243

Extraction of Related Genes in the Vesicular Transport by Simultaneous Comparison of Two-Hybrid and Microarray Data

Akiyasu C. Yoshizawa   **Akihiro Nakaya**
`acyshzw@kuicr.kyoto-u.ac.jp`   `nakaya@kuicr.kyoto-u.ac.jp`

Susumu Goto   **Minoru Kanehisa**
`goto@kuicr.kyoto-u.ac.jp`   `kanehisa@kuicr.kyoto-u.ac.jp`

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

Keywords: gene cluster, intracellular vesicular transport, binary relationships

1 Introduction

We have presented a new method to extract a set of correlated genes with respect to multiple biological features that are given as graph structures [3]. By multiple graph comparison, this algorithm extracts sets of genes which are considered biologically related (details in [3]). For example, when we analyze two datasets and if a set of genes makes a cluster in both datasets, the relationships among these genes are considered to be biological relevant at high probability. Using this method, we analyzed protein-protein interaction data by two-hybrid experiments and gene coexpression data by microarrays in *S.cerevisiae*, and searched putative gene-gene relationships which participate in the intracellular vesicular transport.

Table 1: Correlated gene clusters related to the vesicular transport.

YBL050W	SEC17; vesicular-fusion protein [SP:SC17_YEAST]
YDR178W	SDH4; succinate dehydrogenase membrane anchor subunit [EC:1.3.5.1] [SP:SDH4_YEAST]
YMR197C	VTI1; vesicle transport V-snare protein VTI1 [SP:VTI1_YEAST]
YOR036W	PEP12; syntaxin (T-SNARE), vacuolar [SP:PE12_YEAST]
YDR468C	TLG1; tSNARE that affects a late Golgi compartment
YGL044C	RNA15; component of the cleavage and polyadenylation factor CF I involved in pre-mRNA 3'-end processing [SP:RN15_YEAST]
YBL102W	SFT2; SFT2 protein [SP:SFT2_YEAST]
YNL133C	unknown [SP:YNN3_YEAST]
YOR220W	unknown
YIL004C	BET1; protein transport protein [SP:BET1_YEAST]
YIL005W	protein disulfide-isomerase [EC:5.3.4.1] [SP:YIA5_YEAST]
YMR039C	SUB1, TSP1; transcriptional coactivator [SP:SUB1_YEAST]
YLR324W	unknown
YDR453C	YDR453C; probable thiol-specific antioxidant protein 2 [SP:TSA2_YEAST]
YNL044W	YIP3; unknown [SP:YIPC_YEAST]
YGR192C	TDH3, GPD3; glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12] [SP:G3P3_YEAST]
YJL052W	TDH1, GPD1, SSS2; glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12] [SP:G3P1_YEAST]
YDR313C	PIB1; phosphatidylinositol(3)-phosphate binding protein
YGR042W	unknown [SP:YG1T_YEAST]
YGL198W	unknown [SP:YGU8_YEAST]
YNL216W	RAP1, GRF1; DNA-binding protein with repressor and activator activity [SP:RAP1_YEAST]
YJL036W	SNX4; unknown [SP:YJD6_YEAST]
YDR473C	PRP3; essential splicing factor
YKR014C	YPT52; GTP-binding protein of the rab family [SP:YP52_YEAST]
YKL035W	UGP1; UTP–glucose-1-phosphate uridylyltransferase [EC:2.7.7.9] [SP:UDPG_YEAST]
YFL054C	unknown [SP:YFF4_YEAST]
YDR425W	unknown
YPL280W	unknown

2 Methods

We used the following datasets: (1)the two-hybrid protein-protein interaction dataset developed by Ito et al. [2](3,280 genes and 4,549 bait-prey interactions), and (2)the coexpression dataset derived from

the results of microarray analysis by DeRisi et al. [1](time-course measurements of gene expressions). The number of inter-graph links that connect the identical genes in the two-hybrid dataset and the coexpression dataset was 1,547. The results were examined manually with the annotations in the KEGG/GENES database entries.

3 Results and Discussion

Because these two datasets share only one gene-gene relationship, we allow gaps; namely, we include indirect gene-gene relationships through one intermediate gene. We found a total of 249 correlated gene clusters, and within these clusters, we found 12 sets of putative gene-gene relationships with respect to the intracellular vesicular transport (Table 1).

Fig. 1 shows the graph which represents a part of the relationships that we found. In this figure, ovals represent the genes belonging to the clusters. Single/double rectangles are the intermediate genes extracted from the co-expression/two-hybrid data, respectively. Lines with small letters *th* are

Figure 1: The genes related to the vesicular transport.

the relations extracted from the two-hybrid data. Lines with no letters are the relations extracted from the coexpression data. Genes connected by the same kind of lines belong to the same cluster. For example, two proteins YLR305C and YIL177C function as the intermediate genes in multiple clusters like "hubs". YLR305C, which participates in two clusters, is the gene of STT4 which is a PI4-kinase, and YIL002C in the other cluster is the gene of INP51 which is a $PI(4,5)P_2$ 5-phosphatase. These results suggest their roles in the PI turnover during the intracellular vesicular transport. YIL177C is a putative protein and its function is not yet assigned, but it may play some important role in the vesicular transport, because it participates in three clusters.
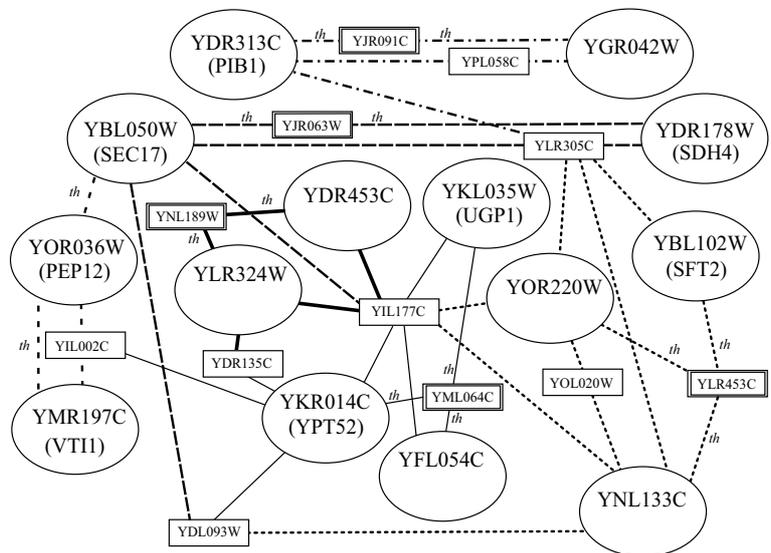
Acknowledgements

References

[1] DeRisi, J.L., Iyer, V.R., and Brown, P.O., Exploring the metabolic and genetic control of gene expression on a genome scale, *Science*, 278(5338):680–686, 1997.

[2] Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y., A comprehensive two-hybrid analysis to explore the yeast protein interactome, *Proc. Natl. Acad. Sci. USA*, 98(8):4569–4574, 2001.

[3] Nakaya, A., Goto, S., and Kanehisa, M., Extraction of correlated gene clusters by multiple graph comparison, *Genome Informatics*, 12, 2001. submitted.