Table S1: Lists of downloaded data sources with links.

Table S2: Gene set enrichment.

S2A: Promoters that exhibit changes from Bivalent (TssBiv, BivFlnk, EnhBiv), PC-Repressed (RepPC, ReprPCWk), Heterochromatin, or Quiescent to an active (TssA, TssaFlnk, TXFlnk, Tx, TxWk, EnhG, Enh, ZNF/Rpts) chromatin structure. S2B: Promoters that exhibit a changes from an active (TssA, TssaFlnk, TXFlnk, Tx, TxWk, EnhG, Enh, ZNF/Rpts) to a Bivalent (TssBiv, BivFlnk, EnhBiv), PC-Repressed (RepPC, ReprPCWk), Heterochromatin, or Quiescent chromatin structure. EnrichR enrichment analysis was perfomed on 113 newly expressed or 796 genes newly silenced genes using 25 gene set libraries provided by EnrichR (see Table). The results were filtered on abs(z-score) >1.9 and q-value <0.01

Table S3: DMRs between HSPC and BasoE

Table includes the location of the DMRs, their level of methylation and their overlap with chromHMM categories, promoters, CpGclusters, UMR and LMR in both HSPC and BasoE. For instance, column BasoE.UMR contains the number of bp overlap of the given DMR with UMR(s) from BasoE.

Supplementary figure legends

Figure S1: Density of transcription start sites (TSS) in HMDs and PMDs.

A: Bar plots showing the average TSS density as a function of PMD length in 250kb bins. A cutoff between short PMDs and long PMDs was selected at 250kb, indicated by the light and dark green colors, respectively.

B: Bar plots representing the density of TSS per million base pairs of HMDs, PMDs, as well as of short and long PMDs separately. In all cells, long PMDs are poor in TSS but HMDs and short PMDs are rich in TSS. In differentiated and transformed cells, this results in PMDs being generally poorer in TSS than HMDs because long PMDs represent a much larger fraction of the genome than short PMDs. By contrast, in SPCs PMDs are generally richer in TSS than HMDs because short PMDs are predominant since these cells contain virtually no long PMDs.

Figure S2: Promoter methylation and gene expression.

A: Correlation between CAGE and RNA-seq signal. The adjusted Pearson coefficient of correlation (adj. R^2) of $\log_{10}$-transformed transcripts per million (TPM) values is indicated. Histone genes were removed (because RNA-seq data did not include polyA- genes). As previously reported, a moderately strong correlation between the CAGE and expression signal measured by RNA-seq was observed for all cell types tested (r^2 between 0.32 and 0.52).

B: Most active promoter regions are CpG-rich. Bar plots illustrating expression of CpG-rich and CpG-poor promoters in 7 cell types. Between 89-92% of all active promoter regions were CpG-rich and approximately two-thirds (10,000 of

15,000) of the CpG-rich promoters were active in any given cell type. By contrast only about 3-4% of the 20,000 CpG-poor promoters were active (500-to 800 promoter regions depending on cell type).

C: Heat maps illustrating the chromatin states of HMDs, short PMDs and long PMDs in HSPC, differentiated, and transformed cells based on the 15-state chromHMM core model from the Roadmap Epigenome Project. In all cells, short and long-PMDs were composed in large majority of various amounts of quiescent, Polycomb-repressed and heterochromatin, but short PMDs also contained a significant amount of bivalent chromatin (TssBiv, BivFk, EnhBiv) in stem cells, and of active chromatin in transformed cells. HMDs are enriched in transcriptionally active states. Percent coverage of the indicated regions by the 15 states are represented in heat map colors.

Figure S3: Gene body methylation.

Meta-gene analysis illustrating gene body methylation as a function of expression and location in HMDs or PMDs. The average percentage of methylated CpG dinucleotides of aggregated signal from all genes with indicated coordinates relative to the transcription start (TSS) or end site (TES) is shown for silent genes (Q1) and for expressed genes in tertiles of expression (Q2-Q4). Flanking sequence from 10kb upstream to 10kb downstream of the gene body is also shown.

Figure S4: ChromHMM analysis of CpG-rich promoters.

Heat map illustrating the chromatin states of unmethylated (overlapping with UMR or LMR) and methylated (not overlapping with UMR or LMR) promoter regions as a function of expression and their location in HMD or PMDs. Silent promoters were defined as promoters associated with genes exhibiting no or very low RNA-seq (< 1 TPM) and CAGE signal; expressed promoters as promoters associated with genes exhibiting both an RNA-seq and CAGE signal. Numbers in each blue box represent the number of promoters in a given chromatin state. The fraction of the number of bases of promoters overlapping the indicated chromHMM state are shown.

In all cells, the majority of expressed genes was unmethylated, located in HMDs and in TssA or TssFlnk chromatin states. Unmethylated silent promoters were generally associated with active or bivalent chromatin states when located in HMDs, but were generally associated with Polycomb-repressed chromatin when in PMDs. Methylated promoters were mainly in the quiescent chromatin state.

Figure S5: DNA methylation and chromatin changes during erythroid differentiation in CpG-poor promoters.

A: Circos plots illustrating the changes in DNA methylation status of newly silenced (left) and newly expressed (right) CpG-poor promoters during differentiation from HSPC (top half of the circle) to BasoE (bottom half) as a function of expression levels. The bands connecting the half-circles indicate the redistribution of genes from HSPC into the indicated categories in BasoE. The outer circle represents the expression levels (Q1 (silent), dark blue; Q2, blue; Q3 orange; Q4, (highly expressed) red). The inner circle segments represent the methylation status of the promoters (grey: unmethylated (UMR-containing); purple: methylated (no-UMR)). The numbers on the outside indicate the number of

promoters according to their methylation status split by expression quartile. The plot illustrates that there is little change in methylation of these promoter regions during differentiation.

B: Circos plots illustrating the change in chromatin status of promoters as a function of their location in HMDs or PMDs and as a function of their change in expression during differentiation from HSPC to BasoE, (clockwise): newly silenced in BasoE, newly expressed in BasoE, expressed in both HSPC and BasoE and not expressed in either HSPC or BasoE during differentiation from HSPC to BasoE. The bands connecting the half-circles indicate the redistribution of genes from HSPC into the indicated categories in BasoE. The outer circle segments represent the location of the promoters in HMDs (H, brown), short-PMDs (S, light green) or long-PMDs (L, dark green) in BasoE. The inner circle represents the chromatin states. For clarity, the 15 chromHMM classes were summarized into 5 larger categories: active (TssA, TssaFlnk, TXFlnk, Tx, TxWk, EnhG, Enh, ZNF/Rpts; red), Bivalent (TssBiv, BivFlnk, EnhBiv; green), PC-Repressed (RepPC, ReprPCWk, blue), Heterochromatin (purple), and Quiescent (grey). The circle segments above the dotted line represent the methylation and chromatin status of promoters in HSPC, those below the dotted line the methylation and chromatin status of promoters in BasoE. The numbers on the outside indicate the number of promoters in each category.

