Proteonano™: a novel deep proteomics platform with picogram sensitivity and its application in Alzheimer's disease

Dan Liu[1, 2, 3], Yonghao Zhang[4], Xiehua Ouyang[4], Shanshan Lv[4], Yanting Meng[4], Libing Wang[4], Zhiquan Cao[4], Yi Wang[4], Hao Wu[4#], Yan Zeng[1,2,3#]

1.Brain Science and Advanced Technology Institute, Wuhan University of Science and Technology, Wuhan, China, 2. Geriatric Hospital Affiliated to Wuhan University of Science and Technology, Wuhan, China, 3. School of Public Health, Wuhan University of Science and Technology, Wuhan, China, 4. Nanomics Biotechnology Co., Ltd., Hangzhou, Zhejiang, China

Nanomics
The Next-Gen Proteomics Company

武漢科技大学
WUHAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

ABSTRACT

Introduction: Alzheimer's disease (AD), the predominant form of dementia, manifests as a protracted progression exceeding a decade. It is documented that concomitant cerebral damage accrues prior to the manifestation of overt behavioral symptoms. Timely disease detection is paramount for impeding progression and augmenting patient quality of life. However, extant biomarkers, notably amyloid β (Aβ40, Aβ42), demonstrate suboptimal performance in identifying individuals at early stage of disease development. Thus, it is critical to identify novel biomarkers that distinguish patients at distinct stages of disease development. To meet this demand, we developed a deep, untargeted plasma proteome profiling technology (Proteonano™ platform) to facilitate early detection of AD.

Methods: Proteonano™ technology was developed as an affinity-selective mass spectrometry platform, including usage of nanoparticle-based affinity protein binders (nanobinders) to enrich low abundance proteins and employment of an automated pre-treatment workstation for parallel sample preparation. Patients were serially enrolled under approved ethical review. Plasma samples were collected at time of enrollment, and participants were stratified as normal (N), mild cognitive impairment (MCI), and dementia (D). Plasma samples were processed through the Proteonano™ pipeline and analyzed by a ThermoFisher Orbitrap Astral mass spectrometer at data independent acquisition mode. Raw data were analyzed by using DIA-NN, normalized, further processed by using a customized biostastic and bioinformatic pipeline.

Results: Plasma samples from 206 serially enrolled participants (N=142, MCI=35, D=29) were used for untargeted proteomic analysis by using the Proteonano™ platform. 4347 protein groups were identified, with 2344 ± 37 (AVG ± SE) proteins groups identified in each sample. 2704 of the protein groups were mapped to human plasma protein project published plasma protein database. Concentrations of these proteins spanned nine orders of magnitude, and the lowest abundant protein had a reported plasma concentration of 1.6 pg/mL. Differential protein expression analyses showed 64, 91, and 159 proteins had different abundance between MCI and N, D and MCI, and D and N groups, respectively. The most upregulated protein in MCI group relative to N group was matrix Gla protein (MGP) and most downregulated protein was non-erythrocytic β spectrin (SPTBN1). Eight feature selection methods, including least absolute shrinkage and selection operator (LASSO), and random forest (RF), were employed, and top features selected in each method were subjected to Akaike information criteria (AIC) based model selection. Identified features were then combined for another round of selection. Differentiating powers of these models were assessed by receiver operating curve (ROC) and precision and recall (PR) methods. The best model differentiating MCI and N groups contains nine proteins, with a ROC-AUC of 0.92 (95 % confidence interval: 0.89-0.98), and PR-AUC of 0.84. This model is superior to the discriminative power of Aβ 40, with ROC-AUC=0.80. Similarly, excellent models differentiating D and N (8 proteins, ROC-AUC=0.99, PR-AUC=0.96) and D and MCI groups (9 proteins, ROC-AUC=0.98, PR-AUC=0.97) were identified. These results indicate a nanobinder assisted untargeted proteomics approach can effectively identify protein features differentiating patients at different cognitive states, which will eventually lead to improved early detection of patients with elevated risk for AD and dementia.

