Identification of Novel Diagnostic Biomarkers for Asthma and Chronic Obstructive Pulmonary Disease

Nicole M. Verrills[1,4,7†], Jennifer A. Irwin[2,4,5,7†], Xiao Yan He[2,4,5,7], Lisa G. Wood[1,4,5], Heather Powell[4,5], Jodie L Simpson[2,4,5], Vanessa M. McDonald[2,3,4,5], Alistair Sim[1,4] and Peter G. Gibson[2,4,5,6,7].

[1]School of Biomedical Sciences and Pharmacy, [2]School of Medicine and Public Health and [3]School of Nursing and Midwifery, Faculty of Health, University of Newcastle, NSW, Australia; [4]Hunter Medical Research Institute, Newcastle, NSW, Australia; [5]Department of Respiratory and Sleep Medicine, John Hunter Hospital, New Lambton, NSW, Australia; [6]Woolcock Institute of Medical Research, Sydney NSW, Australia; [7]CRC for Asthma and Airways, Australia.

[†]These authors have made an equal contribution to this manuscript.

Corresponding author: Prof. Peter G. Gibson, Department of Respiratory and Sleep Medicine, John Hunter Hospital, Lookout Rd, New Lambton 2305, Australia Ph: +61 2 49213470; Fax: +61 2 49855850; Email: Peter.Gibson@hnehealth.nsw.gov.au

This work was supported by the CRC for Asthma and Airways.

Short running heading: Diagnostic Biomarkers of Respiratory Disease.

Descriptor number: 3.06

Word Count (body): 4013

CONTRIBUTORSHIP

Nicole M. Verrills: 2D-DIGE study design, spot analysis and interpretation.

Jennifer A. Irwin: Data acquisition, design, analysis and interpretation for 2D-DIGE and two validation studies; writing and compilation of the manuscript.

Xiao Yan He: Interpretation of data and revision of manuscript.

Lisa G. Wood: Involvement in study conception, coordination of sample collection and the acquisition/compilation of clinical data for the discovery and primary validation population.

Heather Powell: Logistic regression analysis and statistical analysis of all clinical data.

Jodie L Simpson: Coordination of sample collection and the acquisition/compilation of clinical data for the secondary validation population.

Vanessa M. McDonald: Coordination of sample collection and the acquisition/compilation of clinical data for the secondary validation population.

Alistair Sim: Project conception.

Peter G. Gibson: Project conception, study design, data interpretation and writing/revision of manuscript.

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

The respiratory diseases asthma and chronic obstructive pulmonary disease (COPD) have many similar symptoms and are typically diagnosed by performing a number of clinical tests to assess an individual's lung function and response to reliever medication. To date, some studies have attempted to identify biomarkers of COPD or asthma, however, no study has attempted to identify non-invasive, blood-based, diagnostic biomarkers that can discriminate between healthy controls, asthmatics and individuals with COPD.

What This Study Adds to the Field

Using a proteomics approach, we have identified a panel of four blood-based biomarkers that, when used in combination, can discriminate between healthy controls, asthmatics and individuals with COPD and has the potential to be a valuable tool in the clinical diagnosis of respiratory disease. These markers implicate the anti-inflammatory iron metabolism pathways in the pathogenesis of asthma and COPD.

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org.

ABSTRACT

Rationale: Proteomics may identify a useful panel of biomarkers for identification of asthma and chronic obstructive pulmonary disease (COPD).

Objectives: To conduct an unsupervised analysis of peripheral blood proteins in well characterised subjects with asthma and COPD, and identify and validate a biomarker panel for disease discrimination.

***Methods*:** Two-dimensional difference gel electrophoresis (2D-DIGE) was used to separate plasma proteins from healthy controls, stable asthmatics and individuals with COPD. Candidate protein markers were identified using matrix assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF MS) and subsequently validated in two populations via immunoassay. A panel of four biomarkers was selected and their ability to distinguish between groups assessed in isolation and in combination in 2 separate validation populations.

Results: Seventy-two protein spots displayed significantly different expression levels between the three subject groupings ($p<0.05$). 58 were positively identified representing 20 unique proteins. A panel of four biomarkers (α-2 macroglobulin, haptoglobin, ceruloplasmin and hemopexin) was able to discriminate with statistical significance between the clinical groups of asthma, COPD, and controls, and these results were confirmed in a second clinical population of older adults with airflow obstruction.

Conclusions: Proteomics has identified novel biomarkers for asthma and COPD, and shown that the iron metabolism pathways and acute phase response may be involved in the pathogenesis of airway disease. The panel of peripheral blood biomarkers has the potential to become an extremely useful addition to the clinical diagnosis and management of respiratory disease.

Word count: 239

Key words: Plasma proteomics, biomarkers, asthma, COPD, logistic regression.

INTRODUCTION

Word count: 434

The obstructive airway diseases asthma and chronic obstructive pulmonary disease (COPD) are significant and increasing health problems throughout the world. The management of these diseases could be improved by better diagnosis and recognition, and better understanding of their pathogenesis. New technologies for investigating human diseases now offer significant potential to address the need for better diagnosis and improved understanding of asthma and COPD. Proteomics can simultaneously identify multiple proteins associated with different disease states (1) and potentially discover novel proteins not previously associated with particular disease states. Several studies have utilised proteomics for the discovery of protein changes in lung tissue from mouse models of asthma (2-6), and bronchoalveolar lavage fluid (BALF) (7) and $CD3^+$ T-lymphocytes (8) from asthmatic and healthy subjects. However, for diagnostic and prognostic purposes, identification of biomarkers in readily obtainable samples, such as blood, is preferred. A selected analysis of multiple known blood markers in COPD using protein array methodology has shown the potential for proteomics (9), whereby a panel of biomarkers were found to associate with COPD patients versus healthy controls. This type of study used a predefined array of markers, and is potentially limited by the pool of analytes available on the protein array and by our current knowledge of disease pathology. An open, unsupervised proteomic study should identify a broader panel of candidate markers, including proteins not previously associated with respiratory disease. The need for such unbiased approaches has recently been emphasised and called for in the chronic inflammatory diseases of asthma and COPD (10).

There is also general agreement that a panel of independent disease-related proteins considered in aggregate should be less prone to the influence of genetic and environmental 'noise' than is the level of a single marker protein (11), and that proteomics has the power to

identify such panels of proteins in a high-throughput manner. There remains however a need to apply methods to quantify the added benefit of biomarker panels for disease assessment. Thus identification of a panel of biomarkers that are differentially expressed between asthmatics and COPD patients, and between patients with and without these airway diseases is required. In this study we have utilised 2D-difference gel electrophoresis (2D-DIGE) to conduct an unsupervised analysis of circulating proteins in well characterised subjects with asthma and COPD, and applied logistic regression to evaluate the power of combining markers compared with the markers in isolation. The biomarker panel was then applied to a second validation population with obstructive airway disease to demonstrate the power of this approach.

Some of the results of this study have been previously reported in the form of an abstract (12).

METHODS

Subjects

We confirmed 2D-DIGE reproducibility (on-line supplement, Figure E1) and conducted a proteomic discovery study and 2 validation studies. The discovery study assessed mutually exclusive clinical groups (n=43, Table 1 and Table E3) of stable non-smoking asthma (n=21), COPD (n=5), and matched healthy controls (n=17). The biomarkers were then validated by immunoassay in this population, supplemented by 7 asthmatics and 9 COPD patients (n=58, Table E6). A clinical validation study then assessed the biomarker panel in older adults (>55 years) with obstructive airway disease (FEV_1 <80% predicted and FEV_1/VC<70%) consecutively recruited from the clinic (n=50) (Table 2, asthma (n = 14), COPD (n = 22), overlap asthma/COPD (n = 14)), and in age matched controls (n=23) recruited by advertisement (12). The study was approved by institutional ethics committees and subjects gave written informed consent.

The proteomic discovery study and immunoassay validation study recruited subjects from mutually exclusive clinical groups of asthma, COPD, and controls without airway disease. None of the subjects were current smokers. Asthma was defined as a compatible history of episodic symptoms responsive to asthma treatment, together with airway hyperresponsiveness to hypertonic (4.5%) saline. Asthmatics had fully reversible airflow obstruction. and DLCO was >70% predicted. The COPD group were selected to have COPD (GOLD stage II or greater) with predominant airway disease with minimal emphysema (D_{LCO}>60% predicted) and no asthma. They were ex-smokers with a significant smoking history, compatible symptoms and a doctors diagnosis of COPD, and incompletely reversible airflow obstruction (FEV_1<80% predicted, FEV_1/FVC < 0.7 postbronchodilator) and a negative bronchodilator response (BDR). The clinical validation study applied these same definitions, and subjects with features of both asthma (variable airflow obstruction) and COPD (incomplete reversibility of airflow obstruction) were labelled mixed asthma/COPD.

Sample Preparation & 2D-DIGE & Mass Spectrometry

Platelet depleted plasma was processed on a ProteomeLab IgY-12 LC2 column (Beckman Coulter, CA, USA), precipitated via TCA/acetone and protein quantified (2D quant; GE Biosciences). Protein (100µg) was labelled with Cy3 or Cy5 dyes (GE Biosciences) and separated on 24cm pH 4-7 IPG strips followed by 4-18% polyacrylamide gels in the second dimension, was visualised on a Typhoon 9410 Variable Mode Imager as described (13). A pooled internal control consisting of individuals from each clinical group (n=10) was labelled with Cy2 and included on every gel. Image analysis and comparison was performed using the Batch Processor and Biological Variation Analysis modules of the DeCyder software version 6.5 (GE Healthcare, Australia). Average expression ratios of individual protein spots were compared using one way analysis of variance (1-ANOVA) and proteins with an expression ratio of ±1.2 (assigned following reproducibility analysis) and $p<0.05$ were then manually

inspected for densitometric Gaussian distribution and match quality. Proteins of interest were excised from preparative 2D gels, trypsin digested, and analysed on an Ettan MALDI-ToF/Pro or an Applied Biosystems 4700 Proteomics Analyser with TOF/TOF optics in MS mode as previously described (13). Data was submitted to the database search program Mascot (Matrix Science Ltd, London, UK) for protein identification as described (13).

Immunoassay

Serum IgA, complement factor H, haptoglobin, α-2 macroglobulin, ceruloplasmin, hemopexin and antithrombin III concentrations were assayed with commercially available ELISAs. For Western blotting, proteins were separated on 10% or 4-18% polyacrylamide gels, transferred to nitrocellulose, and probed with anti-prothrombin , anti-fibrinogen gamma chain , or anti-inter-alpha-trypsin inhibitor heavy chain H4 (anti-ITI-H4) primary antibodies followed by appropriate HRP-linked secondary antibodies. Reactive bands were visualised using ECL and a Fujifilm Luminescent Image Analyser LAS-300 and quantified using Multi Gauge V3.0 software. The chemiluminescent intensity was normalised to a healthy control serum sample run on every gel.

Statistical Analysis

Differential expression between groups was assessed using Students t-test (GraphPad Prism 4 for Windows, GraphPad Software Inc.) with Bonferoni corrected p values. Logistic regression (Stata 9, StataCorp) was used to calculate the predicted value of an individual having the disease based on their level of a single marker or combination of markers. Receiver-Operating Characteristic (ROC) curves were generated, and sensitivity, specificity and the area under the curve were calculated. Significance was accepted when $p<0.05$. The possibility of biomarker serum concentrations varying with inhaled corticosteroid (ICS) use and age was assessed using either a simple logistic regression or Spearman's rank correlation, as appropriate.

RESULTS

Differential Protein Expression Between Healthy Controls, Asthmatics and COPD Patients.

Plasma proteins were separated by 2D-DIGE for 43 subjects across the 3 clinical groups: healthy controls (n=17), asthma (n=21) and COPD (n=5). Over 1900 proteins were detected on each gel, 72 proteins displayed significantly altered expression ($p<0.05$, 1-ANOVA), and 58 were identified by mass spectrometry (Table 3; Figure 1). While 58 individual protein spots were differentially expressed, a number of these represented variant charge and/or molecular weight isoforms of the same protein (e.g. 10 isoforms of IgA (Table 3)), thus a total of 20 unique proteins were identified as candidate markers. These proteins clustered in groups of known function including iron metabolism (ceruloplasmin, haptoglobin, hemopexin), the coagulation cascade (α-2-macroglobulin, prothrombin, fibrinogen gamma, fibrin beta), immunoglobulins (IgA, IgM), and complement pathways (complement factor H, complement factor B).

Validation of Candidate Markers in the Diagnostic Immunoassay Validation Assessment (DIVA) Group

In order to assess the performance of the markers identified by 2D-DIGE analysis, we determined the relative expression of selected marker candidates in serum by immunoassay (ELISA or immunoblotting). Markers were prioritised for validation according to three criteria (i) multiple isoforms identified with consistent trends observed, (ii) biologically plausible in the context of respiratory disease and (iii) the availability of commercially available ELISA kits or primary antibodies. Serum concentrations of ceruloplasmin ($p = 0.0002$), haptoglobin ($p = 0.0003$) and antithrombin III ($p = 0.0140$) were significantly higher in asthmatics compared with healthy controls. Additionally, serum levels of prothrombin

(high molecular weight isoform) were lower in COPD patients than asthmatics ($p = 0.0024$) and healthy controls ($p < 0.0001$) (Figure 2).

Validation of Candidate Markers in the Clinical Validation Population

We then proceeded to validate the candidate biomarkers in an unselected clinical population consisting of older adults with asthma, COPD, and asthma-COPD overlap, and who also exhibited significant co morbidity. Serum concentrations of ceruloplasmin ($p = 0.0010$), haptoglobin ($p = 0.0073$) and hemopexin ($p = 0.0004$) were significantly higher in asthmatics compared with healthy controls. When compared to controls, the COPD patients showed trends for elevations in ceruloplasmin, α-2 macroglobulin, haptoglobin and hemopexin. Haptoglobin ($p = 0.0014$) and hemopexin ($p = 0.0024$) exhibited significantly elevated serum levels in individuals with overlap asthma-COPD compared with healthy controls. None of the markers investigated showed statistically significant differences between COPD patients and individuals with overlap asthma-COPD and asthmatics (Figure 3).

Analysis of a panel of Proteomic Biomarkers

The performance of a biomarker panel was evaluated in the two populations and the candidate markers ceruloplasmin, α-2 macroglobulin, haptogobin and hemopexin were selected for use in a combinatorial analysis due to their statistically significant differences in serum marker levels between disease groups and their consistent performance over the two distinct populations.

DIVA Group: Several biomarker combinations were capable of distinguishing between all three clinical groups (asthma, COPD, healthy control) in a manner superior to any one marker in isolation (Table 4, Table E4). The combination of ceruloplasmin and haptoglobin was the best for discriminating between asthmatics and healthy controls, whereas serum ceruloplasmin and hemopexin concentrations in combination best distinguished between COPD patients and healthy controls. For differential diagnosis between COPD patients and

asthmatics several marker combinations were able to discriminate between the two disease states. The combination α-2 macroglobulin and either haptoglobin or hemopexin provided a good balance of sensitivity and specificity whereas α-2 macroglobulin, haptoglobin and hemopexin delivered the best sensitivity (92%) and α-2 macroglobulin, hemopexin and ceruloplasmin gave the best specificity (89.3%). It is interesting to note that the serum concentration of α-2 macroglobulin alone performed almost as well, indicating that it was the predominant contributor to the power of the combinations tested.

Clinical Validation Population:. The combination of haptoglobin and hemopexin was the best for discriminating between asthmatics and healthy controls in this population (Table 5, Table E5, Figure 4). Ceruloplasmin either alone, or combined with haptoglobin also performed well. α-2 macroglobulin and hemopexin concentrations in combination best distinguished between COPD patients and healthy controls. The combination of ceruloplasmin and hemopexin also demonstrated statistically significant discrimination for COPD and controls, in agreement with the analysis of the primary validation group.

For differential diagnosis between COPD patients and asthmatics the combination α-2 macroglobulin, haptoglobin and ceruloplasmin delivered the best sensitivity whereas α-2 macroglobulin, haptoglobin and hemopexin provided the best specificity. The combination of the markers α-2 macroglobulin, haptoglobin and hemopexin was the best for discriminating between individuals with asthma-COPD overlap and healthy controls α-2 macroglobulin and ceruloplasmin serum levels in combination could discriminate asthmatics and individuals with overlap asthma/COPD No marker combinations or markers in isolation were able to discriminate between COPD patients and individuals with overlap asthma-COPD with statistical significance.

Effect of Co-morbid Conditions in the Clinical Validation Population

A literature review identified 5 medical conditions as potential confounders of the blood based marker diagnosis panel: hepatobiliary disease (14), ischemic coronary disease (15), obesity (16, 17), endocrine and metabolic disorders (18, 19) as well as psychiatric conditions (20). Table E7 summarises a co-morbidity analysis of the clinical validation population and indicates no significant effect of these co-morbid conditions on the marker levels, apart from an effect of psychiatric disease on levels of α-2 macroglobulin.

Potential Effect of ICS Use and Age on Biomarker Serum Concentrations

For the DIVA group, stable asthmatics (but not COPD patients) were prescribed ICS. The potential effect of ICS use on serum marker concentrations was assessed and it was found that ICS dosage did not significantly alter serum marker concentrations and is thus unlikely to be a cause of the observed elevation of serum marker concentrations in this study. [Hemopexin: p = 0.5135; haptoglobin: p=0.083; ceruloplasmin: p = 0.6723; α-2 macroglobulin: p = 0.6761].

For the clinical validation population many COPD patients were also using ICS. In order to evaluate the effect of ICS use on serum marker levels in COPD patients, individuals from the clinical validation and DIVA (none using ICS) groups were combined and reclassified into two groups (with and without ICS). Logistic regression analysis identified no statistically significant effect of ICS use on marker concentration.

Additionally, we evaluated the potential effect of age on serum marker concentrations and found that there was no significant correlation between age and serum marker concentration for the DIGE, DIVA and clinical validation groups ($p>0.05$ for all panel markers).

DISCUSSION

This proteomic discovery programme has identified a panel of protein markers whose serum concentrations are significantly altered in asthma and COPD compared with age and sex matched healthy controls, and whose function points to novel mechanistic pathways

indicating the involvement of the iron metabolism pathways, complement pathways, and the coagulation cascade in asthma and COPD. In addition, we have identified a panel of 4 serum biomarkers that includes ceruloplasmin, haptoglobin, hemopexin and α-2-macroglobulin that can be used in combination to accurately identify asthma and COPD, and have validated these results in a second clinical population of older adults with obstructive airway diseases, including asthma and COPD. The proteins in the diagnostic biomarker panel are all involved in the regulation of inflammation, and usually function as anti-inflammatory proteins.

We used an unbiased analysis design in well characterised groups of patients with asthma and COPD to discover differentially expressed proteins in these groups. Ten of the twenty candidate markers identified were subsequently validated in serum and a biomarker panel including the markers ceruloplasmin, α-2 macroglobulin, haptoglobin and hemopexin provided significant discrimination between subject groups in both validation studies. Although individual markers can differentiate between particular clinical groupings, logistic regression analysis has shown that the consideration of marker combination for each specific comparison yields vastly superior performance. ROC curves constructed for each comparison are of high quality (Figure 4) and specific cut-off points may be chosen to tailor the test for either maximum sensitivity or specificity, depending upon the diagnostic requirements. Indeed, apart from mixed airways disease and COPD (for which no combination can differentiate) cut-points may be chosen where sensitivity and specificity are well balanced (Tables 4 and 5). Thus our study has identified a panel of highly discriminatory proteins that could be extremely useful in a clinical context.

The biomarker panel comprises three positive acute-phase proteins (α-2 macroglobulin, ceruloplasmin and haptoglobin) and one type II acute-phase protein (hemopexin). These are predominantly liver-synthesised proteins that can have important anti-inflammatory activity through inhibition of oxidative stress, and iron sequestration resulting in antimicrobial

activity. As such, they may function to modulate the systemic inflammatory response to inflammation (21) and be involved in tissue repair through fibrosis and angiogenesis.

The acute phase response is an innate body defence observed during infection, physical trauma, malignancy and tissue damage that aims to minimise on-going tissue damage by isolating and destroying infective agents whilst activating repair processes. It is an antigen non-specific, innate response which aims to eliminate microbes and hence prevent infection. The innate immune response involves the recruitment and activation of macrophages and leukocytes that release inflammatory cytokines upon recognition of a PAMP. These cytokines travel through the bloodstream and stimulate hepatocytes in the liver to synthesize and secrete acute-phase proteins which we have in this case identified as differentially expressed between our clinical groupings in circulating blood.

Ceruloplasmin, also called ferroxidase, is a multi-functional, copper protein synthesised primarily in the liver and by activated macrophages. It has important roles in iron homeostasis, inflammation, and it has antimicrobial activity via regulation of iron availability to microorganisms. Serum levels have previously been reported to be elevated (22, 23) or unchanged in children with allergic asthma (24). Engstrom reported that ceruloplasmin and haptoglobin were weakly correlated with lung function in COPD, and were associated with an increased future risk of hospitalisation in COPD (25). Our results extend these observations by showing elevated ceruloplasmin in adults with asthma and COPD, as well as asthma-COPD overlap. In addition, we show that the elevation of ceruloplasmin is part of an increase in iron metabolism proteins as part of asthma and COPD, probably as an anti-inflammatory response to the airway inflammation that characterises these conditions.

Haptoglobin and hemopexin were other iron–related proteins that were altered in this study.

The inhibition of heme release from globin by haptoglobin and sequestration of heme by hemopexin suppress hemoglobin-mediated oxidative stress, attenuates endothelial

cytotoxicity and protects cells from heme toxicity. Iron sequestration may also have antimicrobial effects by depriving microbes of essential iron. Additionally, hemoglobin and its derivative heme are often released into tissue compartments where there is infection and inflammation, in the presence of degrading blood, and hemoglobin synergizes with multiple TLR agonists to induce release of high levels of tumor necrosis factor and interleukin-6 from macrophages, an effect that is attenuated by hemopexin (26). Hemopexin also down-regulates LPS-induced proinflammatory cytokine release from macrophages (27). Hemopexin has not been previously been associated with asthma or COPD. However, a 2D-DIGE based proteomic study of BAL in sarcoidosis patients, chronic beryllium disease (CBD) patients and controls showed a significant difference in BAL hemopexin concentration between controls and CBD patients. CBD is a granulomatous disorder that can lead to chronic lung inflammation and fibrosis (28).

