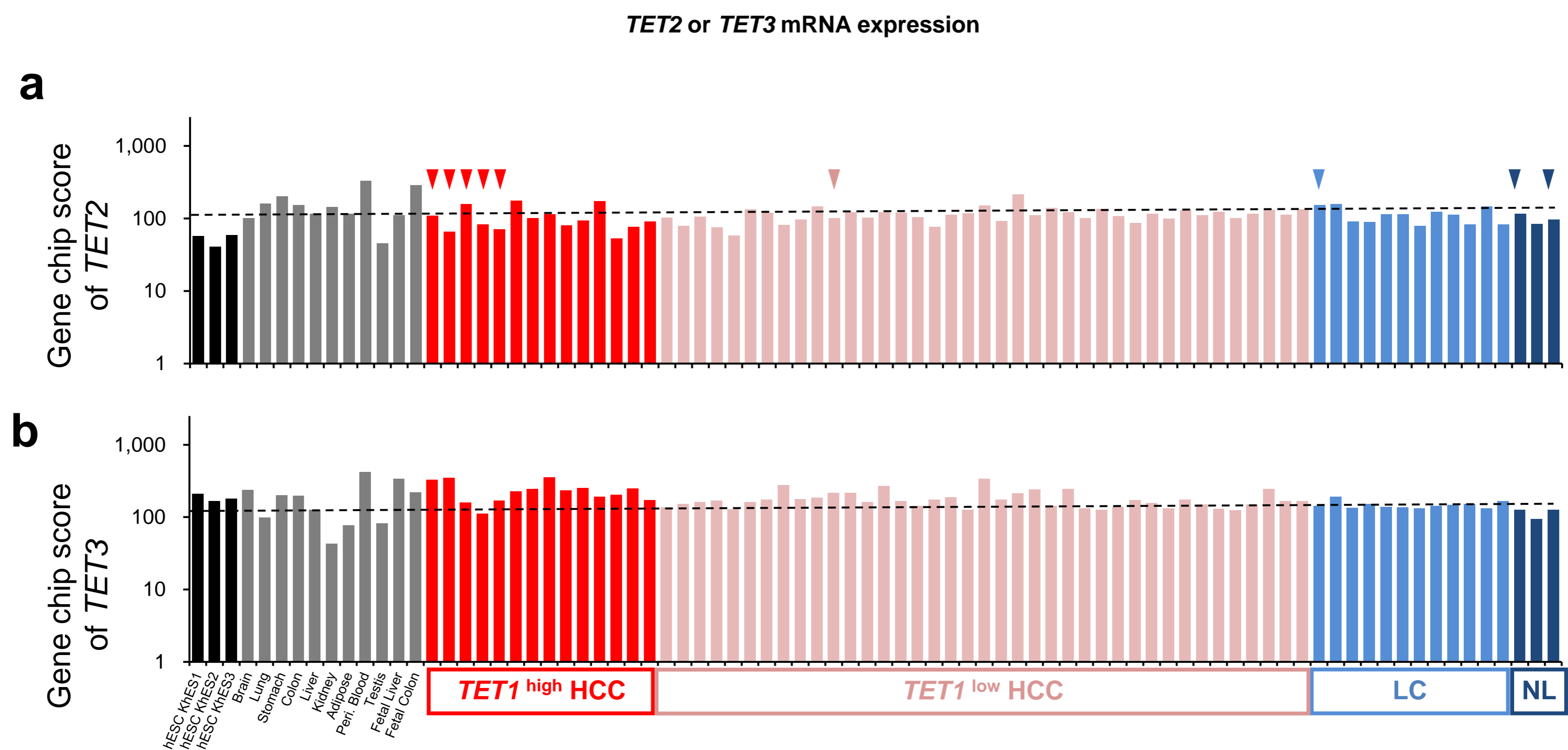


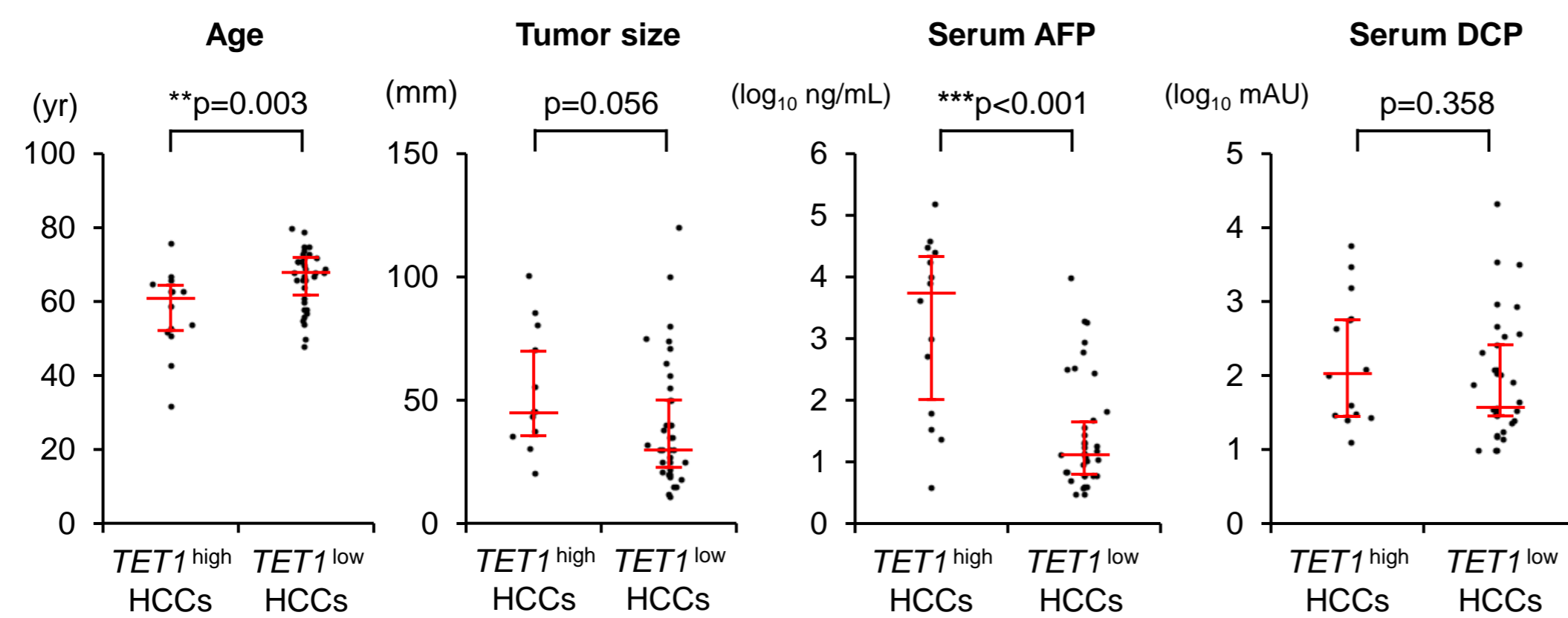
# Figure S1



## Figure S1 | mRNA expression level of *TET2* and *TET3*.

*TET2* (a) and *TET3* (b) expression level among hESCs, somatic tissues and HCCs in expression microarray (U133 plus 2.0, Affymetrix). *TET2*, Gene Name- KIAA1546, Probe Name: 227624at. Lower: *TET3*, Gene Name: hCG\_40378, Probe Name: 235542at. hESC, human embryonic stem cell ; HCC, hepatocellular