SUPPLEMENTARY TABLES

Supplementary Table 1. Publicly available gene sets and signatures used in the study.

Name	Reference
HCC classification	
Sia HCC immune class	Sia D, et al. Gastroenterology 2017;153:812-826
Chiang classification	Chiang D, et al. Cancer Res 2008;68:6779-6788
Hoshida classification	Hoshida Y, et al. Cancer Res 2009;69:7385-7392
Cluster A signature	Lee JS, et al. Hepatology 2004;40:667–76
Molecular pathways	
TGFB late signature	Coulouarn C, et al. Hepatology 2008: 47 2059-2067
MET signature	Kaposi-Novak P, et al. J Clin Invest 2006;116:1582-95
NOTCH signature	Villanueva A, et al. Gastroenterology 2012;143: 1660-1669
RB1 signature	Bollard J, et al. Gut 2017;66: 1286-1296
Antitumor immune response	
PD1 signaling signature	Quigley M, et al. Nat Med 2010;16:1147-51
Exhaustion signature	Quigley M, et al. Nat Med 2010;16:1147-51
Stromal enrichment score	Yoshihara K, et al. Nat Commun 2013;4:2612
Immune enrichment score	Yoshihara K, et al. Nat Commun 2013;4:2612
IFN signature	Ayers M, et al. J Clin Invest 2017; 127:2930–40
Immune cell infiltrate gene sets	Bindea G, et al. Immunity 2013;39:782-95

Supplementary Table 2. Baseline characteristics of the Western cohort by sample origin.

	Western Europe (n=117)	Western USA (n=70)*	P value
Age (years)	67 (40-83)	65 (29-91)	ns
< 60 years (n, %)	19 (16)	13 (24)	ns
Gender (male, %)	92 (79)	45 (82)	ns
Etiology:			
HBV+ (n, %)	23 (19)	18 (26)	
HBV/HDV+ (n, %)	1 (1)	2 (3)	0.001
HCV+ (n, %)	56 (48)	13 (19)	
Non-infected (n, %)	37 (32)	37 (53)	
Bilirubin (mg/dL)	1 (0.4-3.2)	0.7 (0.3-3.8)	ns
Albumin (g/L)	40 (25-54)	42 (22-49)	ns
Platelets (10^9/L)	161 (29-493)	154 (27-460)	ns
< 150x10^9/L (n, %)	55 (47)	26 (47)	ns
AFP > 400 IU/mL (n, %)	18 (16)	8 (17)	ns
Tumor size (cm)	4.2 (1.5-20)	4.4 (1-19)	ns
> 5 cm (n, %)	45 (41)	22 (42)	ns
BCLC stage (0-A, %)	87 (80)	35 (66)	ns
Multinodular disease (n, %)	29 (26)	13 (25)	ns
Advanced liver fibrosis (F3-4, %)	72 (88)	34 (64)	0.002
Cirrhosis (F4, %)	57 (70)	24 (45)	0.007
Microvascular invasion (yes, %)	44 (38)	36 (66)	0.001
Tumor grade (G3-4, %)	24 (27)	17 (31)	ns

*Baseline characteristics, except etiology, are missing for 15 (21%) patients in the Western USA subcohort

Supplementary Table 3. Viral genotypes in the Mongolian and Western cohorts.

	Mongolian Cohort (N=106)	Western Cohort (N=44)	p value
HBV genotype			
Genotype A (n, %)	0 (0)	2 (4.5)	
Genotype B (n, %)	0 (0)	2 (4.5)	
Genotype C (n, %)	0 (0)	12 (27.3)	<0.001
Genotype D (n, %)‡	95 (89.6)	19 (43.2)	
Non-genotypable* (n, %)	11 (10.4)	9 (20.5)	
HDV genotype			
Genotype 1 (n, %)	85 (95.5)	1 (33.3)	
Genotype 2 (n, %)	0 (0)	1 (33.3)	<0.001
Non-genotypable* (n, %)	4 (4.5)	1 (33.4)	
HBV mutations			
BCP A1762T (yes, %)†	10 (13.7)	21 (60)	<0.001
BCP G1764A (yes, %)†	15 (20.5)	23 (65.7)	<0.001
Precore G1896A (yes, %)†	21 (28.8)	16 (45.7)	ns

HBV, hepatitis B virus; HDV, hepatitis delta virus; BCP, basal pre-core
‡2 patients showed recombinant forms of the C and D genotypes
* non-genotypable due to technical failure
†HBV mutations were evaluated in 73 (69%) patients in the Mongolian cohort and 35 (80%) patients in the Western cohort

Supplementary Table 4. Baseline characteristics of HBV-infected Mongolian patients according to the 25th quartile of HBV-DNA load.

	Low HBV-DNA (<4 log copies/µg total DNA) (n=30)	High HBV-DNA (≥4 log copies/µg total DNA) (n=75)	p value
Age (years)	57.7 (18-71.3)	56,2 (41.1- 75.8)	ns
Gender (male, %)‡	15 (51.7)	37 (52.1)	ns
Etiology			
HBV (n, %)	8 (26.7)	7 (9.3)	
HBV/HDV (n, %)	20 (66.7)	56 (74.7)	ns
HBV/HCV/HDV (n, %)	1 (3.3)	11 (14.7)	
HBV/HCV (n, %)	1 (3.3)	1 (1.3)	
HDV+ (n, %)	21 (70)	67 (89.3)	0.021
Region			
Western (n, %)	3 (11.5)	17 (25.8)	
Central (n, %)	12 (46.2)	25 (37.9)	ns
Eastern (n, %)	4 (15.4)	4 (6.1)	
Ulaanbaatar (n, %)	7 (26.9)	20 (30.3)	
Liver fibrosis (F3-4, %)*	6 (20.7)	34 (47.2)	0.015
Tumor size (cm)	6 (3-14.9)	6.5 (1.6-20)	ns
Multinodular (yes, %)	0 (0)	12 (17.4)	0.032
BCLC stage (0-A, %)	24 (96)	50 (72.5)	0.02
BCP A1762T (yes, %)†	2 (11.1)	8 (14,5)	ns
BCP G1764A (yes, %)†	2 (11.1)	13 (23.6)	ns
Precore G1896A (yes, %)†	7 (38.9)	14 (25,5)	ns

HBV, hepatitis B virus; HCV, hepatitis C virus; HDV, hepatitis delta virus; BCLC, Barcelona Clinic Liver Cancer; BCP, basal pre-core
‡Gender information was missing in 9 individuals in the Mongolia cohort
*Fibrosis stage was evaluated in 168 (88%) in the Mongolia cohort
†HBV mutations were evaluated in 73 (69%) patients in the Mongolian cohort (18 HBV-DNA low and 55 HBV-DNA high)

Supplementary Table 5. Baseline characteristics of HDV-infected Mongolian patients according to the 25th quartile of HDV-RNA load.

	Low HDV-RNA (<2.4 log IU/ng total RNA)(n=23)	High HDV-RNA (≥2.4 log IU/ng total RNA) (n=66)	p value
Age (years)	57 (41.1-72.6)	56 (44.2-75.6)	ns
Gender (male, %)	8 (36.4)	33 (53.2)	ns
Etiology			
HBV/HDV (n, %)	17 (73.9)	60 (90.9)	ns
HBV/HCV/HDV (n, %)	6 (26.1)	6 (9.1)	
Region			
Western (n, %)	7 (35)	7 (12.3)	
Central (n, %)	5 (25)	28 (49.1)	ns
Eastern (n, %)	3 (15)	5 (8.8)	
Ulaanbaatar (n, %)	5 (25)	17 (29.8)	
Liver fibrosis (F3-4, %)*	8 (34.8)	30 (45.5)	ns
Multinodular (yes, %)	1 (4.8)	12 (20.3)	ns
BCLC stage (0-A, %)	18 (85.7)	43 (72.9)	ns
AFP > 400 IU/mL (n, %)	7 (35)	9 (19.6)	ns
ALT (IU/L)	49.5 (14-426)	72.3 (10-452)	0.013
BCP A1762T (yes, %)	2 (14.3)	6 (13.3)	ns
BCP G1764A (yes, %)	3 (21.4)	8 (17.8)	ns
Precore G1896A (yes, %)	4 (28.6)	8 (17.8)	ns

HBV, hepatitis B virus; HCV, hepatitis C virus; HDV, hepatitis delta virus; BCLC, Barcelona Clinic Liver Cancer; AFP, alfa-fetoprotein; ALT, alanine aminotransferase;BCP, basal pre-core
†‡Gender information was missing in 9 individuals in the Mongolia cohort
*Fibrosis stage was evaluated in 168 (88%) in the Mongolia cohort
†HBV mutations were evaluated in 68 (64%) patients in the Mongolia cohort

Supplementary Table 6. Baseline characteristics of HBV-infected Western patients classified according to the median HBV-DNA load.

	Low HBV-DNA (<5 log copies/µg total DNA) (n=20)	High HBV-DNA (≥5 log copies/µg total DNA) (n=21)	p value
Age (years)	63 (29-87)	62 (41-78)	ns
Gender (male, %)‡	12 (70.5)	17 (100%)	0.016
HDV+ (n, %)	1 (5)	2 (10)	ns
Liver fibrosis (F3-4, %)*	5 (41.7)	13 (76.5)	ns
Tumor size (cm)	5 (1.8-16)	5.5 (2-18)	ns
Multinodular (yes, %)	4 (23.5)	4 (21.1)	ns
BCLC stage (0-A, %)	11 (64.7)	16 (84.2)	ns
BCP A1762T (yes, %)†	8 (50)	13 (72.2)	ns
BCP G1764A (yes, %)†	9 (56.3)	14 (77.8)	ns
Precore G1896A (yes, %)†	7 (43.8)	9 (50)	ns

HBV, hepatitis B virus; HCV, hepatitis C virus; HDV, hepatitis delta virus; BCLC, Barcelona Clinic Liver Cancer; BCP, basal pre-core
‡Gender information was missing in 15 individuals in the Western cohort
*Fibrosis stage was evaluated in 135 (72%) in the Western cohort
†HBV mutations were evaluated in 35 (80%) patients in the Western cohort

Supplementary Table 7. Focal copy number alterations. P values correspond to Fisher test comparing Mongolian and Western cohorts. There was no difference in overall CNV burden between cohorts.

Chromosome arm	Gain, n (%)				Loss, n (%)			
	Mongolia (n=151)	Western (n=112)	Total (n=263)	p value	Mongolia (n=151)	Western (n=112)	Total (n=263)	p value
8p	18 (12%)	11 (10%)	29 (11%)	ns	59 (39%)	66 (59%)	125 (48%)	**0.002**
9q	11 (7%)	2 (2%)	13 (5%)	**0.047**	14 (9%)	24 (21%)	38 (14%)	**0.007**
1q	72 (48%)	55 (49%)	127 (48%)	ns	2 (1%)	6 (5%)	8 (3%)	**0.01**
1p	29 (19%)	9 (8%)	38 (14%)	**0.013**	8 (5%)	17 (15%)	25 (10%)	ns
2p	17 (11%)	10 (9%)	27 (10%)	ns	7 (5%)	8 (7%)	15 (6%)	ns
2q	15 (10%)	8 (7%)	23 (9%)	ns	9 (6%)	8 (7%)	17 (6%)	ns
3p	8 (5%)	5 (4%)	13 (5%)	ns	7 (5%)	8 (7%)	15 (6%)	ns
3q	13 (9%)	5 (4%)	18 (7%)	ns	3 (2%)	8 (7%)	11 (4%)	ns
4p	6 (4%)	5 (4%)	11 (4%)	ns	26 (17%)	20 (18%)	46 (17%)	ns
4q	2 (1%)	3 (3%)	5 (2%)	ns	37 (25%)	29 (26%)	66 (25%)	ns
5p	35 (23%)	31 (28%)	66 (25%)	ns	5 (3%)	5 (4%)	10 (4%)	ns
5q	24 (16%)	27 (24%)	51 (19%)	ns	12 (8%)	5 (4%)	17 (6%)	ns
6p	40 (26%)	28 (25%)	68 (26%)	ns	4 (3%)	4 (4%)	8 (3%)	ns
6q	19 (13%)	11 (10%)	30 (11%)	ns	22 (15%)	17 (15%)	39 (15%)	ns
7p	53 (35%)	30 (27%)	83 (32%)	ns	1 (1%)	3 (3%)	4 (2%)	ns
7q	55 (36%)	30 (27%)	85 (32%)	ns	1 (1%)	3 (3%)	4 (2%)	ns
8q	66 (44%)	61 (54%)	127(48%)	ns	8 (5%)	5 (4%)	13 (5%)	ns
9p	14(9%)	4 (4%)	18 (7%)	ns	23 (15%)	25 (22%)	48 (18%)	ns
10p	12 (8%)	10 (9%)	22 (8%)	ns	11 (7%)	10 (9%)	21 (8%)	ns
10q	6 (4%)	4 (4%)	10 (4%)	ns	26 (17%)	19 (17%)	45 (17%)	ns
11p	7 (5%)	4 (4%)	11 (4%)	ns	11 (7%)	14	25 (10%)	ns
11q	7 (5%)	4 (4%)	11 (4%)	ns	12 (8%)	15	27 (10%)	ns
12p	14 (9%)	9 (8%)	23 (9%)	ns	13 (9%)	14	27 (10%)	ns
12q	14 (9%)	9 (8%)	23 (9%)	ns	9 (6%)	11	20 (8%)	ns
13q	2 (1%)	5 (4%)	7 (3%)	ns	26(17%)	26	52 (20%)	ns
14q	7 (5%)	6 (5%)	13 (5%)	ns	16 (11%)	13	29 (11%)	ns
15q	6 (4%)	2 (2%)	8 (3%)	ns	18 (12%)	10	28 (11%)	ns
16p	5 3%)	6 (5%)	11 (4%)	ns	35 (23%)	28	63 (24%)	ns
16q	2 (1%)	6 (5%)	8 (3%)	ns	47 (31%)	41	88 (33%)	ns
17p	9 (6%)	6 (5%)	15 (6%)	ns	36 (24%)	23	59 (22%)	ns
17q	21(14%)	18 (16%)	39 (15%)	ns	7 (5%)	9	16 (6%)	ns
18p	11 (7%)	3 (3%)	14 (5%)	ns	20 (13%)	13	33 (13%)	ns
18q	8 (5%)	2 (2%)	10 (4%)	ns	23 (15%)	16	39 (15%)	ns
19p	19 (13%)	8 (7%)	27 (10%)	ns	18 (12%)	19	37 (14%)	ns
19q	20 (13%)	12 (11%)	32 (12%)	ns	13 (9%)	16	29 (11%)	ns
20p	33 (22%)	22 (20%)	55 (21%)	ns	5 (3%)	6	11 (4%)	ns
20q	37 (25%)	23 (21%)	60 (23%)	ns	4 (3%)	2	6 (2%)	ns
21q	5 (3%)	10 (9%)	15 (6%)	ns	39 (26%)	24	63 (24%)	ns
22q	12 (8%)	8 (7%)	20 (8%)	ns	23 (15%)	21	44(17%)	ns

Supplementary Table 8. Protein-coding mutations and tumor mutational burden. (TMB) in the in-house and external cohorts. TMB is shown as mutations/30 MB (Alexandrov, Nature 2013) and mutations/50 MB (Schulze, Nat Gen 2015) for comparison with previously published data.

	In-house cohorts		External cohorts			
	Mongolian	Western	Mongolian NCI	TCGA	European (Schulze)	Korean (Ahn)
Mutations	121	70	111	76	61	63
TMB (Mutations/30 Mb)	4.0	2.3	3.7	2.5	2.0	2.1
TMB (Mutations/50 Mb)	2.4	1.4	2.2	1.5	1.2	1.3

Supplementary Table 9. Mutations in DNA damage repair (DDR) genes in the Mongolian and Western cohorts.