Novel aspect: Proteonano™ platform is a powerful biomarker discovery tool, which enabled identifying a highly discriminative protein panel for AD early detection.

METHOD

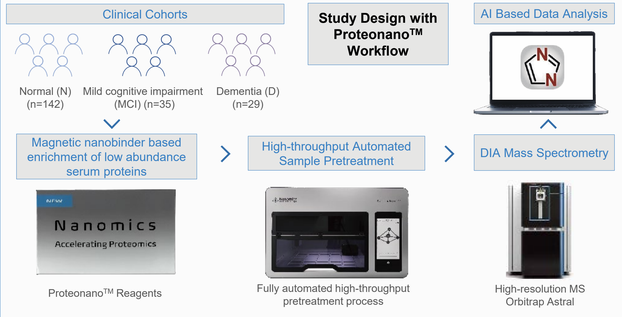

Figure 1. Study outline. A community cohort of elderly individuals were enrolled and plasma samples collected . Samples were pre-processed by using Proteonano™ magnetic nanobinders by using an automatic workflow, and subjected to mass spectrometry (Orbitrap Astral mass spectrometer at data independent acquisition mode).

RESULTS

Proteonano™ enables deep proteomic analysis of plasma samples from a community cohort with different levels of cognitive defects

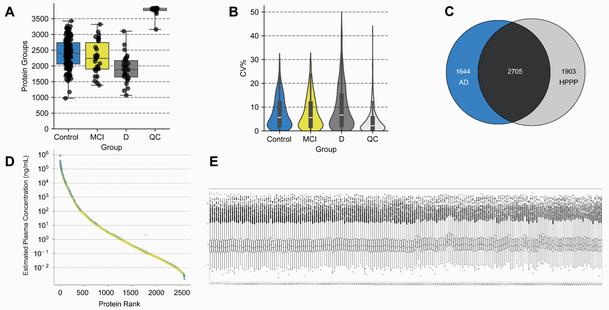

Figure 2. Sample preprocessing and LC-MS analysis. (A) Protein groups detected in plasma samples from an elderly community cohort (Hubei Memory and Aging Cohort Study, HMACS) of patient samples with different levels of cognitive defects. Normal (N) patients do not have observable cognitive defect while mild cognitive decline (MCI) and dementia (D) groups have varying levels of cognitive defects. A separate normal plasma sample was used for assay quality control (QC). (B). Distribution of coefficients of variation (CV) for detected protein groups in different patient subpopulations and QC samples. (C) Overlap between identified protein groups in the patient cohort (blue) and and those reported by Human Plasma Proteome Project (HPPP, grey). (D) Distribution of protein abundance of identified protein groups in all samples. Protein concentrations were referenced to the HPPP database. (E) Distribution of protein abundance in each of the patient within the cohort prior to data normalization.

Differentially expressed proteins between patient groups

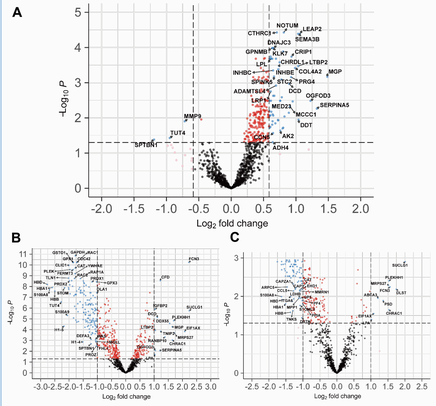

Figure 3. Differentially expressed proteins (DEP) in different patient groups. After data normalization, DEPs were determined. (A) 64 DEPs between N and MCI groups. (B) 159 DEPs between N and D groups. (C) 91 DEPs between D and MCI groups.

