TRANSCRIPTION, 2018
VOL. 9, NO. 3, 171–176
https://doi.org/10.1080/21541264.2017.1373209

Taylor & Francis
Taylor & Francis Group

POINT-OF-VIEW

Orphan CpG islands as alternative promoters

Shrutii Sarda and Sridhar Hannenhalli

Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD, USA

ABSTRACT
CpG islands (CGIs) are associated with ~60% of mammalian promoters. Most unmethylated CGIs exhibit transcriptional activity, which has led to their co-option as promoters by retrogenes. CGIs may also serve as alternative promoters for downstream genes with methylated promoters, with implications on aberrant activation of oncogenes in cancer phenotypes.

ARTICLE HISTORY
Received 25 July 2017
Revised 23 August 2017
Accepted 24 August 2017

KEYWORDS
CpG islands; DNA methylation; alternative promoters; evolutionary selection; human

Background

Vertebrate genomes are typically depleted of CG dinucleotides due to spontaneous deamination of cytosines at methylated CG dinucleotides ($^{5m}C_pG$) resulting in a CG to TG mutation. Yet, a disproportionately large fraction of CGs are concentrated in so-called CpG islands (CGIs).[1] For instance, in the human genome, CGIs have on average 1 CpG every 10 base pairs (bps), which is about 10 times more frequent than the surrounding DNA. CGIs are 200–1000 bps long, exhibit an elevated G+C base composition as well as frequent absence of DNA methylation.[2]

The largely unmethylated state of CGs in a CGI explains, in part, how they escape deamination and mutation, and it is likely that their evolutionary maintenance may be a consequence of their functional importance. Since their discovery, CGIs have been implicated in a vast array of fundamental processes such as DNA replication (they act as the origins of replication; the sequences themselves are potentially genomic footprints left on the chromosome by replication events[3]), imprinting (they are differentially methylated in an allele-specific manner[4]), as well as transcriptional regulation (they primarily act as sites for RNA Pol II recruitment and transcription initiation[5]). The last of these affects a broad range of cellular functions and will be the focus of this review.

Genomic landscape of CGIs

Based on restriction fragment length distributions upon treatment with a methyl-sensitive restriction endonuclease Hpa II, Antequera and Bird estimated ~45000 and 37000 CGIs in human and mouse genomes respectively.[6] However, later studies found the number of CGIs in human and mouse to be comparable.[7] Based on these initial experimental observations, CGIs are operationally defined as a genomic regions with length $\geq$ 200 bp, CG composition $\geq$ 50%, and observed over expected CpG ratio $\geq$ 0.60,[2] however, other variants of this criteria have been used as well.

CGIs are associated with the transcription start sites (TSS) of 50–70% of mammalian genes, and these CGIs have historically served as genomic markers of genes and promoters.[8,9] However, about half of CGIs in mammalian genomes are not associated with a known gene promoter (i.e., either intra- or intergenic) and are referred to as *orphan* CGIs.[7] The orphan CGIs nevertheless are bound by RNA polymerase II and exhibit transcriptional activity in a conserved fashion between human and mouse.[7] Further, unlike CGIs associated with gene promoters, which are generally unmethylated, orphan CGIs exhibit a more dynamic pattern of methylation and are preferentially methylated during early development.[7] Interestingly, most of the evolutionarily conserved methylation differences between tissues are enriched at orphan CGIs, suggesting that

CONTACT Sridhar Hannenhalli sridhar@umiacs.umd.edu 3104G Biomolecular Sciences Building, University of Maryland, College Park, MD 20742;
Shrutii Sarda ssarda@umiacs.umd.edu 3104E Biomolecular Sciences Building, University of Maryland, College Park, MD 20742.

they are regulated and may be involved in tissue specification.[10]

