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 †

Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, and Biodynamic Optical Imaging Center, Peking University, Beijing 100871, China

*These authors contributed equally to this work.

† To whom correspondence should be addressed.

Corresponding author

Professor Yi Qin Gao

Address: College of Chemistry and Molecular Engineering, Peking University, Beijing, 100871, China.

Phone: 86-10-6275-2431.

Email: gaoyq@pku.edu.cn

Running title

DNA Methylation Reflects Chromatin Structure

Keywords

DNA methylation, chromatin structure, Hi-C, power law, correlation

Abstract

It has been thought that DNA methylation can remodel chromatin structure, and modulate the transcriptional level of genes by altering chromatin density and accessibility of DNA to cellular machinery, which is still largely unknown. Here we report that both the long-range and local correlations in DNA methylation reflect the spatial organization of chromatin. We discover a long-range power law correlation that implies a scale-free property of DNA methylation and the cell-class-specific scaling exponents reflect the global change of DNA methylation landscape during cellular differentiation or oncogenesis which can be further related with chromatin structure. By analyzing on the local correlation which correlates with nucleosome positioning, Hi-C data and molecular modeling, we show that partially methylated domain (PMD) prefers a compact structure. We further demonstrate that chromatin structures reflect the DNA sequence property and are intimately related to gene expression. Our study therefore provides a novel view of the spatial organization of chromatin structure from a perspective of DNA methylation.

Introduction

The spatial organization of chromatin plays an essential role in many genomic functions, including gene expression, DNA replication and cell mitosis (Naumova et al. 2013; Levine et al. 2014; Galupa and Heard 2015; Pombo and Dillon 2015). Composed of DNA and histone proteins, chromatin has a three-dimensional structure with different hierarchical levels (Gibcus and Dekker 2013). Several lines of evidence show that epigenetics can remodel chromatin structure at different levels (Bell and Felsenfeld 2000; Jenuwein and Allis 2001; Cedar and Bergman 2009; Chodavarapu et al. 2010; Aranda et al. 2015; Boettiger et al. 2016). A variety of histone modifications are involved in chromatin remodeling which result in the histone code hypothesis of gene regulation (Jenuwein and Allis 2001). The different histone tail modifications can change chromatin structure by recruiting other proteins (Aranda et al. 2015). Super-resolution imaging also recently showed that chromatin folding varies for different epigenetic states (Boettiger et al. 2016).

As the most abundant epigenetic modification in eukaryotic chromosomes, DNA methylation is also thought to influence chromatin structure (Cedar and Bergman 2009). DNA methylation has a close relationship with nucleosome positioning (Chodavarapu et al. 2010), and the binding of CTCF can be partly influenced by DNA methylation and thus changes chromatin structure (Bell and Felsenfeld 2000). Recently, DNA methylation was also used to reconstruct A/B compartments of chromatin revealed by Hi-C experiments (Fortin and Hansen 2015). Nevertheless, how DNA methylation influences chromatin structure remains largely unexplored.

Here we investigate the long-range and local correlations in DNA methylation landscape which are shown to be informative of the underlying chromatin structure.

Results

DNA methylation shows long-range power law correlation

We compared the Pearson correlation coefficients of DNA methylation levels (beta values) within the methylome across a wide-range of human cells, including normal somatic cells, cancer cells, brain cells, gland cells and stem cells. The sources of relative whole-genome bisulfite sequencing data were summarized in Supplementary Information. In these human cells, DNA methylation occurs mainly in the CpG context.

Taking chromosome 1 as representative, all the methylation correlations strikingly present a long-range power law decay as the genomic distance increases (Fig. 1). The scale-invariant genomic segment lies in the kilobase-to-megabase scale that can be important for a variety of genomic functions, which is also the sizes of genes and chromatin domains (Schneider and Grosschedl 2007; Rao et al. 2014).

Power law scaling is of general interest (Clauset et al. 2009) and often noticed in evolving systems that may be produced by hierarchical structure of several length-scales (Eugene et al. 2006). Such a general phenomenon across different cells indicates that our findings here could have important biological meanings.

