Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Shotgun proteomics differentiates COPD from asthma patients, smokers, and non-smoking controls

Tatiane Martins Lima Crosara Bastos
Federal University of Uberlândia
Luciana Oliveira e Silva
Federal University of Uberlândia
Bianca Uliana Picolo
Federal University of Uberlândia
Hebreia Oliveira Almeida-Souza
Federal University of Uberlândia
Mario Machado Martins
Federal University of Uberlândia
Robinson Sabino-Silva
Federal University of Uberlândia
Dave Singh
University of Manchester
Thulio Marquez Cunha
thulio.cunha@ufu.br

Federal University of Uberlândia

Article

Keywords: Shotgun, proteomics, COPD, asthma, smokers, biomarkers

Posted Date: May 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9441451/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

A biomarker panel that improves chronic obstructive pulmonary disease (COPD) diagnosis, even in the presence of confounding factors such as asthma and smoking, is of great importance. Label-free proteomic analysis was conducted in four groups (10 subjects each): COPD patients, asthma patients, smoker controls, and non-smoker controls. A total of 1,469 proteins were identified in the 40 serum samples, of which 29 were differentially expressed in the COPD group ($p < 0.05$ and > 2.0-fold change) compared with asthma and control groups. Five differentially expressed proteins (DEPs) stood out in volcano plots and heatmaps: ubiquitin carboxyl-terminal hydrolase 31 (USP31), RING finger protein 10 (RNF10) and transforming acidic coiled-coil containing protein 2 (TACC2) differentiated COPD from both control groups and asthma patients, while mitogen-activated protein 3-kinase 10 (MAP3K10) and collagen type V alpha 1 chain (COL5A1) differentiated COPD from healthy controls. Logistic regression models with 5-fold cross-validation, based on this protein set, achieved over 80% accuracy in diagnosing COPD. DEPs were explored for pathway enrichment using Reactome and Protein Analysis THrough Evolutionary Relationships (PANTHER), revealing pathways related to COPD pathophysiology. Knowledge of these alterations is important for identifying novel biomarkers to aid diagnosis and therapeutic strategies, contributing to precision medicine in COPD.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is the fourth leading cause of global mortality[1]. As a heterogeneous condition, its diverse phenotypes can complicate the differential diagnosis from other respiratory diseases, particularly asthma[2]. Both COPD and asthma share similar clinical and biological features, such as chronic inflammation and airflow obstruction[3]. However, COPD is characterized by persistent and incompletely reversible airflow limitation, the presence of emphysema and epithelial barrier dysfunction[4], while asthma is characterized by reversible airflow obstruction and airway hyperresponsiveness[5].

Currently, clinical features and spirometry tests are used to diagnose both diseases. However, lung function measurements do not provide detailed information regarding underlying pathological processes that may differentiate COPD from asthma [6].

The pathological mechanisms of COPD are primarily due to small airway disease by bronchiolar remodeling, mucus obstruction, peribronchial fibrosis, structural changes in the airway epithelium, and immune cell involvement. Emphysema is a destruction of the alveolar septum and occurs due to protease-antiprotease disbalance[7]. Characterizing the biological pathways and inflammatory mechanisms underlying COPD pathogenesis is important for identifying new biomarkers and effective treatments using drug targets [8]. Biomarker research has been boosted by high-performance sequencing techniques such as transcriptomics, metabolomics and proteomics [9].

Untargeted proteomics can identify proteins not previously associated with respiratory diseases such as COPD and asthma [10]. Liquid-chromatography tandem mass spectrometry (LC-MS/MS), which

combines two of the most powerful high-throughput techniques, facilitates the identification of candidate diagnostic and prognostic biomarkers when associated with bioinformatic tools [11]. Blood fractions (serum, plasma) stand as key sources of biological markers, offering minimal risk and allowing repeated sample collection. Previous untargeted LC-MS/MS-based proteomic studies of serum from COPD patients have demonstrated the ability of this approach to distinguish COPD from healthy controls [11, 12], and lung cancer [13], based on the comparison of differentially expressed proteins (DEPs).

The present study analyzed serum samples from well-defined COPD and asthma patients as well as smoking and non-smoking controls using label-free LC-MS/MS. This design extends beyond previous studies including smoking controls and asthma patients addresses an important clinical limitation. The identified proteins were evaluated for their potential to differentiate the COPD from asthma and control groups. Differentially expressed proteins were further analyzed using bioinformatics tools to explore altered signaling pathways in COPD.

RESULTS

Participants and clinical outcomes

Demographics and clinical characteristics of the subjects are shown in Table 1. There was no significant difference in age between COPD patients and smoking controls. COPD patients were older than asthma patients and the non-smoking control group (p = 0.006). Mostly serum proteins were similar (p > 0.05) in all groups, such as CRP and α1 antitrypsin concentration. Although homocysteine presented levels in physiological values, it was higher in COPD than in asthma patients (p=0.047). Asthma patients presented increased blood eosinophil cells compared to the non-smoking group (p=0.016), and serum total IgE was higher than COPD and both control groups (p=0.002).

Spirometry showed airflow obstruction in both asthma and COPD patients, with worse airflow in COPD.

LC-MS/MS Analysis and Identification of DEPs

Shotgun proteomic analysis identified 1,469 proteins in serum samples from 40 participants and 4 analyzed groups . A total of 29 were differentially expressed (DEPs) between the comparison groups (COPD *vs.* asthma; COPD *vs*. smoking, and COPD *vs*. non-smoking). The number of DEPs for each compared group and their status are summarized in Table 2. Heatmaps and volcano plots (Figure 1) were constructed to better visualize DEPs. The proteins showing substantial differential expression are highlighted in volcano plots for all comparisons with COPD patients.

Each identified DEP was named and linked to its gene code/ UniProt accession and is listed in Table 3. The proteins were listed in order of importance to differentiate COPD from all compared groups, using the number of distinct groups first and analysis of heatmaps after (highest fold change (FC) and lowest *p*-value) as criteria. The top five listed proteins were upregulated in the COPD group: ubiquitin carboxyl-terminal hydrolase 31 (USP31), mitogen-activated protein 3- kinase 10 (MAP3K10), RING finger protein 10 (RNF10), transforming acidic coiled-coil containing protein 2 (TACC2), and collagen, type V alpha 1(

COL5A1). Three of these proteins (USP31, RNF10, and TACC2) were capable in differentiating COPD from both control groups and asthma patients.

Top five DEPs capable to diagnose COPD.

The five proteins demonstrating significant differential expression in COPD patients compared to control groups were selected to construct a Principal Component Analysis (PCA) model to distinguish COPD group from both smoking and non-smoking controls (Figure 2 b, c).Among these proteins, USP31, RNF10, and TACC2 effectively differentiated COPD patients from all groups, including asthma patients. Subsenquently, another PCA model was evaluated to distinguish COPD group from asthma group, using these three proteins (Figure 2a). The explained variance for group separation was 97% for COPD vs. asthma (using 2 components), 93% for COPD vs. smoking (3 components), and 97% for COPD vs. non-smoking (2 components). Additionally, supervised evaluation using Logistic Regression model with five-fold cross-validation was performed, and sensitivity and specificity were derived from the confusion matrix. Subsequently, ROC curves were generated for each comparison to illustrate the diagnostic capacity of these proteins (Figure 2 d, e, and f). For the COPD *vs*. asthma comparison, the logistic regression model differentiated COPD patients with a sensitivity of 70% and a specificity of 90%, and ROC curve showed an area under curve (AUC) of 80.2%. In the COPD *vs*. smoking comparison, sensitivity was 90%, specificity 80%, with an AUC of 81,5%. In the COPD *vs*. non-smoking comparison, sensitivity was 80%, specificity of 90%, and AU of 82.5%.

