

Regulatory patterns of histone modifications to control the DNA methylation status at CpG islands

Inkyung Jung¹ and Dongsup Kim^{1,2,*}

¹Department of Bio and Brain Engineering and

²KAIST Institute for BioCentury, KAIST, Daejeon 305-701, S. Korea.

Subject areas: Bioinformatics, Epigenetics

Author contribution: IJ and DK designed the methods and experimental setup. IJ carried out the implementation of the various methods. IJ and DK wrote the manuscript. All authors have read and approved the final manuscript.

***Correspondence** and requests for materials should be addressed to D.S.K. (kds@kaist.ac.kr)

Reviewer: Tae-Young Roh, Postech, Republic of Korea. Sanghyuk Lee, Ewha Womans University, Republic of Korea.

Editor: Sun Shim Choi, Kangwon National University, Republic of Korea

Received March 24, 2009;

Accepted March 30, 2009;

Published March 30, 2009

Citation: Jung, I., et al. Regulatory patterns of histone modifications to control the DNA methylation status at CpG islands. IBC 2009, 1(1):4, 1-7. doi:10.4051/ibc.2009.1.0004

Funding: This work is supported by CHUNG Moon Soul Center for BioInformation and BioElectronics (CMSC), and by Korea Science and Engineering Foundation (grant number: 2008-05442).

Competing interest: All authors declare no financial or personal conflict that could inappropriately bias their experiments or writing.

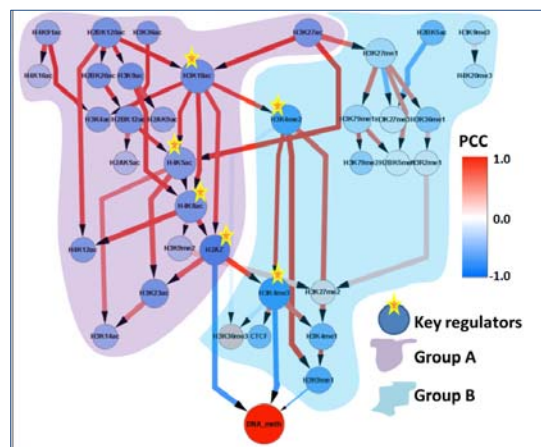
Copyright: This article is licensed under a Creative Commons Attribution License, which freely allows to download, reuse, reprint, modify, distribute, and/or copy articles as long as a proper citation is given to the original authors and sources.

SYNOPSIS

Introduction: Histone modifications and DNA methylation are the major factors in epigenetic gene regulation. Especially, revealing how histone modifications are related to DNA methylation is one of the challenging problems in this field. In this paper, we address this issue and propose several plausible mechanisms for precise controlling of DNA methylation status at CpG islands

Materials and Methods: To establish the regulatory relationships, we used 38 histone modification types including H2A.Z and CTCF, and DNA methylation status at CpG islands across chromosome 6, 20, and 22 of human CD4+ T cell. We utilized Bayesian network to construct regulatory network.

Results and Discussion: We found several meaningful relationships supported by previous studies. In addition, our results show that histone modifications can be clustered into several groups with different regulatory properties. Based on those findings we predicted the status of methylation level at CpG islands with high accuracy, and suggested core-regulatory network to control DNA methylation status.



Keywords: Histone modification, epigenetics, DNA methylation

Introduction

Histone modifications and DNA methylation are responsible for epigenetic regulation. Thousands of genes in an organism are regulated systematically by many different kinds of factors, and epigenetic research aims to understand how those genes are precisely controlled and how epigenetic factors are connected with each other.

Among them DNA methylation and histone modifications are known to play a major role in epigenetic regulation. Modification of histones can provide systematic changes for gene activation or repression through acetylation, methylation, phosphorylation, ubiquitination, and other covalent modifications at their N-terminal tails. In addition, DNA methylation restricts gene expression by addition of methyl group to the cytosine in CpG sites. Thus, DNA methylation and histone modifications modulate the packaging of the DNA in the nucleus, which results in blocking or enhancing the access of transcription factors to DNA. Despite of many intensive studies, the mechanism of how histone modifications affect DNA methylation is not clear.

Recently, development of several genome-wide analysis techniques such as ChIP(chromatin immunoprecipitation)-chip and ChIP-Seq opened the new era for rapid progress in high-throughput studies for histone modifications. The examples of using ChIP-chip are profiling histone modifications in *Saccharomyces cerevisiae* (Bernstein et al., 2002; Robyr et al., 2002) and in mammals (Bernstein et al., 2005; Bernstein et al., 2006). Another technology, ChIP-Seq, gives quantitative modification intensities by counting matched tags on genome. The first application of ChIP-Seq was profiling histone methylation patterns for human (Barski et al., 2007) and mouse stem cell (Mikkelsen et al., 2007), and was extended to profiling histone acetylation patterns (Wang et al., 2008).