Haptoglobin can be expressed by eosinophils, and variable serum levels have been reported in asthma, where both elevated (29) and reduced (30) serum haptoglobin levels are described. Increases in haptoglobin are seen in uncontrolled asthma, such as asthma exacerbation (31) and 24 hours after allergen challenge in late responders (32). In asthma, haptoglobin has also been correlated with FEV_1 (29). As part of its tissue repair function, haptoglobin can induce differentiation of fibroblast progenitor cells into lung fibroblasts (33), and angiogenesis, potentially implicating haptoglobin in remodelling and fibrosis in asthma and COPD. Haptoglobin has not previously been linked to COPD.

α-2 macroglobulin binds host or foreign peptides and particles, thereby serving as a humoral defense barrier against pathogens in the plasma and tissues of vertebrates. It interacts with and captures virtually any proteinase including serine, cysteine, aspartic and metalloproteinases e.g. gelatinase (MMP-2 and MMP-9), both self and foreign, suggesting a function as a unique "panproteinase inhibitor" (34). Its structure contains a 35 amino acid

"bait" region. Proteinases binding and cleaving the bait region become bound to α- 2 macroglobulin and the proteinase-α-2 macroglobulin complex is recognised by macrophage receptors and cleared from the system (35). α-2 macroglobulin provides negative feedback on the inflammatory response by inhibiting thrombin (coagulation) and plasmin (fibrinolysis). It has been studied in airway secretions as a marker of plasma exudation, and is increased in sputum samples in asthma and COPD (36). Plasma levels may be increased in asthma (37), and are normal in emphysema but reduced in chronic bronchitis (38).

In agreement with these proteins being positive acute-phase proteins, their serum levels are elevated in the asthmatic and COPD groups (independent of age and ICS use) in both our validation populations relative to the healthy controls in our study. More importantly, the differential expression of the markers in our panel makes sense in the context of asthma and COPD which both manifest inflammatory and fibrotic components during their progression.

Some limitations to our study relate to subject selection in the discovery population. The COPD group was small in number, and selected to have airway predominant disease. The primary reason for studying airway predominant COPD was to rigorously test the hypothesis that there would be differential markers between asthma and COPD reflective of small airway inflammation which is a relevant lesion in both asthma and COPD. This approach eliminates the potentially confounding effects of airspace disease. By designing the study this way, we believe that the information about potential disease mechanisms is not confounded by the site of pathology. In addition, because COPD is a heterogeneous disease, we thought it was important to minimize this heterogeneity by studying a recognized and clinically relevant group of COPD subjects. A larger group or one with more emphysema may identify additional proteins that are associated with COPD, and this would be useful future work. Similarly, the COPD subjects in the discovery group were not using inhaled corticosteroid (ICS), whereas, most of the patients in the clinical validation group were prescribed ICS. This

has the potential to influence results. We conducted some analyses to test for these effects and none were apparent, however this requires further study. The 2D-DIGE technology is not well suited to analysis of small molecular weight proteins, eg <10kDa. This means that many cytokines would not be detected by this approach, and additional methods such as used by PintoPlata are useful here (9).

The strengths of the results in this study come from the approach used and the combination of a panel of markers to use for disease discrimination. A panel of independent disease related biomarkers, as can be identified by proteomics, is generally considered to be more powerful and less prone to the influence of genetic and environmental 'noise' than a single marker protein (11). For example, Rai *et al.* identified three potential biomarkers that could differentiate ovarian cancer from healthy individuals and compared their performance against the tumour marker, cancer antigen 125 (CA125) (39). Each biomarker individually did not out-perform CA125, however the combination of two of the new biomarkers together with CA125 significantly improved their performance (39, 40). Similarly, we show here that the combination of protein biomarkers significantly improves the performance as a diagnostic marker than each individual protein alone.

Proteomic analysis, utilizing high-resolution 2D-gel electrophoresis coupled with mass spectrometry, is a powerful means to identify differential protein expression between biological samples. However, a major limitation in traditional 2D-gel technology is the reproducibility, and thus statistical comparison of protein expression between individual gels is difficult. A recent advance in this area has come from the introduction of Cy dye fluorophores for pre-labelling of protein samples. Two-dimensional difference gel electrophoresis (2D-DIGE) technology adds a quantitative component to conventional 2D-gel analyses, allowing for comparison of protein expression changes across multiple samples simultaneously without gel-to-gel variation, and hence with statistical confidence (41, 42).

Our reproducibility study (Online supplement) clearly shows that this technology, together with immunodepletion of abundant plasma proteins, can reproducibly separate thousands of proteins. Combined with well defined clinical groups and advanced statistical analyses, we have shown that this technology is a powerful tool for the identification of novel disease biomarkers. As the biomarkers are detectable in blood, a readily obtainable biological sample, and reagents are currently available for testing the abundance of these proteins, this panel of biomarkers has the potential to become an extremely useful addition to the clinical diagnosis and management of respiratory disease.

Acknowledgements

This work was supported by the CRC for Asthma and Airways. NMV and LGW were supported by Postdoctoral Fellowships from the National Health and Medical Research Council, Australia. PGG was supported by a Practitioner Fellowship from the National Health and Medical Research Council, Australia. The authors would like to acknowledge the technical assistance of Michelle Gleeson and Anne-Marie Gibson. This research has been facilitated by access to the Australian Proteome Analysis Facility established under the Australian Government's Major National Research Facilities Program.

References

1. Verrills, N. M. Clinical proteomics: present and future prospects. *Clin Biochem Rev* 2006; 27(2): 99-116.
2. Fehniger TE, Sato-Folatre JG, Malmstrom J, Berglund M, Lindberg C, Brange C, Lindberg H, Marko-Varga G. Exploring the context of the lung proteome within the airway mucosa following allergen challenge. *J Proteome Res* 2004; 3:307-320.
3. Fajardo I, Svensson L, Bucht A, Pejler G. Increased levels of hypoxia-sensitive proteins in allergic airway inflammation. *Am J Respir Crit Care Med* 2004; 170:477-484.
4. Jeong H, Rhim T, Ahn MH, Yoon PO, Kim SH, Chung IY, Uh S, Kim SI, Park CS. Proteomic analysis of differently expressed proteins in a mouse model for allergic asthma. *J Korean Med Sci* 2005; 20:579-585.
5. Roh GS, Shin Y, Seo SW, Yoon BR, Yeo S, Park SJ, Cho JW, Kwack K. Proteome analysis of differential protein expression in allergen-induced asthmatic mice lung after dexamethasone treatment. *Proteomics* 2004; 4:3318-3327.
6. Houtman R, Krijgsveld J, Kool M, Romijn EP, Redegeld FA, Nijkamp FP, Heck AJ, Humphery-Smith I. Lung proteome alterations in a mouse model for nonallergic asthma. *Proteomics* 2003; 3:2008-2018.
7. Wu J, Kobayashi M, Sousa EA, Liu W, Cai J, Goldman SJ, Dorner AJ, Projan SJ, Kavuru MS, Qiu Y, Thomassen MJ. Differential proteomic analysis of bronchoalveolar lavage fluid in asthmatics following segmental antigen challenge. *Mol Cell Proteomics* 2005; 4:1251-1264.

8. Jeong HC, Lee SY, Lee EJ, Jung KH, Kang EH, Kim JH, Park EK, Lee SH, Uhm CS, Cho Y, Shin C, Shim JJ, Kim HK, In KH, Kang KH, Yoo SH. Proteomic analysis of peripheral T-lymphocytes in patients with asthma. *Chest* 2007; 132:489-496.

9. Pinto-Plata V, Toso J, Lee K, Park D, Bilello J, Mullerova H, De Souza MM, Vessey R, Celli B. Profiling serum biomarkers in patients with COPD: associations with clinical parameters. *Thorax* 2007; 62: 595-601.

10. Auffray C, Adcock IM, Chung KF, Djukanovic R, Pison C, Sterk PJ. An integrative systems biology approach to understanding pulmonary diseases. *Chest* 2010; 137:1410-1416.

11. Anderson L. Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J Physiol* 2005; 563:23-60.

12. McDonald VM, Simpson JL, Higgins I, Gibson PG. Multidimensional assessment of older people with asthma and COPD: clinical management and health status. *Age Ageing*, in press.

13. Verrills NM, Liem NL, Liaw TY, Hood BD, Lock RB, Kavallaris M. Proteomic analysis reveals a novel role for the actin cytoskeleton in vincristine resistant childhood leukemia--an in vivo study. *Proteomics* 2006; 6(5):1681-1694.

14. Meliconi R, Parracino O, Facchini A, Morselli-Labate AM, Bortolotti F, Tremolada F, Martuzzi M, Miglio F, Gasbarrini G. Acute phase proteins in chronic and malignant liver diseases. *Liver* 1988; 8(2):65-74.

15. Correale M, Brunetti ND, De Gennaro L, Di Biase M. Acute phase proteins in atherosclerosis. *Cardiovasc Hematol Agents Med Chem* 2008; 6:272-277.

16. Chiellini C, Santini F, Marsili A, Berti P, Bertacca A, Pelosini C, Scartabelli G, Pardini E, López-Soriano J, Centoni R, Ciccarone AM, Benzi L, Vitti P, Del Prato S,

Pinchera A, Maffei M. Serum haptoglobin: A novel marker of adiposity in humans. *J Clin. Endocrinol. Metab.* 2004; 89:2678-2683.

17. Cignarelli M, De Pergola G, Picca G, Sciaraffia M, Pannac Turrisi E, Giorgino R. Relationship of obesity and body fat distribution with ceruloplasmin serum levels. *Int J Obes* .1996: 20(9): 809-813.

18. Kim CH, Park JY, Choi CS, Kim YI, Chung YE, Lee MS, Hong SK, Lee KU. Elevated serum ceruloplasmin levels in subjects with metabolic syndrome: a population-based study. *Metabolism* 2002; 51(7):838-842.

19. Jonsson A, Wales JK. Blood glycoprotein levels in diabetes mellitus. *Diabetologia* 1976; 12(3):245-250.

20. Maes M, Delange J, Ranjan R, Meltzer HY, Desnyder R, Cooremans W, Scharpe S. Acute phase proteins in schizophrenia, mania and major depression: modulation by psychotropic drugs. *Psychiatry Res* 1997: 66(1):1-11.

21. Fink MP Hemopexin: newest member of the anti-inflammatory mediator club. *J Leukoc Biol* 2009; 86(2):203-4.

22. Vural H, Uzun K. Serum red blood cell antioxidant status in patients with bronchial asthma. 2000; *Can Respir J* 7(6): 476-480.

23. el-Kholi MS, Gas Allah MA, el-Shimi S, el-Baz F, el-Tayeb H, Abdel-Hamid MS. Zinc and copper status in children with bronchial asthma and atopic dermatitis. 1990; *J Egypt Public Health Assoc*. 65(5-6):657-668.

24. Di Toro R, Galdo Capotorti G, Gialanella G, Miraglia del Giudice M, Moro R, Perrone L. Zinc and copper status of allergic children. *Acta Paediatr Scand* 1987; 76(4) 612-617.

25. Engstrom G, Segelstorm N, Ekberg-Aronsson M, Nilsson PM, Lingarde F, Lofdahl CG. Plasma markers of inflammation and incidence of hospitalisations for COPD: results from a population-based study. *Thorax* 2009; 64(3):211-215.

26. Lin T, Kwak YH, Sammy F, He P, Thundivalappil S, Sun G, Chao W, Warren HS. Synergistic Inflammation Is Induced by Blood Degradation Products with Microbial Toll-Like Receptor Agonists and Is Blocked by Hemopexin. *J Infect Dis* 2010; Jul 9. (Epub ahead of print)

27. Liang X, Lin T, Sun G, Beasley-Topliffe L, Cavaillon JM, Warren HS. Hemopexin down-regulates LPS-induced proinflammatory cytokines from macrophages. *J Leukoc Biol* 2009; 86(2):229-35.

28. Silva, E. Bourin, S. Sabounchi-Schutt, F. Laurin, Y. Barker, E. Newman, L.Eriksson, H. Eklund, A. Grunewald, J. A quantitative proteomic analysis of soluble bronchoalveolar fluid proteins from patients with sarcoidosis and chronic beryllium disease. *Sarcoidosis Vasc Diffuse Lung Dis* 2007; 24(1):24-32.

29. Kauffmann FC, Frette I, Annesi M-F, Dore, Neukirch F. Relationships of haptoglobin level to FEV_1, wheezing, bronchial hyper-responsiveness and allergy. *Clin. Exp. Allergy* 1991; 21: 669-674

30. Piessens MF, Marien G, Stevens E. Decreased haptoglobin levels in respiratory allergy. *Clin. Allergy* 1984; 14: 287-293

31. Koh YY, Kim YW, Park JD, Oh JW. A comparison of serum haptoglobin levels between acute exacerbation and clinical remission in asthma. *Clin Exp Allergy* 1996; 10:1202-9.

32. Kim CK, Chung CY, Koh YY Changes in serum haptoglobin level after allergen challenge test in asthmatic children. *Allergy* 1998; 53(2):184-9.

33. Larsen K, Macleod D, Nihlberg K, Gurcan E, Bjermer L, Marko-Varga G, Westergren-Thorsson G. Specific haptoglobin expression in bronchoalveolar lavage during differentiation of circulating fibroblast progenitor cells in mild asthma. *J Proteome Res* 2006; 5:1479-1483.

34. Borth, W. α-2- macroglobulin, a multifunctional binding protein with targeting characteristics. *The FASEB Journal* 1992; 6:3345-3353.

35. Kurdowska A, Alden SM, Noble JM, Stevens MD, Carr FK (). Involvement of α-2-macroglobulin receptor in clearance of interleukin-8-α-2-macroglobulin complexes by human alveolar macrophages. *Cytokine* 2000; 12(7):1046-1053.

36. Schoonbrood DF, Lutter R, Habets FJ, Roos CM, Jansen HM, Out TA. Analysis of plasma-protein leakage and local secretion in sputum from patients with asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1994; 150(6 Pt 1):1519-27.

37. Gaillard MC, Reichberg SB, Nogueira CM, Kilroe-Smith TA. Differences in elastase-binding activity of alpha 1-protease inhibitor and alpha 2-macroglobulin for asthma patients and control subjects with various alpha1-protease inhibitor phenotypes. *Clin Chem* 1993; 39(4):675-9.

38. Burnett D, Stockley RA. Serum and sputum alpha 2 macroglobulin in patients with chronic obstructive airways disease. *Thorax* 1981; 36(7):512-6.

39. Rai AJ, Zhang Z, Rosenzweig J, Shih Ie M, Pham T, Fung ET, Sokoll LJ, Chan DW. Proteomic approaches to tumor marker discovery. *Arch Pathol Lab Med* 2002; 126:1518-1526.

40. Zhang Z, Bast RC, Jr, Yu Y, Li J, Sokoll LJ, Rai AJ., Rosenzweig JM, Cameron B, Wang YY, Meng XY, Berchuck A, Van Haaften-Day C, Hacker NF, de Bruijn HW,

van der Zee AG, Jacobs IJ, Fung ET, Chan, DW. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res* 2004; 64:5882-5890.

41. Von Eggeling F, Gawriljuk A, Fielder W, Ernst G, Claussen U, Klose J, Romer I. Fluorsecent dual colour 2D-protein gel electrophoresis for rapid detection of differences in protein pattern with standard image analysis software. *Int J Mol Med* 2001; 8:373-377.

42. Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Surrie I, Davison M. Validation and development of fluorescent two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* 2001; 1:377-396.

TABLE 1. CLINICAL CHARACTERISTICS OF PEOPLE IN THE 2D-DIGE MARKER DISCOVERY STUDY

Group Description	1 Healthy Controls	2 Stable Asthma	3 COPD	P value
Number of patients	17	21	5	
Age (Years) [a]	44.2 ± 14.4	48.1±12.7	§*65.7 ± 10.6	0.012
Sex (Male/ Female) [c]	8/9	11/10	2/3	1.0
Smoking, n(%)[c]:				0.127
Never	8 (47%	11 (52%)	0	
Ex	9 (53%)	10 (48%)	5 (100%)	
Pack years [a]	20.3 ± 17.4	20.7 ± 13.3	§*72.7 ± 36.9	0.003
Atopy, n(%)[c]	6 (35.3%)	*19 (90.5%)	§1 (20.0%)	<0.0001
%predicted FEV_1 [a,e]	97.7 ± 9.0	*81.3 ± 16.7	*65.0 ± 17.5	0.0001
%predicted FEV_1 [a,f]	101.0 ± 8.7	*85.9 ± 14.4	*69.0 ± 18.3	<0.0001
%predicted FVC[a,e]	101.3 ± 11.3	95.1 ± 14.9	*79.9 ± 10.9	0.010
FEV_1/FVC %[a,e]	79.5± 8.0	*70.1 ± 7.9	*63.8 ± 10.0	0.0003
PD_{15} (mL) [b,d]	N/A	5.3 (3.6, 15.1).	4.8 (4.7, 8.4)	0.926
D_{LCO} [a]	N/A	85.9 ± 15.0	65.6 ± 11.5	0.012
ICS use, n(%)[c]	N/A	21(100%)	0	<0.0001
ICS (µg beclomethasone equivalents /day) [b]	N/A	1464 ± 1228	N/A	
ACQ score	N/A	1.1 ± 0.8	N/A	
GOLD Classification, n (%)				
I	N/A	N/A	0 (0%)	
II	N/A	N/A	5 (100%)	
III	N/A	N/A	0 (0%)	
IV	N/A	N/A	0 (0%)	
Induced Sputum Cell Counts				
Quality [b]	19 (17, 20)	19 (17, 20)	19 (18, 20)	0.976
Total cell count x 10^6/ml [b]	3.96 (2.4, 6.7)	3.47 (1.6, 5.1)	6.03 (3.7, 10.1)	0.3224
Neutrophils% [b]	34.5 (17.8, 61.0)	42.0 (27.5, 49.3)	§64.5 (57.8, 67.8)	0.028
Eosinophils% [b]	0 (0, 0.25)	*1.0 (0.25, 6.5)	1.25 (0.25, 1.5)	0.002
Macrophages% [b]	59.8 (36.8, 79.8)	51.5 (41.3, 62.8)	§*31.8 (31.3, 32.0)	0.032
Lymphocytes% [b]	0.75 (0.25, 1.25)	0.5 (0.25, 0.75)	0.75 (0.5, 1.25)	0.551
Columnar epithelial% [b]	2.0 (0.5, 5.5)	2.25 (1.0, 6.5)	1.75 (0.25, 2.25)	0.442
Squamous% [b]	2.7 (1.7, 5.7)	4.5 (0.7, 7.2)	2.2 (0.99, 3.9)	0.453

[a]Values are Mean ± SD, oneway ANOVA; [b]Values are median (interquartile range), Kruskall-Wallis test; [c]Chi Square or Fisher's exact test; [d]PD_{15} is provocation dose resulting in 15% drop in baseline FEV_1 expressed as geometric mean (log SD); FEV_1 is forced expiratory volume in 1 second either pre[e]- or post-[f] bronchodilator; FVC is forced vital capacity; D_{LCO} is carbon monoxide diffusing capacity; ICS is inhaled corticosteroids; Bonferroni post hoc test significant compared to: *Healthy Controls; [§]Stable Asthma. Asthma: symptoms with fully reversible variable airflow obstruction [airway hyperresponsiveness and/or increased bronchodilator reversibility]. COPD: symptoms, incomplete reversibility of airflow obstruction [postbronchodilator FEV<80%predicted].

TABLE 2. CLINICAL DATA FOR THE CLINICAL VALIDATION POPULATION

Group Description	1 Healthy Controls	2 Stable Asthma	3 COPD	4 Overlap Asthma-COPD	P value
Number of patients	23	14	22	14	
Age (Years) [a]	49.9 ± 17.6	*67.7 ± 6.7	*68.0 ± 7.9	*68.7 ± 9.0	<0.0001
Sex (Male/ Female) [c]	11/12	8/6	8/14	5/9	0.566
Smoking, n(%)[c]:					0.759
Never	12 (52.2%)	7 (50%)	8 (36.4%)	5 (35.7%)	
Ex	11 (47.8%)	7 (50%)	13 (59.1%)	8 (57.1%)	
Current	0	0	1 (4.6%)	1 (7.1%)	
Pack years [a]	13.9 ± 15.6	39.5 ± 28.6	39.9 ± 36.9	37.2 ± 37.9	0.151
Atopy, n(%)[c]	6 (26.1%)	*11 (78.6%)	8 (36.4%)	9 (64.3%)	0.006
%predicted FEV_1 [a,e]	103.9 ± 13.6	*56.9 ± 22.4	*54.1 ± 21.5	*50.6 ±14.9	<0.0001
%predicted FEV_1 [a,f]	109.0 ± 13.8	*61.6 ± 23.8	*60.6 ± 22.7	*55.0 ± 17.1	<0.0001
%predicted FVC[a,e]	110.10 ± 14.3	*75.1 ± 18.6	*75.1 ± 22.6	*76.6 ± 14.8	<0.0001
FEV_1/FVC %[a,e]	77.5 ± 5.3	*59.2 ± 15.2	*55.9 ± 13.6	*52.1 ± 12.9	<0.0001
PD_{15} (mL) [b, d]	N/A	7.4 (3.5, 19.6)	5.9 (4.9, 8.7)	6.3 (2.9, 11.2)	0.935
D_{LCO} [a]	N/A	73.3 ± 17.3	72.5 ± 23.2	84.1 ± 21.2	0.323
ICS use, n(%)[c]	0	11 (78.6%)	21 (95.5%)	14 (100%)	0.08
ICS (µg beclomethasone equivalents /day) [b]	N/A	1600 ± 780	1460 ± 646	1657 ± 511	0.659
GOLD Classification, n (%)					
I	N/A	N/A	4 (20%)	N/A	
II	N/A	N/A	9 (45%)	N/A	
III	N/A	N/A	4 (20 %)	N/A	
IV	N/A	N/A	3 (15%)	N/A	
Sputum Cell Counts					
Quality [b]	18 (16, 19)	17 (16, 19)	18.5 (17.5, 19)	17.5 (16, 18)	0.250
Total cell count x 10^6/ml [b]	1.7 (1.3, 3.1)	4.0 (2.8, 8.5)	3.3 (1.9, 7.6)	3.7 (1.4, 9.5)	0.039
Neutrophils% [b]	28.5 (13.3, 54.5)	62.9 (45.5, 78.5)	*56.0 (35.5, 88.3)	*77.0 (45.8, 85.5)	0.003
Eosinophils% [b]	0.25 (0, 0.5)	*1.65 (0.5, 3.25)	*2.0 (0.75, 7.25)	*1.0 (0.25, 5.5)	0.0001
Macrophages% [b]	63.8 (41.5, 77.8)	29.9 (16.8, 52.0)	*15.0 (8.0, 44.0)	*17.5 (7.3, 37.3)	0.0003

Lymphocytes% [b]	0.75 (0.25, 1.75)	0 (0, 0.5)	0.25 (0, 1.0)	0.25 (0, 1.0)	0.062
Columnar epithelial% [b]	1.0 (0.25, 7.5)	1.5 (0.5, 2.0)	0.75 (0.25, 2.25)	1.0 (0.5, 2.25)	0.659
Squamous% [b]	4.5 (1.96, 8.3)	1.6 (0.7, 18.4)	2.7 (0.99, 8.9)	3.9 (2.2, 8.5)	0.725

[a]Values are Mean ± SD, oneway ANOVA; [b]Values are median (interquartile range), Kruskall-Wallis test; [c]Chi Square or Fisher's exact test; [d] PD_{15} is provocation dose resulting in 15% drop in baseline FEV_1 expressed as geometric mean (log SD); FEV_1 is forced expiratory volume in 1 second either pre[e]- or post-[f] bronchodilator; FVC is forced vital capacity; D_{LCO} is carbon monoxide diffusing capacity; ICS is inhaled corticosteroids. Bonferroni post hoc test significant compared to: *Healthy Controls. Asthma: symptoms and reversible variable airflow obstruction [airway hyperresponsiveness and/or increased bronchodilator reversibility]. COPD: symptoms, incomplete reversibility of airflow obstruction [postbronchodilator FEV<80%predicted]; Overlap asthma/COPD: symptoms with increased variability and incomplete reversibility of airflow obstruction.

