S1, S3 subclass

Figure S4A

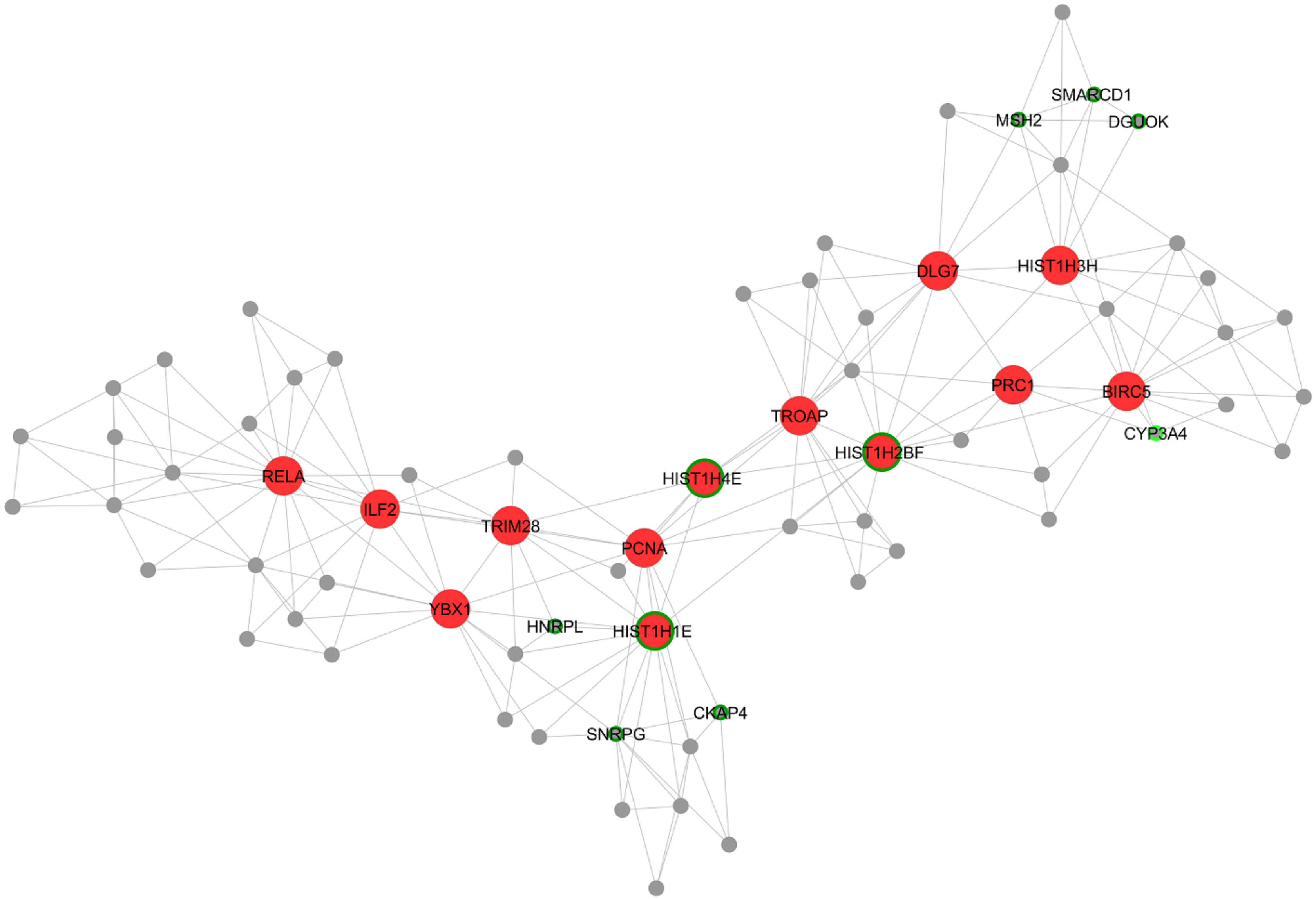

SMARCD1
MSH2
DGUOK
DLG7
HIST1H3H
TROAP
PRC1
BIRC5
HIST1H2BF
CYP3A4
HIST1H4E
RELA
ILF2
TRIM28
PCNA
YBX1
HNRPL
HIST1H1E
CKAP4
SNRPG

