BIOINFORMATICS APPLICATIONS NOTE Vol. 23 no. 12 2007, pages 1559–1561
doi:10.1093/bioinformatics/btm126

Gene expression

SCEPTRANS: an online tool for analyzing periodic transcription in yeast

Andrzej Kudlicki, Maga Rowicka and Zbyszek Otwinowski*

University of Texas Southwestern Medical Center, Department of Biochemistry, 5323 Harry Hines Blvd. Dallas, TX 75390-8816, USA

Received on March 2, 2007; revised on March 22, 2007; accepted on March 24, 2007

Advance Access publication March 30, 2007

Associate Editor: Chris Stoeckert

ABSTRACT

Summary: SCEPTRANS is designed for analysis of microarray timecourse data related to periodic phenomena in the budding yeast. The server allows for easy viewing of temporal profiles of multiple genes in a number of datasets. Additional functionality includes searching for coexpressed genes, periodicity and correlation analysis, integrating functional annotation and localization data as well as advanced operations on sets of genes.

Availability: Available online at http://sceptrans.org/

Contact: andrzej@work.swmed.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

Periodic processes in the living organisms (e.g. cell cycle), involve precisely orchestrated transcription of hundreds of genes. Studying whole-genome time-course expression data provides insight into regulation and function of individual genes in context of cellular processes. A number of microarray datasets related to periodic phenomena in yeast are available (Table 1). The functionality of the existing servers (Pramila *et al*., 2006; Spellman *et al*., 1998, SGD) does not allow for precise comparison of expression profiles within a user-defined group of genes or for comprehensive comparison of expression patterns of a group of genes across conditions, both essential for in-depth analysis. The methods of assessing periodicity also vary substantially, from relying on visual examination (Cho *et al*., 1998) or assuming all genes in such a system are periodic (Klevecz *et al*., 2004) to various complex approaches (Pramila *et al*., 2006; Tu *et al*., 2005), making comparisons between datasets difficult for general users not willing to completely re-analyze each dataset. Here, we present an integrative resource applying the same set of tools to data from different experiments.

2 IMPLEMENTATION

The interface is build on an apache/CGI server and consists of two functional modules: gene selection and analysis.

*To whom correspondence should be addressed.

2.1 Gene selection

Gene selection is done by subsequent modification of the current set of genes, by applying the AND, OR, NOT and REPLACE set operators. Genes may be thus added or removed from the current set by one of the following criteria:

- Explicit list of gene names or identifiers (e.g. YNL300w, cln3)
- Regular expression search of gene name or description (e.g. to select autophagy-related genes, one can either search names for 'ATG[0–9]' or search descriptions for 'autophag' and then remove 'non-autophag')
- Correlation with a given gene in one of the experiments
- Peak-to-trough ratio in each experiment
- Periodicity score in one of the experiments
- Cellular localization (e.g. mitochondrion or spindle pole)
- Classified as periodic by different studies
- Functional annotation (general, e.g. metabolism, or detailed, e.g. nitrogen and sulfur metabolism)

2.2 Gene analysis

The gene analysis module allows the user to inspect expression profiles of chosen genes in selected datasets. The following analysis tools are available:

- **Table of gene features.** The table includes the common and systematic names of a gene, its cellular localization (Huh *et al*., 2003), functional annotation (Comprehensive Yeast Genome Database), description (Saccharomyces Genome Database), peak-to-trough transcription ratios, cell cycle regulated flags (Cho *et al*., 1998; Pramila *et al*., 2006; Rowicka *et al*., 2007; Spellman *et al*., 1998) and metabolic regulated flag (Tu *et al*., 2005).
- **Plots of temporal expression profiles of selected genes.** Profiles are shown in the linear scale, normalized to unit average or unit maximum. With a limited number of plots displayed at a time, the user can highlight each expression profile for clarity.
- **Periodicity analysis.** Two measures of periodicity are implemented: the magnitude of the Fourier mode and