Gene Ontology Annotation (GOA) and Pathway analysis

The frequencies of the GO annotations and protein classes obtained among DEPs were analyzed by PANTHER and illustrated in Figure 3. Considering all DEPs that differentiated COPD, within the molecular function (MF) category, binding and catalytic activity were the most and second-most frequent annotations. In the biological process (BP) category, cellular process and biological regulation were, in this sequency, the most observed. Most GO annotations for the cellular component (CC) category corresponded to cellular anatomical entities. The annotated protein classes (PC) were distributed equally among extracellular matrix proteins, gene-specific transcriptional regulators, and protein-modifying enzymes followed by transporters.

Some DEPs (n=22) were found annotated in PANTHER for biological pathways. The Reactome tool was also used to evaluate pathways related to DEPs, which were statistically important in the interaction of these proteins ($p < 0,05$). The pathways related with DEPs are listed in Figure 4. We found two pathways that were a common result from the assessment of both tools: *wnt* and integrin pathways.

DISCUSSION

In contrast to prior research, which primarily reported protein frequencies for individual diseases, this study identified a panel of DEPs with discriminatory ability to distinguish COPD from both controls and

asthma groups. Within this set of discriminative proteins, USP31, RNF10, and TACC2 demonstrated the capacity to differentiating COPD from control groups as well as asthma patients.

The USP31 protein hydrolyzes the peptide bond at the C-terminal glycine of ubiquitin and is involved in the processing of polyubiquitin precursors and ubiquitinated proteins[14]. Therefore, USP31 is a member of the deubiquitinase family.

Ubiquitination is a cellular process that consists of adding ubiquitin to the proteins to be degraded. Ligases (E1, E2 and E3) and deubiquitinases (DEUs) are responsible for controlling protein ubiquitination. Ubiquitin-specific proteases (USPs) are the main class of deubiquitinase. Both processes are important for regulating the activity and levels of proteins under physiological conditions. USP31 has been found in the nucleus and cytoplasm and acts by activating the nuclear factor NFκB [14], which is important for inflammation and cancer progression. Recent studies have shown that ubiquitination regulates not only protein degradation but also protein activity by modulating the interaction between proteins. These proteostasis imbalances in the pathogenesis of COPD and severe emphysema had already been measured in a study [15]. Authors found that NFκB activation, the accumulation of ubiquitinated proteins, and the deubiquitinating system (UCH-L1 enzyme) were associated with increased severity of emphysema in COPD patients. Thus, proteasome activity, in addition to exposure to toxic smoke particles, could be related to increased inflammation, apoptosis, and emphysema [16].

We found USP31 upregulated in the COPD group with elevated FC(Table 3) compared to asthma (FC = 2,065.4), smoking (FC = 7,209.7), and non-smoking (FC = 2,859.1) groups, corroborating that proteostasis imbalance described in COPD pathophysiology.

RNF10 was the second DEP that differentiated COPD from all other groups. This protein is involved in controlling the ubiquitination system specifically as an E3 ligase (E3-RING). This subtype facilitates substrate binding and regulates specificity [17], influencing the posttranscriptional regulation of inflammation, apoptosis and fibrosis genes [18].

The RNF10 protein is associated with the innate immune response and has been shown to play a suppressive role in inflammatory responses, exerting negative feedback on the production of proinflammatory cytokines, hyperproliferation [19], regulation of apoptosis[19],[20], cell differentiation [21], and neuronal development and morphology [22]. It also has been reported that RNF10 overexpression reduced TNF-α and NF-κB levels and the proliferation of the carotid intimal layer, after experimental mechanical balloon injury [20] and after hyperglycemia injury[19]. The mechanism may be related to the regulation of the cell cycle and anti-inflammatory and anti-hyperproliferation activity. We found RNF10 upregulated in our COPD patients probably as an effort to control the production of proinflammatory cytokines. To our knowledge, this is the first study linking RNF10 to COPD.

The third protein that differentiated the COPD group from the other three comparison groups was TACC2, a member of the TAAC family of acidic proteins with a coiled-coil domain that regulates microtubule homeostasis through interactions with the centrosome and mitotic spindle[23]. Because of

TACCs function and local action, researchers have explored their relationship with malignancies in various human cancers using a variety of approaches [24]. However, both the up- and downregulation of TACCs are observed in various types of human cancer, making their roles as oncogenes or tumor suppressors uncertain. What seems possible is that they may contribute directly or indirectly to tumorigenesis. TACC2 has been identified as a putative breast tumor suppressor[25], reversing a malignant phenotype to benign [26]. However, it has been associated with the risk of low-grade breast cancer [27], and TACC2 mRNA expression has been significantly correlated with shorter disease-free survival in patients with breast carcinoma [28]. Similarly, TACC2 expression is associated with worse prognosis in prostate cancer[29], infant acute lymphoblastic leukemia [30] and hepatocellular carcinoma [31].

In our study, the COPD group had TACC2 upregulated when compared with all other comparison groups. None of our COPD patients had any clinical suspicion of cancer. To date, only TACC3 has been evaluated in lung cancer, revealing a high level of TACC3 expression in 14.8% of cases in non-small cell lung cancer [32].

A relationship between TACC2 protein and the pathophysiology of COPD was previously reported [33]; suppression of TACC2 in bronchial epithelial cell cultures increased DNA damage, especially when exposed to cigarette smoke extract. On the other hand, overexpression of TACC2 attenuated these effects. This research also indicated that the degradation of TACC2 by cigarette smoke extract is regulated by the ubiquitin—proteasome system. In TACC2 knockout mice, exposure to cigarette smoke extract resulted in accumulated DNA damage, apoptosis of the alveolar epithelium and lung inflammation [33].

In our study, TACC2 was overexpressed in the serum of COPD patients compared with asthma patients and both smoking and non-smoking control groups. This represents the first serum-based analysis of this protein in this patient population.

MAP3K10 was among the top five DEPs listed in Table 3 and was upregulated in COPD patients compared with smoker and nonsmoker controls. Proteins of the MAP3Ks or MLKs (mixed lineage kinase) family activate receptor 1 of tumor necrosis factor (TNF r1) triggering noncanonical inflammatory pathways, including N-terminal c-Jun (JNK), p38 mitogen-activated kinase (p38 MAP kinase) and extracellular signal-regulated kinases (ERKs)[34]. The MAP3K10 and MAP3k4 proteins are sufficient to mediate p38 MAPK pathway activation[35] stimulating proinflammatory cytokine production (IL-6 and TNF-α) in COPD [36]. Consistently, PANTHER analysis highlighted p38 MAPK pathway as altered in our results. These findings align with previous studies implicating p38 MAPK cascade in COPD pathogenesis[37],[38],[39], and its consideration as a potential therapeutic target, like selective blockade of p38 MAPK [40],[41]. Murine experimental models further suggest a role for p38α/β MAPK in IL-33-induced ILC2 activation[42], supporting upstream inhibition as a mechanism of action form anti-inflammatory biologics, such as anti-IL-33 [43],[44]. These known may influence the finding of new molecules in the future to improve COPD strategies treatment with target therapies.