C and D: Bar graphs illustrating the fraction of CpG-rich (C) or CpG-poor (D) promoters that exhibit chromatin change during differentiation from HSPC to BasoE classified by changes in expression. The numbers above each bar represent the number of promoters that exhibited a change in chromatin structure in BasoE as compared to HSPC. The number in parenthesis represents the total number of promoters in each chromatin category in HSPC. The y-axis represents the ratio of the number of promoters that exhibit a change over the total number of promoters in each category. The stars indicate the categories that contained less than ten promoters; the low number made generation of ratios uninformative. Most changes in chromatin states occurred in genes that were either silent in both cell types or newly silent in BasoE (blue and light blue bars).

E: Changes in size and edge location of DNA methylation canyons upon differentiation of HSPC into BasoE. Right: Bar graph illustrating the number of DNA methylation canyons according to their change in edge position in HSPC and BasoE (left) and scatter plot (right) illustrating the size changes of the DNA methylation canyons called in either HSPC or BasoE. The majority of canyons with changes in edge locations expanded but a few contracted, and canyon sizes generally became larger upon differentiation of HSPC into BasoE.

F: Circos plot illustrating the change of chromatin structure during differentiation of HSPC into BasoE in DNA methylation canyons (UMRs > 5kb) identified in HSPC. The graph is organized as in B.

Figure S6: Mutation spectra and Ts/Tv ratio are similar in a-DMRs and the rest of the genome.

Mutation spectra and Transition to Transversion ratio (Ts/Tv) in a-DMRs and in the rest of the genome were determined using SNPeffect. While the SNPs/kb ratio is considerably higher in a-DMRs, the mutation spectra and Ts/Tv ratio are very similar.

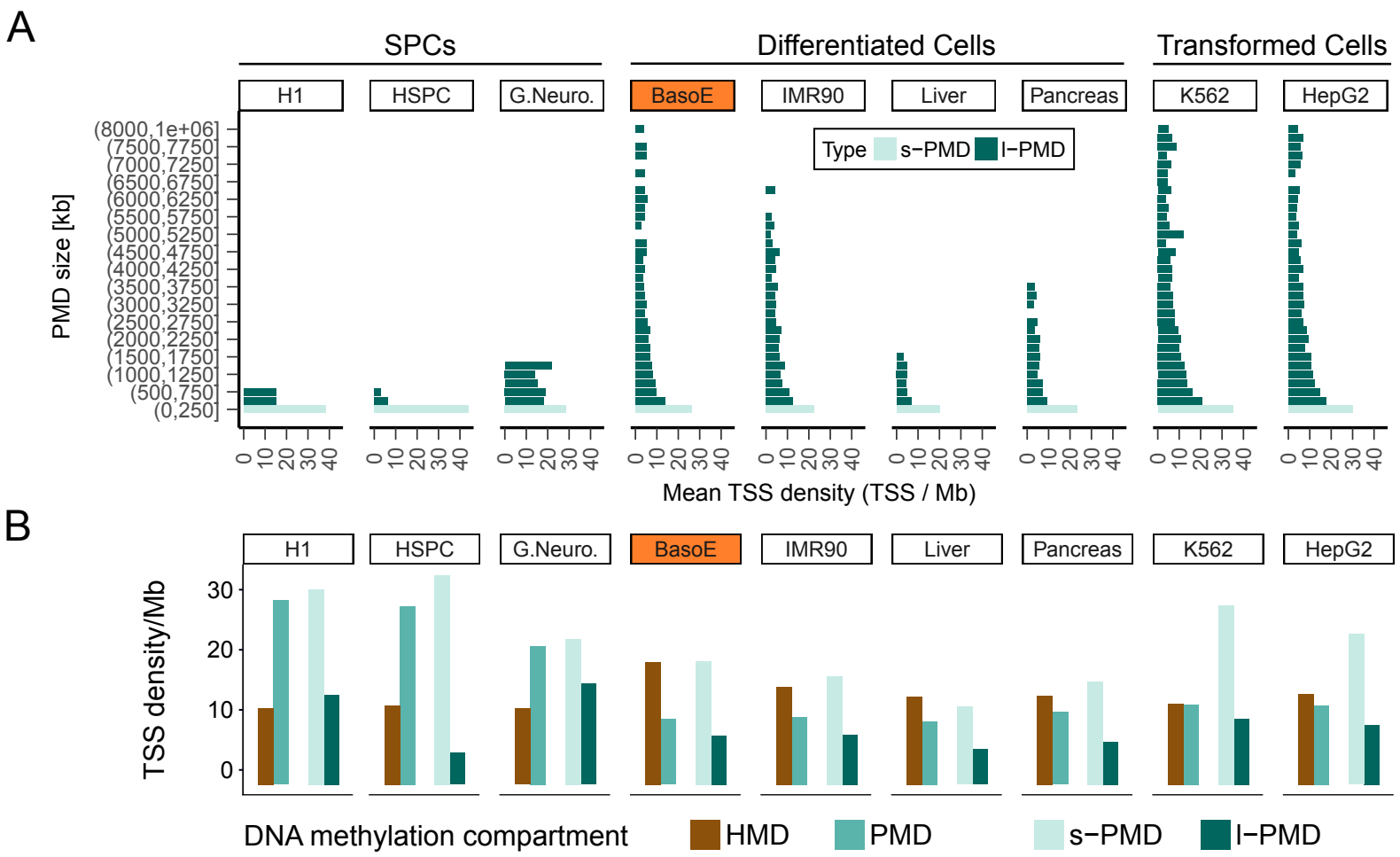
Figure S7: Reduced representation sequencing of cycling and G0/G1 blocked p51R mesenchymal cells.

A: Diagram summarizing the large fragment reduced representation approach. B: P51R cells were encapsulated in agarose plugs, DNA was extracted and restricted with PacI, and DNA between about 60-95 kb was extracted after Pulse-Field Gel Electrophoresis (PFGE). The picture illustrates an ethidium bromide stain of a gel after purification of the 60-95 kb fraction. C: GenPlay browser screenshot illustrating the result of large fragments reduced representation bisulfite sequencing. A two million base pair region is represented. Track 1 indicates Refseq genes; track 2 the 60-100kb predicted PacI fragments; tracks 3 and 5 the percent unmethylation of CpGs in cells blocked in G0/G1 and in cycling cells, respectively. Tracks 4 and 6 show called PMD and HMD regions. D: Plot Illustrating the correlation between the percent CpG methylation of each individual CpG measured in cells blocked in G0/G1 (y-axis) and in cycling cells (x-axis).