Gene	Western cohort		Mongolian cohort	
	Patients harboring mutations (n)	**Patients harboring mutations (%)**	**Patients harboring mutations (n)**	**Patients harboring mutations (%)**
TP53	36	32.14	70	46.36
ATM	7	6.25	13	8.61
BRCA2	4	3.57	3	1.99
ATR	3	2.68	4	2.65
HERC2	3	2.68	3	1.99
POLE	3	2.68	2	1.32
ATRX	2	1.79	9	5.96
POLD1	2	1.79	4	2.65
REV3L	2	1.79	4	2.65
TP53BP1	2	1.79	4	2.65
PTEN	2	1.79	3	1.99
CUL3	2	1.79	2	1.32
HELQ	2	1.79	1	0.66
PER1	2	1.79	1	0.66
SMARCA4	1	0.89	7	4.64
SHPRH	1	0.89	6	3.97
HFM1	1	0.89	5	3.31
RIF1	1	0.89	5	3.31
FANCA	1	0.89	3	1.99
POLA1	1	0.89	3	1.99
SLX4	1	0.89	3	1.99
SMARCAD1	1	0.89	2	1.32
MSH6	1	0.89	1	0.66
PARP4	1	0.89	1	0.66
SMC5	1	0.89	1	0.66
SMC6	1	0.89	1	0.66
WRN	1	0.89	1	0.66
FANCD2	0	0.00	6	3.97
ASCC3	0	0.00	5	3.31
FANCM	0	0.00	5	3.31
POLQ	0	0.00	5	3.31
BLM	0	0.00	4	2.65
MDC1	0	0.00	3	1.99
PALB2	0	0.00	2	1.32
RAD50	0	0.00	2	1.32
DDB1	0	0.00	1	0.66
MLH3	0	0.00	1	0.66
RFC1	0	0.00	1	0.66
TOPB1	0	0.00	0	0.00
LIGA4	0	0.00	0	0.00

Supplementary Table 10. 250 genes most frequently mutated genes in the in-house HCC cohort (Mongolia plus Western).

Gene	Western cohort		Mongolian cohort	
	Patients harboring mutations (n)	Patients harboring mutations (%)	Patients harboring mutations (n)	Patients harboring mutations (%)
TP53	36	32	70	46
CTNNB1	48	43	55	36
TTN	19	17	55	36
ALB	13	12	26	17
ARID1A	11	10	26	17
KMT2A	3	3	16	11
RYR2	8	7	24	16
HMCN1	7	6	23	15
APOB	5	4	23	15
CSMD3	14	13	22	15
SYNE1	6	5	20	13
CUBN	8	7	19	13
LAMA1	3	3	18	12
AXIN1	8	7	17	11
ABCA13	3	3	16	11
FSIP2	10	9	16	11
KMT2B	5	4	11	7
DNAH7	4	4	16	11
KMT2C	3	3	13	9
FLG	1	1	16	11
LRP1B	11	10	15	10
NOTCH3	1	1	10	7
OBSCN	11	10	15	10
DNAH8	1	1	14	9
KRT7	0	0	14	9
TSC2	1	1	14	9
PTPN13	2	2	14	9
AHNAK2	1	1	14	9
DNAH9	3	3	14	9
ATM	7	6	13	9
DYNC2H1	5	4	13	9
COL12A1	6	5	13	9
ZNF469	4	4	13	9
COL6A3	2	2	13	9

FREM2	9	8	10	7
IGSF9B	9	8	1	1
FAT4	6	5	12	8
LRBA	1	1	12	8
AKAP9	4	4	12	8
DNAH5	8	7	12	8
KIAA1109	2	2	12	8
ANK3	2	2	12	8
ALMS1	2	2	11	7
CELSR1	3	3	11	7
RYR1	6	5	11	7
KMT2D	7	6	16	11
PLXNA4	1	1	11	7
VCAN	3	3	11	7
PCDH15	4	4	11	7
EYS	5	4	11	7
RB1	5	4	11	7
CMYA5	0	0	11	7
MYO15A	2	2	11	7
TNR	4	4	11	7
RNF213	1	1	11	7
LRRIQ1	5	4	11	7
PRSS3	8	7	4	3
TAS2R43	8	7	6	4
MACF1	8	7	7	5
SPTA1	8	7	8	5
DNAH3	3	3	10	7
DNAH10	5	4	10	7
GPR112	1	1	10	7
MGAM	1	1	10	7
CHD7	2	2	10	7
COL11A1	5	4	10	7
CNTNAP5	6	5	10	7
SORCS3	2	2	10	7
RELN	3	3	10	7
FREM1	4	4	10	7
WDFY4	3	3	10	7
PTPRZ1	3	3	10	7
MTOR	2	2	10	7

PTPRB	2	2	10	7
SYNE2	3	3	10	7
MUC6	6	5	10	7
KEAP1	6	5	10	7
MYH13	1	1	10	7
RYR3	5	4	10	7
SLC7A8	0	0	10	7
USH2A	6	5	10	7
PLEC	7	6	6	4
FLNB	7	6	1	1
GPR98	7	6	9	6
HSPG2	2	2	9	6
CACNA2D1	5	4	9	6
NAV3	1	1	9	6
DNAH2	3	3	9	6
DCHS1	3	3	9	6
CHD9	3	3	9	6
NFE2L2	1	1	9	6
PCDH10	2	2	9	6
ATRX	2	2	9	6
SDK2	0	0	9	6
MKI67	1	1	9	6
MYH15	6	5	9	6
NBEA	5	4	9	6
ZNF462	2	2	9	6
FBN2	4	4	9	6
OTOG	1	1	9	6
ARID2	4	4	9	6
TRPM6	5	4	9	6
NEB	3	3	9	6
LAMC3	1	1	9	6
ZNF831	3	3	9	6
PXDNL	1	1	9	6
FMN2	4	4	9	6
MUC17	2	2	9	6
PIEZO2	6	5	5	3
ASH1L	6	5	4	3
STARD9	6	5	5	3
TRPS1	6	5	2	1

GRM5	6	5	4	3
TENM3	6	5	6	4
LRP2	6	5	5	3
OR2L8	6	5	2	1
PTPN21	6	5	1	1
FRG1B	6	5	0	0
PREX2	6	5	2	1
DOCK2	6	5	4	3
ILDR2	6	5	3	2
KIF26B	6	5	5	3
LRP1	6	5	8	5
JMY	0	0	8	5
COL4A5	1	1	8	5
AMER3	2	2	8	5
KDM3B	1	1	8	5
MYH7	1	1	8	5
EPHA3	3	3	8	5
CNTN6	3	3	8	5
PTPRS	0	0	8	5
NID2	1	1	8	5
CCDC168	3	3	8	5
AKAP6	1	1	8	5
NIPBL	1	1	8	5
UTRN	3	3	8	5
KIF21B	1	1	8	5
PLXNA3	0	0	8	5
HCN1	0	0	8	5
IGSF10	2	2	8	5
KIAA1731	1	1	8	5
NLRP3	2	2	8	5
NLRP8	3	3	8	5
GTF3C1	0	0	8	5
CHD6	0	0	8	5
UNC80	2	2	8	5
NAV2	2	2	8	5
NOS1	0	0	8	5
KRT6A	0	0	8	5
TTC40	4	4	8	5
IRX4	1	1	8	5

PKHD1L1	4	4	8	5
ITGB8	2	2	8	5
BTAF1	1	1	8	5
ARAP3	2	2	8	5
TRIP12	2	2	8	5
WDFY3	3	3	8	5
FLNC	1	1	8	5
NWD1	1	1	8	5
GABRA1	1	1	8	5
PDZD2	0	0	8	5
TMEM132C	1	1	8	5
DCC	3	3	8	5
SCN2A	1	1	8	5
DLEC1	0	0	8	5
SPTBN4	5	4	8	5
FAT2	1	1	8	5
MDN1	3	3	8	5
RNF17	1	1	8	5
MAP2	0	0	8	5
TELO2	0	0	8	5
ERBB4	4	4	8	5
EPHA7	2	2	8	5
CTNNA2	4	4	8	5
PKD1L1	4	4	8	5
NRXN1	2	2	8	5
TNC	3	3	8	5
USP34	4	4	7	5
PCDHB5	1	1	7	5
PCDH7	4	4	7	5
DOCK10	3	3	7	5
PTPRT	5	4	7	5
CCDC30	0	0	7	5
SVEP1	3	3	7	5
PCDHA6	0	0	7	5
SLIT1	0	0	7	5
ADAM21	3	3	7	5
NIN	1	1	7	5
DMXL2	2	2	7	5
DMXL1	0	0	7	5

HERC1	5	4	7	5
LRRC7	3	3	7	5
TG	5	4	7	5
ATP10A	2	2	7	5
BPTF	1	1	7	5
CDHR2	1	1	7	5
WNK2	1	1	7	5
CDKN2A	2	2	7	5
JMJD1C	1	1	7	5
PLXNA1	2	2	7	5
BMPER	2	2	7	5
LAMC2	3	3	7	5
SI	3	3	7	5
NLRP2	1	1	7	5
PIK3R4	1	1	7	5
RIMS1	1	1	7	5
SPEG	3	3	7	5
CACNA1E	2	2	7	5
KIAA1217	4	4	7	5
IPO9	0	0	7	5
DYNC1H1	4	4	7	5
NALCN	3	3	7	5
PXDN	4	4	7	5
DSCAM	3	3	7	5
UNC5D	4	4	7	5
ZFP36L1	0	0	7	5
PRRC2A	0	0	7	5
KIAA1462	1	1	7	5
ZFPM2	2	2	7	5
CACNA1D	3	3	7	5
KIAA1549L	2	2	7	5
BIRC6	2	2	7	5
PRDM5	3	3	7	5
COL6A6	3	3	7	5
PCDH18	4	4	7	5
ARID1B	5	4	7	5
MYO3A	5	4	7	5
MAP1B	0	0	7	5
NCOA6	1	1	7	5

UNC13C	2	2	7	5
SACS	5	4	7	5
SLC4A3	3	3	7	5
SYCP2	0	0	7	5
RBMXL1	2	2	7	5
SMARCA4	1	1	7	5
TPO	4	4	7	5
TMEM132D	0	0	7	5
ANK2	3	3	7	5
ZNF804B	5	4	7	5
ABCA1	0	0	7	5
CNOT1	1	1	7	5
TOPAZ1	3	3	7	5
NPAP1	2	2	7	5
COL22A1	2	2	7	5
COL4A1	1	1	7	5
SCN4A	4	4	7	5
ZNF521	5	4	7	5
MUC12	3	3	7	5
SND1	1	1	7	5
SDK1	1	1	7	5

Supplementary Table 11. 100 genes with statistical differences between Mongolian and Western cohorts.

Gene	% mutations Mongolian	% mutations Western	Odds ratio Mongolian vs Western	P value Mongolian vs Western	% mutations Europe	% mutations USA	P value Europe vs USA
TP53	46.4%	32.1%	0.548	0.022	26.1%	41.9%	0.098
TTN	36.4%	17.0%	0.357	0.001	20.3%	11.6%	0.305
RYR2	15.9%	7.1%	0.407	0.036	8.7%	4.7%	0.708
APOB	15.2%	4.5%	0.26	0.005	1.4%	9.3%	0.071
HMCN1	15.2%	6.3%	0.371	0.03	1.4%	14.0%	0.013
SYNE1	13.2%	5.4%	0.371	0.038	4.3%	7.0%	0.674
LAMA1	11.9%	2.7%	0.203	0.006	2.9%	2.3%	1
FLG	10.6%	0.9%	0.076	0.001	1.4%	0.0%	1
ABCA13	10.6%	2.7%	0.232	0.015	4.3%	0.0%	0.284
KMT2A	10.6%	2.7%	0.232	0.015	2.9%	2.3%	1
DNAH7	10.6%	3.6%	0.313	0.036	5.8%	0.0%	0.296
NOTCH3	6.6%	0.9%	7.828	0.027	1.4%	0.0%	1
KRT7	9.3%	0.0%	0	0	0.0%	0.0%	NA
DNAH8	9.3%	0.9%	0.088	0.003	0.0%	2.3%	0.384
TSC2	9.3%	0.9%	0.088	0.003	0.0%	2.3%	0.384
AHNAK2	9.3%	0.9%	0.088	0.003	0.0%	2.3%	0.384
PTPN13	9.3%	1.8%	0.178	0.016	2.9%	0.0%	0.523
DNAH9	9.3%	2.7%	0.269	0.041	4.3%	0.0%	0.284
COL6A3	8.6%	1.8%	0.193	0.028	1.4%	2.3%	1
LRBA	7.9%	0.9%	0.104	0.009	0.0%	2.3%	0.384
ANK3	7.9%	1.8%	0.211	0.029	1.4%	2.3%	1
KIAA1109	7.9%	1.8%	0.211	0.029	1.4%	2.3%	1
CMYA5	7.3%	0.0%	0	0.003	0.0%	0.0%	NA
RNF213	7.3%	0.9%	0.115	0.015	0.0%	2.3%	0.384
PLXNA4	7.3%	0.9%	0.115	0.015	0.0%	2.3%	0.384
ALMS1	7.3%	1.8%	0.231	0.047	1.4%	2.3%	1
MYO15A	7.3%	1.8%	0.231	0.047	1.4%	2.3%	1
SLC7A8	6.6%	0.0%	0	0.006	0.0%	0.0%	NA
MYH13	6.6%	0.9%	0.127	0.027	1.4%	0.0%	1
MGAM	6.6%	0.9%	0.127	0.027	0.0%	2.3%	0.384
GPR112	6.6%	0.9%	0.127	0.027	1.4%	0.0%	1
SDK2	6.0%	0.0%	0	0.011	0.0%	0.0%	NA
NAV3	6.0%	0.9%	0.142	0.047	1.4%	0.0%	1

MKI67	6.0%	0.9%	0.142	0.047	1.4%	0.0%	1
NFE2L2	6.0%	0.9%	0.142	0.047	1.4%	0.0%	1
OTOG	6.0%	0.9%	0.142	0.047	1.4%	0.0%	1
LAMC3	6.0%	0.9%	0.142	0.047	1.4%	0.0%	1
PXDNL	6.0%	0.9%	0.142	0.047	0.0%	2.3%	0.384
GTF3C1	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
NOS1	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
KRT6A	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
JMY	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
PTPRS	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
HCN1	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
MAP2	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
DLEC1	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
CHD6	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
PDZD2	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
PLXNA3	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
TELO2	5.3%	0.0%	0	0.023	0.0%	0.0%	NA
TMEM132D	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
PCDHA6	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
DMXL1	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
PRRC2A	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
ZFP36L1	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
SLIT1	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
MAP1B	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
SYCP2	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
CCDC30	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
IPO9	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
ABCA1	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
TNN	4.6%	0.0%	0	0.022	0.0%	0.0%	NA
LMTK3	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
PDGFRA	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
CD1C	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
FAM184B	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
SLC23A1	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
TTLL5	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
LILRA2	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
PCSK5	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
ASTN1	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
ZRSR2	4.0%	0.0%	0	0.04	0.0%	0.0%	NA

DLGAP3	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
ABHD17A	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
BAZ2B	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
FAR2	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
KCNT1	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
GABRB2	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
HEPACAM2	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
SLC44A5	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
ANKRD31	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
PHACTR4	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
FANCD2	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
COL16A1	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
CYP2A13	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
SPEN	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
PABPC5	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
STK31	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
CCDC146	4.0%	0.0%	0	0.04	0.0%	0.0%	NA
IGSF9B	0.7%	8.0%	13.107	0.002	0.0%	0.0%	NA
FLNB	0.7%	6.3%	10	0.012	0.0%	0.0%	NA
PTPN21	0.7%	5.4%	8.491	0.044	0.0%	0.0%	NA
FRG1B	0.0%	5.4%	inf	0.006	0.0%	0.0%	NA
KAT6A	0.0%	4.5%	inf	0.013	0.0%	0.0%	NA
TDO2	0.0%	3.6%	inf	0.032	0.0%	0.0%	NA
TAF1A	0.0%	3.6%	inf	0.032	0.0%	0.0%	NA
ASB14	0.0%	3.6%	inf	0.032	0.0%	0.0%	NA
KIF20A	0.0%	3.6%	inf	0.032	0.0%	0.0%	NA
SFSWAP	0.0%	3.6%	inf	0.032	0.0%	0.0%	NA
IGSF3	0.0%	3.6%	inf	0.032	0.0%	0.0%	NA

Supplementary Table 12. Genes more frequently mutated in Mongolia versus other cohorts. Genes significantly mutated in 1 or more external non-Mongolian cohorts are shown.

