A path from discovery to a targeted proteomics approach for the verification and validation of tissue derived biomarkers in coronary artery diseases

Chi D.L. Nguyen[1] , Austin L. Seal[3] , Annie Moradian[1] , Dragana Noe[1] , Stephen A. Whelan[1] , Mitra Mastali[2] , Sarah J. Parker[2] , Susan Mockus[1] , Jennifer E. Van Eyk[2] , David M. Herrington[3]

(1) Precision Biomarker Laboratories, Cedars Sinai Medical Center, Los Angeles, California 90048, United States (2) Heart Institute & Advanced Clinical Biosystems Research Institute, Cedars Sinai Medical Center, Los Angeles, California 90048, United States (3) Department of Cardiovascular Medicine, Wake Forest University, Winston-Salem, North Carolina 27101, United States

Cedars Sinai
Precision Biomarker Laboratories
www.cs-pbl.com

Introduction

Over the past decades, there has been significant improvements on mass spectrometry (MS) - based technology, such that proteomics is now a significant tool for biomarker research including discovery, analytical validation, and clinical validation. Discovery proteomics results in a tremendous amount of information which is valuable not only in the discovery of biomarkers but also in characterizing the chemical properties of those biomarkers. In this study, we describe in detail the path from collecting and utilizing the comprehensive information from various sources of discovery proteomics analysis to the creation of the targeted proteomics assays for the verification and validation of biomarkers in application on coronary artery disease (CAD). CAD is the most common type among the heart diseases and is one of the leading causes of death in current times. One major challenge with current therapeutic approaches for CAD is the lack of an early diagnostic tool, which can dramatically impact treatment outcomes.

Discovery Proteomics ProteoSweep™

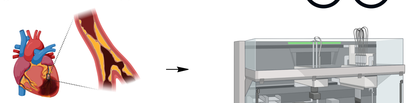
Targeted Proteomics ProteoMarker™

Methods

The high-throughput discovery proteomics platform (ProteoSweep™) was developed for the global proteome measurement in a variety of sample types delivering relative quantification information of the whole biological system with high sensitivity and reproducibility. This technology is not only significant in the discovery of biomarkers, but also in the development of targeted proteomics approaches.

Discovery Proteomics ProteoSweep™

Proteomics discovery mass spectrometry of approximately 200 aorta and left anterior descending (LAD) artery samples was used for the detection of potential biomarkers and as a spectral library for MRM (multiple reaction monitoring) method development. Publicly available spectral libraries from PRIDE and PlasmaPilot were also used as reference information for the development of the MRM targeted proteomics assay. Spectral library information (elution time, fragmentation pattern, etc.) was inspected and investigated using Skyline. PeCorA* was used to detect any discordant peptides among experimental groups.



>200 aorta and LAD artery samples

Fully automated sample preparation platform

high-throughput discovery platform

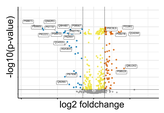
PRIDE
PRoteomics IDEntifications Database

PlasmaPilot
PlasmaPilot is based on experimentally defined bottom-up human plasma proteomics.
Our goal is provide a prospective sample preparation guide for the quantitative MS based analysis of LFQ proteins.

PeCorA

Peptide Correlation Analysis*

ProEpic™



Spectral library

Skyline

- Elution profile
- Fragmentation pattern
- MS responsive peptides
- Matrix background information

Detection of discordant peptides among experiment groups

Detection of potential biomarker proteins

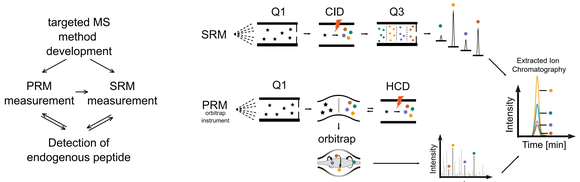
36 proteins - 156 peptides

* Peptide Correlation Analysis (PeCorA) Reveals Differential Proteoform Regulation. Maria Dermit, Trenton M. Peters-Clarke, Evgenia Shishkova, and Jesse G. Meyer. Journal of Proteome Research 2021 20 (4), 1972-1980. DOI: 10.1021/acs.jproteome.0c00602

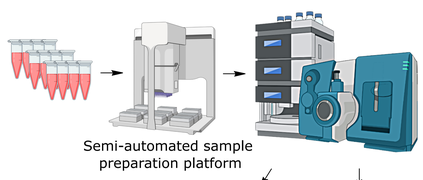
The targeted proteomics platform (ProteoMarker™) delivers quantitative information for a selection of proteins of interest with high sensitivity, selectivity and precision. This platform is designed for the verification and validation of biomarkers in large number of clinical samples.

Targeted Proteomics ProteoMarker™

Targeted proteomics method development



Targeted proteomics large scale measurement

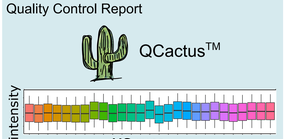


Semi-automated sample preparation platform

For the targeted proteomics measurement in large scale, the sample preparation step was performed in semi-automated mode utilizing the Bravo Automated Liquid Handling Platform (Agilent). MRM measurement was performed using the analytical flow LC system coupled with AB Sciex™ triple quadruple mass spectrometer was applied for the robustness and stability maintaining the high standard of sensitivity and selectivity. Quality control (QC) report was generated using the in-house developed QC software (QCactus™). The MRM raw data analysis report was performed using ProEpic™ platform.

