

Prediction and 3D visualization of biological networks using cytological disease mapping

Björn Sommer¹, Vladimir Ivanisenko², Patrizio Arrigo³, Ralf Hofestädt¹✉

¹AG Bioinformatics and Medical Informatics, University Bielefeld, Bielefeld, Germany

²Institute of Genetics, Siberian Branch of the Russian Academy of Sciences Novosibirsk, Novosibirsk, Russian Federation

³CNR ISMAC, Genoa, Italy

Motivation and Objectives

The subcellular localization of biochemical networks is one of the new challenges of the post-genomic era. It may be used in future as an additional criteria to judge the probability of potential proteomic interaction partners, to choose the appropriate experimental constellations or to create virtual cell environments. Today we can use database integration tools and text mining for the prediction of biological networks. However, we can use the same tools for the prediction of the localization values of each component of biological networks. Having this information calculated we can realize a 3D (2D) visualization of biochemical networks based on a 3D (2D) virtual cell. Therefore we extended our CELLmicrocosmos (<http://Cm4.CELLmicrocosmos.org>) project.

Methods

The CELLmicrocosmos 4.2 PathwayIntegration (short: CmPI) is a module developed based on the CELLmicrocosmos 1.1 CellExplorer (CmCX). CmCX provides a virtual cell environment. This cell environment, containing different cell components, can be associated with protein-related networks using localization information parsed from DAWIS-M.D. as well as ANDCell. Figure 1 shows the workflow between CmPI and the two aforementioned systems. CmPI connects now directly to ANDCell, which is a data mining tool for pathway reconstruction, and DAWIS-M.D., which is a bio data warehouse. The input is a set of proteins or a protein-related network. Based on the data mining and bio data warehouse approach, CmPI will search for localization information for each network component. Based on this data and an already created virtual cell, CmPI will semi-automatically map this information into the three-dimensional cell environment. The user-friendly mapping process is supported by different visualization techniques.

DAWIS-M.D. which is based on the Bioinformatics Data Warehouse (BioDWH) toolkit, is a platform-independent data warehouse approach developed by the Bio-/Medical Informatics Department of Bielefeld University. It contains a number of metabolic-disease-related databases. For this work, only a subset of those databases was used for the localization of proteins: Brenda, Reactome, The Gene Ontology (GO), and UniProt.

The ANDCell (Associative Network Discovery) system contains a large variety of different databases. It is a commercial product of PBSOFT Ltd., a start-up company from the Institute of Cytology and Genetics (Novosibirsk). It was developed for the automatic extraction of facts and knowledge of interactions between proteins, genes, metabolites, microRNA, cellular components, molecular processes, and their association with diseases from the texts of scientific publications and databases. For this work, only those information were taken into account which were ex-

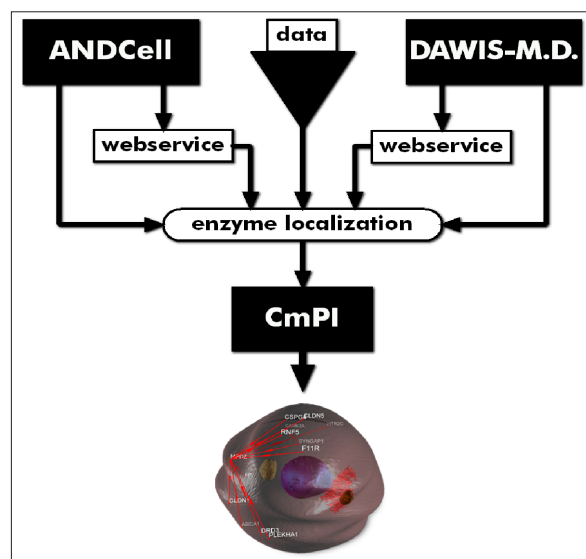


Figure 1: The Workflow between the databases and the CELLmicrocosmos 4.2 PathwayIntegration (CmPI)