TABLE 3. DIFFERENTIALLY EXPRESSED PROTEINS BETWEEN THE 3 DISEASE GROUPS (HEALTHY CONTROLS, ASTHMATICS AND COPD PATIENTS) IDENTIFIED BY 2D-DIGE AND MASS SPECTROMETRY.

Master spot no.	ANOVA p-value	Protein ID	Accession #	Differential Expression Ratio (*p<0.05, t-test)		
				Asthmatics / Controls	COPD / Controls	COPD / Asthmatics
Iron Metabolism						
480	0.002	Ceruloplasmin (ferroxidase)	P00450	1.01	-1.34*	-1.35*
1437	0.0066	Haptoglobin	P00738	1.94*	1.48	-1.31
1394	0.016	Haptoglobin	P00738	2.13*	1.59	-1.34
1444	0.022	Haptoglobin	P00738	1.83*	1.37	-1.33
1403	0.027	Haptoglobin	P00738	1.88*	1.59	-1.18
1381	0.03	Haptoglobin	P00738	2.13*	1.37	-1.55
968	0.024	Hemopexin	P02790	-1.2*	NA	NA
965	0.029	Hemopexin	P02790	-1.12	-1.23*	-1.1
977	0.033	Hemopexin	P02790	-1.34*	-1.2	-1.12
Coagulation cascade						
822	0.04	Prothrombin	P00734	-1.08	-1.24*	-1.15
824	0.0054	Prothrombin	P00734	-1.1	-1.31*	-1.19
818	0.016	Prothrombin	P00734	-1.14*	-1.26*	-1.1
819	0.018	Prothrombin	P00734	-1.13*	-1.19*	-1.05
1242	0.0019	Fibrinogen gamma	P02679	1.06	1.31*	1.24*
1235	0.03	Fibrinogen gamma	P02679	1.18	1.48*	1.26
1153	0.011	Fibrinogen gamma	P02679	1.25*	1.28*	1.02
859	0.0041	Heparin cofactor II	P05546	1.18*	-1.25	-1.47*
865	0.0097	Heparin cofactor II	P05546	1.1	-1.2*	-1.32*
611	0.037	Heparin cofactor II	P05546	1.31*	1.05	-1.24
911	0.0028	Protein S precursor	P07225	-1.28*	-1.19	1.07
1152	0.0084	Fibrin beta		1.2*	1.51*	1.26
559	0.025	ITI heavy chain H4	Q14624	-1.24*	-1.15	1.08
557	0.031	ITI heavy chain H4	Q14624	-1.14*	-1.25*	-1.09
558	0.049	ITI heavy chain H4	Q14624	-1.14	-1.29*	-1.13
1284	0.036	alpha-2-antiplasmin	P08697	1.19	1.27*	1.07
967	0.0085	alpha-2 antiplasmin	P08697	1.18	2.07*	1.75*
857	0.041	Histidine-rich glycoprotein	P04196	1.07	-1.22	-1.3*
1116	0.048	Antithrombin-III	P01008	-1.23*	-1.15	1.08
1106	0.019	Antithrombin-III	P01008	-1.18*	-1.22*	-1.04
737	0.042	Gelsolin	P06396	-1.19	-1.43*	-1.21
236	0.0034	Alpha-2 macroglobulin	P01023	-1.31*	NA	NA
495	0.022	Alpha-2-macroglobulin	P01023	-1.23*	-1.02	1.2

498	0.032	Alpha-2 macroglobulin	P01023	-1.34*	-1.17	1.14
249	0.041	Alpha-2-macroglobulin	P01023	-1.3*	NA	NA
Complement Pathways						
349	0.028	Complement factor H	P08603	1.3*	1.34	1.03
338	0.00017	Complement factor H	P08603	-1.53*	1.19	1.82*
336	0.0031	Complement factor H	P08603	-1.24*	1.59*	1.97*
340	0.0045	Complement factor H	P08603	-1.3*	1.27	1.66*
339	0.0078	Complement factor H	P08603	1.16	1.81*	1.56*
680	0.015	Complement factor B	P00751	-1.25	-1.5*	-1.2
688	0.029	Complement factor B	P00751	1.12	-1.2	-1.34*
547	0.032	Complement 3	P01024	-1.18	-1.76*	-1.49
691	0.0034	Complement component C4-A	P0C0L4	-1.29*	-1.59*	-1.23*
Immunoglobulins						
1001	0.0019	IgA	P01876	1.45*	2.38*	1.64*
992	0.0055	IgA	P01876	1.19	1.98*	1.66*
1009	0.0084	IgA	P01876	1.24	1.85*	1.49*
1005	0.016	IgA	P01876	1.33	2.11*	1.58
1017	0.017	IgA	P01876	1.2	1.94*	1.62*
1018	0.022	IgA	P01876	1.11	1.86*	1.67*
1022	0.022	IgA	P01876	1.16	1.86*	1.61*
996	0.024	IgA	P01876	-1.36	1.3	1.77*
1011	0.034	IgA	P01876	1.22	1.92*	1.57
1010	0.039	IgA	P01876	1.3	1.99*	1.53
806	0.011	IgM heavy chain	P01871	-2.17*	-2.04	1.07
796	0.02	IgM heavy chain	P01871	-1.8*	-2.68*	-1.49
794	0.026	IgM heavy chain	P01871	-2.09*	-2.2	-1.06
811	0.018	IgM heavy chain	P01871	-1.74*	-2.22*	-1.27
804	0.04	IgM heavy chain	P01871	-1.87*	-2.42	-1.3

COPD = chronic obstructive pulmonary disease; ITI heavy chain H4 = Inter-alpha-trypsin inhibitor heavy chain H4; IgA = immunoglobulin A; IgM = immunoglobulin M.

TABLE 4. ANALYSIS OF DIAGNOSTIC VALUES OF MARKER AND MARKER IN COMBINATIONS IN DIAGNOSTIC IMMUNOASSAY VALIDATION GROUP.

Comparison	Marker combination	*Logistic Regression Model p value*	*Minimal False Negatives*			*Minimal False Positives*			AUC (%)
			Cut-off, Predicted value	*Sensitivity*	*Specificity*	*Cut-off Predicted value*	*Sensitivity*	*Specificity*	
Asthma v Healthy	**Combination**								
n = 44	***Ceruloplasmin Haptoglobin***	***<0.00001***	***>=0.448***	***92.86***	***75.0***	***>=0.785***	***67.86***	***93.75***	***89.29***
	Ceruloplasmin Hemopexin	0.0006	>=0.442	89.29	62.5	>=0.753	67.86	93.75	83.04
	Haptoglobin Hemopexin	0.0006	>=0.590	85.71	68.75	>=0.676	78.57	81.25	82.59
COPD v Healthy	**Combination**								
n = 30	Ceruloplasmin Haptoglobin	0.0004	>=0.342	84.62	68.75	>=0.590	69.23	87.50	88.94
	Ceruloplasmin Hemopexin	***<0.00001***	***>=0.426***	***92.86***	***75.0***	***>=0.480***	***85.71***	***87.50***	***91.07***
	Haptoglobin Hemopexin	0.009	>=0.444	76.92	62.50	>=0.547	61.54	87.50	81.25
COPD v Asthma	**Combination**								
n = 42	α-2-Macroglobulin Haptoglobin	0.045	>=0.289	84.62	64.29	>=0.382	61.54	78.57	74.73
	α-2-Macroglobulin Hemopexin	0.019	>=0.314	78.57	67.86	>=0.435	57.14	78.57	75.51
	α-2-Macroglobulin Haptoglobin Hemopexin	***0.046***	***>=0.285***	***92.31***	***64.29***	>=0.355	76.92	75.0	***76.10***
	α-2-Macroglobulin Hemopexin Ceruloplasmin	***0.048***	>=0.338	78.57	71.43	***>=0.493***	***50.0***	***89.29***	***75.77***

COPD = chronic obstructive pulmonary disease; AUC = area under the curve.

TABLE 5. ANALYSIS OF THE DIAGNOSTIC VALUES OF MARKER AND MARKER IN COMBINATIONS IN THE CLINICAL VALIDATION POPULATION:

Comparison	Marker combination	*Logistic Regression Model p value*	*Minimal false negatives*			*Minimal false positives*			AUC (%)
			Cut-off Predicted value	*Sensitivity*	*Specificity*	*Cut-off Predicted value*	*Sensitivity*	*Specificity*	
Asthma v Healthy	**Combination**								
n = 36	Ceruloplasmin Haptoglobin	0.0002	>=0.342	85.71	72.73	>=0.561	64.29	95.45	87.01
	Ceruloplasmin Hemopexin	0.0001	>=0.226	92.31	68.18	>=0.421	69.23	81.82	89.16
	Haptoglobin Hemopexin	***<0.00001***	***>=0.289***	***100.0***	***90.91***	***>=0.457***	***92.31***	***95.45***	96.85
COPD v Healthy	**Combination**								
n = 44	Ceruloplasmin Haptoglobin	0.016	>=0.460	72.73	59.09	>=0.573	63.64	81.82	75.41
	Ceruloplasmin Hemopexin	0.008	>=0.521	72.73	77.27	>=0.617	63.64	90.91	76.65
	Haptoglobin Hemopexin	***0.014***	>=0.439	72.73	63.64	***>=0.567***	***59.09***	***95.45***	73.97
	α-2-Macroglobulin Hemopexin	***0.006***	***>=0.412***	***86.36***	***68.18***	***>=0.631***	***54.55***	***95.45***	80.17
Mixed v Healthy	**Combination**								
n = 36	Ceruloplasmin Haptoglobin	0.002	>=0.321	92.86	59.09	>=0.488	64.29	81.82	80.84
	Ceruloplasmin Hemopexin	0.007	>=0.321	92.86	68.18	>=0.464	64.29	86.36	81.49
	Ceruloplasmin Hemopexin Haptoglobin	***0.0007***	>=0.277	92.86	59.09	***>=0.530***	***71.43***	***95.45***	86.04
	α-2-Macroglobulin Hemopexin Haptoglobin	***0.0002***	***>=0.295***	***100.0***	***77.27***	***>=0.435***	***92.86***	***95.45***	94.16
COPD v Asthma	**Individual**								
n = 36	α-2-Macroglobulin	0.013	>=0.585	72.73	71.43	>=0.672	54.55	92.86	71.43
	Combination								
	α-2-Macroglobulin Haptoglobin	0.008	>=0.556	72.73	64.29	>=0.621	68.18	92.86	79.55
	α-2-Macroglobulin Hemopexin	0.028	>=0.523	81.82	61.54	>=0.667	59.09	84.62	74.48
	α-2-Macroglobulin Ceruloplasmin	0.012	>=0.585	72.73	71.43	>=0.665	68.18	92.86	78.25
	α-2-Macroglobulin Haptoglobin Hemopexin	***0.003***	>=0.512	81.82	61.54	***>=0.686***	***68.18***	***100.0***	84.27
	α-2-Macroglobulin Haptoglobin Ceruloplasmin	***0.011***	***>=0.520***	***86.36***	***64.29***	>=0.641	63.64	85.71	79.87
	α-2-Macroglobulin Hemopexin Ceruloplasmin	0.024	>=0.477	81.82	61.54	>=0.684	68.18	84.62	78.67
Mixed v Asthma	**Individual**								
n = 28	α-2-Macroglobulin	0.025	>=0.392	78.57	64.29	>=0.544	57.14	71.43	70.92
	Combination								
	α-2-Macroglobulin Haptoglobin	0.069	>=0.340 or >=0.450	85.71 or 78.57	64.29 or 71.43	>=0.490	64.29	78.57	77.04
	α-2-Macroglobulin Hemopexin	0.063	>=0.380	78.57	61.54	>=0.602	57.14	84.62	75.82
	α-2-Macroglobulin Ceruloplasmin	***0.045***	***>=0.368***	***85.71***	***57.14***	***>=0.583 or >=0.509***	***57.14 or 64.29***	***92.86 or 71.43***	75.00
	α-2-Macroglobulin Haptoglobin Ceruloplasmin	0.069	>=0.390	85.71	71.43	>=0.503	71.43	78.57	78.57
	α-2-Macroglobulin Hemopexin Ceruloplasmin	0.089	>=0.428	71.43	61.54	>=0.517	64.29	76.92	73.63

COPD = chronic obstructive pulmonary disease; AUC = area under the curve.

Figure Legends

Figure 1. Representative Cy3-labelled 2D-gel of immunodepleted human plasma proteins from a healthy control. Proteins were separated on pH 4-7 IPG strips in the first dimension, and by 4-18%T SDS-PAGE in the second dimension. Numbered spots are differentially expressed ($p<0.05$; 1-ANOVA) across the 3 disease groups (healthy controls, asthmatics and COPD) and were positively identified by mass spectrometry (Table 3).

Figure 2. Validation data for **A** prothrombin, **B** ceruloplasmin, **C** haptoglobin **D** antithrombinIII in the diagnostic immunoassay validation group. Statistical significance determined using un-paired, two-tailed students t-test with multiple comparisons correction using the Bonferroni method (significant $p < 0.017$). Triplicate analyses were performed using either two-site ELISA or immunoblotting for each patient, and the mean value used for analysis. Error bars represent ± SEM.

Figure 3. Validation data for **A** ceruloplasmin, **B** haptoglobin and **C** hemopexin in the clinical validation population. Statistical significance determined using un-paired, two-tailed students t test with multiple comparisons correction using the Bonferroni method (significant $p < 0.0083$). A single analysis was performed for this population using either two-site ELISA or immunoblotting. Error bars represent ± SEM.

Figure 4. Selected ROC curves for **A** asthma versus healthy controls, **B**. COPD versus healthy controls and **C** COPD versus asthma in the clinical validation population.

Identification of Novel Diagnostic Biomarkers for Asthma and Chronic Obstructive Pulmonary Disease

Online Supplement

Nicole M. Verrills[1,4,7†], Jennifer A. Irwin[2,4,5,7†], Xiao Yan He[2,4,5,7], Lisa G. Wood[1,4,5], Heather Powell[2,4,5], Jodie L Simpson[2,4,5], Vanessa M. McDonald[2,3,4,5], Alistair Sim[1,4] and Peter G. Gibson[2,4,5,6,7].

ADDITIONAL METHODS

Chemicals & Reagents

Cy2, Cy3 and Cy5 (minimal dye), Immobiline Drystrips (pH 4-7, 24 cm) PlusOne™ drystrip cover fluid, Pharmalyte™ 3-10 for IEF, Bind-silane solution, 2D quant kit, ECL anti-mouse IgG HRP-linked whole antibody (from sheep), ECL anti-rabbit IgG HRP-linked whole antibody (from sheep), Hybond® C-extra nitrocellulose and ECL Plus western blotting detection system were purchased from GE Healthcare Bio-Sciences AB, Uppsala, Sweden. SDS, 40% acrylamide solution, 30% acrylamide/Bis solution (37.5:1, 2.6%C), SyproRuby™ protein stain and SDS PAGE standards (broad range, unstained) were all electrophoresis purity and purchased from BioRad, NSW, Australia. Agarose (Type I-A: Low EEO), tributyl phosphine solution (TBP), N,N,N',N'-tetramethylethylenediamine (TEMED, >99%), CF_3COOH (TFA, Fluka Biochimika, >99.5%) NH_4HCO_3 (Reagent Plus®, ≥ 99%), CCl_3COOH (TCA, 99%), *L*-lysine monochloride (> 98%), thiourea (ACS reagent), CHAPS (> 98%), bovine serum albumin (BSA) Fraction V, protease inhibitor cocktail, and anti-goat/sheep-HRP were purchased from Sigma-Aldrich, NSW, Australia. $(NH_4)_2S_2O_8$ (>98 %), urea (>98 %), glycine (>98.5%) were sourced from Chem. Supply, Gillman, SA, Australia. Tris(hydroxymethyl)-aminomethane (ACS reagent), glycerol (BDH AnalR, > 99.5%), CH_3CN (BDH HiPer Solv for HPLC, >99.9 %), NaOH (BDH AnalR, 99%) and CH_3COOH (BDH, 100%) were purchased from Merck Australia, VIC, Australia. Methanol (CH_3OH, > 99.8%) was purchased from Ajax Finechem., NSW, Australia. Dithiothreitol (DTT, >99.5%) was purchased from Applichem, Darmstadt, Germany. Bromophenol blue (BPB) was purchased as the sodium salt from Research Organics, Cleveland, OH, USA. $(CH_3)_2NCOH$ (DMF, ≥99.5%) was purchased from USB Corporation, Cleveland, OH, USA and was stored in the dark

under a N_2 atmosphere. α-cyano-4-hydoxycinnamic acid (α-CHCA, recrystallised) and peptide calibration mix 1 (1000-2500 Da) were purchased from Laser BioLabs, Sophia-Antipolis Cedex, France. Sequencing-grade modified porcine trypsin was purchased from Promega (Madison, WI, USA). Ziptip® μ-C18 pipette tips and 0.45 μm white nylon filters were purchased from Millipore. MA, USA. Human IgA ELISA quantitation kit and ELISA starter accessory package were purchased from Bethyl Laboratories, Inc., TX, USA. Human haptoglobin, α-2 macroglobulin, and hemopexin ELISA kits were purchased from Genway Biotech San Diego, CA, USA. Human ceruloplasmin and antithrombin III ELISA kits were purchased from Assay Pro, St. Charles, MO, USA. Human complement factor H ELISA kit was purchased from Hycult biotechnology, Uden, The Netherlands. Anti-prothrombin (ab48627) was purchased from Abcam, Cambridge, UK. Anti-fibrinogen γ chain monoclonal antibody (M01), clone 1F2 was purchased from Abnova Corporation, Taipei, Taiwan. Anti-inter-α-trypsin inhibitor heavy chain H4 (45A12) monoclonal antibody was purchased from Ab Frontier, Seoul, Korea. All reagents were used as received without further purification.

Clinical Assessment

The recruitment criteria for the proteomic discovery study were established to clinically characterise the subject groups of interest with the aim of establishing mutually exclusive clinical diagnostic groups of asthma, COPD, and controls without airway disease. None of the subjects were current smokers. The selection criteria ensured there was control for smoking and that there was control for age and sex effects by matching subjects within defined age ranges. Stable airway disease was defined as no increase in bronchodilator use, no use of oral corticosteroids, no limitation in activities, no doctor's visit and no hospitalisation due to asthma in the

past 4 weeks. Subjects with recent respiratory tract infection (past 4 weeks) were excluded.

Asthma was defined as a compatible history of episodic wheeze, cough and dyspnea that were responsive to asthma treatment, together with airway hyperresponsiveness to hypertonic (4.5%) saline. Asthmatics had fully reversible airflow obstruction and no evidence of COPD as FEV_1/FVC ratio after bronchodilator was normal (>70%) and DLCO was >70% predicted.

The COPD group were selected to have COPD (GOLD stage II or greater) with predominant airway disease with minimal emphysema (D_{LCO}>60% predicted) and no asthma. They were ex-smokers with a significant smoking history, compatible symptoms and a doctors diagnosis of COPD, and incompletely reversible airflow obstruction (FEV_1<80% predicted, FEV_1/FVC < 0.7 postbronchodilator). In addition, there was a negative bronchodilator response (BDR).

Healthy controls had no respiratory symptoms, nor a diagnosis of respiratory disease, together with normal measures of airway function.

Sample Collection and Clinical Measurements

Subjects were assessed following an overnight fast (minimum 12 hours). Peripheral blood was collected from a vein in the forearm. Sputum was induced during hypertonic saline challenge as described (E1, E2). Atopy was assessed using skin prick testing to common allergen extracts (Dome/Hollister-Steir; Bayer Pharmaceuticals, Sydney, Australia) for house dust mites (*Dermatophagoides pteronyssinus*), mold mix (*Alternaria, Tenuis, Aspergillus mix*), mixed grasses, and cockroach, together with positive (histamine) and negative (glycerine) controls. Participants were asked to withhold antihistamine 5 days prior to testing. A skin-prick test was defined as positive if the wheal diameter was 3 mm or greater at 15

min. Participants were considered atopic if a positive skin-prick test was recorded for any allergen, with no reaction to the negative control.

Airway responsiveness was assessed from spirometry (KoKo spirometer, PDS instrumentation, Louisville, Co, 80027, USA) with hypertonic saline (4.5%) provocation challenge as described (1). Airflow obstruction was assessed in each participant using spirometry (KoKo K313100 PDS Instrumentation, Louisville, CO, USA) to measure pre and post bronchodilator FEV_1, FVC and FEV1/FVC%, Participants withheld bronchodilators for their duration of action before testing. Three reproducible measurements of FEV1 and FVC were obtained before and after inhalation of 200 mg albuterol via a metered dose inhaler with valved holding chamber (Volumatic, Allen and Hanbury's, Melbourne Victoria, Australia) using predicted values according to Knudson *et al*. Airflow obstruction was defined as an FEV_1 < 80% predicted and an FEV_1/FVC% as <70% and performed using the American Thoracic Society/European Respiratory Society standards of lung function testing.