Mechanistic link between CGI and transcription

Interpreting how DNA methylation modulates gene expression is complicated,[11] as the epigenetic mark co-occurs with distinct histone post-translational modifications (PTMs) in a site-specific manner, which are themselves associated positively or negatively with gene expression. In combination with them, DNA methylation establish several unique chromatin signatures throughout the genome,[11] such as heterochromatin domains in repeat regions (with H3K9me3), bivalent activity domains at gene promoters (with H3K27me3), high expression domains at gene-bodies (with H3K36me3), etc. In CGI promoters, broadly speaking, unmethylated CGIs are associated with transcriptional activity while in the methylated state they are associated with transcriptional silencing. Unmethylated CGIs serve as a substrate for several activating transcription factors (TF)[12] and chromatin modifying enzymes. Upon methylation, the methyl moieties on the DNA interfere with these interactions,[13] and additionally facilitate recruitment of proteins associated with histone deacetylase leading to condensed chromatin.[14,15] While CGI methylation has been widely shown to be associated with lack of transcriptional activity, the causal links between methylation and transcription are not entirely clear, and it appears that methylation is likely to be one component in a more complex regulatory hierarchy[16] in vertebrates. Moreover, the details of the role of methylation in regulating non-CGI promoter activity are not well understood.

Notwithstanding, recent advances in high-throughput transcriptome quantification have revealed, strikingly, that even the unmethylated orphan CGIs exhibit at least a basal level of transcriptional activity,[5] due to their inherent ability to recruit and bind transcriptionally engaged RNA polymerase II. These findings emphasize the strong correlation between CGIs and transcription initiation, which is a key property relevant to co-option of orphan CGIs by retrogenes and otherwise silenced genes, as discussed below.

Origin, evolution, and maintenance of CGIs

The events leading to the appearance of CGIs in vertebrate genomes are not clear. Further, what are the forces that protect them from *de novo* methylation, to spontaneous deamination, and to ultimately being lost across evolutionary time? There have been a few different theories, including (i) the intrinsic ability to refract methylation, and (ii) active demethylation,[17] neither of which gained much traction as it has been shown that (i) CGIs can and do become methylated during normal development,[7] and (ii) mechanisms of active demethylation in animals are still unresolved.[17] Recent studies have suggested that the binding of transcription factors, or the act of transcription itself during early development is required for establishment of the methylation-free state of CGIs.[5]

During embryonic development, *de novo* DNA methylation facilitated by Dnmt3a and Dnmt3b methyltransferases occurs in waves, and its appropriate placement in the genome is essential for normal development and cellular function. While there is very little mechanistic understanding of how targeting of *de novo* methylation occurs, it has been shown that Dnmt3a/b targeting is dependent, in some contexts, on pre-existing histone methylation.[18] For example, H3K36 methylation correlates with enrichment of DNA methylation; Dnmt3a/b is recruited through its PWWP domain,[18] which preferentially binds to H3K36me3. Another histone modification, H3K4me3, might block *de novo* methylation, as the Dnmt3-associated protein Dnmt3L[17] specifically interacts only with unmodified H3K4. It has been speculated therefore that the presence of H3K4me3 at CGIs, via repulsion of Dnmt3,[17] might be responsible for the persistence of their hypomethylated state.

A bioinformatic analysis of promoter GC composition distribution across invertebrates, chordates, and vertebrates shows the appearance of divergent GC composition with a clear bimodal distribution specifically in warm blooded vertebrates, linking their evolution to CGIs.[19] Further, many genes with CGI promoters are expressed in the germ line and during early embryonic development, which may provide selection pressure to preserve the CGIs. CGIs are linked not only to active gene promoters in the germ line, but they are also linked with the origin of replication, leading to speculations that occupancy of these loci by transcriptional and replication machinery might protect them from methylation and eventual loss.[20] Curiously, even amongst closely related species, such as human and mouse, promoters of key conserved genes can have different CGI status, notably,

α-globin gene in human has a CGI promoter while its mouse counterpart does not.[6] Thus, the details of the origin, evolution, and maintenance of CGIs are currently not fully resolved.