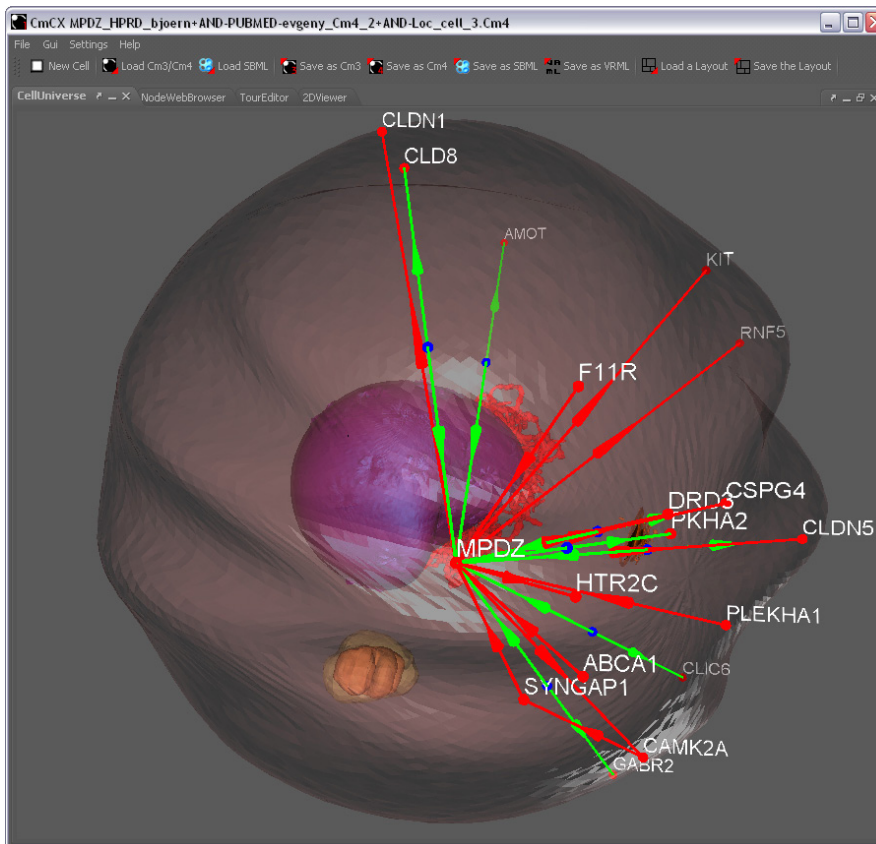


Figure 2: The MPDZ-protein-protein interaction network mapped onto a cell model

tracted from PubMed-abstracts which are freely available for CmPI.

Results and Discussion

These data integration techniques will underscore our understanding of the molecular mechanisms of diseases. In particular, this approach is important to investigate which cellular compartment is mainly involved in a specific pathological process. Based on this integrated knowledge, biological network prediction and reconstruction can be done for many investigated diseases even if the molecular knowledge of such disease is only rudimentary.

Figure 2 shows a MPDZ-related protein-protein interaction network associated with a three-dimensional cell model. This application case was discussed in Sommer et al. 2010. Experimental results suggested that MPDZ is involved in dilated cardiomyopathy. The interacting proteins were found to be associated with the tight junction complex of the cell membrane. Therefore it can

be suggested that this cell localization is affected by dilated cardiomyopathy and should be taken into account in future computational and experimental analyses.

Acknowledgements

This work was supported in part by: BMBF Internationale Zusammenarbeit in Bildung und Forschung mit Russland, RUS 08/005, EU-FP6, CARDIOWORKBENCH project, RAS Program "Living Nature: Current State & the Problems of Development", 14.740.11.0001 (PD, VI, NK); EU-FP7-260429-SYSPATHO.

References

1. B. Sommer, E. S. Tiys, B. Kormeier, K. Hippe, S. J. Janowski, T. V. Ivanisenko, A. O. Bragin, P. Arrigo, P. S. Demenkov, A. V. Kochetov, V. A. Ivanisenko, N. A. Kolchanov, R. Hofestädt. Visualization and Analysis of a Cardiovascular Disease-and MUPPI-related Biological Network combining Text Mining and Data Warehouse Approaches. *Journal of Integrative Bioinformatics*, 7(1):148, 2010.