Table 1. Source datasets

Dataset	Synchronization method	Reference	Points/cycles
YMC	Yeast metabolic cycle	Tu *et al.* (2005)	36/3
MC-40	Metabolic cycle (40 min)	Klevecz *et al.* (2004)	32/3
cdc28	Cell cycle (*cdc28* mutants)	Cho *et al.* (1998)	17/2.2*
cdc15	Cell cycle (*cdc15* mutants)	Spellman *et al.* (1998)	24/2.4*
alpha	Cell cycle (alpha pheromone)	Spellman *et al.* (1998)	18/2.2*
eluct	Cell cycle (elutration)	Spellman *et al.* (1998)	14/3.4**
alpha-26	Cell cycle (alpha pheromone)	Pramila *et al.* (2006)	13/2
alpha-30	Cell cycle (alpha pheromone)	Pramila *et al.* (2006)	25/2
alpha-38	Cell cycle (alpha pheromone)	Pramila *et al.* (2006)	25/2

*our estimation, **doubtful periodicity.

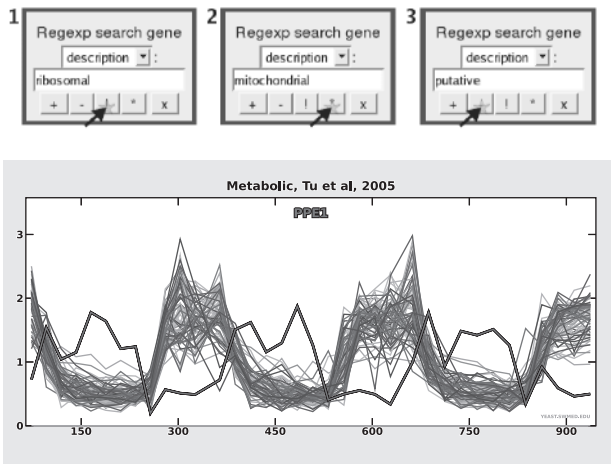

Fig. 1. Example of SCEPTRANS functionality. The list of genes currently annotated as coding for components of mitochondrial ribosomes is easily obtained (top). Note the unique expression pattern of PPE1 (bottom). Colour version of this figure is available as Supplementary material online.

autocorrelation at different time shifts. The computed Fourier mode (Lomb, 1976; Scargle, 1982), is presented as a periodogram (i.e. using period length as the abscissa). Autocorrelations are computed on normalized profiles interpolated with cubic splines (Press *et al.*, 1992). Periodograms or autocorellograms can be displayed graphically as an alternative to time series plots. Statistical significance of periodograms (Horne and Baliunas, 1986) can also be displayed.

- **Correlation table.** Correlation tables are available to distinguish co-regulation modules in the group of genes. To facilitate interpretation, correlation tables are displayed both as heat maps and numerical values of Pearson correlation coefficients. Genes may be sorted according to time of expression in each dataset, thus visually separating different expression patterns in the correlation table.

2.3 Limitations

Our implementation of the interface relies on the use of JavaScript, which has to be enabled in the user's browser.

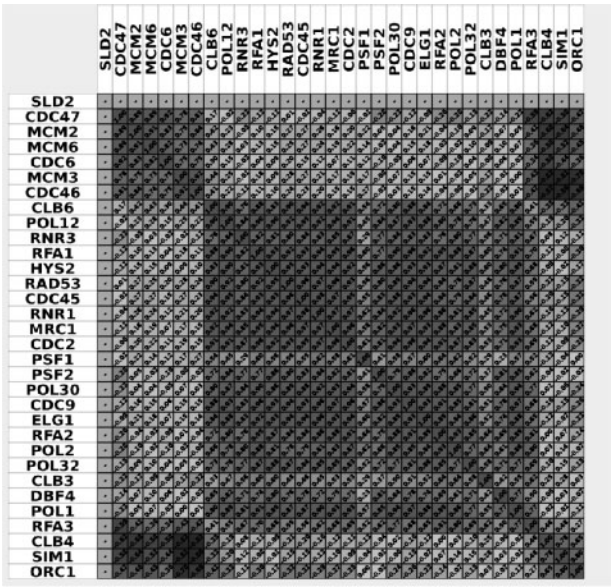

Fig. 2. Correlation table of genes with 'DNA replication' function which are cell-cycle periodic (Pramila *et al.*, 2006; Rowicka *et al.*, 2007) displayed for 'alpha-30' data. Unlike time-course plots, the table shows clearly that expression of MCM subunits (top left corner) precedes that of other genes.