Gene	% Mongolian in-house cohort	Number of cohorts with Significant diff	% Western in-house cohort	P VALUE Mongolian vs Western	% Korean	P VALUE Mongolian vs Korean	% European	P VALUE Mongolian vs European	% TCGA	P VALUE Mongolian vs TCGA	% Mongolian NCI	P VALUE in-house Mongolian vs Mongolian NCI
	Western and asian non-Mongolian cohorts										**Mongolian NCI cohort**	
TP53	46	4	32	0.022	31	0.00	22	0.000	28	0.000	30	0.020
LAMA1	12	4	3	0.006	6	0.03	3	0.001	4	0.003	8	0.496
KRT7	9	4	0	0.000	0	0.00	1	0.000	1	0.000	0	0.006
PTPN13	9	4	2	0.016	1	0.00	1	0.000	4	0.015	3	0.099
COL6A3	9	4	2	0.028	3	0.04	3	0.036	4	0.026	8	1.000
GPR112	7	4	1	0.027	0	0.00	1	0.007	0	0.000	0	0.033
ALMS1	7	4	2	0.047	3	0.04	2	0.006	3	0.026	3	0.233
MYO15A	7	3	2	0.047	3	0.04	2	0.006	1	0.001	4	0.556
PRRC2A	5	3	0	0.022	0	0.00	1	0.0497	0	0.001	0	0.100
DLEC1	5	4	0	0.023	1	0.02	1	0.016	1	0.025	4	1.000
JMY	5	4	0	0.023	0	0.00	0	0.003	0	0.000	0	0.057
KRT6A	5	4	0	0.023	1	0.03	0	0.000	1	0.008	0	0.057
TELO2	5	4	0	0.023	0	0.00	1	0.016	1	0.001	3	0.508
PLXNA3	5	4	0	0.023	0	0.00	1	0.016	1	0.025	3	0.508
ARID1A	17	3	10	0.107	3	0.00	10	0.043	7	0.001	4	0.002
KMT2C	9	3	3	0.062	3	0.01	0	0.000	3	0.008	6	0.591
CSMD3	15	3	13	0.718	8	0.04	6	0.008	8	0.034	13	0.674
AGRN	4	3	1	0.244	0	0.02	0	0.015	0	0.001	6	0.730
CUBN	13	3	7	0.217	6	0.04	3	0.000	5	0.005	10	0.496
AKAP9	8	3	4	0.193	3	0.02	2	0.023	1	0.000	0	0.011
PRDM5	5	3	3	0.524	0	0.01	1	0.031	1	0.018	1	0.441
PIK3R4	5	3	1	0.143	1	0.03	1	0.031	1	0.009	3	0.722
CELSR1	7	3	3	0.163	3	0.04	1	0.001	2	0.009	1	0.109
MUC6	7	3	5	0.797	2	0.02	1	0.002	2	0.012	1	0.181
ARAP3	5	3	2	0.197	0	0.00	0	0.003	2	0.047	3	0.508
BTAF1	5	3	1	0.083	1	0.02	1	0.016	2	0.033	6	1.000
GABRA1	5	3	1	0.083	0	0.00	0	0.003	2	0.033	0	0.057
KIAA1731	5	3	1	0.083	1	0.03	1	0.026	0	0.000	0	0.057
BAI1	5	3	1	0.143	0	0.01	1	0.031	0	0.000	0	0.100
BEND5	3	3	0	0.074	0	0.01	0	0.033	1	0.025	0	0.180
C10orf118	3	3	0	0.074	0	0.04	0	0.033	0	0.002	0	0.180
CCDC147	3	3	0	0.074	0	0.04	0	0.033	0	0.002	0	0.180
CCT8	3	3	0	0.074	0	0.04	0	0.033	0	0.002	1	0.667
CDC42EP4	3	3	0	0.074	0	0.04	0	0.008	0	0.002	1	0.667
FAM83E	3	3	1	0.244	0	0.04	0	0.033	0	0.010	1	0.667
FERD3L	3	3	1	0.244	0	0.04	0	0.033	0	0.010	0	0.180
GART	3	3	1	0.244	0	0.04	0	0.008	1	0.025	0	0.180
HOMER3	3	3	0	0.074	0	0.04	0	0.033	0	0.002	1	0.667
LMNA	3	3	1	0.244	0	0.04	0	0.008	1	0.025	1	0.667
LRRC36	3	3	0	0.074	0	0.01	0	0.033	0	0.010	3	1.000
MAP6	3	3	0	0.074	0	0.01	0	0.033	0	0.002	1	0.667
AADAT	3	3	1	0.398	0	0.02	0	0.021	0	0.028	1	1.000
ARFIP1	3	3	0	0.139	0	0.02	0	0.021	0	0.028	1	1.000
ATP6AP1	3	3	0	0.139	0	0.02	0	0.021	0	0.028	0	0.309
PRKAR1B	3	3	1	0.398	0	0.02	0	0.021	0	0.007	0	0.309
PNPLA2	3	3	0	0.139	0	0.02	0	0.021	0	0.007	0	0.309
OSMR	3	3	0	0.139	0	0.02	0	0.021	0	0.007	1	1.000

Gene											
CSMD1	0	3	0	1.000	4	0.01	8	0.000	6	0.002	10	0.001
DMD	0	3	0	1.000	6	0.00	3	0.047	4	0.008	8	0.003
DNAH17	0	3	0	1.000	5	0.00	3	0.026	4	0.013	4	0.032
DST	0	3	0	1.000	10	0.00	5	0.004	4	0.013	3	0.101
FRAS1	0	3	0	1.000	4	0.01	5	0.002	5	0.003	6	0.010
MUC16	0	3	0	1.000	23	0.00	12	0.000	16	0.000	30	0.000
MUC2	0	3	0	1.000	3	0.045	3	0.047	5	0.002	0	1.000
PCLO	0	3	0	1.000	15	0.000	9	0.000	11	0.000	15	0.000
NLRP8	5	3	3	0.363	1	0.03	1	0.016	1	0.001	0	0.057
PHACTR4	4	3	0	0.040	1	0.06	0	0.015	1	0.022	0	0.180
HMCN1	15	3	6	0.030	13	0.55	5	0.000	7	0.008	11	0.302
SYNE1	13	3	5	0.038	10	0.32	2	0.000	4	0.000	13	1.000
LRBA	8	3	1	0.009	5	0.29	1	0.002	1	0.000	4	0.397
KMT2A	11	3	3	0.015	5	0.07	2	0.001	3	0.001	6	0.316
CMYA5	7	3	0	0.003	3	0.08	1	0.003	2	0.009	1	0.109
RNF213	7	3	1	0.015	4	0.25	2	0.006	2	0.009	1	0.109
MGAM	7	3	1	0.027	6	0.83	2	0.030	2	0.018	4	0.558
LAMC3	6	3	1	0.047	2	0.09	2	0.038	2	0.017	0	0.061
CHD6	5	3	0	0.023	3	0.18	1	0.026	2	0.047	0	0.057
CCDC30	5	3	0	0.022	1	0.06	0	0.001	1	0.009	1	0.441
PCDHA6	5	3	0	0.022	1	0.06	0	0.006	1	0.047	0	0.100
ABHD17A	4	3	0	0.040	1	0.06	0	0.003	0	0.001	0	0.180
CD1C	4	3	0	0.040	1	0.06	0	0.015	0	0.003	0	0.180
CYP2A13	4	3	0	0.040	1	0.06	0	0.003	1	0.010	1	0.435
FAM184B	4	3	0	0.040	1	0.16	0	0.015	1	0.022	1	0.435
FANCD2	4	3	0	0.040	2	0.20	0	0.003	1	0.022	1	0.435
ZRSR2	4	3	0	0.040	1	0.16	0	0.015	0	0.003	1	0.435
CCDC146	4	3	0	0.040	0	0.00	0	0.003	1	0.071	4	1.000
NOTCH3	7	3	1	0.027	0	0.00	2	0.064	3	0.047	1	0.181
SLC23A1	4	3	0	0.040	0	0.00	1	0.059	0	0.003	0	0.180
OTOG	6	3	1	0.047	2	0.04	2	0.105	1	0.006	6	1.000
HCN1	5	3	0	0.023	0	0.00	2	0.067	2	0.033	0	0.057
MYH13	7	3	1	0.027	2	0.02	2	0.064	1	0.003	4	0.558
TSC2	9	3	1	0.003	3	0.01	5	0.088	3	0.008	7	0.798
FAR2	4	3	0	0.040	0	0.00	1	0.059	0	0.003	1	0.435
HEPACAM2	4	3	0	0.040	0	0.02	1	0.059	1	0.022	0	0.180
TTLL5	4	3	0	0.040	0	0.00	1	0.059	0	0.003	3	1.000

Supplementary Table 13. Substitution frequency in the Mongolian, Western, and Mongolian NCI cohorts. Differences between Mongolian vs Western and Mongolian NCI vs Western are also indicated.

Substitution	Substitution frequency Mongolian	Substitution frequency Western	Difference Mongolian vs Western	Substitution frequency Mongolian NCI	Difference Mongolian NCI vs Western
A[C>A]A	0.016	0.011	0.164	0.014	0.083
A[C>A]C	0.013	0.012	0.054	0.014	0.079
A[C>A]G	0.007	0.01	-0.174	0.006	-0.247
A[C>A]T	0.009	0.007	0.123	0.009	0.107
C[C>A]A	0.016	0.017	-0.041	0.018	0.016
C[C>A]C	0.017	0.015	0.043	0.018	0.078
C[C>A]G	0.01	0.012	-0.082	0.013	0.027
C[C>A]T	0.013	0.011	0.091	0.014	0.151
G[C>A]A	0.014	0.012	0.079	0.017	0.197
G[C>A]C	0.015	0.016	-0.035	0.017	0.031
G[C>A]G	0.012	0.013	-0.059	0.016	0.079
G[C>A]T	0.01	0.009	0.033	0.015	0.235
T[C>A]A	0.016	0.018	-0.063	0.017	-0.028
T[C>A]C	0.025	0.021	0.096	0.025	0.096
T[C>A]G	0.01	0.01	0.026	0.009	-0.023
T[C>A]T	0.018	0.017	0.03	0.019	0.076
A[C>G]A	0.005	0.005	-0.002	0.005	0.04
A[C>G]C	0.005	0.005	0.045	0.006	0.077
A[C>G]G	0.002	0.002	0.105	0.003	0.239
A[C>G]T	0.006	0.004	0.192	0.005	0.149
C[C>G]A	0.006	0.007	-0.044	0.005	-0.145
C[C>G]C	0.008	0.006	0.105	0.008	0.097
C[C>G]G	0.005	0.006	-0.078	0.005	-0.131
C[C>G]T	0.01	0.009	0.05	0.008	-0.022
G[C>G]A	0.003	0.004	-0.083	0.003	-0.158
G[C>G]C	0.006	0.006	0.021	0.005	-0.08
G[C>G]G	0.003	0.004	-0.108	0.002	-0.256
G[C>G]T	0.005	0.005	0.06	0.006	0.129
T[C>G]A	0.007	0.007	-0.006	0.006	-0.085
T[C>G]C	0.009	0.01	-0.016	0.009	-0.019
T[C>G]G	0.003	0.003	0.071	0.002	-0.137
T[C>G]T	0.012	0.014	-0.069	0.012	-0.057

A[C>T]A	0.012	0.013	-0.074	0.01	-0.13
A[C>T]C	0.01	0.011	-0.043	0.011	0.008
A[C>T]G	0.019	0.025	-0.135	0.028	0.05
A[C>T]T	0.007	0.009	-0.111	0.008	-0.058
C[C>T]A	0.019	0.024	-0.123	0.02	-0.092
C[C>T]C	0.018	0.022	-0.097	0.02	-0.056
C[C>T]G	0.023	0.032	-0.162	0.033	0.01
C[C>T]T	0.018	0.023	-0.119	0.017	-0.148
G[C>T]A	0.018	0.018	-0.017	0.021	0.08
G[C>T]C	0.022	0.021	0.022	0.031	0.197
G[C>T]G	0.027	0.035	-0.118	0.045	0.132
G[C>T]T	0.016	0.018	-0.063	0.017	-0.036
T[C>T]A	0.02	0.019	0.043	0.021	0.055
T[C>T]C	0.018	0.024	-0.123	0.021	-0.061
T[C>T]G	0.013	0.019	-0.176	0.019	0.006
T[C>T]T	0.013	0.017	-0.164	0.013	-0.145
A[T>A]A	0.004	0.003	0.049	0.002	-0.19
A[T>A]C	0.005	0.006	-0.083	0.005	-0.068
A[T>A]G	0.01	0.01	0.006	0.006	-0.198
A[T>A]T	0.004	0.004	-0.038	0.003	-0.168
C[T>A]A	0.007	0.005	0.163	0.004	-0.073
C[T>A]C	0.016	0.012	0.141	0.012	0.019
C[T>A]G	0.037	0.026	0.179	0.024	-0.036
C[T>A]T	0.01	0.008	0.086	0.009	0.017
G[T>A]A	0.004	0.003	0.213	0.003	0.081
G[T>A]C	0.005	0.005	0.05	0.006	0.125
G[T>A]G	0.009	0.008	0.012	0.007	-0.102
G[T>A]T	0.004	0.003	0.085	0.002	-0.207
T[T>A]A	0.004	0.004	0.075	0.003	-0.077
T[T>A]C	0.008	0.007	0.057	0.006	-0.087
T[T>A]G	0.009	0.009	-0.009	0.006	-0.195
T[T>A]T	0.005	0.005	-0.015	0.004	-0.164
A[T>C]A	0.016	0.017	-0.049	0.015	-0.086
A[T>C]C	0.007	0.009	-0.113	0.007	-0.164
A[T>C]G	0.024	0.025	-0.014	0.022	-0.056
A[T>C]T	0.011	0.013	-0.07	0.011	-0.089
C[T>C]A	0.006	0.006	-0.046	0.005	-0.101
C[T>C]C	0.012	0.011	0.055	0.011	-0.011
C[T>C]G	0.018	0.018	0.013	0.016	-0.04

C[T>C]T	0.01	0.009	0.017	0.009	0.005
G[T>C]A	0.01	0.01	-0.001	0.008	-0.122
G[T>C]C	0.009	0.009	-0.023	0.008	-0.08
G[T>C]G	0.012	0.012	0.019	0.01	-0.08
G[T>C]T	0.01	0.01	0.037	0.009	-0.036
T[T>C]A	0.006	0.007	-0.067	0.006	-0.049
T[T>C]C	0.007	0.008	-0.108	0.007	-0.068
T[T>C]G	0.01	0.009	0.048	0.01	0.045
T[T>C]T	0.006	0.006	0.038	0.006	-0.023
A[T>G]A	0.003	0.002	0.162	0.002	-0.19
A[T>G]C	0.004	0.002	0.282	0.003	0.137
A[T>G]G	0.012	0.004	0.505	0.005	0.123
A[T>G]T	0.006	0.003	0.272	0.005	0.234
C[T>G]A	0.002	0.002	0.103	0.002	0.126
C[T>G]C	0.005	0.006	-0.111	0.006	-0.053
C[T>G]G	0.011	0.007	0.202	0.009	0.085
C[T>G]T	0.011	0.01	0.019	0.009	-0.098
G[T>G]A	0.002	0.001	0.323	0.001	-0.062
G[T>G]C	0.003	0.002	0.144	0.003	0.14
G[T>G]G	0.007	0.002	0.464	0.004	0.23
G[T>G]T	0.004	0.004	0.012	0.004	0.013
T[T>G]A	0.004	0.002	0.327	0.002	-0.023
T[T>G]C	0.006	0.005	0.071	0.005	-0.005
T[T>G]G	0.009	0.004	0.361	0.005	0.092
T[T>G]T	0.007	0.007	-0.013	0.005	-0.14

Supplementary Table 14. Assignment of the four de novo extracted HCC signatures to all single and linear combinations of two COSMIC v3 signatures. Signature mapping highlighted in yellow indicated the assignment used in our analysis, selected based on cosine similarity (Cos sim).

