Poster L-28

The role of membrane proteins in sample classification inferred by genome-wide expression profiling data

Authors:
Paolo Uva (*IRBM (MRL-Rome)*)
Armin Lahm (*IRBM (MRL-Rome)*)
Emanuele de Rinaldis (*IRBM (MRL-Rome)*)

Short Abstract: Membrane proteins play a crucial role in cell biology. We define the “Human Membranome” as the set of in-silico predicted membrane protein genes, and we investigate its expression in tumors, cell lines and normal tissue samples. We illustrate here that the largest part of the global transcriptional differences between different samples can be ascribed to the Membranome expression levels.

Long Abstract:

Cell membrane proteins play an essential role in cell biology. In fact they are involved in several functions, including the signal transduction, the mechanisms of cell-cell adhesion and the response to external stimuli. Given their crucial role in many aspects of the cell life, deregulation of expression/activity of this class of proteins has been shown to be associated to several diseases, from cancer to autoimmune diseases. Recently, several efforts have been made to characterize the protein components in different cancer cell membranes in order to identify membrane proteins with a tumor specific expression profile (Adams et al., 2003; Canelle et al., 2006). In this work we focus on the differences, at the transcriptional level, of the expression of membrane protein genes in different human derived biological samples. In particular, we try to assess the “discriminative power” of the membrane protein genes to distinguish between groups of samples obtained from different cell lines, tumor and normal tissue samples (Shadt et al., 2004; Staunton et al., 2001). To identify the set of human plasma membrane proteins we started from the NCBI Gene database and extracted the membrane proteins by means of a combined analysis based on the Gene Ontology annotations and the prediction of transmembrane domains. The list was manually revised to add known membrane associated proteins from the literature and to remove proteins associated with intracellular compartments. As result, we obtained a list of about 4,600 genes, representing about 20% of all the all human genome. This set of genes was defined as the Membranome. We investigated the role of the Membranome in biological human samples discrimination studies. We used different genome-wide microarray expression profiling data sets and assessed the statistical significance of the results by the application of a random permutations procedure. For each dataset, the subset of probes corresponding to the Membranome genes was identified and a Principal Component Analysis (PCA) was carried, using the selected probe values and the samples as variables. In this way each sample was represented by a small set of new variables (principal components) composed of a linear combination of the microarray probe values. PCA is an unsupervised clustering approach; samples are therefore grouped according to the expression values without any a-priori user classification. By using the first three principal components, representing more than 70% of the total variance, we compared the separation between groups of samples

obtained with the Membranome subset genes against random subsets of not-Membranome genes of the same size. The distances between sample classes were computed using the average distance within clusters / average distance between clusters. The analytical procedure was implemented into an automatic workflow running on the R statistical platform. By applying this workflow we assessed the contribution of the Membranome genes to the global transcriptional differences residing into different biological samples. We applied this approach independently to three different microarray data sets, respectively containing data about cell lines, tumor and normal samples. Our results indicate that in all cases, membrane protein genes contribute, more than not-membrane protein genes, to the global transcriptional differences residing into biological samples. In addition, by the analysis of the functional annotations of the genes falling into each principal component, functionally related membrane protein genes, showing correlated expression across samples, could be identified.

References:

Adam, P. J., Boyd, R., Tyson, K. L., Fletcher, G. C., Stamps, A., Hudson, L., Poyser, H. R., Redpath, N., Griffiths, M., Steers, G., Harris, A. L., Patel, S., Berry, J., Loader, J. A., Townsend, R. R., Daviet, L., Legrain, P., Parekh, R., and Terrett, J. A. (2003) Comprehensive proteomic analysis of breast cancer cell membranes reveals unique proteins with potential roles in clinical cancer. J. Biol. Chem. 278, 6482 –6489

Canelle, L, Bousquet, J., Pionneau, C., Hardouin, J., Choquet-Kastylevsky, G., Joubert-Caron, R., and Caron, M. (2006). A proteomic approach to investigate potential biomarkers directed against membrane-associated breast cancer proteins. Electrophoresis 27: 1609-1616

Schadt, EE., Edwards, SW., GuhaThakurta, D., Holder D., Ying L., Svetnik V., Leonardson A., Hart KW., Russell A., Li G., Cavet, G., Castle, G., McDonagh, P., Kan, Z,. Chen, R., Kasarskis, A,. Margarint, M., Caceres, RM., Johnson, JM., Armour, CD., Garrett-Engele, PW., Tsinoremas, NF., and Shoemaker, DD., 2004. A comprehensive transcript index of the human genome generated using microarrays and computational approaches. Genome Biol 2004, 5: R73

Staunton, JE., Slonim, DK., Coller, HA., Tamayo, P., Angelo MJ., Park, J., Sherf, U., Lee, JK., Reinhold, WO., Weinstein, JN., Mesirov, JP., Lander, ES., and Golub, TR., 2001. Chemosensitivity prediction by transcriptional profiling. Proc Natl Acad Sci USA 2001: 10787-10792

Poster L-28

Phospholipase D in Citrus EST database - Structure and function studies

Authors:
Martins NF (*ab3c*)
Mehta A (*ab3c*)

Short Abstract: Phospholipases in plants are involved in cellular responses, including plant growth, development, stress and defense. They are classified according to their site of

hydrolysis on phospholipids PLA1, PLA2, PLC .The in silico analyze of Citrus ESTs database revealed 457 sequences related to phospholipases and provided 13 contigs and 12 singlets.

Long Abstract:
Phospholipases in plants are known to be involved in cellular responses, including plant growth, development, stress and defense. Phospholipases are classified according to their site of hydrolysis on phospholipids PLA1, PLA2, PLC and PLD. Each class is divided into subfamilies based on sequences and biochemical properties. Using data mining techniques in the CitrusESTs database it was identified 12 isoforms of PLD with references with other PLD in plants. The higher conserved regions were identified and a theoretical model was built through homology modeling. The validation used WHAT IF and Procheck servers. It is the first evidence of PLD in Citrus.