

A HIGH-THROUGHPUT PLATFORM FOR PROTEOME AND PHOSPHO-PROTEOME PROFILING OF TUMOR TISSUES

Jakob Vowinckel¹, Karel Novy¹, Thomas Corwin², Tobias Treiber¹, Vito Dozio¹, Roland Bruderer¹, Lukas Reiter¹, Eike-Christin von Leitner², Oliver Rinner¹, Claudia Escher¹

1) Biognosys AG, Wagistrasse 21, 8952 Schlieren, Switzerland 2) Indivumed GmbH, Falkenried 88, 20251 Hamburg, Germany



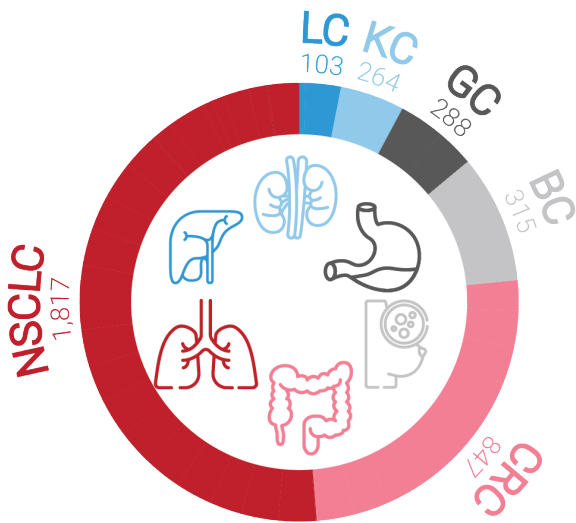
Jakob Vowinckel, PhD
Senior Scientific Project Manager
jakob.vowinckel@biognosys.com
www.biognosys.com



INTRODUCTION

Precision oncology requires a detailed molecular understanding of tumor biology. Phenotype and underlying cellular functions are best characterized by the study of the proteome. However, MS-based proteome profiling is underrepresented in precision medicine compared to DNA/RNA sequencing techniques. Limitations in instrument stability, reproducibility, sample throughput and data analysis have prevented large-scale proteome characterization

experiments. Recent developments in data-independent acquisition (DIA) LC-MS/MS and robust chromatographic separation now present the opportunity to make proteomics available to routine analysis. Here we present a workflow that is capable of routine profiling 850 whole proteome (WP) or 650 phospho-proteome (PP) tumor samples per month and platform with an average depth of 6,000 proteins (WP) or 30,000 phospho-peptides (PP).



CONCLUSIONS

- An optimized, semi-automated workflow enables high throughput deep proteome and phospho-proteome profiling of matching tumor and normal tissue samples from Indivumed's high quality collection of fresh-frozen bio-specimens.
- Rigorous quality control during sample collection, sample processing, data acquisition and analysis allows reproducible generation of data sets consisting of thousands of samples.
- In NSCLC patient lung tissue on average 5,903 protein groups and 28,819 phospho-peptides are quantified in each sample, with up to 7,346 protein groups quantified.
- Observed protein expression and phosphorylation status of known surrogates are in accordance with established knowledge and provide valuable insights to previously unknown markers relevant for tumor biology