The fifth top-ranked DEP (Table 3), type V collagen alpha 1 chain (COL5A1) was also upregulated in COPD group compared with both control groups (smoking and non-smoking). Pathway analysis (PANTHER and Reactome) linked COL5A1 to Integrin signaling pathway. Some integrins bind to type 1 and 5 collagen, modulating formation and remodeling of the extracellular matrix, as well as contributing to tissue integrity [45],[46]. Other collagen-related pathways were significantly expressed in Reactome pathway analysis, as described on Fig. 4.

Normally sequestered in lung interstitium, type V collagen becomes exposed and overexpressed during remodeling. In a murine model after induction of pulmonary fibrosis, the expression of COL5A1 and antibodies against this protein were significantly higher compared to the normal lungs, and treatment with the same protein (COL5A1) suppressed systemic levels of anti-collagen V antibodies, IL-6 and TNF-α [47]. In COPD, increased formation of type V collagen as well as fragments of types I and IV collagen have been observed, especially during exacerbations [48]. These results are consistent with the findings of our study, which also identified upregulation of fragments type V collagen in COPD patients.

Fibroblast proliferation and increased collagen production were related to activation of the ERK1/2 pathways (MAPKs) [49], a positive feedback loop between inflammation and remodeling. Additionally, eukaryotic translation initiation factor 3a (EIF3a), which was upregulated in our COPD group, modulates the ERK (MAPK) pathway, signaling and enhancing global protein synthesis including proliferation, such as c-Myc [50]. All these proteins cited before, the MAPK pathway protein (MAP3K10), collagen V (COL5A1), the EIF3a and l-Myc were overexpressed in our COPD group compared to controls, suggesting an interaction between these proteins and cell proliferation, fibrosis and lung remodeling, activated by inflammation.

As intended, the upregulated l-Myc protein in the COPD group was responsible for the *wnt* signaling pathway being found to be related, both by PANTHER and Reactome tools, in our study. In general, canonical *wnt* pathway is negatively regulated, while noncanonical signaling is positively regulated in COPD [51].

Although the shotgun proteomic approach identified a wide range of proteins, the fact that several differentially expressed ones were already implicated in COPD pathophysiology lends greater robustness to our findings. This reinforces current knowledge of altered pathways, consolidates the understanding of disease mechanisms, and highlights opportunities for more targeted and personalized therapeutic strategies.

On the other hand, some limitations of the present study include the relatively small sample size, typical of exploratory proteomics studies, limit the generalizability of our findings and should be addressed in future investigations with larger cohorts. In addition, our results have not yet been validated using complementary direct quantification methods, such as immunological tests, which are necessary to confirm the true magnitude of the observed differences. Importantly, the target selection of a protein panel for statistical testing and logistic regression with five-fold cross validation ensured the statistical rigor of the predictive model and improved the robustness of our results.

CONCLUSION

Proteomics has demonstrated its effectiveness in identifying low-abundance proteins that distinguish COPD patients from those with asthma and healthy controls (smokers and non-smoking). The main differentially expressed proteins (DEPs) identified are directly linked to the pathophysiology of COPD. Pathway analyses using PANTHER and Reactome further validated associations between key biological altered pathways and COPD.

Since a panel of protein biomarkers is more effective than individual protein to distinguish well defined patient groups, we first identified proteins suitable for inclusion in such a biomarker panel. Using learning machine, our panel was capable to diagnose COPD patients with an accuracy over 80% in comparison with asthma patients and both smoking and non-smoking control groups. The next study phase is to perform immunological testing and validation in a larger COPD population utilizing our selected protein set, with the aim of developing a biosensor to enhance diagnostic accuracy in COPD patients.

The investigation of differential proteins and pathways provides valuable insights into the development of targeted therapies that address the specific molecular mechanisms underlying this complex disease. In this way, this approach, which combines proteomic analysis and pathway exploration, not only could improve our diagnostic capabilities but also has potential to advance precision medicine in COPD treatment strategies.

MATERIALS AND METHODS

Patients

Patients were recruited during follow-up visits to the pulmonology outpatient clinic at the Clinical Hospital of Uberlandia Federal University (COPD group, n = 10 and asthma group, n = 10). Age- and sex-matched healthy control subjects with a negative history of pulmonary or systemic disease were recruited. Ten healthy participants in the control group were never smokers (non-smoking, n = 10), and the other 10 participants were smokers or former smokers (smoking, n = 10). Patients had already been diagnosed with asthma or COPD by the pulmonologist team, following GINA or GOLD reports criteria. In addition to COPD or asthma diagnosis, the inclusion criteria were age over 18 years and, for the asthma group, having never smoked. The general exclusion criteria were any serious systemic disease, such as heart failure, chronic renal dysfunction, liver dysfunction, diabetes mellitus, other pulmonary disease and current or previous malignancy diagnosis. To ensure the exclusion criteria, patients and control individuals underwent blood tests, echocardiograms and chest X-rays, in addition to evaluating their clinical records.

All study participants underwent spirometry before and after bronchodilator administration, according to American Thoracic Society (ATS) standards[52], and blood samples were collected for proteomics analyses. Blood was drawn in the morning after 12 h of fasting in a vacuum dry tube. Serum was

obtained by immediate centrifugation at 3000 rpm for 10 min and stored at -80°C. Later, a 100 µl sample aliquot was stored in an Eppendorf tube until analysis.

This study was approved by Uberlandia Federal University Human Research Ethics Committee (protocol number 61266816.1.0000.5152) and followed the principles of the Helsinki Declaration. All participants signed an informed consent form.

PROTEOMIC ANALYSIS

Depletion of Abundant Serum Proteins

Abundant serum proteins were depleted by High-Performance Liquid Chromatography (HPLC) using Multi Affinity Removal Column Human 14-4.6x100mm (*Amershan*) according to the manufacturer's instructions. The material elutes was pooled and subsequently desalted using centrifugation filters and 5000Da cut-off (*Agilant Technologies*), according to the manufacturer's instructions.

Enzymatic Hydrolysis of Protein by Trypsin

The protein concentrations of samples were quantified using a Pierce™ BCA Protein Assay Kit (*Thermo Fisher Scientific Inc. Waltham, MA, USA)* according to the manufacturer's instructions. For in-solution protein digestion, 50µg of proteins was treated with 1% RapiGest SF (*Waters, MA*, USA), 0.5 M dithiotrheitol (*Sigma –Aldrich*) for reduction, and 0.5 M iodoacetamide (*Sigma –Aldrich*) for alkylation. Trypsin (*Promega)* was added (20 ng/µL), and the samples were digested overnight at 37°C. The peptide mixture was desalted and concentrated using a pipette tip Omix (*Agilent, CA, USA*) with a C18 stationary phase inside. Eluted peptides were dried in a vacuum concentrator speedVac (Thermo Scientific™, MA, USA) and then solubilized in 0.1% trifluoracetic acid. Samples were analyzed using LC—MS/MS.

