Visualization Software for Scientific Discovery in Toponome Research

Andreas Dress
MPI Leipzig/Shanghai

Ever more profound and thorough elucidation of cell-state specific changes of protein distributions is currently possible by new biotechnology developments in concert with simultaneous advances in computing, signal processing, and mathematics. Though we understand that protein concentrations indicate functional states and potential dynamic developments of cells, details are not well understood. The ability to take proteomics to the next level through localizing any reasonable number of distinct proteins (up to more than one hundred at present), and then analyzing and displaying this information in a graphical format that can be understood and used by biologists and physicians alike, is vital to edge-cutting research in molecular and cell biology, and for its application in pharmacology and medicine.

While the aim of proteomics is to elucidate the multifaceted, complex, and highly specific relations between the spatial distribution of proteins at any given moment in a given cell and that cell's functional state, most of the existing approaches to proteomics rely on first homogenizing any sample under consideration and then supplying the resulting mixture to chip-based expression-profiling technologies and 2D-gel electrophoresis combined with mass spectrometry. All of these techniques provide highly valuable insights into the overall collection of proteins present in a given piece of tissue at a given moment. Yet, they do not provide insights into the *biological code*, i.e., the rules and mechanisms according to which the actual spatial distribution of proteins within an individual cell in a given piece of tissue determines its specific functional state.

In contrast, fluorescence microscopy has been used to locate individual proteins in cell tissue, and a new variant of this technology, developed by Walter Schubert and dubbed MELK (Multi Epitope Liganden Kartographie), now allows cell biologists to locate up to more than one hundred distinct proteins within the same cell, or piece of tissue, linking molecular events in a cell directly with resulting phenotypical consequences.

More specifically, this technology allows its user to recognize, in subcellular resolution, specific patterns of protein arrangements in situ, thus laying the foundations for topological proteomics, a new and very promising branch of proteomics that deals with recognizing, analyzing, and classifying the modes of protein arrangements within cells and tissue, and relating these modes to a cell's functional state.

This complements existing expression-profiling technologies as it is not based on the destruction of the biological code that is written in terms of the spatial organization of protein networks, including colocalization and "anti-colocalization" events.

I will present samples of the data generated by this technology and various data analysis methods that have been developed to investigate these data and to separate signal from noise.