Another major epigenetic marker, DNA methylation is known to be involved with gene regulation and chromosomal stability. The GC-rich regions (CpG islands) are resistant to methylation and are associated with gene expression level. Genome-wide studies of DNA methylation have been conducted by using restriction enzymes (Lippman et al., 2004; Selker et al., 2003), bisulphite treatment (Adorjan et al., 2002; Gitan et al., 2002), and MeDIP (Keshet et al., 2006; Weber et al., 2005). Especially, Human Epigenome Project (HEP) produced DNA methylation profiles across human chromosome 6, 20 and 22 of 12 different human tissues by combining sequencing and bisulphite treatment (Eckhardt et al., 2006).

Application of new techniques has led to great advance in our understanding for the epigenetic control mechanism. In addition, a role of several modification patterns has been revealed through genome-wide studies. In computational approaches, analyses of histone modification patterns with gene expression and regulatory networks have suggested new insight for epigenetic regulation. Furthermore, prediction algorithms for methylation status at CpG islands have been developed with sequence-based features (Bock et al., 2006; Das et al., 2006; Fan et al., 2008; Fang et al., 2006), and by combining histone modification information with sequence features (Fan et al., 2008). In several studies, Bayesian network was employed to analyze DNA methylome (Down et al., 2008), and to identify causal relationships between histone methylations and gene expression (Yu et al., 2008).

Amid all these efforts, necessity for establishing a complete regulatory system that includes all relevant factors including histone modifications, DNA methylation, and gene expression becomes more urgent. There were several previous studies to find how DNA methylation is linked to histone modifications. The linker histone binding has been suggested to be one of factors which link DNA

methylation and histone modifications (Gilbert et al., 2007). Also the role of histone modifications has been suggested as cofactors (Tamaru and Selker, 2001). They revealed that the histone deacetylase inhibitor trichostatin A (TSA) causes selective loss of DNA methylation in *Neurospora*, suggesting that histone modification can influence DNA methylation.

In this regard, we focus on identifying relationships between various histone modifications and DNA methylation at CpG islands. By using Bayesian network we identified several known relationships and suggested regulatory patterns of histone modifications that were believed to regulate DNA methylation. Lastly, we predicted the status of DNA methylation with only histone modifications, which supports the close relationship among histone modifications and DNA methylation.

Results and Discussion

Constructing regulatory network among different histone modifications together with DNA methylation at CGIs

Although the relationship between histone modifications and DNA methylation is not clear, many studies support that DNA methylation is regulated by several combinatorial patterns of histone modifications as 'histone codes'. However there are too many potential combinatorial patterns, thereby it is impossible to validate all relationships by wet experiment. For this reason computational approach is needed to build comprehensive model for regulatory network. Thereby, we used Bayesian network to identify relationships among histone methylations, histone acetylations, and DNA methylation. In this study, we focused on uncovering the specific patterns of histone modifications that are related with the regulation of methylation levels at CpG islands (CGIs).

Figure 1 shows the Bayesian network we constructed, showing regulatory relationships between all epigenetic modification factors. From this Figure, we first notice that the regulatory network can be grouped into two parts; one is enriched with histone acetylations

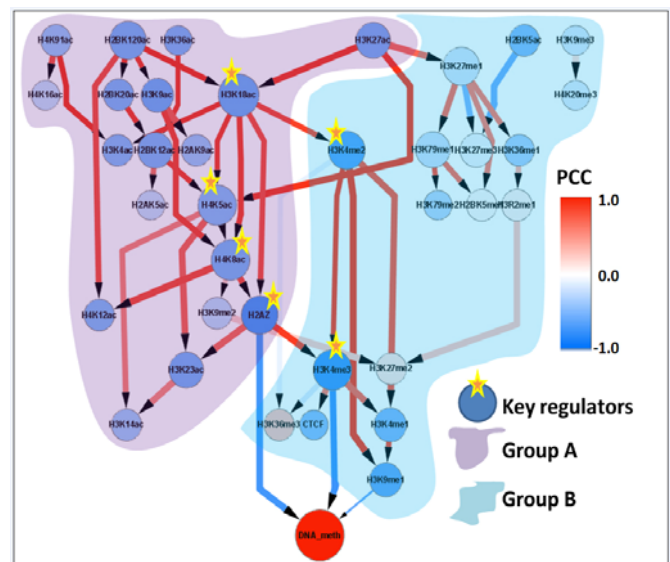


Figure 1. Regulatory relationships among histone modifications and DNA methylation. The regulatory network can be grouped into two parts; one is enriched with histone acetylations (group A), while another mainly consists of histone methylations (group B). The modifications annotated by star are the key-regulators. The color of edges represents the Pearson correlation coefficient (PCC) between two nodes, and the width of edges indicates confidence level of a causal relationship between two nodes. The color of nodes indicates the Pearson correlation coefficient between the modification pattern and DNA methylation, and the size of nodes represents number of edges.