Evolutionary co-option of CGIs as promoters

Recent work has established a biochemical connection between the abundance of CpG in CGIs and H3K4me3, mediated by a CXXC domain protein, called Cfp1 that binds specifically to nonmethylated CpGs.[21] In fact, insertion of an artificial CGI-like DNA sequence into the genome leads to recruitment of Cfp1 resulting in H3K4me3 deposition.[5] This is particularly important, as H3K4me3 is a signature histone mark of transcriptionally active gene promoters.[22] Thus, equipped with the ability to independently set up an active chromatin state, recruit RNA polymerase II and initiate transcription, CGIs globally appear to be predisposed for active promoter function. Indeed, as mentioned earlier, about half of all CGIs in the genome are associated with proximal gene promoters. In fact, there is evidence to suggest that even orphan CGIs can be utilized for their 'promoter-potential' across evolution. Very recently, a fascinating study[23] that mapped the processes underlying the evolution of retrogenes (intronless and promoterless copies of reverse transcribed RNA inserted into the genome) into new *bona fide* functional genes, discovered that only a marginal fraction (~11%) of these piggybacked on existing gene promoters for their expression, while the majority of novel retrogene promoters (~86%) actually overlapped orphan CGIs and other proto-promoter elements.[23] Furthermore, as these retrogenes emerged into fully functional genes, most (75%–93%) gained new exons from their upstream flanking sequences; and this overrepresentation of novel 5' exons suggests that such a gain served to place retrogenes under the transcriptional control of upstream orphan CGIs (see Fig. 1A).

CGIs as alternative cell-type-specific promoters

In addition to co-option as promoters by retrogenes, orphan CGIs function as promoters in other regulatory contexts as well. Several specific examples, showing promoter activity at orphan CGIs in the context of critical functions like genomic imprinting, development and cell differentiation have been reported.[10,24,25] Analysis of the imprinting control region (ICR) of a telomeric gene cluster in mouse, and its orthologous

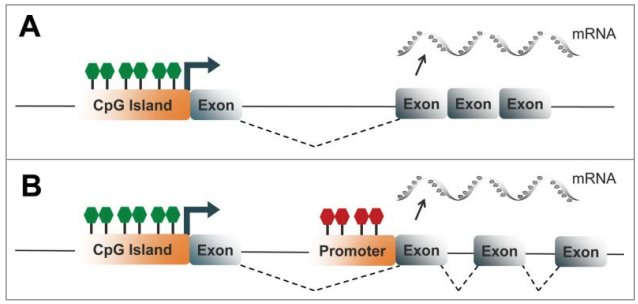

Figure 1. Illustration depicting usage of upstream CGIs as alternative promoters by retrogenes, and genes with methylated proximal promoters. (A) A retrocopy comprising 3 original concatenated exons gained a new 5′ exon as it comes under the transcriptional control of an upstream orphan CGI. This co-opted CGI promoter allows the inactive retrocopy to emerge as a functional retrogene. The CGI is depicted in an unmethylated state (hypomethylation indicated by green flags) (B) A gene with a silenced proximal promoter (hypermethylation indicated by red flags) uses an upstream orphan CGI as an alternative promoter.

region in human revealed that a ~2kb CpG island located inside intron 10 of one these genes (Kcnq1) was a previously unknown promoter for a large non-coding anti-sense RNA (Kcnq1ot1).[25] This RNA, whose transcription is controlled by the intragenic CGI is required for imprinting of the gene cluster; deletion of the CGI on a paternally inherited human chromosome *in vitro* led to increased expression of several genes in the cluster implying loss of imprinting.[25] Additionally, a recent study showed that promoter-distal CGIs exhibited the most tissue-specific methylation patterns, and were linked to the tissue-specific initiation sites of alternative transcripts within gene bodies and non-coding RNAs in intergenic regions.[26]

Alternative promoter activity was detected at yet another upstream CGI that expresses a specific isoform of the Epac2 gene, namely Epac2A in a tissue-specific fashion.[27] This isoform is expressed specifically in the brain, pituitary and pancreatic islets, although bisulfite sequencing revealed that the Epac2A-promoter at the upstream CGI was virtually completely demethylated in all tissues tested, i.e., in Epac2A expressing tissues, in tissues that express other Epac2 isoforms as well as in non-expressing tissues. This finding is in agreement with the repeated observation that the majority of CGI promoters are free of methylation regardless of the expression level of the associated gene.[28,29] Yet, there are many instances where epigenetic control of transcriptional activity through differential methylation of CGI promoters has been observed. For example, tissue-specific activity

of ubiquitin-specific peptidase USP44 was correlated with DNA methylation of a CGI close to its primary promoter.[30] Thus, although the functional importance of orphan CGIs as cell-type specific promoters is now established, both the landscape of these events, as well as the mechanisms underlying their regulation are not well understood.