de novo signature 1			*de novo* signature 2			*de novo* signature 3			*de novo* signature 4 (SBSM)		
Reference signature	Weight	Cos sim	Reference signature	Weight	Cos sim	Reference signature	Weight	Cos sim	Reference signature	Weight	Cos sim
SBS22	SBS22=1	0.9735	SBS6 + SBS40	SBS6=0.12; SBS40=0.88	0.91406	SBS16 + SBS26	SBS16=0.45; SBS26=0.55	0.92379	SBS28 + SBS40	SBS28=0.09; SBS40=0.91	0.81766
SBS3 + SBS22	SBS3=0.12; SBS22=0.88	0.97496	SBS15 + SBS40	SBS15=0.09; SBS40=0.91	0.90401	SBS12 + SBS16	SBS12=0.55; SBS16=0.45	0.91218	SBS17b + SBS40	SBS17b=0.07; SBS40=0.93	0.81288
SBS22 + SBS25	SBS22=0.86; SBS25=0.14	0.97488	SBS30 + SBS40	SBS30=0.13; SBS40=0.87	0.90294	SBS16 + SBS54	SBS16=0.67; SBS54=0.33	0.89886	SBS9 + SBS40	SBS9=0.3; SBS40=0.7	0.80978
SBS8 + SBS22	SBS8=0.09; SBS22=0.91	0.97469	SBS23 + SBS40	SBS23=0.1; SBS40=0.9	0.9012	SBS16 + SBS46	SBS16=0.59; SBS46=0.41	0.88026	SBS40 + SBS55	SBS40=0.9; SBS55=0.1	0.8047
SBS4 + SBS22	SBS4=0.07; SBS22=0.93	0.97468	SBS19 + SBS40	SBS19=0.09; SBS40=0.91	0.8984	SBS16 + SBS33	SBS16=0.76; SBS33=0.24	0.87846	SBS40 + SBS43	SBS40=0.93; SBS43=0.07	0.78171
SBS22 + SBS40	SBS22=0.9; SBS40=0.1	0.97462	SBS40 + SBS42	SBS40=0.85; SBS42=0.15	0.89789	SBS16 + SBS21	SBS16=0.74; SBS21=0.26	0.86621	SBS40 + SBS60	SBS40=0.98; SBS60=0.02	0.77619
SBS22 + SBS46	SBS22=0.95; SBS46=0.05	0.97415	SBS40 + SBS84	SBS40=0.89; SBS84=0.11	0.89751	SBS5 + SBS16	SBS5=0.59; SBS16=0.41	0.86563	SBS40 + SBS54	SBS40=0.97; SBS54=0.03	0.77316
SBS18 + SBS22	SBS18=0.04; SBS22=0.96	0.97412	SBS32 + SBS40	SBS32=0.11; SBS40=0.89	0.89374	SBS5 + SBS26	SBS5=0.55; SBS26=0.45	0.85408	SBS40 + SBS41	SBS40=0.93; SBS41=0.07	0.77314
SBS22 + SBS45	SBS22=0.97; SBS45=0.03	0.97409	SBS1 + SBS40	SBS1=0.05; SBS40=0.95	0.89348	SBS16 + SBS25	SBS16=0.56; SBS25=0.44	0.85369	SBS37 + SBS40	SBS37=0.06; SBS40=0.94	0.77238
SBS22 + SBS39	SBS22=0.94; SBS39=0.06	0.97407	SBS40 + SBS44	SBS40=0.88; SBS44=0.12	0.89218	SBS16 + SBS37	SBS16=0.57; SBS37=0.43	0.85156	SBS40 + SBS51	SBS40=0.97; SBS51=0.03	0.77162
SBS22 + SBS36	SBS22=0.97; SBS36=0.03	0.97406	SBS24 + SBS40	SBS24=0.13; SBS40=0.87	0.89124	SBS3 + SBS16	SBS3=0.46; SBS16=0.54	0.84754	SBS40 + SBS57	SBS40=0.98; SBS57=0.02	0.77099
SBS22 + SBS38	SBS22=0.98; SBS38=0.02	0.97392	SBS11 + SBS40	SBS11=0.08; SBS40=0.92	0.89086	SBS5 + SBS12	SBS5=0.55; SBS12=0.45	0.84635	SBS1 + SBS40	SBS1=0; SBS40=1	0.7703
SBS17a + SBS22	SBS17a=0.02; SBS22=0.98	0.97391	SBS7b + SBS40	SBS7b=0.07; SBS40=0.93	0.88954	SBS16 + SBS44	SBS16=0.74; SBS44=0.26	0.84225	SBS2 + SBS40	SBS2=0; SBS40=1	0.7703
SBS22 + SBS35	SBS22=0.96; SBS35=0.04	0.97391	SBS7a + SBS40	SBS7a=0.05; SBS40=0.95	0.88584	SBS16 + SBS40	SBS16=0.62; SBS40=0.38	0.83552	SBS3 + SBS40	SBS3=0; SBS40=1	0.7703
SBS22 + SBS56	SBS22=0.98; SBS56=0.02	0.9739	SBS29 + SBS40	SBS29=0.1; SBS40=0.9	0.88549	SBS16 + SBS17a	SBS16=0.89; SBS17a=0.11	0.83266	SBS4 + SBS40	SBS4=0; SBS40=1	0.7703
SBS5 + SBS22	SBS5=0.06; SBS22=0.94	0.97389	SBS5 + SBS40	SBS5=0.22; SBS40=0.78	0.88366	SBS26 + SBS40	SBS26=0.62; SBS40=0.38	0.83076	SBS5 + SBS40	SBS5=0; SBS40=1	0.7703
SBS22 + SBS54	SBS22=0.98; SBS54=0.02	0.97388	SBS31 + SBS40	SBS31=0.07; SBS40=0.93	0.88193	SBS9 + SBS16	SBS9=0.27; SBS16=0.73	0.83046	SBS6 + SBS40	SBS6=0; SBS40=1	0.7703
SBS22 + SBS29	SBS22=0.97; SBS29=0.03	0.97387	SBS18 + SBS40	SBS18=0.07; SBS40=0.93	0.88079	SBS8 + SBS26	SBS8=0.28; SBS26=0.72	0.8304	SBS7a + SBS40	SBS7a=0; SBS40=1	0.7703
SBS22 + SBS24	SBS22=0.97; SBS24=0.03	0.97383	SBS20 + SBS40	SBS20=0.05; SBS40=0.95	0.88025	SBS4 + SBS16	SBS4=0.23; SBS16=0.77	0.8302	SBS7b + SBS40	SBS7b=0; SBS40=1	0.7703
SBS22 + SBS53	SBS22=0.98; SBS53=0.02	0.97381	SBS2 + SBS40	SBS2=0.02; SBS40=0.98	0.88024	SBS26 + SBS58	SBS26=0.78; SBS58=0.22	0.82948	SBS7c + SBS40	SBS7c=0; SBS40=1	0.7703
SBS13 + SBS22	SBS13=0.01; SBS22=0.99	0.97381	SBS21 + SBS40	SBS21=0.04; SBS40=0.96	0.8797	SBS7d + SBS16	SBS7d=0.13; SBS16=0.87	0.82906	SBS7d + SBS40	SBS7d=0; SBS40=1	0.7703
SBS20 + SBS22	SBS20=0.02; SBS22=0.98	0.97378	SBS36 + SBS40	SBS36=0.04; SBS40=0.96	0.87903	SBS3 + SBS26	SBS3=0.38; SBS26=0.62	0.82839	SBS8 + SBS40	SBS8=0; SBS40=1	0.7703
SBS10a + SBS22	SBS10a=0.01; SBS22=0.99	0.97378	SBS4 + SBS40	SBS4=0.07; SBS40=0.93	0.87889	SBS8 + SBS16	SBS8=0.24; SBS16=0.76	0.82801	SBS10a + SBS40	SBS10a=0; SBS40=1	0.7703
SBS9 + SBS22	SBS9=0.04; SBS22=0.96	0.97375	SBS40 + SBS52	SBS40=0.98; SBS52=0.02	0.87874	SBS16 + SBS20	SBS16=0.86; SBS20=0.14	0.82791	SBS10b + SBS40	SBS10b=0; SBS40=1	0.7703
SBS22 + SBS57	SBS22=0.98; SBS57=0.02	0.97371	SBS7d + SBS40	SBS7d=0.03; SBS40=0.97	0.87816	SBS4 + SBS26	SBS4=0.24; SBS26=0.76	0.82751	SBS11 + SBS40	SBS11=0; SBS40=1	0.7703
SBS22 + SBS52	SBS22=0.99; SBS52=0.01	0.97369	SBS40 + SBS50	SBS40=0.96; SBS50=0.04	0.87811	SBS16 + SBS42	SBS16=0.82; SBS42=0.18	0.82731	SBS12 + SBS40	SBS12=0; SBS40=1	0.7703
SBS22 + SBS33	SBS22=0.99; SBS33=0.01	0.97369	SBS35 + SBS40	SBS35=0.05; SBS40=0.95	0.87799	SBS16 + SBS35	SBS16=0.81; SBS35=0.19	0.82729	SBS13 + SBS40	SBS13=0; SBS40=1	0.7703
SBS22 + SBS44	SBS22=0.98; SBS44=0.02	0.97368	SBS14 + SBS40	SBS14=0.03; SBS40=0.97	0.87768	SBS25 + SBS26	SBS25=0.33; SBS26=0.67	0.82666	SBS14 + SBS40	SBS14=0; SBS40=1	0.7703
SBS22 + SBS41	SBS22=0.97; SBS41=0.03	0.97368	SBS10b + SBS40	SBS10b=0.02; SBS40=0.98	0.87748	SBS16 + SBS18	SBS16=0.84; SBS18=0.16	0.82605	SBS15 + SBS40	SBS15=0; SBS40=1	0.7703
SBS22 + SBS55	SBS22=0.99; SBS55=0.01	0.97366	SBS40 + SBS46	SBS40=0.96; SBS46=0.04	0.87704	SBS14 + SBS16	SBS14=0.11; SBS16=0.89	0.82554	SBS16 + SBS40	SBS16=0; SBS40=1	0.7703
SBS22 + SBS49	SBS22=0.99; SBS49=0.01	0.97365	SBS40 + SBS45	SBS40=0.98; SBS45=0.02	0.87686	SBS16 + SBS24	SBS16=0.83; SBS24=0.17	0.82548	SBS17a + SBS40	SBS17a=0; SBS40=1	0.7703
SBS22 + SBS31	SBS22=0.98; SBS31=0.02	0.97365	SBS40 + SBS48	SBS40=0.99; SBS48=0.01	0.87654	SBS16 + SBS31	SBS16=0.85; SBS31=0.15	0.82525	SBS18 + SBS40	SBS18=0; SBS40=1	0.7703

SBS14 + SBS22	SBS14=0.01; SBS22=0.99	0.97364	SBS40 + SBS49	SBS40=0.99; SBS49=0.01	0.87642	SBS16 + SBS29	SBS16=0.85; SBS29=0.15	0.82461	SBS19 + SBS40	SBS19=0; SBS40=1	0.7703
SBS22 + SBS42	SBS22=0.98; SBS42=0.02	0.97363	SBS40 + SBS59	SBS40=0.99; SBS59=0.01	0.8761	SBS16 + SBS36	SBS16=0.88; SBS36=0.12	0.8245	SBS20 + SBS40	SBS20=0; SBS40=1	0.7703
SBS22 + SBS26	SBS22=0.98; SBS26=0.02	0.97362	SBS40 + SBS56	SBS40=0.99; SBS56=0.01	0.87605	SBS16 + SBS43	SBS16=0.89; SBS43=0.11	0.82397	SBS21 + SBS40	SBS21=0; SBS40=1	0.7703
SBS7d + SBS22	SBS7d=0.01; SBS22=0.99	0.97362	SBS40 + SBS53	SBS40=0.99; SBS53=0.01	0.87588	SBS16 + SBS57	SBS16=0.87; SBS57=0.13	0.82355	SBS22 + SBS40	SBS22=0; SBS40=1	0.7703
SBS21 + SBS22	SBS21=0.01; SBS22=0.99	0.9736	SBS3 + SBS40	SBS3=0.05; SBS40=0.95	0.87583	SBS16 + SBS41	SBS16=0.84; SBS41=0.16	0.82332	SBS23 + SBS40	SBS23=0; SBS40=1	0.7703
SBS22 + SBS30	SBS22=0.99; SBS30=0.01	0.97359	SBS38 + SBS40	SBS38=0.01; SBS40=0.99	0.87583	SBS16 + SBS22	SBS16=0.89; SBS22=0.11	0.82315	SBS24 + SBS40	SBS24=0; SBS40=1	0.7703
SBS22 + SBS28	SBS22=0.99; SBS28=0.01	0.97359	SBS25 + SBS40	SBS25=0.01; SBS40=0.99	0.87559	SBS16 + SBS58	SBS16=0.87; SBS58=0.13	0.82201	SBS25 + SBS40	SBS25=0; SBS40=1	0.7703
SBS7b + SBS22	SBS7b=0.01; SBS22=0.99	0.97358	SBS33 + SBS40	SBS33=0.01; SBS40=0.99	0.87558	SBS7c + SBS16	SBS7c=0.09; SBS16=0.91	0.82177	SBS26 + SBS40	SBS26=0; SBS40=1	0.7703
SBS15 + SBS22	SBS15=0.01; SBS22=0.99	0.97358	SBS7c + SBS40	SBS7c=0; SBS40=1	0.87558	SBS16 + SBS39	SBS16=0.82; SBS39=0.18	0.82168	SBS27 + SBS40	SBS27=0; SBS40=1	0.7703
SBS22 + SBS37	SBS22=0.98; SBS37=0.02	0.97358	SBS8 + SBS40	SBS8=0; SBS40=1	0.87558	SBS16 + SBS45	SBS16=0.93; SBS45=0.07	0.82148	SBS29 + SBS40	SBS29=0; SBS40=1	0.7703
SBS22 + SBS50	SBS22=0.99; SBS50=0.01	0.97357	SBS9 + SBS40	SBS9=0; SBS40=1	0.87558	SBS16 + SBS50	SBS16=0.9; SBS50=0.1	0.82125	SBS30 + SBS40	SBS30=0; SBS40=1	0.7703
SBS22 + SBS23	SBS22=0.99; SBS23=0.01	0.97356	SBS10a + SBS40	SBS10a=0; SBS40=1	0.87558	SBS16 + SBS17b	SBS16=0.95; SBS17b=0.05	0.82121	SBS31 + SBS40	SBS31=0; SBS40=1	0.7703
SBS12 + SBS22	SBS12=0.02; SBS22=0.98	0.97356	SBS12 + SBS40	SBS12=0; SBS40=1	0.87558	SBS12 + SBS58	SBS12=0.79; SBS58=0.21	0.82118	SBS32 + SBS40	SBS32=0; SBS40=1	0.7703
SBS22 + SBS32	SBS22=0.99; SBS32=0.01	0.97355	SBS13 + SBS40	SBS13=0; SBS40=1	0.87558	SBS26 + SBS35	SBS26=0.8; SBS35=0.2	0.82115	SBS33 + SBS40	SBS33=0; SBS40=1	0.7703
SBS6 + SBS22	SBS6=0.01; SBS22=0.99	0.97353	SBS16 + SBS40	SBS16=0; SBS40=1	0.87558	SBS16 + SBS56	SBS16=0.95; SBS56=0.05	0.82099	SBS34 + SBS40	SBS34=0; SBS40=1	0.7703
SBS22 + SBS84	SBS22=0.99; SBS84=0.01	0.97353	SBS17a + SBS40	SBS17a=0; SBS40=1	0.87558	SBS10a + SBS16	SBS10a=0.04; SBS16=0.96	0.82091	SBS35 + SBS40	SBS35=0; SBS40=1	0.7703
SBS19 + SBS22	SBS19=0.01; SBS22=0.99	0.97353	SBS17b + SBS40	SBS17b=0; SBS40=1	0.87558	SBS16 + SBS53	SBS16=0.93; SBS53=0.07	0.82086	SBS36 + SBS40	SBS36=0; SBS40=1	0.7703
SBS22 + SBS43	SBS22=0.99; SBS43=0.01	0.97353	SBS22 + SBS40	SBS22=0; SBS40=1	0.87558	SBS7b + SBS16	SBS7b=0.07; SBS16=0.93	0.82053	SBS38 + SBS40	SBS38=0; SBS40=1	0.7703
SBS11 + SBS22	SBS11=0.01; SBS22=0.99	0.9735	SBS26 + SBS40	SBS26=0; SBS40=1	0.87558	SBS16 + SBS38	SBS16=0.95; SBS38=0.05	0.8205	SBS39 + SBS40	SBS39=0; SBS40=1	0.7703
SBS1 + SBS22	SBS1=0; SBS22=1	0.9735	SBS27 + SBS40	SBS27=0; SBS40=1	0.87558	SBS16 + SBS30	SBS16=0.92; SBS30=0.08	0.82029	SBS40 + SBS42	SBS40=1; SBS42=0	0.7703
SBS2 + SBS22	SBS2=0; SBS22=1	0.9735	SBS28 + SBS40	SBS28=0; SBS40=1	0.87558	SBS16 + SBS51	SBS16=0.91; SBS51=0.09	0.82005	SBS40 + SBS44	SBS40=1; SBS44=0	0.7703
SBS7a + SBS22	SBS7a=0; SBS22=1	0.9735	SBS34 + SBS40	SBS34=0; SBS40=1	0.87558	SBS15 + SBS16	SBS15=0.06; SBS16=0.94	0.81994	SBS40 + SBS45	SBS40=1; SBS45=0	0.7703
SBS7c + SBS22	SBS7c=0; SBS22=1	0.9735	SBS37 + SBS40	SBS37=0; SBS40=1	0.87558	SBS16 + SBS23	SBS16=0.94; SBS23=0.06	0.81989	SBS40 + SBS46	SBS40=1; SBS46=0	0.7703
SBS10b + SBS22	SBS10b=0; SBS22=1	0.9735	SBS39 + SBS40	SBS39=0; SBS40=1	0.87558	SBS16 + SBS32	SBS16=0.92; SBS32=0.08	0.81977	SBS40 + SBS47	SBS40=1; SBS47=0	0.7703
SBS16 + SBS22	SBS16=0; SBS22=1	0.9735	SBS40 + SBS41	SBS40=1; SBS41=0	0.87558	SBS18 + SBS26	SBS18=0.17; SBS26=0.83	0.81968	SBS40 + SBS48	SBS40=1; SBS48=0	0.7703
SBS17b + SBS22	SBS17b=0; SBS22=1	0.9735	SBS40 + SBS43	SBS40=1; SBS43=0	0.87558	SBS11 + SBS16	SBS11=0.05; SBS16=0.95	0.81942	SBS40 + SBS49	SBS40=1; SBS49=0	0.7703
SBS22 + SBS27	SBS22=1; SBS27=0	0.9735	SBS40 + SBS47	SBS40=1; SBS47=0	0.87558	SBS12 + SBS40	SBS12=0.65; SBS40=0.35	0.81938	SBS40 + SBS50	SBS40=1; SBS50=0	0.7703
SBS22 + SBS34	SBS22=1; SBS34=0	0.9735	SBS40 + SBS51	SBS40=1; SBS51=0	0.87558	SBS6 + SBS16	SBS6=0.05; SBS16=0.95	0.81914	SBS40 + SBS52	SBS40=1; SBS52=0	0.7703
SBS22 + SBS47	SBS22=1; SBS47=0	0.9735	SBS40 + SBS54	SBS40=1; SBS54=0	0.87558	SBS16 + SBS84	SBS16=0.94; SBS84=0.06	0.81912	SBS40 + SBS53	SBS40=1; SBS53=0	0.7703
SBS22 + SBS48	SBS22=1; SBS48=0	0.9735	SBS40 + SBS55	SBS40=1; SBS55=0	0.87558	SBS12 + SBS25	SBS12=0.67; SBS25=0.33	0.8191	SBS40 + SBS56	SBS40=1; SBS56=0	0.7703
SBS22 + SBS51	SBS22=1; SBS51=0	0.9735	SBS40 + SBS57	SBS40=1; SBS57=0	0.87558	SBS16 + SBS55	SBS16=0.96; SBS55=0.04	0.81907	SBS40 + SBS58	SBS40=1; SBS58=0	0.7703
SBS22 + SBS58	SBS22=1; SBS58=0	0.9735	SBS40 + SBS58	SBS40=1; SBS58=0	0.87558	SBS16 + SBS85	SBS16=0.93; SBS85=0.07	0.81886	SBS40 + SBS59	SBS40=1; SBS59=0	0.7703
SBS22 + SBS59	SBS22=1; SBS59=0	0.9735	SBS40 + SBS60	SBS40=1; SBS60=0	0.87558	SBS16 + SBS19	SBS16=0.96; SBS19=0.04	0.81872	SBS40 + SBS84	SBS40=1; SBS84=0	0.7703
SBS22 + SBS60	SBS22=1; SBS60=0	0.9735	SBS40 + SBS85	SBS40=1; SBS85=0	0.87558	SBS16 + SBS59	SBS16=0.97; SBS59=0.03	0.81868	SBS40 + SBS85	SBS40=1; SBS85=0	0.7703
SBS22 + SBS85	SBS22=1; SBS85=0	0.9735	SBS40	SBS40=1	0.87558	SBS7a + SBS16	SBS7a=0.03; SBS16=0.97	0.81845	SBS40	SBS40=1	0.7703
SBS25 + SBS27	SBS25=0.9; SBS27=0.1	0.7411	SBS3 + SBS30	SBS3=0.8; SBS30=0.2	0.86242	SBS16 + SBS52	SBS16=0.98; SBS52=0.02	0.81845	SBS3 + SBS9	SBS3=0.51; SBS9=0.49	0.74308
SBS25 + SBS34	SBS25=0.96; SBS34=0.04	0.72535	SBS5 + SBS29	SBS5=0.77; SBS29=0.23	0.85806	SBS22 + SBS26	SBS22=0.14; SBS26=0.86	0.81825	SBS3 + SBS28	SBS3=0.86; SBS28=0.14	0.74092
SBS25 + SBS35	SBS25=0.91; SBS35=0.09	0.72431	SBS5 + SBS18	SBS5=0.77; SBS18=0.23	0.85756	SBS10b + SBS16	SBS10b=0.02; SBS16=0.98	0.81815	SBS4 + SBS9	SBS4=0.27; SBS9=0.73	0.72368
SBS8 + SBS25	SBS8=0.04; SBS25=0.96	0.72021	SBS4 + SBS5	SBS4=0.27; SBS5=0.73	0.85722	SBS16 + SBS49	SBS16=0.98; SBS49=0.02	0.81808	SBS9 + SBS55	SBS9=0.85; SBS55=0.15	0.72333