Figure S4B

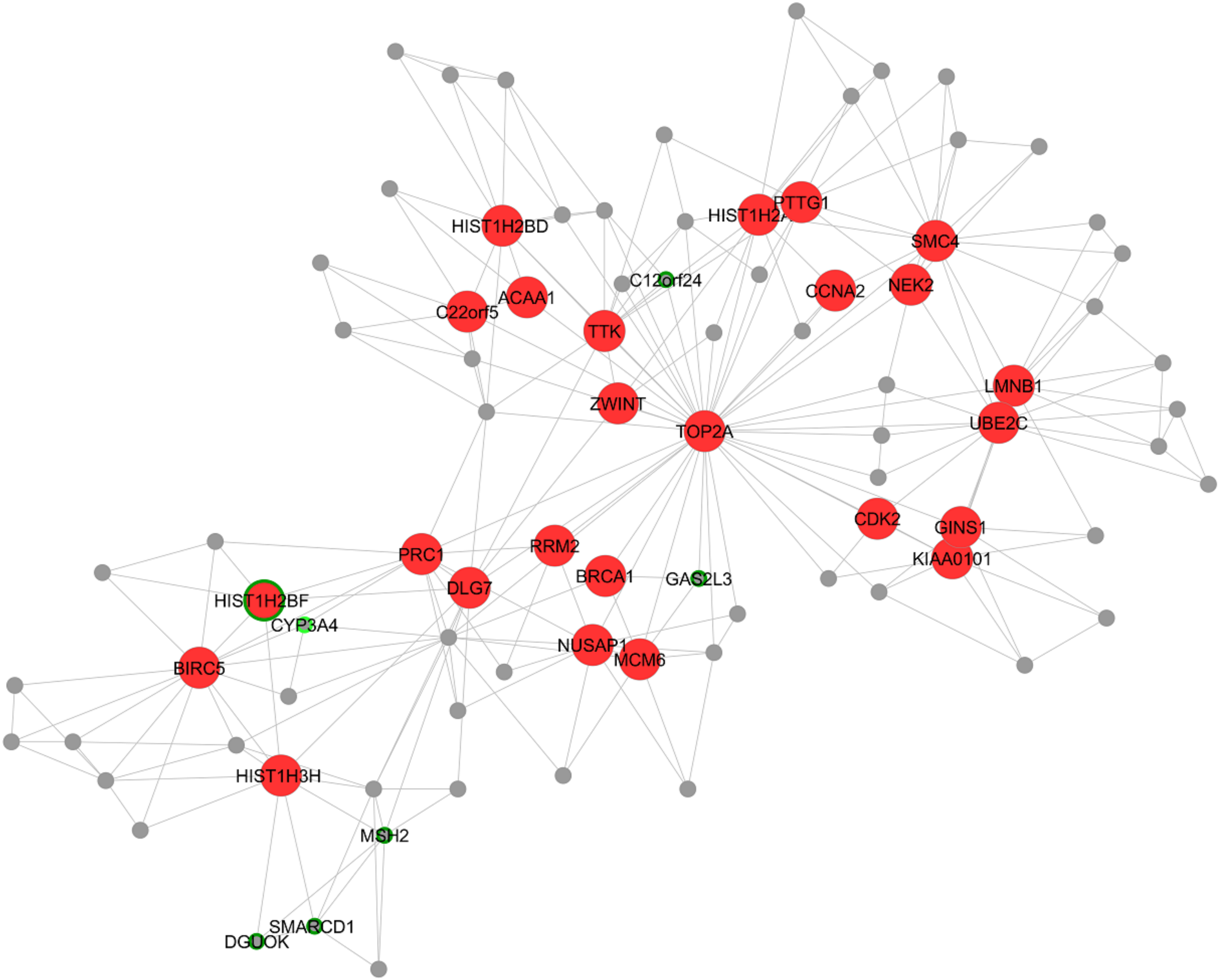

HIST1H2BD
PTTG1
HIST1H2A
SMC4
C12orf24
ACAA1
CCNA2
NEK2
C22orf5
TTK
LMNB1
ZWINT
UBE2C
TOP2A
CDK2
GINS1
PRC1
RRM2
KIAA0101
BRCA1
GAS2L3
DLG7
HIST1H2BF
CYP3A4
NUSAP1
MCM6
BIRC5
HIST1H3H
MSH2
SMARCD1
DGUOK

