BIOGNOSYS
NEXT GENERATION PROTEOMICS

Biognosys AG
Wagistrasse 21
8952 Schlieren (Zurich)
Switzerland

t: +41 (0) 44 738 20 40
f: +41 (0) 44 738 20 49
info@biognosys.com
www.biognosyscom

Deepest Profiling of Healthy Human Tissues to Date Sheds Light on Expressed Proteins

Abstract ID number: 312863

Presenter: ***Sandra Schär***

Introduction

Proteins are the biochemical actors in virtually all cellular processes in health and disease. As diseases predominantly manifest at the protein level, the analysis of proteins is of great importance in biological and medical research. The comprehensive assessment of the human proteome poses a major analytical challenge given the temporal variability and the dynamic range of expression. Furthermore, there is a large number of distinct proteoforms arising from alternative splicing, allelic variations, isoforms and protein prost translational modifications adding exponentially to the complexity of the 20,359 protein coding gene [1]. The aim of this work was to generate a high quality, uniform and semi quantitative proteome abundance atlas of different human tissues to generate a resource library.

Methods

Firstly, the most efficient two-dimensional fractionation method consisting of a combination of high pH reverse phase chromatography (HPRP) and high field asymmetric waveform ion mobility spectrometry (FAIMS) was determined. Secondly, the selected fractionation method was applied to 22 human tissue samples (brain, thyroid, breast, lung, lymph node, liver, gallbladder, adrenal gland, kidney, pancreas, colorectal cancer, colon, bladder, prostate, rectum, ovary, testis, testicular cancer, fallopian tube, melanoma, small intestine, tonsil), each separated in 60 pooled HPRP fractions and half of which measured via LC-MS/MS using an optimized ion mobility data dependent acquisition method. Thirdly, the human tissues were profiled in single-shot FAIMS data independent acquisition (DIA) using Biognosys' UltraDeep method to investigate the relative protein abundance across tissues thereby providing both quantitative as well as semi-quantitative datasets.

Preliminary Data

The generated deep tissue library encompassed 20'025 protein groups that could be mapped onto 17'715 coding genes, corresponding to roughly 90% of all protein-coding genes annotated within the neXtProt repository [1]. 17,230 measured proteins were categorized at the PE1 level, representing 93.6% of the proteins ever identified by mass spectrometry [2]. The number of identified protein groups was mostly consistent across all tissues and following known trends, indicating the uniformity of the generated dataset. Across all tissues, 10'755 protein groups were shared and detected in every profiled tissue. Gene ontology enrichment of the ubiquitously

Biognosys AG
Wagistrasse 21
8952 Schlieren (Zurich)
Switzerland

t: +41 (0) 44 738 20 40
f: +41 (0) 44 738 20 49
info@biognosys.com
www.biognosyscom

expressed proteins revealed common biological process repository including “cellular localization”, “RNA processing” and “catabolic processes”. On the other hand, when a GO enrichment was performed on the proteins found only in individual tissues specific biological processes indicative of the identity of the tissue were identified. Hence, the proteins found uniquely in specific tissues indicate the representative function of their biological role. We identified 37 proteins mapping to putative genes categorized as PE5 with more than three proteotypic peptides. Among these, the putative spermatogenesis-associated protein was identified by six peptides and the putative testis protein 13C was identified with 36 peptides found only in testis tissue.

DIA analysis quantified 15’772 protein groups in a single shot across the 22 tissues. Clustering-based analysis segregated the tissues into two main groups, with brain tissue being the most different. While several proteins exhibited tissue-specific expression patterns, housekeeping proteins were uniformly expressed across tissues. Overall, the generated uniform and comprehensive tissue protein atlas holds the potential to aid the selection of peptides for targeted assays, while creating an unbiased batch correction pool together with being a valuable training repository for deep learning models (with over 60 million PSMs identified) and could also be used as a launching pad for biomarker discovery.

Novel aspect

The deepest tissue library to date provides insights into protein expression, while accelerating targeted assays generation for biomarker discovery.