Cell line	Data type	WGBS data
BasoE	**WGBS data**	produced in house
IMR90	**WGBS data**	ftp://ftpuser3:s3qu3nc3@neomorph.salk.edu/mc/imr90_c_basecalls.tar.gz
Pancreas	**WGBS data**	http://egg2.wustl.edu/roadmap/data/byDataType/dnamethylation/WGBS/FractionalMethylation_bigwig/E098_WGBS_FractionalMethylation.bigwig and http://egg2.wustl.edu/roadmap/data/byDataType/dnamethylation/WGBS/ReadCoverage_bigwig/E098_WGBS_ReadCoverage.bigwig
Liver	**WGBS data**	http://egg2.wustl.edu/roadmap/data/byDataType/dnamethylation/WGBS/FractionalMethylation_bigwig/E066_WGBS_FractionalMethylation.bigwig and http://egg2.wustl.edu/roadmap/data/byDataType/dnamethylation/WGBS/ReadCoverage_bigwig/E066_WGBS_ReadCoverage.bigwig
K562	**WGBS data**	ftp://ftp-trace.ncbi.nlm.nih.gov/sra/sra-instant/reads/ByExp/sra/SRX/SRX840/SRX840166/SRR1754892/SRR1754892.sra
HepG2	**WGBS data**	https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSM1204463&format=file&file=GSM1204463_BiSeq_cpgMethylation_BioSam_1500_HepG2_304072.BiSeq.bed.gz
H1	**WGBS data**	ftp://ftpuser3:s3qu3nc3@neomorph.salk.edu/mc/h1_c_basecalls.tar.gz
HSPC	**WGBS data**	http://egg2.wustl.edu/roadmap/data/byDataType/dnamethylation/WGBS/FractionalMethylation_bigwig/E050_WGBS_FractionalMethylation.bigwig and http://egg2.wustl.edu/roadmap/data/byDataType/dnamethylation/WGBS/ReadCoverage_bigwig/E050_WGBS_ReadCoverage.bigwig
G. Neurosphere	**WGBS data**	http://egg2.wustl.edu/roadmap/data/byDataType/dnamethylation/WGBS/FractionalMethylation_bigwig/E054_WGBS_FractionalMethylation.bigwig and http://egg2.wustl.edu/roadmap/data/byDataType/dnamethylation/WGBS/ReadCoverage_bigwig/E054_WGBS_ReadCoverage.bigwig
Cell line	**Data type**	**chromHMM data**
BasoE	**chromHMM**	chromHMM according to 15_coreMarks model was generated using chromHMM and the same marks obtained from the sources below and lifted over from hg18 to hg19 with UCSC liftOver: http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/data/CD36-H3K4me1.bed.gz http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/data/CD36-H3K4me3.bed.gz http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/data/CD36-H3K9me3.bed.gz http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/data/CD36-H3K27me3.bed.gz http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/data/CD36-H3K36me3.bed.gz
IMR90	**chromHMM**	http://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/E017_15_coreMarks_segments.bed
Pancreas	**chromHMM**	http://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/E098_15_coreMarks_segments.bed
Liver	**chromHMM**	http://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/E066_15_coreMarks_segments.bed
K562	**chromHMM**	http://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/E118_15_coreMarks_segments.bed
HepG2	**chromHMM**	http://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/E118_15_coreMarks_segments.bed
H1	**chromHMM**	http://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/E003_15_coreMarks_segments.bed
HSPC	**chromHMM**	http://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/E050_15_coreMarks_segments.bed
G. Neurosphere	**chromHMM**	http://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/E054_15_coreMarks_segments.bed
Cell line	**Data type**	**RNA-seq data**
BasoE	**RNA-seq**	Produced in house
IMR90	**RNA-seq**	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeCshlLongRnaSeq/wgEncodeCshlLongRnaSeqImr90CellPapFastqRd1Rep1.fastq.gz, http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeCshlLongRnaSeq/wgEncodeCshlLongRnaSeqImr90CellPapFastqRd1Rep2.fastq.gz, http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeCshlLongRnaSeq/wgEncodeCshlLongRnaSeqImr90CellPapFastqRd2Rep1.fastq.gz, http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeCshlLongRnaSeq/wgEncodeCshlLongRnaSeqImr90CellPapFastqRd2Rep2.fastq.gz
Pancreas	**RNA-seq**	RNA-seq from pancreas was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM1120309 (SRX263865)
Liver	**RNA-seq**	
K562	**RNA-seq**	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeCshlLongRnaSeq/wgEncodeCshlLongRnaSeqK562CellPapFastqRd1Rep1.fastq.gz (etc.)
HepG2	**RNA-seq**	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeCshlLongRnaSeq/wgEncodeCshlLongRnaSeqHepg2CellPapFastqRd1Rep1.fastq.gz (etc.)
H1	**RNA-seq**	wgEncodeCshlLongRnaSeqH1hescCellPapFastqRd1Rep1.fastq.gz (etc.)
HSPC	**RNA-seq**	RNA-seq from HSPC was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM909310 (SRX135562)

G. Neurosphere	**RNA-seq**	Not available
Cell line	**Data type**	**CAGE data**
BasoE	**CAGE**	http://fantom.gsc.riken.jp/5/datafiles/phase2.0/basic/human.primary_cell.hCAGE/CD34%2520cells%2520differentiated%2520to%2520erythrocyte%2520lineage%252c%2520biol_%2520rep1.CNhs13552.11931-125I5.hg19.ctss.bed.gz and CD34%20cells%20differentiated%20to%20erythrocyte%20lineage%2c%20biol_%20rep2
IMR90	**CAGE**	extracted from R data package ENCODEprojectCAGE_1.0/ENCODEprojectCAGE/data/IMR90.RData
Pancreas	**CAGE**	http://fantom.gsc.riken.jp/5/datafiles/phase2.0/basic/human.tissue.hCAGE/pancreas%252c%2520adult%252c%2520donor1.CNhs11756.10049-101G4.hg19.ctss.bed.gz
Liver	**CAGE**	http://fantom.gsc.riken.jp/5/datafiles/phase2.0/basic/human.tissue.hCAGE/liver%252c%2520adult%252c%2520pool1.CNhs10624.10018-101C9.hg19.ctss.bed.gz
K562	**CAGE**	extracted from R data package ENCODEprojectCAGE_1.0/ENCODEprojectCAGE/data/K562.RData
HepG2	**CAGE**	extracted from R data package ENCODEprojectCAGE_1.0/ENCODEprojectCAGE/data/HepG2.RData
H1	**CAGE**	extracted from R data package ENCODEprojectCAGE_1.0/ENCODEprojectCAGE/data/H1-hESC.RData
HSPC	**CAGE**	http://fantom.gsc.riken.jp/5/datafiles/phase2.0/basic/human.primary_cell.hCAGE/CD34%252b%2520stem%2520cells%2520-%2520adult%2520bone%2520marrow%2520derived%252c%2520donor1%252c%2520tech_rep1.CNhs12588.12225-129F2.hg19.ctss.bed.gz and http://fantom.gsc.riken.jp/5/datafiles/phase2.0/basic/human.primary_cell.LQhCAGE/CD34%252b%2520stem%2520cells%2520-%2520adult%2520bone%2520marrow%2520derived%252c%2520donor1%252c%2520tech_rep2.CNhs12553.12225-129F2.hg19.ctss.bed.gz
Cell line	**Data type**	**DNase HSS data**
IMR90	**DNase HSS**	https://www.encodeproject.org/files/ENCFF001UWF/@@download/ENCFF001UWF.bed.gz
K562	**DNase HSS**	https://www.encodeproject.org/files/ENCFF941ITD/@@download/ENCFF941ITD.bed.gz
HepG2	**DNase HSS**	https://www.encodeproject.org/files/ENCFF873IZM/@@download/ENCFF873IZM.bed.gz
H1	**DNase HSS**	https://www.encodeproject.org/files/ENCFF001UVM/@@download/ENCFF001UVM.bed.gz
Cell line	**Data type**	**Replication timing**
BasoE	**Replication timing**	Produced in house
H1	**Replication timing**	Produced in house
IMR90	**Replication timing**	Gilbert's lab
HepG2	**Replication timing**	Gilbert's lab