carcinoma; LC, liver cirrhosis (corresponding non-cancerous liver); NL, normal liver.

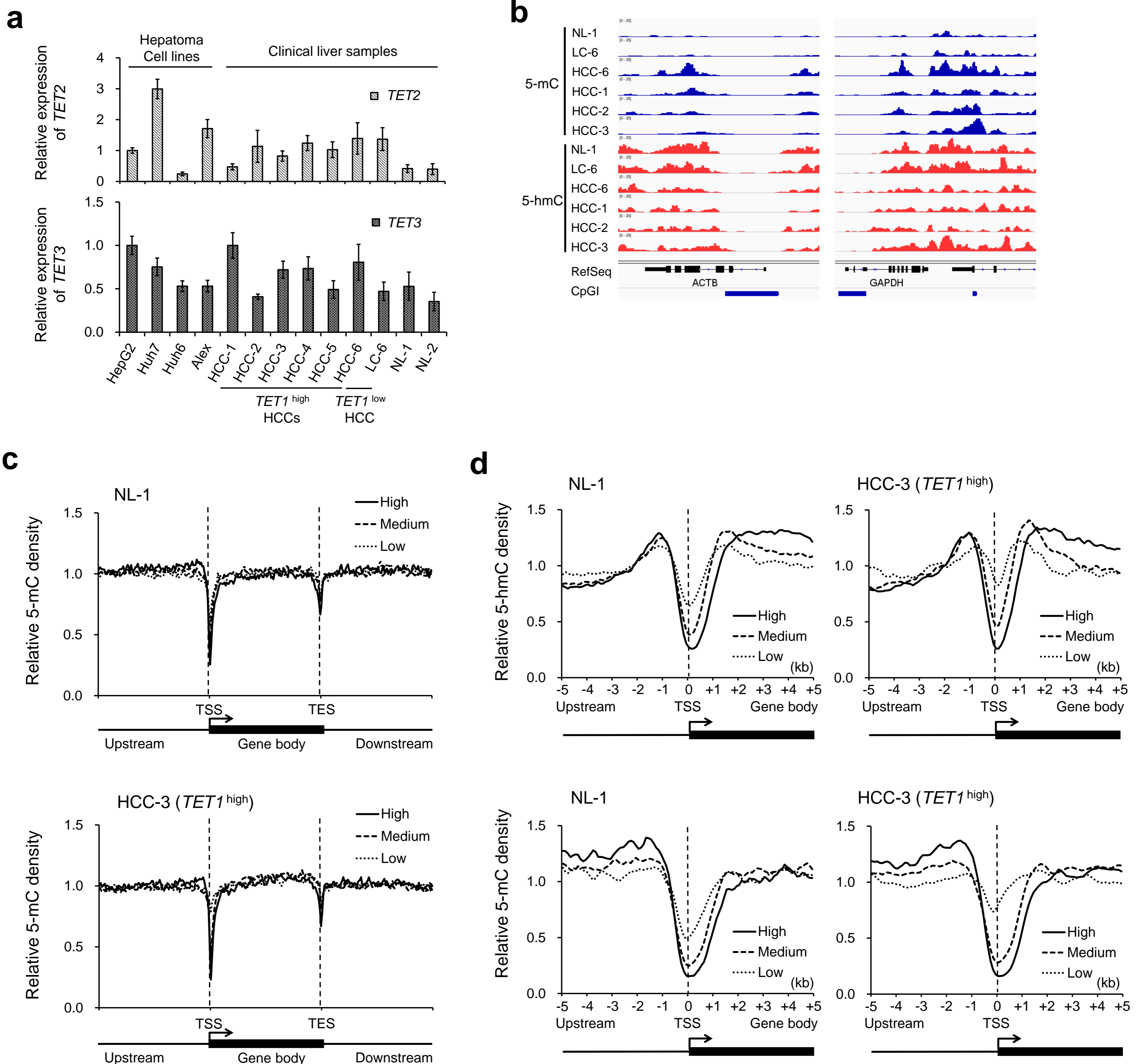
# Figure S2



**Figure S2 | Clinicopathological characteristics of HCC patients associated with *TET1* expression.**

Clinical characteristics of HCCs. AFP,  $\alpha$ -fetoprotein; DCP, des-gamma-carboxy prothrombin; Data are shown as median  $\pm$  quartile. p values are measured by Mann–Whitney U test. \*\*p < 0.01; \*\*\*p < 0.001.

