Biophysical Journal, Volume *113*

Supplemental Information

DNA Methylation Landscape Reflects the Spatial Organization of Chromatin in Different Cells

Ling Zhang, Wen Jun Xie, Sirui Liu, Luming Meng, Chan Gu, and Yi Qin Gao

Supporting Material

DNA Methylation Landscape Reflects the Spatial Organization of Chromatin in Different Cells

Ling Zhang[1,2†], Wen Jun Xie[1†], Sirui Liu[1], Luming Meng[1], Chan Gu[1,2] and Yi Qin Gao[1,2*]

[1] Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.
[2] Biodynamic Optical Imaging Center (BIOPIC), School of Life Sciences, Peking University, Beijing 100871, China.

*Corresponding author. Tel: 86-10-6275-2431; E-mail: gaoyq@pku.edu.cn
[†]These authors contributed equally to this work

The Supporting Material includes:

1. Sources of whole-genome bisulfite sequencing data
2. Detrended Fluctuation Analysis (DFA) for different cell classes
3. The luad_t5 and stad_n4 sample used in Fig. 2B
4. Gene analysis

1. Sources of whole-genome bisulfite sequencing data

Methylomes of human somatic cells

Reference: M. D. Schultz *et al*. (1)

URL: http://neomorph.salk.edu/human_tissue_methylomes.html

Data for individual 2 were used in the text (Fig. 1 and Fig. 2). Data for individual 1 and 3 were used to test the robustness of long-range correlations among individuals (Fig. S2A).

Table S1. Sample Details for Human Somatic Cells

number	name	symbol	individual	gender	age(year)
1	bladder	BL_1			
2	fat	FT_1			
3	gastric	GA_1			
4	lung	LG_1			
5	left ventricle	LV_1			
6	psoas	PO_1	1	male	3
7	right ventricle	RV_1			
8	thymus	TH_1			
9	small bowel	SB_1			
10	sigmoid colon	SG_1			
11	spleen	SX_1			
12	adrenal	AD_2			
13	aorta	AO_2			
14	esophagus	EG_2			
15	fat	FA_2			
16	gastric	GA_2			
17	lung	LG_2	2	female	30
18	ovary	OV_2			
19	pancreas	PA_2			
20	psoas	PO_2			
21	small bowel	SB_2			
22	spleen	SX_2			
23	adrenal	AD_3			
24	aorta	AO_3			
25	esophagus	EG_3			
26	fat	FT_3			
27	gastric	GA_3			
28	lung	LG_3	3	male	34
29	left ventricle	LV_3			
30	pancreas	PA_3			
31	psoas	PO_3			
32	right atrium	RA_3			
33	right ventricle	RV_3			

34	small bowel	SB_3
35	sigmoid colon	SG_3
36	spleen	SX_3

Methylomes of human cancer cells

The results shown here are partly based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/.

URL: https://portal.gdc.cancer.gov/legacy-archive/search/f

Reference for colon cells: B. P. Berman *et al*. (2)

The following typical samples were used to represent different cancers in Fig. 1B and Fig. S2: brca_t5, coad_t1, gbm_t2, luad_t1, lusc_t4, read_t2, stad_t1, ucec_t3 and colon_t1.

Table S2. Sample Details for Human Cancer Cells

number	TCGA barcode	symbol	cancer type
1	TCGA-A2-A04X-01A-21D-A19F-05	brca_t1	breast invasive carcinoma
2	TCGA-A8-A07I-01A-11D-A19F-05	brca_t2	
3	TCGA-A2-A0YG-01A-21D-A19F-05	brca_t3	
4	TCGA-E2-A15H-01A-11D-A19F-05	brca_t4	
5	TCGA-A7-A0CE-01A-11D-A148-05	brca_t5	
6	TCGA-A7-A0CE-11A-21D-A148-05	brca_n5	
7	TCGA-AA-A00R-01A-01D-A22T-05	coad_t1	colon adenocarcinoma
8	TCGA-AA-3518-01A-02D-1518-05	coad_t2	
9	TCGA-AA-3518-11A-01D-1518-05	coad_n2	
10	TCGA-06-0128-01A-01D-2294-05	gbm_t1	glioblastoma multiforme
11	TCGA-14-1454-01A-01D-2294-05	gbm_t2	
12	TCGA-14-3477-01A-01D-2294-05	gbm_t3	
13	TCGA-14-1401-01A-01D-2294-05	gbm_t4	
14	TCGA-16-1460-01A-01D-2294-05	gbm_t5	
15	TCGA-19-1788-01A-01D-2294-05	gbm_t6	
16	TCGA-38-4630-01A-01D-2365-05	luad_t1	lung adenocarcinoma
17	TCGA-67-6215-01A-11D-2365-05	luad_t2	
18	TCGA-78-7156-01A-11D-2365-05	luad_t3	
19	TCGA-91-6840-01A-11D-2365-05	luad_t4	
20	TCGA-44-6148-01A-11D-2365-05	luad_t5	
21	TCGA-44-6148-11A-01D-2365-05	luad_n5	
22	TCGA-34-2600-01A-01D-1871-05	lusc_t1	lung squamous cell carcinoma
23	TCGA-60-2695-01A-01D-1871-05	lusc_t2	
24	TCGA-21-1078-01A-01D-2365-05	lusc_t3	
25	TCGA-60-2722-01A-01D-1871-05	lusc_t4	
26	TCGA-60-2722-11A-01D-1871-05	lusc_n4	
27	TCGA-AG-3593-01A-01D-2294-05	read_t1	rectum adenocarcinoma
28	TCGA-AF-2689-01A-01D-2294-05	read_t2	