Table S1: Data sources


A
SPCs
Differentiated Cells
Transformed Cells
H1
HSPC
G.Neuro.
BasoE
IMR90
Liver
Pancreas
K562
HepG2
Type s−PMD l−PMD
PMD size [kb]
(8000,1e+06]
(7500,7750]
(7000,7250]
(6500,6750]
(6000,6250]
(5500,5750]
(5000,5250]
(4500,4750]
(4000,4250]
(3500,3750]
(3000,3250]
(2500,2750]
(2000,2250]
(1500,1750]
(1000,1250]
(500,750]
(0,250]
0 10 20 30 40
Mean TSS density (TSS / Mb)

B
H1
HSPC
G.Neuro.
BasoE
IMR90
Liver
Pancreas
K562
HepG2
TSS density/Mb
0
10
20
30
DNA methylation compartment
HMD
PMD
s−PMD
l−PMD

Figure S1

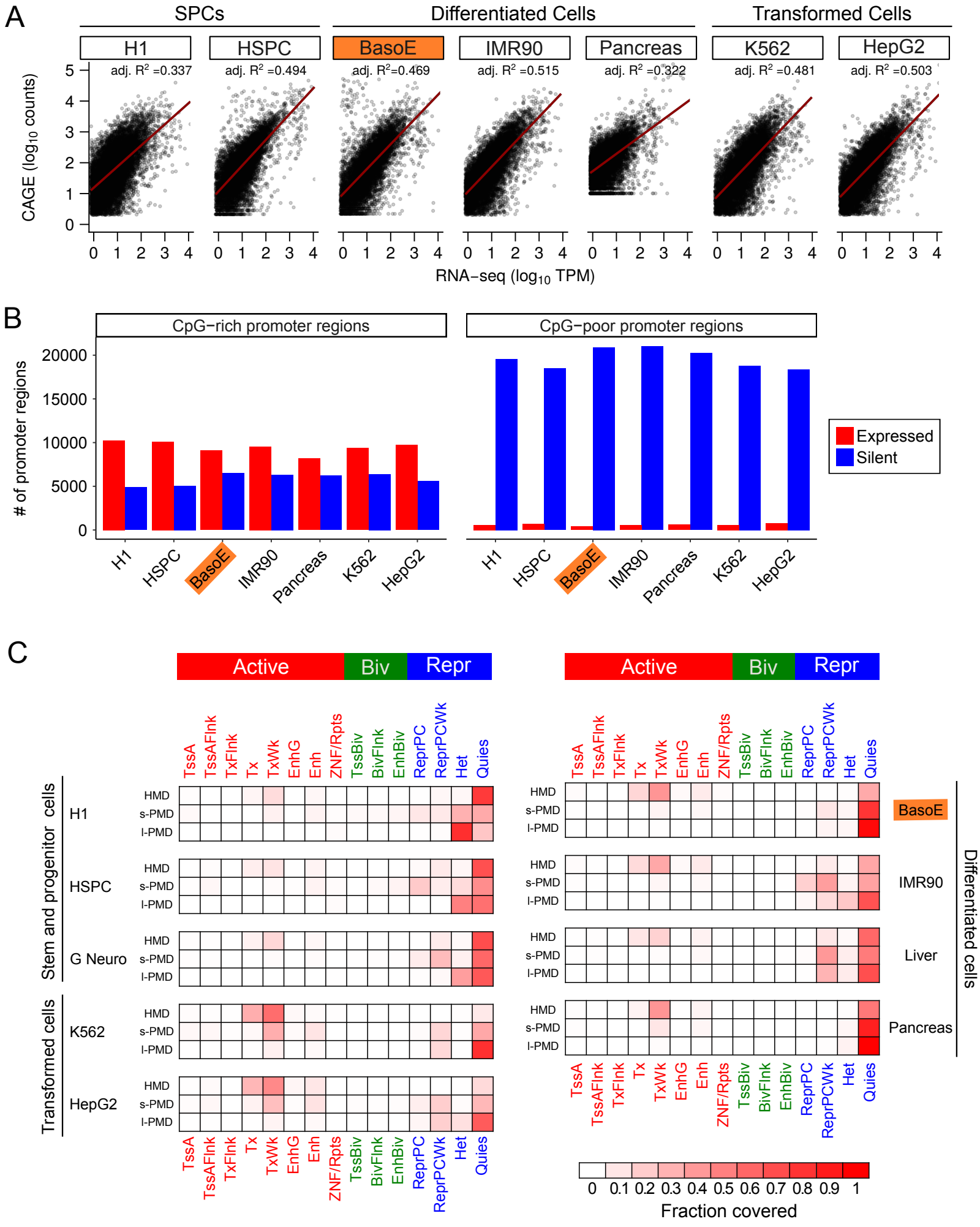

A
SPCs
Differentiated Cells
Transformed Cells
H1
HSPC
BasoE
IMR90
Pancreas
K562
HepG2
adj. R² =0.337
adj. R² =0.494
adj. R²=0.469
adj. R² =0.515
adj. R² =0.322
adj. R² =0.481
adj. R² =0.503
CAGE (log10 counts)
RNA−seq (log10 TPM)
B
CpG−rich promoter regions
CpG−poor promoter regions
# of promoter regions
Expressed
Silent
C
Active
Biv
Repr
TssA
TssAFlnk
TxFlnk
Tx
TxWk
EnhG
Enh
ZNF/Rpts
TssBiv
BivFlnk
EnhBiv
ReprPC
ReprPCWk
Het
Quies
HMD
s-PMD
l-PMD
Stem and progenitor cells
Transformed cells
Differentiated cells
G Neuro
Liver
Fraction covered

