# Conflicts of CpG density and DNA methylation are proximally and distally involved in gene regulation in human and mouse tissues

## Abstract

The relationship between CpG content and DNA methylation has attracted considerable interest in recent years. Direct or indirect methods have been developed to investigate their regulatory functions based on various hypotheses, large cohort studies, and meta-analyses. However, all of these analyses were performed at units of CpG blocks and, thus, the influence of finer genome structure has been neglected. Herein, we present a novel algorithm of base-pair resolution to systematically investigate the relationship between CpG contents and DNA methylation. By introducing the concept of 'complementary index' we examined the methylomes of 34 adult and 7 embryonic tissues and successfully fitted the relationship of DNA methylation and CpG density into a nonlinear mathematical model. A further algorithm was developed to locate the regions where CpG density does not match expectations from the model, termed 'conflict of gap' (COG) regions. Interestingly, COGs are highly concordant in human and mouse and their distributions display a tissue-specific pattern. Based on COG methylation patterns we correctly classified tissues according to their function or origin. We demonstrate that COGs based on our method can reveal more and deeper information than traditional differential methylation region (DMR) approaches. We also found that when COGs are located near to transcription start site (TSS), these regions can determine which promoters will be utilized for initiating gene transcription. Furthermore, COGs located far from the TSS perform as enhancers in terms of histone modification, sequence conservation, transcription factor binding, and DNase I-hypersensitivity..

## Keywords