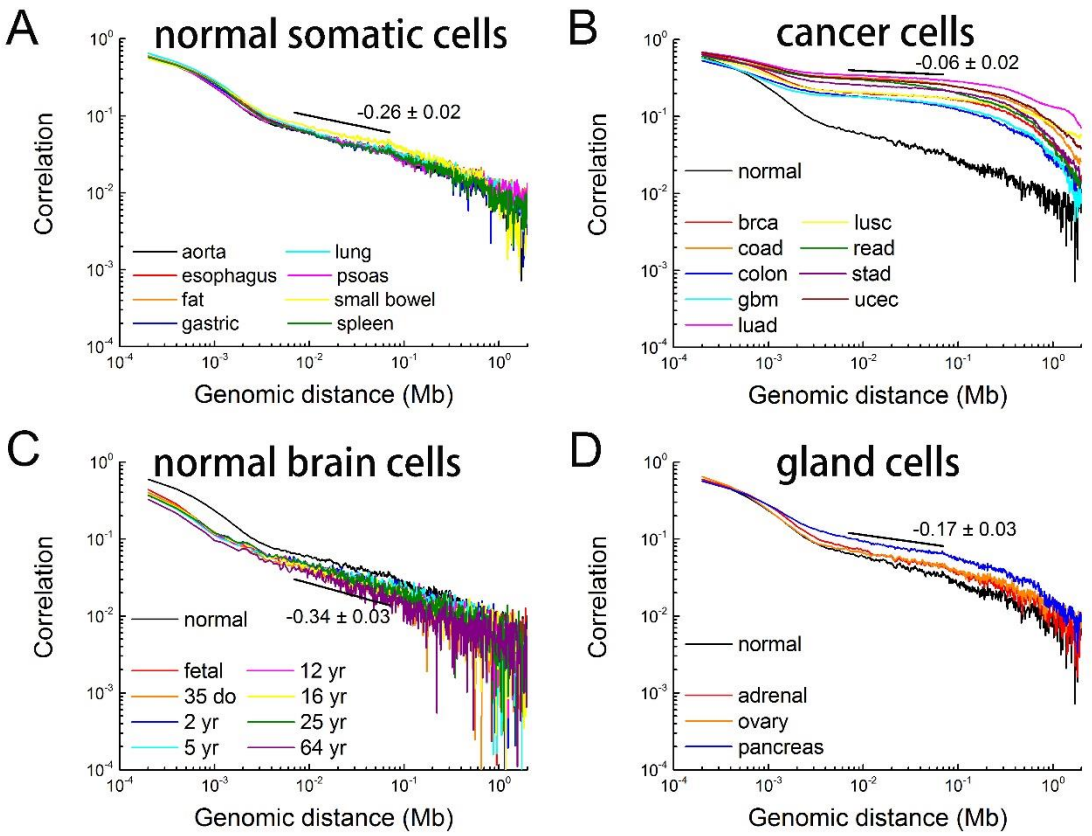

Figure 1. Long-range correlations in DNA methylation are distinctly different among different cell classes. The Pearson correlation coefficients for chromosome 1 of different cells are shown in log-log plots. The average scaling exponents are annotated in the figure. **A**, 8 different somatic cells: aorta, esophagus, fat, gastric, lung, psoas, small bowel and spleen. **B,** 9 different cancer cells: brca, coad, colon, gbm, luad, lusc, read, stad and ucec. The cells are labelled following TCGA except for colon cancer (colon). **C,** Normal brain cells of different ages (fetal, 35 days old, 2-64 years old). **D,** 3 different gland cells (adrenal, ovary, pancreas). Correlation for normal aorta cells (normal) is also plotted for comparison in **B, C and D**.

Cell-class-specific scaling exponents reflect the global change of chromatin during cellular differentiation

The scaling exponents differ substantially between normal somatic cells and cancer cells, and the respective values are -0.26±0.02 and -0.06±0.02, which is obviously smaller for cancer cells than normal somatic cells as a result of more pronounced and clearer differences in the methylation level of different domains in the former.

The smaller power law scaling in cancer cells is not caused by the lower average methylation level which has been subtracted in the calculation of the methylation level correlation. In fact, although cells like human inner cell mass and primordial germ cells have methylation levels that are even much lower than cancer cells, they possess power law correlation decays very similar to the normal but not cancer cells (Supplemental Fig. S1).

On the other hand, small standard deviations show that the scaling exponents are conserved among either normal somatic cells or cancer cells (Fig. 1A and B), although the methylation levels of individual CpGs vary greatly (Schultz et al. 2015). Such a result indicates that the locally averaged methylation level and its correlation reflect the large scale structural property of the chromatin.

Similarly, the differences between normal somatic cells and brain cells or gland cells but consistency within each class suggest that cellular differentiation causes systematic variations of DNA methylation landscape (Fig. 1C and D). These properties are robust among individuals (Supplemental Fig. S2). Detrended

fluctuation analyses (Peng et al. 1992) were also used which again show the long-range correlation in DNA methylome (Supplemental Figs. S5-S9).

The significant difference between different cell classes demonstrates that the long-range correlations in DNA methylome can-not be simply originated from DNA sequence (Peng et al. 1992; Arneodo et al. 2011). We also checked that the scaling exponents of methylation correlations are not affected by copy number variations in cancer cells (Supplemental Fig. S10).

DNA methylation was previously demonstrated to have long-range correlations by establishing a firm link with A/B compartment (Fortin and Hansen 2015), suggesting the scale-free property found here for DNA methylation to be originated from chromatin structure.