(group A), while another mainly consists of histone methylations (group B). This indicates that the histone acetylations and methylations are densely connected to the same types of modifications, suggesting that acetylations and methylations have different roles for epigenetic control. More precisely, histone acetylations are known to be enriched in CGIs and be related to gene activation, while methylation is related to both activation and repression.

Several relationships identified in our network were consistent with experimental findings of previous studies. For example, in our network one of the main sequential signal propagation pipelines to DNA methylation is 'H3K18ac → H4K5ac → H4K8ac → H2A.Z → DNA methylation'. Here, H2A.Z, a histone variant, is known to be enriched in euchromatin for promoter regions of both active and inactive genes in *S. cerevisiae* (Raisner et al., 2005). H2A.Z is deposited in euchromatin at the flanks of silent heterochromatin to prevent its ectopic spread. This experiment reported that histone tail acetylation is required for efficient recruitment of H2A.Z. The component of the Swr1 complex, Bdf1, is responsible for H2A.Z deposition binding to acetylated histone tails (Ladurner et al., 2003). They revealed a dependence upon histone tail acetylation for robust H2A.Z enrichment and found that the strongest defect was observed in cells harboring H4K5R mutant. Those facts are consistent with our finding that there is a strong connection from a series of acetylations, H3K18ac, H4K5ac, and H4K8ac, to H2A.Z.

Our result also suggests that the group A shown in Figure 1, enriched with several acetylation types densely connected to each other, precisely control H2A.Z, which then antagonistically regulates DNA methylation, which is consistent with the experimental finding that H2A.Z and DNA methylation are mutually antagonistic (Zilberman et al., 2008). There are several experimental evidences for the functional relationships between acetylations (group A) and DNA methylation. The absence of DNA methylation globally elevated levels of acetylations at H3K9, H4K5, and H4K16 (Gilbert et al., 2007). Both Dnmts and methyl-CpG-binding proteins can physically associate with histone deacetylases, so the altered histone acetylation patterns that completely lack CpG methylation may reflect the loss of these interactions (Rountree et al., 2000).

Histone H3K4 trimethylation is a well known factor for gene activation and correlates with transcriptional rate (Liang et al., 2004). The fact that H3K4me3 can be reversed to di- and mono- but not unmethylated products by a JmjC-domain protein (Iwase et al., 2007) is consistent with strong connection among mono-, di-, and trimethylation in our network, suggesting robust control of DNA methylation by H3K4me3. Indeed, our network coincides with the result that CGIs are associated with high levels of histone acetylation and H3K4me3 in human T cells (Roh et al., 2006; Roh et al., 2005).

Many previous studies reported that H2A.Z and H3K4me3 are closely related. Both factors are highly related to repression of DNA methylation. A genome-wide study (Barski et al., 2007) reported that there is a co-occurrence pattern between H2A.Z and H3K4me3. Our current findings, however, suggest a distinguished role of H2A.Z and H3K4me3. In our network, they belong to the different group (H2A.Z to Group A, H3K4me3 to Group B), and connected to largely different groups of factors. It is known that H3K4me3 is closely related with gene activation, while H2A.Z is not. Also, the previous result that their relative abundance varies according to position, with preferring the upstream and downstream promoter regions, respectively (Schmid and Bucher, 2007), suggests the distinguished role between H2A.Z and H3K4me3. All these experimental results are consistent with our network.

Previous studies (Boyer et al., 2006) showed that H3K27 methylations have strong silencing effect on gene expression.

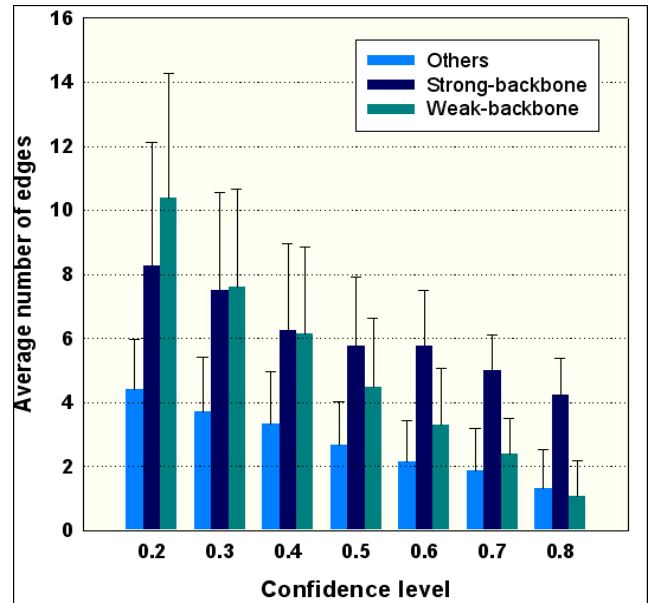


Figure 2. The distribution of the average number of edges as a function of various cutoff values on confidence level. The 'strong-backbone' is defined as the modifications that have more than 4 edges when the confidence level is set to as high as 0.8. The other remaining 'backbone' modifications belong to the 'weak-backbone'. In addition, 'others' is defined as all other modifications. Each group shows different pattern of average number edges according to various cutoff values on confidence level.