Figure S5

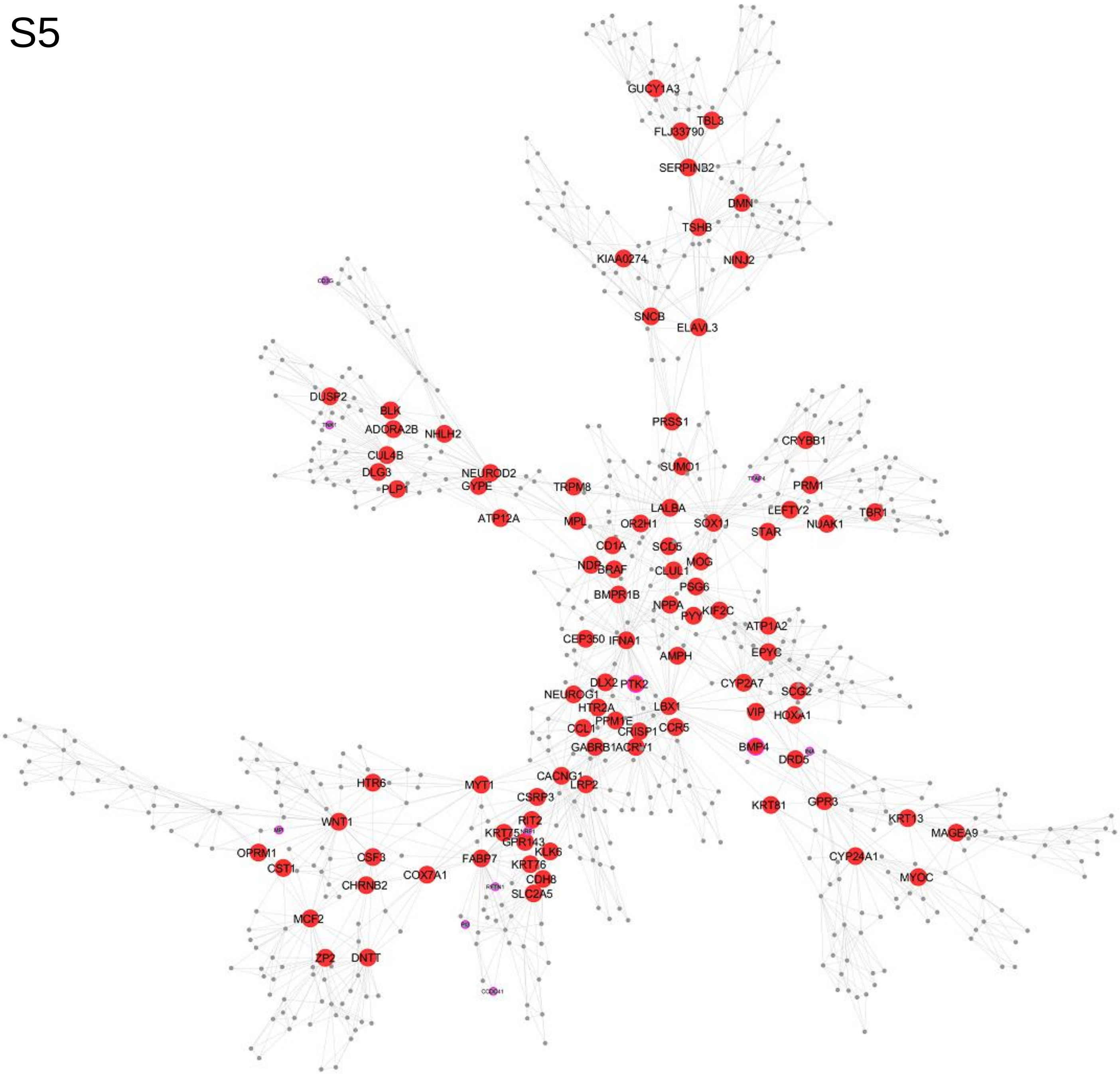

GUCY1A3
TBL3
FLJ33790
SERPINB2
DMN
TSHB
KIAA0274
NINJ2
SNCB
ELAVL3
CD3G
DUSP2
BLK
ADORA2B
NHLH2
CUL4B
DLG3
PLP1
NEUROD2
GYPE
ATP12A
TRPM8
MPL
PRSS1
SUMO1
LALBA
OR2H1
SOX11
CRYBB1
PRM1
LEFTY2
TBR1
STAR
NUAK1
CD1A
SCD5
NDP
BRAF
CLUL1
MOG
PSG6
BMPR1B
NPPA
PYY
KIF2C
ATP1A2
CEP350
IFNA1
AMPH
EPYC
DLX2
PTK2
NEUROG1
CYP2A7
SCG2
HTR2A
LBX1
VIP
HOXA1
CCL1
PRM1E
CRISP1
CCR5
GABRB1
ACRV1
BMP4
DRD5
CACNG1
LRP2
HTR6
MYT1
CSRP3
KRT81
GPR3
WNT1
RIT2
KRT13
KRT75
GPR143
MAGEA9
KLK6
OPRM1
CSF3
FABP7
CYP24A1
CST1
KRT76
COX7A1
CDH8
MYOC
CHRNB2
SLC2A5
MCF2
ZP2
DNTT