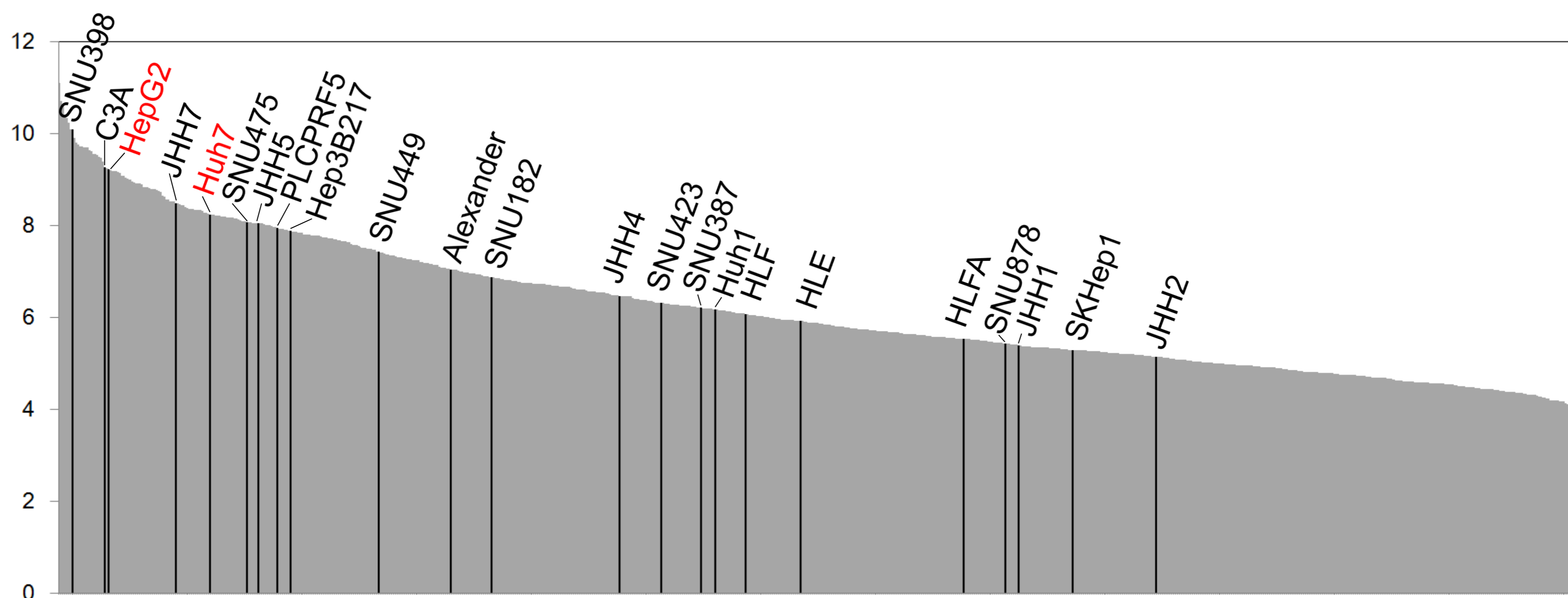
# Figure S3



## Figure S3 | 5mC and 5hmC profiles of clinical liver tissues.

(a) Relative *TET2* (upper) and *TET3* (lower) mRNA expression of hepatoma cells and clinical liver samples by RT-qPCR. *TET1*<sup>high</sup> HCCs, HCCs of high level *TET1* (GeneChip score > 50); *TET1*<sup>low</sup> HCC, HCC of low level *TET1* (GeneChip score < 50). Data are shown as mean ± SD from triplicate experiments. (b) Genome-wide profiling of 5mC (upper) and 5hmC (lower) for clinical liver tissues in *ACTB* and *GAPDH* locus. (c) 5mC distributions relative to human RefSeq gene position by MeDIP-seq at different expression levels in clinical liver samples. (d) 5hmC and 5mC distributions relative to each gene position around TSS (±5kb) by hmeDIP-seq (upper) and MeDIP-seq (lower) for clinical liver samples according to their gene expression levels. Expression level, High; GeneChip score > 250, Low; GeneChip score < 25, Medium; GeneChip score 25-250

# Figure S4



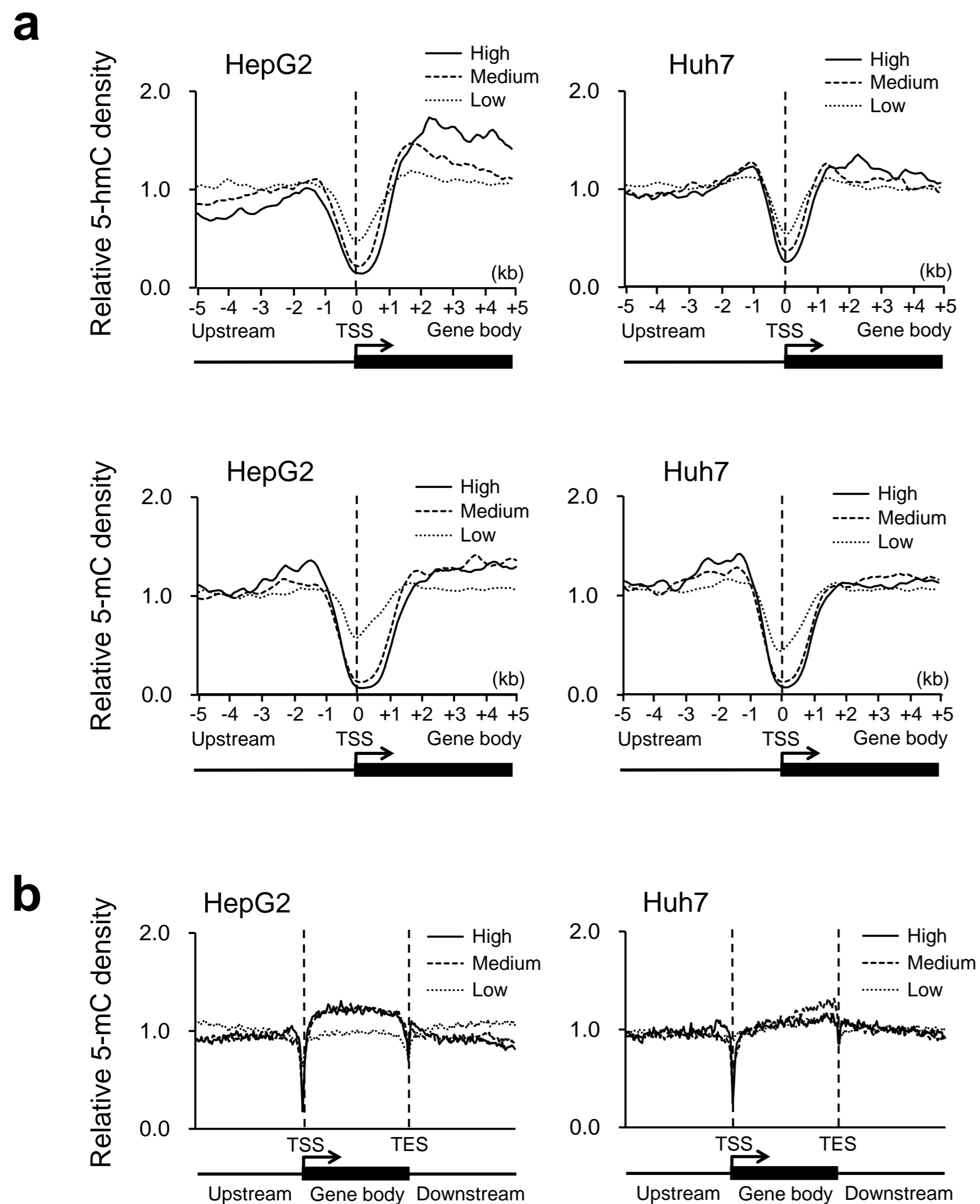
## Figure S4 | Expression level of TET1 in 807 CCLE cancer cell lines.

