The Construction of a Database on the Intracellular Vesicular Transport

Akiyasu C. Yoshizawa acyshzw@kuicr.kyoto-u.ac.jp
Shuichi Kawashima shuichi@kuicr.kyoto-u.ac.jp
Minoru Kanehisa kanehisa@kuicr.kyoto-u.ac.jp

Institute for Chemical Research, Kyoto University Gokasho, Uji, Kyoto 611-0011, Japan

Keywords: intracellular vesicular transport, membrane traffic, SNARE, database

1 Introduction

One of the most conspicuous morphological differences of eucaryotic cells from procaryotic cells is the presence of several membrane compartments within the cell. In order to maintain these organelles, eucaryotic cells transfer membrane structural elements, such as proteins and lipids, between the organelles. In this transfer system, small-size membrane vesicles containing structural elements bud from the 'donor' membrane of the donor organelle, move to the vicinity of the target organelle, and fuse to the 'target' membrane to transfer the structural elements. This three-stage system consisting of budding, moving, and fusion is named 'vesicular transport' a.k.a. 'membrane traffic.'

In order for the transport vesicle to fuse only to the precise target membrane, unique mutual recognitions are required between the membranes. Following the 'SNARE hypothesis' [1], this mechanism can be described as follows. When a proper vesicle SNARE (v-SNARE) protein and a target SNARE (t-SNARE) protein exist on the transport vesicle and the target membrane, respectively, at an appropriate region in the cell, their unique binding interactions make it possible to recognize mutually and to trigger the membrane fusion. SNARE (Soluble N-ethylmaleimide-sensitive factor Attachment protein REceptor) is the general term for a protein whose function is unique recognition and membrane fusion, so a single cell contains multiple SNARE pairs representing the variety of paths between organelles.

In order to understand genomic contents and evolutionary origins of intracellular traffic systems, we have constructed a SNARE database which contains a basic data set of t-SNAREs and v-SNAREs. We have then collected other candidates of SNARE proteins from the protein sequence databases and performed the cluster analysis.

2 Method

1. We collected known SNAREs in *Saccaromyces cerevisiae* from experimental papers and the yeast database at MIPS [2], because *S. cerevisiae* is best studied with genetic analysis. The amino acid sequences of these proteins were extracted from the KEGG/GENES database in the GenomeNet.

2. These sequences were used as queries against the nr-aa (non-redundant amino acid sequence) database in the GenomeNet with BLAST2, and similar sequences to each query were extracted (E value < 0.001). By combining all the BLAST2 searches and removing duplicate hits, a data set of SNARE homologs was obtained.

3. The sequence similarity for each pair of sequences in the data set was examined with SSEARCH3.1 program, which is an implementation of the Smith-Waterman algorithm. Then, the hierarchical cluster analysis was performed using the SSEARCH similarity scores.

4. The clustering results were examined manually with the experimental papers and the annotations in the database entries.

3 Results and Discussion

The numbers of SNAREs obtained in *S. cerevisiae* were: 11 for t-SNAREs and 10 for v-SNAREs. With each of these 21 SNAREs as a query sequence, the nr-aa database in the GenomeNet were searched, and the total of 632 SNARE homologs (including experimentally known SNAREs) were found.

When the complete-linkage cluster analysis was performed for the 632 proteins, a clear separation was observed between the cluster of t-SNAREs and the cluster of v-SNAREs. It was most interesting to observed a correlation with the transport pathway. When the position of two organelles are close on the transport pathway, such as TGN (Trans Golgi Network) and lysosomes, their SNAREs are more closely related especially in the case of t-SNAREs.

SNAREs nearby the plasma membrane formed large clusters, but it may be on the ground that the synaptic secretion is studied well and many sequences are reported. On the contrary, t-SNAREs of ER (Endoplasmic Reticulum) formed no cluster because there were only distant homologs. These observations may lead to an evolutional origin of the vesicular transport system which exists only in the eucaryote. We will further make analyses on the relations between the position on the transport pathway and the sequence similarity and/or the domain structure.

Acknowledgements

This work was supported by grants from the Ministry of Education, Science, Sports and Culture, the Science and Technology Agency, and the Japan Society for the Promotion of Science. The computational resource was provided by the Supercomputer Laboratory, Institute for Chemical Research, Kyoto University.

References

[1] Soellner, T., Whiteheart, S.W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J.E., SNAP receptors implicated in vesicle targeting and fusion, *Nature*, 362:318-324, 1993.

[2] Mewes, H.W., Frishman, D., Gruber, C., Geier, B., Haase, D., Kaps, A., Lemcke, K., Mannhaupt, G., Pfeiffer, F., Schoeller, C., Stocker, S., and Weil, B., MIPS: a database for genomes and protein sequences, *Nucleic Acids Res.*, 28:37–40, 2000.