The carbon monoxide transfer co-efficient (KCO) was determined according to ATS guidelines (Med- Graphics Elite DX Pulmonary function testing system Medical Graphics Corporation, Minnesota, MN, USA) (4).

Induced Sputum Processing

All sputum samples were processed within 2 hours of collection. Mucus portions were selected from saliva and divided into two portions; the first portion was processed for RNA extraction and the second portion was mixed with a dispersing agent (dithiothreitol, DTT, 0.1%, Calbiochem, La Jolla Ca USA). The tube was capped and rocked for 30 minutes at room temperature, after which PBS was added and the dispersed suspension was filtered (60µm, Millipore, Australia). A leukocyte total cell

count (TCC) and cell viability (trypan blue exclusion) was performed using a haemocytometer. After centrifugation at 400 × *g* for 10 minutes, the cell pellet was resuspended to 1 × 10^6 cells/mL using phosphate buffered saline (PBS). The supernatant was aspirated and stored at -80°C. Cytospins were prepared from the resuspended cell pellet.

Sputum Differential

Cytospins were fixed in methanol and stained with May and Grunwald stain and subsequently with Giemsa stain. A total of 400 non-squamous cells were counted, with the squamous cells proportion recorded separately. Cells were identified by their morphology and the differential cell count was expressed as a percentage of non-squamous cells. Cytospin quality was determined using a scale to evaluate squamous contamination, cell and nuclear integrity, presence of airway macrophages and number of cells present on the slide.

Sample collection and immunodepletion

Blood samples were collected from a peripheral vein in the forearm. Serum samples were collected in tubes in the absence of anti-coagulant and plasma with anti-coagulant. For serum isolation, blood was allowed to clot and serum collected via ultrafiltration at 1400g. For plasma collection, samples were centrifuged at 2000g for 10 minutes at room temperature. Plasma was removed and centrifuged at 2500g for 15 minutes at room temperature to deplete platelets. Protease inhibitor cocktail (Sigma-Aldrich, Missouri, USA) (1% v/v) was added to the platelet depleted plasma prior to storage at -80°C. Each plasma sample was subjected to immuno-depletion using an IgY-12 LC-2 Proteome Partitioning Kit (Beckman Coulter™, CA, USA) and an AKTA P920 FPLC system (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Using this methodology, the twelve most abundant plasma proteins (albumin, IgG,

α1-antitrypsin, IgA, IgM, transferrin, haptoglobin, α1-acid glycoprotein, α2-macroglobulin, HDL (apolipoproteins A-I and A-II) and fibrinogen) were depleted from samples to yield a sub-sample for analysis that was enriched with respect to the other proteins present (E3). Briefly, human plasma (50 µL) was diluted in dilution buffer (200 µL, 10 mM Tris-HCl pH 7.4, 150 mM NaCl), centrifuged at 8300 rpm for 1 min and loaded into an injection loop. The IgY column was pre-equilibrated (dilution buffer, 1 column volume (CV), 0.35 mL min^{-1}) prior to automated sample injection. The unbound protein fraction was collected (fraction 1, 5 mL) using a Frac 900 fraction collector. Dilution buffer was passed through the column (3.5 CV, 0.35 mL min^{-1}) prior to elution of the bound (high abundance) proteins with stripping buffer (50 mM glycine-HCl, pH 2.5, 8 CV, 0.35 mL min^{-1}). The eluate was subsequently neutralised with 10 × neutralisation buffer (1M Tris-HCl, pH 8.0) and stored at -80 °C until required. The column was then immediately re-equilibrated with dilution buffer (32 CV, 0.35 mL min^{-1}) to neutralise the column pH. As 100µg of protein was required per sample hence multiple depletion runs (three to four) were performed and pooled in order to obtain the required quantity of protein. Eluted proteins were stored immediately at -80 ºC until required. The unbound column fractions were thawed, replicates pooled and subsequently concentrated by TCA precipitation. Briefly, an aqueous stock solution of TCA (90% ^{w}/v) was added to each of the pooled replicates to yield a final concentration of 9% and the samples were left overnight at 4 °C to facilitate protein precipitation. Each tube was then centrifuged at 15,000 × g (15 min, 4 °C) and the supernatant discarded. The protein pellet was washed sequentially in acidified ethanol (40 mM CH_3COOH in CH_3OH) and ice-cold acetone (precipitate pelleted via centrifugation at 15,000 ×g (15 min, 4 °C)), dried in a rotary vacuum concentrator. Proteins were allowed to precipitate overnight at 4 ºC,

centrifuged for 15 min at 15000 g (4 ºC) and the resultant pellet resuspended in DIGE lysis buffer (30 mM Tris pH 8.5, 2M thiourea, 7M urea, 4%(w/v) CHAPS, pH 8.5). Samples were vortexed briefly and pulse sonicated (2 × 2 second pulse at an amplitude of ≤ 2 microns whilst on ice) prior to analysis. Sample pH was adjusted to 8.5 with NaOH solution (50 mM) and the quantity of protein recovered estimated using a 2-D Quant Kit (GE Biosciences) according to the manufacturer's recommendations.

2D-DIGE & Image analysis

Reproducibility study: An experimental design was formulated in order to test the reproducibility of the sample preparation and 2D-DIGE technique. Plasma was collected from one healthy control (#9173) and divided into 3 separate aliquots (labelled A, B, C). Each sample was immunodepleted, TCA precipitated, and quantitated as described previously. The reproducibility of the FPLC separation is shown in Figure E1. Separate aliquots were then labelled with Cy2, Cy3 or Cy5 dyes. Equal protein from the Cy2 samples were pooled for the internal control to be run on every gel. The remaining samples were separated by 2D-DIGE in two independent gel runs, as shown in Table E1. The variability between dye labelling, sample preparation, and gel runs was determined across all spots using the DeCyder generated volume ratios (Table E2).

Discovery study: An experimental design (Table E3) was formulated with the 43 samples assigned to a particular gel and CyDye™ label. Within each clinical group, some samples were labelled with Cy3 and others the Cy5. An internal pooled control was produced from equal quantities of ten of the samples (two per clinical group). Proteins were CyDye labelled and separated on 24cm pH 4-7 IPG strips in the first dimension and 4-18% gradient acrylamide gels in the second dimension as previously

described. Protein separations were visualised on a Typhoon 9410 Variable Mode Imager with Ettan DALT alignment guides using excitation/emission wavelengths specific for Cy2 (488/520 nm), Cy3 (532/580 nm) and Cy5 (633/670 nm) The photomultiplier tube voltage (PMT) was adjusted to preclude spot intensity saturation. The resultant image files were cropped using the program ImageQuant™ Tools 2.1 and saved using the DIGE file naming format.

Image analysis was performed using the Batch Processor and Biological Variation Analysis (BVA) modules of the DeCyder 2D software version 6.5 (GE Healthcare, Australia). This analysis initially normalizes each sample to its respective in-gel Cy2 internal standard, and then matches all controls and samples between different gels. A total of 1918 spots were detected in the master gel (automatically assigned by DeCyder) and matched across the gel images. Comparing each disease group in the BVA module generated average expression ratios and Student's *t* tests of individual protein spots. One way analysis of variance (1-ANOVA) was also used to identify protein spots that showed a significant change across the three groups (healthy control, asthmatic and COPD). Proteins with an expression ratio of +/-1.2 (assigned following reproducibility analysis – see results) and $p<0.05$ by 1-ANOVA or individual group comparison Student's *t*-test, were assigned as a protein of interest (POI). Each of these proteins spots were manually inspected for densitometric Gaussian distribution and match quality.

Protein Identification

For protein identification, five preparative 2D-gels were prepared, one from each disease group. Immuno-depleted plasma (600 µg) was separated by 2D-PAGE as described above. Gels were stained with SyproRuby™ protein stain or Colloidal Coomassie G250 and protein spots of interest excised and subjected to tryptic

digestion. Tryptic digests were analysed by matrix assisted laser desorption ionisation time of flight (MALDI-TOF) mass spectrometry. Peptide mass fingerprint data was acquired using an Ettan MALDI-ToF/Pro mass spectrometer and analysed using the NCBI FASTA human database provided with the system. The instrument was operated in positive reflectron mode acquiring duplicate spectra (700 – 4000 *m/z*), each comprising signal from 400 laser shots (fixed laser power, 337 nm nitrogen laser). Positive ions were extracted into the mass analyser at 20 kV using pulsed extraction. The instrument was calibrated using Pepmix 1 (Laser Biolabs, Ang II (1046.542 Da) and hACTH 18-39 (2465.199 Da)) according to the manufacturers directions. Each sample spectrum was further internally calibrated using Trypsin I (842.508 Da) and trypsin III (2211.108 Da) autolytic peaks prior to database searching. Tryptic digestion was specified with only 1 missed cleavage allowed. The partial amino acid modifications oxidation (M) and propionamide (C) were considered and a fragment mass tolerance of 0.2 *m/z* allowed. PMF data was searched against Homo Sapien entries in the NCBI FASTA database (20060322) with the protein database update tool v1.30 using the PROFOUND peptide mass fingerprinting search engine licensed from ProteoMetrics (NY, USA). Bayesian statistics were used to rank the protein sequences in the database according to their probability of generating the experimental data (expt < 0.010 significant with $p<0.05$). Proteins were assigned as positive identifications if they showed an expectation value <0.05. Samples that were unable to be identified by MALDI-TOF were analysed by MALDI-TOF/TOF at the Australian Proteome Analysis Facility on an Applied Biosystems 4700 Proteomics Analyser with TOF/TOF optics in MS mode. Data was submitted to the database search program Mascot (Matrix Science Ltd, London, UK) and searched against *Homo Sapien* entries in the Swisprot database. Significant Mascot scores in

the database search indicated a likely match, and were confirmed or qualified by operator inspection.

ELISA analysis

Non-depleted serum from the primary validation population was tested for IgA, complement factor H, haptoglobin, α-2 macroglobulin, ceruloplasmin, hemopexin and antithrombin III concentrations using commercially available two-site ELISA kits following the manufacturer's instructions. Both samples and standards were analysed in duplicate wells for each determination and the quantification was performed in triplicate (three separate runs) for the primary validation population and duplicate wells singly for the secondary validation population. Standard curves were fit using a four parameter logistic regression and individual sample concentrations calculated via interpolation. Triplicate determinations (primary validation population) for each sample were averaged and the mean data analysed for differentiation between known subject groupings using GraphPad Prism® 4.02 for Windows. Unpaired t-test analysis was used to test for statistical significance of the observed experimental trends.

Western Blotting

Blood serum was analysed via Western blotting for the candidate markers prothrombin, inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4) and fibrinogen γ chain . Proteins were separated on 10 % or 4-18% acrylamide gels and transferred to nitrocellulose. Membranes were blocked with 2 % BSA in Tris-buffered-saline with Tween (TBST, 20 mM Tris-Cl, 140 mM NaCl, pH 7.5, 0.05 % Tween-20) for one hour at room temperature or overnight at 4 °C. Primary antibodies anti-prothrombin (1:200), anti-ITIH4 (1/1000) and anti-fibrinogen gamma chain (1:1000) in TBST were incubated with the appropriate blocked membranes with agitation for 60 min at

room temperature, the membranes washed in TBST (triplicate, 20 min) and then incubated with secondary antibody (rabbit anti-mouse IgG, HRP-linked whole antibody, 1:10000 in TBST, 60 min) with agitation. Reactive bands were visualised using ECL chemiluminescent reagents and a Fujifilm Luminescent Image Analyser LAS-300 and quantified using Multi Gauge V3.0 software. The chemiluminescent intensity was normalised to a healthy control serum sample run on every gel. Blots were run in triplicate (primary validation group) or singly (secondary validation group). Unpaired Students *t*-test was used to test for statistical significance between the clinical groups.

Statistical Analysis for quantitative assessment of multiple biomarkers

In this programme logistic regression was used to calculate the predicted value of an individual having the disease based on their level of a particular marker (simple logistic regression) or combination of markers (multiple logistic regression). The regression equation utilised was:

$$\ln(p/(1-p)) = \beta_0 + \beta_i x_i$$

$$p = \exp(\beta_0 + \beta_i x_i)/(1 + \exp(\beta_0 + \beta_i x_i)$$

The predicted value is the probability that the person has the disease given the level of the markers. Receiver-Operating Characteristic (ROC) Curves were generated from the predicted values from the regression analyses. The ROC curves were then used to determine the sensitivity and specificity of both individual or combinations of markers at specified cut points of the predicted values. An individuals' protein measurements can be entered into the logistic regression equation to determine their predicted value, a value above or below the cut point would determine their likelihood of disease with the defined sensitivity and specificity calculated from the

ROC curve. The area under the curve (AUC) was also calculated. The AUC value gives a measure of the ability of the test to correctly classify people.

The significance of the regression models were used to assist in determining the best combination of markers. The combinations with the greatest significance and greatest AUC are highlighted in red in Table E4 (primary validation population) and Table E5 (secondary validation population). Sensitivity and specificity were used to describe the discriminatory power of the protein combinations. We selected cut-points in order to maximise the discriminatory power in any given diagnostic situation. Different batches of ELISA kits were used in the in the two validations and this is why there were some differences in the cut offs for the two datasets. We are now able to source kits from the same company on a consistent basis.

Statistically significant combinations are presented in the two tables in red. The combinations highlighted (bold + italics) display highest sensitivity (rule out) and specificity (rule in). High AUC was also considered.

SUPPLEMENTARY RESULTS

Reproducibility analysis of immuno-depletion and 2D-DIGE

The reproducibility of the sample preparation procedure and 2D-DIGE separation was examined. This involved evaluation of the reproducibility of 2D DIGE separations from one healthy control sample in order to test the variability of three parameters: (i) dye labelling; (ii) sample preparation; (iii) independent gel runs (temporal changes). The densitometric volume relative to the internal control (volume ratio) was determined for each protein using the program DeCyder, and the variation between parameters determined by linear regression analysis (Figure. E1). There is a strong correlation between the spot volume ratio of samples labelled with Cy3 and Cy5, slope = 1.07 ± 0.01, r^2 = 0.9744, $p < 0.0001$. (Figure E1A). Similarly, a good

correlation is observed between two samples that have been independently prepared (IgY column, TCA precipitation, protein quantitation, Cy dye labelling), slope = 0.786 ± 0.02, $r^2 = 0.8155$, $p < 0.0001$ (Figure E1B) and between the same sample Cy-labelled and separated by 2D-DIGE in independent experimental gel runs one week apart, slope = 091 ± 0.02, $r^2 = 0.8495$, $p < 0.0001$ (Figure E1C). Additionally, DeCyder was used to analyse the differences between all possible comparator combinations across the 12 gel images, and the analysis of variance determined (Table E2). For downstream analyses, we arbitrarily set proteins of interest as statistically significant changes that show a fold change of at least the mean ± 1 standard deviation. From the reproducibility analysis (Table E2) the mean ± 1SD correlates to a 1.2 fold up- or down-regulation. Thus only protein changes greater than or equal to ±1.2 fold were considered in the subsequent analyses.

The clinical details of the diagnostic immunoassay validation group are shown in Table E6. These results are similar to the discovery population.

Effect of Co-morbid Conditions in the Clinical Validation Population

A literature review identified 5 medical conditions as potential confounders of the blood based marker diagnosis panel: hepatobiliary disease, ischemic coronary disease, obesity , endocrine and metabolic disorders as well as psychiatric conditions. Table E7 summarises a co-morbidity analysis of the clinical validation population and indicates no significant effect of these co-morbid conditions on the marker levels, apart from an effect of psychiatric disease on levels of α-2 macroglobulin.

REFERENCES

E1. Gibson PG, Wlodarczyk JK, Hensley MJ, Gleeson M, Henry RL, Cripps AW, Clancy RL. Epidemiological association of airway inflammation with asthma

symptoms and airway hyperresponsiveness in childhood. *Am J Respir Crit Care Med* 1998; 158:36-41.

E2. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology* 2006; 11:54-61.

E3. Liu T, Qian W-J, Mottaz HM, Gritsenko MA, Norbeck AD, Moore RJ, Purvine SO, Camp II DG, Smith RD. Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol Cell Proteomics* 2006; 5(11):2167-2174.

E4. American Thoracic Society. Single-breath carbon monoxide diffusing capacity (transfer factor). *Am J Respir Crit Care Med* 1995;152:2185-98.

TABLE E1. 2D-DIGE EXPERIMENTAL DESIGN FOR REPRODUCIBILITY STUDY

Gel #	Cy3	Cy5	Run #
1	Sample A	Sample A	1
2	Sample B	Sample B	1
3	Sample C	Sample C	1
4	Sample A	Sample A	2
5	Sample B	Sample B	2
6	Sample C	Sample C	2

TABLE E2. STATISTICS OF SPOT VOLUME RATIOS ACROSS ALL REPRODUCIBILITY ANALYSES

Number of values	3126
Minimum	0.2128
25% Percentile	0.9259
Median	1
75% Percentile	1.08
Maximum	3.68
Mean	1.008
Std. Deviation	0.1745
Std. Error	0.003121
Lower 95% CI of mean	1.002
Upper 95% CI of mean	1.014

TABLE E3. EXPERIMENTAL DESIGN FOR BIOMARKER DISCOVERY STUDY

Gel #	Cy3 (group #)	Cy5 (group #)	Cy2
1	9028 (1)	9787 (4)	Pooled Control
2	9295 (3)	10336 (4)	Pooled Control
3	9404 (5)	9428 (3)	Pooled Control
4	9629 (4)	9620 (3)	Pooled Control
5	9197 (2)	9750 (5)	Pooled Control
6	9210 (2)	9738 (5)	Pooled Control
7	9559 (4)	9311 (2)	Pooled Control
8	9239(1)	9651 (2)	Pooled Control
9	9694 (5)	9249 (1)	Pooled Control
10	9371 (3)	9305 (1)	Pooled Control
11	9444 (1)	10359 (2)	Pooled Control
12	9476 (3)	10404 (4)	Pooled Control
13	10928 (4)	9506 (3)	Pooled Control
14	9612 (3)	10662 (2)	Pooled Control
15	10911 (2)	9666 (3)	Pooled Control
16	9770 (3)	9450 (1)	Pooled Control
17	11173 (4)	9903 (3)	Pooled Control
18	10855 (2)	10540 (4)	Pooled Control
19	9525 (4)	9494 (1)	Pooled Control
20	11165 (2)	10570 (4)	Pooled Control
21*	9197 (2)	9750 (5)	Pooled Control
22*	11108 (3)	10336 (4)	Pooled Control

* represents repeat gels

TABLE E4. ANALYSIS OF DIAGNOSTIC VALUES OF MARKER AND MARKER IN COMBINATIONS IN PRIMARY VALIDATION POPULATION.

Comparison	Marker combination	Logistic regression: Constant	Logistic regression: Coefficient	Logistic regression: p value	Logistic regression: Model p value	Minimal false negatives/rule disease out: Cut off measurement	Minimal false negatives/rule disease out: Cut-off, Predicted value	Minimal false negatives/rule disease out: Sensitivity	Minimal false negatives/rule disease out: Specificity	Minimal false positives/ rule disease in: Cut off measurement	Minimal false positives/ rule disease in: Cut-off, Predicted value	Minimal false positives/ rule disease in: Sensitivity	Minimal false positives/ rule disease in: Specificity	AUC (%)
Asthma v Healthy														
	Individual													
n = 44	Ceruloplasmin	-4.018952	0.005213	0.002	0.0002	>=761.22mcg/ml	>=0.487	85.71	62.5	>=1017.89mcg/ml	>=0.784	64.29	93.75	81.47
	Haptoglobin	-3.276754	1.109178	0.003	0.0001	>=3.37mg/ml	>=0.613	85.71	75.0	>=3.84mg/ml	>=0.727	67.86	87.50	82.37
	Hemopexin	0.307188	0.0002305	0.904	0.904									50.89
	α-2-Macroglobulin	1.891332	0.0005327	0.190	0.181									65.63
	Combination													
	Ceruloplasmin	***-6.880227***	***0.0045298***	***0.015***	***<0.00001***		***>=0.448***	***92.86***	***75.0***		***>=0.785***	***67.86***	***93.75***	***89.29***
	Haptoglobin		***1.036785***	***0.014***										
	Ceruloplasmin	-2.27678	0.005567	0.002	0.0006		>=0.442	89.29	62.5		>=0.753	67.86	93.75	83.04
	Hemopexin		-0.0018857	0.436										
		-2.215251	1.144461	0.003	0.0006		>=0.590	85.71	68.75		>=0.676	78.57	81.25	82.59
	Haptoglobin		-0.0010982	0.626										
	Hemopexin													
COPD v Healthy														
	Individual													
n = 30	Ceruloplasmin	-7.004728	0.0079928	0.005	0.0003	>=794.79mcg/ml	>=0.343	85.71	75.0	>=972.71mcg/ml	>=0.684	64.29	93.75	86.61
	Haptoglobin	-3.684739	1.104089	0.034	0.013	>=2.96mg/ml	>=0.398	76.92	56.25	>=3.15mg/ml	>=0.448	61.54	68.75	73.56
	Hemopexin	4.024576	-0.004047	0.093	0.073									
	α-2-Macroglobulin	-1.084965	0.0003458	0.535	0.531									
	Combination													
	Ceruloplasmin	-9.850005	0.0073615	0.013	0.0004		>=0.342	84.62	68.75		>=0.590	69.23	87.50	88.94
	Haptoglobin		1.063352	0.072										
	Ceruloplasmin	***-0.8243742***	***0.0118777***	***0.007***	***<0.00001***		***>=0.426***	***92.86***	***75.0***		***>=0.480***	***85.71***	***87.50***	***91.07***
	Hemopexin		***-0.0092461***	***0.024***										
		0.9626632	1.420984	0.035	0.009		>=0.444	76.92	62.50		>=0.547	61.54	87.50	81.25
	Haptoglobin		-0.0055647	0.104										
	Hemopexin													
	α-2-Macroglobulin	3.912215	0.0000273	0.964	0.199									
	Hemopexin		-0.0040107	0.114										
COPD v Asthma														
	Individual													

n = 42	Ceruloplasmin	0.0630528	-0.0007368	0.576	0.575									
	Haptoglobin	1.20993	-0.5029125	0.102	0.078									
	Hemopexin	4.72965	-0.0052496	0.032	0.02	<=1090.89mg/ml	>=0.270	85.71	57.14	<=1009.75mg/ml	>=0.361	64.29	75.0	71.68
	α-2-Macroglobulin	-2.927178	0.0008705	0.066	0.050	>=2518.45mg/ml	>=0.324	85.71	75.0	>=2689.38mg/ml	>=0.358	64.29	82.14	76.28
	Combination													
	α-2-Macroglobulin	-0.9782056	0.000794	0.092	0.045		>=0.289	84.62	64.29		>=0.382	61.54	78.57	74.73
	Haptoglobin		-0.4664416	0.137										
	α-2-Macroglobulin	2.392418	0.0007116	0.133	0.019		>=0.314	78.57	67.86		>=0.435	57.14	78.57	75.51
	Hemopexin		-0.0047623	0.058										
	α-2-Macroglobulin	***2.530823***	***0.0006869***	***0.147***	***0.046***		***>=0.285***	***92.31***	***64.29***		>=0.355	76.92	75.0	***76.10***
	Haptoglobin		***-0.3750563***	***0.251***										
	Hemopexin		***-0.0034578***	***0.193***										
	α-2-Macroglobulin	***2.668874***	***0.0007084***	***0.134***	***0.048***		>=0.338	78.57	71.43		***>=0.493***	***50.0***	***89.29***	***75.77***
	Hemopexin		***-0.0047018***	***0.063***										
	Ceruloplasmin		***-0.0003201***	***0.829***										

COPD = chronic obstructive pulmonary disease; AUC = area under the curve.