Long-range methylation correlations suggest the different chromatin structure in PMD and non-PMD

The CpG sites are overall highly methylated in human somatic cells (Meissner et al. 2008). Extensive changes in DNA methylation take place during tumourgenesis and genome-scale studies have shown that PMDs may serve to silence tumor-suppressor genes (Berman et al. 2012; Hon et al. 2012). These PMDs were identified as regions of long-range DNA hypomethylation which can cover almost half of the genome and the widespread PMDs were previously suggested to change the chromatin structure (Berman et al. 2012).

We compared DNA methylation in cancer PMD and non-PMD (genomic region

that isn't classified as PMD) to establish a possible relation between methylation and chromatin structure. Breast cancer was used as an example because of its extensive available data.

The long-range correlations inside cancer PMDs or non-PMDs again show power law decays at the kilobase-to-megabase scale (Fig. 2A). The scaling exponents for PMD and non-PMD are -0.23 and -0.43, respectively.

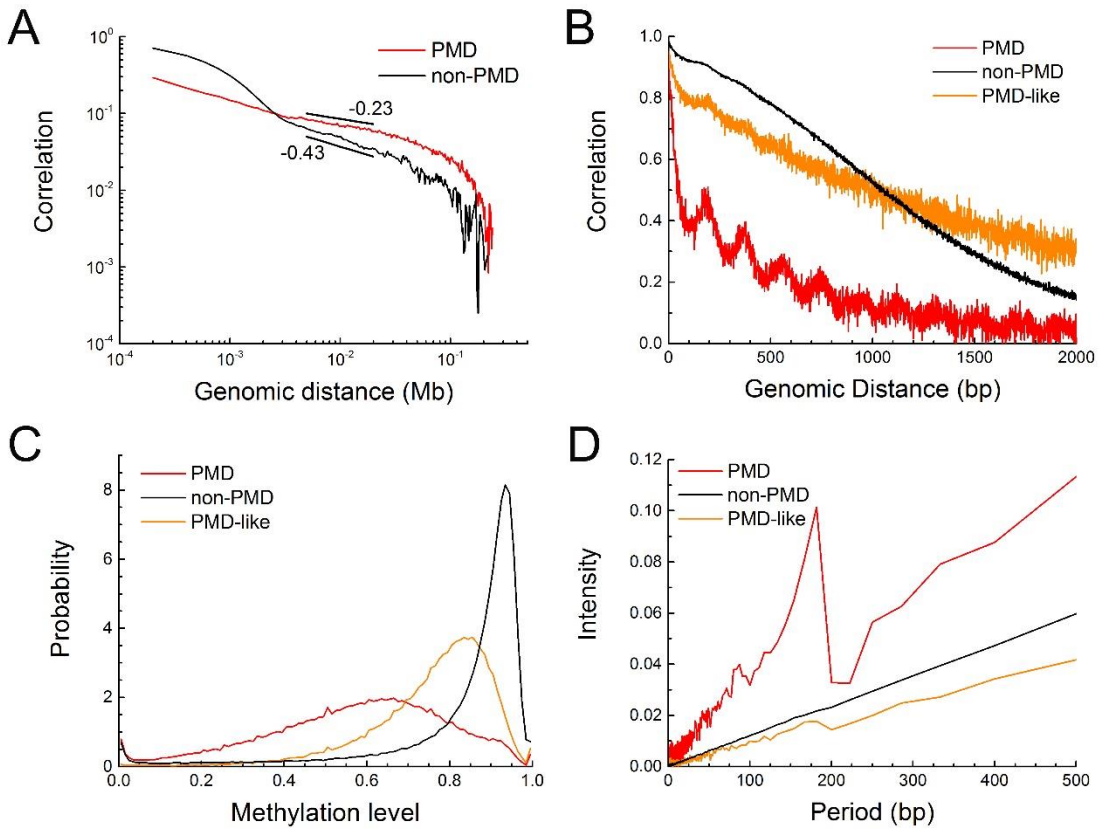

Figure 2. DNA methylation correlations and methylation levels of different genomic regions in breast cells. A, DNA methylation for PMD and non-PMD exhibit distinct scaling behavior. **B**, **C** and **D**, Methylation properties of PMD, non-PMD and PMD-like genomic regions. **B**, Local correlations of DNA methylation. **C**, The distribution of methylation levels. **D**, FFT is used to deconstruct inherent frequencies of the local correlations.

Interestingly, these values match well with the exponents of power law scaling between the radius of gyration (Rg) and genomic distance for different epigenetic states partly defined by the enrichment of histone modifications in fluorescence *in situ* hybridization (FISH) experiment, which are 0.22±0.02 for repressed domains, 0.30±0.02 for inactive domains and 0.37±0.02 for active domains (Boettiger et al. 2016). Such an observation suggests that the spatial organization of PMD is more compact than other genomic regions.

