



Library Indian Institute of Science Education and Research Mohali



DSpace@IISERMohali / Thesis & Dissertation / Master of Science / MS-16

Please use this identifier to cite or link to this item: <http://hdl.handle.net/123456789/3870>

Title:	Developing a framework to study the effects of enhancer-like promoters on gene regulation
Authors:	Omprakash, Arpit.
Keywords:	Framework Gene Regulation
Issue Date:	28-Jul-2021
Publisher:	IISERM
Abstract:	<p>GWAS (Genome Wide Association Studies) have been crucial to identifying genetic loci associated with diseased phenotype. The hypothesis-free nature of GWAS studies have been a success at predicting specific cancer markers. However, this hypothesis-free nature has also led to one of its main issues, i.e., the large number of distant SNPs discovered with no biological link to the known genetic pathways of the diseased phenotype. Recent advancements in chromatin interaction mapping techniques have identified long-ranged promoter-promoter interactions that regulate gene expression pathways in eukaryotes. The presence of regulatory enhancer-like activity in some promoters and differences in the epigenetic features associated with the promoters and enhancer-like promoters (ELPs) have also been described. It led us to hypothesize that studying such long-range promoter-promoter contacts using ELPs may provide insights into biological links between distant SNPs and genetic pathways of disease in a population. Here we explore possible histone markers and transcription factor bindings (epigenetic factors) that can distinguish between promoters and ELPs. We also build machine learning models that can predict the magnitude of enhancer-like activity (enhancer potential) of a promoter given its epigenetic factors. Regression models to predict the enhancer potential values were made, but the models' accuracy was not up to the mark. Improvements have been suggested for the models, including better feature extraction methods using machine learning classifiers. In the case of HeLa cancer cells, biologically significant epigenetic factors are identified via the classifiers that distinguish between promoters and ELPs. However, the models did not exhibit sufficient accuracy to get relevant features in K562 cancer cells. In the later part of the thesis, spatial interactions between distant promoters have been characterised using Hi-C data. A mathematical framework incorporating the enhancer potentials and spatial interactions between promoters has been proposed to study the propagation of gene regulation in promoter-promoter networks. Initial results from the framework indicate that it can be used to identify distant upstream interacting promoters of a given promoter of interest and model time-course gene expression data to identify novel pathways of gene regulation. Keywords – GWAS, Hi-C, enhancer-like promoters (ELPs), gene regulation, machine learning</p>
URI:	http://hdl.handle.net/123456789/3870
Appears in Collections:	MS-16

Files in This Item:

File	Description	Size	Format	
MS16124 MS Thesis - print.pdf		1.78 MB	Adobe PDF	View/Open

Show full item record



Items in DSpace are protected by copyright, with all rights reserved, unless otherwise indicated.