METHODS & RESULTS

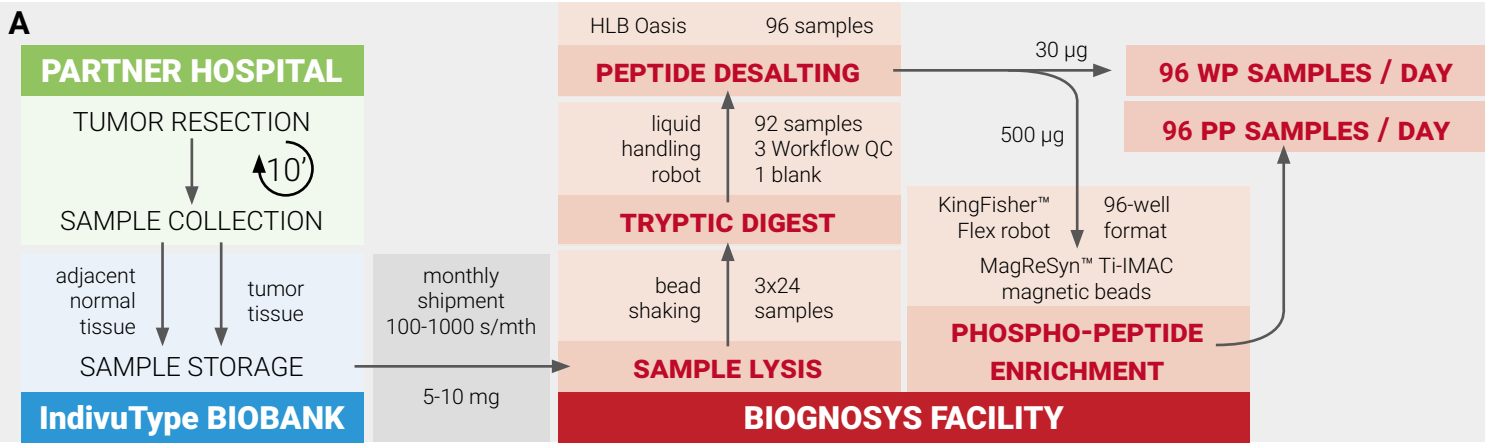


Figure 1: Sample Preparation Workflow Diagram

(A) Sample preparation of high quality fresh-frozen IndivuType samples in batches of 92 samples (+ 3 workflow QC and 1 blank) using partial robotic automation. **(B)** Automated QC validation and DIA data analysis by a workflow involving directDIA™ search using SpectroMine™ software and pre-set thresholds, and subsequent data analysis of individual raw files with tissue-specific spectral libraries using Spectronaut™.

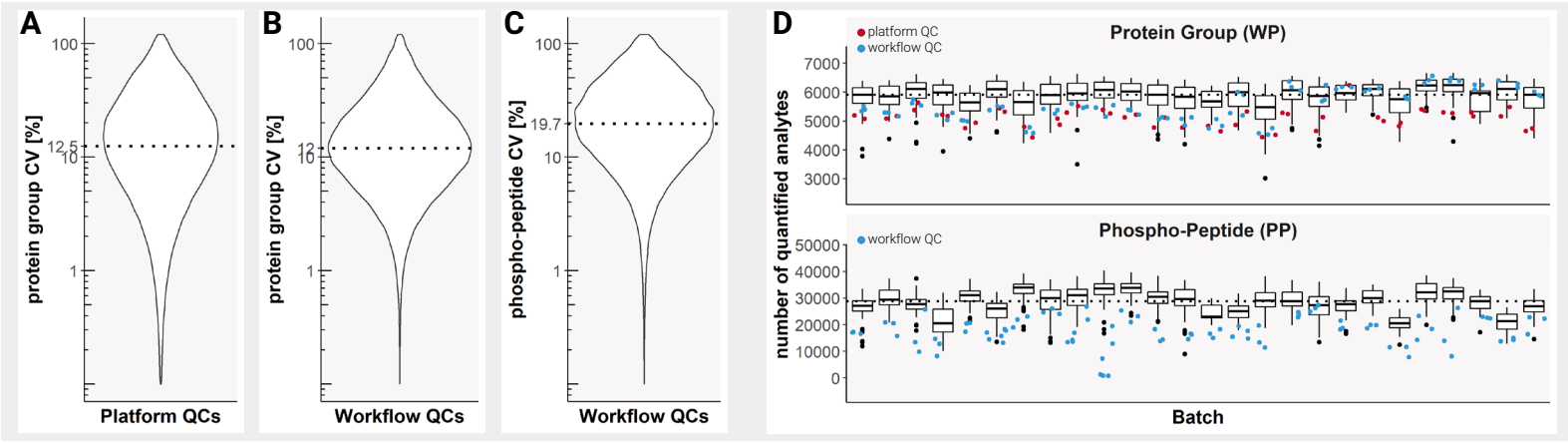


Figure 2: Reproducibility of Data Acquisition and Analysis

Over 43 batches, median protein group coefficients of variation (% CV) are 12.5 % for WP platform QC **(A)**, 12.0 % for WP workflow QC **(B)**, and median phospho-peptide % CV are 19.7 % **(C)**. In 26 NSCLC batches (1,755 samples) on average 5,903 protein groups and 28'819 phospho-peptides are quantified with peptide and protein FDR < 0.01, and site localization probability > 0.75 **(D)**.