Liquid Chromatography with mass spectrometry (LC—MS/MS)

Mass spectrometry analyses were performed using a liquid chromatograph (*Agilent Infinity 1260*) coupled to a high-resolution mass spectrometer with an electrospray ionization source *(Agilant*, *USA-6520B Q-TOF*). This device is allocated at the Nanobiotechnology Laboratory of Uberlandia Federal University. Chromatographic parameters were as follows: 1) column AdvanceBio Peptide Mapping (*Agilent, CA, USA*) – 2,1 mm internal diameter, 10 cm length and 2,7 µm particle size 2) mobile phase constituents - water (A) and acetonitrile (B), both acidified with formic acid (0,1% vv-1). The following B phase gradient parameters were used: 2% B (0 min), 2% B (10 min), 15% B (40 min), 50% B (150 min), 70% B (200 min), 98% B (220 min), 98% B (300 min), 100% B (301 min) and 100% B (400 min) in 400 µL/min flow. The ionization patterns were as follows: 45 psi nebulizer pressure, 8 L/min drying gas flow, 325°C temperature, and 4 kV of energy applied to the capillary.

Protein Identification

Mass data were obtained through LC–MS/MS analyses, considering high resolution mass (mass error less than 10 ppm) and mass spectrum (MS/MS). The results were matched with the database program Spectrum Mil MS Proteomics Workbench (*Agilent, CA, USA*) for protein identification using the UniProt database, accessed online (https://www.uniprot.org/). The false discovery rate (FDR) was set to 1.2%.

Statistical analysis

The data identified by the Spectrum Mill MS program was exported and submitted to the Mass Hunter-Mass Profiler Professional v. B.13.1.1 (MPP) software for statistical analysis. The MPP software was used to filter and align the extracted molecular compounds. The filters used: minimum absolute abundance = 5000 counts; minimum number of ions allowed = 2; all charges allowed. The alignment parameters: retention time tolerance of 0.15 min; mass window 15 ppm + 2mDa. Proteins present in 100% of at least one group were considered. Statistical analysis of the data generated was carried out using log2 transformed values. The unpaired T-test was applied to pairwise analyses with a p-value < 0.05 and a fold change equal to or greater than 2.00. DEP's were used to compare COPD patients with all other groups (smoking and non-smoking controls, and asthma patients).

Clinical and demographic differences between groups were analyzed using one-way analysis of variance (ANOVA), Tukey's test and Kruskal-Wallis test for multiple comparisons, and Mann–Whitney U test, Student's t test or Fisher's exact test. Shapiro–Wilk and Levene tests were used to analyze the normality of sampling errors and the homogeneity of variances for choosing a parametric or nonparametric test. Statistical significance was defined as $p < 0.05$. The software used for statistical analyses was SPSS *v.* 23 (IBM SPSS® Statistics, USA) and GraphPad Prism 7.0 (GraphPad Software, USA).

PCA models

After Mass Hunter Profiler (MPP- v. B. 13.1.1) was used to analyze data, the intensities of the main significant proteins for each comparison were subjected to multivariate analysis using the Orange 3.34.1 software to create unsupervised models (PCA) and supervised models - Logistic Regression - to classify the groups. Sensitivity, specificity and accuracy calculations were performed using the MEDCALC online platform (https://www.medcalc.org).

Gene Ontology Annotation (GOA)

Gene Ontology considers how a gene or its products, such as proteins, can participate in an organism biological domain. Three aspects are used for functional classification: biological process (BP), cellular component (CC)and molecular function (MF). UniProt® Knowledge Base Annotation (https://www.uniprot.org/uniprotkb)- (last access on 2023/October) was used for collection of functional information about proteins and provided statistical information about differentially expressed proteins (DEPs). Analyses of the Clusters of Orthologous Groups of proteins (prokaryotic homologous protein clusters/eukaryotic homologous protein clusters, COG/KOG) functional classification statistics for DEPs were conducted through database comparisons.

Enrichment Analysis of PANTHER and Reactome Knowledgebase

The Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification system (https://www.pantherdb.org) combines gene function, ontology, pathways and statistical analysis tools. The enrichment test tool determines whether the numerical values of the genes associated with a particular ontology class or pathway were drawn randomly from the overall distribution of values [53]. The enrichment analysis was performed using GO for differentially expressed proteins (DEPs) between comparison groups using panther. Pathway enrichment analysis was performed using Reactome Knowledgebase *v.* 86 on 11/24/2023 (https://reactome.org/), and the *p* value was corrected for FDR (false discovery rate) with the Benjamini–Hochberg method. The core unit of the Reactome data model is the reaction. Entities (nucleic acids, proteins, complexes, vaccines, anticancer therapeutics and small molecules) participating in reactions form a network of biological interactions and are grouped into pathways. Examples of biological pathways in Reactome include classical intermediary metabolism, signaling, transcriptional regulation, apoptosis and disease [54].

Declarations

Data availability statement

The proteomics datasets generated during the current study are deposited in the PRIDE repository (https://www.ebi.ac.uk/pride/login) under accession number PXD068041 (Reviewer Token: LZ1HRMUxzcDW). Alternatively, reviewers can access the dataset by logging in to the PRIDE website using the following account details:

Username: reviewer_pxd068041@ebi.ac.uk / Password: gytt6Pe3H5q5

Data are available from the corresponding author upon reasonable request.

Acknowledgements

We would like to thank Dr Luiz Ricardo Goulart Filho (*in memoriam*) and Nanobiotechnology Laboratory from the Federal University of Uberlandia/ Brazil for performing LC-MS/ MS proteomic experiments.

Author contributions

T.M.L.C.B collected the patient data. T.M.L.C.B, T.M.C conceived study design, T.M.L.C.B, B.U.P**,** H.O.A.S, M.M.M contributed to performing experiments. T.M.L.C.B, T.M.C, B.U.P**,** H.O.A.S**,** M.M.M, L.O.S contributed to data analysis. T.M.L.C.B, T.M.C, R.S.S, D.S contributed to drafting and reviewing manuscript intellectual content. All authors reviewed the manuscript.

Competing interests and Funding

The authors declare no competing interests. Chiesi Foundation Onlus and National Institute of Science and Technology in Theranostics and Biotechnology - INCT-TeraNano (CNPq Process number

403193/2022-2) provided grant support for equipment and supplies.