The coincidence also indicates that genes in PMDs are more repressive while those in non-PMDs are more active, resulted from the possible crosstalk between DNA methylation and histone modifications (Cedar and Bergman 2009). It was previously found that the PMDs in IMR90 correlate with repressive and anti-correlate with active histone marks (Berman et al. 2012). Consistent with earlier studies, we also find that genes within PMDs are transcriptionally repressed (Schultz et al. 2015), and discover that these genes relate with membrane, glycoprotein, disulfide bond and receptor instead of housekeeping genes (Supplemental Table S6). Furthermore, the average gene expression level of non-PMDs is much higher than that of cancer PMDs and corresponding genomic regions of cancer PMDs in normal cells (defined as “PMD-like” regions here, Supplemental Table S7). The binding proportion of all the transcription factors to PMDs is only 20%, in accordance with gene sparsity in PMD regions.

Local methylation correlations suggest the different chromatin structure in

PMD and non-PMD

Next, we compared the base-resolution correlation of CpG methylation in PMDs and non-PMDs in breast cells (Fig. 2B). Consistent with previous studies on IMR90 cell line (Gaidatzis et al. 2014), the decay of PMD correlation clearly shows an obvious periodic behavior at a base-resolution. The non-PMD regions, in contrast, do not show any apparent periodic behavior. PMD-like regions have an average methylation level higher than PMD and lower than non-PMD (Fig. 2C and 3E), and a less obvious methylation correlation periodicity (Fig. 2B).

Fast Fourier transforms (FFTs) of these correlation functions are used to characterize the periodicities which show at 181 bp a strong peak for PMD (Fig. 2D) and at the similar position a weaker one for PMD-like regions. The periodicity is consistent with the nucleosome repeat length, which suggests that the periodicity may come from the regular organization of nucleosomes. The different organization or depletion of nucleosomes that wrap DNA influence chromatin structure (Diesinger and Heermann 2009; Arneodo et al. 2011; Ricci et al. 2015), which imply the different chromatin structures in PMD and non-PMD.

Hi-C data confirm the different chromatin structure in PMD and non-PMD

In recent years, molecular methods based on chromosome conformation capture techniques, combined with molecular modeling, have dramatically expanded our understanding of chromatin structures (Dekker et al. 2002; Dostie et al. 2006; Lieberman-Aiden et al. 2009; Giorgetti et al. 2014; Zhang and Wolynes 2015; Dixon et

al. 2016). By measuring the frequency of physical interactions between any different genomic loci, 3C-based approaches including Hi-C technique have identified many evolutionary conserved topologically associating domains (TADs) and chromatin loops in mammalian genomes, which enable the distal *cis*-regulatory elements such as promoters and enhancers to interact with a specific gene (Dixon et al. 2012; Rao et al. 2014).

Considering that IMR90 has a high fraction of PMDs and its high-resolution Hi-C data are available (Lister et al. 2009; Rao et al. 2014), this cell line was used to check the difference of spatial organization between PMD and non-PMD. We first demonstrated that the majority of PMDs coincide with TADs (Supplemental Figs. S11 and S12).

We then examined how the contact probability derived from Hi-C experiments depends on genomic distance (Fig. 3A). This dependence is significantly different between PMD and non-PMD. In contrast to non-PMD, PMD chromosomes display a much slower decrease in contact probability which means that PMDs are more compact than non-PMDs in kilobase-megabase scale. We also showed that our conclusion is robust using Hi-C data from separate sources (Supplemental Fig. S13).

The Hi-C data analysis highlights the necessity of considering the heterogeneous character of chromatin in Hi-C based polymer modeling.