TABLE E5. ANALYSIS OF DIAGNOSTIC VALUES OF MARKER AND MARKER IN COMBINATIONS IN SECONDARY VALIDATION POPULATION.

Comparison	Marker combination	Logistic regression				*Minimal false negatives/rule disease out*				*Minimal false positives/ rule disease in*				AUC (%)
		Constant	*Coefficient*	*p value*	*Model p value*	*Cut off measurement*	*Cut-off, Predicted value*	*Sensitivity*	*Specificity*	*Cut off measurement*	*Cut-off, Predicted value*	*Sensitivity*	*Specificity*	
Asthma v Healthy	*Individual*													
n = 36	Ceruloplasmin	-7.266969	0.0085751	0.008	0.0007	>=744.93 mcg/ml	>=0.293	85.71	68.18	>=898.31 mcg/ml	>=0.607	50.0	95.45	82.47
	Haptoglobin	-3.976472	0.5013821	0.014	0.002	>=6.36 mg/ml	>=0.313	85.71	54.55	>=8.70 mg/ml	>=0.596	50.0	86.36	79.87
	Hemopexin	-9.184255	0.0187517	0.004	0.0002	>=402.35 mg/ml	>=0.162	100.00	50.0	>=479.65 mg/ml	>=0.453	69.23	86.36	82.17
	α-2-Macroglobulin	0.1041666	-0.1956642	0.677	0.672	<=2.89 mg/ml	>=0.386	64.29	65.22	<=2.85 mg/ml	>=0.388	57.14	69.57	53.73
	Combination													
	Ceruloplasmin	-10.4399	0.0087462	0.022	0.0002		>=0.342	85.71	72.73		>=0.561	64.29	95.45	87.01
	Haptoglobin		0.434046	0.046										
	Ceruloplasmin	-11.97446	0.0052252	0.074	0.0001		>=0.226	92.31	68.18		>=0.421	69.23	81.82	89.16
	Hemopexin		0.015727	0.024										
		-22.76312	***1.248297***	***0.014***	***<0.00001***		***>=0.289***	***100.0***	***90.91***		***>=0.457***	***92.31***	***95.45***	96.85
	Haptoglobin		***0.0276208***	***0.015***										
	Hemopexin													
COPD v Healthy	*Individual*													
n = 44	Ceruloplasmin	-3.5269	0.0045915	0.032	0.014	>=692.70 mcg/ml	>=0.414	81.82	50.0	>=800.27 mcg/ml	>=0.537	54.55	72.73	71.07
	Haptoglobin	-1.58826	0.245888	0.047	0.034	>=5.17 mg/ml	>=0.422	77.27	40.91	>=7.81 mg/ml	>=0.582	50.0	81.82	65.91
	Hemopexin	-4.232852	0.0093937	0.019	0.009	>=462.65 mg/ml	>=0.523	72.73	86.36	>=517.64 mg/ml	>=0.652	50.0	95.45	73.76
	α-2-Macroglobulin	-2.46742	0.6973598	0.038	0.016	>=3.29 mg/ml	>=0.456	72.73	69.57	>=3.79 mg/ml	>=0.544	50.0	91.3	71.15
	Combination													
	Ceruloplasmin	-4.246169	0.0039499	0.079	0.016		>=0.460	72.73	59.09		>=0.573	63.64	81.82	75.41
	Haptoglobin		0.1850704	0.150										
	Ceruloplasmin	-6.018214	0.0034743	0.110	0.008		>=0.521	72.73	77.27		>=0.617	63.64	90.91	76.65
	Hemopexin		0.0074654	0.071										
		-4.623185	***0.1674288***	***0.200***	***0.014***		>=0.439	72.73	63.64		***>=0.567***	***59.09***	***95.45***	73.97
	Haptoglobin		***0.007912***	***0.061***										
	Hemopexin													
	α-2-Macroglobulin	***-6.03456***	***0.6182299***	***0.090***	***0.006***		***>=0.412***	***86.36***	***68.18***		***>=0.631***	***54.55***	***95.45***	80.17
	Hemopexin		***0.0086636***	***0.039***										

Mixed v Healthy	**Individual**													
n = 36	Ceruloplasmin	-3.238282	0.0037112	0.119	0.092	>=694.91 mcg/ml	>=0.341	71.43	50.0	>=785.22 mcg/ml	>=0.420	50.0	72.73	64.61
	Haptoglobin	-5.139459	0.6587639	0.009	0.0005	>=6.64 mg/ml	>=0.318	85.71	68.18	>=7.74 mg/ml	>=0.490	71.43	81.82	81.49
	Hemopexin	-8.831491	0.0187215	0.010	0.002	>=435.08 mg/ml	>=0.335	85.71	63.64	>=454.53 mg/ml	>=0.420	64.29	81.82	79.87
	α-2-Macroglobulin	-2.941042	0.7262588	0.089	0.071	>=3.07 mg/ml	>=0.330	71.43	56.52	>=3.50 mg/ml	>=0.401	50.0	82.61	68.63
	Combination													
	Ceruloplasmin	-6.705737	0.0024152	0.468	0.002		>=0.321	92.86	59.09		>=0.488	64.29	81.82	80.84
	Haptoglobin		0.6268258	0.016										
	Ceruloplasmin	-9.419774	0.0017056	0.552	0.007		>=0.321	92.86	68.18		>=0.464	64.29	86.36	81.49
	Hemopexin		0.017201	0.022										
		-12.00865	***0.0006532***	***0.866***	***0.0007***		>=0.277	92.86	59.09		***>=0.530***	***71.43***	***95.45***	86.04
	Ceruloplasmin		***0.0156691***	***0.069***										
			0.5881633	***0.030***										
	Hemopexin													
	Haptoglobin													
	α-2-macroglobulin	***-15.55849***	***1.120276***	***0.106***	***0.0002***		***>=0.295***	***100.0***	***77.27***		***>=0.435***	***92.86***	***95.45***	94.16
	Hemopexin		***0.0153898***	***0.078***										
	Haptoglobin		***0.63758***	***0.029***										
COPD v Asthma	**Individual**													
n = 36	Ceruloplasmin	1.67445	-0.0014039	0.414	0.408									
	Haptoglobin	1.521738	-0.1376352	0.323	0.310									
	Hemopexin	2.222555	-0.0033549	0.346	0.334									
	α-2-Macroglobulin	-2.734693	0.9372688	0.038	0.013	>=3.29 mg/ml	>=0.585	72.73	71.43	>=3.68 mg/ml	>=0.672	54.55	92.86	71.43
	Combination													
	Ceruloplasmin	2.242203	-0.0010095	0.575	0.510									
	Haptoglobin		-0.1169128	0.422										
	Ceruloplasmin	3.0697	-0.0011649	0.506	0.502									
	Hemopexin		-0.0030206	0.404										
		3.373483	-0.2144594	0.193	0.247									
	Haptoglobin		-0.0022269	0.557										
	Hemopexin													
	α-2-Macroglobulin	-1.485353	1.281106	0.024	0.008		>=0.556	72.73	64.29		>=0.621	68.18	92.86	79.55
	Haptoglobin		-0.304815	0.088										
	α-2-Macroglobulin	-1.175074	0.9805878	0.044	0.028		>=0.523	81.82	61.54		>=0.667	59.09	84.62	74.48
	Hemopexin		-0.0031468	0.417										
	α-2-Macroglobulin	-0.9186874	1.211829	0.026	0.012		>=0.585	72.73	71.43		>=0.665	68.18	92.86	78.25
	Ceruloplasmin		-0.0031181	0.118										
	α-2-Macroglobulin	***-0.1642794***	***1.898634***	***0.019***	***0.003***		>=0.512	81.82	61.54		***>=0.686***	***68.18***	***100.0***	84.27
	Haptoglobin		***-0.5863676***	***0.034***										
	Hemopexin		***-0.0015585***	***0.749***										
	α-2-Macroglobulin	***-0.0511131***	***1.439518***	***0.019***	***0.011***		***>=0.520***	***86.36***	***64.29***		>=0.641	63.64	85.71	79.87
	Haptoglobin		***-0.2743187***	***0.153***										
	Ceruloplasmin		***-0.002518***	***0.234***										
	α-2-Macroglobulin	0.0666635	1.279262	0.030	0.024		>=0.477	81.82	61.54		>=0.684	68.18	84.62	78.67

	Hemopexin		-0.0023254	0.576										
	Ceruloplasmin		-0.0030108	0.141										
Mixed v Asthma	**Individual**													
n = 28	Ceruloplasmin	2.635702	-0.0031127	0.183	0.154									
	Haptoglobin	-0.225885	0.0273526	0.888	0.888									
	Hemopexin	3.613586	-0.0070479	0.193	0.177									
	α-2-Macroglobulin	-4.068517	1.232145	0.047	0.025	>=2.94 mg/ml	>=0.392	78.57	64.29	>=3.45 mg/ml	>=0.544	57.14	71.43	70.92
	Combination													
	Ceruloplasmin	1.894628	-0.0035794	0.144	0.294									
	Haptoglobin		0.1383806	0.528										
	Ceruloplasmin	5.38602	-0.0025878	0.260	0.202									
	Hemopexin		-0.0062152	0.271										
		4.881399	-0.141323	0.567	0.341									
	Haptoglobin		-0.0071849	0.186										
	Hemopexin													
	α-2-Macroglobulin	-5.300792	1.308788	0.046	0.069		>=0.340	85.71	64.29		>=0.490	64.29	78.57	77.04
	Haptoglobin		0.1190857	0.586			or	or	or					
							>=0.450	78.57	71.43					
	α-2-Macroglobulin	-3.045245	1.231603	0.083	0.063		>=0.380	78.57	61.54		>=0.602	57.14	84.62	75.82
	Hemopexin		-0.0018068	0.773										
	α-2-Macroglobulin	***-1.67291***	***1.150547***	***0.068***	***0.045***		***>=0.368***	***85.71***	***57.14***		***>=0.583***	***57.14***	***92.86***	75.00
	Ceruloplasmin		***-0.0024947***	***0.302***							***or***	***or***	***or***	
											>=0.509	***64.29***	***71.43***	
	α-2-Macroglobulin	-2.289593	1.202923	0.094	0.133									
	Haptoglobin		-0.067749	0.803										
	Hemopexin		-0.0019742	0.755										
	α-2-Macroglobulin	-3.341737	1.276604	0.063	0.069		>=0.390	85.71	71.43		>=0.503	71.43	78.57	78.57
	Haptoglobin		0.2225828	0.371										
	Ceruloplasmin		-0.0031672	0.206										
	α-2-Macroglobulin	-1.267479	1.18816	0.100	0.089		>=0.428	71.43	61.54		>=0.517	64.29	76.92	73.63
	Hemopexin		-0.0011687	0.858										
	Ceruloplasmin		-0.0022912	0.337										
Mixed v COPD	**Individual**													
n = 36	Ceruloplasmin	0.3701332	-0.0009998	0.567	0.558									
	Haptoglobin	-1.829842	0.1760592	0.243	0.227									
	Hemopexin	-0.1913662	-0.0005365	0.884	0.884									
	α-2-Macroglobulin	0.3487275	-0.2102742	0.474	0.458									
	Combination													
	Ceruloplasmin	-0.6409984	-0.0019024	0.326	0.285									
	Haptoglobin		0.225127	0.157										
	Ceruloplasmin	0.3984816	-0.0009929	0.578	0.843									
	Hemopexin		-0.0000701	0.985										
		-0.5461736	0.2702018	0.162	0.316									
	Haptoglobin		-0.0041748	0.370										
	Hemopexin													
	α-2-Macroglobulin	-0.9540765	-0.2721244	0.366	0.308									
	Haptoglobin		0.1971568	0.196										

α-2-Macroglobulin	0.6977482	-0.2143809	0.471	0.747
Hemopexin		-0.0006869	0.853	
α-2-Macroglobulin	0.8839842	-0.1852103	0.538	0.689
Ceruloplasmin		-0.0007668	0.664	
α-2-Macroglobulin	0.794288	-0.3306576	0.309	0.322
Haptoglobin		0.3095113	0.119	
Hemopexin		-0.0049681	0.302	
α-2-Macroglobulin	-0.0225753	-0.2367757	0.441	0.369
Haptoglobin		0.238172	0.136	
Ceruloplasmin		-0.0016801	0.388	
α-2-Macroglobulin	1.022907	-0.1882617	0.535	0.861
Hemopexin		-0.0003211	0.933	
Ceruloplasmin		-0.0007321	0.686	

COPD = chronic obstructive pulmonary disease; AUC = area under the curve.

TABLE E6. CLINICAL DATA FOR THE DIAGNOSTIC IMMUNOASSAY VALIDATION (DIVA) GROUP

Group Description	1 Healthy Controls	2 Stable Asthma	3 COPD	P value
Number of patients	16	28	14	
Age (Years) [a]	45.2 ± 14.3	50.4 ± 13.2	[§]*65.5 ± 7.3	0.0001
Sex (Male/ Female) [c]	7/9	13/15	8/6	0.737
Smoking, n(%)[c]:			*	0.012
Never	7 (43.8%)	14 (50.0%)	1 (7.1%)	
Ex	9 (56.3%)	14 (50.0%)	11 (78.6%)	
Current	0	0	2 (14.3%)	
Pack years [a]	20.3 ± 17.4	15.9 ± 14.5	[§]*67.7 ± 31.4	<0.0001
Atopy, n(%)[c]	6 (37.5%)	*24 (85.7%)	[§]5 (35.7%)	0.001
%predicted FEV_1 [a,e]	97.2 ± 9.0	*76.9 ± 18.3	*71.2 ± 14.7	<0.0001
%predicted FEV_1 [a,f]	100.4 ± 8.6	*85.9 ± 14.4[g]	*72.7 ± 14.4	<0.0001
%predicted FVC[a,e]	101.4 ± 11.6	91.8 ± 15.5	*85.2 ± 14.8	0.011
FEV_1/FVC %[a,e]	79.0± 8.0	*68.2 ± 9.3	*62.5 ± 7.6	<0.0001
PD_{15} (mL) [b, d]	N/A	5.1 (2.7, 8.8)	4.8 (2.2, 8.4)	0.941
D_{LCO} [a]	N/A	85.5 ± 14.6	65.3 ± 13.2	0.01
ICS use, n(%)[c]	0	27 (96.4%)	[§]0	0.0003
ICS (µg beclomethasone equivalents /day) [b]	N/A	1406 ± 1121	N/A	
ACQ score	N/A	1.13 ± 0.82	N/A	
GOLD Classification, n (%)				
I	N/A	N/A	4 (28%)	
II	N/A	N/A	10 (71,4%)	
III	N/A	N/A	0 (0%)	
IV	N/A	N/A	0 (0%)	
Sputum Cell Counts				
Quality [b]	18.5 (17, 20)	19 (17, 20)	19 (17, 20)	0.926
Total cell count x 10^6/ml [b]	3.5 (2.4, 6.7)	4.1 (2.3, 9.7)	5.7 (2.9, 10.4)	0.728
Neutrophils% [b]	28.7 (17.8, 54.8)	48.9 (35.3, 74.5)	57.0 (53.5, 64.5)	0.065
Eosinophils% [b]	0 (0, 0.25)	*1.5 (0.5, 2.75)	[§]0.25 (0, 0.75)	0.0001
Macrophages% [b]	62.0 (38.3, 79.87)	42.8 (22.8, 54.8)	40.5 (31.8, 44.5)	0.066
Lymphocytes% [b]	0.75 (0.25, 1.25)	0.5 (0.25, 0.75)	1.0 (0.75, 1.5)	0.083
Columnar epithelial% [b]	2.25 (0.5, 5.5)	2.0 (0.5, 4.0)	1.75 (0.75, 2.5)	0.815
Squamous% [b]	2.3 (1.7, 5.7)	3.6 (0.99, 7.2)	2.2 (0.25, 3.9)	0.190

[a]Values are Mean ± SD, oneway ANOVA; [b]Values are median (interquartile range), Kruskall-Wallis test; [c]Chi Square or Fisher's exact test; [d] PD_{15} is provocation dose resulting in 15% drop in baseline FEV_1 expressed as geometric mean (log SD); FEV_1 is forced expiratory volume in 1 second either pre[e]- or post-[f] bronchodilator ([g] data only available for 20/28 individuals); FVC is forced vital capacity; D_{LCO} is carbon monoxide diffusing capacity; ICS is inhaled corticosteroids. Bonferroni post hoc test significant compared to: *Healthy Controls; [§]Stable Asthma. Asthma: symptoms with fully reversible variable airflow obstruction[airway hyperresponsiveness and/or increased bronchodilator reversibility]. COPD: symptoms, incomplete reversibility of airflow obstruction [postbronchodilator FEV<80%predicted].

TABLE E7 BIOMARKER LEVELS BY COMORBIDITIES

Biomarker	No Diabetes N= 66	Diabetes N=6	P value
Ceruloplasmin, mcg/ml [b]	785(668, 906)	717 (695, 749)	0.328
Haptoglobin, mg/ml [a]	7.0 ± 2.6	8.3 ± 3.0	0.255
Hemopexin, mcg/ml [a]	469 ± 95.5	493 ± 39.6	0.549
α-2-macroglobulin, mg/ml [b]	3.2 (2.8, 3.8)	3.8 (2.4, 4.1)	0.962
	BMI<=30 N=54	BMI>30 N=19	
Ceruloplasmin mcg/ml [b]	800 (668, 898)	732 (680, 875)	0.754
Haptoglobin, mg/ml [a]	6.9 ± 2.8	7.6 ± 2.1	0.354
Hemopexin, mcg/ml [a]	470 ± 94	476 ± 90	0.806
α-2-macroglobulin, mg/ml [b]	3.2 (2.8, 3.7)	3.0 (2.7, 4.1)	0.880
	No Cardiac disease N=60	Cardiac disease n=12	
Ceruloplasmin, mcg/ml [b]	793 (666, 910)	721 (681, 783)	0.257
Haptoglobin, mg/ml [a]	6.9 ± 2.7	8.1 ± 2.3	0.158
Hemopexin, mcg/ml [a]	474 ± 95	460 ± 83	0.632
α-2-macroglobulin, mg/ml [b]	3.1 (2.8, 3.7)	3.7 (3.2, 4.1)	0.181
	No Liver disease N=67	Liver disease N=5	
Ceruloplasmin, mcg/ml [b]	766 (668, 898)	800 (708, 801)	0.938
Haptoglobin, mg/ml [a]	7.2 ± 2.7	7.0 ± 3.0	0.867
Hemopexin, mcg/ml [a]	477 ± 85	400 ± 158	0.072
α-2-macroglobulin, mg/ml [b]	3.2 (2.8, 3.8)	3.7 (2.7, 5.0)	0.600
	No Psychiatric disorder, n=55	Psychiatric disorder, n=17	
Ceruloplasmin, mcg/ml [b]	785 (680, 913)	708 (652, 820)	0.210
Haptoglobin, mg/ml [a]	7.2 ± 2.5	6.9 ± 3.2	0.626
Hemopexin, mcg/ml [a]	464 ± 88	493 ± 103	0.262
α-2-macroglobulin, mg/ml [b]	3.1 (2.7, 3.6)	3.7 (3.5, 4.5)	0.031

[a]Values are Mean ± SD, Student's tests; [b]Values are median (interquartile range), Wicoxon rank sum test;

FIGURE LEGENDS

Figure E1. Reproducibility of plasma sample preparation and 2D-DIGE. Correlation of protein spot ratios between (A) the same sample labelled with Cy3 or Cy5; (B) the same plasma sample independently immunodepleted and TCA precipitated; and (C) the same sample run on a different gel one week apart.

