P15

Identifying cooperative transcription factors in yeast using multiple data sources

Fu-Jou Lai 1, Mei-Huei Jhu 2, Chia-Chun Chiu 2, Yueh-Min Huang 1 and Wei-Sheng Wu 2

1 Department of Engineering Science, National Cheng Kung University, Taiwan

2 Department of Electrical Engineering, National Cheng Kung University, Taiwan

Abstract

Background

Transcriptional regulation of gene expression is usually accomplished by multiple interactive transcription factors (TFs). Therefore, it is crucial to understand the precise cooperative interactions among TFs. Various kinds of experimental data including ChIP-chip, TF binding site (TFBS), gene expression, TF knockout and protein-protein interaction data have been used to identify cooperative TF pairs in existing methods. The nucleosome occupancy data is not yet used for this research topic despite that several researches have revealed the association between nucleosomes and TFBSs.

Results

In this study, we developed a novel method to infer the cooperativity between two TFs by integrating the TF-gene documented regulation, TFBS and nucleosome occupancy data. TF-gene documented regulation and TFBS data were used to determine the target genes of a TF, and the genome-wide nucleosome occupancy data was used to assess the nucleosome occupancy on TFBSs. Our method identifies cooperative TF pairs based on two biologically plausible assumptions. If two TFs cooperate, then (i) they should have a significantly higher number of common target genes than random expectation and (ii) their binding sites (in the promoters of their common target genes) should tend to be co-depleted of nucleosomes in order to make these binding sites simultaneously accessible to TF binding. Each TF pair is given a cooperativity score by our method. The higher the score is, the more likely a TF pair has cooperativity. Finally, a list of 27 cooperative TF pairs has been predicted by our method. Among these 27 TF pairs, 19 pairs are also predicted by existing methods. The other 8 pairs are novel cooperative TF pairs predicted by our method. The biological relevance of these 8 novel cooperative TF pairs is justified by the existence of protein-protein interactions and co-annotation in the same MIPS functional categories. Moreover, we adopted three performance indices to compare our predictions with 11 existing methods' predictions. We show that our method performs better than these 11 existing methods in identifying cooperative TF pairs in yeast. Finally, the cooperative TF network constructed from the 27 predicted cooperative TF pairs shows that our method has the power to find cooperative TF pairs of different biological processes.

Conclusion

Our method is effective in identifying cooperative TF pairs in yeast. Many of our predictions are validated by the literature, and our method outperforms 11 existing methods. We believe that our study will help biologists to understand the mechanisms of transcriptional regulation in eukaryotic cells.