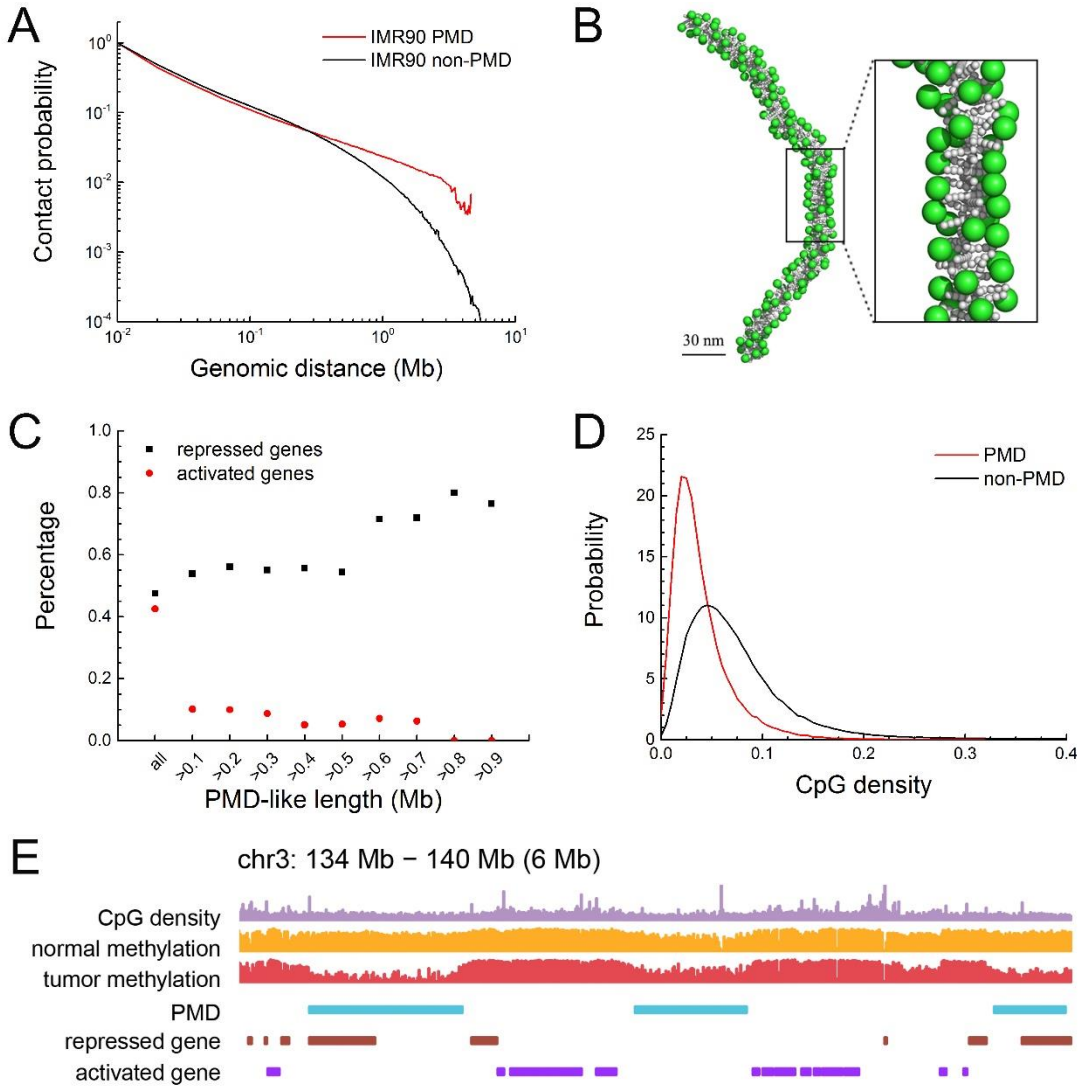

Figure 3. Relation between structure, gene expression and sequence property of different genomic regions. A, Contact probability derived from the Hi-C data as a function of genomic distance in IMR90. **B,** Representative conformation of the simulated chromatin structure for PMD. Green and gray spheres represent nucleosomes and the linker DNA chain, respectively. **C,** Percentage of genes that are transcriptionally repressed or activated in oncogenesis as a function of PMD length. **D,** The CpG density correlates with the formation of PMD. **E,** Plot of a representative region showing the relation among CpG density, methylation levels before and after oncogenesis, PMD, and repressed and activated genes in oncogenesis. In **C, D and**

E, the data of breast cells were used.

Molecular modeling confirms the different chromatin structure in PMD and non-PMD

To further contrast the molecular structure of PMD and non-PMD, we reconstructed the spatial organization of PMD chromatin using molecular modeling (see Methods).

In light of the periodicity in PMD, we infer that chromatin in PMD prefer nucleosomes with similar repeat length, at least more so than non-PMD. We sampled the conformation of chromatin composed of 300 nucleosomes with uniform repeat length using self-avoiding random walk simulation. The analysis of the sampled structure ensemble shows that PMD has a structure resembling a 30 nm chromatin fiber (Fig. 3B and Supplemental Fig. S15).

We removed nucleosomes randomly from the above models to undermine the periodic behavior of chromatin organization to model chromatin structure of non-PMD. It can be clearly seen that PMD has a more compact structure than non-PMD (Supplemental Fig. S16).

Discussion

Under the notion that the correlations of DNA methylation are affected by the spatial organization of the chromatin, one would expect that the spatial organization of PMD is more ordered and compact than non-PMD. Such a conjecture is tested using

Hi-C data and molecular modeling.

Therefore, the methylation correlation power law scaling and periodicity, gene expression levels, Hi-C data and molecular modeling all suggest that PMD has a more regular and compact structure than non-PMD. Since the periodic behavior of PMD-like region lies between PMD and non-PMD, one would expect that it has a medium compactness.

Consistent to its resemblance to PMD, it was noted before that PMD-like domains tend to contain genes with repressed expression levels compared to non-PMDs (Schultz et al. 2015). Here we found that genes are more likely to be repressed in long PMD regions after oncogenesis (Fig. 3C). Regulation of gene activity involves a wide range of mechanisms, including the interplay between transcription factors, chromatin structures and regulatory RNA (Carlberg and Molnár 2014).

Our results suggest that with the increasing of PMD or PMD-like length, the middle genomic regions are buried to result in gene repression (which probably also impedes the binding with TFs, RNA polymerase or other regulators). Interestingly, PMD or PMD-like domains tend to lie in genomic regions with lower CpG density (Fig. 3D and E).