29	TCGA-AF-2689-11A-01D-2294-05	read_n2	
30	TCGA-CG-5730-01A-11D-2365-05	stad_t1	stomach adenocarcinoma
31	TCGA-D7-6519-01A-11D-2365-05	stad_t2	
32	TCGA-F1-6177-01A-11D-2365-05	stad_t3	
33	TCGA-BR-6452-01A-12D-2365-05	stad_t4	
34	TCGA-BR-6452-11A-01D-2365-05	stad_n4	
35	TCGA-B5-A0K6-01A-11D-A23D-05	ucec_t1	uterine corpus endometrial carcinoma
36	TCGA-AX-A1CK-01A-11D-A23D-05	ucec_t2	
37	TCGA-AP-A05J-01A-11D-A23D-05	ucec_t3	
38	TCGA-A5-A0G2-01A-11D-A23D-05	ucec_t4	
39	TCGA-AX-A1CI-01A-11D-A17H-05	ucec_t5	
40	TCGA-AX-A1CI-11A-11D-A17H-05	ucec_n5	
41	TCGA-DK-A1AA-01A-11D-A23D-05	blca_t1	bladder urothelial carcinoma
42	TCGA-DK-A1AG-01A-11D-A23D-05	blca_t2	
43	TCGA-BL-A13J-01A-11D-A23D-05	blca_t3	
44	TCGA-BT-A2LA-01A-11D-A23D-05	blca_t4	
45	TCGA-H4-A2HQ-01A-11D-A23D-05	blca_t5	
46	TCGA-BT-A20V-01A-11D-A23D-05	blca_t6	
47	TCGA-BT-A20V-11A-11D-A23D-05	blca_n6	
	colorectal cancer		
number	name	symbol	cancer type
48	colon tumor	colon_t1	colorectal cancer
49	colon normal	colon_n1	

Methylomes of human and mouse brain cells

Reference: R. Lister *et al.* (3)

URL: http://neomorph.salk.edu/brain_methylomes/

Table S3. Sample Details for Human Brain Cells

number	species	symbol	brain region	cell type	gender	age
1	human	fetal	cerebral cortex	tissue	male	20 week
2	human	35do	middle frontal gyrus	tissue	male	35 day
3	human	2yr	middle frontal gyrus	tissue	male	2 year
4	human	5yr	middle frontal gyrus	tissue	male	5 year
5	human	12yr	middle frontal gyrus	tissue	male	12 year
6	human	16yr	middle frontal gyrus	tissue	male	16 year
7	human	25yr	middle frontal gyrus	tissue	male	25 year
8	human	64yr	frontal cortex	grey matter	female	64 year
9	mouse	10wk	frontal cortex	tissue	male	10 week

Methylomes of human stem cells

Reference: R. Lister *et al.* (4)

URL: http://neomorph.salk.edu/ips_methylomes/data.html

Table S4. Sample Details for Human Stem Cells

number	name	symbol
1	ADS(adipose-derived stem cells)	ads
2	adipocytes derived from the ADS cells	ads_adipose
3	ADS iPSCs	ads_ipsc
4	foreskin fibroblast(FF)	ff
5	FF iPSC 6.9	ff_ipsc_6_9
6	FF iPSC 19.7	ff_ipsc_19_7
7	FF iPSC 19.11	ff_ipsc_19_11
8	IMR90(fetal lung fibroblast)	imr90
9	IMR90-iPSC	imr90_ipsc
10	H1	h1
11	H9	h9
12	HUES6	hues6

*For hues6, the reference is R. Lister *et al.(3)*

Methylomes for human neurodegenerative diseases

Reference: J.V. Sanchez-Mut *et al.* (5)

FastQ format reads of neurodegenerative diseases methylome were kindly provided by M. Esteller and the reads were aligned to the hg19 human reference genome with the Bowtie alignment algorithm(6).

Table S5. Sample Details for Human Neurodegenerative Diseases

number	name	symbol	disease	region	age (year)	gender
1	A09	a09	Alzheimer's disease	Brodmann area 9 gray matter	81	female
2	DBL2	dbl2	Dementia with Lewy bodies	Brodmann area 9 gray matter	77	female
3	BK1207	bk1207	Parkinson's disease	Brodmann area 9 gray matter	77	female
4	31_08	31_08	Down syndrome with Alzheimer's disease	Brodmann area 9 gray matter	49	male
5	G145	g145	Control gray matter	Brodmann area 9 gray matter	64	female
6	W145	w145	Control white matter	Brodmann area 9 white matter	64	female

2. Detrended Fluctuation Analysis (DFA) for different cell

classes

DFA has been used to show the long-range correlation in DNA sequence (7). Here we use the DFA to demonstrate the long-range correlation in DNA methylome. Root mean square fluctuation F(r) of a one-dimensional sequence is an important statistical quantity. It is typically defined as

$$\tilde{F}(r)^2 = \overline{[\Delta s(r)]^2} - \overline{\Delta s(r)^2} \quad (1)$$

where $s(r) = \sum_{i=1}^{r} u(r)$ is the sum of the methylation level for the first r*th* units, $\Delta s(r) = s(r_0 + r) - s(r_0)$, the bars indicate an average over all possible r_0 in the sequence. To make comparisons simple, here we normalize $\tilde{F}(r)$ as

$$F(r)^2 = \tilde{F}(r)^2 / Var \quad (2)$$

so that all detrended fluctuations start from the same point F(1)=1. Here *Var* is the variance of the methylation level of the whole sequence. This normalized $F(r)$ is directly related to the correlation function $C(r)$ through the equation

$$F(r)^2 \approx \prod_{j,k=1}^{r} C(j-k). \quad (3)$$

The '≈' can be replaced by a '=' as long as r is much smaller than sequence length L, which is often the case.

