Identification and Characterization of Yeast Synergistic Regulatory Interaction from High Throughput Data

CAI Chunhui

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Principal Supervisor: Prof. TANG Lei Han
Hong Kong Baptist University
July 2010
Abstract

Synergistic regulation of gene expression is a commonly used strategy in eukaryotic cells. We have developed a computational method, by integrating gene expression data and Chip-on-chip data, to generate a genome-scale list of synergistic transcription factor pairs and make associations with particular growth conditions. A significant part of our identified TF pairs are shown to have literature evidence or predicted to be functionally related. These TF pairs can be further classified by the nature of their interactions, i.e. cooperative and competitive. The TF-TF synergy network is shown to be modular, and disassortative. A common theme of implementing regulatory logic is to have a global regulator working with specific regulators. Furthermore, we have identified the synergistic TF pairs of three main cellular systems, and study their regulatory roles in cell’s functional context i.e. metabolism, cell cycle and stress response.
Contents

Declaration i

Abstract ii

Acknowledgements iii

Contents iv

List of Figures viii

List of Tables x

1 Introduction and Literature Review 1

1.1 Introduction to Systems Biology 1

1.2 Yeast - A Model System in Biological Research 3

1.3 Overview of Yeast Transcriptional Regulation 5

1.3.1 Cell’s regulatory systems – the signal transduction system and the gene regulatory system 5

1.3.2 Gene transcriptional regulatory system – the most fundamental regulatory processes 6

1.4 Overview of Transcriptional Regulatory Control in Yeast’s Three Main Cellular Systems 7

1.4.1 Transcriptional regulation in yeast metabolic network 8

1.4.2 Transcriptional regulation in yeast cell cycle progression 9

1.4.2.1 Transcriptional regulation in yeast stress response system 11

1.5 Experimental and Computational Studies of Transcriptional Regulation 13

1.5.1 Pioneer work: Lac operon 13

1.5.2 Experimental developments – high throughput technology 14

1.5.2.1 DNA microarray technology 14

1.5.2.2 ChIP-on-chip 15
1.5.3 Computational developments – reverse engineering of transcriptional regulatory network 17
   1.5.3.1 Gene clustering 17
   Cluster analysis 18
   Bi-cluster analysis 18
   1.5.3.2 Computational identification of TF binding sites 19
   1.5.3.3 Integrative data analysis 20
1.5.4 Operation of yeast TR network on genome-scale 20
1.6 Synergistic Regulation – A Rule Rather Than An Exception 22
   1.6.1 Literature evidences 22
   1.6.2 Previous works 24
   1.6.3 Overall statistics 26
1.7 Research Purpose 28
1.8 Organization of Thesis 29

2 Data Description, Collection and Preprocessing 31
   2.1 MIPS gene functional categories 31
   2.2 Microarray gene expression data 34
   2.3 ChIP-on-chip data 37

3 Identification of Yeast Synergistic TF Pairs 40
   3.1 Computational Method 40
   3.2 Statistics Characteristics of the Results 42
      3.2.1 Powerlaw behavior of conditional TR network 42
      3.2.2 Distribution of P values 43
      3.2.3 Additive effect 44
      3.2.4 Selection criterion of statistically most significant TF pairs 47
      3.2.5 Comparison with random networks 48
   3.3 Most Prominent TF Pairs and Comparison with Literature Knowledge 49
      3.3.1 Experimental evidences 49
      3.3.2 Pair overlaps with other computational studies 51
   3.4 A Critique of the Method: Comparison With Mutual Information analysis 52
      3.4.1 Model description of mutual information analysis 53
         3.4.1.1 2-way mutual information analysis 53
         3.4.1.2 3-way mutual information analysis 56
      3.4.2 Comparison of TF pairs identified from our model with those from MI2 and MI3 57
   3.5 Summary 58

4 TF-TF Synergy Network: Classification and Topology 59
   4.1 Classification of TF-TF Synergy Types 59
      4.1.1 TF pairs are formed mostly by condition variant TFs 59
      4.1.2 Synergy takes different forms – cooperation and competition 60
         4.1.2.1 Cooperative pairs 61