The 807 CCLE cancer cell lines are arranged in order of *TET1* expression level and shown in a gray bar graph.

The 23 liver cancer cell lines are shown in black, and their names are appended above.

CCLE: Broad Institute Cancer Cell Line Encyclopedia, <https://portals.broadinstitute.org/ccle>

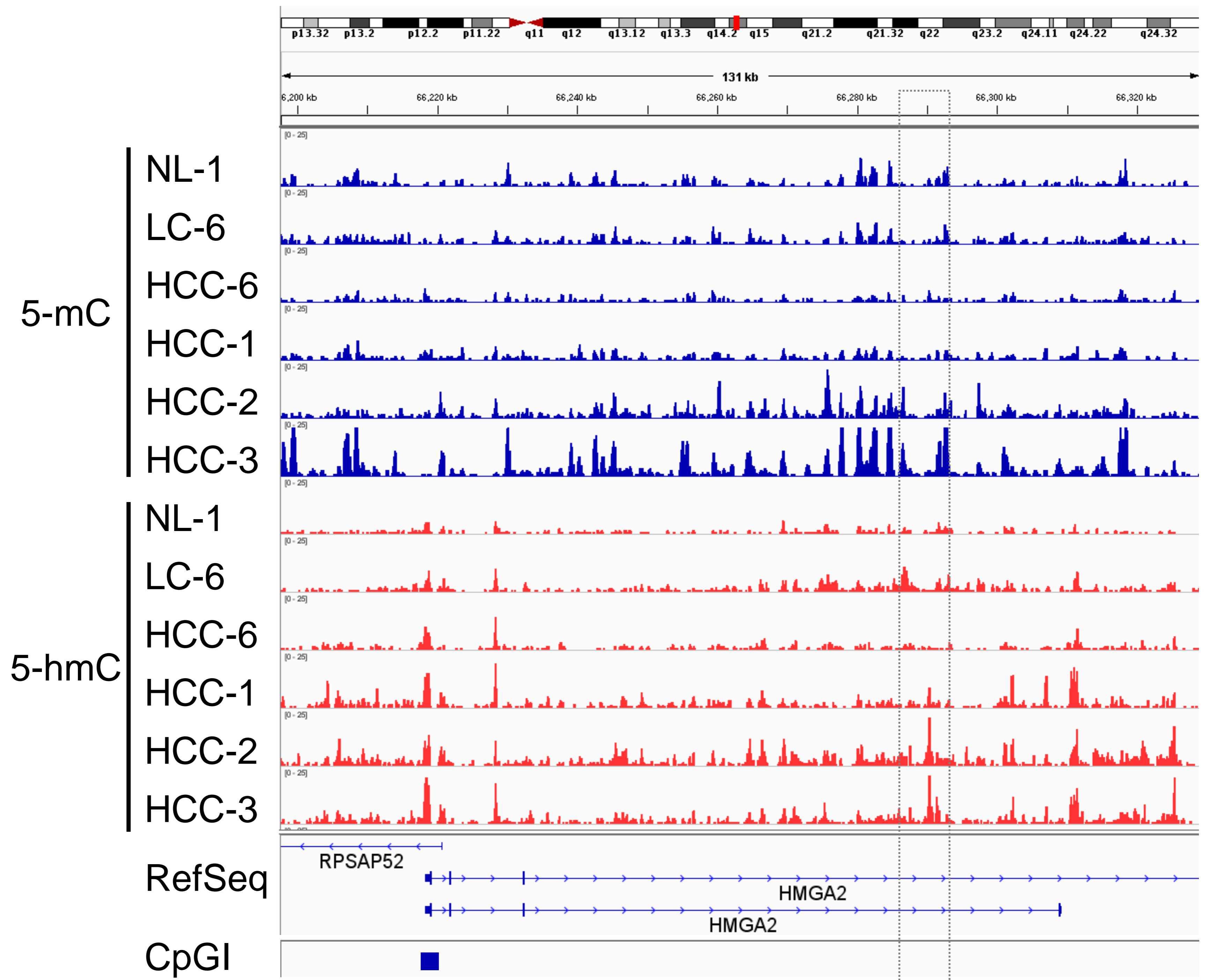
# Figure S5



## Figure S5 | 5hmC and 5mC distributions of liver cancer cell lines.

(a) 5hmC and 5mC distributions relative to each gene position around TSS by hmeDIP- (upper) and MeDIP-seq (lower) with regard to the expression levels in HepG2 and Huh7 cells. (b) 5mC distributions relative to human RefSeq gene position by MeDIP-seq according to the expression levels in HepG2 and Huh7. Expression level, High; GeneChip score > 250, Low; GeneChip score < 25, Medium; GeneChip score 25-250