We recently reported the ability of intergenic orphan CGIs located tens of kilobases upstream of methylated-promoter genes to serve as their cell-type specific alternative promoters[31] (see Fig. 1B). Such CGI-initiated transcription explains the expression of about half of the genes with highly methylated proximal promoters, as observed across 34 human cell types. We showed that in these cases, there was an explicit lack of transcription initiation at the annotated promoter accompanied by strong initiation signal at the upstream CGI. Further, these CGI-initiated transcripts were associated with signals of stable elongation and splicing that extend into the gene body, supporting the observation that these were true alternative transcripts. Genes that exhibit such orphan CGI alternative promoter usage, appear to use their proximal promoter for expression in other tissues when it is unmethylated. While the mechanism regulating the choice of use of proximal (which was non-CGI type in most of these cases) versus distal CGI promoters remains undetermined, based on our preliminary analyses, we are inclined to believe that this occurs independent of the methylation status of the proximal promoter. It appears that an already active orphan CGI was co-opted as an alternative promoter, analogous to the co-option by retrogenes discussed above. Our analysis of conserved synteny between these genes and their upstream CGIs suggest that this co-option is likely to have occurred relatively recently close to the divergence of mammals from the vertebrates.[31] Additional analyses explicitly testing the potential of these CGIs to serve as alternative promoters for orthologous genes in a wide range of vertebrate species would be necessary to further substantiate this assertion.

We also assessed the effect of global demethylation on the activity of proximal promoters of genes transcribed by alternative distal CGIs. Promoter usage patterns of genes expressed by distal CGIs in wild-type (WT) mouse embryonic stem cells were analyzed in cells without DNA methyltransferase activity (DNMT triple knock out; TKO). Proximal promoter methylation is lost in the DNMT TKO phenotype, and if the distal CGI was the only promoter of these genes, then removing methylation at the annotated promoters should not lead to a change in their activity status. Instead, we observed that there was increased transcription in TKO immediately downstream of these proximal promoters relative to WT (based on evidence from relative RNAseq read densities), suggesting some form of transcriptional activation at these sites. However, as these insights were not derived from data that directly quantifies transcriptional initiation levels (such as CAGE, GRO-seq or Ser5 Pol II ChIP-seq), they are not conclusive. It is not clear if increased read density is merely a reflection of noisy transcription due to de-repression of proximal promoters, or if they correspond to true gene transcripts, and therefore these results hinting at a potential switch from the usage of distal CGIs to the proximal promoters should be considered with caution.

Implication in cancer

CGI methylation has been previously studied in the context of cancer as it pertains to aberrant gene expression. Even though *de novo* methylation during development predominantly affects orphan CGIs, this is not the case in cancer. Illingworth et al. found that cancer-specific *de novo* methylation affects gene-associated CGIs as well as orphan CGIs equally,[7] and these events do not recapitulate the methylation changes seen during normal development. Interestingly, upon contrasting matched normal and colorectal cancer samples, it became apparent that a significantly large fraction of the cancer-specific CGI methylation events[7] was shared by all tumors (39%), suggesting that there might exist some yet unexplored cancer-specific regulatory program affecting the activities of CGIs and their associated genes, as indicated by the common methylation patterns exhibited by cancer samples.

Aside from the above, it has also been widely recognized that cancer phenotypes are associated with global patterns of hypomethylation.[32] We investigated in breast and kidney tumors the extent to which upstream CGIs are used as alternative promoters (as a result of global loss of methylation in cancer) by genes with methylated proximal promoters and found that the aberrant gene expression patterns of several clinically important genes[31] can be explained by this phenomenon. We further found that, on average, the

transcription of approximately 10% of all oncogenes active in a breast cancer patient is initiated by upstream orphan CGIs acting as alternative promoters, rather than by their proximal promoters. Thus, activation of orphan CGIs due to global hypomethylation in cancer may contribute to cancer progression.

