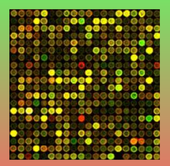
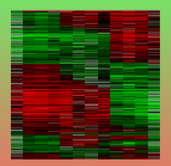


Extracting digital signals from microarray time-course data

Debashis Sahoo, David L. Dill, Rob Tibshirani, Sylvia K. Plevritis
Stanford University, Stanford, CA



Abstract

This is an algorithm for mining microarray time course data that extracts digital signals consisting of sequences of instantaneous transitions between discrete levels. The algorithm uses adaptive regression to select the best data fit from a collection of curves with varying numbers and times of transitions. It produces a list of the genes that change at a particular time step. We applied the algorithm to published microarray data for budding yeast and identified genes that are changing at a particular time step; the results have biologically relevant GO annotations and are consistent with published results. Application to human cancer data is open to interpretation and further validation.

Introduction

Goal:

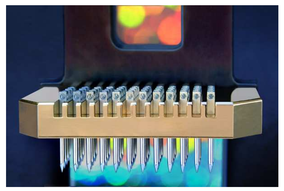
- Understand ordering of events in transcriptional network from microarray time-course data.

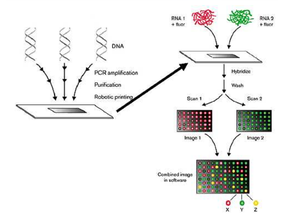
Challenges:

- The data is voluminous.
- Time courses are relatively short.
- The data is noisy, and there are many other sources of errors.

Approach:

- Find genes that turn on and off at particular times.
- Method is appropriate for time courses of 7-30 points of response to a stimulus.





Method

- Test for "single step" and "binary two step"
- Adaptive regression using step functions
- Step functions are shifted to find the best fit (Figure 1)
- Estimates the degrees of freedom for the F-statistic
- Get a p-value from the F-statistic
- Group genes by the type and time of changes in the expression
- Use GO-TermFinder[2] to find enriched Gene Ontology annotations.

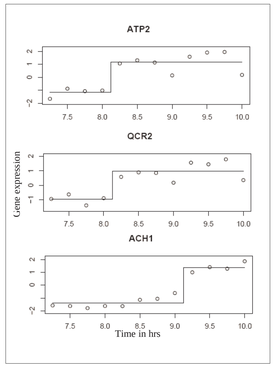


Figure 1: Fitting Step function to diauxic shift time course

Results

- Evaluated on a publicly available time course of microarrays from Brauer *et al.* [1].
- Heat map is shown in Figure 2. GO annotations are shown in Table 1.
- Compared against the clustering method used by Brauer *et al.* [1].
- The method is also applied on time course of response of human prostate tumor cells to treatment with MSA from Zhao *et al.* [3](Figure 3).

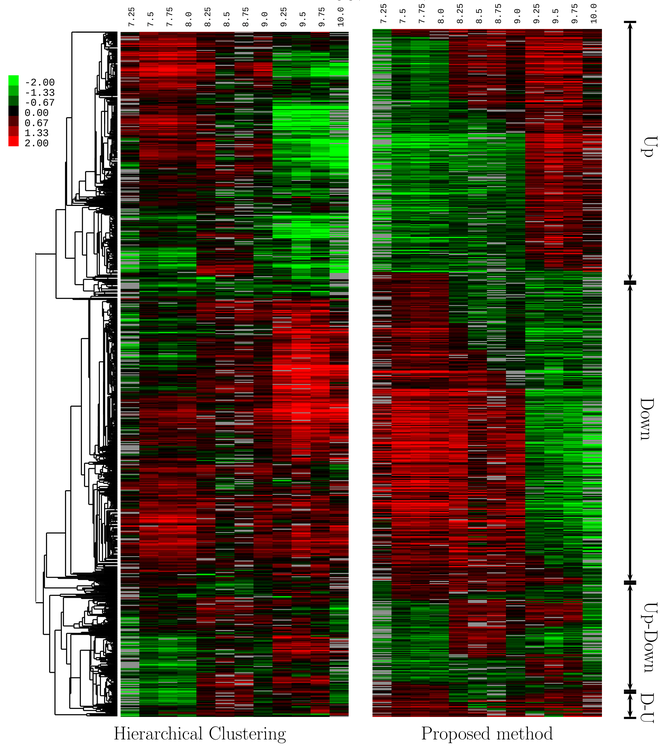
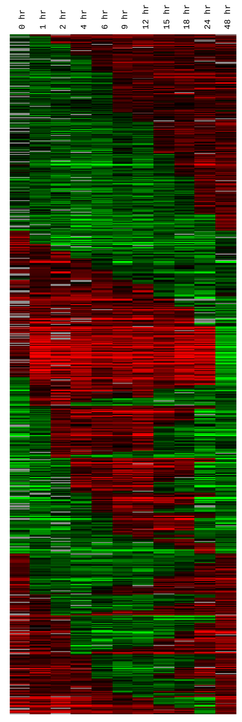


Figure 2: Analysis on diauxic shift time-course microarray data [1] on glucose limited budding yeast.

GO Annotations	p-value	Group	p-value[1]
protein biosynthesis	3.4e-51	All Down at 9.25 hrs	9.7e-33
ribosome biogenesis and assembly	1.2e-39	All Down	1.4e-33
generation of precursor metabolites and energy	7.4e-24	All Up	6.1e-14
oxidative phosphorylation	4.9e-14	All Up	6e-08
amino acid and derivative metabolism	1.7e-11	All Up-Down	6.2e-25
amine biosynthesis	1.7e-12	Up-Down - Down at 9 hrs	1.1e-24
hexose catabolism	0.00046	Up-Down - Up at 8.25 hrs	0.044
monosaccharide catabolism	0.0012	Up-Down - Up at 8.25 hrs	0.091
Siderophore transport	-	-	0.013
Intracellular Transport	-	-	1.6e-08
Secretory pathways	-	-	1.5e-06

TABLE 1: GO annotations and p-values according to GO-TermFinder. "p-value[1]" is the p-value using the list of genes from the clusters reported by Brauer *et al.*.



GO Annotations	p-value	Group
cell cycle	2.6e-29	All Down
metabolism	4.2e-16	All Down-Up
protein transport	1.5e-13	All Up
cell proliferation	4.8e-08	All Down
cell death	1.2e-05	All Up-Down

GO annotations and p-values according to GO-TermFinder.

Figure 3: Analysis on the human prostate cancer data from Zhao*et al.* [3].

Conclusion

The algorithm has been very useful in listing genes fully automatically with GO annotations relevant to the experiment in yeast microarray time-course data from Brauer *et al.*. The method has potential to explore microarray time-course data on cancer cells.

References

[1] Matthew J. Brauer, Alok J. Saldanha, Kara Dolinski, and David Botstein. Homeostatic Adjustment and Metabolic Remodeling in Glucose-limited Yeast Cultures. *Mol. Biol. Cell*, 16(5):2503–2517, 2005.

[2] Gavin Sherlock. GO-TermFinder. *The Comprehensive Perl Archive Network*, 0.7ed, 2003.

[3] Hongjuan Zhao, Michael L. Whitfield, Tong Xu, David Botstein, and James D. Brooks. Diverse Effects of Methylseleninic Acid on the Transcriptional Program of Human Prostate Cancer Cells. *Mol. Biol. Cell*, 15(2):506–519, 2004.