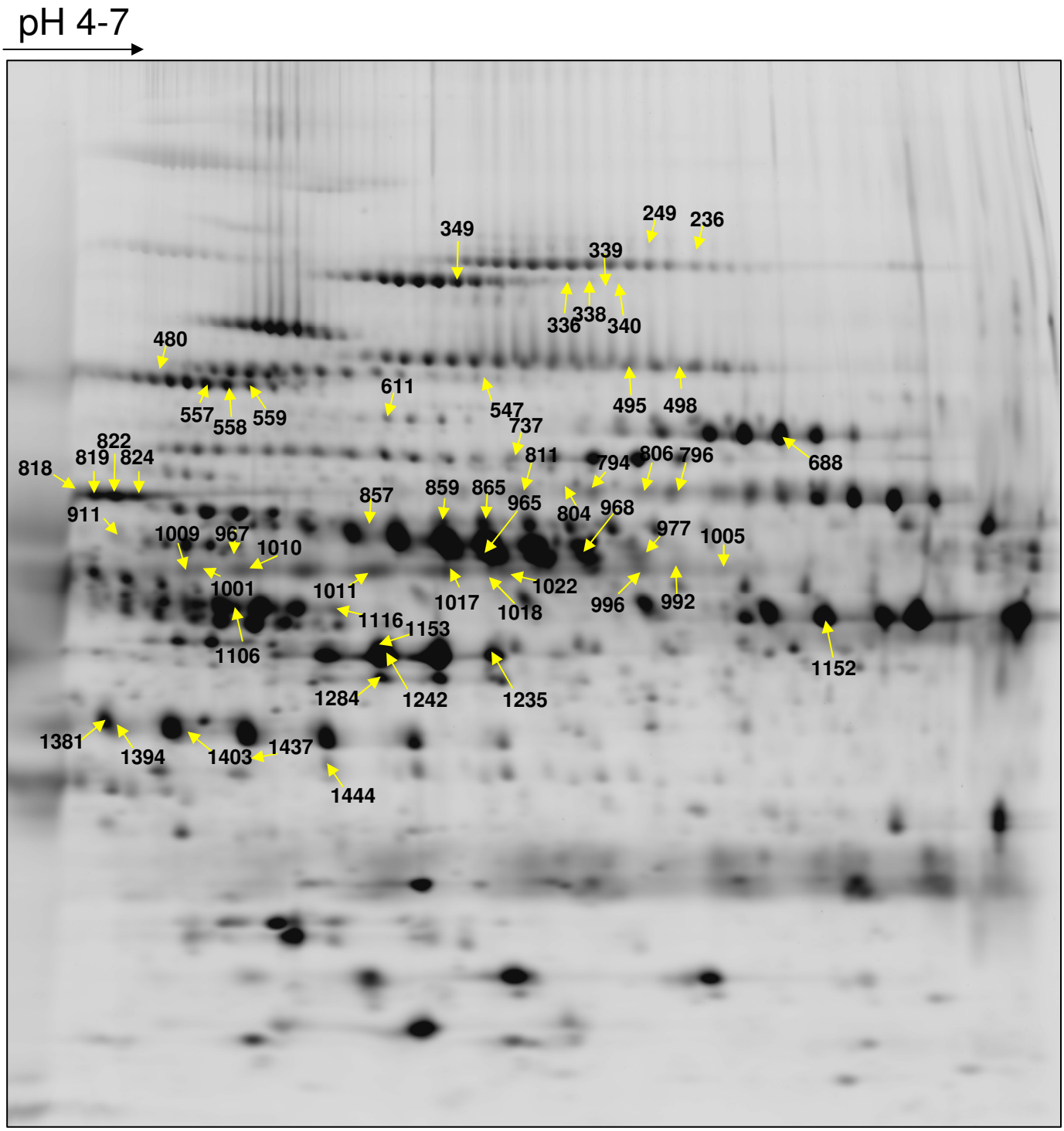
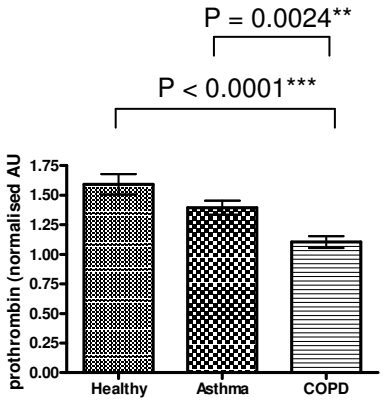
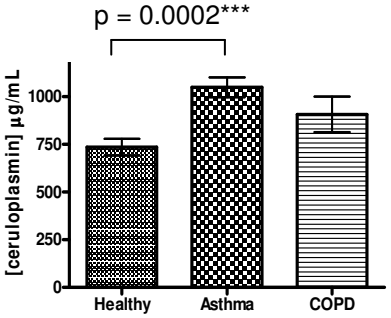

pH 4-7
349
249
236
339
336
338
340
480
611
547
557
558
559
495
498
737
822
819
824
818
811
794
806
796
688
911
857
859
865
965
804
968
977
1005
1009
967
1010
1001
1011
1017
1022
1018
996
992
1116
1153
1106
1152
1284
1242
1235
1381
1394
1403
1437
1444

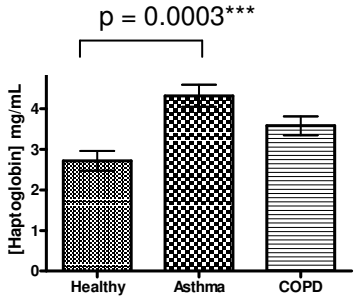
Figure 1.



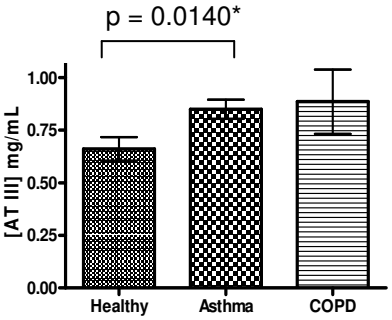
A



B



C



D

Figure 2.

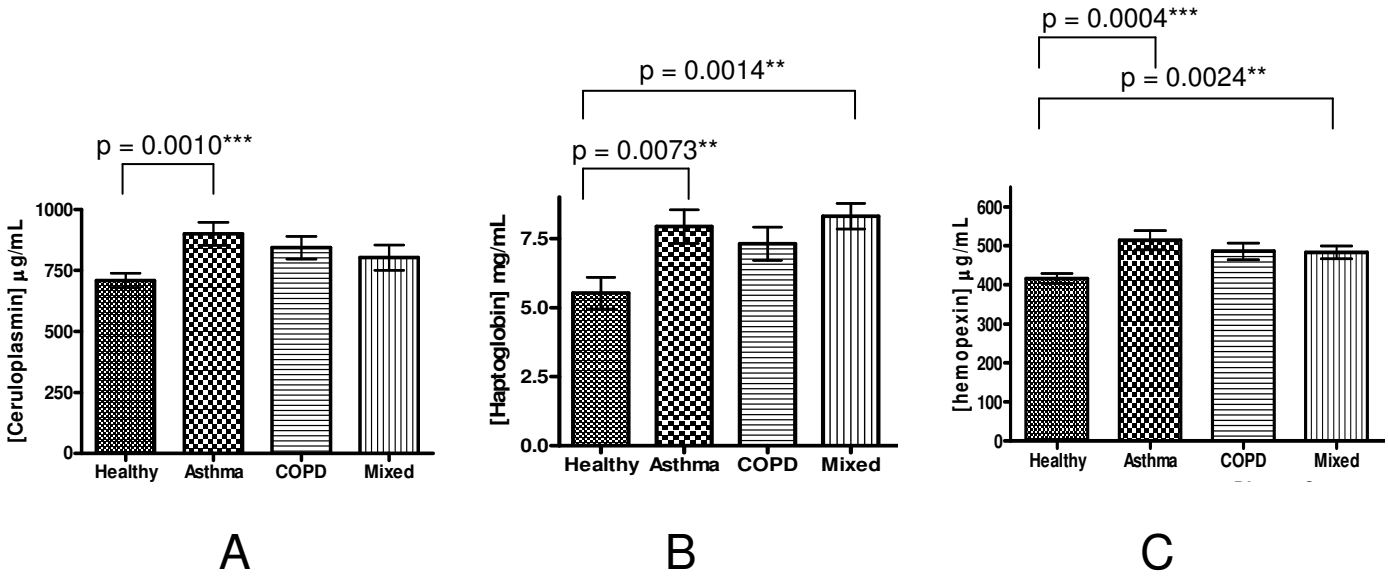

p = 0.0010***
[Ceruloplasmin] µg/mL
1000
750
500
250
0
Healthy
Asthma
COPD
Mixed
A
p = 0.0014**
p = 0.0073**
[Haptoglobin] mg/mL
7.5
5.0
2.5
0.0
Healthy
Asthma
COPD
Mixed
B
p = 0.0004***
p = 0.0024**
[hemopexin] µg/mL
600
500
400
300
200
100
0
Healthy
Asthma
COPD
Mixed
C

Figure 3.

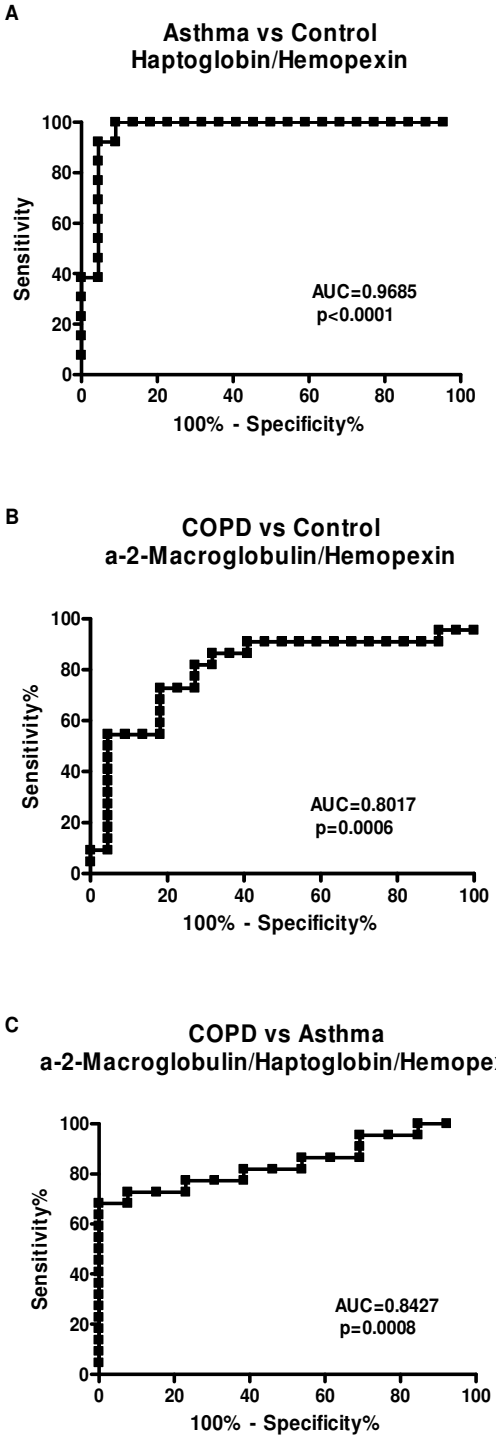

A
Asthma vs Control
Haptoglobin/Hemopexin
Sensitivity
100% - Specificity%
AUC=0.9685
p<0.0001
B
COPD vs Control
a-2-Macroglobulin/Hemopexin
Sensitivity%
100% - Specificity%
AUC=0.8017
p=0.0006
C
COPD vs Asthma
a-2-Macroglobulin/Haptoglobin/Hemopexin
Sensitivity%
100% - Specificity%
AUC=0.8427
p=0.0008

Figure 4.

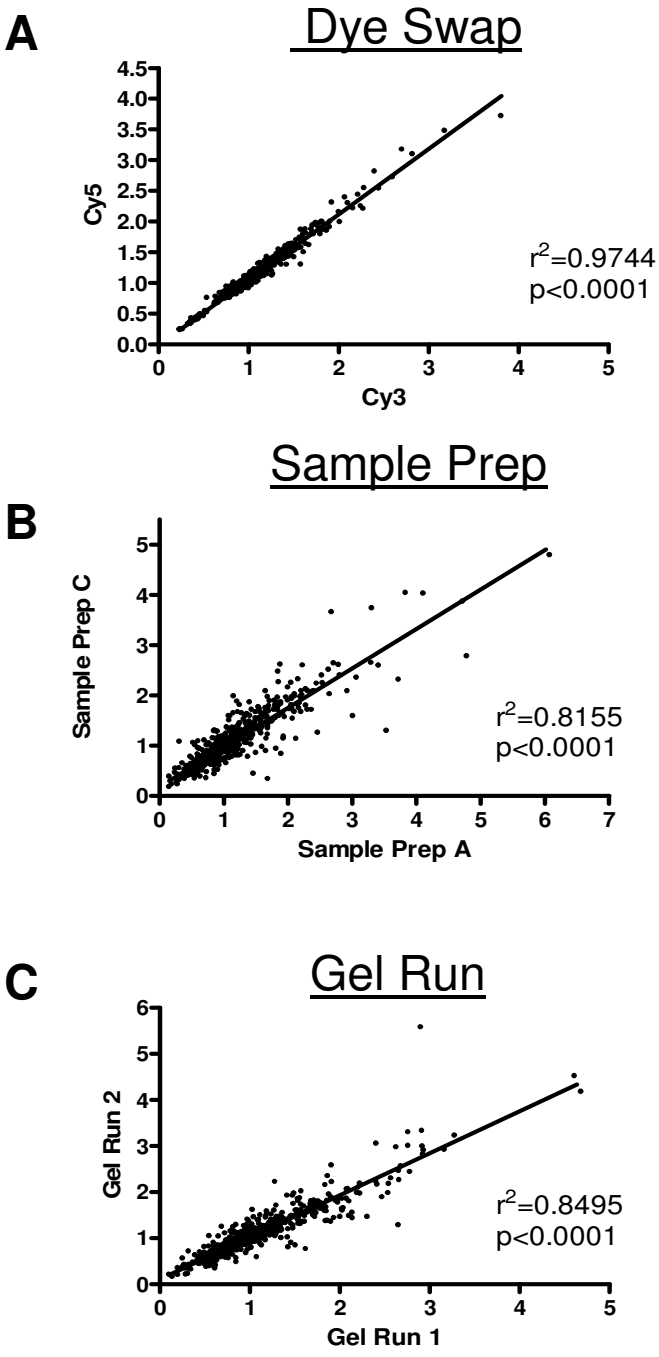

A
Dye Swap
Cy5
Cy3
r2=0.9744
p<0.0001
B
Sample Prep
Sample Prep C
Sample Prep A
r2=0.8155
p<0.0001
C
Gel Run
Gel Run 2
Gel Run 1
r2=0.8495
p<0.0001

Figure E1

Identification of Novel Diagnostic Biomarkers for Asthma and Chronic Obstructive Pulmonary Disease

Nicole M. Verrills[1,4,7†], Jennifer A. Irwin[2,4,5,7†], Xiao Yan He[2,4,5,7], Lisa G. Wood[1,4,5], Heather Powell[4,5], Jodie L Simpson[2,4,5], Vanessa M. McDonald[2,3,4,5], Alistair Sim[1,4] and Peter G. Gibson[2,4,5,6,7].

[1]School of Biomedical Sciences and Pharmacy, [2]School of Medicine and Public Health and [3]School of Nursing and Midwifery, Faculty of Health, University of Newcastle, NSW, Australia; [4]Hunter Medical Research Institute, Newcastle, NSW, Australia; [5]Department of Respiratory and Sleep Medicine, John Hunter Hospital, New Lambton, NSW, Australia; [6]Woolcock Institute of Medical Research, Sydney NSW, Australia; [7]CRC for Asthma and Airways, Australia.

[†]These authors have made an equal contribution to this manuscript.

Corresponding author: Prof. Peter G. Gibson, Department of Respiratory and Sleep Medicine, John Hunter Hospital, Lookout Rd, New Lambton 2305, Australia Ph: +61 2 49213470; Fax: +61 2 49855850; Email: Peter.Gibson@hnehealth.nsw.gov.au

This work was supported by the CRC for Asthma and Airways.

Short running heading: Diagnostic Biomarkers of Respiratory Disease.

Descriptor number: 3.06

Word Count (body): 4013

CONTRIBUTORSHIP

Nicole M. Verrills: 2D-DIGE study design, spot analysis and interpretation.

Jennifer A. Irwin: Data acquisition, design, analysis and interpretation for 2D-DIGE and two validation studies; writing and compilation of the manuscript.

Xiao Yan He: Interpretation of data and revision of manuscript.

Lisa G. Wood: Involvement in study conception, coordination of sample collection and the acquisition/compilation of clinical data for the discovery and primary validation population.

Heather Powell: Logistic regression analysis and statistical analysis of all clinical data.

Jodie L Simpson: Coordination of sample collection and the acquisition/compilation of clinical data for the secondary validation population.

Vanessa M. McDonald: Coordination of sample collection and the acquisition/compilation of clinical data for the secondary validation population.

Alistair Sim: Project conception.

Peter G. Gibson: Project conception, study design, data interpretation and writing/revision of manuscript.

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

The respiratory diseases asthma and chronic obstructive pulmonary disease (COPD) have many similar symptoms and are typically diagnosed by performing a number of clinical tests to assess an individual's lung function and response to reliever medication. To date, some studies have attempted to identify biomarkers of COPD or asthma, however, no study has attempted to identify non-invasive, blood-based, diagnostic biomarkers that can discriminate between healthy controls, asthmatics and individuals with COPD.

What This Study Adds to the Field

Using a proteomics approach, we have identified a panel of four blood-based biomarkers that, when used in combination, can discriminate between healthy controls, asthmatics and individuals with COPD and has the potential to be a valuable tool in the clinical diagnosis of respiratory disease. These markers implicate the anti-inflammatory iron metabolism pathways in the pathogenesis of asthma and COPD.

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org.

ABSTRACT

Rationale: Proteomics may identify a useful panel of biomarkers for identification of asthma and chronic obstructive pulmonary disease (COPD).

Objectives: To conduct an unsupervised analysis of peripheral blood proteins in well characterised subjects with asthma and COPD, and identify and validate a biomarker panel for disease discrimination.

***Methods*:** Two-dimensional difference gel electrophoresis (2D-DIGE) was used to separate plasma proteins from healthy controls, stable asthmatics and individuals with COPD. Candidate protein markers were identified using matrix assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF MS) and subsequently validated in two populations via immunoassay. A panel of four biomarkers was selected and their ability to distinguish between groups assessed in isolation and in combination in 2 separate validation populations.

Results: Seventy-two protein spots displayed significantly different expression levels between the three subject groupings ($p<0.05$). 58 were positively identified representing 20 unique proteins. A panel of four biomarkers (α-2 macroglobulin, haptoglobin, ceruloplasmin and hemopexin) was able to discriminate with statistical significance between the clinical groups of asthma, COPD, and controls, and these results were confirmed in a second clinical population of older adults with airflow obstruction.

Conclusions: Proteomics has identified novel biomarkers for asthma and COPD, and shown that the iron metabolism pathways and acute phase response may be involved in the pathogenesis of airway disease. The panel of peripheral blood biomarkers has the potential to become an extremely useful addition to the clinical diagnosis and management of respiratory disease.

Word count: 239

Key words: Plasma proteomics, biomarkers, asthma, COPD, logistic regression.

INTRODUCTION

Word count: 434

The obstructive airway diseases asthma and chronic obstructive pulmonary disease (COPD) are significant and increasing health problems throughout the world. The management of these diseases could be improved by better diagnosis and recognition, and better understanding of their pathogenesis. New technologies for investigating human diseases now offer significant potential to address the need for better diagnosis and improved understanding of asthma and COPD. Proteomics can simultaneously identify multiple proteins associated with different disease states (1) and potentially discover novel proteins not previously associated with particular disease states. Several studies have utilised proteomics for the discovery of protein changes in lung tissue from mouse models of asthma (2-6), and bronchoalveolar lavage fluid (BALF) (7) and $CD3^+$ T-lymphocytes (8) from asthmatic and healthy subjects. However, for diagnostic and prognostic purposes, identification of biomarkers in readily obtainable samples, such as blood, is preferred. A selected analysis of multiple known blood markers in COPD using protein array methodology has shown the potential for proteomics (9), whereby a panel of biomarkers were found to associate with COPD patients versus healthy controls. This type of study used a predefined array of markers, and is potentially limited by the pool of analytes available on the protein array and by our current knowledge of disease pathology. An open, unsupervised proteomic study should identify a broader panel of candidate markers, including proteins not previously associated with respiratory disease. The need for such unbiased approaches has recently been emphasised and called for in the chronic inflammatory diseases of asthma and COPD (10).

There is also general agreement that a panel of independent disease-related proteins considered in aggregate should be less prone to the influence of genetic and environmental 'noise' than is the level of a single marker protein (11), and that proteomics has the power to

identify such panels of proteins in a high-throughput manner. There remains however a need to apply methods to quantify the added benefit of biomarker panels for disease assessment. Thus identification of a panel of biomarkers that are differentially expressed between asthmatics and COPD patients, and between patients with and without these airway diseases is required. In this study we have utilised 2D-difference gel electrophoresis (2D-DIGE) to conduct an unsupervised analysis of circulating proteins in well characterised subjects with asthma and COPD, and applied logistic regression to evaluate the power of combining markers compared with the markers in isolation. The biomarker panel was then applied to a second validation population with obstructive airway disease to demonstrate the power of this approach.

Some of the results of this study have been previously reported in the form of an abstract (12).

METHODS

Subjects

We confirmed 2D-DIGE reproducibility (on-line supplement, Figure E1) and conducted a proteomic discovery study and 2 validation studies. The discovery study assessed mutually exclusive clinical groups (n=43, Table 1 and Table E3) of stable non-smoking asthma (n=21), COPD (n=5), and matched healthy controls (n=17). The biomarkers were then validated by immunoassay in this population, supplemented by 7 asthmatics and 9 COPD patients (n=58, Table E6). A clinical validation study then assessed the biomarker panel in older adults (>55 years) with obstructive airway disease (FEV_1 <80% predicted and FEV_1/VC<70%) consecutively recruited from the clinic (n=50) (Table 2, asthma (n = 14), COPD (n = 22), overlap asthma/COPD (n = 14)), and in age matched controls (n=23) recruited by advertisement (12). The study was approved by institutional ethics committees and subjects gave written informed consent.

The proteomic discovery study and immunoassay validation study recruited subjects from mutually exclusive clinical groups of asthma, COPD, and controls without airway disease. None of the subjects were current smokers. Asthma was defined as a compatible history of episodic symptoms responsive to asthma treatment, together with airway hyperresponsiveness to hypertonic (4.5%) saline. Asthmatics had fully reversible airflow obstruction. and DLCO was >70% predicted. The COPD group were selected to have COPD (GOLD stage II or greater) with predominant airway disease with minimal emphysema (D_{LCO}>60% predicted) and no asthma. They were ex-smokers with a significant smoking history, compatible symptoms and a doctors diagnosis of COPD, and incompletely reversible airflow obstruction (FEV_1<80% predicted, FEV_1/FVC < 0.7 postbronchodilator) and a negative bronchodilator response (BDR). The clinical validation study applied these same definitions, and subjects with features of both asthma (variable airflow obstruction) and COPD (incomplete reversibility of airflow obstruction) were labelled mixed asthma/COPD.

Sample Preparation & 2D-DIGE & Mass Spectrometry

Platelet depleted plasma was processed on a ProteomeLab IgY-12 LC2 column (Beckman Coulter, CA, USA), precipitated via TCA/acetone and protein quantified (2D quant; GE Biosciences). Protein (100µg) was labelled with Cy3 or Cy5 dyes (GE Biosciences) and separated on 24cm pH 4-7 IPG strips followed by 4-18% polyacrylamide gels in the second dimension, was visualised on a Typhoon 9410 Variable Mode Imager as described (13). A pooled internal control consisting of individuals from each clinical group (n=10) was labelled with Cy2 and included on every gel. Image analysis and comparison was performed using the Batch Processor and Biological Variation Analysis modules of the DeCyder software version 6.5 (GE Healthcare, Australia). Average expression ratios of individual protein spots were compared using one way analysis of variance (1-ANOVA) and proteins with an expression ratio of ±1.2 (assigned following reproducibility analysis) and p<0.05 were then manually

inspected for densitometric Gaussian distribution and match quality. Proteins of interest were excised from preparative 2D gels, trypsin digested, and analysed on an Ettan MALDI-ToF/Pro or an Applied Biosystems 4700 Proteomics Analyser with TOF/TOF optics in MS mode as previously described (13). Data was submitted to the database search program Mascot (Matrix Science Ltd, London, UK) for protein identification as described (13).