Figure S6

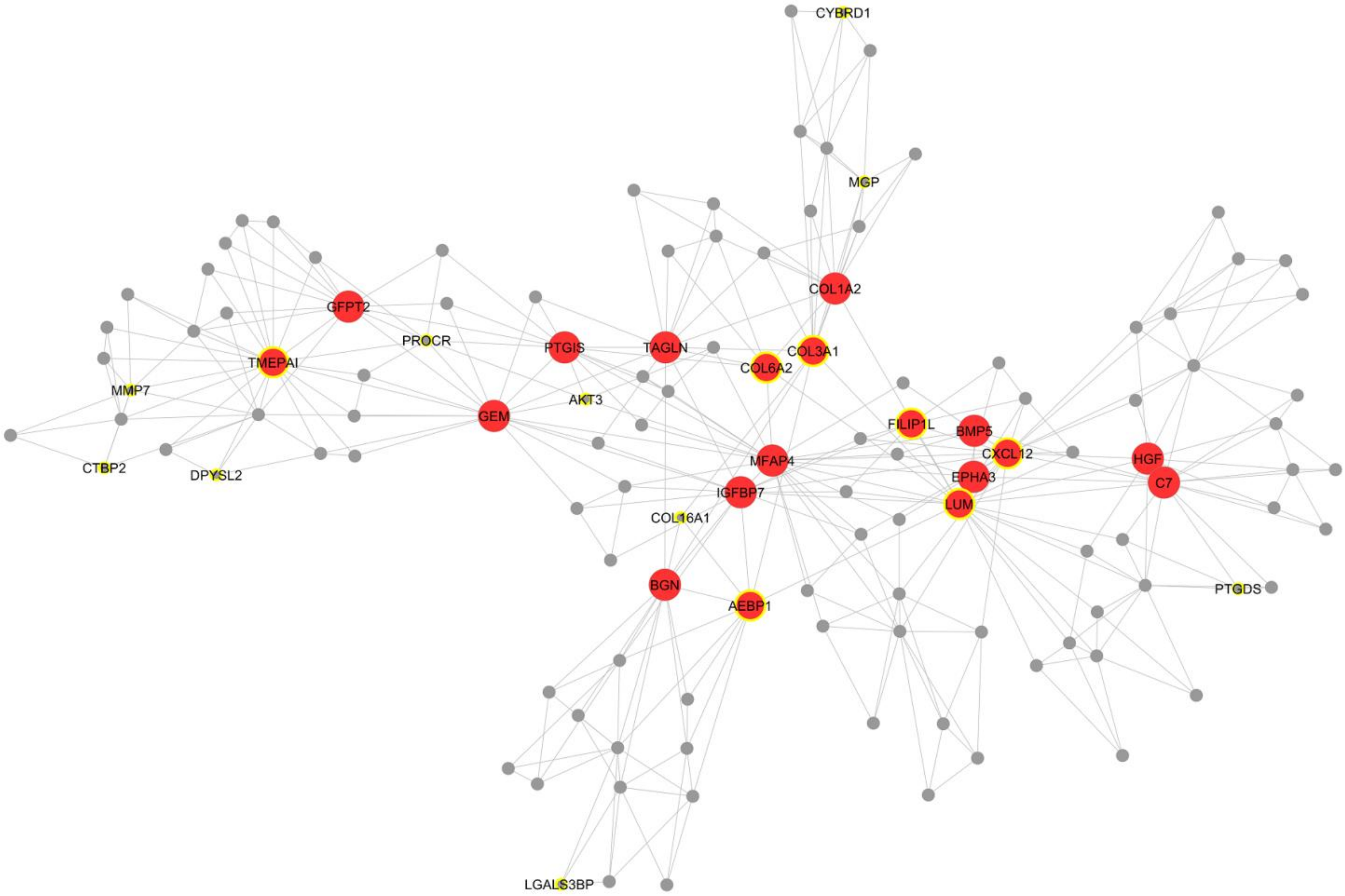

CYBRD1
MGP
COL1A2
GFPT2
PROCR
PTGIS
TAGLN
COL3A1
COL6A2
TMEPAI
MMP7
AKT3
GEM
FILIP1L
BMP5
CXCL12
HGF
MFAP4
EPHA3
C7
IGFBP7
LUM
CTBP2
DPYSL2
COL16A1
BGN
AEBP1
PTGDS
LGALS3BP