However, from our network we could not recognize the significant role of H3K27 methylation, which suggests that H3K27 methylations regulate gene expression but are not directly involved in DNA methylation.

In conclusion, we suggest the existing of two different groups for regulation of DNA methylation. One is histone acetylation enriched regulation with H2A.Z (group A) and another is histone methylation dependent regulation with mono-, di-, and trimethylation of H3K4 (group B). Also in each group H3K18ac and H3K4me2 can be considered as hub regulators, respectively since the two modifications are highly interconnected to others and provide redundant pathways.

'Backbone' modifications in control of DNA methylation

'Backbone' modifications were reported by previous study (Wang et al., 2008) after analyzing histone modification patterns. The 'backbone' modifications refer to a group of 17 modifications which tend to coexist, including H2A.Z, H2BK5ac, H3K4me, H4K8ac, etc. Our results indicate that those multiple histone modifications are associated with critical regulation of DNA methylation. When we gradually increased the confidence level for edges, the majority of remaining relationships are overlapped with those 'backbone' modifications. For this reason, we expected that 'backbone' could have specific properties in our regulatory network. Thereby, we grouped each modification into 'strong-backbone', 'weak-backbone', and 'others'. The 'strong-backbone' is defined as the modifications that have more than 4 edges when the confidence level is set to as high as 0.8. They are H3K4me3, H2A.Z, H4K5ac, and H4K8ac. Those modifications tend to be strongly connected with many other modifications. The other remaining 'backbone' modifications belong to the 'weak-backbone'. In addition, 'others' is defined as all other modifications.

We elucidated the average number of edges for each group. Figure 2 shows the distribution of the average number of edges as a function of various cutoff values on confidence level. From this Figure we notice that each group has distinguished patterns.

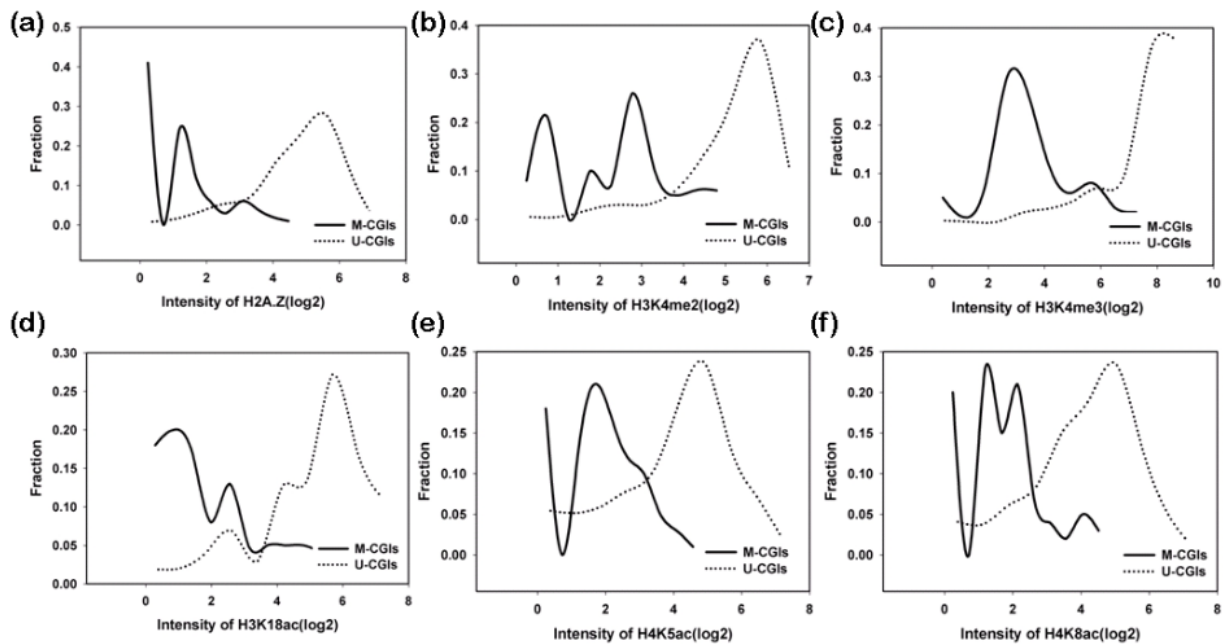


Figure 3. The distribution of modification intensity according to DNA methylation status in key regulators. All key regulators show inverse correlation between U-CGIs and M-CGIs. The x-axis represents intensity of modifications with log scale and y-axis indicates fraction of corresponding intensity. The Pearson correlation coefficient for each figure is -0.67 (a), -0.42 (b), -0.60 (c), -0.67 (d), -0.39 (e), and -0.40 (f), respectively.