The above analyses show that the DNA methylome can be used to infer the chromatin structure of different cell classes, which shows clear differences among them. We thus propose a simple method to discern cell classes by clustering on the different scaling exponents of chromosomes.

Hierarchical clustering for scaling exponents on all autosomal chromosomes

demonstrates that most chromosomes behave similarly within each cell class, while chromosome 14 and 21 tend to always have a higher scaling exponent (Fig. 4). It can be clearly seen that cancer cells segregate from normal somatic cells (Fig. 4A). Systematic differences are also clearly seen among normal brain cells, glioblastoma and neurodegenerative diseases (Fig. 4B).

The observation that different types of neurodegenerative diseases behave similarly points to similarities in their chromatin structure changes. On the other hand, the neurodegenerative diseases behave significantly different from glioblastoma, highlighting their different pathogenesis.

More detailed analyses of the structure change of each individual chromatin is expected to provide more information on the pathological changes at the gene levels. In addition, the scaling exponent also clearly distinguishes embryonic stem (ES) cells and induced pluripotent stem cells (iPSCs) from somatic cell lines and adult stem cell lines (Fig. 4C), suggesting a large chromatin structure difference between them.

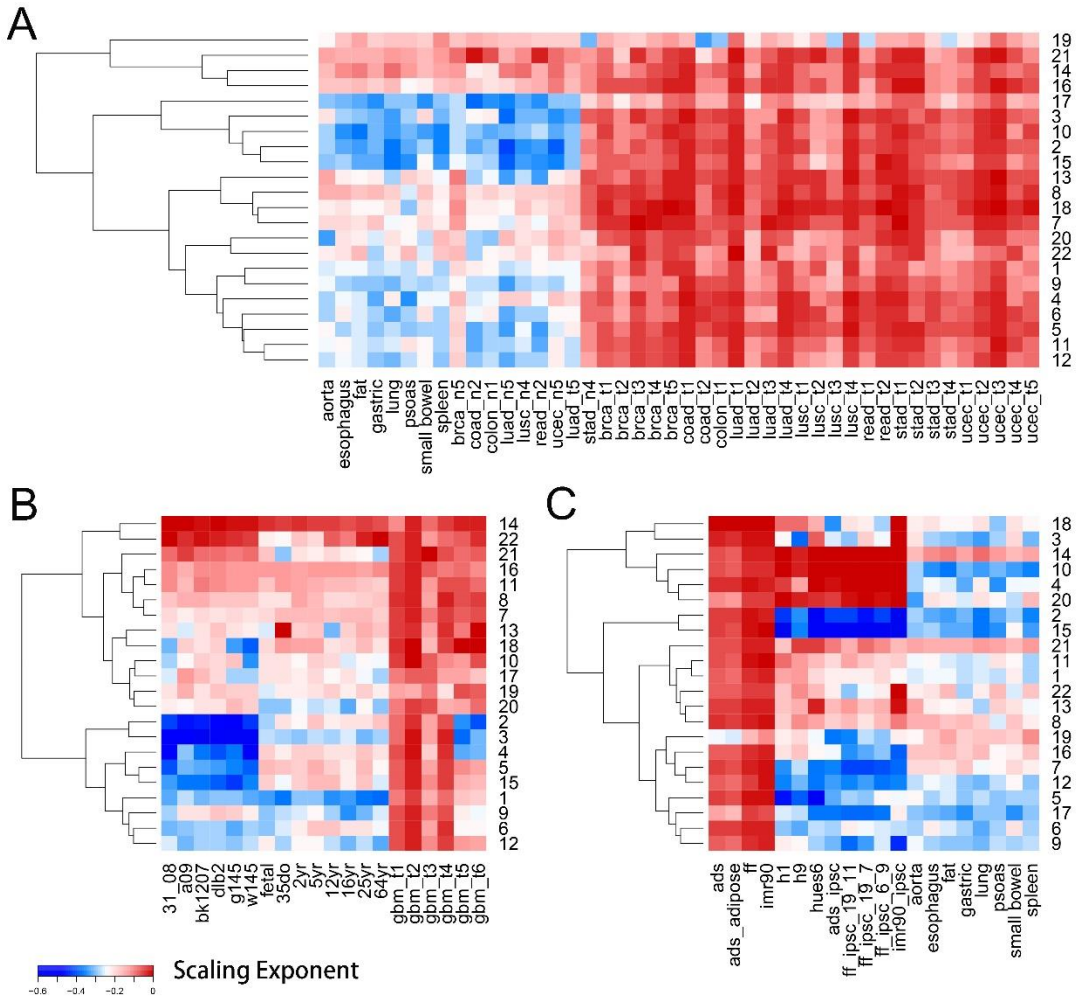

Figure 4. Heatmap clustering for the scaling exponents of all autosomal chromosomes shows differences between different cell classes. Sample labels are summarized in Supplementary Information. **A**, Normal somatic cells (left to right: aorta to ucec_n5) segregate from cancer cells (brca_t1 to ucec_t5). luad_t5 and stad_n4 are further analyzed in Supplementary Information. **B**, Normal brain cells (fetal to 64yr) segregate from glioblastoma (gbm_t1 to gbm_t6) or neurodegenerative diseases (31_08 to w145). **C**, ES cells and iPSCs (h1 to imr90_ipsc) segregate from adult stem cell line (ads) and somatic cell lines (ads_adipose, ff, imr90). Normal somatic cells (aorta to spleen) were shown again for comparison.