Immunoassay

Serum IgA, complement factor H, haptoglobin, α-2 macroglobulin, ceruloplasmin, hemopexin and antithrombin III concentrations were assayed with commercially available ELISAs. For Western blotting, proteins were separated on 10% or 4-18% polyacrylamide gels, transferred to nitrocellulose, and probed with anti-prothrombin , anti-fibrinogen gamma chain , or anti-inter-alpha-trypsin inhibitor heavy chain H4 (anti-ITI-H4) primary antibodies followed by appropriate HRP-linked secondary antibodies. Reactive bands were visualised using ECL and a Fujifilm Luminescent Image Analyser LAS-300 and quantified using Multi Gauge V3.0 software. The chemiluminescent intensity was normalised to a healthy control serum sample run on every gel.

Statistical Analysis

Differential expression between groups was assessed using Students t-test (GraphPad Prism 4 for Windows, GraphPad Software Inc.) with Bonferoni corrected p values. Logistic regression (Stata 9, StataCorp) was used to calculate the predicted value of an individual having the disease based on their level of a single marker or combination of markers. Receiver-Operating Characteristic (ROC) curves were generated, and sensitivity, specificity and the area under the curve were calculated. Significance was accepted when $p<0.05$. The possibility of biomarker serum concentrations varying with inhaled corticosteroid (ICS) use and age was assessed using either a simple logistic regression or Spearman's rank correlation, as appropriate.

RESULTS

Differential Protein Expression Between Healthy Controls, Asthmatics and COPD Patients.

Plasma proteins were separated by 2D-DIGE for 43 subjects across the 3 clinical groups: healthy controls (n=17), asthma (n=21) and COPD (n=5). Over 1900 proteins were detected on each gel, 72 proteins displayed significantly altered expression ($p<0.05$, 1-ANOVA), and 58 were identified by mass spectrometry (Table 3; Figure 1). While 58 individual protein spots were differentially expressed, a number of these represented variant charge and/or molecular weight isoforms of the same protein (e.g. 10 isoforms of IgA (Table 3)), thus a total of 20 unique proteins were identified as candidate markers. These proteins clustered in groups of known function including iron metabolism (ceruloplasmin, haptoglobin, hemopexin), the coagulation cascade (α-2-macroglobulin, prothrombin, fibrinogen gamma, fibrin beta), immunoglobulins (IgA, IgM), and complement pathways (complement factor H, complement factor B).

Validation of Candidate Markers in the Diagnostic Immunoassay Validation Assessment (DIVA) Group

In order to assess the performance of the markers identified by 2D-DIGE analysis, we determined the relative expression of selected marker candidates in serum by immunoassay (ELISA or immunoblotting). Markers were prioritised for validation according to three criteria (i) multiple isoforms identified with consistent trends observed, (ii) biologically plausible in the context of respiratory disease and (iii) the availability of commercially available ELISA kits or primary antibodies. Serum concentrations of ceruloplasmin ($p = 0.0002$), haptoglobin ($p = 0.0003$) and antithrombin III ($p = 0.0140$) were significantly higher in asthmatics compared with healthy controls. Additionally, serum levels of prothrombin

(high molecular weight isoform) were lower in COPD patients than asthmatics ($p = 0.0024$) and healthy controls ($p < 0.0001$) (Figure 2).

Validation of Candidate Markers in the Clinical Validation Population

We then proceeded to validate the candidate biomarkers in an unselected clinical population consisting of older adults with asthma, COPD, and asthma-COPD overlap, and who also exhibited significant co morbidity. Serum concentrations of ceruloplasmin ($p = 0.0010$), haptoglobin ($p = 0.0073$) and hemopexin ($p = 0.0004$) were significantly higher in asthmatics compared with healthy controls. When compared to controls, the COPD patients showed trends for elevations in ceruloplasmin, α-2 macroglobulin, haptoglobin and hemopexin. Haptoglobin ($p = 0.0014$) and hemopexin ($p = 0.0024$) exhibited significantly elevated serum levels in individuals with overlap asthma-COPD compared with healthy controls. None of the markers investigated showed statistically significant differences between COPD patients and individuals with overlap asthma-COPD and asthmatics (Figure 3).

Analysis of a panel of Proteomic Biomarkers

The performance of a biomarker panel was evaluated in the two populations and the candidate markers ceruloplasmin, α-2 macroglobulin, haptogobin and hemopexin were selected for use in a combinatorial analysis due to their statistically significant differences in serum marker levels between disease groups and their consistent performance over the two distinct populations.

DIVA Group: Several biomarker combinations were capable of distinguishing between all three clinical groups (asthma, COPD, healthy control) in a manner superior to any one marker in isolation (Table 4, Table E4). The combination of ceruloplasmin and haptoglobin was the best for discriminating between asthmatics and healthy controls, whereas serum ceruloplasmin and hemopexin concentrations in combination best distinguished between COPD patients and healthy controls. For differential diagnosis between COPD patients and

asthmatics several marker combinations were able to discriminate between the two disease states. The combination α-2 macroglobulin and either haptoglobin or hemopexin provided a good balance of sensitivity and specificity whereas α-2 macroglobulin, haptoglobin and hemopexin delivered the best sensitivity (92%) and α-2 macroglobulin, hemopexin and ceruloplasmin gave the best specificity (89.3%). It is interesting to note that the serum concentration of α-2 macroglobulin alone performed almost as well, indicating that it was the predominant contributor to the power of the combinations tested.

Clinical Validation Population:. The combination of haptoglobin and hemopexin was the best for discriminating between asthmatics and healthy controls in this population (Table 5, Table E5, Figure 4). Ceruloplasmin either alone, or combined with haptoglobin also performed well. α-2 macroglobulin and hemopexin concentrations in combination best distinguished between COPD patients and healthy controls. The combination of ceruloplasmin and hemopexin also demonstrated statistically significant discrimination for COPD and controls, in agreement with the analysis of the primary validation group.

For differential diagnosis between COPD patients and asthmatics the combination α-2 macroglobulin, haptoglobin and ceruloplasmin delivered the best sensitivity whereas α-2 macroglobulin, haptoglobin and hemopexin provided the best specificity. The combination of the markers α-2 macroglobulin, haptoglobin and hemopexin was the best for discriminating between individuals with asthma-COPD overlap and healthy controls α-2 macroglobulin and ceruloplasmin serum levels in combination could discriminate asthmatics and individuals with overlap asthma/COPD No marker combinations or markers in isolation were able to discriminate between COPD patients and individuals with overlap asthma-COPD with statistical significance.

Effect of Co-morbid Conditions in the Clinical Validation Population

A literature review identified 5 medical conditions as potential confounders of the blood based marker diagnosis panel: hepatobiliary disease (14), ischemic coronary disease (15), obesity (16, 17), endocrine and metabolic disorders (18, 19) as well as psychiatric conditions (20). Table E7 summarises a co-morbidity analysis of the clinical validation population and indicates no significant effect of these co-morbid conditions on the marker levels, apart from an effect of psychiatric disease on levels of α-2 macroglobulin.

Potential Effect of ICS Use and Age on Biomarker Serum Concentrations

For the DIVA group, stable asthmatics (but not COPD patients) were prescribed ICS. The potential effect of ICS use on serum marker concentrations was assessed and it was found that ICS dosage did not significantly alter serum marker concentrations and is thus unlikely to be a cause of the observed elevation of serum marker concentrations in this study. [Hemopexin: p = 0.5135; haptoglobin: p=0.083; ceruloplasmin: p = 0.6723; α-2 macroglobulin: p = 0.6761].

For the clinical validation population many COPD patients were also using ICS. In order to evaluate the effect of ICS use on serum marker levels in COPD patients, individuals from the clinical validation and DIVA (none using ICS) groups were combined and reclassified into two groups (with and without ICS). Logistic regression analysis identified no statistically significant effect of ICS use on marker concentration..

Additionally, we evaluated the potential effect of age on serum marker concentrations and found that there was no significant correlation between age and serum marker concentration for the DIGE, DIVA and clinical validation groups (p>0.05 for all panel markers).

DISCUSSION

This proteomic discovery programme has identified a panel of protein markers whose serum concentrations are significantly altered in asthma and COPD compared with age and sex matched healthy controls, and whose function points to novel mechanistic pathways

indicating the involvement of the iron metabolism pathways, complement pathways, and the coagulation cascade in asthma and COPD. In addition, we have identified a panel of 4 serum biomarkers that includes ceruloplasmin, haptoglobin, hemopexin and α-2-macroglobulin that can be used in combination to accurately identify asthma and COPD, and have validated these results in a second clinical population of older adults with obstructive airway diseases, including asthma and COPD. The proteins in the diagnostic biomarker panel are all involved in the regulation of inflammation, and usually function as anti-inflammatory proteins.

We used an unbiased analysis design in well characterised groups of patients with asthma and COPD to discover differentially expressed proteins in these groups. Ten of the twenty candidate markers identified were subsequently validated in serum and a biomarker panel including the markers ceruloplasmin, α-2 macroglobulin, haptoglobin and hemopexin provided significant discrimination between subject groups in both validation studies. Although individual markers can differentiate between particular clinical groupings, logistic regression analysis has shown that the consideration of marker combination for each specific comparison yields vastly superior performance. ROC curves constructed for each comparison are of high quality (Figure 4) and specific cut-off points may be chosen to tailor the test for either maximum sensitivity or specificity, depending upon the diagnostic requirements. Indeed, apart from mixed airways disease and COPD (for which no combination can differentiate) cut-points may be chosen where sensitivity and specificity are well balanced (Tables 4 and 5). Thus our study has identified a panel of highly discriminatory proteins that could be extremely useful in a clinical context.

The biomarker panel comprises three positive acute-phase proteins (α-2 macroglobulin, ceruloplasmin and haptoglobin) and one type II acute-phase protein (hemopexin). These are predominantly liver-synthesised proteins that can have important anti-inflammatory activity through inhibition of oxidative stress, and iron sequestration resulting in antimicrobial

activity. As such, they may function to modulate the systemic inflammatory response to inflammation (21) and be involved in tissue repair through fibrosis and angiogenesis.

The acute phase response is an innate body defence observed during infection, physical trauma, malignancy and tissue damage that aims to minimise on-going tissue damage by isolating and destroying infective agents whilst activating repair processes. It is an antigen non-specific, innate response which aims to eliminate microbes and hence prevent infection. The innate immune response involves the recruitment and activation of macrophages and leukocytes that release inflammatory cytokines upon recognition of a PAMP. These cytokines travel through the bloodstream and stimulate hepatocytes in the liver to synthesize and secrete acute-phase proteins which we have in this case identified as differentially expressed between our clinical groupings in circulating blood.

Ceruloplasmin, also called ferroxidase, is a multi-functional, copper protein synthesised primarily in the liver and by activated macrophages. It has important roles in iron homeostasis, inflammation, and it has antimicrobial activity via regulation of iron availability to microorganisms. Serum levels have previously been reported to be elevated (22, 23) or unchanged in children with allergic asthma (24). Engstrom reported that ceruloplasmin and haptoglobin were weakly correlated with lung function in COPD, and were associated with an increased future risk of hospitalisation in COPD (25). Our results extend these observations by showing elevated ceruloplasmin in adults with asthma and COPD, as well as asthma-COPD overlap. In addition, we show that the elevation of ceruloplasmin is part of an increase in iron metabolism proteins as part of asthma and COPD, probably as an anti-inflammatory response to the airway inflammation that characterises these conditions.

Haptoglobin and hemopexin were other iron–related proteins that were altered in this study.

The inhibition of heme release from globin by haptoglobin and sequestration of heme by hemopexin suppress hemoglobin-mediated oxidative stress, attenuates endothelial

cytotoxicity and protects cells from heme toxicity. Iron sequestration may also have antimicrobial effects by depriving microbes of essential iron. Additionally, hemoglobin and its derivative heme are often released into tissue compartments where there is infection and inflammation, in the presence of degrading blood, and hemoglobin synergizes with multiple TLR agonists to induce release of high levels of tumor necrosis factor and interleukin-6 from macrophages, an effect that is attenuated by hemopexin (26). Hemopexin also down-regulates LPS-induced proinflammatory cytokine release from macrophages (27). Hemopexin has not been previously been associated with asthma or COPD. However, a 2D-DIGE based proteomic study of BAL in sarcoidosis patients, chronic beryllium disease (CBD) patients and controls showed a significant difference in BAL hemopexin concentration between controls and CBD patients. CBD is a granulomatous disorder that can lead to chronic lung inflammation and fibrosis (28).

Haptoglobin can be expressed by eosinophils, and variable serum levels have been reported in asthma, where both elevated (29) and reduced (30) serum haptoglobin levels are described. Increases in haptoglobin are seen in uncontrolled asthma, such as asthma exacerbation (31) and 24 hours after allergen challenge in late responders (32). In asthma, haptoglobin has also been correlated with FEV_1 (29). As part of its tissue repair function, haptoglobin can induce differentiation of fibroblast progenitor cells into lung fibroblasts (33), and angiogenesis, potentially implicating haptoglobin in remodelling and fibrosis in asthma and COPD. Haptoglobin has not previously been linked to COPD.

α-2 macroglobulin binds host or foreign peptides and particles, thereby serving as a humoral defense barrier against pathogens in the plasma and tissues of vertebrates. It interacts with and captures virtually any proteinase including serine, cysteine, aspartic and metalloproteinases e.g. gelatinase (MMP-2 and MMP-9), both self and foreign, suggesting a function as a unique "panproteinase inhibitor" (34). Its structure contains a 35 amino acid

"bait" region. Proteinases binding and cleaving the bait region become bound to α- 2 macroglobulin and the proteinase-α-2 macroglobulin complex is recognised by macrophage receptors and cleared from the system (35). α-2 macroglobulin provides negative feedback on the inflammatory response by inhibiting thrombin (coagulation) and plasmin (fibrinolysis). It has been studied in airway secretions as a marker of plasma exudation, and is increased in sputum samples in asthma and COPD (36). Plasma levels may be increased in asthma (37), and are normal in emphysema but reduced in chronic bronchitis (38).

In agreement with these proteins being positive acute-phase proteins, their serum levels are elevated in the asthmatic and COPD groups (independent of age and ICS use) in both our validation populations relative to the healthy controls in our study. More importantly, the differential expression of the markers in our panel makes sense in the context of asthma and COPD which both manifest inflammatory and fibrotic components during their progression.

Some limitations to our study relate to subject selection in the discovery population. The COPD group was small in number, and selected to have airway predominant disease. The primary reason for studying airway predominant COPD was to rigorously test the hypothesis that there would be differential markers between asthma and COPD reflective of small airway inflammation which is a relevant lesion in both asthma and COPD. This approach eliminates the potentially confounding effects of airspace disease. By designing the study this way, we believe that the information about potential disease mechanisms is not confounded by the site of pathology. In addition, because COPD is a heterogeneous disease, we thought it was important to minimize this heterogeneity by studying a recognized and clinically relevant group of COPD subjects. A larger group or one with more emphysema may identify additional proteins that are associated with COPD, and this would be useful future work. Similarly, the COPD subjects in the discovery group were not using inhaled corticosteroid (ICS), whereas, most of the patients in the clinical validation group were prescribed ICS. This

has the potential to influence results. We conducted some analyses to test for these effects and none were apparent, however this requires further study. The 2D-DIGE technology is not well suited to analysis of small molecular weight proteins, eg <10kDa. This means that many cytokines would not be detected by this approach, and additional methods such as used by PintoPlata are useful here (9).

The strengths of the results in this study come from the approach used and the combination of a panel of markers to use for disease discrimination. A panel of independent disease related biomarkers, as can be identified by proteomics, is generally considered to be more powerful and less prone to the influence of genetic and environmental 'noise' than a single marker protein (11). For example, Rai *et al.* identified three potential biomarkers that could differentiate ovarian cancer from healthy individuals and compared their performance against the tumour marker, cancer antigen 125 (CA125) (39). Each biomarker individually did not out-perform CA125, however the combination of two of the new biomarkers together with CA125 significantly improved their performance (39, 40). Similarly, we show here that the combination of protein biomarkers significantly improves the performance as a diagnostic marker than each individual protein alone.

Proteomic analysis, utilizing high-resolution 2D-gel electrophoresis coupled with mass spectrometry, is a powerful means to identify differential protein expression between biological samples. However, a major limitation in traditional 2D-gel technology is the reproducibility, and thus statistical comparison of protein expression between individual gels is difficult. A recent advance in this area has come from the introduction of Cy dye fluorophores for pre-labelling of protein samples. Two-dimensional difference gel electrophoresis (2D-DIGE) technology adds a quantitative component to conventional 2D-gel analyses, allowing for comparison of protein expression changes across multiple samples simultaneously without gel-to-gel variation, and hence with statistical confidence (41, 42).

Our reproducibility study (Online supplement) clearly shows that this technology, together with immunodepletion of abundant plasma proteins, can reproducibly separate thousands of proteins. Combined with well defined clinical groups and advanced statistical analyses, we have shown that this technology is a powerful tool for the identification of novel disease biomarkers. As the biomarkers are detectable in blood, a readily obtainable biological sample, and reagents are currently available for testing the abundance of these proteins, this panel of biomarkers has the potential to become an extremely useful addition to the clinical diagnosis and management of respiratory disease.

Acknowledgements

This work was supported by the CRC for Asthma and Airways. NMV and LGW were supported by Postdoctoral Fellowships from the National Health and Medical Research Council, Australia. PGG was supported by a Practitioner Fellowship from the National Health and Medical Research Council, Australia. The authors would like to acknowledge the technical assistance of Michelle Gleeson and Anne-Marie Gibson. This research has been facilitated by access to the Australian Proteome Analysis Facility established under the Australian Government's Major National Research Facilities Program.

References

1. Verrills, N. M. Clinical proteomics: present and future prospects. *Clin Biochem Rev* 2006; 27(2): 99-116.
2. Fehniger TE, Sato-Folatre JG, Malmstrom J, Berglund M, Lindberg C, Brange C, Lindberg H, Marko-Varga G. Exploring the context of the lung proteome within the airway mucosa following allergen challenge. *J Proteome Res* 2004; 3:307-320.
3. Fajardo I, Svensson L, Bucht A, Pejler G. Increased levels of hypoxia-sensitive proteins in allergic airway inflammation. *Am J Respir Crit Care Med* 2004; 170:477-484.
4. Jeong H, Rhim T, Ahn MH, Yoon PO, Kim SH, Chung IY, Uh S, Kim SI, Park CS. Proteomic analysis of differently expressed proteins in a mouse model for allergic asthma. *J Korean Med Sci* 2005; 20:579-585.
5. Roh GS, Shin Y, Seo SW, Yoon BR, Yeo S, Park SJ, Cho JW, Kwack K. Proteome analysis of differential protein expression in allergen-induced asthmatic mice lung after dexamethasone treatment. *Proteomics* 2004; 4:3318-3327.
6. Houtman R, Krijgsveld J, Kool M, Romijn EP, Redegeld FA, Nijkamp FP, Heck AJ, Humphery-Smith I. Lung proteome alterations in a mouse model for nonallergic asthma. *Proteomics* 2003; 3:2008-2018.
7. Wu J, Kobayashi M, Sousa EA, Liu W, Cai J, Goldman SJ, Dorner AJ, Projan SJ, Kavuru MS, Qiu Y, Thomassen MJ. Differential proteomic analysis of bronchoalveolar lavage fluid in asthmatics following segmental antigen challenge. *Mol Cell Proteomics* 2005; 4:1251-1264.

8. Jeong HC, Lee SY, Lee EJ, Jung KH, Kang EH, Kim JH, Park EK, Lee SH, Uhm CS, Cho Y, Shin C, Shim JJ, Kim HK, In KH, Kang KH, Yoo SH. Proteomic analysis of peripheral T-lymphocytes in patients with asthma. *Chest* 2007; 132:489-496.

9. Pinto-Plata V, Toso J, Lee K, Park D, Bilello J, Mullerova H, De Souza MM, Vessey R, Celli B. Profiling serum biomarkers in patients with COPD: associations with clinical parameters. *Thorax* 2007; 62: 595-601.

10. Auffray C, Adcock IM, Chung KF, Djukanovic R, Pison C, Sterk PJ. An integrative systems biology approach to understanding pulmonary diseases. *Chest* 2010; 137:1410-1416.

11. Anderson L. Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J Physiol* 2005; 563:23-60.

12. McDonald VM, Simpson JL, Higgins I, Gibson PG. Multidimensional assessment of older people with asthma and COPD: clinical management and health status. *Age Ageing*, in press.

13. Verrills NM, Liem NL, Liaw TY, Hood BD, Lock RB, Kavallaris M. Proteomic analysis reveals a novel role for the actin cytoskeleton in vincristine resistant childhood leukemia--an in vivo study. *Proteomics* 2006; 6(5):1681-1694.

14. Meliconi R, Parracino O, Facchini A, Morselli-Labate AM, Bortolotti F, Tremolada F, Martuzzi M, Miglio F, Gasbarrini G. Acute phase proteins in chronic and malignant liver diseases. *Liver* 1988; 8(2):65-74.

15. Correale M, Brunetti ND, De Gennaro L, Di Biase M. Acute phase proteins in atherosclerosis. *Cardiovasc Hematol Agents Med Chem* 2008; 6:272-277.

16. Chiellini C, Santini F, Marsili A, Berti P, Bertacca A, Pelosini C, Scartabelli G, Pardini E, López-Soriano J, Centoni R, Ciccarone AM, Benzi L, Vitti P, Del Prato S,

Pinchera A, Maffei M. Serum haptoglobin: A novel marker of adiposity in humans. *J Clin. Endocrinol. Metab.* 2004; 89:2678-2683.

17. Cignarelli M, De Pergola G, Picca G, Sciaraffia M, Pannac Turrisi E, Giorgino R. Relationship of obesity and body fat distribution with ceruloplasmin serum levels. *Int J Obes* .1996: 20(9): 809-813.