Pair	Values	Score	Pair	Values	Score	Pair	Values	Score	Pair	Values	Score
SBS25 + SBS47	SBS25=0.98; SBS47=0.02	0.72016	SBS5 + SBS36	SBS5=0.84; SBS36=0.16	0.85037	SBS16 + SBS60	SBS16=0.99; SBS60=0.01	0.81805	SBS9 + SBS39	SBS9=0.67; SBS39=0.33	0.71522
SBS1 + SBS25	SBS1=0; SBS25=1	0.71986	SBS5 + SBS24	SBS5=0.76; SBS24=0.24	0.84979	SBS4 + SBS12	SBS4=0.23; SBS12=0.77	0.8179	SBS9 + SBS18	SBS9=0.78; SBS18=0.22	0.71505
SBS2 + SBS25	SBS2=0; SBS25=1	0.71986	SBS3 + SBS7a	SBS3=0.89; SBS7a=0.11	0.84274	SBS16 + SBS47	SBS16=0.98; SBS47=0.02	0.81789	SBS8 + SBS9	SBS8=0.29; SBS9=0.71	0.71393
SBS3 + SBS25	SBS3=0; SBS25=1	0.71986	SBS2 + SBS3	SBS2=0.07; SBS3=0.93	0.83827	SBS1 + SBS16	SBS1=0; SBS16=1	0.81776	SBS9 + SBS29	SBS9=0.79; SBS29=0.21	0.71238
SBS4 + SBS25	SBS4=0; SBS25=1	0.71986	SBS3 + SBS32	SBS3=0.82; SBS32=0.18	0.83301	SBS2 + SBS16	SBS2=0; SBS16=1	0.81776	SBS5 + SBS9	SBS5=0.4; SBS9=0.6	0.71213
SBS5 + SBS25	SBS5=0; SBS25=1	0.71986	SBS3 + SBS5	SBS3=0.49; SBS5=0.51	0.83275	SBS13 + SBS16	SBS13=0; SBS16=1	0.81776	SBS9 + SBS50	SBS9=0.82; SBS50=0.18	0.71142
SBS6 + SBS25	SBS6=0; SBS25=1	0.71986	SBS5 + SBS52	SBS5=0.92; SBS52=0.08	0.83155	SBS16 + SBS27	SBS16=1; SBS27=0	0.81776	SBS9 + SBS51	SBS9=0.8; SBS51=0.2	0.71029
SBS7a + SBS25	SBS7a=0; SBS25=1	0.71986	SBS3 + SBS6	SBS3=0.87; SBS6=0.13	0.83149	SBS16 + SBS28	SBS16=1; SBS28=0	0.81776	SBS9 + SBS24	SBS9=0.79; SBS24=0.21	0.70762
SBS7b + SBS25	SBS7b=0; SBS25=1	0.71986	SBS5 + SBS8	SBS5=0.75; SBS8=0.25	0.83072	SBS16 + SBS34	SBS16=1; SBS34=0	0.81776	SBS3 + SBS17b	SBS3=0.9; SBS17b=0.1	0.7051
SBS7c + SBS25	SBS7c=0; SBS25=1	0.71986	SBS5 + SBS30	SBS5=0.83; SBS30=0.17	0.82942	SBS16 + SBS48	SBS16=1; SBS48=0	0.81776	SBS9 + SBS36	SBS9=0.85; SBS36=0.15	0.70343
SBS7d + SBS25	SBS7d=0; SBS25=1	0.71986	SBS5 + SBS45	SBS5=0.9; SBS45=0.1	0.82939	SBS16	SBS16=1	0.81776	SBS5 + SBS28	SBS5=0.83; SBS28=0.17	0.70299
SBS9 + SBS25	SBS9=0; SBS25=1	0.71986	SBS3 + SBS19	SBS3=0.87; SBS19=0.13	0.82841	SBS26 + SBS36	SBS26=0.87; SBS36=0.13	0.81735	SBS2 + SBS9	SBS2=0.07; SBS9=0.93	0.70113
SBS10a + SBS25	SBS10a=0; SBS25=1	0.71986	SBS3 + SBS18	SBS3=0.81; SBS18=0.19	0.82689	SBS12 + SBS54	SBS12=0.82; SBS54=0.18	0.81686	SBS9 + SBS30	SBS9=0.86; SBS30=0.14	0.6986
SBS10b + SBS25	SBS10b=0; SBS25=1	0.71986	SBS5 + SBS35	SBS5=0.81; SBS35=0.19	0.82575	SBS8 + SBS12	SBS8=0.25; SBS12=0.75	0.81658	SBS9 + SBS25	SBS9=0.73; SBS25=0.27	0.69554
SBS11 + SBS25	SBS11=0; SBS25=1	0.71986	SBS5 + SBS42	SBS5=0.8; SBS42=0.2	0.82529	SBS26 + SBS29	SBS26=0.84; SBS29=0.16	0.81623	SBS9 + SBS42	SBS9=0.84; SBS42=0.16	0.69521
SBS12 + SBS25	SBS12=0; SBS25=1	0.71986	SBS5 + SBS50	SBS5=0.85; SBS50=0.15	0.82463	SBS3 + SBS12	SBS3=0.35; SBS12=0.65	0.81582	SBS9 + SBS35	SBS9=0.82; SBS35=0.18	0.69516
SBS13 + SBS25	SBS13=0; SBS25=1	0.71986	SBS5 + SBS19	SBS5=0.88; SBS19=0.12	0.82431	SBS24 + SBS26	SBS24=0.17; SBS26=0.83	0.81564	SBS9 + SBS32	SBS9=0.86; SBS32=0.14	0.69422
SBS14 + SBS25	SBS14=0; SBS25=1	0.71986	SBS3 + SBS11	SBS3=0.88; SBS11=0.12	0.82371	SBS26 + SBS42	SBS26=0.84; SBS42=0.16	0.81558	SBS9 + SBS45	SBS9=0.9; SBS45=0.1	0.69358
SBS15 + SBS25	SBS15=0; SBS25=1	0.71986	SBS3 + SBS36	SBS3=0.86; SBS36=0.14	0.82323	SBS26 + SBS31	SBS26=0.85; SBS31=0.15	0.81503	SBS7a + SBS9	SBS7a=0.08; SBS9=0.92	0.69292
SBS16 + SBS25	SBS16=0; SBS25=1	0.71986	SBS3 + SBS23	SBS3=0.88; SBS23=0.12	0.82169	SBS12 + SBS35	SBS12=0.8; SBS35=0.2	0.815	SBS9 + SBS43	SBS9=0.89; SBS43=0.11	0.69124
SBS17a + SBS25	SBS17a=0; SBS25=1	0.71986	SBS8 + SBS30	SBS8=0.61; SBS30=0.39	0.82138	SBS12 + SBS26	SBS12=0.45; SBS26=0.55	0.81435	SBS9 + SBS13	SBS9=0.93; SBS13=0.07	0.69049
SBS17b + SBS25	SBS17b=0; SBS25=1	0.71986	SBS4 + SBS30	SBS4=0.59; SBS30=0.41	0.82111	SBS26 + SBS56	SBS26=0.93; SBS56=0.07	0.81104	SBS9 + SBS28	SBS9=0.92; SBS28=0.08	0.68905
SBS18 + SBS25	SBS18=0; SBS25=1	0.71986	SBS3 + SBS15	SBS3=0.89; SBS15=0.11	0.82079	SBS12 + SBS18	SBS12=0.84; SBS18=0.16	0.8105	SBS9 + SBS38	SBS9=0.94; SBS38=0.06	0.68465
SBS19 + SBS25	SBS19=0; SBS25=1	0.71986	SBS3 + SBS7b	SBS3=0.89; SBS7b=0.11	0.82037	SBS12 + SBS42	SBS12=0.84; SBS42=0.16	0.80975	SBS9 + SBS52	SBS9=0.95; SBS52=0.05	0.68463
SBS20 + SBS25	SBS20=0; SBS25=1	0.71986	SBS3 + SBS42	SBS3=0.8; SBS42=0.2	0.82002	SBS26 + SBS39	SBS26=0.82; SBS39=0.18	0.80972	SBS9 + SBS58	SBS9=0.89; SBS58=0.11	0.68384
SBS21 + SBS25	SBS21=0; SBS25=1	0.71986	SBS3 + SBS24	SBS3=0.8; SBS24=0.2	0.81946	SBS26 + SBS50	SBS26=0.89; SBS50=0.11	0.80971	SBS9 + SBS54	SBS9=0.93; SBS54=0.07	0.68267
SBS23 + SBS25	SBS23=0; SBS25=1	0.71986	SBS5 + SBS7a	SBS5=0.92; SBS7a=0.08	0.81882	SBS26 + SBS45	SBS26=0.93; SBS45=0.07	0.80947	SBS9 + SBS53	SBS9=0.94; SBS53=0.06	0.68235
SBS24 + SBS25	SBS24=0; SBS25=1	0.71986	SBS29 + SBS30	SBS29=0.53; SBS30=0.47	0.81861	SBS14 + SBS26	SBS14=0.08; SBS26=0.92	0.80904	SBS9 + SBS60	SBS9=0.97; SBS60=0.03	0.6821
SBS25 + SBS26	SBS25=1; SBS26=0	0.71986	SBS5 + SBS56	SBS5=0.93; SBS56=0.07	0.81833	SBS12 + SBS22	SBS12=0.87; SBS22=0.13	0.80897	SBS9 + SBS11	SBS9=0.94; SBS11=0.06	0.68133

Supplementary Table 15. The 96 trinucleotide frequency for SBS Mongolia.

trinucleotide change	Relative contribution (SBS Mongolia)
A[C>A]A	0.035955461
A[C>A]C	0.013847095
A[C>A]G	0.001721218
A[C>A]T	0.004252817
C[C>A]A	0.014587099
C[C>A]C	0.012362044
C[C>A]G	0.00118355
C[C>A]T	0.003425482
G[C>A]A	0.014750146
G[C>A]C	0.004370825
G[C>A]G	0.001760215
G[C>A]T	0.008717436
T[C>A]A	0.01448695
T[C>A]C	0.035744186
T[C>A]G	0.005123383
T[C>A]T	0.018882905
A[C>G]A	0.006779861
A[C>G]C	0.002203666
A[C>G]G	0.001205332
A[C>G]T	0.017434403
C[C>G]A	0.006238753
C[C>G]C	0.004975743
C[C>G]G	0.004117009
C[C>G]T	0.007098419
G[C>G]A	0.001594505
G[C>G]C	4.75875E-20
G[C>G]G	0.000473898
G[C>G]T	0.003988173
T[C>G]A	0.003795465
T[C>G]C	0.004487134
T[C>G]G	0.001583128
T[C>G]T	0.023982606

A[C>T]A	0.014945584
A[C>T]C	0.008389485
A[C>T]G	0.004551594
A[C>T]T	0.011651419
C[C>T]A	0.009402314
C[C>T]C	0.006496308
C[C>T]G	0.002564394
C[C>T]T	0.009626158
G[C>T]A	0.010533685
G[C>T]C	0.009922544
G[C>T]G	0.002524279
G[C>T]T	0.001925169
T[C>T]A	0.028736617
T[C>T]C	0.011600356
T[C>T]G	0.004565619
T[C>T]T	0.011004112
A[T>A]A	0.006363384
A[T>A]C	4.75875E-20
A[T>A]G	0.011583259
A[T>A]T	0.001046362
C[T>A]A	4.75875E-20
C[T>A]C	0.008078025
C[T>A]G	0.009598124
C[T>A]T	0.013844478
G[T>A]A	0.004196896
G[T>A]C	0.000210043
G[T>A]G	0.0048592
G[T>A]T	0.000784631
T[T>A]A	0.007414714
T[T>A]C	0.005257449
T[T>A]G	4.75875E-20
T[T>A]T	0.012157859
A[T>C]A	4.75875E-20
A[T>C]C	0.004967699

A[T>C]G	0.030530339
A[T>C]T	6.78691E-16
C[T>C]A	0.004469143
C[T>C]C	0.011343202
C[T>C]G	0.009519282
C[T>C]T	0.004018189
G[T>C]A	0.000606475
G[T>C]C	1.49375E-11
G[T>C]G	0.002786154
G[T>C]T	0.004888446
T[T>C]A	0.009366417
T[T>C]C	0.000592396
T[T>C]G	0.001751501
T[T>C]T	0.014951023
A[T>G]A	0.027779096
A[T>G]C	0.020987142
A[T>G]G	0.056400253
A[T>G]T	0.045994861
C[T>G]A	0.007120867
C[T>G]C	0.00877218
C[T>G]G	0.01775835
C[T>G]T	0.037815563
G[T>G]A	0.00780353
G[T>G]C	0.003159935
G[T>G]G	0.014371422
G[T>G]T	0.012311793
T[T>G]A	0.019668356
T[T>G]C	0.014263082
T[T>G]G	0.037929245
T[T>G]T	0.053111095

Supplementary Table 16. Number of signature occurrences in the Mongolian and Western cohorts. Differences between Western samples from different origins (Europe and USA) are also indicated. Statistical differences between cohorts were assessed by Fisher test corrected by FDR.

