A novel proteogenomics approach identifying key proteins in severe COPD

Brandsma, C.A.; Guryev, V.; Timens, W.; Postma, D.S.; Bischoff, R.; Yakovleva, M.; Marko-Varga, G.; Fehniger, T.; Van Den Berge, M.; Horvatovich, P.

Published in:
American Journal of Respiratory and Critical Care Medicine

A Novel Proteogenomics Approach Identifying Key Proteins In Severe COPD

C. A. Brandsma[1], V. Guryev[2], W. Timens[1], D. S. Postma[3], R. Bischoff[4], M. Yakovleva[5], G. Marko-Varga[5], T. Fehniger[5], M. van den Berge[3], P. Horvatovich[4]

[1]Pathology and Medical Biology, University Medical Center Groningen, Groningen, Netherlands, [2]ERIBA, University Medical Center Groningen, Groningen, Netherlands, [3]Pulmonology, Groningen Research Institute for Asthma and COPD, University Medical Center Gronigen, Groningen, Netherlands, [4]Department of Analytical Biochemistry, University of Groningen, Groningen, Netherlands, [5]Lund University, Center of Excellence in Biological and Medical Mass Spectrometry, Lund, Sweden

Corresponding author's email: c.a.brandsma@umcg.nl

Rationale
COPD is influenced by many genetic and environmental factors. Genetic studies provided insights in the molecular mechanisms underlying COPD, but did not inform on protein changes. We propose that whichever factor contributes to disease, the final common pathway for disease development lies in the proteome. Therefore, we performed a combined proteome and RNA sequencing analysis comparing severe COPD and non-COPD control lung tissue.

Methods
Frozen lung tissue samples from 10 ex-smoking stage IV COPD patients and 10 ex-smoking non-COPD controls were used for protein and RNA extraction using urea and ammonium bicarbonate buffer and Trizol, respectively. The trypsin digested protein extract was analysed with LC-MS/MS approach using the Q-Exactive Plus (Thermo Scientific). RNA sequencing was performed with polyA-selected RNA sequencing using strand-specific quantitative NextFlex qRNA-Seq kit (Bio scientific) on Illumina HiSeq 2500 equipment with 20 million sequencing reads. We constructed patient specific protein reference databases using the RNA sequencing data and these were used to optimise peptide and protein identification. Differential gene and protein expression was assessed using linear models correcting for age, gender, age x gender interaction, and unwanted variation. FDR<0.05 was considered as statistically significant.

Results
The lung proteome analysis identified 1,430 proteins that were consistently expressed across controls and/or COPD sample groups. 61 proteins were up- and 50 proteins downregulated in severe COPD. This includes increased protein levels of microfibrillar associated protein 4 (MFAP4) and A-kinase anchoring protein 12 (AKAP12), and decreased levels of the thymosins beta 4, X-linked and beta 10 (TMSB4X and TSMB10).
RNAseq analysis identified 226 genes with increased and 124 genes with decreased expression in severe COPD. Among the genes with the strongest increase in expression were *AKAP12*, which was also increased at the protein level, and several other genes that are relevant for COPD such as *collagen type 14 alpha 1* (*COL14A1*), *thrombospondin 1 and 2* (*THBS1* and *THBS2*) and *interleukin-8* (*IL8*), and *interleukin-6* (*IL6*).

Conclusions
Using a novel proteogenomics approach we demonstrate differential protein expression of 111 proteins in lung tissue from severe COPD patients compared to non-COPD controls. MFAP4, a protein present in elastic fibres, is of particular interest since we previously found increased *MFAP4* gene expression levels in lung tissue samples from a more heterogeneous group of COPD patients using mRNA arrays. Another interesting candidate is AKAP12, which is increased at both gene and protein level and has been implicated in airway smooth muscle function and COPD previously.

This abstract is funded by: UMCG/ERIBA seeding grant