For purely uncorrelated random sequences, $F(r) \sim r^{1/2}$, corresponding to a ~0.5 slope in double-log plot. If the correlation of a sequence decays exponentially, indicating a finite-range correlation, the fluctuation scaling exponent will also be 0.5. Only when a long-range correlation with an infinite characteristic length is expected, will the scaling exponent deviate from 0.5, thus may be described by a power law. If the sequence holds a power law correlation when extending to infinite length, that is to say, $F(r) \sim r^{\alpha}$ and $C(r) \sim r^{-\gamma}$ when $r \to \infty$, there is a simple relation between fluctuation scaling exponent α and correlation scaling exponent γ

$$\alpha = \frac{2-\gamma}{2} \ . \quad (4)$$

However, in most cases, one can only expect a finite sequence length, thus the quantitative relation described above may not be accurate, but the qualitative property that a higher α corresponds to a lower γ still holds. For DNA methylation, the higher is the fluctuation scaling exponent, the flatter correlation double-log plot is, indicating a slower long-range decay.

We use Equation 3 to calculate the detrended fluctuation from 200-bp resolution methylation correlation. The fluctuation scaling exponent is estimated by linearly fitting the double-log plot of detrended fluctuation in 2kb~0.2Mb range. The average scaling exponents of 0.76 ± 0.01 and 0.92 ± 0.02 are observed in normal somatic cells (Fig. S4A) and cancer cells (Fig. S4B) separately. Their deviation from 0.5 and small variances indicate a uniform power law decay within certain cell states among different types of tissues. Cancer cells hold an obviously higher scaling exponent, in accordance with their flatter double-log correlation curves.

The detrended fluctuation analyses are also performed on gland cells (Fig. S4C) and brain cells (Fig. S4D). Gland cells show similar but smaller positive deviation from

normal somatic cells in fluctuation as cancer cells, while the scaling exponent of brain cells demonstrates no significant difference from that of normal somatic cells, in contrast to the corresponding correlation analysis. Since DFA is based on the sum of correlations, it largely reduces the random fluctuation in correlation. However, it also loses the detailed information through summing with higher cumulative weights for shorter-range correlations and responds much slower to scaling changes than correlation. Thus the methylome landscape differences between brain and normal somatic cells could be concealed by this cumulative operation in fluctuation analyses.

To further investigate the methylome landscape in differentiation, we apply DFA on human stem cells and related samples (Fig. S4E). All the stem cells and related fibroblasts are grouped into two categories according to DFA results, one containing primary somatic cell lines like foreskin fibroblast (ff) and IMR90 as well as adult stem cell lines like adipose-derived stem cells (ads) and adipocytes derived from ads (ads_adipose), the other containing all the human embryonic stem cell (hESC) samples and induced pluripotent stem cell (iPSC) samples. The former category shows an averaged scaling exponent of 0.91 ± 0.03, similar to that in cancer, while the latter gives an exponent of 0.69 ± 0.02, suggesting a long-range correlation with negative deviation from somatic cells.

3. The luad_t5 and stad_n4 sample used in Fig. 2B

As can be seen from Fig. 2B, the luad_t5 sample was clustered into normal cells using the scaling exponents of all chromosomes. The somatic mutations and copy number variations of luad_t5 sample were also analyzed to identify its clinical status. The number of somatic mutations in this sample is 17, which is smaller than that of regular tumor samples. The probability distribution of CNVs in this sample is not a single-peaked distribution. These data indicate that from the perspective of somatic mutations, the sample behaves like a normal one but the CNVs proves that it is actually a tumor sample. The AJCC stage of the patient is Stage IA, so we guess luad_t5 sample has not utterly become tumor sample.

The probability distribution of CNVs in stad_n4 sample is a unimodal distribution with a high peak at 0, indicating that there are very small number of CNVs in this sample and the sample behaves as normal cells. However, the stad_n4 sample's AJCC stage is Stage IIA which is consistent with our clustering that stad_n4 might have some cancer properties.

4. Gene analysis

Table S6. Average Expression Levels of Genes in PMDs, PMD-like Regions and Non-PMDs

Cancer type	Number of genes intersecting with PMDs	PMD	PMD-like regions	Non-PMD in tumor	Non-PMD in normal
brca	473	1.69	5.38	40.69	43.17
coad	642	8.96	24.53	24.63	28.48
luad	133	2.73	3.17	39.88	37.96
ucec	1109	7.94	12.96	23.30	36.24

* Expression levels in this table are TPM (transcripts per million).
* The length of genomic regions used in this table are all greater than 0.1 M.
* The normal samples correspond to the tumor sample of the same patient.

Table S7. Fisher's Exact Test of Gene Expression Level in PMDs, PMD-like Regions and Non-PMDs

Cancer type	Tumor PMD and Tumor non-PMD	Normal PMD-like regions and Tumor non-PMD	Tumor PMD and Normal PMD-like regions
brca	$7.2077*10^{-108}$	$6.8137*10^{-86}$	$9.6417*10^{-04}$
coad	$1.9133*10^{-54}$	$5.2351*10^{-47}$	0.4660
luad	$1.0565*10^{-33}$	$1.6939*10^{-30}$	0.3846
ucec	$1.0527*10^{-93}$	$9.1349*10^{-151}$	$1.2888*10^{-09}$

* The length of genomic regions used in this table are all greater than 0.1 M.

Table S8. Classification of Genes Intersect with PMDs

Genes	Number	Classification Feature	Specific genes	Number	Ratio
genes intersecting with PMDs	473	location	genes within PMDs	305	0.645
			genes within PMD center	156	0.330
		state	specifically expressed in tumor sample	17	0.036
			specifically expressed in normal sample	55	0.116
			repressed genes	167	0.353
			activated genes	13	0.027
		GO term	disulfide bond	190	0.402
			glycoprotein	220	0.465
			membrane	231	0.488
			signal	149	0.315
			housekeeping genes	2	0.004
		Promoter type	Non-CGI promoter	273	0.577
			CGI promoter	200	0.423

* PMDs with genomic lengths greater than 0.1M are considered in this table.
* Total number of housekeeping genes is 3796.
* PMD center is defined as the central 60% regions of PMD.