Identification of pathway differences between MCI and normal patients

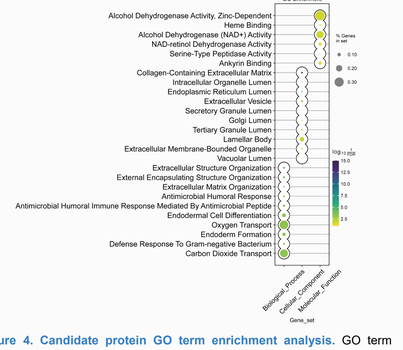

Figure 4. Candidate protein GO term enrichment analysis. GO term analysis identified multiple pathway differences between MCI and N patients.

Functional interactions among differentially expressed proteins between MCI and normal patients

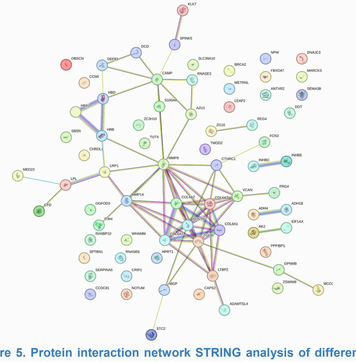

Figure 5. Protein interaction network STRING analysis of differentially expressed proteins between MCI and normal patietnts. Different line colors between proteins represent types of associations. The length of lines represent the confidence of the interaction associations; shorter distances indicate more reliable predictions.

Discriminative power of differentially expressed proteins to distinguish MCI and normal patients

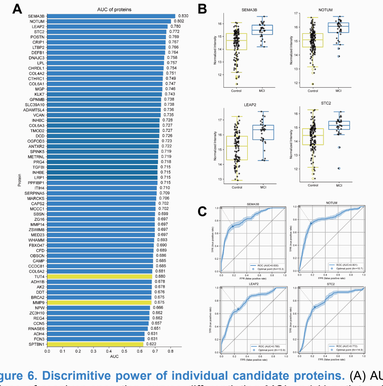

Figure 6. Discrimitive power of individual candidate proteins. (A) AUC values of receiver operation curves differentiating MCI and N patietns (B) Relative abundance of top 4 proteins with highest AUC values. (C) ROC curves of top 4 proteins with highest AUC values.

Identification of effective biomarker combinations to differentiate normal and MCI patients

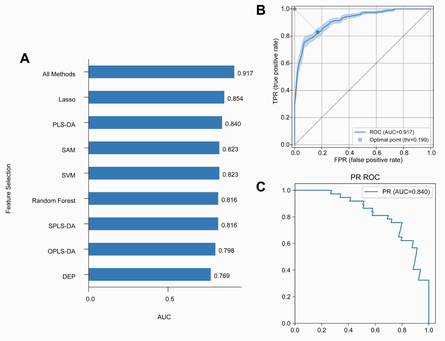

Figure 7. Assessment of multivariate models. (A) AUC values of models created by using different feature selection methods. After feature selection, Akaike information criteria (AIC) was used to identify best combinations of features selected by each method, and effectiveness for each model was assessed by ROC method. (B) ROC curve of the best model with 9 proteins based on all feature selection methods. (C) Precision recall curve of the best model with 9 proteins based on all feature selection methods.

CONCLUTION AND DISCUSSION

- Utilization of Proteonano™ platform allowed deep profiling of plasma proteomics in plasma sample from a community based cognitive study, enabling enabled identification of multiple differentially expressed proteins.
- While individual proteins have limited predictive power, use of multivariate analysis identified a 9 protein panel that effectively differentiate normal individuals and patients with mild cognitive decline (ROC-AUC=0.92, PR-AUC=0.84).
- Subsequent validation studies are needed to assess the effectiveness of this multivariate biomarker panel for identifying MCI patients.

ACKNOWLEDGEMENTS

This work was supported by grants from Science and Technology Innovation 2030 Major Projects of China (2022ZD0211600), and National Natural Science Foundation of China (81870901, 82071272, and 72174159), and internal fundings from Nanomics Biotech. YZ, XO, SL, YM, LW, ZC. YW, and HW are employees of Nanomics Biotechnology.