



COMPUTATIONAL DETECTION AND ANALYSIS OF TRANSCRIPTIONAL CONTROL ELEMENTS IN LYMPHOCYTE DEVELOPMENT

[Login \(/login\)](#)

- [IUPUI ScholarWorks Repository](#)
- →
- [School of Informatics and Computing](#)
- →
- [Informatics Theses and Dissertations](#)
- →
- [Informatics Graduate Theses and PhD Dissertations](#)
- →
- [View Item](#)

COMPUTATIONAL DETECTION AND ANALYSIS OF TRANSCRIPTIONAL CONTROL ELEMENTS IN LYMPHOCYTE DEVELOPMENT

[Singh, Vandana](#)



Name: Vandana_MasterThe ...
Size: 408.4Kb
Format: PDF

[View/Open](#)

Permanent Link: <http://hdl.handle.net/1805/1999>
Date: 2009-11-05

Abstract:

ABSTRACT BACKGROUND: Lymphocyte development and differentiation in mammals follow complex gene regulatory mechanisms with control at the transcriptional stage playing a major role. B and T cells, the two large subsets of lymphocytes, develop differentially due to the varying expression patterns of a variety of genes. Computational tools and methods are becoming increasingly useful in the elucidation of various mechanisms in this process, which has traditionally been studied by experimentation. Wet laboratory experimentation invariably consists of studying one gene at a time although recent advances in microarray and chromatin immunoprecipitation (ChIP) technologies have made available large data sets for informatics analysis. Another impetus for computational approaches has been the explosion of annotated mammalian genomic data in various databases. Traditionally, DNA sequences upstream of the expressed genes (cis-acting) and transcription factor molecules binding to these DNA sequences (transacting) have been explored. We have been employing a computational regimen to identify transcriptional control elements in the DNA (promoters) of genes that may differentiate the development of B and T cells. Towards this goal, our scheme involves the collection and analysis of four different data sets specific for genes involved in B and T cell development with the focus being on the sequences upstream to the transcription start site (TSS) of the relevant genes. **RESULTS:** Using datasets of B and T cell specific genes (Immunoglobulin and T cell receptor genes respectively) from RefSeq, we have identified two predominant consensus

patterns in their upstream regions using the Gibbs Recursive Sampler software. With the help of transcription factor binding site (TFBS) prediction software, different TFBS were obtained for B and T cell genes on the same datasets. A few of them are biologically important, for example, in the case of B cell specific genes we obtained Oct-1, a known immune-specific TFBS. We employed MEME and Gibbs Recursive Sampler software on two different data sets of B and T cell specific regulatory sequences and found different motifs, which are carried by genes common in both software predictions and further used the EZ-Retrieve tool on different motifs to find TFBS. We predicted several immunologically relevant TFBS, such as E47, Oct-1 and GATA-1, at different locations and on both strands in these motifs. In addition, k-means clustering was performed on the datasets in order to classify the B and T cell genes based on the frequencies of TFBS in their upstream sequences. Applying several computational methods, we are able to find additional information on B and T cell genes in terms of TFBS, which may help in the understanding of B and T cell development. CONCLUSIONS: Performing computational approaches like MEME, Gibbs Recursive Sampler, statistical analysis and k-means clustering on different DNA (promoter) sequences does not always identify biologically meaningful transcriptional control elements involved in lymphocyte development. On the other hand, our predictions of conserved motifs in upstream regulatory regions of target genes, and in particular, the identification of immune-specific TFBS in these motifs are biologically relevant. We hope that they will provide a guide for the experimental biologist to focus on certain elements for biological validation. In summary, this informatics approach to detect transcriptional control elements may efficiently and effectively aid the biologist to study transcriptional regulation that distinguishes B and T cell development.

This item appears in the following Collection(s)

- [Informatics Graduate Theses and PhD Dissertations \(/handle/1805/303\)](/handle/1805/303)
- [Informatics School Theses and Dissertations \(/handle/1805/954\)](/handle/1805/954)



[Show Statistical Information \(#\)](#)

My Account

- [Login](#)
- [Register](#)

Statistics

- [Most Popular Items](#)
- [Statistics by Country](#)
- [Most Popular Authors](#)

[About Us \(/page/about\)](/page/about) | [Contact Us \(/contact\)](/contact) | [Send Feedback \(/feedback\)](/feedback)

[_\(/htmlmap\)](#)

FULLFILLING *the* PROMISE

[Privacy Notice \(http://ulib.iupui.edu/privacy_notice\)](http://ulib.iupui.edu/privacy_notice)



Copyright (<http://www.iu.edu/copyright/index.shtml>) ©2015

The Trustees of Indiana University (<http://www.iu.edu/>),

Copyright Complaints (<http://www.iu.edu/copyright/complaints.shtml>)