Rhythm of Epigenetics
Dancing to the Bead of DNA Methylation

Flemming S[1], Grüning BA[1], Bohleber S[1], Häupl T[2], Günther S[1]

stephan.flemming@pharmazie.uni-freiburg.de

[1] *Pharmaceutical Bioinformatics, Institute of Pharmaceutical Sciences, University of Freiburg*
[2] *Department of Rheumatology and Clinical Immunology, Charité University Hospital, Berlin*

Methylation of cytosins within a CpG dinucleotide is a common epigenetic DNA modification and may arrest cells in a pathogenic state in complex disorders, e.g. cancer [1] or rheumatoid arthritis [2]. CpGs occur mainly in clusters, called CpG islands (CPIs), being present in nearly 70% of the human genes' promotor region [3]. The Illumina HumanMethylation450 Beadchip platform provides a genome-wide coverage of 485,577 CpGs. Analysis of these CpGs reveals a correlation between changes in DNA methylation and gene expression, even though not all sites have the same impact.

The methylation state of one CpG or a whole CPI may influence the expression of the corresponding gene due to binding of Methylation-Binding-Domains (MBD) and other methylation dependent proteins. To identify CpGs influencing gene expression and common methylation patterns we used several approaches, e.g. network analysis and machine learning techniques. The results help to understand differential methylation analyses with the Illumina BeadChip platform and other techniques.

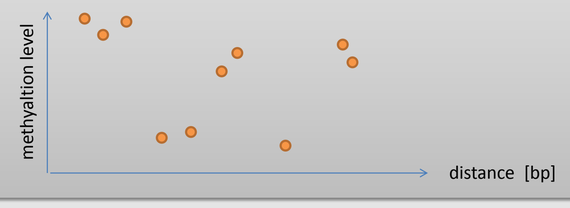
1 Methylation Analysis

Analysis of methylation states of 485,477 CpG sites – approximately 23,000 genes – reveals that not all CpG dinucleotides within a gene have an equal level of methylation. Correlation with expression data shows that not all sites might have the same impact on gene expression regulation.

"What are the reasons for differences in methylation within a gene promotor?"

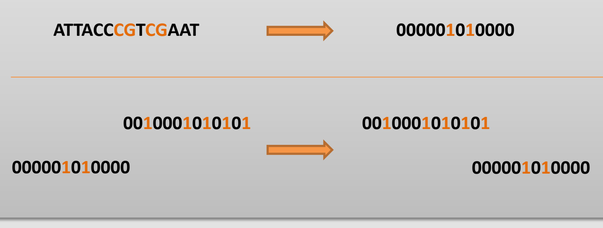
2 Mapping

Analysis of the CpG sites provided by the Illumina platform leads to differences in methylation levels of CpG sites closely located.



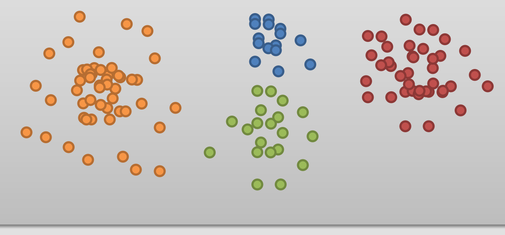
3 Similarity Search

A bitwise comparison approach was used to identify methylation patterns in human gene promotor regions.



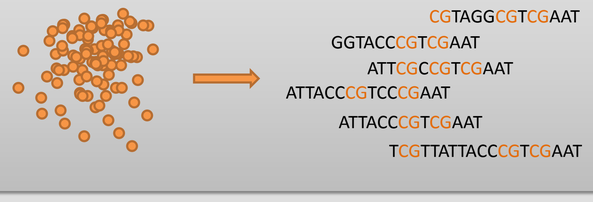
4 Network Analysis

Application of clustering algorithms leads to the identification of sets of gene promotors with similar CpG patterns.



5 Pattern Identification

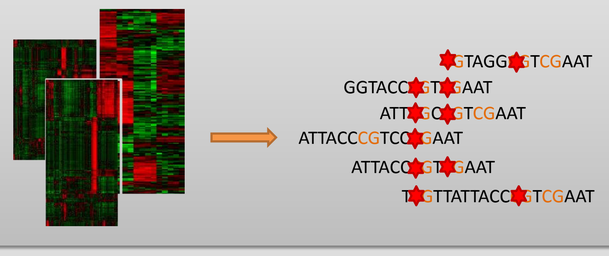
Calculation of multiple alignments for each cluster to obtain CpG consensus sequences.



"Is there a pattern in the CpG order?"

6 Tissue Analysis

Each consensus sequence will be analysed in different tissues to identify common methylated CpG dinucleotides.

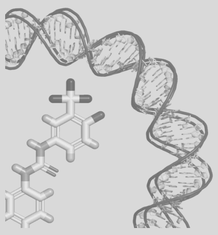


"Is it possible to compare methylation data from different techniques?"

We analysed the promotor sequences of approximately 23,000 genes of the human genome and mapped all CpG sites provided by the Illumina HumanMethylation450 platform. Unfortunately, there is only a small amount of CpG covered by this technique (~19 sites per gene). Therefore, all CpG dinucleotides located in promotor regions were analysed. This led to ~309 CpG per gene. The developed algorithm calculated approximately $5.3*10^8$ pairwise CpG alignments using $1.2*10^{13}$ bit comparisons.

Cluster of gene promotors with similar CpG patterns were obtained and each cluster led to a CpG consensus sequence. Experimental data obtained with the HumanMethylation450 platform will be applied to the computed alignments. Further development will include different methylation techniques (e.g. Illumina HumanMethylation27, NGS approaches). The aim will be to provide a comprehensive overview of the CpG coverage of those platforms and their dependency from location within the promotor region.

Albert-Ludwigs-Universität Freiburg



References

[1] Hansen KD et al. Increased methylation variation in epigenetic domains across cancer types. Nat Genet., 43(8):768, 2011
[2] Karouzakis E et al. DNA methylation regulates the expression of CXCL12 in rheumatoid arthritis synovial fibroblasts. Genes and Immunity, 12:643–652, 2011
[3] Nakano K et al. DNA methylome signature in rheumatoid arthritis. Ann Rheum Dis., ahead of print, 2012

Pharmaceutical Bioinformatics (www.pharmaceutical-bioinformatics.org)

UNI FREIBURG