CIS REGULATORY MODULE DISCOVERY IN TH1 CELL DEVELOPMENT

Satishkumar Ranganathan Ganakammal

Submitted to the faculty of the School of Informatics
in partial fulfillment of the requirements
for the degree of
Master of Science in Bioinformatics,
Indiana University

December 2010

Accepted by the Faculty of Indiana University,
in partial fulfillment of the requirements for the degree of Master of Science
in Bioinformatics

Master's Thesis Committee

Narayanan B. Perumal, Ph.D., Chair

Mark H. Kaplan, Ph.D.

Golnaz Vahedi, Ph.D.

Dedicated to my parents

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ACKNOWLEDGEMENTS

I would like to take the opportunity to acknowledge some of the people who made my graduate study a memorable experience and made this thesis possible. Foremost, it is my sincere pleasure to express my deep and sincere gratitude to my advisor, Dr. Narayanan B. Perumal, for his guidance, motivation, feedback, encouragement, support, and patience during the course of my thesis. His input and efforts have been of great value for me.

I would like to thank the other members of my thesis committee, Dr. Mark H. Kaplan and Dr. Golnaz Vahedi for their time, encouragement, insightful comments, critical feedback and hard questions. I must appreciate their efforts to review my work. I also render my sincere thanks to member of Dr.Kaplan's lab for helping me with their wet lab cross validation and Dr. John O'Shea and his laboratory at NIH for providing me with his ChIP-Seq data.

I owe my sincere thanks to Indiana University for providing the financial support throughout my Master's program. Without the adequate academic preparation, my studies could not have been a successful experience. Hence, I would like to add my thanks to Dr. Mathew J. Palakal, Dr. Meeta Pradhan, Dr. Karl F. MacDorman, Dr.Malika Mahoui and other faculty and staff of the Department of Bioinformatics for their support in the course work.

I owe my loving thanks to my parents and sister for their encouragement and understanding. I would also like to thank Seth R. Good, Rahul Krishna Kollipara and James Scherschel for help in various stages of my work. My loving thanks to Rini Pauly for her help in my thesis writing and presentation. I would also like to thank my friends, Abhinita, Deepali, Anusha and the whole Indian student community at IUPUI for their support and all the fun we have had in the last three years.

ABSTRACT

Immune response enables the body to resist foreign invasions. The Inflammatory response is an important aspect in the immune response which is articulated by elements such as cytokines, APC, T-cell and B-cell, effector cell or natural killer. Of these elements, T-cells especially T-helper cells; a sub class of T-cells plays a pivotal role in stimulating the immune response by participating in various biological reactions such as, the transcription regulatory network. Transcriptional regulatory mechanisms are mediated by a set of transcription factors (TFs), that bind to a specific region (motifs or transcription factor binding sites, TFBS), on the target gene(s) controlling the expression of genes that are involved in T-helper cell mediated immune response. Eukaryotic regulatory motifs, referred to as *cis* regulatory modules (CRMs) or cistrome, co-occur with the regulated gene's transcription start site (TSS) thus, providing all the essential components for building the transcriptional regulatory networks that depends on the relevant TF-TFBS interactions. Here, we study IL-12 stimulated transcriptional regulators in STAT4 mediated T helper 1 (Th1) cell development by focusing on the identification of TFBS and CRMs using a set of Stat4 ChIP-on-chip target genes. A region containing 2000 bases of *Mus musculus* sequences with the Stat4 binding site, derived from the ChIP-on-chip data, has been characterized for enrichment of other motifs and, thus CRMs. Our experiments identify some potential motifs, (such as NF-κB and PPARγ/RXR) being enriched in the Stat4 binding sequences compared to neighboring background sequences. Furthermore, these predicted CRMs were observed to be associated with biologically relevant target genes in the ChIP-on-chip data set by meaningful gene ontology annotations. These analyses will enable us to comprehend the complicated transcription regulatory network and at the same time categorically analyze the IL-12 stimulated Stat4 mediated Th1 cell differentiation.