Comprehensive Analysis of Delay in Transcriptional Regulation Using Expression Profiles

Koji Ota ota@kuicr.kyoto-u.ac.jp
Takuji Yamada takuji@kuicr.kyoto-u.ac.jp
Yoshihiro Yamanishi yoshi@kuicr.kyoto-u.ac.jp
Susumu Goto goto@kuicr.kyoto-u.ac.jp
Minoru Kanehisa kanehisa@scl.kyoto-u.ac.jp
Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

Keywords: regulatory delay, transcriptional regulation, time series expression profiles

1 Introduction

Regulation of gene expression is mainly conducted by transcription factors (TFs) binding to upstream regions of their target genes (Targets). Many algorithms have been developed to extract such regulatory relationships between TFs and Targets from time series expression profiles. Here we focus our attention on the time delay of TF expression and its Targets expression, which may not be uniform because of the fluctuation of regulatory interactions. We analyzed time series expression profiles in yeast cell cycles and observed such regulatory delays within known pairs of TFs and Targets, providing the evidence that the distribution of such delays is not uniform.

2 Materials and Methods

2.1 Materials

We used the dataset of direct regulatory relationships between TFs and Targets from chromation immunoprecipitation (ChIP) by Lee et al [3]. This includes about 4000 relationships between 106 TFs and 2363 Targets. We also used the dataset of time series genomic expression profiles in yeast cell cycle synchronized by alpha factor by Spellman et al [4]. This includes 18 time points, with 7 minute intervals, corresponding to two cell cycles.

2.2 Methods

First, we applied cubic smoothing spline [2] to the expression profiles to remove the noise, where the degree of freedom is set to 10. The original time series of each gene expression was converted into a smooth time series.

Then we observed TFs-Targets relationships from the ChIP dataset, exploring corresponding expression profiles. For each of the 4000 relationships, we calculated the correlation coefficients between expression profiles of a TF and its Targets with -1 to 7 time point delays, because one cell cycle is achieved by 8-9 time points in this dataset and we assume that the time delay between TF and Targets expressions would not exceed one cell cycle.

Finally, for each relationship, we chose the time delay with the highest absolute value of the correlation coefficient, for we assumed that fluctuations of TFs would resemble those of Targets with a time delay. We focused on the clear relationship whether it reflects positive correlation or negative correlation, with the threshold value of $>= 0.7$ or $<= -0.7$.

3 Results and Discussion

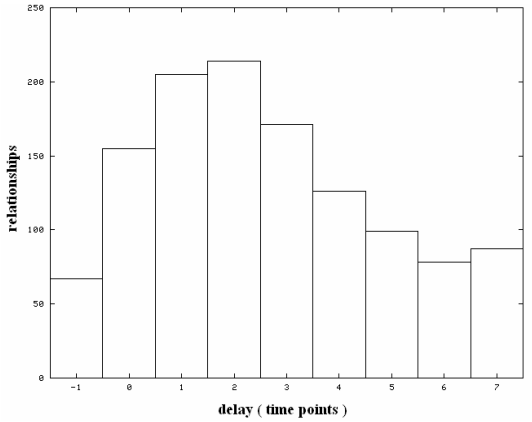

Figure 1: Distribution of regulatory delays.

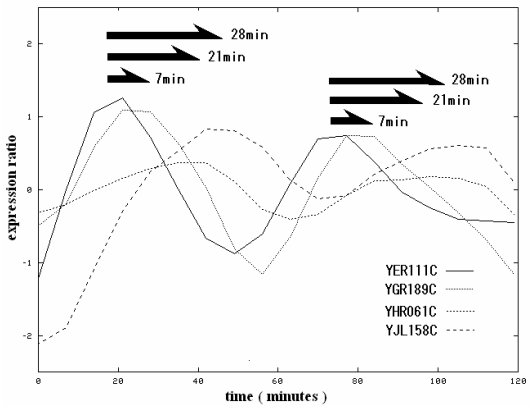

Figure 2: An example of relationships with delay. YER111C activates YGR189C, YHR061C, and YJL158C. Each delay is 7, 14, and 28 minutes.

The total of 2158 relationships, including 1202 positive and 956 negative ones, are extracted from about 4000 relationships in the ChIP dataset, exploring the expression profiles. Here, we discuss positive relationships because negative ones seem to contain many false relationships. In the distribution of 1202 positive relationships, relationships with 2 points (14 minutes) delay are the most frequent (Fig. 1). However, many relationships with long delays also exist. Although these relationships may include many false ones owing to the preliminary stage of developing our method, it is certain that the time delay of a TF and its Targets is not uniform (for example, Fig. 2). We also classified the relationships into 7 biological process groups according to Gene Ontology classification by Saccharomyces Genome Database [1], and observed the average and the standard deviation of the delays in each group (Table 1).

Table 1: Biological process and average delay.

Biological process	Average delay; time points	(std. dev.)
DNA replication	2.32	(2.45)
Signal transduction	2.37	(2.33)
Transcription	2.39	(2.22)
Growth & maintenance	2.39	(2.25)
Transport	2.41	(2.39)
Cell cycle	2.53	(2.39)
Metabolism	2.73	(2.13)
All	2.62	(2.23)

These results suggest that transcriptional regulatory relationships contain diverse regulatory delays, and we should take such delays into account to develop algorithms for inferring transcriptional regulatory networks from time series expression profiles. Now, we are developing further improvements of our method in order to correlate regulatory delays with different molecular and cellular events.

Acknowledgments

We would like to thank Masumi Itoh and Michihiro Tanaka for their helpful discussion.

References

[1] Dweit, S.S., Haris, M.A., Dolinski, K., Ball, C.A., Binkley, G., Christie, K.R., Fisk, D.G., Issel-Tarver, L., Schroeder, M., Serlock, G., Sthuraman, A., Weng, S, Botstein, D., and Cherry, J.M., Saccharomyces Genome Database (SGD) provides secondary gene annotation using the Gene Ontology (GO), *Nucleic Acids Res.*, 30:69–72, 2002.

[2] Green, P.J. and Silverman, B.W., *Nonparametric regression and generalized linear modes: a roughness penalty approach*, Chapman and Hall, 1994.

[3] Lee, T.I., Rinaldi, N.J., Rovert, F., Odom, D.T., Bar-Joseph Z., Gerver, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., Zeitlinger, J., Jennings, E.G., Mrray, H.L., Gordon, D.B., Ren, B, Wyrick, J.J., Tagne, J.B., Volkert, T.L., Fraenkel E., Gifford, D.K., and Young, R.A., Transcriptional regulatory networks in *Saccharomyces cerevisiae*, *Science*, 298:799–804, 2002.

[4] Spellman, P.T., Sheriock, G., Zhang, M.Q., Iyer V., Anders, K., Eisen, M.B, Brown, P.O., Bostein, D., and Futcher, B., Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization, *Molecular Biology of the Cell*, 9:3273–97, 1998.