Finding the needle in the haystack - how statistics can help to find mild co-regulation effects in high-throughput microarray data

Dr. Katharina Zweig (University of Heidelberg)

Microarrays are very helpful in getting an overview on how various biological entities react to different experimental conditions. They often trade-off time and money with the achievable quality of the data, thus, many statistical methods have been introduced to make sense of these massive data sets. One important question is whether it is possible to identify groups of proteins/mRNAs/or other biological entities that are co-regulated in a statistical significant way, even if the regulation itself is only mild. For this kind of data, e.g., miRNA induced protein regulations, we have developed a new 'more-than-random' model that tries to identify those co-regulated entities that are statistically significantly often co-regulated - even if only in a mild way. We will demonstrate this new method both theoretically and on a real data set of 810 human miRNAs on the EGFR/cell cycle network in an especially malevolent type of breast cancer. With this method we were able to identify three possible targets for this tumour type which showed very good results in subsequently conducted cell viability essays.