Figure S7

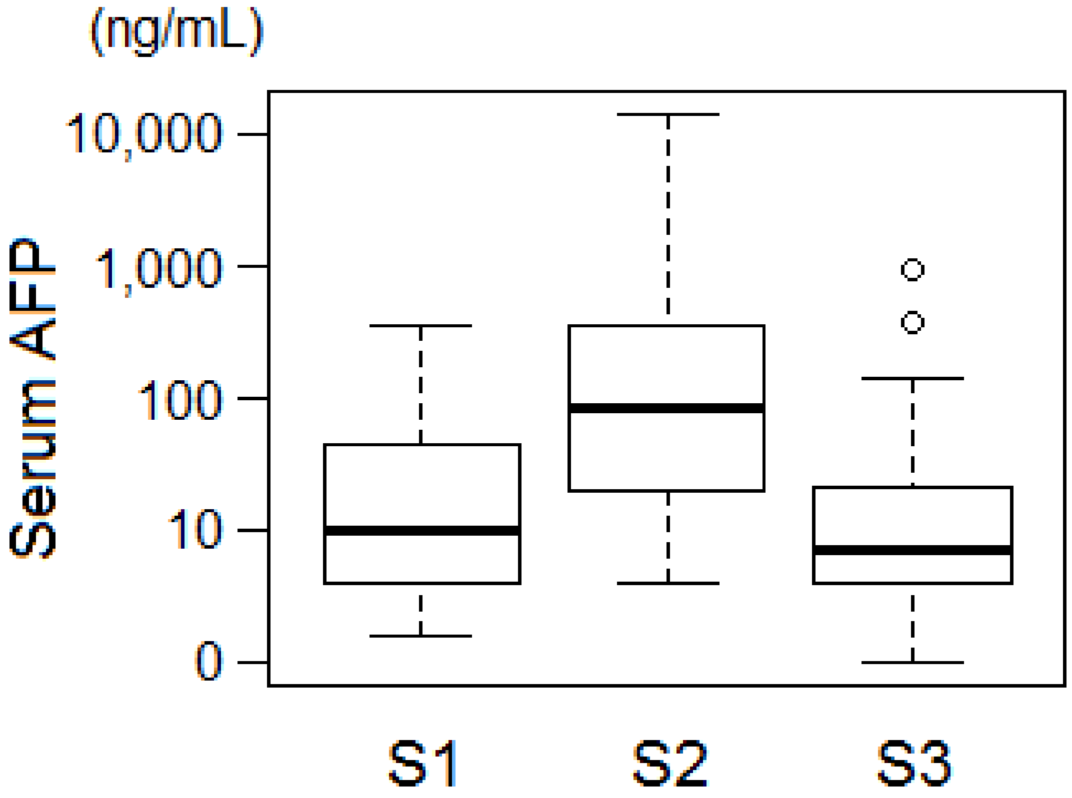


Figure S8A

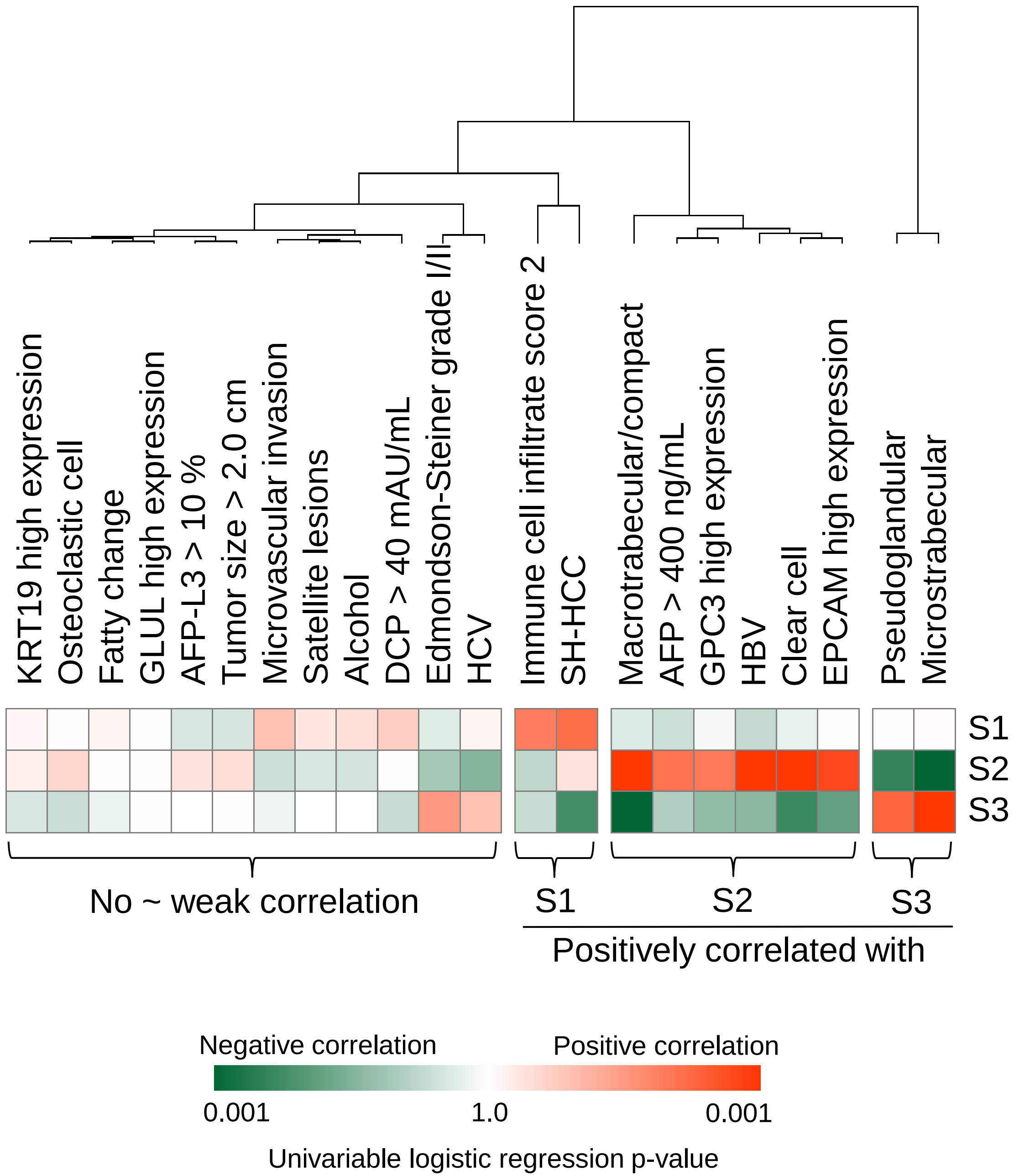

KRT19 high expression
Osteoclastic cell
Fatty change
GLUL high expression
AFP-L3 > 10 %
Tumor size > 2.0 cm
Microvascular invasion
Satellite lesions
Alcohol
DCP > 40 mAU/mL
Edmondson-Steiner grade I/II
HCV
Immune cell infiltrate score 2
SH-HCC
Macrotrabecular/compact
AFP > 400 ng/mL
GPC3 high expression
HBV
Clear cell
EPCAM high expression
Pseudoglandular
Microstrabecular
S1
S2
S3
No ~ weak correlation
S1
S2
S3
Positively correlated with
Negative correlation
Positive correlation
0.001
1.0
0.001
Univariable logistic regression p-value