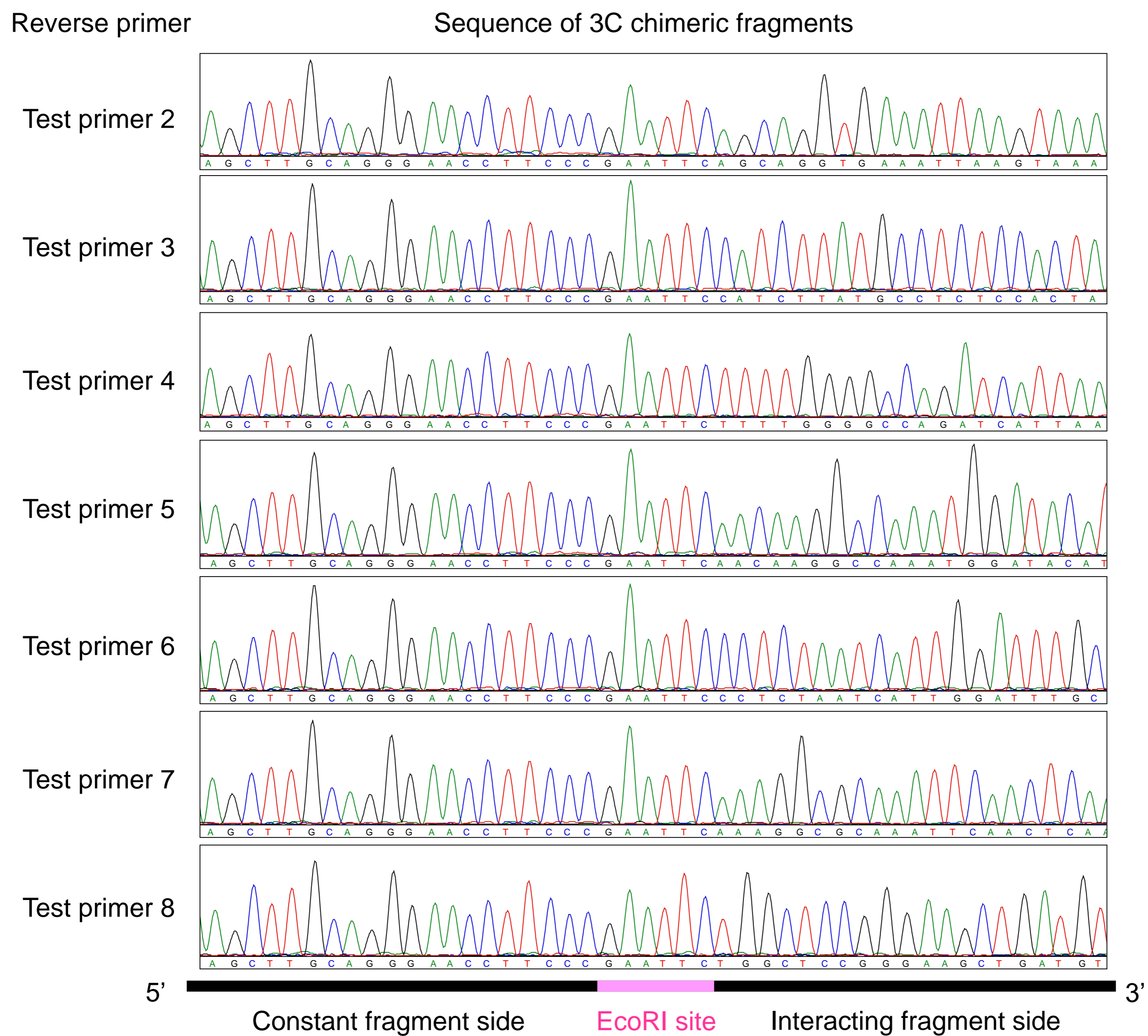
# Figure S6



**Figure S6 | 5mC and 5hmC mapping of HMGA2 locus for clinical liver tissues.**

5mC- (blue) and 5hmC- (red) enriched regions of HMGA2 locus are demonstrated using Integrative Genomics Viewer (IGV). The box with dotted lines indicates intragenic enhancer regions identified by CHIP-sequencing analysis of liver cancer cells