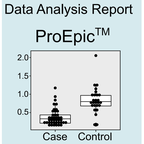
Quality Control Report

QCactus™



Data Analysis Report

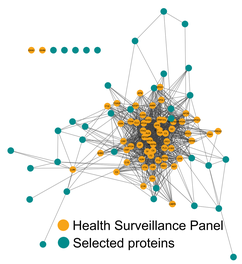
ProEpic™

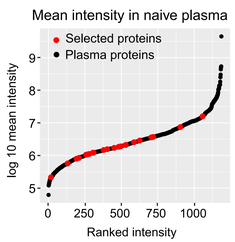


Results

Targeted proteomics method development and verification

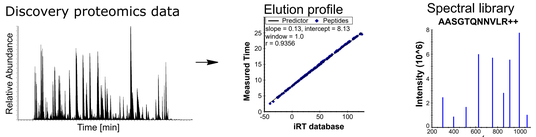
1. Selection of potential tissue-derived biomarkers for CAD



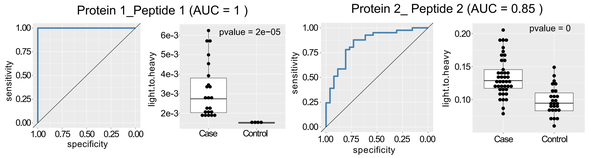


Left panel: 36 potential biomarker proteins were selected from the discovery mass spectrometry runs. Protein-protein interaction network shows the separation of the selected proteins (blue dots) to each other in contrast to the tight connection between the previously developed health surveillance panel (HSP) orange dots). Right panel: The majority of the selected proteins are in the mid - low abundance range of the plasma proteom, which is challenging using the analytical flowrate for MRM method development.

2. Targeted proteomics method development based on discovery information



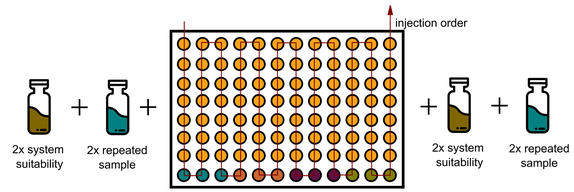
3. Evaluation of the MRM targeted assay in plasma samples of CAD patients



28 proteins out of the 36 potential biomarker proteins could be detected in plasma samples with a minimum of one unique peptide. Among them, 14 proteins show significant differences between case and control and 5 proteins have strong predictive power (AUC >= 0.8). In this cohort, 42 plasma samples derived from female donors, ages from 52 to 76 years old with CAD and 26 healthy matched controls were used.

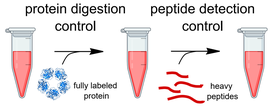
Targeted proteomics large scale measurement

1. Qualitity control for large scale measurement of plasma samples



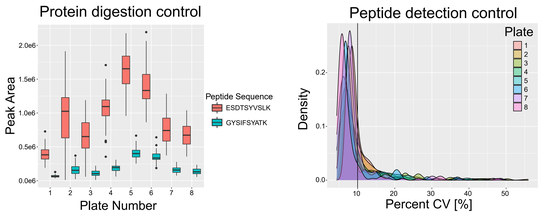
For the quantification of selected biomarkers in a large number of samples, a 5-levels of QC system was applied including system suitability control, repeated samples measurement, protein digestion control, peptide detection control and defined ratio analysis.

- Batch quality samples = digested human plasma
- 1 Human : 2 Chicken plasma
- 2 Human : 1 Chicken plasma
- Digested chicken plasma
- Tested samples



Protein digestion control: An intact full-length recombinant protein was spiked in each sample prior to trypsin digestion to determine the digestion efficiency of each sample in a batch measurement.
Peptide detection control: A collection of isotopic labeled heavy peptide was spiked in each sample not only for the quantification of the endogenous peptides but also for the quality control of the sample preparation and detection process.

2. Protein digestion control and peptide detection control for large scale MRM measurement



Protein digestion control: Two tryptic peptides of the spiked-in recombinant protein were monitored over the total number of eight 96-well plates (~660 samples).
Peptide detection control: The detection of isotopic labeled peptides (heavy peptides) was monitored in all samples of eight 96-well plates. The median values of percent coefficient of variation (CV) in all plates are less than 10 %.

Conclusion

- This study illustrates an efficient workflow to develop a targeted proteomics approach from all the possible discovery data sources relevant to the disease of interest.
- The detection of 5 potential tissue-derived biomarker proteins showing high predictive values in a small CAD cohort demonstrate the potential utility of using these proteins as diagnostic biomarker panel for the early detection of CAD.
- Further validation in a larger cohort size of ~660 plasma samples was performed to investigate the potential utility for clinical use.
- For the large scale measurement, a 5-levels QC system was developed including system suitability control, repeated samples measurement, protein digestion control, peptide detection control and defined ratio analysis.