Figure S2

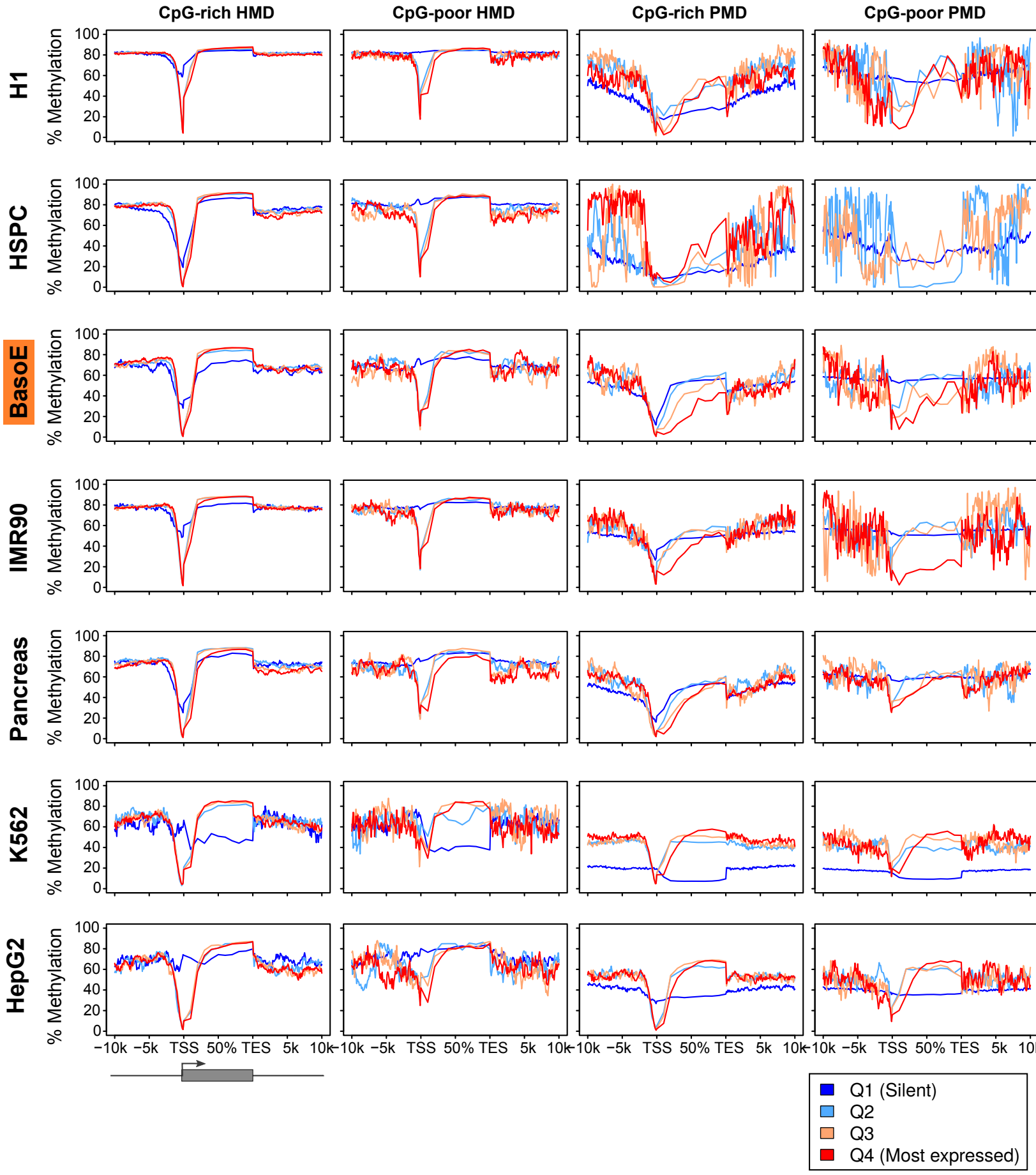

CpG-rich HMD
CpG-poor HMD
CpG-rich PMD
CpG-poor PMD
H1
HSPC
BasoE
IMR90
Pancreas
K562
HepG2
% Methylation
100
80
60
40
20
0
−10k
−5k
TSS
50% TES
5k
10k
Q1 (Silent)
Q2
Q3
Q4 (Most expressed)