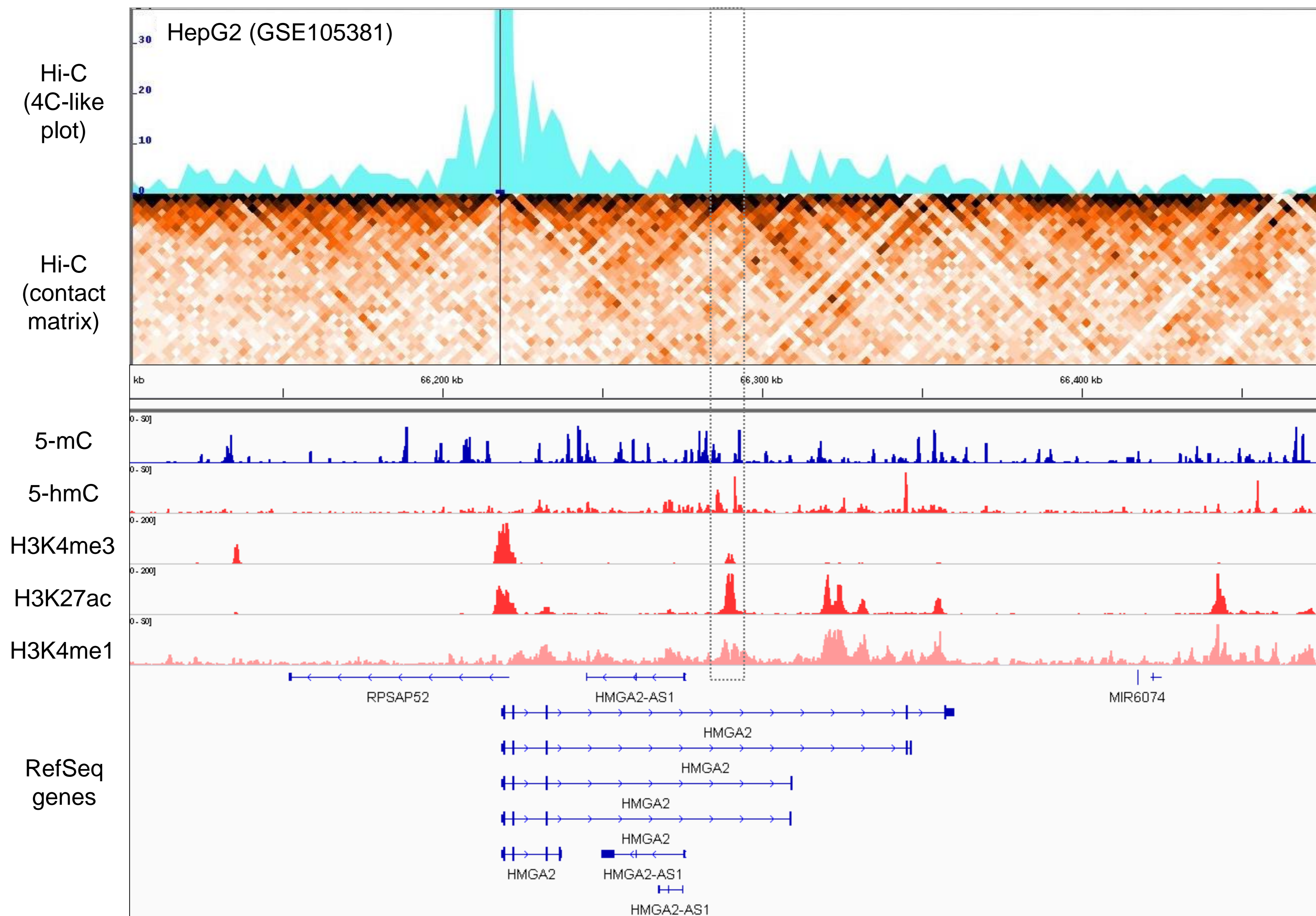
# Figure S7



**Figure S7 | Sanger Sequence of 3C chimeric fragments.**

The PCR product is cloned and sequenced with primers that anneal to the vector backbone.

# Figure S8

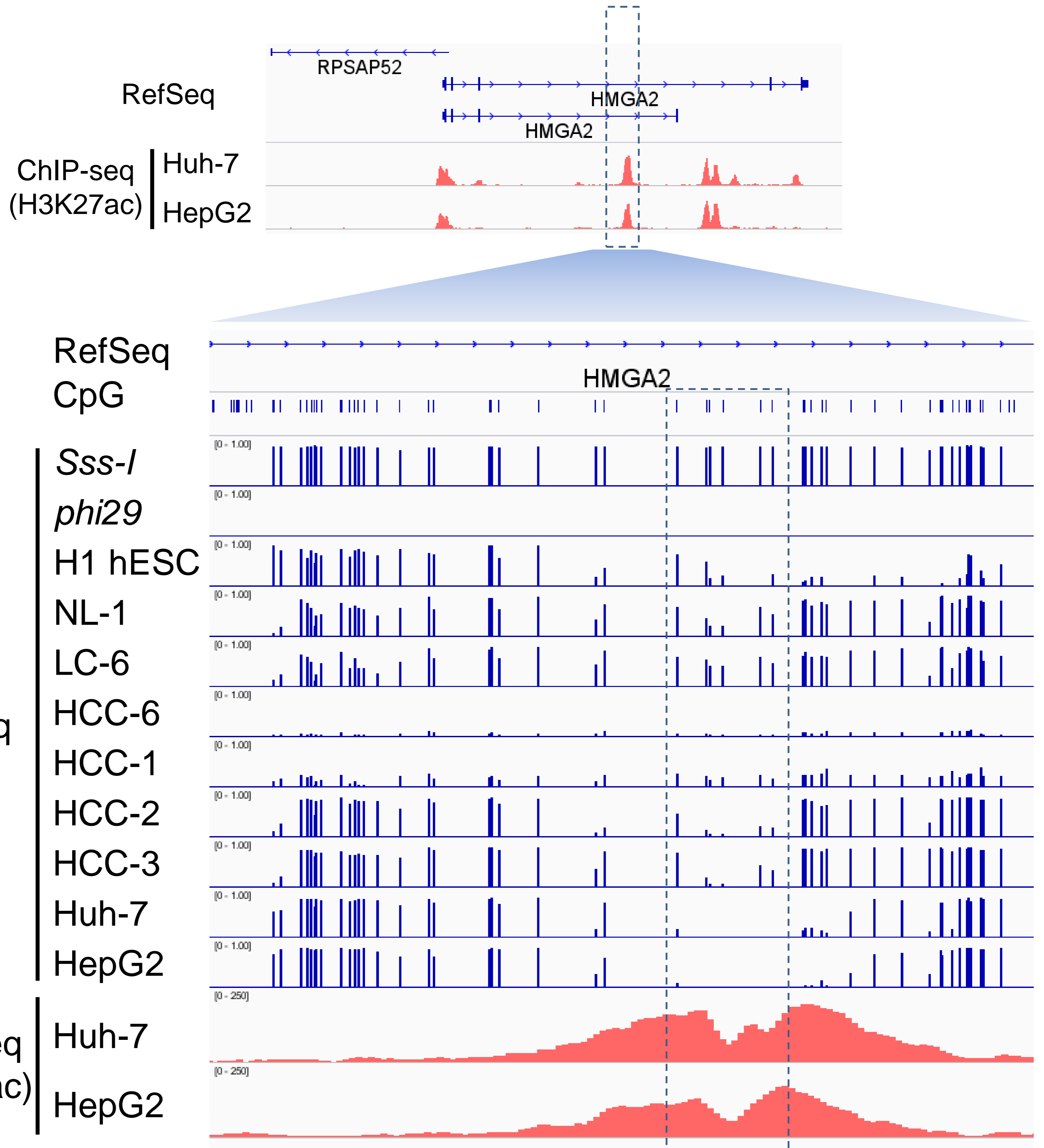


**Figure S8 | 5mC and 5hmC mapping of HMGA2 locus for clinical liver tissues.**

5mC- (blue) and 5hmC- (red) enriched regions of HMGA2 locus are demonstrated using Integrative Genomics Viewer (IGV). The box with dotted lines indicates intragenic enhancer regions identified by ChIP-seq analysis of liver cancer cells