References

1. Naghavi, M. et al. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **403**, 2100–2132 (2024).
2. Marcon, A. et al. The coexistence of asthma and COPD: risk factors, clinical history and lung function trajectories. *Eur. Respir J.* **58**, 2004656 (2021).
3. Yawn, B. P. & Han, M. K. Practical Considerations for the Diagnosis and Management of Asthma in Older Adults. *Mayo Clin. Proc.* 92, 1697–1705 (2017).
4. Global Initiative For Chronic Obstructive Lung Disease (GOLD). Global strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease (2025 REPORT) (2025). https://goldcopd.org/2025-gold-report/(
5. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention. (2025). https://ginasthma.org/2025-gina-strategy-report/(
6. Regard, L., Roche, N. & Burgel, P. R. The Ongoing Quest for Predictive Biomarkers in Chronic Obstructive Pulmonary Disease. *Am. J. Respir Crit. Care Med.* **208**, 511–513 (2023).
7. Lange, P., Ahmed, E., Lahmar, Z. M., Martinez, F. J. & Bourdin, A. Natural history and mechanisms of COPD. *Respirology* **26**, 298–321 (2021).
8. Franssen, F. M. et al. Personalized medicine for patients with COPD: where are we? *Int. J. Chron. Obstruct Pulmon Dis.* **14**, 1465–1484 (2019).
9. Regan, E. A. et al. Omics and the Search for Blood Biomarkers in Chronic Obstructive Pulmonary Disease. Insights from COPDGene. *Am. J. Respir Cell. Mol. Biol.* **61**, 143–149 (2019).
10. Rossi, R. et al. Biomarker discovery in asthma and COPD by proteomic approaches. *Proteom. – Clin. Appl.* **8**, 901–915 (2014).
11. Li, R. et al. Differential Expression of Serum Proteins in Chronic Obstructive Pulmonary Disease Assessed Using Label-Free Proteomics and Bioinformatics Analyses. *Int. J. Chron. Obstruct Pulmon Dis.* **17**, 2871–2891 (2022).
12. Zhang, Z. et al. Proteomics and metabolomics profiling reveal panels of circulating diagnostic biomarkers and molecular subtypes in stable COPD. *Respir Res.* **24**, 73 (2023).
13. Bracht, T. et al. Plasma Proteomics Enable Differentiation of Lung Adenocarcinoma from Chronic Obstructive Pulmonary Disease (COPD). *Int. J. Mol. Sci.* **23**, 11242 (2022).
14. Tzimas, C. et al. Human ubiquitin specific protease 31 is a deubiquitinating enzyme implicated in activation of nuclear factor-κB. *Cell. Signal.* **18**, 83–92 (2006).
15. Min, T., Bodas, M., Mazur, S. & Vij, N. Critical role of proteostasis-imbalance in pathogenesis of COPD and severe emphysema. *J. Mol. Med.* **89**, 577–593 (2011).

16. Young, M. J., Hsu, K. C., Lin, T. E., Chang, W. C. & Hung, J. J. The role of ubiquitin-specific peptidases in cancer progression. *J. Biomed. Sci.* **26**, 42 (2019).
17. Woodsmith, J., Jenn, R. C. & Sanderson, C. M. Systematic Analysis of Dimeric E3-RING Interactions Reveals Increased Combinatorial Complexity in Human Ubiquitination Networks. *Mol. Cell. Proteomics* 11, M111.016162-1-M111.016162-11 (2012).
18. Shi, H. et al. RING Finger Protein 10 Regulates AP-1/Meox2 to Mediate Pirarubicin-Induced Cardiomyocyte Apoptosis. *Oxid. Med. Cell. Longev.* 1–11 (2023). (2023).
19. Li, S. et al. RING finger protein10 is a potential drug target for diabetic vascular complications. *Mol. Med. Rep.* https://doi.org/10.3892/mmr.2019.10358 (2019).
20. Yu, G. et al. RING finger protein 10 prevents neointimal hyperplasia by promoting apoptosis in vitro and in vivo. *Life Sci.* **208**, 325–332 (2018).
21. Huang, P. et al. Putative regulators for the continuum of erythroid differentiation revealed by single-cell transcriptome of human BM and UCB cells. *Proc. Natl. Acad. Sci.* 117, 12868–12876 (2020).
22. Carrano, N. et al. The Synaptonuclear Messenger RNF10 Acts as an Architect of Neuronal Morphology. *Mol. Neurobiol.* **56**, 7583–7593 (2019).
23. Gergely, F., Centrosomal & TACCtics *BioEssays* 24, 915–925 (2002).
24. Ha, G. H., Kim, J. L. & Breuer, E. K. Y. Transforming acidic coiled-coil proteins (TACCs) in human cancer. *Cancer Lett.* **336**, 24–33 (2013).
25. Chen, H. M. et al. AZU-1: A Candidate Breast Tumor Suppressor and Biomarker for Tumor Progression. *Mol. Biol. Cell.* **11**, 1357–1367 (2000).
26. Lauffart, B., Gangisetty, O. & Still, I. H. Molecular cloning, genomic structure and interactions of the putative breast tumor suppressor TACC2. *Genomics* 81, 192–201 (2003).
27. Purrington, K. S. et al. Genetic variation in mitotic regulatory pathway genes is associated with breast tumor grade. *Hum. Mol. Genet.* **23**, 6034–6046 (2014).
28. Cheng, S., Douglas-Jones, A., Yang, X., Mansel, R. E. & Jiang, W. G. Transforming acidic coiled-coil-containing protein 2 (TACC2) in human breast cancer, expression pattern and clinical/prognostic relevance. *Cancer Genomics Proteom.* **7**, 67–73 (2010).
29. Takayama, K. et al. TACC2 is an androgen-responsive cell cycle regulator promoting androgen-mediated and castration-resistant growth of prostate cancer. *Mol. Endocrinol. Baltim. Md.* **26**, 748–761 (2012).
30. Kang, H. et al. Gene expression profiles predictive of outcome and age in infant acute lymphoblastic leukemia: a Children's Oncology Group study. *Blood* **119**, 1872–1881 (2012).
31. Shakya, M. et al. High expression of TACC2 in hepatocellular carcinoma is associated with poor prognosis. *Cancer Biomark.* **22**, 611–619 (2018).
32. Jung, C. K. et al. Expression of transforming acidic coiled-coil containing protein 3 is a novel independent prognostic marker in non-small cell lung cancer. *Pathology International* 56, *503–509 ()*. (2006).