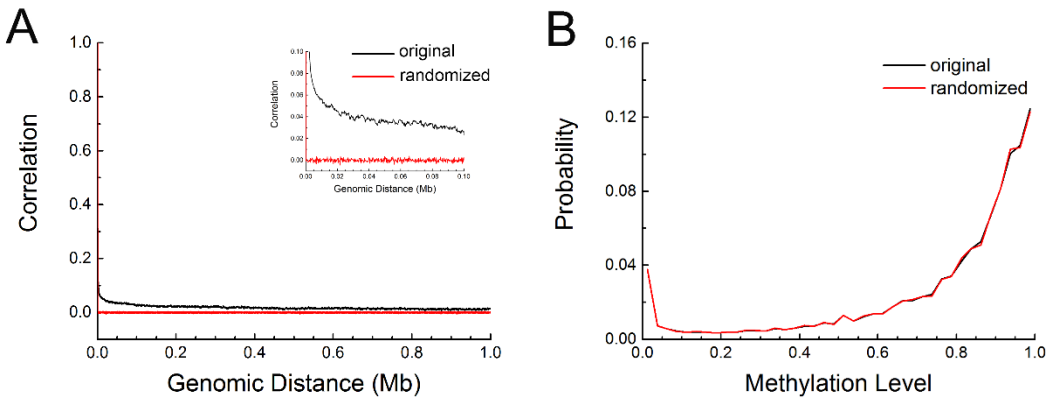

Figure S1. The non-random nature of DNA methylation. (A) The DNA methylation correlations from original experimental data and randomized data for chromosome 1 of human aorta cell (sample label: AO_2). The genomic distances below 0.10 Mb are enlarged and shown in the inset. (B) Methylation level distribution of original experimental data and randomized data. The randomized data was produced by assigning each CpG site with a random value following the overall distribution of DNA methylation level. We first generated a random number (y) following the uniform distribution between 0 and 1, and found the highest x satisfying $F(x) \leq y$, where F(x) is the cumulative distribution function of methylation. The value of x was assigned to each CpG site as its methylation level.

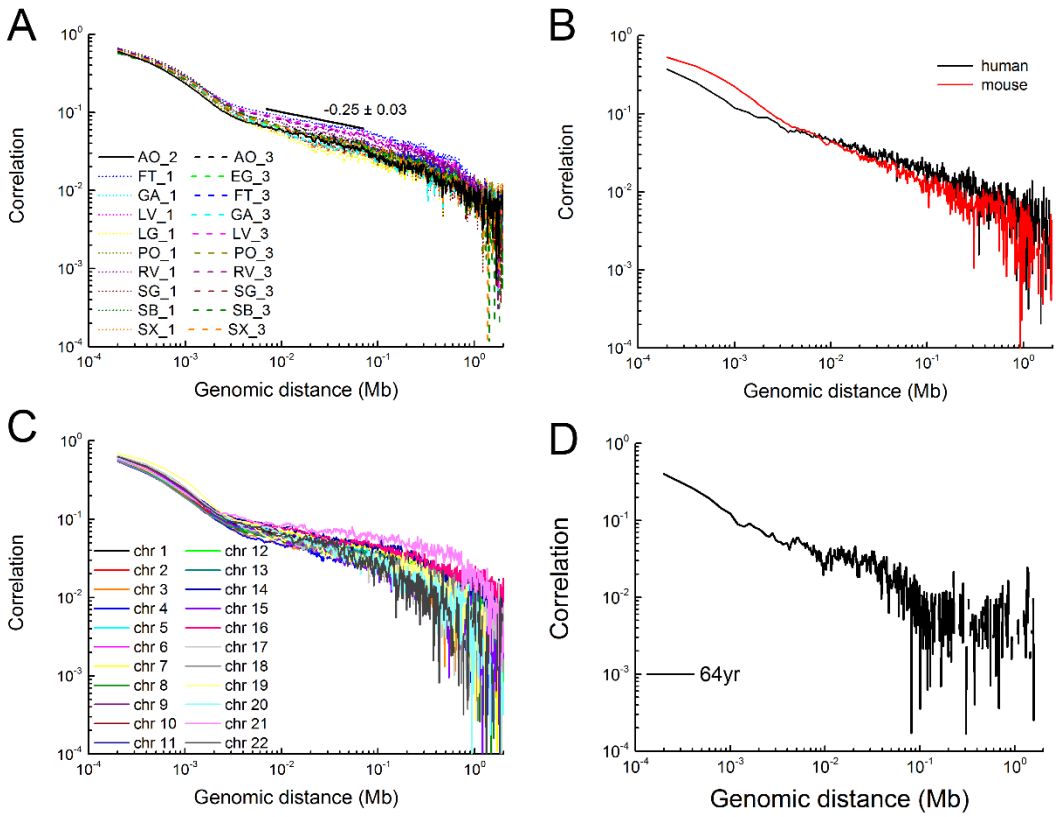

Figure S2. Power law scaling of methylation correlation in different individuals, species, chromosomes and exceptions.

(A) Robustness of the scaling exponents among different individuals. The scaling exponents for chromosome 1 of human somatic cells in three different individuals. The small standard deviations show that the scaling exponents are conserved among different individuals. (B) The power law scaling is also present in mouse brain cell. 25-year-old human brain sample and 10-week-old mouse brain sample are used as examples of human and mouse brain, respectively. All the brain data are summarized in Table S3. (C) The power law scaling behavior is observed in different chromosomes. The chromosomes in aorta from individual 2 (sample label: AO_2) are used. (D) The scaling exponents in the concerned genomic region (kilobase to megabase) are not well-defined in some chromosomes. The chromosome 22 of the 64yr human brain is taken as an example. In the kilobase to megabase region, the large fluctuation of methylation correlation makes it not feasible to calculate the scaling exponent.

