A high-throughput automated proteomic approach to explore plasma biomarkers: application in a cohort of hospitalized COVID-19 Patients

Luciana Godoy Viana[1], Andrea Tedesco Faccio[1], Caroline Lacerra de Souza[1], Jessica Bassani Borges[2], Gisele Medeiros Bastos[2], Maria Carolina Tostes Pintão[1], Karina Helena Morais Cardozo[1], Rozana Mesquita Ciconelli[2], Valdemir Melechco Carvalho[1]

[1.] GF, Grupo Fleury, Av. Morumbi, 8860 - São Paulo-SP;
[2.] BP, Hospital Beneficiência Portuguesa de São Paulo, Rua Maestro Cardim 769, São Paulo-SP;

Introduction:

Plasma proteomics, a highly sought-after substract for clinical research and biomarker exploration, encounters several obstacles. Its intrinsic complexity, characterized by a vast dynamic range of protein abundance necessitates labor-intensive enrichment strategies that may introduce variability in protein quantitation. Building upon our previous work —a fully sample preparation automated test followed by targeted detection of SARS-CoV-2 viral proteins using mass spectrometry — we extended this approach to untargeted proteomics analysis. We applied this approach in a cohort of hospitalized patients to explore a set of biomarkers that are differentially expressed according to the severity of COVID-19 infection.

Materials and Methods:

Samples were collected from patients of all sexes, aged 18 years or older, diagnosed with COVID-19 through the real-time RT-PCR molecular test at the Beneficência Portuguesa Hospital of São Paulo. Two sample extraction methods were evaluated: the first utilized magnetic particles based on a modified solid phase-enhanced sample preparation (SP3) protocol, adapted for protein reduction, alkylation, and tryptic digestion automated on a robotic platform. The second method was based on perchloric acid precipitation, followed by solid phase extraction. Analyses were performed using high-resolution mass spectrometry in data-independent acquisition (DIA) mode, with data processed using DIA-NN software (v.1.9). The statistical analysis and graph construction were performed using R (v.4.3.1), following the MS-DAP pipeline (v.1.3.0).

Results and Discussion:

Patients were classified based on disease severity. Differential expression analysis (DEA) of samples processed by the SP3 method revealed 294 specific proteins, including complement factors (C8A, CFH) and inflammation modulators (APOA1, SAA2), with a q-value threshold of <0.01 and log2 fold-change >0.49. Conversely, the perCA protocol identified 650 proteins, including those involved in inflammation (ITIH4) and coagulation (TF), with a q-value threshold of <0.01 and log2 fold-change >0.41. Notably, 203 proteins were common to both methodologies. The perCA method selectively precipitates abundant proteins, allowing each technique to identify distinct sets of proteins, highlighting their complementary nature in proteomic analysis.

Conclusion:

This study demonstrates the effectiveness of an automated sample preparation approach that combines the SP3 and perCA methods within a high-capacity pipeline, marking a significant advancement in large-scale proteomic analysis for COVID-19 patients. The successful application of this automated workflow in a cohort of hospitalized COVID-19 patients serves as a proof of concept for the capabilities of this platform, which enables robust detection and monitoring of plasma biomarkers, thereby enhancing our understanding of the disease\'s pathogenesis. Future

studies in independent cohorts will be essential to validate the identified biomarkers and establish their clinical relevance.

Agradecimentos: BP and GF