Evolutionary advantage of orphan CGIs

Aside from the knowledge that about half of all CpG islands (CGIs) are remote from gene promoters, the biological significance of these 'orphan' CGIs remains largely elusive. Orphan CGIs resemble promoter CGIs in their promoter-like state and general lack of DNA methylation in the germline.[7] Their evolutionary maintenance, as well as their enhanced propensity[7] to become dynamically methylated during development is suggestive of their functional importance. We suggest a potential function[31] for a small subset of these as tissue-specific alternative promoters for downstream genes. For this subset of orphan CGIs, a higher relative conservation, both across species and within the human population, as well as conserved synteny with the downstream gene[31] is suggestive of their functional importance. The tissue-specific nature of their usage as alternative promoters leads us to speculate that these CGIs may have been selected to drive robust expression profiles of their associated genes in specific tissues. Perhaps, placing the control of expression of certain genes under CGI alternative promoters, rather than their proximal non-CGI promoters, allows for finer control of downstream activation (while not necessarily achieving higher levels of expression). Also, given that these genes typically have CpG-poor proximal promoters which tend to elicit transcription in relatively narrow contexts, it is possible that usage of distal orphan CGI promoters might allow relatively broader, and more robust, expression profiles. Although based on current knowledge of CGI alternative promoters and their usage patterns, it is not possible to deduce the implications of this phenomenon.

In summary, even though the roles of CpG methylation, in the context of CGIs in particular, in a variety of functions ranging from transcription, DNA replication, and imprinting are widely recognized, the functions of individual CGIs remain underappreciated. Many orphan CGIs may play critical, hitherto undiscovered roles in tissue-specific transcriptional control, and may underlie pathogenic transcription, especially in cancer with broad epigenomic changes.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

References

1. Bird A, Taggart M, Frommer M, Miller OJ, Macleod D. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell. 1985; 40(1):91–99. doi:10.1016/0092-8674(85)90312-5. PMID:2981636
2. Gardiner-Garden M, Frommer M. CpG Islands in vertebrate genomes. J Mol Biol. 1987; 196(2):261–82. doi:10.1016/0022-2836(87)90689-9. PMID:3656447
3. Antequera F, Bird A. CpG islands as genomic footprints of promoters that are associated with replication origins. Curr Biol. 1999; 9(17):R661–7. doi:10.1016/S0960-9822(99)80418-7. PMID:10508580
4. Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature. 1997; 389(6652):745–9. doi:10.1038/39631. PMID:9338788
5. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011; 25(10):1010–22. doi:10.1101/gad.2037511. PMID:2157626
6. Antequera F, Bird A. Number of CpG islands and genes in human and mouse. Proc Natl Acad Sci. 1993; 90(24):11995–9. doi:10.1073/pnas.90.24.11995
7. Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr ARW, James KD, Turner DJ, Smith C, Harrison DJ, Andrews R, Bird AP. Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet. 2010; 6(9):e1001134. doi:10.1371/journal.pgen.1001134. PMID:20885785
8. Hannenhalli S, Levy S. Promoter prediction in the human genome. Bioinformatics. 2001; 17(Suppl 1):S90–S96. doi:10.1093/bioinformatics/17.suppl_1.S90. PMID:11472997
9. Larsen F, Gundersen G, Lopez R, Prydz H. CpG islands as gene markers in the human genome. Genomics. 1992; 13(4):1095–107. doi:10.1016/0888-7543(92)90024-M. PMID:1505946
10. Illingworth R, Kerr A, DeSousa D, Jørgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, et al. A novel CpG island set identifies tissue-specific methylation at developmental gene Loci. PLoS Biol. 2008; 6(1):e22. doi:10.1371/journal.pbio.0060022. PMID:1823273
11. Rothbart SB, Strahl BD. Interpreting the language of histone and DNA modifications. Biochim Biophys Acta – Gene Regul Mech. 2014; 1839(8):627–43. doi:10.1016/j.bbagrm.2014.03.001
12. Weinmann AS, Yan PS, Oberley MJ, Huang TH-M, Farnham PJ. Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island