Especially, ‘strong-backbone’ modifications are not severely influenced by confidence level, while ‘weak-backbone’ modifications showed a rapid decline. Also, the ‘others’ were not highly or strongly connected to other histone modifications. Based on this finding, we suggest that ‘strong-backbone’ may be the key regulators and should maintain their connections in any circumstance to regulate DNA methylation, whereas ‘weak-backbone’ modifications may operate collectively together with ‘strong-backbone’ modifications to stabilize the signals among histone modifications.

We designate those four ‘strong-backbone’ modifications as ‘key regulators’. In addition, although H3K18ac and H3K4me2 are included in ‘weak-backbone’, both can be considered as hub regulators of each acetylation group (group A) and methylation group (group B), respectively (see Figure 1). Due to this reason, we also include H3K18ac and H3K4me2 as key regulators for DNA methylation. The sequential connections among key regulators with direct regulation of DNA methylation in Figure 1 support that the key regulators are in the main stream of signals to DNA methylation, supporting the experiment evidences described in the above section.

Next, we investigated the distribution of modification intensity according to DNA methylation status in each group. The Pearson correlation coefficients between methylated CGIs (M-CGIs) and unmethylated CGIs (U-CGIs) are shown in Table 1. We notice that ‘others’ show positive correlation, indicating that ‘others’ do not have

Table 1. Pearson correlation coefficient (PCC) and standard deviation (SD) for distribution of histone modifications between U-CGIs and M-CGIs

Group	PCC	SD
Strong-backbone	-0.45	0.10
Weak-backbone	-0.36	0.18
Others	0.54	0.37
Key regulators	-0.53	0.14
Weak-backbone without H3K18ac and H3K4me2	-0.30	0.13

distinguished patterns between U-CGIs and M-CGIs. Conversely, in both ‘strong-backbone’ (-0.45) and ‘weak-backbone’ (-0.36) we found anti-correlations. Furthermore in key regulators the distribution of modifications is highly anti-correlated (-0.53). The distribution of intensity of key regulators in U-CGIs and M-CGIs are shown in Figure 3. This indicates that the status of DNA methylation patterns are discriminated by the order of key regulators > ‘strong-backbone’ > ‘weak-backbone’ and no discrimination with ‘others’.

Prediction of DNA methylation level at CGIs

We divided histone modifications into several groups according to their regulatory patterns. We found that some groups showed well distinguished patterns differentiating M-CGIs and U-CGIs. Therefore, we expected that those histone marks might be suitable to predict the DNA methylation status. To validate our assumption we predicted DNA methylation levels based on several groups by using support vector machine (SVM). We benchmarked a previously developed method which used the sequence information in addition to four histone marks such as H3K4me1, H3K4me2, H3K4me3, and H3K9me1. The results are summarized in Table 2 along with specificity (SP), sensitivity (SE), accuracy (ACC), correlation coefficient (CC), ROC_{50} and ROC scores. The best performance was found when features of key-regulators or ‘strong-backbone’ modifications were used for SVM training. In both cases, the accuracy was 0.94. The performance of ‘others’ and ‘weak-backbone’ were less than that of ‘strong-backbone’, key regulators, and a benchmarked result. We also investigated the performance by dividing key regulators into group A and group B. When we measured the performance of SVM with key regulators in group A and those in group B, we observed that both performances were degraded from performance with key regulators. This result indicates that group A and group B have a distinguished mechanism and function to control DNA methylation, implying that both mechanisms synergistically regulate the precise DNA methylation.

We also plot the distribution of SVM output scores for M-CGIs, U-CGIs, and heterogeneously methylated CGIs (H-CGIs). If the SVMs are well trained based on meaningful regulatory patterns, SVM output scores for all range of DNA methylation should be correlated

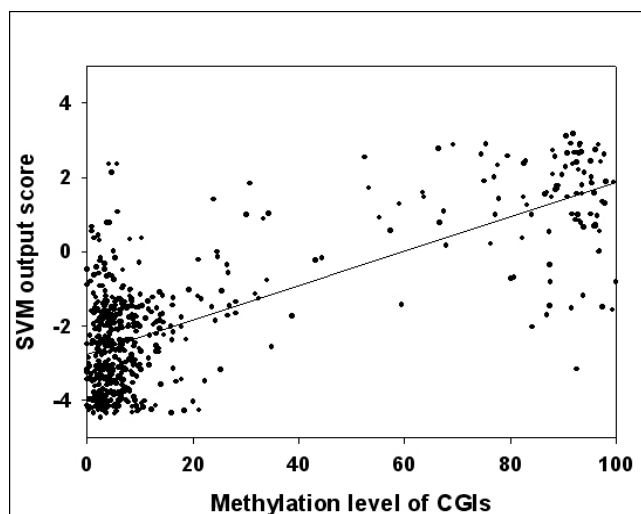


Figure 4. Distribution of SVM output scores according to their methylation level. There exists correlation between SVM output scores and DNA methylation level with key regulators. This remarks that key regulators well represent DNA methylation level even in H-CGIs. In this figure we trained SVM outputs for M-CGI as one.