Figure S8B

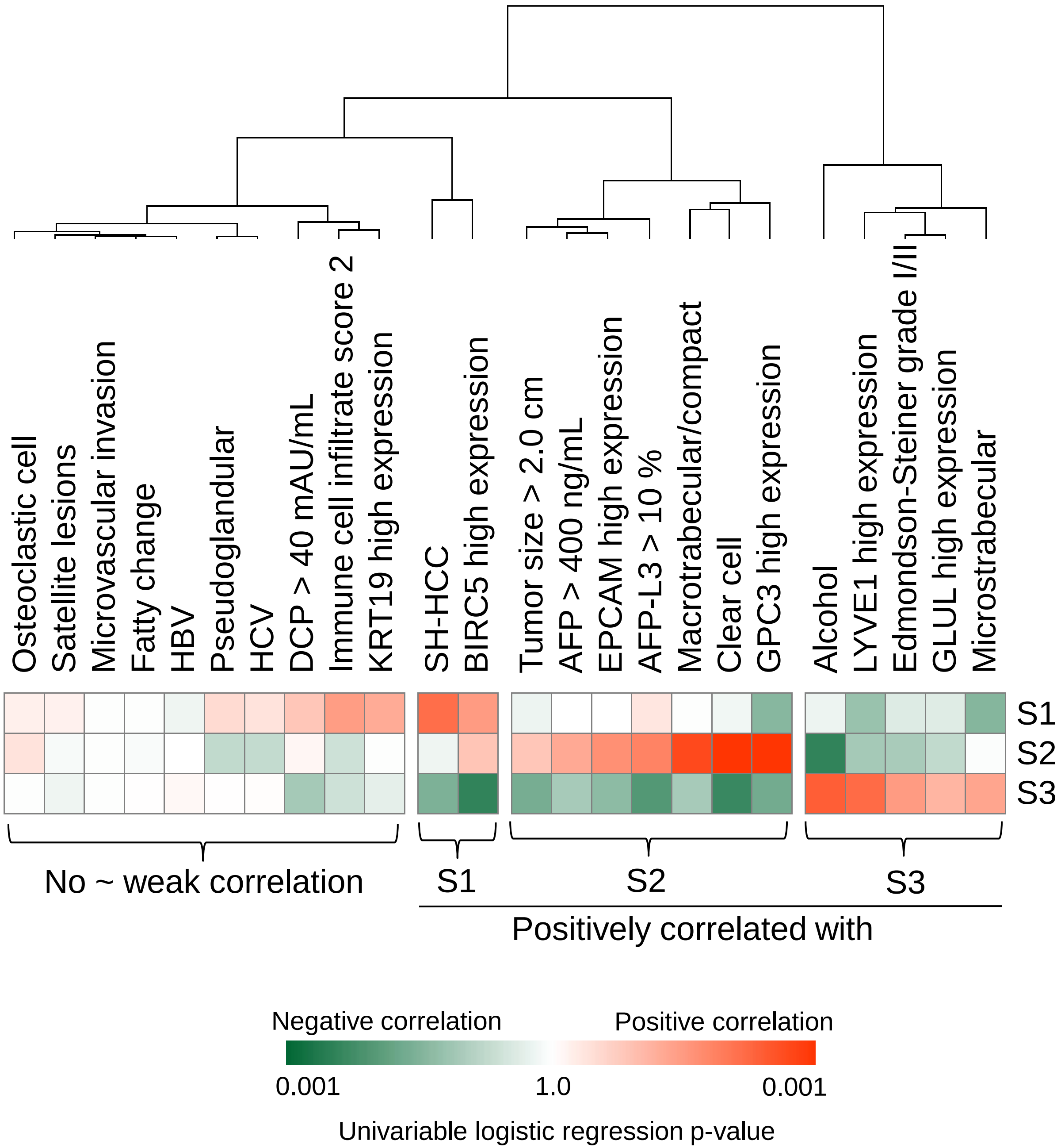

Osteoclastic cell
Satellite lesions
Microvascular invasion
Fatty change
HBV
Pseudoglandular
HCV
DCP > 40 mAU/mL
Immune cell infiltrate score 2
KRT19 high expression
SH-HCC
BIRC5 high expression
Tumor size > 2.0 cm
AFP > 400 ng/mL
EPCAM high expression
AFP-L3 > 10 %
Macrotrabecular/compact
Clear cell
GPC3 high expression
Alcohol
LYVE1 high expression
Edmondson-Steiner grade I/II
GLUL high expression
Microstrabecular
S1
S2
S3
No ~ weak correlation
S1
S2
S3
Positively correlated with
Negative correlation
Positive correlation
0.001
1.0
0.001
Univariable logistic regression p-value