In summary, through exploiting the chromatin structure underlying the long-range and local correlations of DNA methylome, our study provides a comprehensive view of the flow of genetic information, connecting DNA sequence, CpG methylation, local and long-range chromatin structure, and gene expression. In normal somatic cells, DNA sequences with high CpG content correlate with compact chromatin structures with low methylation levels and low expression levels (PMD-like). The development of cancers is associated with further decrease of the average methylation level in PMD-like regions, some of which turn into PMDs containing further suppressed genes. The correlation of methylation shows consistent differences among different classes of cells, including normal somatic cells, cancer cells, brain cells, gland cells and stem cells, but are highly conserved within each class. The clear cell class dependence of the long-range power law scaling in methylation correlation shows that it can serve as a simple measure to discriminate between cells at normal and pathological states. Such a finding points to a new direction in the analysis of development of different diseases, for example, cancers and neurodegenerative diseases at the chromatin level.

Methods

Sources of Whole-genome Bisulfite Sequencing Data

All the methylomes used in this study were summarized in Supplementary Information, including the references, URLs, sample details. Hg18 reference genome was used for human brain cells and human ESCs. The other cells used Hg19 as reference

genome.

Calculation of DNA Methylation Correlation

Pearson correlation was used to describe the correlations in DNA methylation. In calculating the long-range correlation, the methylation level was first averaged using a 200-bp window.

Identification of PMDs

PMDs were identified genome-wide in tumor samples using a sliding window approach (Lister et al. 2009). The window size was set as 10 kb. A region was identified as a PMD if there were at least 10 methylated (beta value greater than 0) CpG dinucleotide within, of which the average methylation level was less than 0.7. The contiguous PMD windows were then merged into a longer PMD. Only PMDs with lengths longer than 100-kb are used in the following analysis.

Identification of PMD-like Regions

PMD-like regions were identified in corresponding adjacent non-tumor samples as the corresponding genomic regions of tumor PMDs.

Identification of non-PMD Regions

Non-PMD regions were identified based on the differences of the methylation level between tumor samples and corresponding adjacent non-tumor samples. Only the CpG sites that were detected in both tumor and non-tumor samples were taken into consideration. The genome was divided into 10-kb windows. If the window contains more than 40 CpG sites with methylation level differences smaller than 0.05, the window is identified as non-PMD. The contiguous non-PMD windows were then

merged into a longer non-PMD.

Calculation of Scaling Exponents

The scaling exponents of the long-range power law correlations were calculated as the maximum slope of the fitted double-log correlation data in the genomic region of 2 kb to 0.2 Mb.

Fast Fourier Transform of the Local Correlation

Fast Fourier transform was performed on the local correlation averaged over all selected regions. To avoid the finite length effect and influences of length distribution of genomic regions, we chose PMDs with a genomic length of 1 to 2 Mb from all breast cancer samples, corresponding PMD-like regions from the breast normal sample, and non-PMDs with a genomic length of 1 to 2 Mb from the brca_t1 sample to calculate PMD, PMD-like and non-PMD Fourier transform, respectively.

Gene Expression Analysis

Gene expression of cancer samples was analyzed. Genes of which the promoter and gene body intersects with tumor PMDs were identified and then the TPM (transcripts per million) differences in expression level between normal and tumor samples were calculated for each gene selected. These genes were divided into three categories: activated (difference > 0.005), repressed (difference < -0.005) and the rest, unchanged. We also defined the gene density of the genome and the specific genomic regions like PMDs as the number of genes per million base pairs. Gene functional classification was carried out using DAVID Functional Annotation Bioinformatics Microarray Analysis (Huang et al. 2009). Housekeeping genes list is

downloaded from https://www.tau.ac.il/~elieis/HKG/ (Eisenberg and Levanon 2013).

Molecular Modeling of Chromatin Structure

Nucleosomes and linker DNA chains were described by Four-Sphere model and Worm-Like-Chain model, respectively. See more details in Supplementary Information.

Hi-C data Analysis

The Hi-C data for IMR90 (10-kb resolution) was used under the GEO accession number GSE63525 (Rao et al. 2014). See more details in Supplementary Information.