with methylation level. Figure 4 shows the correlation between SVM output scores and DNA methylation level with key regulators. We can note that key regulators well represent DNA methylation level even in H-CGIs.

To conclude, combinatorial patterns of histone modifications are sufficient enough to describe the status of DNA methylation level, and the high performance supports that those histone modifications are important for regulating DNA methylation level.

Conclusion and Prospects

The present study suggests the functional link between histone modifications and DNA methylation. We constructed Bayesian network to infer causal relationships among epigenetic factors for the controlling of DNA methylation. From the regulatory network we constructed, we found several distinguished patterns of each modification. We suggested two main regulatory systems for DNA methylation; one is for acetylation group including H2A.Z, and another is methylation group. We suggested that H2A.Z and H3K4me3 regulate DNA methylation directly and each regulation is stabilized by several acetylation types such as H3K18ac or methylation types such as H3K4me2, respectively.

We found the core-regulatory network for the controlling of DNA

methylation status in CGIs consist of 6 key regulators (H2A.Z, H3K4me3, H4K5ac, H3K18ac, H4K8ac, and H3K4me2). Based on our findings, we were able to develop highly accurate SVM model predicting DNA methylation level. In this network the role of H2A.Z, H3K4me3, and H4K5ac are already known to be associated with DNA methylation. However, the role of H3K18ac, H4K8ac, and H3K4me2 for DNA methylation is not yet established experimentally. Judging from the fact that there exist distinct patterns between U-CGIs and M-CGIs among them, and we were able to achieve very high prediction accuracy when those features were used in our prediction model, we expect that those factors have an important role in determining DNA methylation status. Therefore it would be interesting if our findings are experimentally validated.

To conclude, we identified regulatory patterns of histone modifications and suggested those patterns are important for the controlling of DNA methylation status. The regulatory patterns that we found may provide important clues for extending our understanding of epigenetic mechanism in DNA methylation.

Materials and Methods

Network construction by using Bayesian network

The Bayesian network has been widely used to solve many biological problems. It has been used to infer the relationship of biological components through simultaneous observation of multiple variables, which is useful for understanding complex biological system (Needham et al., 2006). Due to these characteristics, Bayesian network would be the ideal method to infer the relationship among histone modifications and DNA methylation. In fact, causal relationships for histone codes have been successfully identified by using Bayesian network (Yu et al., 2008). We applied Bayesian network to understand how histone modifications affect DNA methylation at CpG islands (CGIs). Moreover we investigated the regulatory patterns of histone modifications through network analysis.

Here, we used the in-house Bayesian network package (JUNG et al., 2007), which was downloaded by requesting to one of the authors (Sungwon Jung: bluesky@bisl.kaist.ac.kr). To prepare the inputs for Bayesian structure learning, we first divided the values of each histone modifications into three groups. The new data set consists of 0, 1, and 2 for low, medium, and high level of modification, respectively. The threshold for each level was determined to maximize mutual information between groups. The status of DNA methylation was also grouped into unmethylated CGIs (U-CGIs), heterogeneously methylated CGIs (H-CGIs), and methylated CGIs (M-CGIs), annotated by 0, 1, and 2, respectively (the criteria for each group are described in Datasets Section). Next, to increase the reliability of final result of Bayesian network, we generated 100 subsets by randomly selecting 90% of dataset, and

Table 2. Performance variation with various features

Feature	SP	SE	ACC	CC	ROC ₅₀	ROC
Strong-backbone	0.88	0.96	0.94	0.83	0.95	0.97
Weak-backbone	0.74	0.92	0.88	0.63	0.82	0.90
Others	0.72	0.88	0.86	0.54	0.81	0.90
Key regulators	0.87	0.97	0.94	0.84	0.95	0.97
Key regulators in group A	0.81	0.93	0.90	0.70	0.90	0.95
Key regulators in group B	0.81	0.95	0.92	0.75	0.86	0.93
Fan. et al.	0.94	0.74	0.90	No data	No data	No data

then built Bayesian networks 100 times by hierarchical network learning. Lastly, we measure the confidence level by counting the number of overlapped relationships among 100 networks. We constructed regulatory network with relationships which have over 0.6 confidence score (overlapping more than 60 times among 100 subsets).