Figure S3

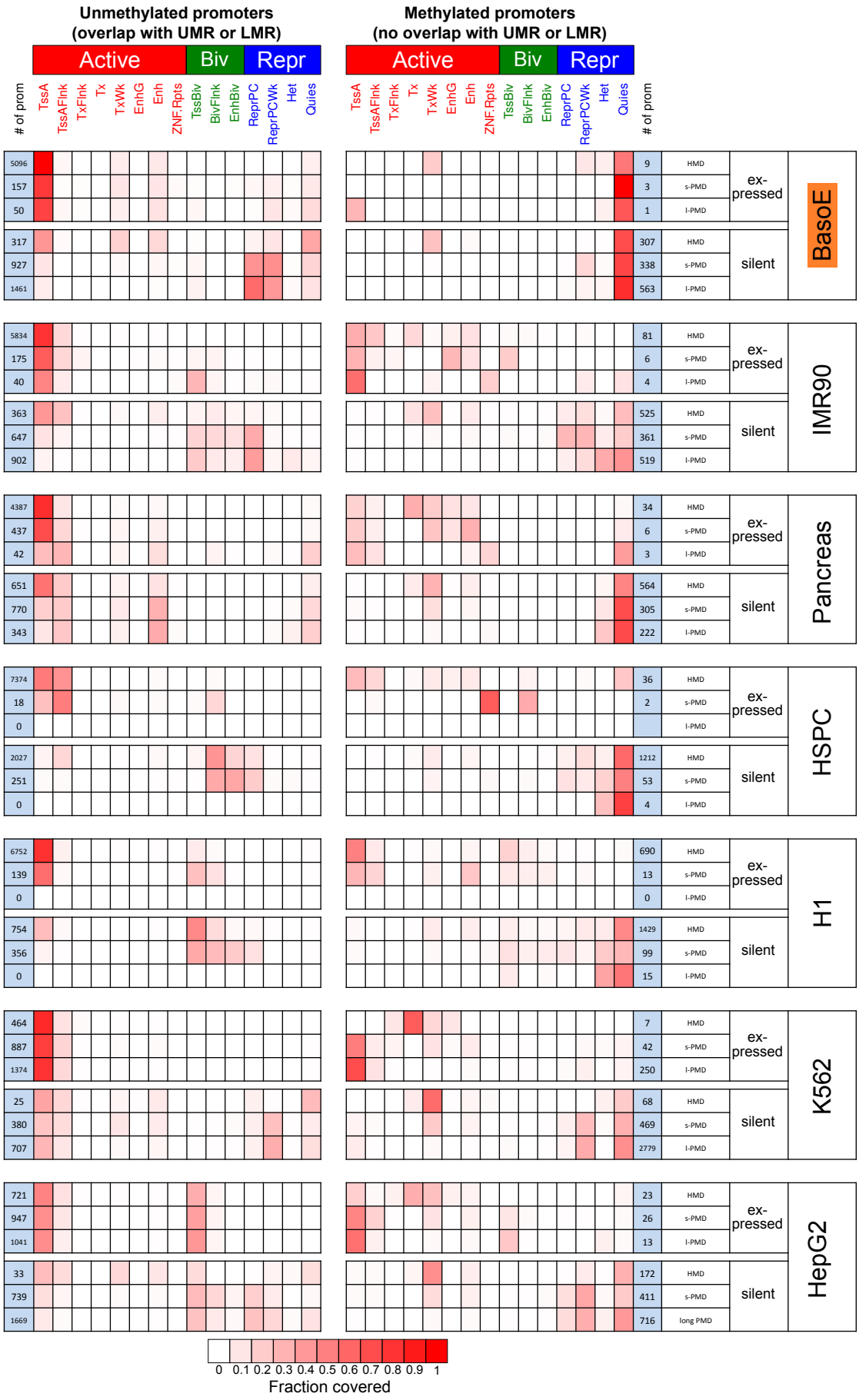

Unmethylated promoters
(overlap with UMR or LMR)
Methylated promoters
(no overlap with UMR or LMR)
Active
Biv
Repr
# of prom
TssA
TssAFlnk
TxFlnk
Tx
TxWk
EnhG
Enh
ZNF.Rpts
TssBiv
BivFlnk
EnhBiv
ReprPC
ReprPCWk
Het
Quies
HMD
s-PMD
l-PMD
long PMD
ex-pressed
silent
BasoE
IMR90
Pancreas
HSPC
H1
K562
HepG2
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Fraction covered

Figure S4

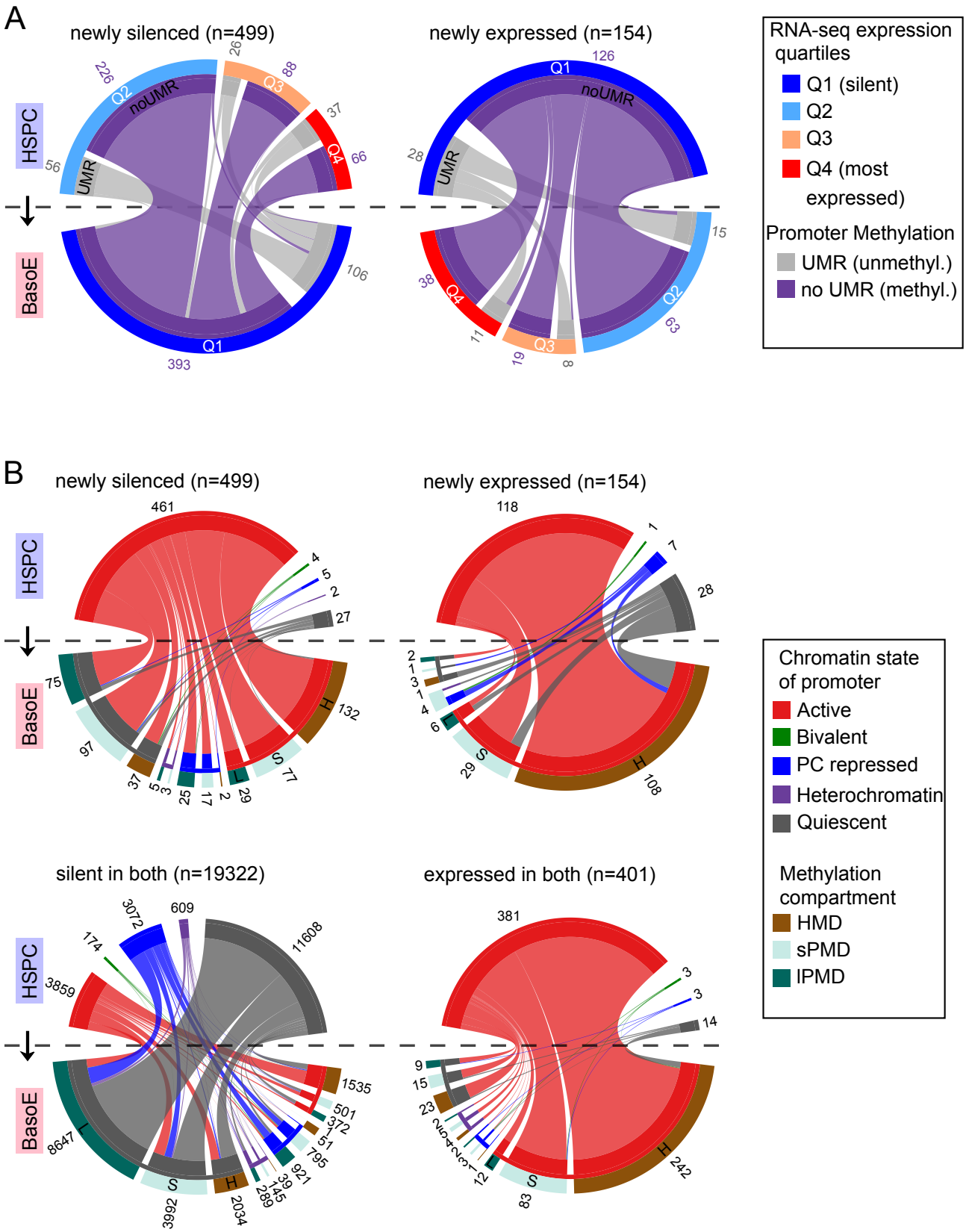
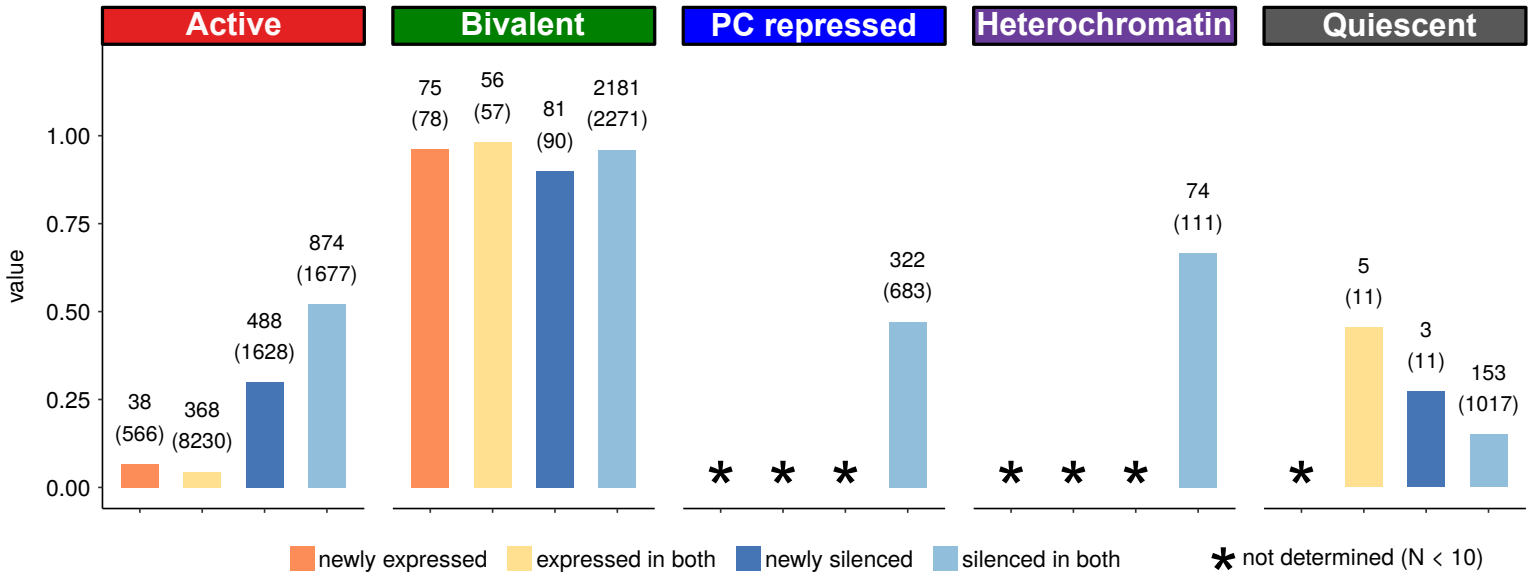