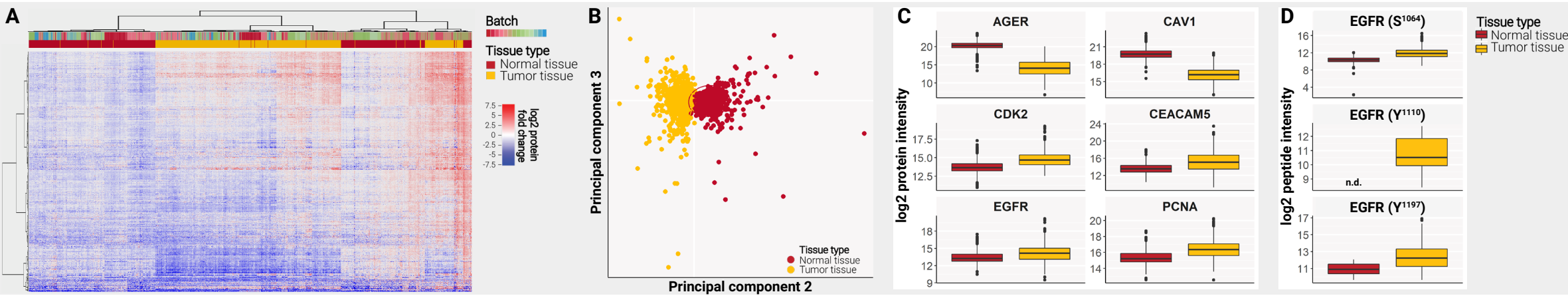
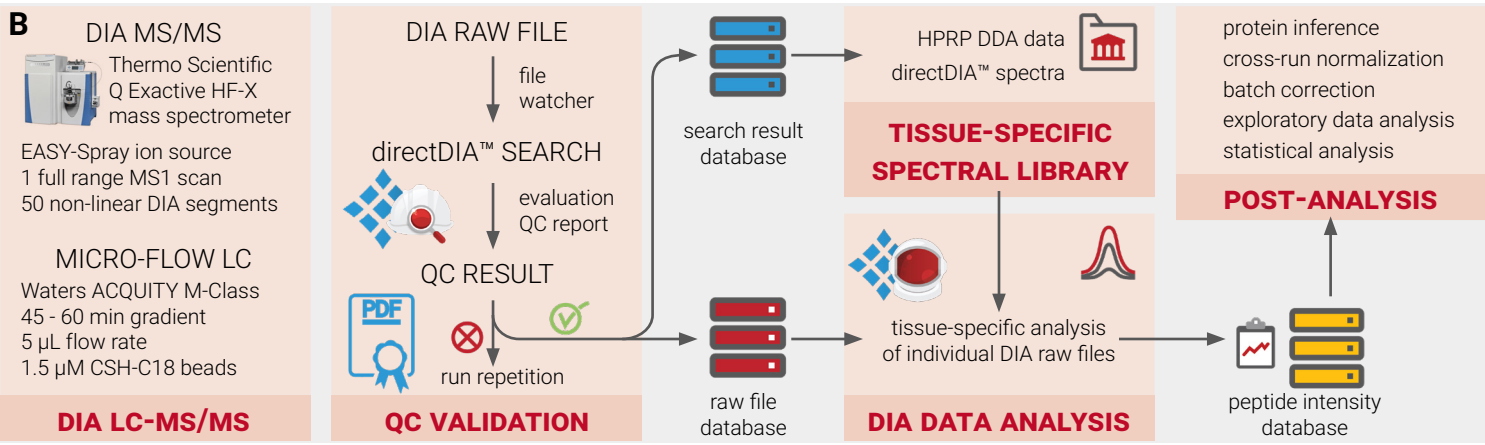


Figure 3: NSCLC Tumor Biology

Hierarchical clustering **(A)** and principal component analysis **(B)** of 7,346 protein intensity values mainly reveals co-clustering according to tissue type in 1,755 lung samples. Six known markers of lung cancer **(C)** show robust change between normal and tumor tissue. EGFR phosphorylation status on three known C-terminal sites **(D)** are consistently elevated in tumor compared to normal tissue.