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank Fu-Chou Tang, Chengqi Yi and Xiaoliang S. Xie for helpful comments on the manuscript. The work was supported by National Science Foundation of China (21233002, 21573006 and 91427304) and Ministry of Science and Technology of China (2012CB917304). The results shown here are partly based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/. We are grateful to M. Esteller for providing the methylome of neurodegenerative diseases.

Author Contributions Y.Q.G designed and supervised the project. L.Z.,W.J.X.,S.L. and L.M. analyzed the DNA methylation data; L.Z. performed gene expression analysis; W.J.X. analyzed the Hi-C data; S.L. performed detrended fluctuation analysis; L.M. performed molecular modeling; G.C. contributed to the analysis of the data. W.J.X. drafted the manuscript.

Competing financial interests The authors declare no conflict of interest.

References

Aranda S, Mas G, Di Croce L. 2015. Regulation of gene transcription by Polycomb proteins. *Sci Adv* **1**: e1500737.

Arneodo A, Vaillant C, Audit B, Argoul F, d'Aubenton-Carafa Y, Thermes C. 2011. Multi-scale coding of genomic information: From DNA sequence to genome structure and function. *Phys Rep* **498**: 45-188.

Bell AC, Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405**: 482-485.

Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, Noushmehr H, Lange CP, van Dijk CM, Tollenaar RA et al. 2012. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat Genet* **44**: 40-46.

Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu CT, Zhuang X. 2016. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**: 418-422.

Carlberg C, Molnár F. 2014. *Mechanisms of Gene Regulation*. Springer.

Cedar H, Bergman Y. 2009. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* **10**: 295-304.

Chodavarapu RK, Feng S, Bernatavichute YV, Chen P-Y, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ et al. 2010. Relationship between nucleosome positioning and DNA

methylation. *Nature* **466**: 388-392.

Clauset A, Shalizi CR, Newman MEJ. 2009. Power-law distributions in empirical data. *SIAM Rev* **51**: 661-703.

Dekker J, Rippe K, Dekker M, Kleckner N. 2002. Capturing chromosome conformation. *Science* **295**: 1306-1311.

Diesinger PM, Heermann DW. 2009. Depletion effects massively change chromatin properties and influence genome folding. *Biophys J* **97**: 2146-2153.

Dixon JR, Gorkin DU, Ren B. 2016. Chromatin domains: the unit of chromosome organization. *Mol Cell* **62**: 668-680.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376-380.

Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C et al. 2006. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res* **16**: 1299-1309.

Eisenberg E, Levanon EY. 2013. Human housekeeping genes, revisited. *Trends Genet* **29**: 569-574.

Eugene VK, Yuri IW, Georgy PK. 2006. *Power Laws, Scale-Free Networks and Genome Biology*. Springer.

Fortin JP, Hansen KD. 2015. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol* **16**: 180.

Gaidatzis D, Burger L, Murr R, Lerch A, Dessus-Babus S, Schubeler D, Stadler MB. 2014. DNA sequence explains seemingly disordered methylation levels in partially methylated domains of Mammalian genomes. *PLOS Genet* **10**: e1004143.

Galupa R, Heard E. 2015. X-chromosome inactivation: new insights into cis and trans regulation. *Curr Opin Genet Dev* **31**: 57-66.

Gibcus JH, Dekker J. 2013. The hierarchy of the 3D genome. *Mol Cell* **49**: 773-782.

Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, Tiana G, Heard E. 2014. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **157**: 950-963.

Hon GC, Hawkins RD, Caballero OL, Lo C, Lister R, Pelizzola M, Valsesia A, Ye Z, Kuan S, Edsall LE et al. 2012. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res* **22**: 246-258.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44-57.

Jenuwein T, Allis CD. 2001. Translating the histone code. *Science* **293**: 1074-1080.

Levine M, Cattoglio C, Tjian R. 2014. Looping back to leap forward: transcription enters a new era. *Cell* **157**: 13-25.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**: 289-293.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM et al. 2009. Human DNA methylomes at base resolution show widespread

epigenomic differences. *Nature* **462**: 315-322.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**: 766-770.

Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J. 2013. Organization of the mitotic chromosome. *Science* **342**: 948-953.

Peng CK, Buldyrev SV, Goldberger AL, Havlin S, Sciortino F, Simons M, Stanley HE. 1992. Long-range correlations in nucleotide sequences. *Nature* **356**: 168-170.

Pombo A, Dillon N. 2015. Three-dimensional genome architecture: players and mechanisms. *Nat Rev Mol Cell Biol* **16**: 245-257.

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**: 1665-1680.

Ricci MA, Manzo C, Garcia-Parajo MF, Lakadamyali M, Cosma MP. 2015. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell* **160**: 1145-1158.

Schneider R, Grosschedl R. 2007. Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes Dev* **21**: 3027-3043.

Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Urich MA, Chen H et al. 2015. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523**: 212-216.

Zhang B, Wolynes PG. 2015. Topology, structures, and energy landscapes of human chromosomes. *Proc Natl Acad Sci USA* **112**: 6062-6067.