Performance and Utility of Illumina Protein Prep Proteomics Assay in Rare Disease Clinical Research

Matthew Brown[1], Athanasios Kousathanas[1], Lily Hoa[1], Greg Elgar[1], Jamie Ellingford[1], Alexander Ho[1], Susan Walker[1], Ali Crawford[2], Simon Freedman[2], Michael Mehan[3], Parul Prabhat[4], Andrew Slatter[4], Nithya Subramanian[4], Yongwen Zhuang[2], Dalia Daujotyte[2].
[1]Genomics England, London, United Kingdom, [2]Illumina, San Diego, CA, USA, [3]Illumina, San Diego, CO, USA, [4]Illumina, Cambridge, United Kingdom.

Background:

Better methods are needed to increase the proportion of patients with rare diseases who are referred to genomic medical services and receive a successful diagnosis.

Material and Methods:

We evaluated novel proteomic assays combining Standard Biotool's (SBI, formerly SomaLogic) SomaScan technology and Illumina Protein Prep to measure 6,056 proteins in serum samples from 1,530 participants in the 100,000 Genomes Project with rare diseases.

Results:

Technical performance was assessed by comparing SBI SomaScan v4.1 with Illumina Protein Prep in 50 samples, yielding a median Spearman r^2 of 0.58. SOMAmer probes exceeded the limit of detection (LoD) in >84% of all samples, >96% in 80% of samples, and >99% in 50% of samples (CLSI method) (Figure 1). The median coefficient of variation was 7.7% (Q90 12.3%).

Proteomics data effectively predicted gender (AUC=0.986) and age (r^2=0.934)(Figure 2), consistent with UK Biobank findings. Across 19 rare disease categories, 470 proteins were differentially abundant (FDR p-value <0.05). Categories with the highest differentially abundant proteins included congenital anomalies of the kidneys (n=178), familial thoracic aortic aneurysm (n=50), familial hypercholesterolaemia (n=35), hypertrophic cardiomyopathy (n=29), and dilated cardiomyopathy (n=27). In contrast, intrauterine growth retardation, intellectual disability, epileptic encephalopathy, and Brugada syndrome showed few or no differentially abundant proteins (Figure 3).

Inspection of specific differentially abundant proteins demonstrates that these are enriched for proteins known to be involved in or to be biomarkers for the disease category concerned, or in some cases their encoding genes to be mutated in conditions in that disease category (Figure 4). This strongly supports further research into the potential for proteomics to assist with diagnosis of rare genetic disorders.

Figure 1:

% of SOMAmers above limit of detection per sample

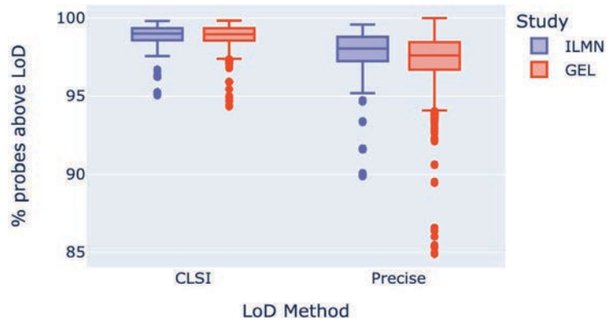


Figure 2:

Proteomics correctly predicts sex and age.

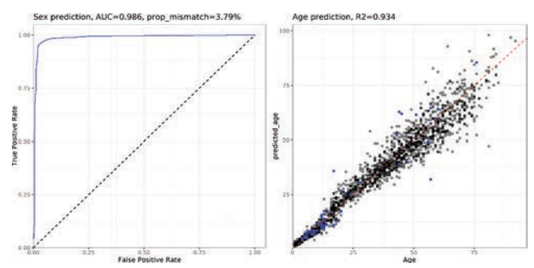


Figure 3:

Differential abundance varied widely by disease category

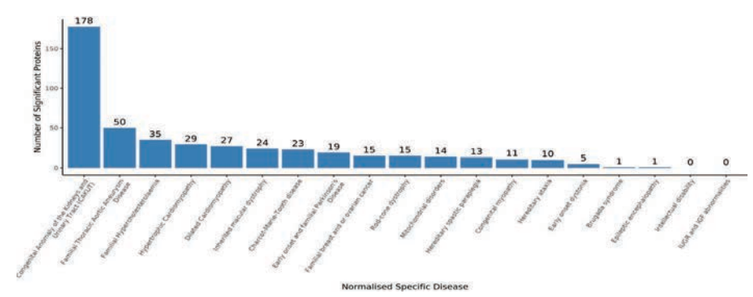


Figure 4:

Differentially abundant proteins link strongly with known factors involved in the disease category concerned.

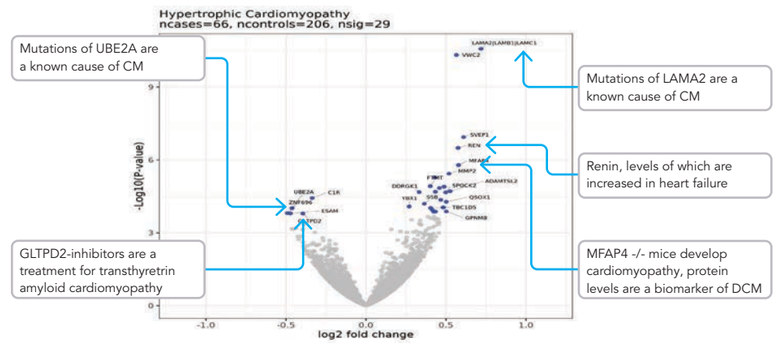
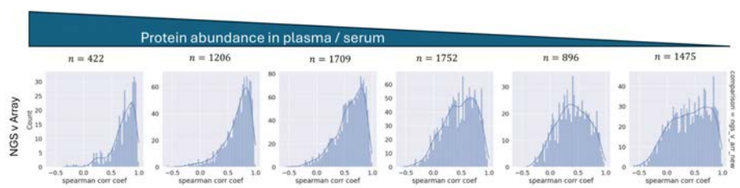


Figure 5:

Correlation stratified abundance / compression groups. The median correlation between NGS and array data was 0.58. This was much better for high than low abundance proteins (Q90 correlation 0.875, Q10 correlation 0.074).



Conclusion:

This study demonstrates the Illumina Protein Prep assay's high sensitivity and strong correlation with SBI SomaScan data. The abundant differential proteins in specific disease categories highlight the potential of proteomics to advance the diagnosis of certain rare genetic diseases by elucidating disease-associated genes and pathways. On completion this data will be made available in the Genomics England National Genome Research Library.