	Mongolian	% Mongolian	Western (all)	% Western (all)	adj. p val	Europe	% Europe	USA	% USA	adj. p val
SBS1	0	0.0%	0	0.0%	1.000	0	0.0%	0	0.0%	NA
SBS4	2	1.3%	2	1.8%	1.000	2	2.9%	0	0.0%	0.834
SBS5	10	6.6%	13	11.6%	0.515	9	13.0%	4	9.3%	0.834
SBS6	1	0.7%	0	0.0%	1.000	0	0.0%	0	0.0%	NA
SBS12	2	1.3%	0	0.0%	0.800	0	0.0%	0	0.0%	NA
SBS16	0	0.0%	2	1.8%	0.515	2	2.9%	0	0.0%	0.834
SBS18	0	0.0%	1	0.9%	0.781	0	0.0%	1	2.3%	0.834
SBS22	21	13.9%	6	5.4%	0.137	0	0.0%	6	14.0%	0.018
SBS26	3	2.0%	0	0.0%	0.580	0	0.0%	0	0.0%	NA
SBS29	1	0.7%	0	0.0%	1.000	0	0.0%	0	0.0%	NA
SBS40	49	32.5%	33	29.5%	0.944	21	30.4%	12	27.9%	0.834
SBSM	38	25.2%	5	4.5%	0.00003	4	5.8%	1	2.3%	0.834

Supplementary Table 17. Frequency of occurrence of COSMIC and de novo mutational signatures in the external NCI Mongolian, and Korean HCC cohorts. Statistical differences compared to the in-house Mongolian cohorts were assessed by Fisher test corrected by FDR.

	NCI	% NCI	adj. p-val	Korean	% Korean	adj. p-val
SBS1	2	2.8%	0.171	0	0%	NA
SBS4	3	4.2%	0.441	36	15.6%	0.000
SBS5	27	38.0%	0.000	28	12.1%	0.153
SBS6	1	1.4%	0.587	0	0.0%	0.435
SBS12	18	25.4%	0.000	10	4.3%	0.213
SBS16	25	35.2%	0.000	81	35.1%	0.000
SBS18	12	16.9%	0.000	28	12.1%	0.000
SBS22	10	14.1%	1.000	23	10.0%	0.349
SBS26	5	7.0%	0.171	2	0.9%	0.435
SBS29	3	4.2%	0.171	2	0.9%	1.000
SBS40	53	74.6%	0.000	120	51.9%	0.000
SBSM	22	31.0%	0.502	10	4.3%	0.000

Supplementary Table 18. Number of environmental signature occurrences in the Mongolian and Western cohorts. Differences between Western samples from different origins (Europe and USA) are also indicated. Statistical differences between cohorts were assessed by Fisher test corrected by FDR.

	Mongolian	% Mongolian	Western	% Western	adj. p val	Europe	% Europe	USA	% USA	adj. p val
MNU_350_uM	61	40.4%	41	36.6%	0.799	28	40.6%	13	30.2%	0.858
DMS_0_078_mM	57	37.7%	21	18.8%	0.0419	14	20.3%	7	16.3%	1.000
AFB1_0_25_uM_plus_S9	46	30.5%	24	21.4%	0.4615	17	24.6%	7	16.3%	0.858
DES_0_938_mM	42	27.8%	33	29.5%	0.9406	23	33.3%	10	23.3%	0.858
DMH_11_6_mM_plus_S9	36	23.8%	32	28.6%	0.7154	19	27.5%	13	30.2%	1.000
ENU_400_uM	36	23.8%	23	20.5%	0.7815	13	18.8%	10	23.3%	1.000
Formaldehyde_120_uM	29	19.2%	32	28.6%	0.4130	19	27.5%	13	30.2%	1.000
S_6_Nitrochrysene_12_5_uM_plus_S9	29	19.2%	31	27.7%	0.463	19	27.5%	12	27.9%	1.000
DBP_0_0039_uM	27	17.9%	12	10.7%	0.4615	7	10.1%	5	11.6%	1.000
AZD7762_1_625_uM	23	15.2%	13	11.6%	0.7588	4	5.8%	9	20.9%	0.384
Mechlorethamine_0_3_uM	22	14.6%	21	18.8%	0.7154	12	17.4%	9	20.9%	1.000
Methyleugenol_1_25_mM	21	13.9%	14	12.5%	0.997	10	14.5%	4	9.3%	1.000
S_1_8_DNP_8_uM	20	13.2%	10	8.9%	0.6916	6	8.7%	4	9.3%	1.000
AAII_37_5_uM	18	11.9%	18	16.1%	0.7154	13	18.8%	5	11.6%	0.907
Temozolomide_200_uM	11	7.3%	16	14.3%	0.4608	13	18.8%	3	7.0%	0.789
S_6_Nitrochrysene_50_uM_plus_S9	10	6.6%	2	1.8%	0.4130	0	0.0%	2	4.7%	0.789
S_5_Methylchrysene_1_6_uM_plus_S9	9	6.0%	5	4.5%	0.9406	2	2.9%	3	7.0%	0.858
Furan_100_mM_plus_S9	7	4.6%	3	2.7%	0.7815	3	4.3%	0	0.0%	0.858
DBPDE_0_000625_uM	7	4.6%	1	0.9%	0.463	1	1.4%	0	0.0%	1.000
MX_7_uM_plus_S9	5	3.3%	11	9.8%	0.3722	7	10.1%	4	9.3%	1.000
OTA_0_08_uM_plus_S9	5	3.3%	6	5.4%	0.7815	4	5.8%	2	4.7%	1.000
AAI_1_25_uM	4	2.6%	4	3.6%	0.9246	0	0.0%	4	9.3%	0.384
S_1_8_DNP_0_125_uM	2	1.3%	1	0.9%	1	1	1.4%	0	0.0%	1.000
SSR_1_25_J	2	1.3%	10	8.9%	0.1081	6	8.7%	4	9.3%	1.000
S_3_NBA_0_1_uM	2	1.3%	2	1.8%	1	2	2.9%	0	0.0%	0.993
DBPDE_0_000156_uM	1	0.7%	0	0.0%	1	0	0.0%	0	0.0%	NA
DBP_0_0313_uM_plus_S9	1	0.7%	0	0.0%	1	0	0.0%	0	0.0%	NA
Benzidine_200_uM	1	0.7%	2	1.8%	0.7815	0	0.0%	2	4.7%	0.789
Cyclophosphamide_18_75_uM_plus_S9	1	0.7%	3	2.7%	0.6916	3	4.3%	0	0.0%	0.858
Potassium_bromate_260_uM	1	0.7%	3	2.7%	0.6916	3	4.3%	0	0.0%	0.858
S_6_Nitrochrysene_0_78_uM	1	0.7%	0	0.0%	1	0	0.0%	0	0.0%	NA
S_1_6_DNP_0_09_uM	1	0.7%	2	1.8%	0.7815	0	0.0%	2	4.7%	0.789
N_Nitrosopyrrolidine_50_mM	1	0.7%	0	0.0%	1	0	0.0%	0	0.0%	NA
Cisplatin_3_125_uM	1	0.7%	3	2.7%	0.6916	3	4.3%	0	0.0%	0.858
Carboplatin_5_uM	1	0.7%	6	5.4%	0.3722	1	1.4%	5	11.6%	0.384
S_6_Nitrochrysene_50_uM	1	0.7%	4	3.6%	0.4736	4	5.8%	0	0.0%	0.858
DBADE_0_109_uM	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA
Potassium_bromate_875_uM	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA
S_4_ABP_300_uM_plus_S9	0	0.0%	5	4.5%	0.1861	3	4.3%	2	4.7%	1.000
Semustine_150_uM	0	0.0%	2	1.8%	0.4736	1	1.4%	1	2.3%	1.000
DBADE_0_0313_uM	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA
BPDE_0_125_uM	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA
BaP_0_39_uM_plus_S9	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA
Ellipticine_0_375_uM_plus_S9	0	0.0%	3	2.7%	0.4130	2	2.9%	1	2.3%	1.000
DBA_75_uM_plus_S9	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA
PhIP_3_uM_plus_S9	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA
S_3_NBA_0_025_uM	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA

Propylene_oxide_10_mM	0	0.0%	2	1.8%	0.4736	2	2.9%	0	0.0%	0.993
DBAC_5_uM_plus_S9	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA
Cisplatin_12_5_uM	0	0.0%	1	0.9%	0.7154	0	0.0%	1	2.3%	0.858
PhIP_4_uM_plus_S9	0	0.0%	1	0.9%	0.7154	1	1.4%	0	0.0%	1.000
BaP_2_uM_plus_S9	0	0.0%	0	0.0%	1	0	0.0%	0	0.0%	NA

Supplementary Table 19. Frequency of environmental signature occurrences in the external NCI Mongolian, and Korean HCC cohorts. Statistical differences between cohorts were assessed by Fisher test corrected by FDR.

	adj. p val	NCI	% NCI	adj. p val	Korea	% Korea	adj. p val
MNU_350_uM	0.799	20	28.17%	0.593	44	19.0%	0.000
DMS_0_078_mM	0.042	17	23.94%	0.593	28	12.1%	0.000
AFB1_0_25_uM_plusS9	0.462	20	28.17%	1.000	0	0.0%	NA
DES_0_938_mM	0.941	13	18.31%	0.702	39	16.9%	0.055
DMH_11_6_mM_plusS9	0.715	22	30.99%	0.742	0	0.0%	NA
ENU_400_uM	0.782	18	25.35%	1.000	46	19.9%	0.625
Formaldehyde_120uM	0.413	14	19.72%	1.000	0	0.0%	NA
S_6_Nitrochrysene_12_5_uM_plusS9	0.463	17	23.94%	0.912	0	0.0%	NA
DBP_0_0039_uM	0.462	7	9.86%	0.704	9	3.9%	0.000
AZD7762_1_625_uM	0.759	11	15.49%	1.000	21	9.1%	0.223
Mechlorethamine_0_3uM	0.715	16	22.54%	0.704	0	0.0%	NA
Methyleugenol_1_25mM	0.997	17	23.94%	0.593	0	0.0%	NA
S_1_8_DNP_8_uM	0.692	10	14.08%	1.000	10	4.3%	0.016
AAII_37_5_uM	0.715	8	11.27%	1.000	35	15.2%	0.673
Temozolomide_200_uM	0.461	3	4.23%	0.912	41	17.7%	0.017
S_6_Nitrochrysene_50_uM_plusS9	0.413	2	2.82%	0.747	0	0.0%	NA
S_5_Methylchrysene_1_6_uM_plusS9	0.941	2	2.82%	0.912	0	0.0%	NA
Furan_100_mM_plusS9	0.782	5	7.04%	0.912	0	0.0%	NA
DBPDE_0_000625_uM	0.463	1	1.41%	0.904	6	2.6%	0.625
MX_7_uM_plus_S9	0.372	2	2.82%	1.000	18	7.8%	0.234
OTA_0_08_uM_plus_S9	0.782	5	7.04%	0.742	15	6.5%	0.429
AAI_1_25_uM	0.925	1	1.41%	1.000	21	9.1%	0.064
S_1_8_DNP_0_125_uM	1.000	1	1.41%	1.000	0	0.0%	0.304
SSR_1_25_J	0.108	4	5.63%	0.593	7	3.0%	0.673
S_3_NBA_0_1_uM	1.000	2	2.82%	0.938	15	6.5%	0.066
DBPDE_0_000156_uM	1.000	0	0.00%	1.000	4	1.7%	0.779
DBP_0_0313_uM_plusS9	1.000	0	0.00%	1.000	0	0.0%	NA
Benzidine_200_uM	0.782	0	0.00%	1.000	14	6.1%	0.029
Cyclophosphamide_18_75_uM_plusS9	0.692	0	0.00%	1.000	0	0.0%	NA
Potassium_br_260uM	0.692	2	2.82%	0.704	0	0.0%	NA
S_6_Nitrochrysene_0_78uM	1.000	1	1.41%	0.912	0	0.0%	NA
S_1_6_DNP_0_09_uM	0.782	0	0.00%	1.000	2	0.9%	1.000
N_Nitrosopyrrolidine_50mM	1.000	2	2.82%	0.704	0	0.0%	NA
Cisplatin_3_125_uM	0.692	0	0.00%	1.000	1	0.4%	1.000
Carboplatin_5_uM	0.372	2	2.82%	0.704	2	0.9%	1.000
S_6_Nitrochrysene_50uM	0.474	2	2.82%	0.704	0	0.0%	NA
DBADE_0_109_uM	1.000	0	0.00%	NA	2	0.9%	0.673
Potassium_bromate_875uM	1.000	0	0.00%	NA	0	0.0%	NA
S_4_ABP_300uM_plusS9	0.186	1	1.41%	0.742	0	0.0%	NA
Semustine_150uM	0.474	1	1.41%	0.742	0	0.0%	NA
DBADE_0_0313_uM	1.000	0	0.00%	NA	2	0.9%	0.673
BPDE_0_125_uM	1.000	0	0.00%	NA	0	0.0%	NA

BaP_0_39_uM_plus_S9	1.000	0	0.00%	NA	0	0.0%	NA
Ellipticine_0_375_uM_plusS9	0.413	0	0.00%	NA	0	0.0%	NA
DBA_75_uM_plusS9	1.000	0	0.00%	NA	0	0.0%	NA
PhIP_3_uM_plusS9	1.000	0	0.00%	NA	0	0.0%	NA
S_3_NBA_0_025_uM	1.000	2	2.82%	0.593	25	10.8%	0.000
Propylene_oxide_1_mM	0.474	2	2.82%	0.593	0	0.0%	NA
DBAC_5_uM_plusS9	1.000	0	0.00%	NA	0	0.0%	NA
Cisplatin_12_5_uM	0.715	0	0.00%	NA	1	0.4%	1.000
PhIP_4_uM_plusS9	0.715	2	2.82%	0.593	0	0.0%	NA
BaP_2_uM_plusS9	1.000	0	0.00%	NA	0	0.0%	NA

Supplementary Table 20. Clinico-pathological characteristics of the Mongolian gene expression-based clusters in the Mongolian cohort.

	MGL1 (n=47, 44%)	MGL2 (n=27, 26%)	MGL3 (n=32, 30%)	p value
Age (years)	62 (43-76)	56 (44-72)	56 (41-75)	0.028
<60 years (n, %)	18 (39.1)	20 (76.9)	19 (63.3)	0.005
Gender (male, %)	29 (63)	8 (30.8)	19 (63.3)	0.017
BMI > 25 kg/m^2 (n, %)	26 (56.5)	12 (46.2)	11 (39.3)	ns
Etiology:				
• HBV (n, %)	4 (8.5)	0 (0)	0 (0)	
• HBV/HDV (n, %)	27 (57.4)	19 (70.4)	26 (81.3)	<0.001
• HBV/HCV/HDV (n, %)	1 (2.1)	6 (22.2)	5 (15.6)	
• HCV (n, %)	15 (31.9)	2 (7.4)	1 (3.1)	
Region				
• Western (n, %)	9 (20.5)	6 (26.1)	4 (14.8)	
• Central (n, %)	18 (40.9)	9 (39.1)	13 (48.1)	ns
• Eastern (n, %)	1 (2.3)	3 (13)	3 (11.1)	
• Ulaanbaatar (n, %)	16 (36.4)	5 (21.7)	7 (25.9)	
BCLC stage (0-A, %)	35 (77.8)	17 (68)	24 (85.7)	ns
AFP >400 IU/mL (n, %)	5 (13.5)	9 (42.9)	4 (16.7)	0.026
Liver fibrosis (F3-4, %)	17 (36.2)	12 (44.4)	12 (37.5)	ns
Microvascular invasion (yes, %)	16 (41)	13 (65)	9 (34.6)	0.099
Tumor grade (G3-4, %)	4 (10)	4 (26.7)	0 (0)	0.045

BMI, body mass index; HBV, hepatitis B virus; HCV, hepatitis C virus; HDV, hepatitis delta virus; BCLC, Barcelona Clinic Liver Cancer, AFP, alfa-fetoprotein
Some variables have missing values for MGL1, MGL2 and MGL3, respectively:
Age, gender: 1, 1, and 2 patients. BMI: 1, 1, and 4 patients. Region: 3, 4, and 5 patients. BCLC stage: 2, 2, and 4 patients. AFP: 10, 6, and 8 patients. Microvascular invasion: 8, 7, and 8 patients. Tumor grade: 7, 12, and 16 patients.