33. Mallampalli, R. K. et al. Cigarette smoke exposure enhances transforming acidic coiled-coil-containing protein 2 turnover and thereby promotes emphysema. *JCI Insight* 5, e125895, 125895 (2020).
34. Kant, S. et al. TNF-stimulated MAP kinase activation mediated by a Rho family GTPase signaling pathway. *Genes Dev.* **25**, 2069–2078 (2011).
35. Sapkota, G. P. The TGFβ-induced phosphorylation and activation of p38 mitogen-activated protein kinase is mediated by MAP3K4 and MAP3K10 but not TAK1. *Open. Biol.* **3**, 130067 (2013).
36. Hu, T. et al. The Activation of M1 Macrophages is Associated with the JNK-m6A-p38 Axis in Chronic Obstructive Pulmonary Disease. *Int. J. Chron. Obstruct Pulmon Dis.* **18**, 2195–2206 (2023).
37. Zhu, L. et al. The antioxidant N-acetylcysteine promotes immune response and inhibits epithelial-mesenchymal transition to alleviate pulmonary fibrosis in chronic obstructive pulmonary disease by suppressing the VWF/p38 MAPK axis. *Mol. Med.* **27**, 97 (2021).
38. Bhavsar, P., Khorasani, N., Baker, J., Johnson, M. & Chung, K. F. Reversal of corticosteroid insensitivity by p38 MAPK inhibition in peripheral blood mononuclear cells from COPD. *Int J. Chron. Obstruct Pulmon Dis* 283 (2015).
39. Gaffey, K., Reynolds, S., Plumb, J., Kaur, M. & Singh, D. Increased phosphorylated p38 mitogen-activated protein kinase in COPD lungs. *Eur. Respir J.* **42**, 28–41 (2013).
40. Barnes, P. J. Kinases as Novel Therapeutic Targets in Asthma and Chronic Obstructive Pulmonary Disease. *Pharmacol. Rev.* **68**, 788–815 (2016).
41. Pelaia, C. et al. Role of p38 Mitogen-Activated Protein Kinase in Asthma and COPD: Pathogenic Aspects and Potential Targeted Therapies. *Drug Des. Devel Ther.* **15**, 1275–1284 (2021).
42. Petrova, T., Pesic, J., Pardali, K., Gaestel, M. & Arthur, J. S. C. p38 MAPK signalling regulates cytokine production in IL-33 stimulated Type 2 Innate Lymphoid cells. *Sci. Rep.* **10**, 3479 (2020).
43. Wang, Y. et al. P38 MAPK/AKT signalling is involved in IL-33-mediated anti-apoptosis in childhood acute lymphoblastic leukaemia blast cells. *Ann. Med.* **53**, 1464–1472 (2021).
44. Calderon, A. A. et al. Targeting interleukin-33 and thymic stromal lymphopoietin pathways for novel pulmonary therapeutics in asthma and COPD. *Eur. Respir Rev. Off J. Eur. Respir Soc.* **32**, 220144 (2023).
45. Agarwal, M. et al. Type I Collagen Signaling Regulates Opposing Fibrotic Pathways through α2β1 Integrin. *Am. J. Respir Cell. Mol. Biol.* **63**, 613–622 (2020).
46. Zoppi, N., Gardella, R., De Paepe, A., Barlati, S. & Colombi, M. Human Fibroblasts with Mutations in COL5A1 and COL3A1 Genes Do Not Organize Collagens and Fibronectin in the Extracellular Matrix, Down-regulate α2β1 Integrin, and Recruit αvβ3 Instead of α5β1 Integrin. *J. Biol. Chem.* **279**, 18157–18168 (2004).
47. Vittal, R. et al. Type V Collagen Induced Tolerance Suppresses Collagen Deposition, TGF-β and Associated Transcripts in Pulmonary Fibrosis. *PLoS ONE*. **8**, e76451 (2013).

48. Schumann, D. M. et al. Collagen Degradation and Formation Are Elevated in Exacerbated COPD Compared With Stable Disease. *Chest* **154**, 798–807 (2018).
49. Xu, T. R. et al. Eukaryotic Translation Initiation Factor 3, Subunit a, Regulates the Extracellular Signal-Regulated Kinase Pathway. *Mol. Cell. Biol.* **32**, 88–95 (2012).
50. Zhang, L., Pan, X. & Hershey, J. W. B. Individual Overexpression of Five Subunits of Human Translation Initiation Factor eIF3 Promotes Malignant Transformation of Immortal Fibroblast *Cells*. *J. Biol. Chem.* **282**, 5790–5800 (2007).
51. Qu, J., Yue, L., Gao, J. & Yao, H. Perspectives on Wnt Signal Pathway in the Pathogenesis and Therapeutics of Chronic Obstructive Pulmonary Disease. *J. Pharmacol. Exp. Ther.* **369**, 473–480 (2019).
52. Miller, M. R. et al. Standardisation of spirometry. *Eur. Respir J.* **26**, 319–338 (2005).
53. Mi, H. et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **45**, D183–D189 (2017).
54. Croft, D. et al. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**, D691–D697 (2011).

Tables

Tables 1 to 3 are available in the Supplementary Files section.

Figures

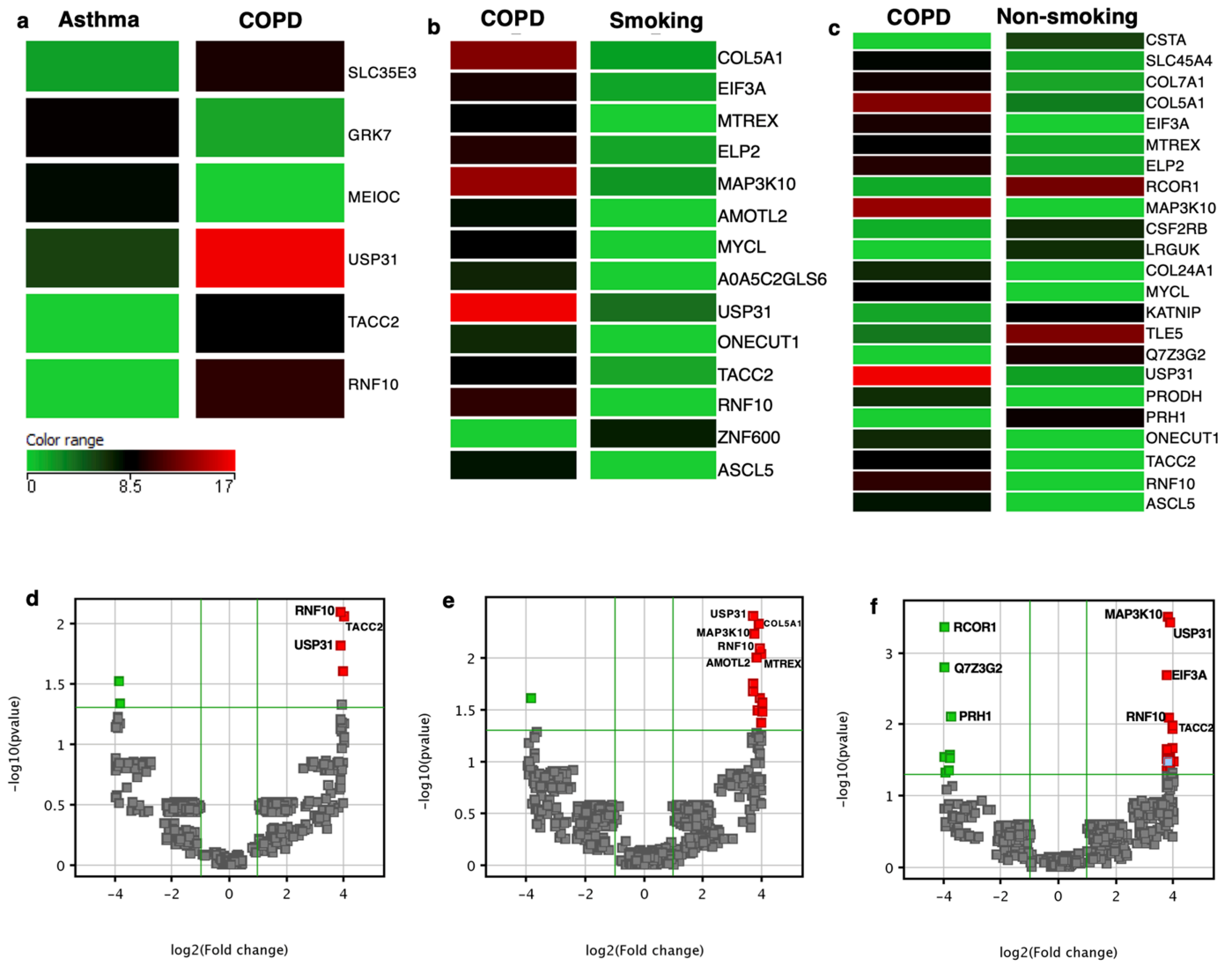


Figure 1

Differentially expressed proteins for differentiate the COPD group.