microarray analysis. Genes Dev. 2002; 16(2):235–44. doi:10.1101/gad.943102. PMID:11799066
13. Becker PB, Ruppert S, Schütz G. Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. Cell. 1987; 51(3):435–43. doi:10.1016/0092-8674(87)90639-8. PMID:2889531
14. Wu X, Johansen JV, Helin K. Fbxl10/Kdm2b Recruits polycomb repressive complex 1 to CpG islands and regulates H2A Ubiquitylation. Mol Cell. 2013; 49(6):1134–46. doi:10.1016/j.molcel.2013.01.016. PMID:23395003
15. Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012; 13(7):484–92. doi:10.1038/nrg3230. PMID:2264101
16. Razin A, Riggs A. DNA methylation and gene function. Science. 1980; 210(4470):604–10. doi:10.1126/science.6254144
17. Illingworth RS, Bird AP. CpG islands – "A rough guide." FEBS Lett. 2009; 583(11):1713–20. doi:10.1016/j.febslet.2009.04.012. PMID:19376112
18. Rose NR, Klose RJ. Understanding the relationship between DNA methylation and histone lysine methylation. Biochim Biophys Acta – Gene Regul Mech. 2014; 1839(12):1362–72. doi:10.1016/j.bbagrm.2014.02.007
19. Sharif J, Endo TA, Toyoda T, Koseki H. Divergence of CpG island promoters: A consequence or cause of evolution? Dev Growth Differ. 2010; 52(6):545–54. doi:10.1111/j.1440-169X.2010.01193.x
20. Antequera F. Structure, function and evolution of CpG island promoters. Cell Mol Life Sci. 2003; 60(8):1647–58. doi:10.1007/s00018-003-3088-6. PMID:14504655
21. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, Kerr AR, Deaton A, Andrews R, James KD, et al. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature. 2010; 464(7291):1082–6. doi:10.1038/nature08924. PMID:20393567
22. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. Cell. 2007; 130(1):77–88. doi:10.1016/j.cell.2007.05.042. PMID:1763205
23. Carelli FN, Hayakawa T, Go Y, Imai H, Warnefors M, Kaessmann H. The life history of retrocopies illuminates the evolution of new mammalian genes. Genome Res. 2016; 26(3):301–14. doi:10.1101/gr.198473.115. PMID:26728716
24. Deaton AM, Webb S, Kerr ARW, Illingworth RS, Guy J, Andrews R, Bird A. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. Genome Res. 2011; 21(7):1074–86. doi:10.1101/gr.118703.110. PMID:21628449
25. Mancini-DiNardo D, Steele SJS, Ingram RS, Tilghman SM. A differentially methylated region within the gene Kcnq1 functions as an imprinted promoter and silencer. Hum Mol Genet. 2003; 12(3):283–94. doi:10.1093/hmg/ddg024. PMID:12554682
26. Mendizabal I, Yi SV. Whole-genome bisulfite sequencing maps from multiple human tissues reveal novel CpG islands associated with tissue-specific regulation. Hum Mol Genet. 2016; 25(1):69–82. doi:10.1093/hmg/ddv449. PMID:2651206
27. Hoivik EA, Witsoe SL, Bergheim IR, Xu Y, Jakobsson I, Tengholm A, Doskeland SO, Bakke M. DNA methylation of alternative promoters directs tissue specific expression of Epac2 Isoforms. PLoS One. 2013; 8(7):e67925. doi:10.1371/journal.pone.0067925. PMID:23861833
28. Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S, Shu J, Chen XA, Waterland RA, Issa JP. Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet. 2007; 3(10):e181. doi:10.1371/journal.pgen.0030181
29. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M, Schübeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet. 2007; 39(4):457–66. doi:10.1038/ng1990. PMID:17334365
30. Tropel P, Jung L, André C, Ndandougou A, Viville S. CpG island methylation correlates with the use of alternative promoters for USP44 gene expression in human pluripotent stem cells and testes. Stem Cells Dev. 2017; 26(15):1100–10. doi:10.1089/scd.2017.0057. PMID:28520534
31. Sarda S, Das A, Vinson C, Hannenhalli S. Distal CpG islands can serve as alternative promoters to transcribe genes with silenced proximal promoters. Genome Res. 2017; 27(4):553–66. doi:10.1101/gr.212050.116. PMID:2822340
32. Ehrlich M. DNA hypomethylation in cancer cells. Epigenomics. 2009; 1(2):239–59. doi:10.2217/epi.09.33. PMID:20495664