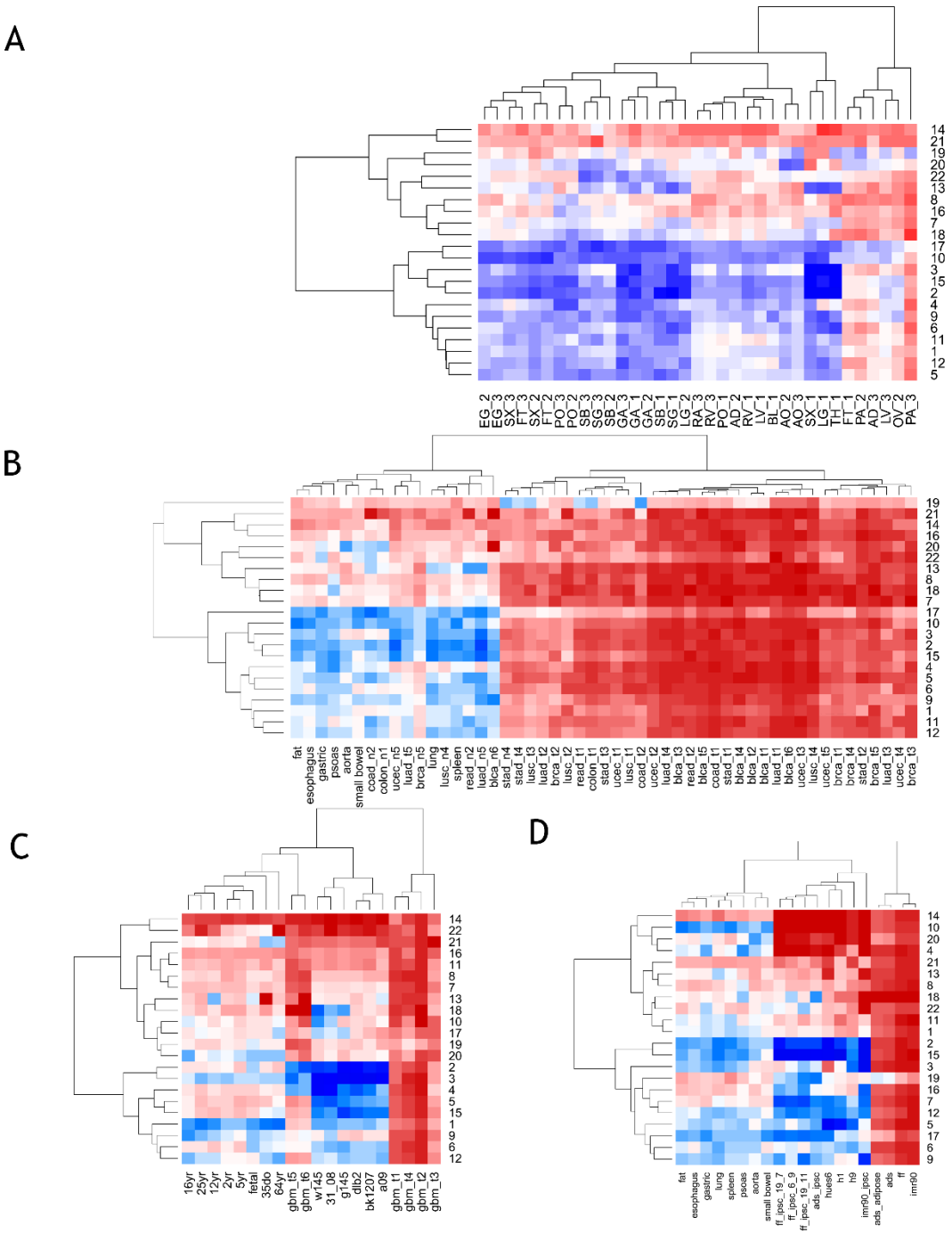

Figure S3. Heatmap clustering of scaling exponents in different chromosomes. The sample labels are the same as Fig. 2 of the main text. In this figure, we also clustered the scaling exponents on the samples. (A) The scaling exponents of normal somatic cells. (B) Normal somatic cells segregate from cancer cells. (C) Normal brain cells segregate from glioblastoma or neurodegenerative diseases. (D) ESCs and iPSCs segregate from cell lines including adult stem cell line and somatic cell lines.