Heatmaps showing the expression pattern of differentially expressed proteins between a) COPD x asthma, b) COPD x smoking, and c) COPD x non-smoking. DEPs are named with gene code (except IG c279_heavy and Uncharacterized protein that are named by uniProtKB accession - A0A5C2GLS6 and Q7Z3G2 respectively). Volcano Plots for same comparisons d) COPD x asthma, e) COPD x smoking, and f) COPD x non-smoking, representing the results of statistical analysis using Student's *t* test, after proteins are filter by frequency with cutt-off 100% (p value <0.05 and FC≥ 2.0). X axis corresponds to log2 FC and Y axis corresponds to −log10 *p* value. The proteins with −log10 p value around 2 or above (*p* value ≤ 0,01) are highlighted. Green indicates proteins with significant downregulation and red indicates those with significant upregulation. The color scale on legend shows intermediated values.

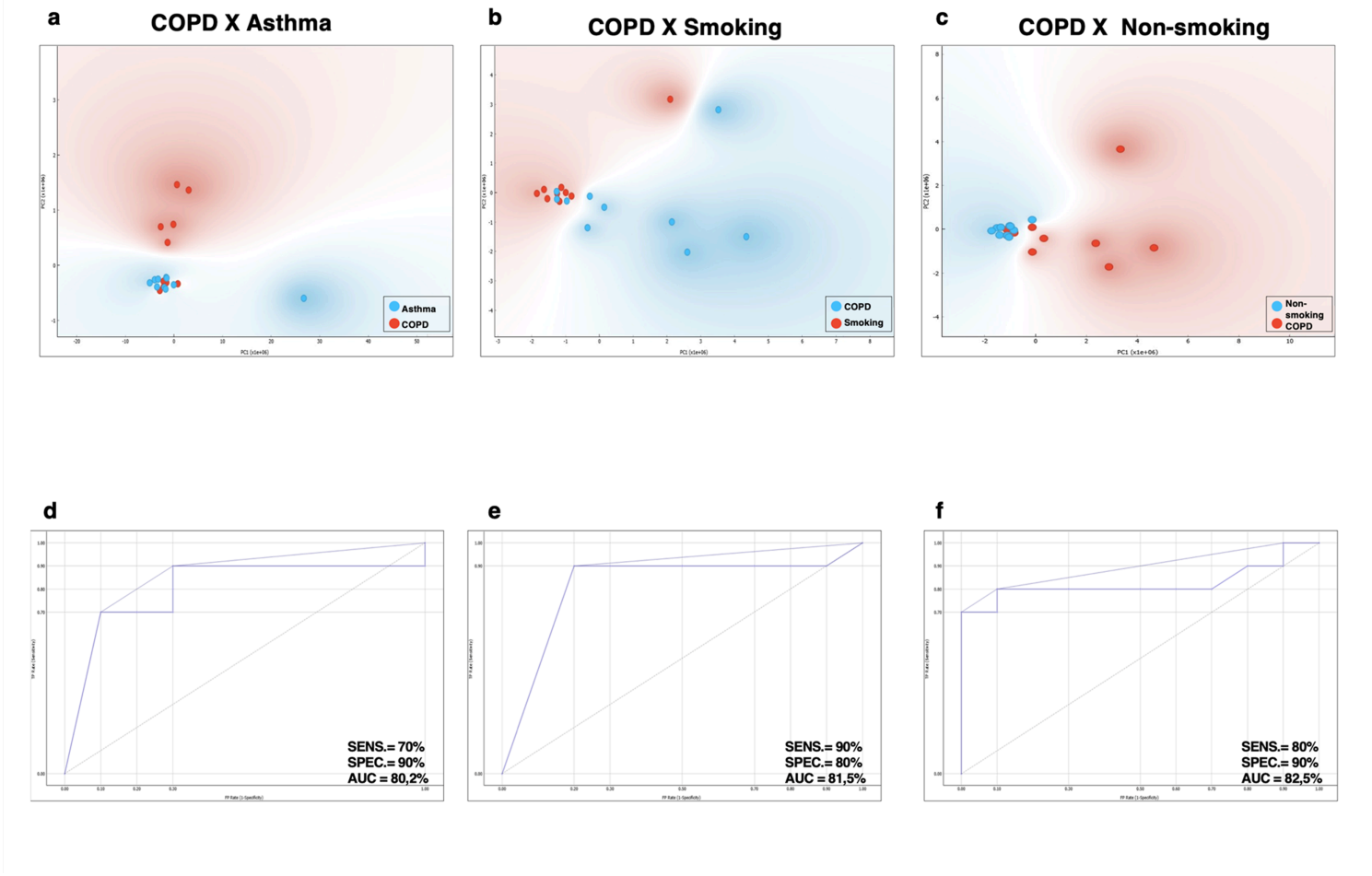


Figure 2

Principal component analysis (PCA) and a supervised logistic regression model to diagnose COPD patients. PCA model graphics to compare the COPD group with asthma (a), smoking (b) and non-smoking (c) groups, showing explained variance model of 97% (2PCs), 93% (3PCs) and 97% (2PCs) respectively. The expressions of the five best-performing proteins listed in Table 3 were used to compare COPD *vs*. smoking and CPOD *vs*. non-smoking. The expression of three best-performing proteins (USP31, RNF10 and TACC2) were used to compare COPD *vs*. asthma. Receiver operating characteristic (ROC) curves for tested logistic regression model (cross validation training with 5fold) were constructed for each comparison group. For COPD *vs*. asthma (d), the model showed 70% of sensitivity and 90% of specificity (AUC = 80,2%). For COPD *vs*. smoking (e), the sensitivity was 90% and specificity 80% (AUC= 81,5%), and COPC*vs*. non-smoking (f) the sensitivity was 80% and specificity was 90% (AUC=82,5%).