A
newly silenced (n=499)
newly expressed (n=154)
HSPC
BasoE
Q1
Q2
Q3
Q4
noUMR
UMR
RNA-seq expression quartiles
Q1 (silent)
Q2
Q3
Q4 (most expressed)
Promoter Methylation
UMR (unmethyl.)
no UMR (methyl.)
B
newly silenced (n=499)
newly expressed (n=154)
silent in both (n=19322)
expressed in both (n=401)
Chromatin state of promoter
Active
Bivalent
PC repressed
Heterochromatin
Quiescent
Methylation compartment
HMD
sPMD
lPMD

Figure S5

C


Fraction of CpG-rich promoters with altered chromatin state between HSPC and BasoE
Active
Bivalent
PC repressed
Heterochromatin
Quiescent
value
1.00
0.75
0.50
0.25
0.00
38 (566)
368 (8230)
488 (1628)
874 (1677)
75 (78)
56 (57)
81 (90)
2181 (2271)
322 (683)
74 (111)
5 (11)
3 (11)
153 (1017)
newly expressed
expressed in both
newly silenced
silenced in both
not determined (N < 10)

D

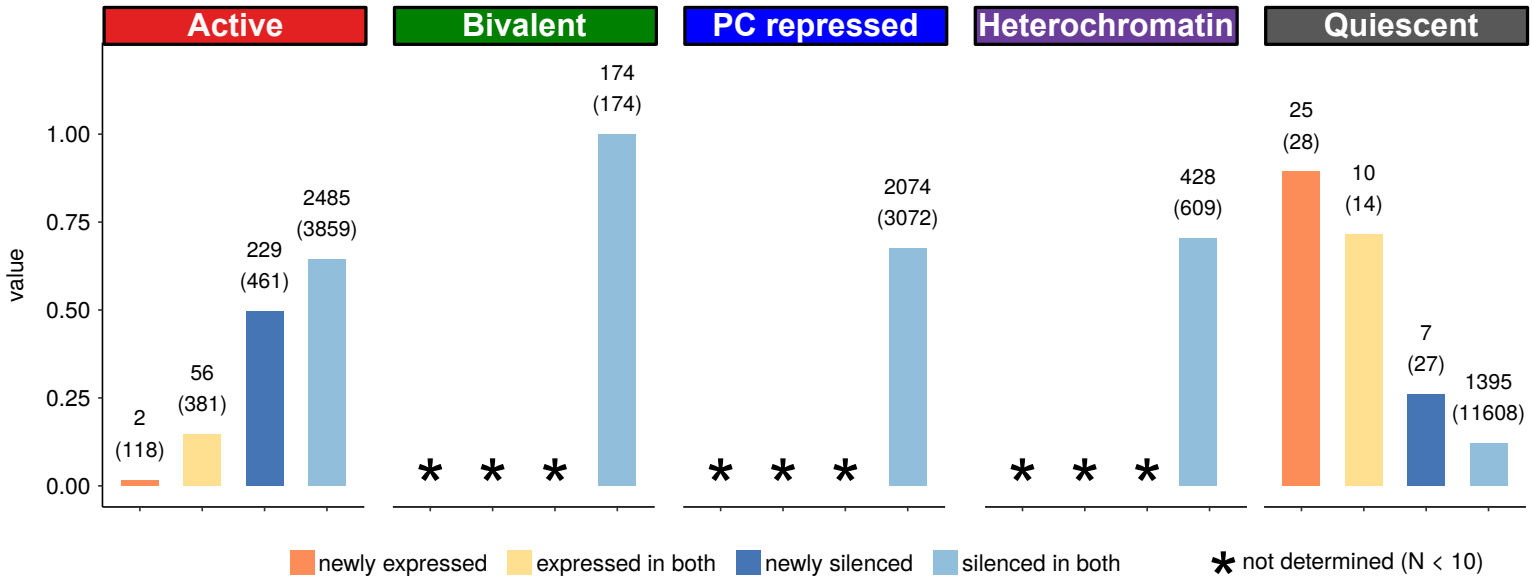
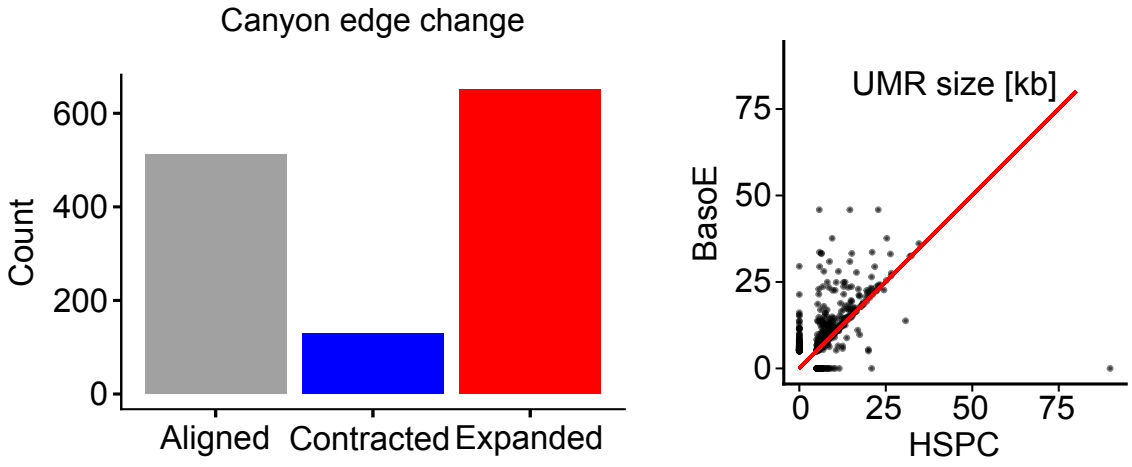