Supplementary Table 21. Pathways enriched in MGL1 cluster according to gene set enrichment analysis (GSEA).

NAME	MSigDB	ES	NES	NOM p-val	FDR q-val
HALLMARK_BILE_ACID_METABOLISM	Hallmark	0.548	2.256	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_DN	C2	0.716	3.167	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_CTNNB1_UP	C2	0.718	3.119	0.000	0.000
LEE_LIVER_CANCER_SURVIVAL_UP	C2	0.699	3.060	0.000	0.000
WOO_LIVER_CANCER_RECURRENCE_DN	C2	0.758	2.929	0.000	0.000
BOYAULT_LIVER_CANCER_SUBCLASS_G123_DN	C2	0.735	2.597	0.000	0.000
KIM_LIVER_CANCER_POOR_SURVIVAL_DN	C2	0.730	2.490	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_INTERFERON_DN	C2	0.689	2.482	0.000	0.000
BOYAULT_LIVER_CANCER_SUBCLASS_G6_UP	C2	0.630	2.458	0.000	0.000
HOSHIDA_LIVER_CANCER_SUBCLASS_S3	C2	0.496	2.310	0.000	0.000
LEE_LIVER_CANCER_DENA_DN	C2	0.579	2.179	0.000	0.001
HSIAO_LIVER_SPECIFIC_GENES	C2	0.479	2.184	0.000	0.001
MINGUEZ_LIVER_CANCER_VASCULAR_INVASION_DN	C2	0.731	2.146	0.000	0.001
NIKOLSKY_BREAST_CANCER_1Q21_AMPLICON	C2	0.637	2.155	0.000	0.001
KEGG_GLYCINE_SERINE_AND_THREONINE_METABOLISM	C2	0.654	2.141	0.000	0.001
ANDERSEN_LIVER_CANCER_KRT19_DN	C2	0.557	2.124	0.000	0.001
CHIANG_LIVER_CANCER_SUBCLASS_POLYSOMY7_UP	C2	0.548	2.112	0.000	0.001
REACTOME_PEROXISOMAL_PROTEIN_IMPORT	C2	0.575	2.098	0.000	0.001
VILLANUEVA_LIVER_CANCER_KRT19_DN	C2	0.552	2.066	0.000	0.002
BOYAULT_LIVER_CANCER_SUBCLASS_G3_DN	C2	0.572	2.057	0.000	0.003
KEGG_PRIMARY_BILE_ACID_BIOSYNTHESIS	C2	0.742	2.046	0.002	0.003
OHGUCHI_LIVER_HNF4A_TARGETS_DN	C2	0.470	2.017	0.000	0.004
REACTOME_GLYOXYLATE_METABOLISM_AND_GLYCINE_DEGRADATION	C2	0.633	2.007	0.000	0.005
KEGG_PEROXISOME	C2	0.512	1.994	0.000	0.006
REACTOME_SYNTHESIS_OF_BILE_ACIDS_AND_BILE_SALTS_VIA_7ALPHA_HYDROXYCHOLESTEROL	C2	0.660	1.995	0.000	0.006
YAMASHITA_LIVER_CANCER_WITH_EPCAM_DN	C2	0.714	1.937	0.000	0.011
REACTOME_BILE_ACID_AND_BILE_SALT_METABOLISM	C2	0.569	1.937	0.002	0.012
LAIHO_COLORECTAL_CANCER_SERRATED_DN	C2	0.491	1.922	0.000	0.013
GINESTIER_BREAST_CANCER_ZNF217_AMPLIFIED_DN	C2	0.413	1.918	0.000	0.013
REACTOME_RECYCLING_OF_BILE_ACIDS_AND_SALTS	C2	0.711	1.908	0.002	0.014
REACTOME_BIOLOGICAL_OXIDATIONS	C2	0.417	1.873	0.000	0.021
KEGG_STEROID_HORMONE_BIOSYNTHESIS	C2	0.516	1.839	0.000	0.027
REACTOME_PEROXISOMAL_LIPID_METABOLISM	C2	0.574	1.844	0.002	0.028
YAMASHITA_LIVER_CANCER_STEM_CELL_DN	C2	0.473	1.840	0.002	0.028
REACTOME_PHASE_I_FUNCTIONALIZATION_OF_COMPOUNDS	C2	0.442	1.844	0.000	0.029
MOOTHA_FFA_OXYDATION	C2	0.643	1.833	0.005	0.029
BOYAULT_LIVER_CANCER_SUBCLASS_G12_DN	C2	0.685	1.828	0.012	0.029
REACTOME_SYNTHESIS_OF_BILE_ACIDS_AND_BILE_SALTS	C2	0.554	1.820	0.005	0.031
BIOCARTA_NUCLEARRS_PATHWAY	C2	0.550	1.810	0.000	0.033
LEE_LIVER_CANCER_ACOX1_DN	C2	0.481	1.812	0.000	0.033
LEE_LIVER_CANCER_CIPROFIBRATE_DN	C2	0.486	1.807	0.003	0.033
REACTOME_SULFUR_AMINO_ACID_METABOLISM	C2	0.610	1.797	0.005	0.036
WENG_POR_TARGETS_GLOBAL_UP	C2	0.653	1.792	0.007	0.036
REACTOME_TRANSPORT_OF_BILE_SALTS_AND_ORGANIC_ACIDS_METAL_IONS_AND_AMINE_COMPOUNDS	C2	0.445	1.765	0.000	0.047

Supplementary Table 22. Pathways enriched in MGL2 cluster according to gene set enrichment analysis (GSEA).

NAME	MSigDB	ES	NES	NOM p-val	FDR q-val
HALLMARK_G2M_CHECKPOINT	Hallmark	-0.379	-1.718	0.000	0.015
HALLMARK_MITOTIC_SPINDLE	Hallmark	-0.382	-1.705	0.000	0.009
GINESTIER_BREAST_CANCER_20Q13_AMPLIFICATION_DN	C2	-0.573	-2.524	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP	C2	-0.550	-2.461	0.000	0.000
GINESTIER_BREAST_CANCER_ZNF217_AMPLIFIED_DN	C2	-0.508	-2.418	0.000	0.000
WHITFIELD_CELL_CYCLE_G1_S	C2	-0.557	-2.344	0.000	0.000
VILLANUEVA_LIVER_CANCER_KRT19_UP	C2	-0.521	-2.332	0.000	0.000
PYEON_CANCER_HEAD_AND_NECK_VS_CERVICAL_UP	C2	-0.522	-2.326	0.000	0.000
WHITFIELD_CELL_CYCLE_S	C2	-0.518	-2.254	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_CTNNB1_DN	C2	-0.497	-2.196	0.000	0.001
ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER	C2	-0.514	-2.193	0.000	0.001
PYEON_HPV_POSITIVE_TUMORS_UP	C2	-0.541	-2.187	0.000	0.001
REACTOME_RESOLUTION_OF_D_LOOP_STRUCTURES	C2	-0.645	-2.110	0.000	0.004
LEE_DIFFERENTIATING_T_LYMPHOCYTE	C2	-0.468	-2.063	0.000	0.008
DUTERTRE_ESTRADIOL_RESPONSE_24HR_UP	C2	-0.428	-2.035	0.000	0.011
KEGG_PHOSPHATIDYLINOSITOL_SIGNALING_SYSTEM	C2	-0.522	-2.024	0.000	0.011
KANG_DOXORUBICIN_RESISTANCE_UP	C2	-0.544	-1.993	0.000	0.017
LAIHO_COLORECTAL_CANCER_SERRATED_DN	C2	-0.497	-1.949	0.000	0.029
CHIANG_LIVER_CANCER_SUBCLASS_UNANNOTATED_UP	C2	-0.483	-1.942	0.000	0.030
ZHAN_MULTIPLE_MYELOMA_PR_UP	C2	-0.565	-1.933	0.000	0.031
REACTOME_HDR_THROUGH_HOMOLOGOUS_RECOMBINATION_HRR	C2	-0.514	-1.916	0.000	0.037
XU_HGF_TARGETS_INDUCED_BY_AKT1_48HR_DN	C2	-0.650	-1.909	0.010	0.039
HOEBEKE_LYMPHOID_STEM_CELL_UP	C2	-0.481	-1.904	0.000	0.040

Supplementary Table 23. Pathways enriched in MGL3 cluster according to gene set enrichment analysis (GSEA).

NAME	MSigDB	ES	NES	NOM p-val	FDR q-val
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	Hallmark	-0.707	-3.128	0.000	0.000
HALLMARK_ALLOGRAFT_REJECTION	Hallmark	-0.618	-2.743	0.000	0.000
HALLMARK_INFLAMMATORY_RESPONSE	Hallmark	-0.579	-2.576	0.000	0.000
HALLMARK_KRAS_SIGNALING_UP	Hallmark	-0.572	-2.522	0.000	0.000
HALLMARK_COMPLEMENT	Hallmark	-0.501	-2.236	0.000	0.000
HALLMARK_TNFA_SIGNALING_VIA_NFKB	Hallmark	-0.495	-2.231	0.000	0.000
HALLMARK_IL2_STAT5_SIGNALING	Hallmark	-0.496	-2.228	0.000	0.000
HALLMARK_ANGIOGENESIS	Hallmark	-0.659	-2.163	0.000	0.000
HALLMARK_TGF_BETA_SIGNALING	Hallmark	-0.541	-1.926	0.000	0.000
HALLMARK_INTERFERON_GAMMA_RESPONSE	Hallmark	-0.409	-1.847	0.000	0.001
SCHUETZ_BREAST_CANCER_DUCTAL_INVASIVE_UP	C2	-0.727	-3.474	0.000	0.000
MCLACHLAN_DENTAL_CARIES_UP	C2	-0.758	-3.471	0.000	0.000
REACTOME_CREATION_OF_C4_AND_C2_ACTIVATORS	C2	-0.883	-3.320	0.000	0.000
REACTOME_SCAVENGING_OF_HEME_FROM_PLASMA	C2	-0.895	-3.305	0.000	0.000
POOLA_INVASIVE_BREAST_CANCER_UP	C2	-0.707	-3.305	0.000	0.000
REACTOME_INITIAL_TRIGGERING_OF_COMPLEMENT	C2	-0.856	-3.300	0.000	0.000
REACTOME_ROLE_OF_LAT2_NTAL_LAB_ON_CALCIUM_MOBILIZATION	C2	-0.895	-3.290	0.000	0.000
REACTOME_BINDING_AND_UPTAKE_OF_LIGANDS_BY_SCAVENGER_RECEPTORS	C2	-0.835	-3.287	0.000	0.000
REACTOME_FCGR_ACTIVATION	C2	-0.880	-3.252	0.000	0.000
REACTOME_ROLE_OF_PHOSPHOLIPIDS_IN_PHAGOCYTOSIS	C2	-0.844	-3.240	0.000	0.000
REACTOME_ANTIGEN_ACTIVATES_B_CELL_RECEPTOR_BCR_LEADING_TO_GENERATION_OF_SECOND_MESSENGERS	C2	-0.829	-3.221	0.000	0.000
REACTOME_CD22_MEDIATED_BCR_REGULATION	C2	-0.907	-3.217	0.000	0.000
REACTOME_IMMUNOREGULATORY_INTERACTIONS_BETWEEN_A_LYMPHOID_AND_A_NON_LYMPHOID_CELL	C2	-0.744	-3.216	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_CTNNB1_DN	C2	-0.736	-3.191	0.000	0.000
REACTOME_FCERI_MEDIATED_MAPK_ACTIVATION	C2	-0.811	-3.188	0.000	0.000
REACTOME_FCERI_MEDIATED_NF_KB_ACTIVATION	C2	-0.764	-3.178	0.000	0.000
BOQUEST_STEM_CELL_UP	C2	-0.693	-3.161	0.000	0.000
REACTOME_FCERI_MEDIATED_CAPLUS2_MOBILIZATION	C2	-0.824	-3.157	0.000	0.000
REACTOME_REGULATION_OF_ACTIN_DYNAMICS_FOR_PHAGOCYTIC_CUP_FORMATION	C2	-0.770	-3.119	0.000	0.000
VECCHI_GASTRIC_CANCER_ADVANCED_VS_EARLY_UP	C2	-0.708	-3.104	0.000	0.000
REACTOME_COMPLEMENT_CASCADE	C2	-0.768	-3.087	0.000	0.000
SMID_BREAST_CANCER_NORMAL_LIKE_UP	C2	-0.627	-3.072	0.000	0.000
REACTOME_CELL_SURFACE_INTERACTIONS_AT_THE_VASCULAR_WALL	C2	-0.696	-3.036	0.000	0.000
WALLACE_PROSTATE_CANCER_RACE_UP	C2	-0.644	-2.995	0.000	0.000
REACTOME_FCGAMMA_RECEPTOR_FCGR_DEPENDENT_PHAGOCYTOSIS	C2	-0.711	-2.977	0.000	0.000
ANASTASSIOU_MULTICANCER_INVASIVENESS_SIGNATURE	C2	-0.799	-2.971	0.000	0.000
NAKAYAMA_SOFT_TISSUE_TUMORS_PCA1_UP	C2	-0.765	-2.948	0.000	0.000
REACTOME_SIGNALING_BY_THE_B_CELL_RECEPTOR_BCR	C2	-0.688	-2.918	0.000	0.000
HOSHIDA_LIVER_CANCER_SUBCLASS_S1	C2	-0.641	-2.916	0.000	0.000
MIYAGAWA_TARGETS_OF_EWSR1_ETS_FUSIONS_DN	C2	-0.649	-2.900	0.000	0.000
ACEVEDO_LIVER_CANCER_WITH_H3K27ME3_UP	C2	-0.643	-2.899	0.000	0.000
REACTOME_FC_EPSILON_RECEPTOR_FCERI_SIGNALING	C2	-0.661	-2.895	0.000	0.000
CHARAFE_BREAST_CANCER_LUMINAL_VS_MESENCHYMAL_DN	C2	-0.591	-2.879	0.000	0.000
FLECHNER_BIOPSY_KIDNEY_TRANSPLANT_REJECTED_VS_OK_UP	C2	-0.733	-2.854	0.000	0.000
RODWELL_AGING_KIDNEY_UP	C2	-0.586	-2.838	0.000	0.000

