Distinct epigenetic architectures in bidirectional promoters revealed by single cell analysis

Fatemeh Behjati Ardakani[a,b,c], Nina Gasparoni[d], Laura Arrigoni[e], Sarah Kinkley[f], Matthias Barann[g], Sebastian Froehler[h], Peter Ebert[b,c], Andreas S. Richter[e], Gilles Gasparoni[d], Karl Nordstroem[d], Florian Schmidt[a,b,c], Stefan Wallner[i], Jan Hengstler[j], Kathrin Gianmoena[j], Cristina Cadenas[j], Barbara Hutter[k], Gideon Zipprich[l], Baerbel Felder[l], Juergen Eils[l], Benedikt Brors[k], Julia K. Polansky[m], Thomas Lengauer[b], Wei Chen[h], Ho-Ryun Chung[f], Thomas Manke[e], Philip Rosenstiel[g], Joern Walter[d], Marcel H. Schulz[a,b]

[a]*Excellence Cluster for Multimodal Computing and Interaction, Saarland University, Saarbruecken Germany*
[b]*Department of Applied Computational Biology and Algorithmics, Max Planck Institute for Informatics, Saarbruecken, Germany*
[c]*Graduate School of Computer Science, Saarland University, Saarbruecken, Germany*
[d]*Department of Genetics, University of Saarland, Saarbruecken, Germany*
eMax Planck Institute of Immunobiology and Epigenetics, Stuebeweg 51, Freiburg, 79108, Germany
[f]*Otto Warburg Laboratories: Epigenomics at Max Planck Institute for Molecular Genetics, Berlin, Germany*
[g]*Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany*
[h]*Berlin Institute for Medical Systems Biology, Max-Delbrueck Center for Molecular Medicine, Berlin, Germany*
[i]*Institute for Clinical Chemistry and Laboratory Medicine, University Hospital Regensburg, Regensburg, Germany*
[j]*Leibniz Research Centre for Working Environment and Human Factors IfADo - Ardeystrasse 67 - D-44139 Dortmund Germany*
[k]*Applied Bioinformatics, Deutsches Krebsforschungszentrum, Heidelberg, Germany*
[l]*Data Management and Genomics IT, Deutsches Krebsforschungszentrum, Heidelberg, Germany*
[m]*Experimental Rheumatology, German Rheumatism Research Centre, Berlin, Germany*

Abstract

Bidirectional promoters (BPs) are prevalent in eukaryotic genomes. It is poorly understood how the cell integrates different epigenomic information, such as transcription factor (TF) binding and chromatin marks, to determine directionality of gene expression.

For example, bimodal distributions of activating histone marks (HMs) are found at BPs, but the question remains unresolved if HMs spread along a BP as part of its regulation.

We utilize single cell RNA-seq data and a novel homogeneity score to discover that BP regulation is more complex than previously described. The two genes at a BP may show concordant (homogeneous) or discordant (heterogeneous) expression distributions.

Using epigenomic datasets we observe distinct patterns of TF binding and HMs in both groups. New computational models show that these patterns reflect positional preferences of binding TFs that regulate the observed differences in gene expression distributions.

Further, we find that the distance between the two transcription start sites (TSS) impacts the correlation of nascent RNA expression, the likelihood of heterogeneous single cell expression, and involvement of upstream enhancer marks in gene regulation.

Despite the bimodal distribution of HMs, we observe that the majority of histone marks associated with gene expression occurs downstream of the gene's TSS, except for upstream enhancer marks that are regulated by tissue-specific TFs. Thus, our results unravel an additional layer of complexity in the analysis of BP regulation. This suggests that future studies investigating the associations of regulatory elements in BPs should consider cell heterogeneity as a confounding factor.