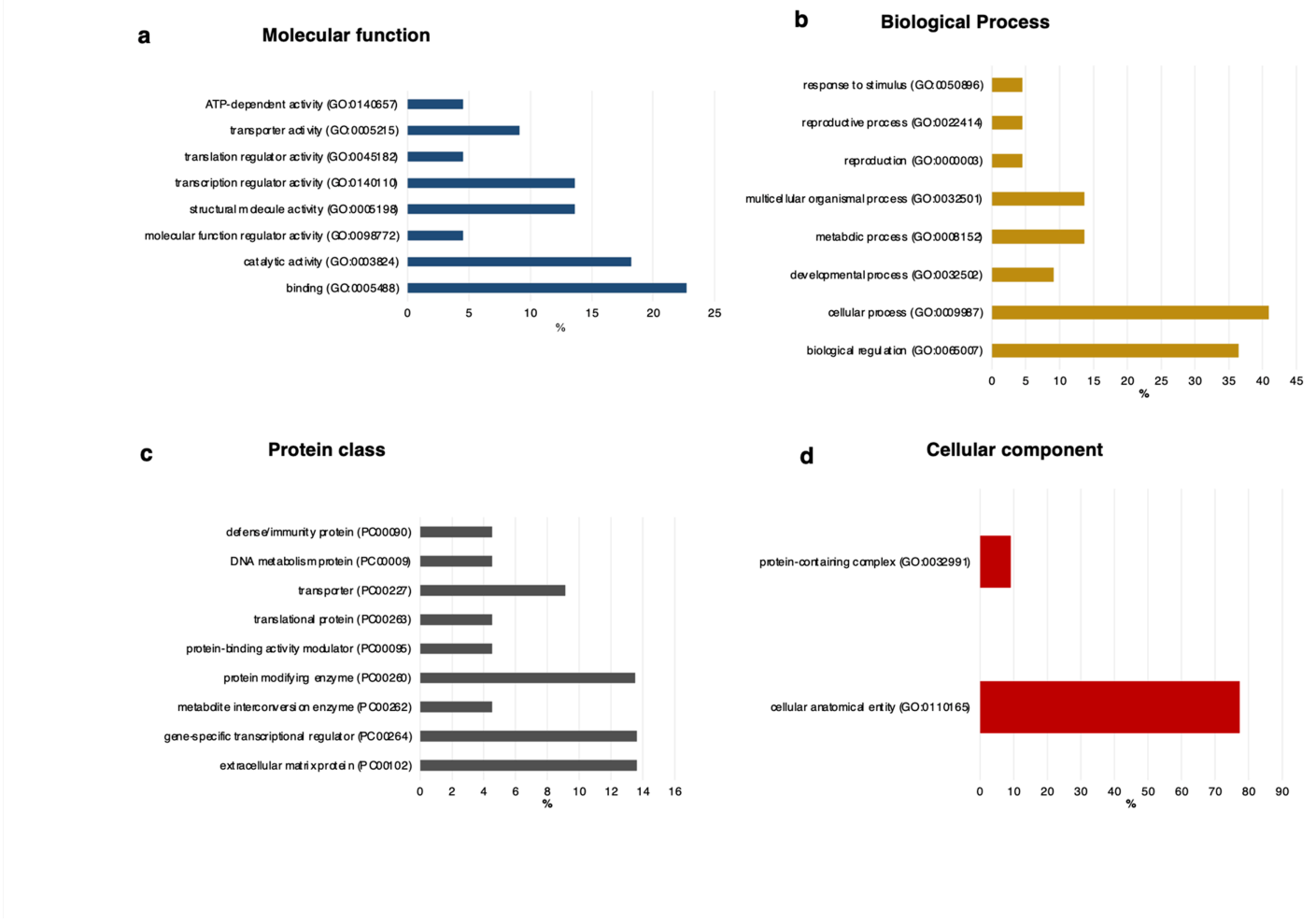


Figure 3

Gene Ontology functional annotation of DEPs. Frequency (%) of matched DEPs in each GO terms for molecular function (a), biological process (b), protein class (c) and cellular component (d), using PANTHER tool.

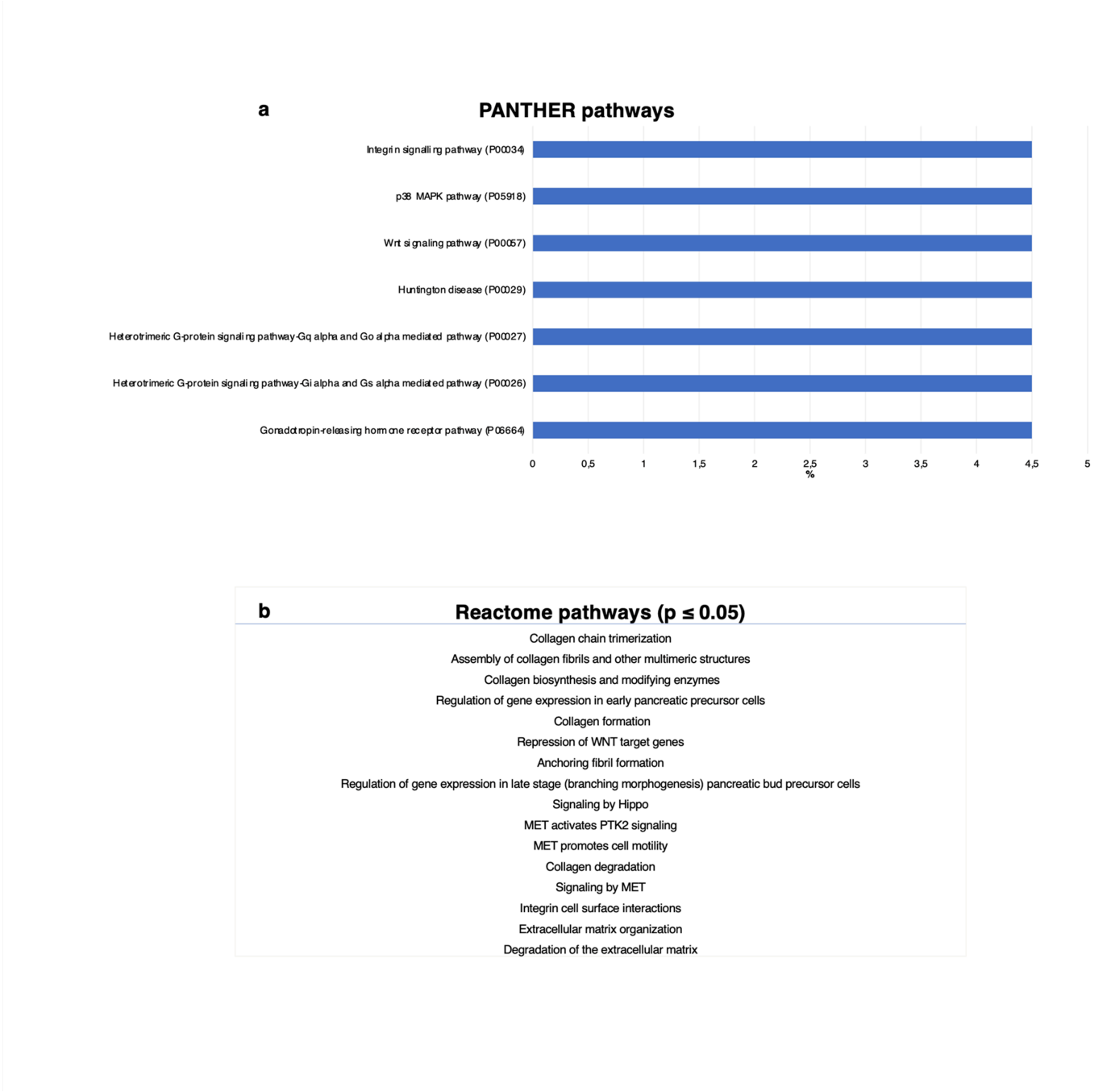


Figure 4

Pathways annotation of DEPs. Frequency (%) of matched DEPs in PANTHER pathways (a). Pathway functional enrichment of DEPs using Reactome tool (b), showing significantly pathways ($p < 0,05$) corrected for FDR (false discovery rate) with the Benjamini–Hochberg method.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.docx

- Table2.docx
- Table3.docx