VERHAAK_GLIOBLASTOMA_MESENCHYMAL	C2	-0.628	-2.830	0.000	0.000
FARMER_BREAST_CANCER_CLUSTER_1	C2	-0.805	-2.829	0.000	0.000
RIGGI_EWING_SARCOMA_PROGENITOR_DN	C2	-0.643	-2.828	0.000	0.000
WANG_SMARCE1_TARGETS_UP	C2	-0.613	-2.812	0.000	0.000
PICCALUGA_ANGIOIMMUNOBLASTIC_LYMPHOMA_UP	C2	-0.635	-2.809	0.000	0.000
LINDGREN_BLADDER_CANCER_CLUSTER_2B	C2	-0.583	-2.802	0.000	0.000
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_8D_DN	C2	-0.626	-2.791	0.000	0.000
ICHIBA_GRAFT_VERSUS_HOST_DISEASE_35D_UP	C2	-0.661	-2.789	0.000	0.000
DELYS_THYROID_CANCER_UP	C2	-0.575	-2.781	0.000	0.000
TURASHVILI_BREAST_LOBULAR_CARCINOMA_VS_DUCTAL_NORMAL_UP	C2	-0.727	-2.780	0.000	0.000
SMID_BREAST_CANCER_RELAPSE_IN_BONE_DN	C2	-0.590	-2.773	0.000	0.000
RASHI_RESPONSE_TO_IONIZING_RADIATION_6	C2	-0.713	-2.766	0.000	0.000
RODWELL_AGING_KIDNEY_NO_BLOOD_UP	C2	-0.618	-2.765	0.000	0.000
WU_CELL_MIGRATION	C2	-0.634	-2.760	0.000	0.000
KIM_GLIS2_TARGETS_UP	C2	-0.719	-2.755	0.000	0.000
WOO_LIVER_CANCER_RECURRENCE_UP	C2	-0.691	-2.742	0.000	0.000
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_10D_DN	C2	-0.644	-2.738	0.000	0.000
TURASHVILI_BREAST_LOBULAR_CARCINOMA_VS_LOBULAR_NORMAL_DN	C2	-0.712	-2.729	0.000	0.000
OISHI_CHOLANGIOMA_STEM_CELL_LIKE_DN	C2	-0.594	-2.728	0.000	0.000
BOYLAN_MULTIPLE_MYELOMA_PCA1_UP	C2	-0.662	-2.724	0.000	0.000
PETROVA_ENDOTHELIUM_LYMPHATIC_VS_BLOOD_DN	C2	-0.637	-2.724	0.000	0.000
JAATINEN_HEMATOPOIETIC_STEM_CELL_DN	C2	-0.602	-2.722	0.000	0.000
ONDER_CDH1_TARGETS_2_UP	C2	-0.591	-2.707	0.000	0.000
BOYLAN_MULTIPLE_MYELOMA_C_D_DN	C2	-0.583	-2.693	0.000	0.000
LINDVALL_IMMORTALIZED_BY_TERT_DN	C2	-0.709	-2.692	0.000	0.000
CROMER_TUMORIGENESIS_UP	C2	-0.747	-2.686	0.000	0.000
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_16D_UP	C2	-0.615	-2.682	0.000	0.000
RICKMAN_HEAD_AND_NECK_CANCER_A	C2	-0.669	-2.682	0.000	0.000
CHEBOTAEV_GR_TARGETS_DN	C2	-0.644	-2.670	0.000	0.000
VERHAAK_AML_WITH_NPM1_MUTATED_DN	C2	-0.578	-2.654	0.000	0.000
MCBRYAN_PUBERTAL_BREAST_4_5WK_UP	C2	-0.585	-2.653	0.000	0.000
PASINI_SUZ12_TARGETS_DN	C2	-0.568	-2.649	0.000	0.000
SABATES_COLORECTAL_ADENOMA_DN	C2	-0.570	-2.644	0.000	0.000
CLASPER_LYMPHATIC_VESSELS_DURING_METASTASIS_DN	C2	-0.800	-2.643	0.000	0.000
CERVERA_SDHB_TARGETS_1_UP	C2	-0.646	-2.628	0.000	0.000
THUM_SYSTOLIC_HEART_FAILURE_UP	C2	-0.547	-2.624	0.000	0.000
CHARAFE_BREAST_CANCER_LUMINAL_VS_BASAL_DN	C2	-0.542	-2.623	0.000	0.000
VECCHI_GASTRIC_CANCER_EARLY_DN	C2	-0.550	-2.614	0.000	0.000
SENESE_HDAC1_AND_HDAC2_TARGETS_DN	C2	-0.580	-2.608	0.000	0.000
SERVITJA_ISLET_HNF1A_TARGETS_UP	C2	-0.600	-2.594	0.000	0.000
CHICAS_RB1_TARGETS_GROWING	C2	-0.562	-2.582	0.000	0.000
CHIARADONNA_NEOPLASTIC_TRANSFORMATION_CDC25_UP	C2	-0.627	-2.581	0.000	0.000
WIELAND_UP_BY_HBV_INFECTION	C2	-0.640	-2.579	0.000	0.000
NABA_CORE_MATRISOME	C2	-0.562	-2.579	0.000	0.000
GRUETZMANN_PANCREATIC_CANCER_UP	C2	-0.542	-2.576	0.000	0.000
LIM_MAMMARY_STEM_CELL_UP	C2	-0.532	-2.575	0.000	0.000
LIAN_LIPA_TARGETS_3M	C2	-0.688	-2.566	0.000	0.000
BOQUEST_STEM_CELL_CULTURED_VS_FRESH_UP	C2	-0.531	-2.556	0.000	0.000
RIGGI_EWING_SARCOMA_PROGENITOR_UP	C2	-0.525	-2.548	0.000	0.000
SWEET_LUNG_CANCER_KRAS_DN	C2	-0.535	-2.543	0.000	0.000
ALTEMEIER_RESPONSE_TO_LPS_WITH_MECHANICAL_VENTILATION	C2	-0.613	-2.538	0.000	0.000
HOELZEL_NF1_TARGETS_UP	C2	-0.598	-2.537	0.000	0.000
REN_ALVEOLAR_RHABDOMYOSARCOMA_DN	C2	-0.529	-2.536	0.000	0.000

NADLER_OBESITY_UP	C2	-0.700	-2.532	0.000	0.000
GAUSSMANN_MLL_AF4_FUSION_TARGETS_E_UP	C2	-0.635	-2.531	0.000	0.000
COWLING_MYCN_TARGETS	C2	-0.745	-2.529	0.000	0.000
RICKMAN_HEAD_AND_NECK_CANCER_D	C2	-0.758	-2.525	0.000	0.000
MARKEY_RB1_ACUTE_LOF_DN	C2	-0.557	-2.520	0.000	0.000
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_10D_UP	C2	-0.568	-2.517	0.000	0.000
REACTOME_INTERLEUKIN_10_SIGNALING	C2	-0.741	-2.514	0.000	0.000
SENESE_HDAC2_TARGETS_DN	C2	-0.613	-2.514	0.000	0.000
VART_KSHV_INFECTION_ANGIOGENIC_MARKERS_UP	C2	-0.580	-2.514	0.000	0.000
NABA_ECM_GLYCOPROTEINS	C2	-0.577	-2.512	0.000	0.000
LENAOUR_DENDRITIC_CELL_MATURATION_DN	C2	-0.591	-2.510	0.000	0.000
GAURNIER_PSMD4_TARGETS	C2	-0.667	-2.508	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP	C2	-0.569	-2.504	0.000	0.000
REACTOME_CHEMOKINE_RECEPTORS_BIND_CHEMOKINES	C2	-0.715	-2.500	0.000	0.000
HADDAD_T_LYMPHOCYTE_AND_NK_PROGENITOR_DN	C2	-0.672	-2.500	0.000	0.000
LIM_MAMMARY_LUMINAL_MATURE_DN	C2	-0.641	-2.499	0.000	0.000
LEE_NEURAL_CREST_STEM_CELL_UP	C2	-0.585	-2.498	0.000	0.000
RUTELLA_RESPONSE_TO_HGF_VS_CSF2RB_AND_IL4_UP	C2	-0.521	-2.497	0.000	0.000
LENAOUR_DENDRITIC_CELL_MATURATION_UP	C2	-0.603	-2.496	0.000	0.000
LIU_VAV3_PROSTATE_CARCINOGENESIS_UP	C2	-0.634	-2.487	0.000	0.000
WINTER_HYPOXIA_DN	C2	-0.709	-2.486	0.000	0.000
ONDER_CDH1_TARGETS_2_DN	C2	-0.512	-2.482	0.000	0.000
WAMUNYOKOLI_OVARIAN_CANCER_LMP_DN	C2	-0.556	-2.480	0.000	0.000
STEGER_ADIPOGENESIS_DN	C2	-0.806	-2.479	0.000	0.000
CHANDRAN_METASTASIS_DN	C2	-0.534	-2.478	0.000	0.000
ANDERSEN_CHOLANGIOCARCINOMA_CLASS2	C2	-0.573	-2.474	0.000	0.000
AZARE_NEOPLASTIC_TRANSFORMATION_BY_STAT3_UP	C2	-0.588	-2.468	0.000	0.000
HUANG_GATA2_TARGETS_UP	C2	-0.574	-2.464	0.000	0.000
PLASARI_TGFB1_TARGETS_10HR_DN	C2	-0.538	-2.464	0.000	0.000
LIAN_LIPA_TARGETS_6M	C2	-0.645	-2.463	0.000	0.000
RUTELLA_RESPONSE_TO_HGF_UP	C2	-0.512	-2.456	0.000	0.000
CAIRO_LIVER_DEVELOPMENT_UP	C2	-0.571	-2.454	0.000	0.000
HUPER_BREAST_BASAL_VS_LUMINAL_UP	C2	-0.690	-2.448	0.000	0.000
DAVICIONI_TARGETS_OF_PAX_FOXO1_FUSIONS_UP	C2	-0.530	-2.444	0.000	0.000
BRUECKNER_TARGETS_OF_MIRLET7A3_DN	C2	-0.647	-2.437	0.000	0.000
TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_ERYTHROCYTE_UP	C2	-0.561	-2.436	0.000	0.000
KAAB_HEART_ATRIUM_VS_VENTRICLE_UP	C2	-0.529	-2.430	0.000	0.000
CHYLA_CBFA2T3_TARGETS_UP	C2	-0.511	-2.427	0.000	0.000
PETROVA_PROX1_TARGETS_DN	C2	-0.668	-2.427	0.000	0.000
ZHENG_BOUND_BY_FOXP3	C2	-0.495	-2.426	0.000	0.000
OSADA_ASCL1_TARGETS_DN	C2	-0.806	-2.424	0.000	0.000
WESTON_VEGFA_TARGETS	C2	-0.604	-2.422	0.000	0.000
MOHANKUMAR_HOXA1_TARGETS_DN	C2	-0.566	-2.421	0.000	0.000
THUM_MIR21_TARGETS_HEART_DISEASE_UP	C2	-0.872	-2.421	0.000	0.000
LIU_PROSTATE_CANCER_DN	C2	-0.494	-2.410	0.000	0.000
HUANG_DASATINIB_RESISTANCE_UP	C2	-0.612	-2.407	0.000	0.000
QI_PLASMACYTOMA_UP	C2	-0.521	-2.405	0.000	0.000
TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_HSC_DN	C2	-0.542	-2.403	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_INTERFERON_UP	C2	-0.792	-2.403	0.000	0.000
TURASHVILI_BREAST_DUCTAL_CARCINOMA_VS_DUCTAL_NORMAL_UP	C2	-0.696	-2.395	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_UNANNOTATED_DN	C2	-0.537	-2.395	0.000	0.000
LINDSTEDT_DENDRITIC_CELL_MATURATION_A	C2	-0.648	-2.388	0.000	0.000
ONDER_CDH1_SIGNALING_VIA_CTNNB1	C2	-0.608	-2.385	0.000	0.000

IZADPANAH_STEM_CELL_ADIPOSE_VS_BONE_DN	C2	-0.587	-2.383	0.000	0.000
VANTVEER_BREAST_CANCER_ESR1_DN	C2	-0.517	-2.382	0.000	0.000
WESTON_VEGFA_TARGETS_6HR	C2	-0.672	-2.378	0.000	0.000
KEGG_TYPE_I_DIABETES_MELLITUS	C2	-0.714	-2.376	0.000	0.000
DURAND_STROMA_S_UP	C2	-0.512	-2.376	0.000	0.000
HELLEBREKERS_SILENCED_DURING_TUMOR_ANGIOGENESIS	C2	-0.612	-2.370	0.000	0.000
REACTOME_ELASTIC_FIBRE_FORMATION	C2	-0.693	-2.370	0.000	0.000
MARKEY_RB1_CHRONIC_LOF_DN	C2	-0.586	-2.370	0.000	0.000
TURASHVILI_BREAST_DUCTAL_CARCINOMA_VS_LOBULAR_NORMAL_DN	C2	-0.629	-2.365	0.000	0.000
KEGG_HEMATOPOIETIC_CELL_LINEAGE	C2	-0.617	-2.364	0.000	0.000
WU_SILENCED_BY_METHYLATION_IN_BLADDER_CANCER	C2	-0.661	-2.361	0.000	0.000
RUTELLA_RESPONSE_TO_CSF2RB_AND_IL4_DN	C2	-0.505	-2.358	0.000	0.000
WILENSKY_RESPONSE_TO_DARAPLADIB	C2	-0.753	-2.356	0.000	0.000
ZHENG_IL22_SIGNALING_UP	C2	-0.662	-2.354	0.000	0.000
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	C2	-0.519	-2.349	0.000	0.000
RICKMAN_HEAD_AND_NECK_CANCER_B	C2	-0.677	-2.349	0.000	0.000
LAIHO_COLORECTAL_CANCER_SERRATED_UP	C2	-0.586	-2.343	0.000	0.000
ICHIBA_GRAFT_VERSUS_HOST_DISEASE_D7_UP	C2	-0.576	-2.342	0.000	0.000
ROSS_AML_WITH_CBFB_MYH11_FUSION	C2	-0.660	-2.341	0.000	0.000
KIM_LIVER_CANCER_POOR_SURVIVAL_UP	C2	-0.800	-2.338	0.000	0.000
BEGUM_TARGETS_OF_PAX3_FOXO1_FUSION_UP	C2	-0.633	-2.335	0.000	0.000
SCHAEFFER_PROSTATE_DEVELOPMENT_48HR_DN	C2	-0.490	-2.333	0.000	0.000
WANG_MLL_TARGETS	C2	-0.502	-2.333	0.000	0.000
WINZEN_DEGRADED_VIA_KHSRP	C2	-0.590	-2.331	0.000	0.000
LEE_LIVER_CANCER_SURVIVAL_DN	C2	-0.543	-2.329	0.000	0.000
IGLESIAS_E2F_TARGETS_UP	C2	-0.550	-2.329	0.000	0.000
CROONQUIST_NRAS_SIGNALING_UP	C2	-0.694	-2.327	0.000	0.000
PAPASPYRIDONOS_UNSTABLE_ATEROSCLEROTIC_PLAQUE_DN	C2	-0.677	-2.327	0.000	0.000
VERHAAK_AML_WITH_NPM1_MUTATED_UP	C2	-0.531	-2.326	0.000	0.000
GRAHAM_CML_DIVIDING_VS_NORMAL_QUIESCENT_DN	C2	-0.585	-2.325	0.000	0.000
REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION	C2	-0.500	-2.321	0.000	0.000
LEE_LIVER_CANCER_DENA_UP	C2	-0.625	-2.316	0.000	0.000
RHEIN_ALL_GLUCOCORTICOID_THERAPY_UP	C2	-0.614	-2.313	0.000	0.000
LU_IL4_SIGNALING	C2	-0.586	-2.312	0.000	0.000
ZHU_CMV_24_HR_DN	C2	-0.589	-2.311	0.000	0.000
KLEIN_PRIMARY_EFFUSION_LYMPHOMA_DN	C2	-0.647	-2.310	0.000	0.000
CROONQUIST_STROMAL_STIMULATION_UP	C2	-0.635	-2.309	0.000	0.000
KUROZUMI_RESPONSE_TO_ONCOCYTIC_VIRUS	C2	-0.680	-2.308	0.000	0.000
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_6HR_UP	C2	-0.584	-2.308	0.000	0.000
LI_INDUCED_T_TO_NATURAL_KILLER_UP	C2	-0.496	-2.305	0.000	0.000
DACOSTA_ERCC3_ALLELE_XPCS_VS_TTD_DN	C2	-0.723	-2.304	0.000	0.000
BOSCO_TH1_CYTOTOXIC_MODULE	C2	-0.561	-2.304	0.000	0.000
VERRECCHIA_EARLY_RESPONSE_TO_TGFB1	C2	-0.643	-2.303	0.000	0.000
PEDERSEN_METASTASIS_BY_ERBB2_ISOFORM_4	C2	-0.565	-2.302	0.000	0.000
REACTOME_PHOSPHORYLATION_OF_CD3_AND_TCR_ZETA_CHAINS	C2	-0.782	-2.302	0.000	0.000
CUI_TCF21_TARGETS_UP	C2	-0.720	-2.300	0.000	0.000
SUNG_METASTASIS_STROMA_UP	C2	-0.560	-2.300	0.000	0.000
LIAO_HAVE_SOX4_BINDING_SITES	C2	-0.669	-2.299	0.000	0.000
DIAZ_CHRONIC_MEYLOGENOUS_LEUKEMIA_DN	C2	-0.572	-2.297	0.000	0.000

Supplementary Table 24. Clinico-pathological characteristics of the Mongolian gene expression-based clusters in the Mongolian NCI cohort.

	MGL1 (n=34, 49%)	MGL2 (n=18, 26%)	MGL3 (n=18, 26%)	p value
Age (years)	62 (41-76)	56.5 (45-77)	59 (23-68)	0.19
<60 years old (n, %)	13 (38.2)	10 (55.6)	9 (50.0)	0.546
Gender (male, %)	21 (61.7)	8 (44.4)	8 (44.4)	0.45
BMI > 25 kg/m2 (n, %)	13 (38.2)	6 (33.3)	7 (38.9)	0.83
Viral infection:				
HBV + (n, %)	22 (64.7)	14 (77.8)	13 (72.2)	0.804
HDV + (n, %)	14 (41.2)	5 (27.8)	8 (44.4)	0.626
HCV + (n, %)	23 (67.6)	10 (55.6)	8 (44.4)	0.473
AFP >20 IU/ml (n, %)	10 (29.4)	10 (55.6)	5 (27.8)	0.023
Liver cirrhosis (F4, %)	14 (41.2)	9 (50.0)	5 (27.8)	0.135

BMI, body mass index; HBV, hepatitis B virus; HCV, hepatitis C virus; HDV, hepatitis delta virus; AFP, alfa-fetoprotein