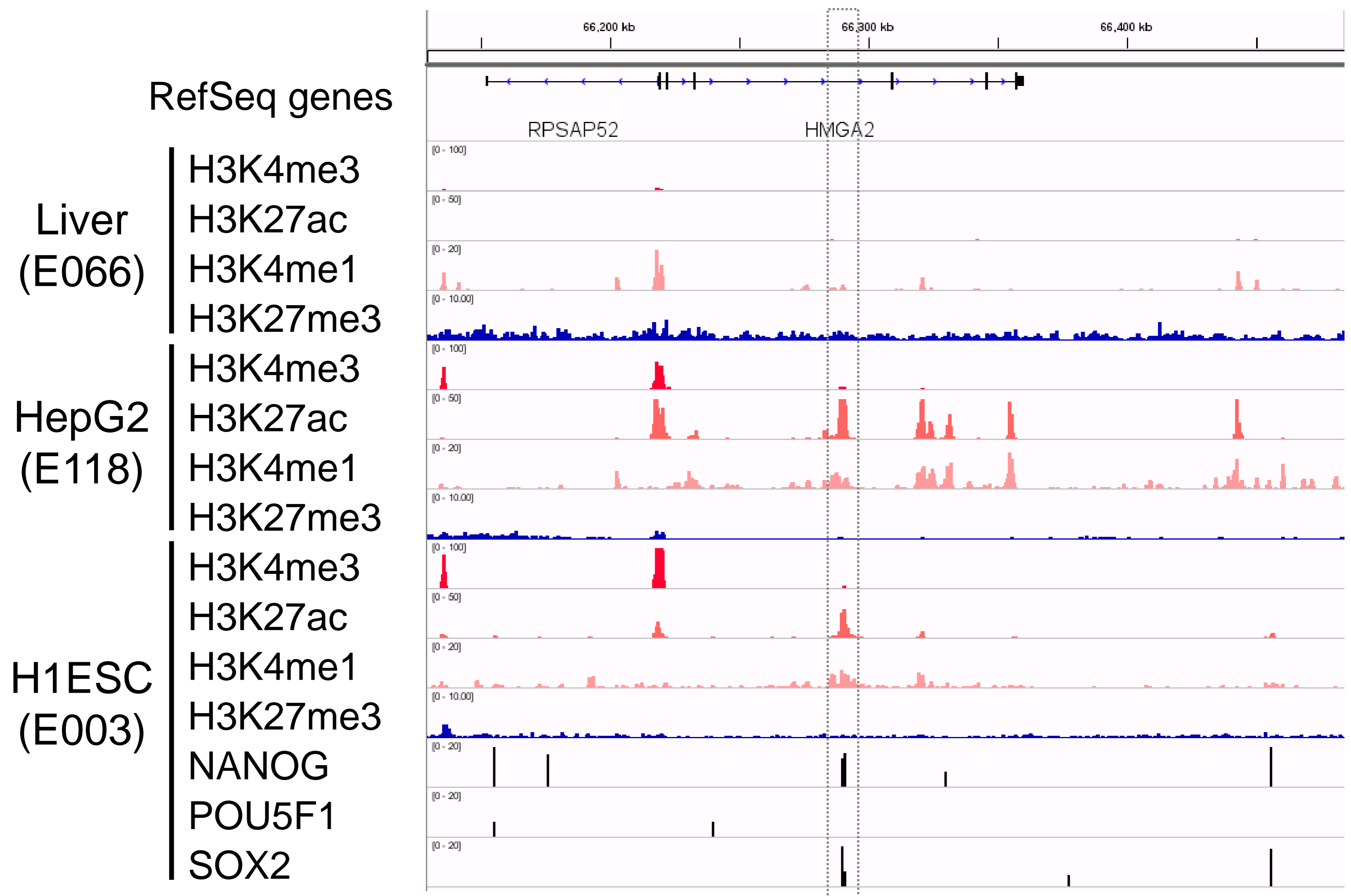
# Figure S9



**Figure S9 | 5mC and 5hmC mapping of HMGA2 locus for clinical liver tissues.**

5mC- (blue) and 5hmC- (red) enriched regions of HMGA2 locus are demonstrated using Integrative Genomics Viewer (IGV). The box with dotted lines indicates intragenic enhancer regions identified by ChIP-sequencing analysis of liver cancer cells

# Figure S10



**Figure S10 | Histone modification status 5mC and 5hmC mapping of HMGA2 locus for clinical liver tissues.** 5mC- (blue) and 5hmC- (red) enriched regions of HMGA2 locus are demonstrated using Integrative Genomics Viewer (IGV). The box with dotted lines indicates intragenic enhancer regions identified by ChIP-sequencing analysis of liver cancer cells

# Table S1

Name	Sequence
si <i>TET1</i> -1	AAUCAAGCGGAAGAAUAACUCAACA
si <i>TET1</i> -2	GGCUACACGAUUAGCUCCAAUUUUAU
si <i>TET1</i> -3	GGAAGCACUGUGGUUUUGUACCUUAA
si <i>HMGA2</i> -1	GAACCAACCGGUGAGCCCUCUCCUA
si <i>HMGA2</i> -2	GGAAACUGAAGAGACAUCCUCACAA

**Table S1 | Sense sequence of siRNAs targeting *TET1* and *HMGA2*.**

# Table S2

Name	Sequence
<i>TET1</i> forward	TGATGACAGAGGTTCTTGACATAAG
<i>TET1</i> reverse	CAGGTTGCACGGTCTCAGTGTT
<i>HMGA2</i> forward	CCCTGTTCCCTTTGAGTGTCTTCTAAC
<i>HMGA2</i> reverse	TTCAACAAGCAAGCGATTCA
<i>ACTB</i> forward	AGAAGGAGATCACTGCCCTGGCACC
<i>ACTB</i> reverse	CCTGCTTGCTGATCCACATCTGCTG

