Poster M-17

Detecting Cell Cycle Regulated Genes of Schizosaccharomyces pombe by using Non-metric Multidimensional Scaling without Sinusoidal Fitting

Authors:
Y-h. Taguchi, (*Department of Physics, Chuo University,*)

Short Abstract: We applied non-metric multidimensional scaling to cell division cycle microarray of Schizosaccharomyces pombe. Genes turn out to be aligned as circular arrangement whose polar angles are associated with cell cycle phases. One of two gene groups selected by consistency among experiments is associated with known cell cycle regulated genes.

Long Abstract:

Recently several groups have studied Schizosaccharomyces pombe's cell division cycle processes using microarray [1-3]. Although the authors estimated cell cycle regulated genes by sinusoidal fittings, this method is sometimes criticized because sinusoidal behavior does not always mean cell cycle regulations [4]. In actual, although three groups selected 750, 407, and 747 cell cycle regulated genes respectively, only 171 genes are commonly selected. Thus, if possible, it is better to select cell cycle regulated genes without sinusoidal fittings. In order to do this, we need some methods in which assumptions are as small as possible. For such methods, we have developed new algorithm for non-metric multidimensional scaling (nMDS) [5,6] and applied it to cell division cycle microarray experiments [7,8]. In this paper, we will report that application of nMDS to S. pombe cell division cycle can select cell cycle regulated genes without sinusoidal fittings.

nMDS is an ordination method where the relationship between objects is represented as geometrical configuration in some metric space (typically, Euclidean space). In this configuration, objects are placed in the way such that rank correlation coefficient between distances (in the embedded space) and predefined dissimilarity is maximized. When applying nMDS to microarray experiments, the correlation coefficients between gene expression profiles are used as predefined similarity between genes [6-8]. Then we typically get circular arrangement of genes in the embedding space.

Thus we have conjectured that polar angle corresponds to cell cycle phases. First of all, distribution of genes along polar angle have two peaks. It is well known behavior for cell division cycle experiments found by sinusoidal fittings. This does not disagree with the above conjecture. In order to see if this conjecture is correct, next we have selected cell cycle regulated genes as follows. If genes are truly cell cycle regulated and polar angle represents cell cycle phases, polar angle between cell cycle regulated genes must be conserved over experiments by different groups. Then we have selected two elutriation experiments from every research (Refs. [1] or [2]) and choose genes groups within which polar angle between pairs of genes is conserved. This results in two independent genes groups each of which consist of 500 genes.

After we compare these mean gene expression profiles in each group with the left one in Supplementary Figure 1b [2], we can notice that these two correspond to DNA synthesis and cell division process (G2 phase) respectively. Also, polar angles agree with cell cycle phases estimated by sinusoidal fittings very well.

For cross-validation, we have embedded other experiments [3] into 2D space again, and confirmed that the above selected genes groups represent DNA synthesis and cell division process for other experiments [3]. Furthermore, polar angles defined by nMDS for the above experiments [1,2] agree well with cell cycle phases decided by sinusoidal fittings for other experiments [3].

Finally, we have selected genes which have conserved polar angle difference from known 38 cell cycle regulated genes over nMDS embeddings for experiments [1,2] (We have named these genes as operationally cell cycle regulated genes; oCCR genes). oCCR genes overlap only the genes group associated with DNA synthesis and have no overlap with the other gene groups. This again confirms that our selection of genes base upon nMDS embeddings is highly biologically informative. In conclusion, nMDS turns out to be useful tool to figure out the hidden relationship between genes. Since nMDS does not assume any sinusoidal behaviors, nMDS is much more suitable tool to select cell cycle regulated genes than sinusoidal fittings. Furthermore, there may turn out to be some problems in the usage of sinusoidal fittings. In our selection of genes, gene expression profiles are grouped together with sign-reversed ones. Since both expressive/depressive genes can play important role, it is very natural to group both of them. However, in sinusoidal fittings, only expressive genes are assigned for corresponding cell cycle phases. Moreover, in contrast to gene selection by sinusoidal fittings, selection by nMDS is based upon consistency. Thus, genes selected are automatically common cell cycle regulated genes for all experiments. nMDS is possibly much better tool than sinusoidal fittings.

One may be doubtful of nMDS because of its requirement for massive computational resources. However, we have also found that principal components analysis (PCA) using normalized gene expression profiles can give us similar polar angle of genes. This means, once we have confirmed that polar angle by nMDS is biologically informative, we can make use of PCA instead of nMDS. Since PCA is much less challenging method for computer resources, massive cpu resources for nMDS is not so important problem from the application point of views.

References

[1] A. Oliva et al, PLoS Biology, vol. 3 (2005) e225.
[2] G. Rustici et al, Nature Genetics, vol. 36 (2004) 809-817.
[3] X. Peng et al, Mol. Biol. Cell, vol. 16 (2005) 1026-1042.
[4] S. Cooper and K. Shedden, Cell and Chrom., vol. 2 (2003) 1.
[5] Y-h. Taguchi and Y. Oono, Advances in Chemical Physics, vol. 130B, (2005) 315-351.
[6] Y-h. Taguchi and Y. Oono, Bioinformatics, vol. 21 (2005), 730-740.
[7] Y-h. Taguchi, IPSJ SIG Technical Report, 2005-BIO-3, (2005) 59-66.
[8] Y-h. Taguchi and Satwik Rajaram, IPSJ SIG Technical Report, 2006-BIO-4 (2006) 9-16.