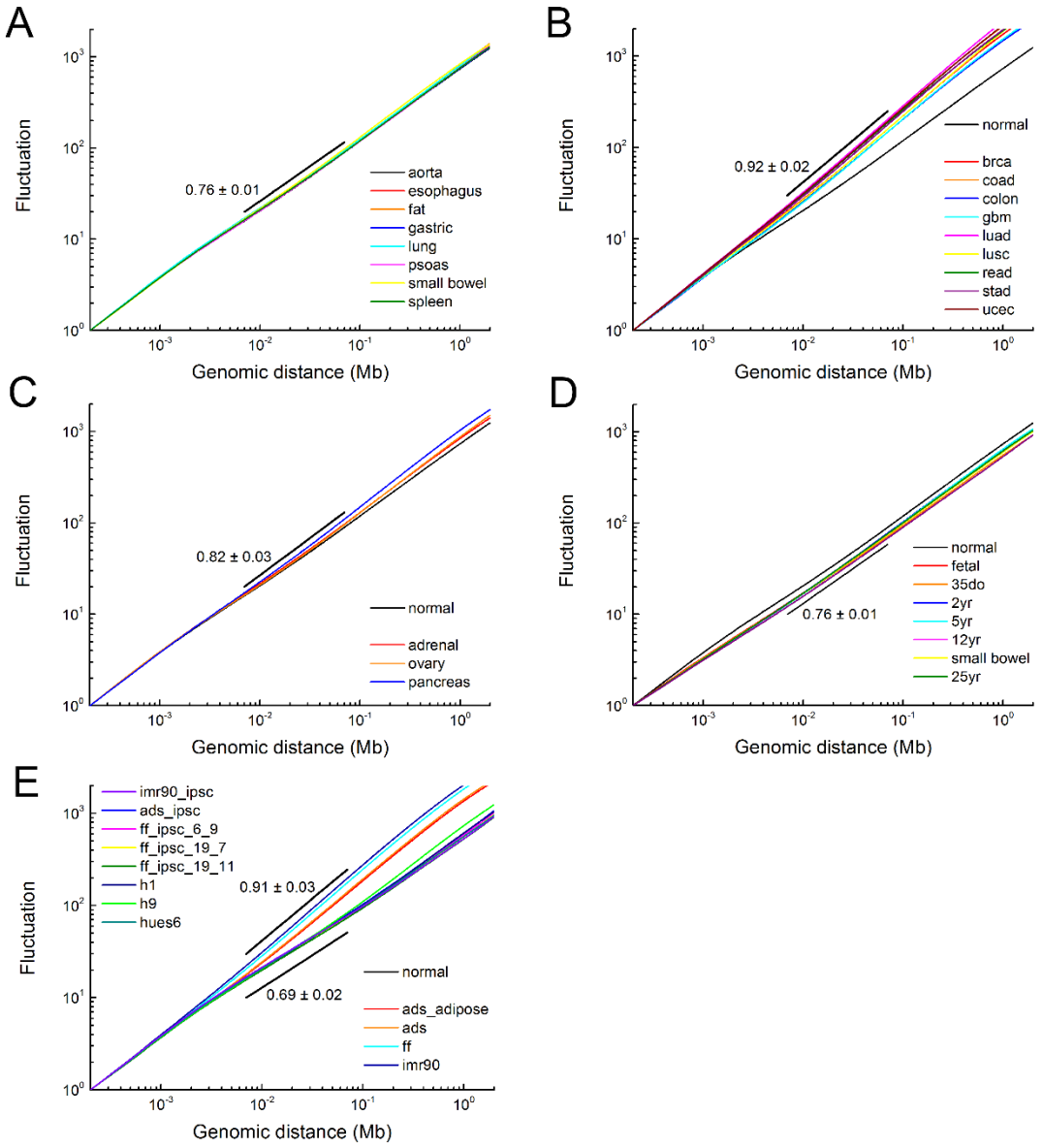

Figure S4. Detrended fluctuation analysis for chromosome 1 in different cell classes.

(A) Normal somatic cells show coherent power law scaling relationship in methylation; (B) Cancer cells. (C) Gland cells. (D) Normal brain cells. (E) Human stem cells and related cells. All these stem cells and related cells are divided into two groups, one with lower-than-normal scaling exponents including all iPSCs and hESCs, the other with cancer-like high scaling exponents including primary somatic cell lines and adult stem cell line. For all cell classes, the average scaling exponent is annotated in the figure and fluctuation for aorta is plotted as normal for comparison.

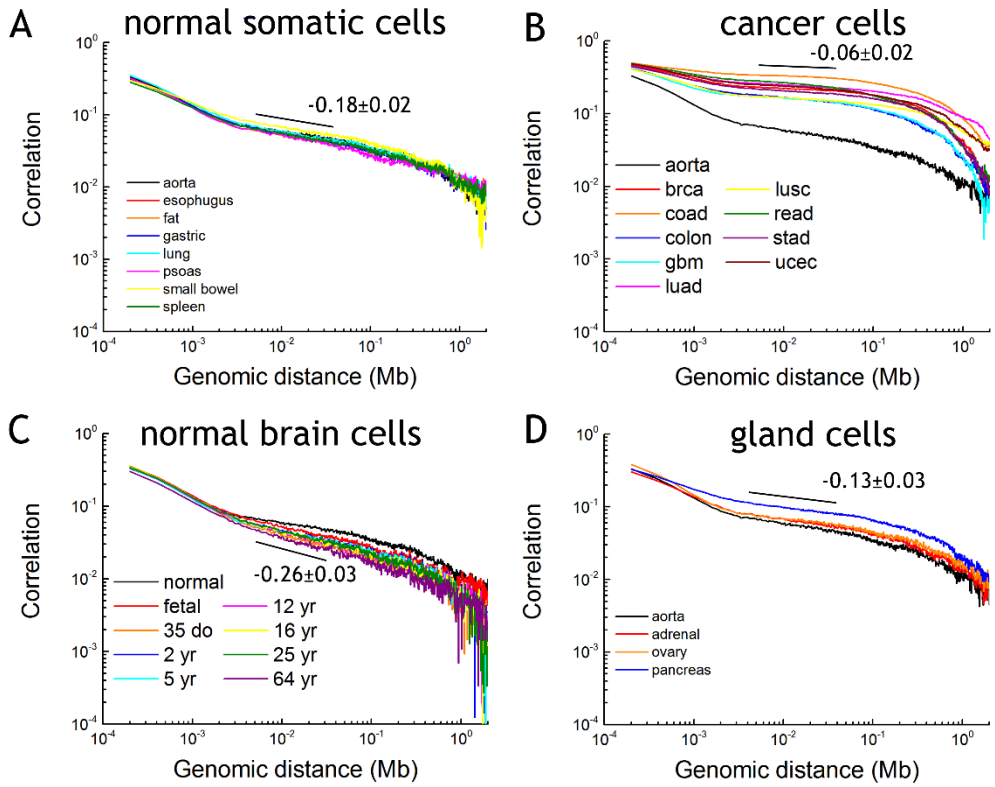

Figure S5. Long-range correlations of DNA methylation using discrete model series. The sample labels are the same as Fig. 1 of the main text. The average scaling exponents are annotated in the figure. Correlation for normal aorta cells (normal) is also plotted for comparison in (B), (C) and (D). We discretized the DNA methylation level of each sample into 1 and 0 with the methylation average as reference value. Specifically, for chromosome 1 of each cell type, we assign a value of 1 to every 200-bp unit with methylation level greater than chromosome average, and 0 to that with methylation level smaller than average.