**Table S2 | Primer sequence for qRT-PCR.**

# Table S3

Name	Method	Pass filtered reads	%PF	Uniquely mapped reads	%Mapped
HepG2 H3K4me3	ChIP-seq	32,990,100	91.81%	29,102,178	88.21%
Huh7 H3K4me3	ChIP-seq	30,934,362	92.29%	26,583,453	85.94%
HepG2 H3K4me1	ChIP-seq	36,466,336	93.02%	30,116,105	82.59%
Huh7 H3K4me1	ChIP-seq	34,450,215	93.12%	29,201,716	84.76%
HepG2 H3K27ac	ChIP-seq	37,973,262	91.67%	33,063,351	87.07%
Huh7 H3K27ac	ChIP-seq	34,881,946	92.47%	30,167,015	86.48%
HepG2 siCTL	hmeDIP-seq	27,656,345	80.15%	15,458,915	55.90%
HepG2 siTET1	hmeDIP-seq	23,980,711	76.72%	14,563,272	60.73%
Huh7 siCTL	hmeDIP-seq	26,189,474	85.26%	14,596,475	55.73%
Huh7 siTET1	hmeDIP-seq	24,907,605	85.12%	13,199,788	53.00%
HCC-1	hmeDIP-seq	37,810,304	67.76%	25,894,022	68.48%
HCC-2	hmeDIP-seq	35,767,631	77.84%	22,767,529	63.65%
HCC-3	hmeDIP-seq	35,882,246	75.74%	24,785,685	69.08%
HCC-6	hmeDIP-seq	22,626,199	82.69%	14,352,956	63.44%
LC-6	hmeDIP-seq	24,539,865	79.37%	16,848,046	68.66%
NL-1	hmeDIP-seq	23,633,662	80.51%	14,220,659	60.17%
HepG2 siCTL	MeDIP-seq	36,090,315	76.35%	14,836,630	41.11%
HepG2 siTET1	MeDIP-seq	36,717,093	78.68%	15,832,332	43.12%
Huh7 siCTL	MeDIP-seq	34,350,531	77.03%	19,973,064	58.14%
Huh7 siTET1	MeDIP-seq	30,282,870	71.81%	18,790,623	62.05%
HCC-1	MeDIP-seq	23,021,818	62.42%	15,331,339	66.59%
HCC-2	MeDIP-seq	24,757,221	77.47%	16,987,547	68.62%
HCC-3	MeDIP-seq	27,363,962	79.30%	16,750,731	61.21%
HCC-6	MeDIP-seq	31,374,006	64.62%	20,666,264	65.87%
LC-6	MeDIP-seq	32,867,789	68.19%	17,932,016	54.56%
NL-1	MeDIP-seq	35,991,252	65.15%	14,859,379	41.29%

**Table S3 | Lists of ChIP-seq, hmeDIP-seq, and MeDIP-seq performed in this study.**

# Table S4

Name	Strand	Sequence
HMGA2_enh1 forward	top	TAGGGAAAATATTGAATTTGGTTTG
HMGA2_enh1 reverse	top	ACACAAAATAACTCTTTCACCAAAAATACT
HMGA2_enh2 forward	top	TTTGGTGAAAGAGTTATTTGTGTTG
HMGA2_enh2 reverse	top	TTATACTTAAAATCAACCCCTTCC
HMGA2_enh3 forward	top	GGAAGGGGGTTGATTTTAAGTATAA
HMGA2_enh3 reverse	top	CATACAAAACCTTATCTCCCCTTC
HMGA2_enh4 forward	top	TATGAAGGAGGGTGAAATTATAG
HMGA2_enh4 reverse	top	CTATTCTACATCCATTTACAAATACT
HMGA2_enh5 forward	top	TATTAGTATTTGTAAATGGATGTAGA
HMGA2_enh5 reverse	top	TAAAACCTCATTTTAACCTCTTTACCA
HMGA2_enh6 forward	top	GGTTTTGAATGTATAGTTTTGATTA
HMGA2_enh6 reverse	top	CTTAAATTTAACCATATAAAAATAATCC
HMGA2_enh7 forward	top	GTAAAGTTGGTTATAAATAGTGATG
HMGA2_enh7 reverse	top	CTATTCACTAACATTTTTCAAAACT
HMGA2_enh8 forward	top	AAATTTTTTGAGAAAAATGAGAAATTTAGA
HMGA2_enh8 reverse	top	ACTAAAACCCCTAACTAACATATAACA
HMGA2_enh8 forward	top	GTATGTATATTTGGATAATGGTG
HMGA2_enh8 reverse	top	TCCTTTATCCTACTAAATTTCTTA
HMGA2_enh9 forward	top	TGAAAAGGAATGTTTTAAGGAGG
HMGA2_enh9 reverse	top	AATCCCTCTAATAATATTACCCAA
HMGA2_enh10 forward	bottom	TAACATTACCAAAAAAATCAC
HMGA2_enh10 reverse	bottom	TAGAAGATTTGTTTTTAGAAATGG
HMGA2_enh11 forward	top	AGAAGTAGATTTTTTGATTTTTG
HMGA2_enh11 reverse	top	CCAAACTCATTCTCACTCAACTAC
HMGA2_enh12 forward	top	TTGGTTGTGTTATTAATGTTAGATTG
HMGA2_enh12 reverse	top	TAAAATAATAAATACCTCATTTACCC
HMGA2_enh13 forward	top	TATTGATTATAGTTATTTTGTGG
HMGA2_enh13 reverse	top	ATCCTCCCTTAACCTAACATTCC
HMGA2_enh14 forward	top	GAGGATGGGTGAAAAGTAATG
HMGA2_enh14 reverse	top	TAAAATTTTACTCTTATCATCCAAAC
HMGA2_enh15 forward	top	GTATTTTAGTTTGGATGATAAGAG
HMGA2_enh15 reverse	top	AATAAAAATTTAAAAACTATTAACAAC

**Table S4 | Primer sequence for 3C-qPCR.**

# Table S5

Name	Site	Sequence
Constant primer	(1)	GTCTCCGATGTGTTTGCTCTC
Test primer 2	2	ATGTGCTGGTGCATTAAGTAAG
Test primer 3	3	TGGGCCTGGGTGATAGTG
Test primer 4	4	CATTCTGTAAAGCTAATTGCACAAC
Test primer 5	5	GTGGTTTGTTCCTTTTCGCATGTATC
Test primer 6	6	AGCTTTGTGATCCCTCTGGTAATG
Test primer 7	7	GTGCTGAAAGGATAAAGGTGAATAG
Test primer 8	8	TGGCTCACTTCATCCACAACCTA
TaqMan probe		ACTGGTGCTAGTGGTCCCAC

**Table S5 | Primer sequence for multiplex targeted sequence of BS-treated amplicons.**