Datasets

We used two datasets of histone modifications and DNA methylation for human CD4+ T cell. We used ChIP-Seq data for 20 histone methylation patterns and 18 histone acetylation patterns together with two other factors, CTCF and H2A.Z (Barski et al., 2007; Wang et al., 2008). The data provide the sequence counts per 200bp (400bp for H2A.Z) intervals. We mapped those modifications onto the genomic coordinate and counted the number of detections in a certain range of CGIs. We obtained the region of CGIs from UCSC genome browser, and counted the number of modifications for ± 1 kb regions flanking the center of CGI.

We obtained DNA methylation data from HEP which provides genome-wide DNA methylation profiling of human chromosome 6, 20 and 22 of 12 different tissues (Eckhardt et al., 2006). We mapped those CpG methylation values (ranging from 0 to 100) to human genome for CD4+ T cell. Then we calculated the average methylation level of each CGI if more than 10% CpG sites were mapped with certain methylation values. We divided those CGIs into three groups. The CGIs which have more than 50 of average methylation level were considered as methylated CGIs (M-CGIs), while unmethylated CGIs (U-CGIs) were defined with less than 10 of average methylation value. Also, CGIs within 10 to 50 of methylation values were regarded as heterogeneously methylated CGIs (H-CGIs). As a result, we generated 100 M-CGIs, 82 H-CGIs, and 367 U-CGIs from human CD4+ T cells over chromosome 6, 20 and 22.

SVM training

SVM was implemented by using SVM-light (<http://svmlight.joachims.org/>), freely available SVM software, with the radial basis function as a kernel. We used default option for SVM except that the value of γ is set to 0.05. We modified original data by taking log2 transformation for SVM input features. We tried to minimize the number of features for each group to avoid potential overtraining because if we used too many features, there may be a danger of overstraining for given dataset. As a result, when we built the SVM classifier for predicting the status of DNA methylation, we selected the top four modifications that were highly linked to other modifications in each group.

Performance assessment for methylation status in CGIs

Performances of methylation status at CGIs were measured by calculating the receiver operating characteristic (ROC) scores and ROC50 scores (Gribskov and Robinson, 1996). ROC score is defined as the areas under the ROC curves, the plot of true positives as a function of the number of false positives. The ROC50 score is the area under the ROC curve up to the first 50 false positives. Also, we assessed the prediction performance with leave-one-out cross validation (LOOCV). We trained SVM with all dataset except one and tested the SVM with left-out one. The specificity (SP), sensitivity (SE), accuracy (ACC), and correlation coefficient (CC) were calculated to measure the performance. The equations are in the following figure. Here, we defined U-CGIs as a positive class, and TP, TN, FP, and FN indicate true-positives, true-negatives, false-positives, and false-negatives, respectively. Also, we benchmarked another methylation prediction method which used both histone modifications and DNA sequence information (Fan et al., 2008). Since our dataset and evaluation scheme were the same, we directly compared the performance presented in that paper.

$$SP = \frac{TN}{TN + FP}$$

$$SE = \frac{TP}{TP + FN}$$

$$ACC = \frac{TP + TN}{TP + FP + TN + FN}$$

$$CC = \frac{TN \times TN - FP \times FN}{\sqrt{(TP + FN) \times (TP + FP) \times (TN + FP) \times (TN + FN)}}$$

Acknowledgement

I.J. thanks Y.L. for many discussions. This work is supported by CHUNG Moon Soul Center for BioInformation and BioElectronics (CMSC), and by Korea Science and Engineering Foundation (Grant number: 2008-05442).

References

- Adorjan, P., Distler, J., Lipscher, E., Model, F., Muller, J., Pelet, C., Braun, A., Florl, A. R., Gutig, D., Grabs, G., et al. (2002). Tumour class prediction and discovery by microarray-based DNA methylation analysis. *Nucleic Acids Res.* 30, e21.
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837.
- Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T., and Schreiber, S. L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc. Natl. Acad. Sci. USA* 99, 8695-8700.
- Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D. K., Huebert, D. J., McMahon, S., Karlsson, E. K., Kulbokas, E. J., 3rd, Gingeras, T. R., et al. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169-181.
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.
- Bock, C., Paulsen, M., Tierling, S., Mikeska, T., Lengauer, T., and Walter, J. (2006). CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. *PLoS Genet.* 2, e26.
- Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349-353.
- Das, R., Dimitrova, N., Xuan, Z., Rollins, R. A., Haghghi, F., Edwards, J. R., Ju, J., Bestor, T. H., and Zhang, M. Q. (2006). Computational prediction of methylation status in human genomic sequences. *Proc. Natl. Acad. Sci. USA* 103, 10713-10716.
- Down, T. A., Rakyian, V. K., Turner, D. J., Flicek, P., Li, H., Kulesha, E., Graf, S., Johnson, N., Herrero, J., Tomazou, E. M., et al. (2008). A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat. Biotechnol.* 26, 779-785.
- Eckhardt, F., Lewin, J., Cortese, R., Rakyian, V. K., Attwood, J., Burger, M., Burton, J., Cox, T. V., Davies, R., Down, T. A., et al. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* 38, 1378-1385.
- Fan, S., Zhang, M. Q., and Zhang, X. (2008). Histone methylation