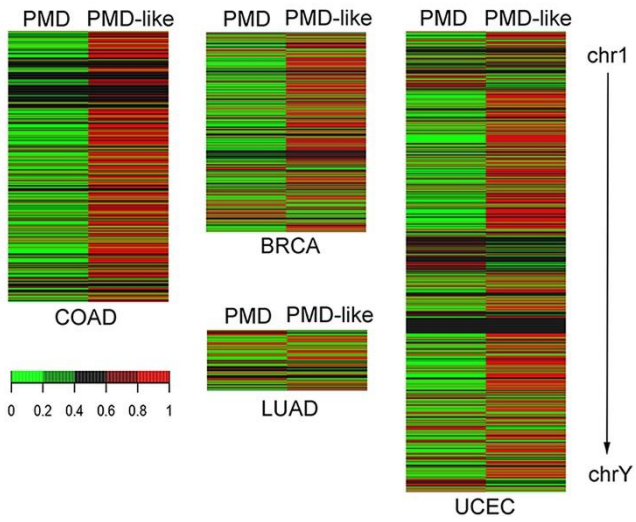

Figure S6. Heatmap of gene expression difference (TPM) in PMD and PMD-like regions in coad_t2-coad_n2, brca_t5-brca_n5, luad_t5-luad_n5 and ucec_t5-ucec_n5 sample pairs. The difference of gene expression in tumor PMDs and normal PMD-like regions for each gene (d = $TPM_{PMD} - TPM_{PMD-like}$). If the difference is greater than 0 we denoted the PMD expression of this gene as 1 and the PMD-like expression as 0. If the difference is smaller than 0 we denoted the PMD expression of this gene as 0 and the PMD-like expression as 1 and if the difference equals 0, both PMD and PMD-like gene expression are denoted as 0.5. The gene expression in PMDs is lower than that in PMD-like regions.

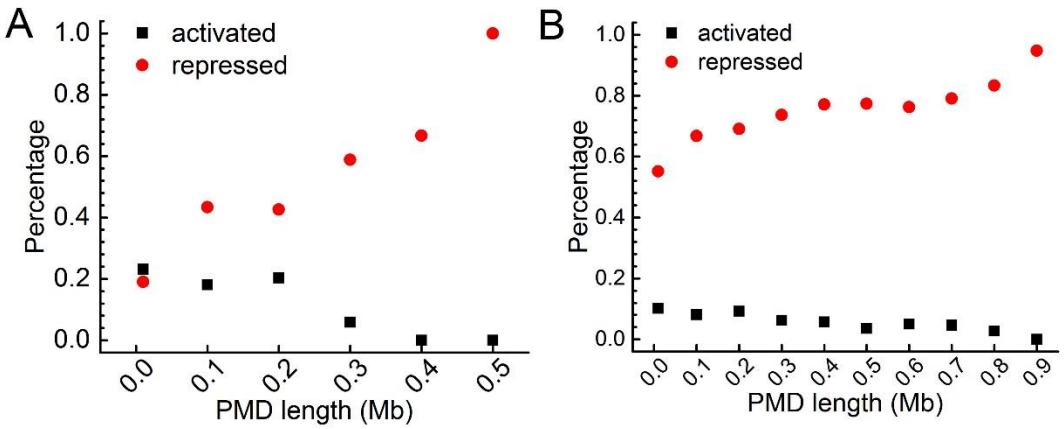

Figure S7. Percentage of genes that are transcriptional activated or repressed in oncogenesis as a function of PMD length.
(A) coad_t2-coad_n2 sample pair. (B) ucec_t5-ucec_n5 sample pair. As the most of PMDs in the luad_t5 sample is short, the luad_t5-luad_n5 sample pair is not shown.

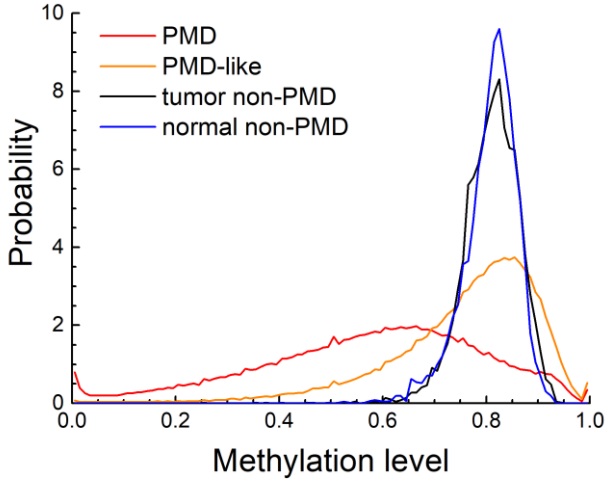

Figure S8. Distribution of methylation level of PMD, non-PMD and PMD-like genomic regions in breast cells. After oncogenesis, the methylation of PMD-like genomic region decreases and turns into the hypomethylated PMD. The methylation of non-PMD doesn't change before and after oncogenesis.

