INTEGRATIVE ANALYSIS OF BREAST CANCER: DISSECTING HETEROGENEITY IN SAMPLES AND SIGNALS

Florian Markowetz
Cancer Research UK
Cambridge Research Institute
http://www.markowetzlab.org

Abstract:
I will talk about computational methods to address heterogeneity of breast cancer at different levels:

(i) At the *sample* level we often find cancer cells mixed with immune cells, stromal cells and others. This mixture of cells leads to a mixture of
signals when DNA, RNA, or proteins are measured in these samples. I will present an automated and quantitative approach to estimate cell mixtures and deconvolute molecular signals.

(ii) On the level of *patients*, different data types (like copy number alterations and gene expression) can offer complementary perspectives on
drivers of disease. When integrating these data to identify homogeneous subpopulations, it is important to distinguish cases where signals are
concordant from cases where they are contradictory. I will describe how patient-specific data fusion based on the hierarchical Dirichlet process can reveal prognostic cancer subtypes.

(iii) On the *population* level, there exist different distinct sub-types of breast cancer and genetic architecture differs between them. When inferring copy-number hotspots and regulatory networks, these sub-types have to be taken into account. In the last part of my talk, I will discuss how penalized regression can elucidate aberration hotspots mediating subtype-specific transcriptional responses in breast cancer.