Tissue Based Proteomics and Biomarker Discovery
– Multivariate Data Mining Strategies for Mass Spectral Imaging –

Raf Van de Plas (K.U. Leuven, ESAT-SCD)

Abstract -
In recent years the molecular mechanisms underlying many of the worlds most pressing diseases have been shown to exhibit a complexity that exceeds earlier expectations. On the diagnostic side, the search for disease-reporting biomarkers in easily accessible fluids such as blood has met with variable success. Most often, these biomarkers are present in extremely low concentrations (only detectable by immunological methods) and they often report secondary disease effects, thereby shedding only limited light on the true cause of the pathology. However, rather than examining fluids, it is the direct analysis of affected tissue that would prove very helpful towards elucidating the underlying patho-mechanism. Tissue analysis reports directly on complex metabolic alterations of a wide array of molecules, and does this close to the source so that the observations include, but are not restricted to, secondary effects. It also presents potential biomarkers undiluted and at their highest concentration, and potentially even tied to a specific tissue region or cell type.

Mass spectral imaging (MSI) has been gaining considerable momentum as a means of examining the spatial distribution of biomolecules throughout tissue. It is a relatively new technology that acquires thousands of individual mass spectra in situ from across a tissue section. MSI fuses the localization information of an imaging method with the molecular specificity and sensitivity of mass spectrometry. The ability to detect biomolecules such as proteins, peptides, metabolites, and lipids throughout tissue without the need for labeling, and its capability of measuring hundreds of molecules concurrently in a single experiment, have made MSI particularly valuable in exploratory settings where there is no prior hypothesis of relevant target molecules. As a result, MSI is rapidly becoming a potent exploratory and confirmatory instrument for tissue-based biomedical studies.

The most common use of MSI is to produce ion images. An ion image shows the spatial distribution throughout the tissue of one particular ion, mass, or molecule. Ion images fill a distinct niche within the molecular imaging field due to their label-free nature and often unrivaled specificity. However, considering ion images as the sole product of a MSI experiment would deny the full potential of the technology. The measurements of a MSI experiment typically comprise a vast amount of concurrently collected biochemical information, spanning both the spatial as well as the molecular mass domain. Although the data sets can become very large as a result, it makes MSI uniquely suited for studying inter-molecule relationships such as co-localization and correlated expression. The applications for this type of information range from insights into metabolic pathways to the exploration of co-localizing biomarkers. Besides univariate ion images, there is currently a lack of multivariate computational methods capable of filtering from this sea of data, the nuggets of information relevant to the biomedical study at hand.

This thesis attempts to answer this challenge with contributions along the entire mass spectral imaging workflow. This includes preprocessing methods (such as normalization and compression methods) to cancel certain noise types, unsupervised methods (such as non-negative matrix factorization and co-localization networks) to unravel the underlying patterns present in the data, and supervised methods (such as spatial querying) to answer specific biological questions. The thesis further adds a pragmatic wavelet compression strategy to counter the difficulties introduced by the data size. The doctorate is concluded by a biomedical case study of diabetes mellitus type 2 in which a combination of mass spectral imaging and the computational techniques developed in this work is used to find out, only by studying the tissue, where the metabolic processing of the diabetes models starts to deviate from normal metabolism.

- Promotors : Prof. dr. ir. B. De Moor & Prof. dr. E. Waelkens