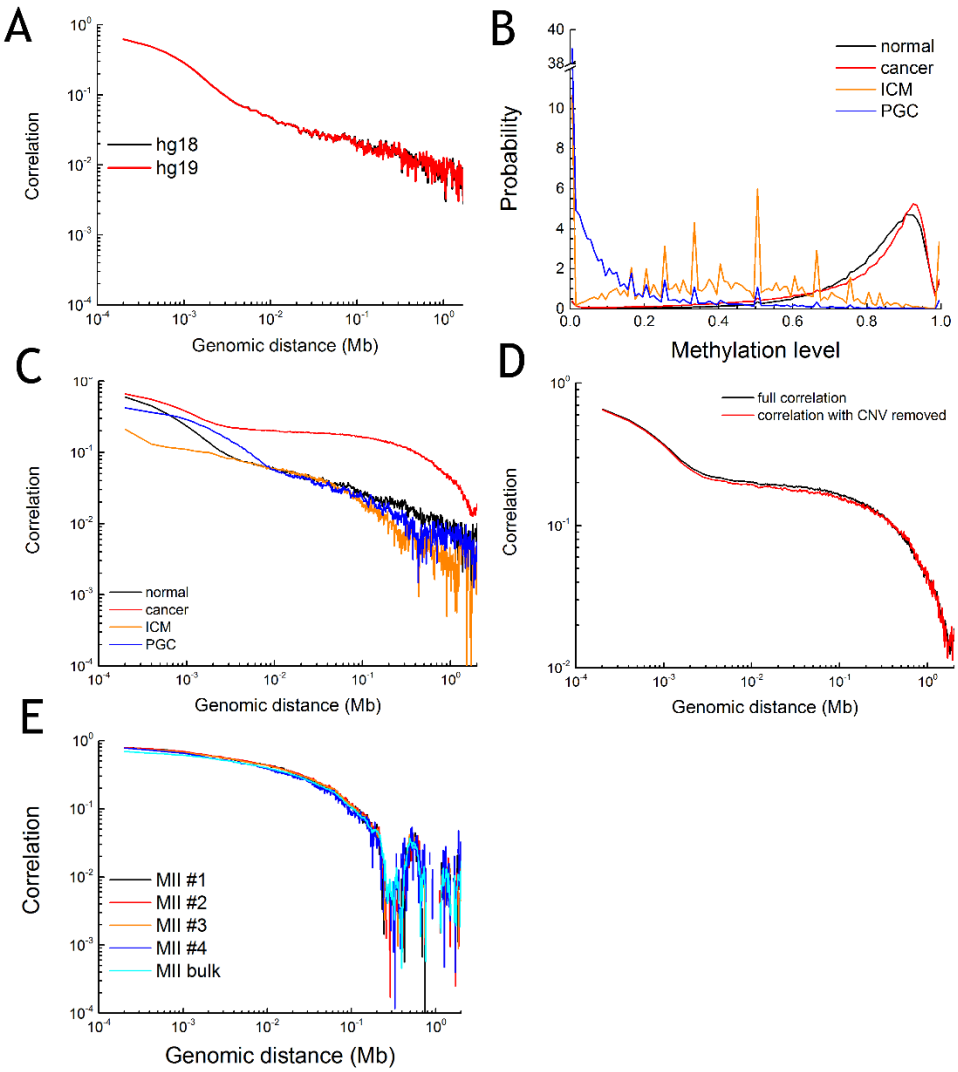

Figure S9. Influence of DNA methylation level, copy number variation and reference genome.

(A) The reference genome has no obvious effect on the methylation correlation. The DNA methylation correlation of chr1 in human brain sample (12yr) using hg18 and hg19 reference genomes respectively. The methylation data under hg19 reference genome was obtained by transforming the original hg18 data using liftover tools. (B) Methylation level of chromosome 1 in four different samples (normal, cancer, PGC and ICM). AO_2 sample is plotted as an example for normal somatic cell and brca_t5 sample for cancer cell. The 10-week female PGC sample and the ICM sequenced by WGBS were used (reference genome: hg19). The average methylation levels for normal, cancer, ICM and PGC are 0.72, 0.74, 0.38, and 0.08, respectively. (C) Methylation correlations for chromosome 1 of the 4 samples in Figure S9B. (D) Long-range DNA methylation correlations are not affected by CNVs in brca_t5 tumor sample. (E) Long-range correlation of DNA methylation is conserved among different single cells. (Data from (8))

Supporting Reference

1. Schultz, M. D., Y. He, J. W. Whitaker, M. Hariharan, E. A. Mukamel, D. Leung, N. Rajagopal, J. R. Nery, M. A. Urich, H. Chen, S. Lin, Y. Lin, I. Jung, A. D. Schmitt, S. Selvaraj, B. Ren, T. J. Sejnowski, W. Wang, and J. R. Ecker. 2015. Human body epigenome maps reveal noncanonical DNA methylation variation. Nature 523:212-216.
2. Berman, B. P., D. J. Weisenberger, J. F. Aman, T. Hinoue, Z. Ramjan, Y. Liu, H. Noushmehr, C. P. Lange, C. M. van Dijk, R. A. Tollenaar, D. Van Den Berg, and P. W. Laird. 2012. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. Nat. Genet. 44:40-46.
3. Lister, R., E. A. Mukamel, J. R. Nery, M. Urich, C. A. Puddifoot, N. D. Johnson, J. Lucero, Y. Huang, A. J. Dwork, M. D. Schultz, M. Yu, J. Tonti-Filippini, H. Heyn, S. Hu, J. C. Wu, A. Rao, M. Esteller, C. He, F. G. Haghighi, T. J. Sejnowski, M. M. Behrens, and J. R. Ecker. 2013. Global epigenomic reconfiguration during mammalian brain development. Science 341:1237905.
4. Lister, R., M. Pelizzola, Y. S. Kida, R. D. Hawkins, J. R. Nery, G. Hon, J. Antosiewicz-Bourget, R. O'Malley, R. Castanon, S. Klugman, M. Downes, R. Yu, R. Stewart, B. Ren, J. A. Thomson, R. M. Evans, and J. R. Ecker. 2011. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471:68-73.
5. Sanchez-Mut, J. V., H. Heyn, E. Vidal, S. Moran, S. Sayols, R. Delgado-Morales, M. D. Schultz, B. Ansoleaga, P. Garcia-Esparcia, M. Pons-Espinal, M. M. de Lagran, J. Dopazo, A. Rabano, J. Avila, M. Dierssen, I. Lott, I. Ferrer, J. R. Ecker, and M. Esteller. 2016. Human DNA methylomes of neurodegenerative diseases show common epigenomic patterns. Transl. Psychiatry 6:e718.
6. Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10:R25.
7. Peng, C. K., S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley. 1992. Long-range correlations in nucleotide sequences. Nature 356:168-170.
8. Smallwood, S. A., H. J. Lee, C. Angermueller, F. Krueger, H. Saadeh, J. Peat, S. R. Andrews, O. Stegle, W. Reik, and G. Kelsey. 2014. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat. Methods 11:817-820.