- marks play important roles in predicting the methylation status of CpG islands. *Biochem. Biophys Res. Commun.* 374, 559-564.
- Fang, F., Fan, S., Zhang, X., and Zhang, M. Q. (2006). Predicting methylation status of CpG islands in the human brain. *Bioinformatics* 22, 2204-2209.
- Gilbert, N., Thomson, I., Boyle, S., Allan, J., Ramsahoye, B., and Bickmore, W. A. (2007). DNA methylation affects nuclear organization, histone modifications, and linker histone binding but not chromatin compaction. *J. Cell Biol.* 177, 401-411.
- Gitan, R. S., Shi, H., Chen, C. M., Yan, P. S., and Huang, T. H. (2002). Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. *Genome Res.* 12, 158-164.
- Gribskov, M., and Robinson, N. L. (1996). Use of receiver operating characteristic (ROC) analysis to evaluate sequence matching. *Comput. Chem.* 20, 25-33.
- Iwase, S., Lan, F., Bayliss, P., de la Torre-Ubieta, L., Huarte, M., Qi, H. H., Whetstone, J. R., Bonni, A., Roberts, T. M., and Shi, Y. (2007). The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell* 128, 1077-1088.
- JUNG, S., LEE, K. H., and LEE, D. (2007). Enabling Large-Scale Bayesian Network Learning by Preserving Intercluster Directionality. *IEICE Transactions on Information and Systems* E90-D(7), 1018-1027.
- Keshet, I., Schlesinger, Y., Farkash, S., Rand, E., Hecht, M., Segal, E., Pikarski, E., Young, R. A., Niveleau, A., Cedar, H., and Simon, I. (2006). Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat. Genet.* 38, 149-153.
- Ladurner, A. G., Inouye, C., Jain, R., and Tjian, R. (2003). Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Mol. Cell* 11, 365-376.
- Liang, G., Lin, J. C., Wei, V., Yoo, C., Cheng, J. C., Nguyen, C. T., Weisenberger, D. J., Egger, G., Takai, D., Gonzales, F. A., and Jones, P. A. (2004). Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc. Natl. Acad. Sci. USA* 101, 7357-7362.
- Lippman, Z., Gendrel, A. V., Black, M., Vaughn, M. W., Dedhia, N., McCombie, W. R., Lavine, K., Mittal, V., May, B., Kasschau, K. D., et al. (2004). Role of transposable elements in heterochromatin and epigenetic control. *Nature* 430, 471-476.
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560.
- Needham, C. J., Bradford, J. R., Bulpitt, A. J., and Westhead, D. R. (2006). Inference in Bayesian networks. *Nat. Biotechnol.* 24, 51-53.
- Raisner, R. M., Hartley, P. D., Meneghini, M. D., Bao, M. Z., Liu, C. L., Schreiber, S. L., Rando, O. J., and Madhani, H. D. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* 123, 233-248.
- Robyr, D., Suka, Y., Xenarios, I., Kurdistani, S. K., Wang, A., Suka, N., and Grunstein, M. (2002). Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. *Cell* 109, 437-446.
- Roh, T. Y., Cuddapah, S., Cui, K., and Zhao, K. (2006). The genomic landscape of histone modifications in human T cells. *Proc. Natl. Acad. Sci. USA* 103, 15782-15787.
- Roh, T. Y., Cuddapah, S., and Zhao, K. (2005). Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev.* 19, 542-552.
- Rountree, M. R., Bachman, K. E., and Baylin, S. B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat. Genet.* 25, 269-277.
- Schmid, C. D., and Bucher, P. (2007). ChIP-Seq data reveal nucleosome architecture of human promoters. *Cell* 131, 831-832; author reply 832-833.
- Selker, E. U., Tountas, N. A., Cross, S. H., Margolin, B. S., Murphy, J. G., Bird, A. P., and Freitag, M. (2003). The methylated component of the *Neurospora crassa* genome. *Nature* 422, 893-897.
- Tamaru, H., and Selker, E. U. (2001). A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* 414, 277-283.
- Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Peng, W., Zhang, M. Q., and Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* 40, 897-903.
- Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L., and Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* 37, 853-862.
- Yu, H., Zhu, S., Zhou, B., Xue, H., and Han, J. D. (2008). Inferring causal relationships among different histone modifications and gene expression. *Genome Res.* 18, 1314-1324.
- Zilberman, D., Coleman-Derr, D., Ballinger, T., and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* 456, 125-129.