Fraction of CpG-poor promoters with altered chromatin state between HSPC and BasoE
Active
Bivalent
PC repressed
Heterochromatin
Quiescent
value
1.00
0.75
0.50
0.25
0.00
2 (118)
56 (381)
229 (461)
2485 (3859)
174 (174)
2074 (3072)
428 (609)
25 (28)
10 (14)
7 (27)
1395 (11608)
newly expressed
expressed in both
newly silenced
silenced in both
not determined (N < 10)

Figure S5

E


Canyon edge change
Count
600
400
200
0
Aligned
Contracted
Expanded
UMR size [kb]
BasoE
75
50
25
0
0
25
50
75
HSPC

F

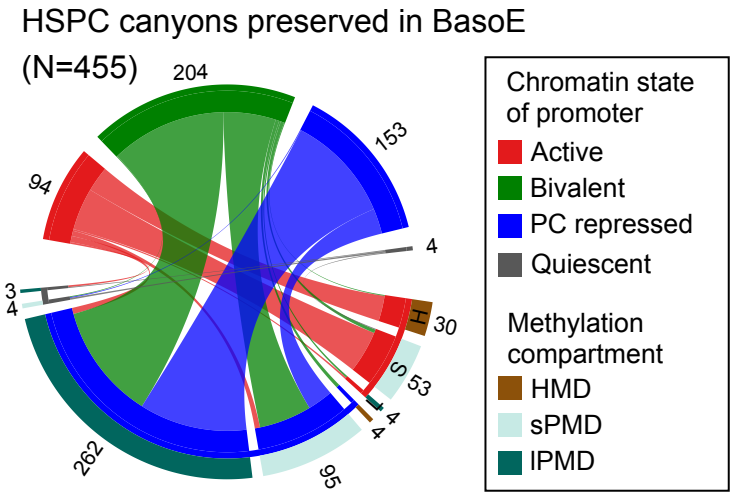

HSPC canyons preserved in BasoE
(N=455)
204
153
94
4
3
4
H
30
S
53
4
4
262
95
Chromatin state of promoter
Active
Bivalent
PC repressed
Quiescent
Methylation compartment
HMD
sPMD
lPMD

Figure S5

FNY 3_2

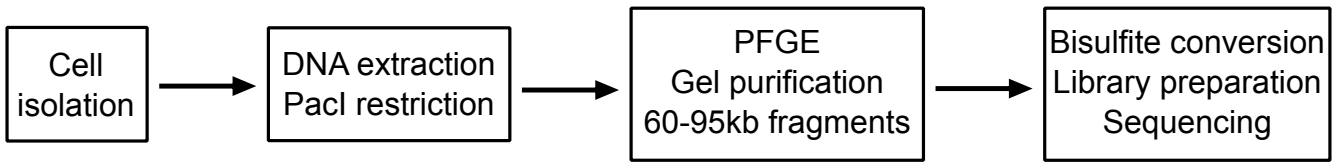
	a-DMRs						non-a-DMRs				
	A	C	G	T	25,624,923	Size (bp)	1,707,490,971	A	C	G	T
A	0	3.78	15.89	3.41	29,939	# Transitions	1,217,456	0	3.96	15.90	3.37
C	4.34	0	4.14	18.79	13,848	# Transversions	571,667	4.30	0	4.34	18.05
G	18.45	4.27	0	4.37	2.16	Ts/Tv	2.13	18.16	4.34	0	4.29
T	3.46	15.25	3.85	0	1.71	SNPs/kb	1.05	3.40	15.94	3.95	0

FNY 3_3

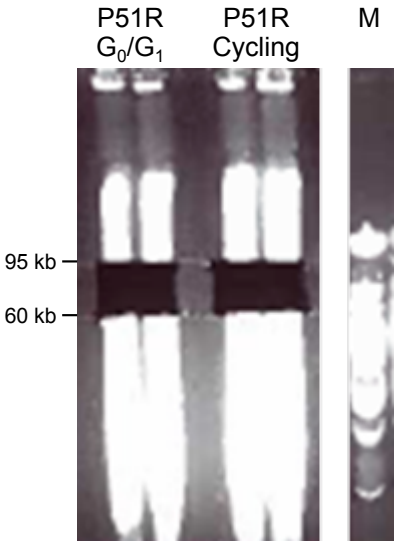
	A	C	G	T	24,841,170	Size (bp)	1,734,892,670	A	C	G	T
A	0	3.87	15.50	3.45	28,054	# Transitions	1,208,042	0	3.96	15.94	3.39
C	4.31	0	4.08	18.63	13,092	# Transversions	568,047	4.31	0	4.34	18.03
G	18.65	4.21	0	4.43	2.14	Ts/Tv	2.13	18.13	4.35	0	4.29
T	3.48	15.41	4.00	0	1.66	SNPs/kb	1.02	3.42	15.92	3.93	0

Figure S6

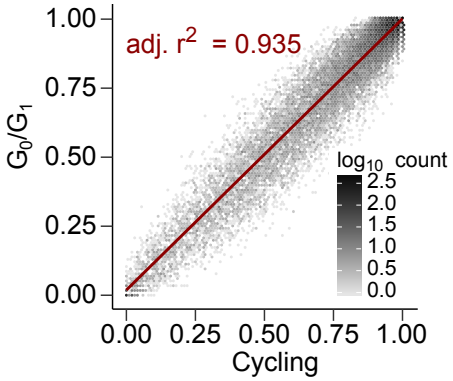
A



B



D



C

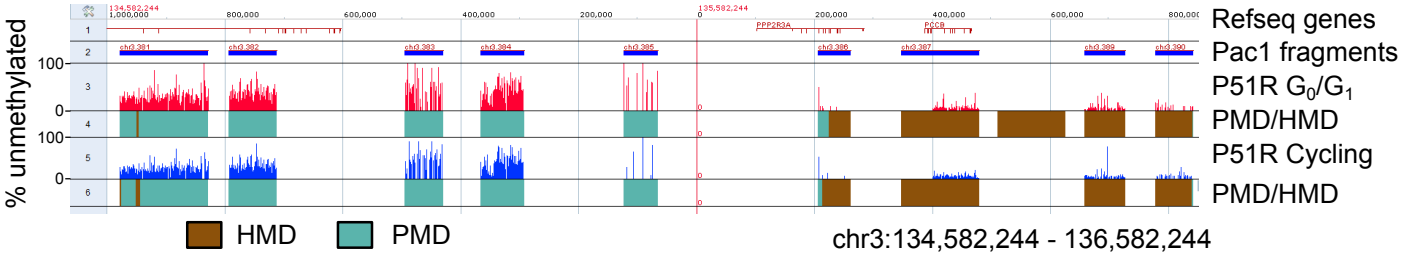


Figure S7