Figure S9

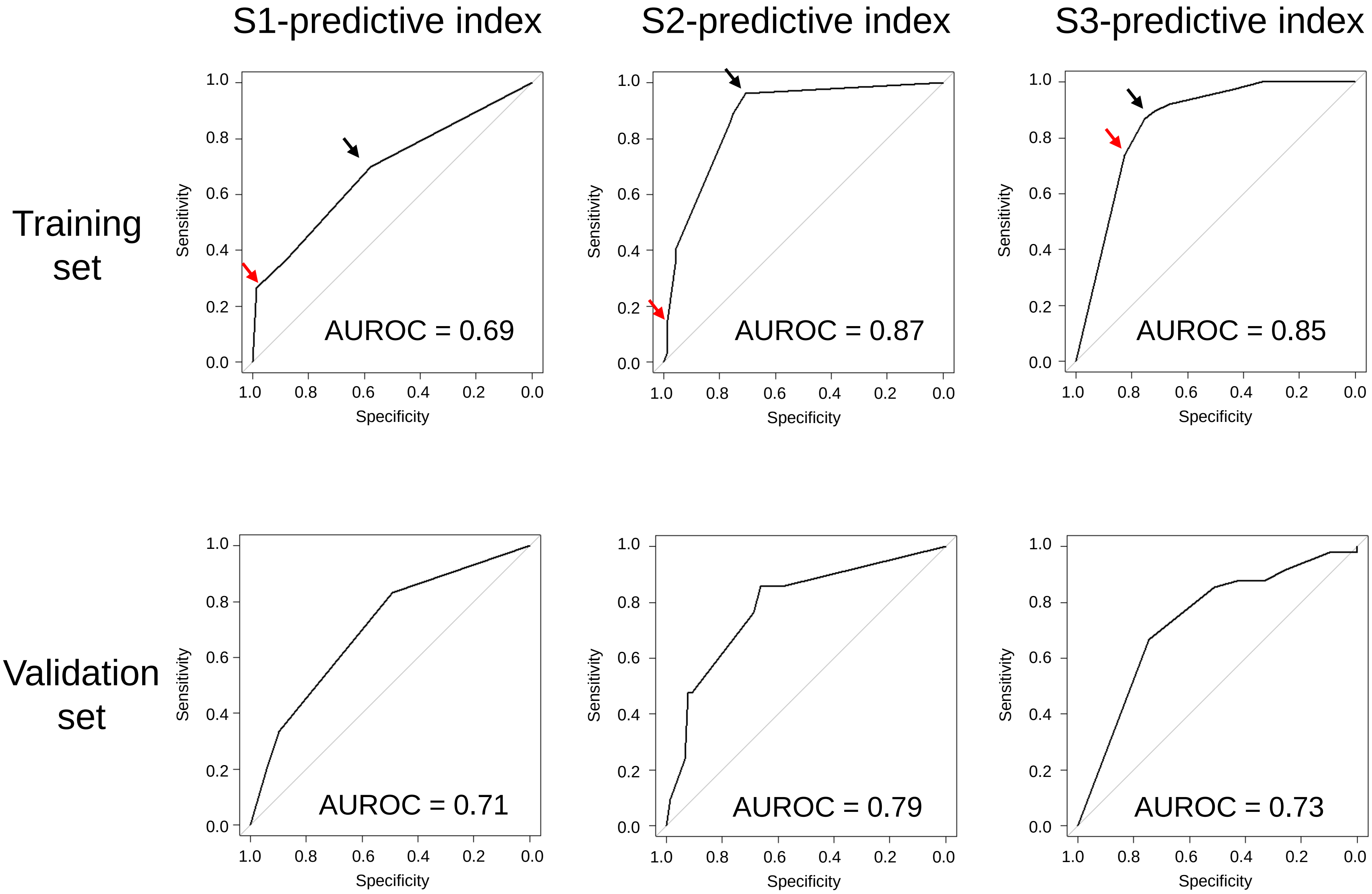
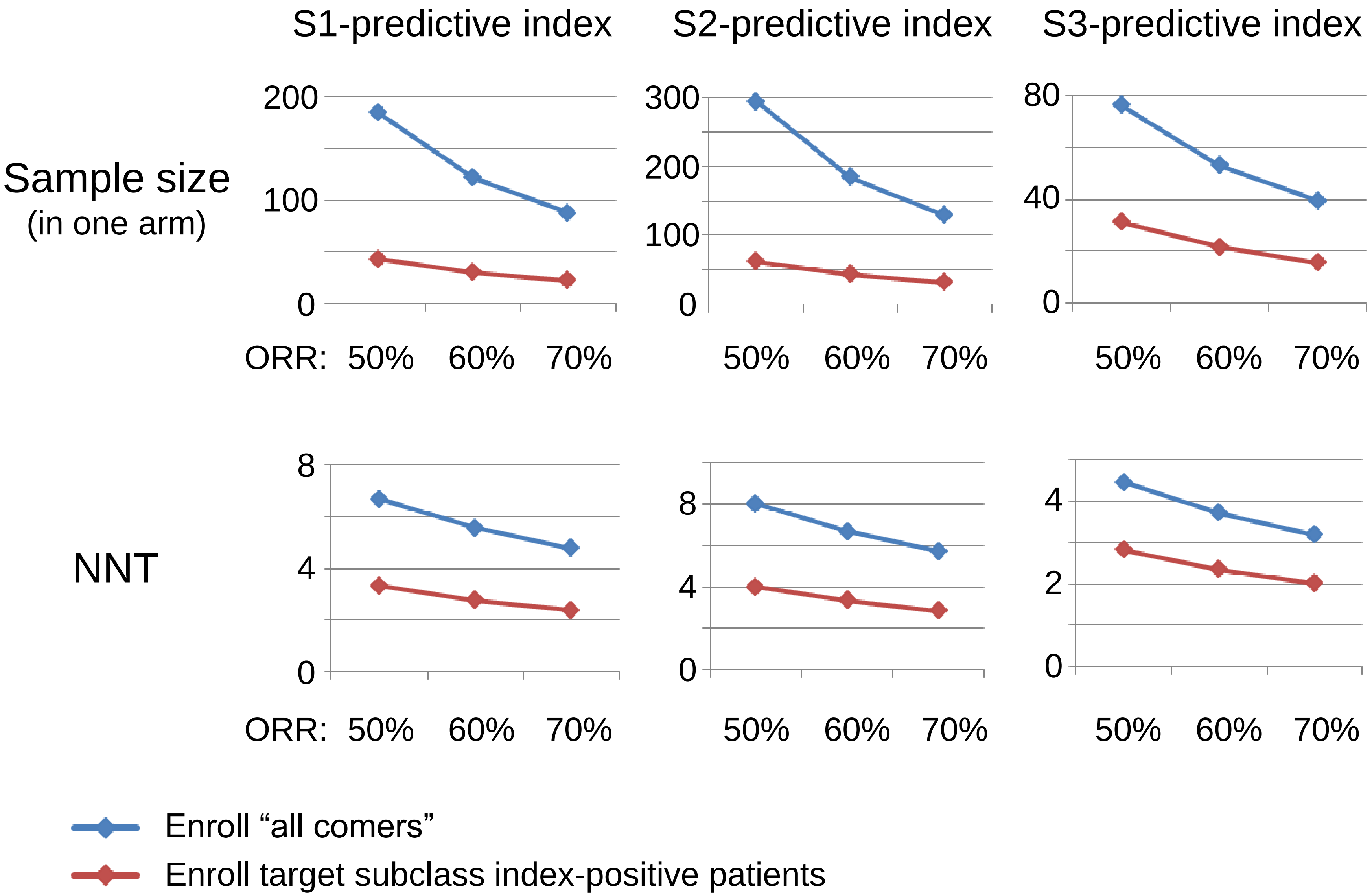

S1-predictive index
S2-predictive index
S3-predictive index
Training set
Validation set
Sensitivity
Specificity
1.0
0.8
0.6
0.4
0.2
0.0
AUROC = 0.69
AUROC = 0.87
AUROC = 0.85
AUROC = 0.71
AUROC = 0.79
AUROC = 0.73

Figure S10


S1-predictive index
S2-predictive index
S3-predictive index
Sample size
(in one arm)
NNT
200
100
0
300
200
100
0
80
40
0
8
4
0
8
4
0
4
2
0
ORR: 50% 60% 70%
50% 60% 70%
50% 60% 70%
ORR: 50% 60% 70%
50% 60% 70%
50% 60% 70%
Enroll “all comers”
Enroll target subclass index-positive patients