To save server-side CPU time and bandwidth, correlation tables will not be displayed for more than 175 transcripts; no more than 100 temporal profiles will be shown at a time (and maximum 50 can be highlighted), and the gene description table will be hidden if it contains more than 1000 lines. These maximum numbers of transcripts may be further reduced if more than one dataset is displayed at a time, or if the current load on the server is too high.

3 DISCUSSION

To exemplify SCEPTRANS functionality, we show how it speeds up the functional analysis we have reported previously (Tu *et al.*, 2005). For instance, lists of genes periodic during yeast metabolic cycle (YMC) with selected cellular localizations can be readily obtained using the 'localized' and 'regulated' boxes in the gene selection panel. This procedure shows

immediately that the most significantly overrepresented localization among the genes found periodic during YMC is mitochondrion (430 out of 523 are periodic in YMC). Similarly, one can obtain that disproportionally large number of genes from the 'energy' and 'metabolism' functional categories are periodic during YMC (80 and 67% respectively, compared to 52% in the whole genome). These results support the view that the observed oscillations are of metabolic nature (Tu *et al*., 2005).

Another task much facilitated by our server is the analysis of expression of mitochondrial ribosomal genes (Fig. 1). Genes with this annotation can be selected by just three clicks of the mouse. In the YMC dataset, the expression pattern of one such transcript, *PPE1*, is significantly different than that of other mitoribosomal transcripts, raising a possibility that *PPE1* is either misannotated, or that it is regulated in a different manner than other mitoribosomal transcripts. Indeed, Ppe1 is the only gene annotated as mitochondrial ribosomal subunit that lacks the regulatory motif characteristic of this group (Tu *et al*., 2005).

Another very useful tool is a correlation table: it allows us to see data structure more easily than time-course plots (Fig. 2).

In summary, SCEPTRANS provides an opportunity for general researchers to perform a comprehensive yet flexible analysis of various microarray time-course data. We expect that SCEPTRANS will facilitate substantially such analysis and will help to formulate and test novel research hypotheses.

ACKNOWLEDGEMENT

This work was supported by NIH GM U54GM074942. We thank Maciej Puzio for technical assistance.

Conflict of Interest: none declared.

REFERENCES

Cho,R. *et al*. (1998) A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell*, **2**, 65–73.

Horne,J.H. and Baliunas,S.L. (1986). *Astrophysical J*., **302**, 757.

Huh,W. *et al*. (2003) Global analysis of protein localization in budding yeast. *Nature*, **425**, 686.

Klevecz,R. *et al*. (2004) A genomewide oscillation in transcription gates DNA replication and cell cycle. *PNAS*, **101**, 1200–1205.

Lomb,N. (1976) *Astrophy. Space Sci*., **39**, 447.

Pramila,T. *et al*. (2006) The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the Sphase gap in the transcriptional circuitry of the cell cycle. *Genes Dev*., **20**, 2266.

Press,W.H. *et al*. (1992) *Numerical Recipes in C. The Art of Scientific Computing*. Cambridge University Press.

Rowicka,M. *et al*. (2007) High-resolution timing of the cell-cycle regulated gene expression. Submitted.

Scargle,J.D. (1982) *Astrophys. J*., **263**, 835.

Spellman,P. *et al*. (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *S. cerevisiae* by microarray hybridization. *Mol. Biol. Cell*, **9**, 3273.

Tu,B.P. *et al*. (2005) Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes. *Science*, **310**, 1152.