18. Kim CH, Park JY, Choi CS, Kim YI, Chung YE, Lee MS, Hong SK, Lee KU. Elevated serum ceruloplasmin levels in subjects with metabolic syndrome: a population-based study. *Metabolism* 2002; 51(7):838-842.

19. Jonsson A, Wales JK. Blood glycoprotein levels in diabetes mellitus. *Diabetologia* 1976; 12(3):245-250.

20. Maes M, Delange J, Ranjan R, Meltzer HY, Desnyder R, Cooremans W, Scharpe S. Acute phase proteins in schizophrenia, mania and major depression: modulation by psychotropic drugs. *Psychiatry Res* 1997: 66(1):1-11.

21. Fink MP Hemopexin: newest member of the anti-inflammatory mediator club. *J Leukoc Biol* 2009; 86(2):203-4.

22. Vural H, Uzun K. Serum red blood cell antioxidant status in patients with bronchial asthma. 2000; *Can Respir J* 7(6): 476-480.

23. el-Kholi MS, Gas Allah MA, el-Shimi S, el-Baz F, el-Tayeb H, Abdel-Hamid MS. Zinc and copper status in children with bronchial asthma and atopic dermatitis. 1990; *J Egypt Public Health Assoc*. 65(5-6):657-668.

24. Di Toro R, Galdo Capotorti G, Gialanella G, Miraglia del Giudice M, Moro R, Perrone L. Zinc and copper status of allergic children. *Acta Paediatr Scand* 1987; 76(4) 612-617.

25. Engstrom G, Segelstorm N, Ekberg-Aronsson M, Nilsson PM, Lingarde F, Lofdahl CG. Plasma markers of inflammation and incidence of hospitalisations for COPD: results from a population-based study. *Thorax* 2009; 64(3):211-215.

26. Lin T, Kwak YH, Sammy F, He P, Thundivalappil S, Sun G, Chao W, Warren HS. Synergistic Inflammation Is Induced by Blood Degradation Products with Microbial Toll-Like Receptor Agonists and Is Blocked by Hemopexin. *J Infect Dis* 2010; Jul 9. (Epub ahead of print)

27. Liang X, Lin T, Sun G, Beasley-Topliffe L, Cavaillon JM, Warren HS. Hemopexin down-regulates LPS-induced proinflammatory cytokines from macrophages. *J Leukoc Biol* 2009; 86(2):229-35.

28. Silva, E. Bourin, S. Sabounchi-Schutt, F. Laurin, Y. Barker, E. Newman, L.Eriksson, H. Eklund, A. Grunewald, J. A quantitative proteomic analysis of soluble bronchoalveolar fluid proteins from patients with sarcoidosis and chronic beryllium disease. *Sarcoidosis Vasc Diffuse Lung Dis* 2007; 24(1):24-32.

29. Kauffmann FC, Frette I, Annesi M-F, Dore, Neukirch F. Relationships of haptoglobin level to FEV_1, wheezing, bronchial hyper-responsiveness and allergy. *Clin. Exp. Allergy* 1991; 21: 669-674

30. Piessens MF, Marien G, Stevens E. Decreased haptoglobin levels in respiratory allergy. *Clin. Allergy* 1984; 14: 287-293

31. Koh YY, Kim YW, Park JD, Oh JW. A comparison of serum haptoglobin levels between acute exacerbation and clinical remission in asthma. *Clin Exp Allergy* 1996; 10:1202-9.

32. Kim CK, Chung CY, Koh YY Changes in serum haptoglobin level after allergen challenge test in asthmatic children. *Allergy* 1998; 53(2):184-9.

33. Larsen K, Macleod D, Nihlberg K, Gurcan E, Bjermer L, Marko-Varga G, Westergren-Thorsson G. Specific haptoglobin expression in bronchoalveolar lavage during differentiation of circulating fibroblast progenitor cells in mild asthma. *J Proteome Res* 2006; 5:1479-1483.

34. Borth, W. α-2- macroglobulin, a multifunctional binding protein with targeting characteristics. *The FASEB Journal* 1992; 6:3345-3353.

35. Kurdowska A, Alden SM, Noble JM, Stevens MD, Carr FK (). Involvement of α-2-macroglobulin receptor in clearance of interleukin-8-α-2-macroglobulin complexes by human alveolar macrophages. *Cytokine* 2000; 12(7):1046-1053.

36. Schoonbrood DF, Lutter R, Habets FJ, Roos CM, Jansen HM, Out TA. Analysis of plasma-protein leakage and local secretion in sputum from patients with asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1994; 150(6 Pt 1):1519-27.

37. Gaillard MC, Reichberg SB, Nogueira CM, Kilroe-Smith TA. Differences in elastase-binding activity of alpha 1-protease inhibitor and alpha 2-macroglobulin for asthma patients and control subjects with various alpha1-protease inhibitor phenotypes. *Clin Chem* 1993; 39(4):675-9.

38. Burnett D, Stockley RA. Serum and sputum alpha 2 macroglobulin in patients with chronic obstructive airways disease. *Thorax* 1981; 36(7):512-6.

39. Rai AJ, Zhang Z, Rosenzweig J, Shih Ie M, Pham T, Fung ET, Sokoll LJ, Chan DW. Proteomic approaches to tumor marker discovery. *Arch Pathol Lab Med* 2002; 126:1518-1526.

40. Zhang Z, Bast RC, Jr, Yu Y, Li J, Sokoll LJ, Rai AJ., Rosenzweig JM, Cameron B, Wang YY, Meng XY, Berchuck A, Van Haaften-Day C, Hacker NF, de Bruijn HW,

van der Zee AG, Jacobs IJ, Fung ET, Chan, DW. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res* 2004; 64:5882-5890.

41. Von Eggeling F, Gawriljuk A, Fielder W, Ernst G, Claussen U, Klose J, Romer I. Fluorsecent dual colour 2D-protein gel electrophoresis for rapid detection of differences in protein pattern with standard image analysis software. *Int J Mol Med* 2001; 8:373-377.

42. Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Surrie I, Davison M. Validation and development of fluorescent two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* 2001; 1:377-396.

TABLE 1. CLINICAL CHARACTERISTICS OF PEOPLE IN THE 2D-DIGE MARKER DISCOVERY STUDY

Group Description	1 Healthy Controls	2 Stable Asthma	3 COPD	P value
Number of patients	17	21	5	
Age (Years) [a]	44.2 ± 14.4	48.1±12.7	§*65.7 ± 10.6	0.012
Sex (Male/ Female) [c]	8/9	11/10	2/3	1.0
Smoking, n(%)[c]:				0.127
Never	8 (47%	11 (52%)	0	
Ex	9 (53%)	10 (48%)	5 (100%)	
Pack years [a]	20.3 ± 17.4	20.7 ± 13.3	§*72.7 ± 36.9	0.003
Atopy, n(%)[c]	6 (35.3%)	*19 (90.5%)	§1 (20.0%)	<0.0001
%predicted FEV_1 [a,e]	97.7 ± 9.0	*81.3 ± 16.7	*65.0 ± 17.5	0.0001
%predicted FEV_1 [a,f]	101.0 ± 8.7	*85.9 ± 14.4	*69.0 ± 18.3	<0.0001
%predicted FVC[a,e]	101.3 ± 11.3	95.1 ± 14.9	*79.9 ± 10.9	0.010
FEV_1/FVC %[a,e]	79.5± 8.0	*70.1 ± 7.9	*63.8 ± 10.0	0.0003
PD_{15} (mL) [b,d]	N/A	5.3 (3.6, 15.1).	4.8 (4.7, 8.4)	0.926
D_{LCO} [a]	N/A	85.9 ± 15.0	65.6 ± 11.5	0.012
ICS use, n(%)[c]	N/A	21(100%)	0	<0.0001
ICS (µg beclomethasone equivalents /day) [b]	N/A	1464 ± 1228	N/A	
ACQ score	N/A	1.1 ± 0.8	N/A	
GOLD Classification, n (%)				
I	N/A	N/A	0 (0%)	
II	N/A	N/A	5 (100%)	
III	N/A	N/A	0 (0%)	
IV	N/A	N/A	0 (0%)	
Induced Sputum Cell Counts				
Quality [b]	19 (17, 20)	19 (17, 20)	19 (18, 20)	0.976
Total cell count x 10^6/ml [b]	3.96 (2.4, 6.7)	3.47 (1.6, 5.1)	6.03 (3.7, 10.1)	0.3224
Neutrophils% [b]	34.5 (17.8, 61.0)	42.0 (27.5, 49.3)	§64.5 (57.8, 67.8)	0.028
Eosinophils% [b]	0 (0, 0.25)	*1.0 (0.25, 6.5)	1.25 (0.25, 1.5)	0.002
Macrophages% [b]	59.8 (36.8, 79.8)	51.5 (41.3, 62.8)	§*31.8 (31.3, 32.0)	0.032
Lymphocytes% [b]	0.75 (0.25, 1.25)	0.5 (0.25, 0.75)	0.75 (0.5, 1.25)	0.551
Columnar epithelial% [b]	2.0 (0.5, 5.5)	2.25 (1.0, 6.5)	1.75 (0.25, 2.25)	0.442
Squamous% [b]	2.7 (1.7, 5.7)	4.5 (0.7, 7.2)	2.2 (0.99, 3.9)	0.453

[a]Values are Mean ± SD, oneway ANOVA; [b]Values are median (interquartile range), Kruskall-Wallis test; [c]Chi Square or Fisher's exact test; [d]PD_{15} is provocation dose resulting in 15% drop in baseline FEV_1 expressed as geometric mean (log SD); FEV_1 is forced expiratory volume in 1 second either pre[e]- or post-[f] bronchodilator; FVC is forced vital capacity; D_{LCO} is carbon monoxide diffusing capacity; ICS is inhaled corticosteroids; Bonferroni post hoc test significant compared to: *Healthy Controls; [§]Stable Asthma. Asthma: symptoms with fully reversible variable airflow obstruction [airway hyperresponsiveness and/or increased bronchodilator reversibility]. COPD: symptoms, incomplete reversibility of airflow obstruction [postbronchodilator FEV<80%predicted].

TABLE 2. CLINICAL DATA FOR THE CLINICAL VALIDATION POPULATION

Group Description	1 Healthy Controls	2 Stable Asthma	3 COPD	4 Overlap Asthma-COPD	P value
Number of patients	23	14	22	14	
Age (Years) [a]	49.9 ± 17.6	*67.7 ± 6.7	*68.0 ± 7.9	*68.7 ± 9.0	<0.0001
Sex (Male/ Female) [c]	11/12	8/6	8/14	5/9	0.566
Smoking, n(%)[c]:					0.759
Never	12 (52.2%)	7 (50%)	8 (36.4%)	5 (35.7%)	
Ex	11 (47.8%)	7 (50%)	13 (59.1%)	8 (57.1%)	
Current	0	0	1 (4.6%)	1 (7.1%)	
Pack years [a]	13.9 ± 15.6	39.5 ± 28.6	39.9 ± 36.9	37.2 ± 37.9	0.151
Atopy, n(%)[c]	6 (26.1%)	*11 (78.6%)	8 (36.4%)	9 (64.3%)	0.006
%predicted FEV_1 [a,e]	103.9 ± 13.6	*56.9 ± 22.4	*54.1 ± 21.5	*50.6 ±14.9	<0.0001
%predicted FEV_1 [a,f]	109.0 ± 13.8	*61.6 ± 23.8	*60.6 ± 22.7	*55.0 ± 17.1	<0.0001
%predicted FVC[a,e]	110.10 ± 14.3	*75.1 ± 18.6	*75.1 ± 22.6	*76.6 ± 14.8	<0.0001
FEV_1/FVC %[a,e]	77.5 ± 5.3	*59.2 ± 15.2	*55.9 ± 13.6	*52.1 ± 12.9	<0.0001
PD_{15} (mL) [b, d]	N/A	7.4 (3.5, 19.6)	5.9 (4.9, 8.7)	6.3 (2.9, 11.2)	0.935
D_{LCO} [a]	N/A	73.3 ± 17.3	72.5 ± 23.2	84.1 ± 21.2	0.323
ICS use, n(%)[c]	0	11 (78.6%)	21 (95.5%)	14 (100%)	0.08
ICS (µg beclomethasone equivalents /day) [b]	N/A	1600 ± 780	1460 ± 646	1657 ± 511	0.659
GOLD Classification, n (%)					
I	N/A	N/A	4 (20%)	N/A	
II	N/A	N/A	9 (45%)	N/A	
III	N/A	N/A	4 (20 %)	N/A	
IV	N/A	N/A	3 (15%)	N/A	
Sputum Cell Counts					
Quality [b]	18 (16, 19)	17 (16, 19)	18.5 (17.5, 19)	17.5 (16, 18)	0.250
Total cell count x 10^6/ml [b]	1.7 (1.3, 3.1)	4.0 (2.8, 8.5)	3.3 (1.9, 7.6)	3.7 (1.4, 9.5)	0.039
Neutrophils% [b]	28.5 (13.3, 54.5)	62.9 (45.5, 78.5)	*56.0 (35.5, 88.3)	*77.0 (45.8, 85.5)	0.003
Eosinophils% [b]	0.25 (0, 0.5)	*1.65 (0.5, 3.25)	*2.0 (0.75, 7.25)	*1.0 (0.25, 5.5)	0.0001
Macrophages% [b]	63.8 (41.5, 77.8)	29.9 (16.8, 52.0)	*15.0 (8.0, 44.0)	*17.5 (7.3, 37.3)	0.0003

Lymphocytes% [b]	0.75 (0.25, 1.75)	0 (0, 0.5)	0.25 (0, 1.0)	0.25 (0, 1.0)	0.062
Columnar epithelial% [b]	1.0 (0.25, 7.5)	1.5 (0.5, 2.0)	0.75 (0.25, 2.25)	1.0 (0.5, 2.25)	0.659
Squamous% [b]	4.5 (1.96, 8.3)	1.6 (0.7, 18.4)	2.7 (0.99, 8.9)	3.9 (2.2, 8.5)	0.725

[a]Values are Mean ± SD, oneway ANOVA; [b]Values are median (interquartile range), Kruskall-Wallis test; [c]Chi Square or Fisher's exact test; [d] PD_{15} is provocation dose resulting in 15% drop in baseline FEV_1 expressed as geometric mean (log SD); FEV_1 is forced expiratory volume in 1 second either pre[e]- or post-[f] bronchodilator; FVC is forced vital capacity; D_{LCO} is carbon monoxide diffusing capacity; ICS is inhaled corticosteroids. Bonferroni post hoc test significant compared to: *Healthy Controls. Asthma: symptoms and reversible variable airflow obstruction [airway hyperresponsiveness and/or increased bronchodilator reversibility]. COPD: symptoms, incomplete reversibility of airflow obstruction [postbronchodilator FEV<80%predicted]; Overlap asthma/COPD: symptoms with increased variability and incomplete reversibility of airflow obstruction.

TABLE 3. DIFFERENTIALLY EXPRESSED PROTEINS BETWEEN THE 3 DISEASE GROUPS (HEALTHY CONTROLS, ASTHMATICS AND COPD PATIENTS) IDENTIFIED BY 2D-DIGE AND MASS SPECTROMETRY.

Master spot no.	ANOVA p-value	Protein ID	Accession #	Differential Expression Ratio (*p<0.05, t-test)		
				Asthmatics / Controls	COPD / Controls	COPD / Asthmatics
Iron Metabolism						
480	0.002	Ceruloplasmin (ferroxidase)	P00450	1.01	-1.34*	-1.35*
1437	0.0066	Haptoglobin	P00738	1.94*	1.48	-1.31
1394	0.016	Haptoglobin	P00738	2.13*	1.59	-1.34
1444	0.022	Haptoglobin	P00738	1.83*	1.37	-1.33
1403	0.027	Haptoglobin	P00738	1.88*	1.59	-1.18
1381	0.03	Haptoglobin	P00738	2.13*	1.37	-1.55
968	0.024	Hemopexin	P02790	-1.2*	NA	NA
965	0.029	Hemopexin	P02790	-1.12	-1.23*	-1.1
977	0.033	Hemopexin	P02790	-1.34*	-1.2	-1.12
Coagulation cascade						
822	0.04	Prothrombin	P00734	-1.08	-1.24*	-1.15
824	0.0054	Prothrombin	P00734	-1.1	-1.31*	-1.19
818	0.016	Prothrombin	P00734	-1.14*	-1.26*	-1.1
819	0.018	Prothrombin	P00734	-1.13*	-1.19*	-1.05
1242	0.0019	Fibrinogen gamma	P02679	1.06	1.31*	1.24*
1235	0.03	Fibrinogen gamma	P02679	1.18	1.48*	1.26
1153	0.011	Fibrinogen gamma	P02679	1.25*	1.28*	1.02
859	0.0041	Heparin cofactor II	P05546	1.18*	-1.25	-1.47*
865	0.0097	Heparin cofactor II	P05546	1.1	-1.2*	-1.32*
611	0.037	Heparin cofactor II	P05546	1.31*	1.05	-1.24
911	0.0028	Protein S precursor	P07225	-1.28*	-1.19	1.07
1152	0.0084	Fibrin beta		1.2*	1.51*	1.26
559	0.025	ITI heavy chain H4	Q14624	-1.24*	-1.15	1.08
557	0.031	ITI heavy chain H4	Q14624	-1.14*	-1.25*	-1.09
558	0.049	ITI heavy chain H4	Q14624	-1.14	-1.29*	-1.13
1284	0.036	alpha-2-antiplasmin	P08697	1.19	1.27*	1.07
967	0.0085	alpha-2 antiplasmin	P08697	1.18	2.07*	1.75*
857	0.041	Histidine-rich glycoprotein	P04196	1.07	-1.22	-1.3*
1116	0.048	Antithrombin-III	P01008	-1.23*	-1.15	1.08
1106	0.019	Antithrombin-III	P01008	-1.18*	-1.22*	-1.04
737	0.042	Gelsolin	P06396	-1.19	-1.43*	-1.21
236	0.0034	Alpha-2 macroglobulin	P01023	-1.31*	NA	NA
495	0.022	Alpha-2-macroglobulin	P01023	-1.23*	-1.02	1.2

498	0.032	Alpha-2 macroglobulin	P01023	-1.34*	-1.17	1.14
249	0.041	Alpha-2-macroglobulin	P01023	-1.3*	NA	NA
Complement Pathways						
349	0.028	Complement factor H	P08603	1.3*	1.34	1.03
338	0.00017	Complement factor H	P08603	-1.53*	1.19	1.82*
336	0.0031	Complement factor H	P08603	-1.24*	1.59*	1.97*
340	0.0045	Complement factor H	P08603	-1.3*	1.27	1.66*
339	0.0078	Complement factor H	P08603	1.16	1.81*	1.56*
680	0.015	Complement factor B	P00751	-1.25	-1.5*	-1.2
688	0.029	Complement factor B	P00751	1.12	-1.2	-1.34*
547	0.032	Complement 3	P01024	-1.18	-1.76*	-1.49
691	0.0034	Complement component C4-A	P0C0L4	-1.29*	-1.59*	-1.23*
Immunoglobulins						
1001	0.0019	IgA	P01876	1.45*	2.38*	1.64*
992	0.0055	IgA	P01876	1.19	1.98*	1.66*
1009	0.0084	IgA	P01876	1.24	1.85*	1.49*
1005	0.016	IgA	P01876	1.33	2.11*	1.58
1017	0.017	IgA	P01876	1.2	1.94*	1.62*
1018	0.022	IgA	P01876	1.11	1.86*	1.67*
1022	0.022	IgA	P01876	1.16	1.86*	1.61*
996	0.024	IgA	P01876	-1.36	1.3	1.77*
1011	0.034	IgA	P01876	1.22	1.92*	1.57
1010	0.039	IgA	P01876	1.3	1.99*	1.53
806	0.011	IgM heavy chain	P01871	-2.17*	-2.04	1.07
796	0.02	IgM heavy chain	P01871	-1.8*	-2.68*	-1.49
794	0.026	IgM heavy chain	P01871	-2.09*	-2.2	-1.06
811	0.018	IgM heavy chain	P01871	-1.74*	-2.22*	-1.27
804	0.04	IgM heavy chain	P01871	-1.87*	-2.42	-1.3

COPD = chronic obstructive pulmonary disease; ITI heavy chain H4 = Inter-alpha-trypsin inhibitor heavy chain H4; IgA = immunoglobulin A; IgM = immunoglobulin M.

TABLE 4. ANALYSIS OF DIAGNOSTIC VALUES OF MARKER AND MARKER IN COMBINATIONS IN DIAGNOSTIC IMMUNOASSAY VALIDATION GROUP.

Comparison	Marker combination	*Logistic Regression Model p value*	*Minimal False Negatives*			*Minimal False Positives*			AUC (%)
			Cut-off, Predicted value	*Sensitivity*	*Specificity*	*Cut-off Predicted value*	*Sensitivity*	*Specificity*	
Asthma v Healthy	**Combination**								
n = 44	***Ceruloplasmin Haptoglobin***	***<0.00001***	***>=0.448***	***92.86***	***75.0***	***>=0.785***	***67.86***	***93.75***	***89.29***
	Ceruloplasmin Hemopexin	0.0006	>=0.442	89.29	62.5	>=0.753	67.86	93.75	83.04
	Haptoglobin Hemopexin	0.0006	>=0.590	85.71	68.75	>=0.676	78.57	81.25	82.59
COPD v Healthy	**Combination**								
n = 30	Ceruloplasmin Haptoglobin	0.0004	>=0.342	84.62	68.75	>=0.590	69.23	87.50	88.94
	Ceruloplasmin Hemopexin	***<0.00001***	***>=0.426***	***92.86***	***75.0***	***>=0.480***	***85.71***	***87.50***	***91.07***
	Haptoglobin Hemopexin	0.009	>=0.444	76.92	62.50	>=0.547	61.54	87.50	81.25
COPD v Asthma	**Combination**								
n = 42	α-2-Macroglobulin Haptoglobin	0.045	>=0.289	84.62	64.29	>=0.382	61.54	78.57	74.73
	α-2-Macroglobulin Hemopexin	0.019	>=0.314	78.57	67.86	>=0.435	57.14	78.57	75.51
	α-2-Macroglobulin Haptoglobin Hemopexin	***0.046***	***>=0.285***	***92.31***	***64.29***	>=0.355	76.92	75.0	***76.10***
	α-2-Macroglobulin Hemopexin Ceruloplasmin	***0.048***	>=0.338	78.57	71.43	***>=0.493***	***50.0***	***89.29***	***75.77***

COPD = chronic obstructive pulmonary disease; AUC = area under the curve.

TABLE 5. ANALYSIS OF THE DIAGNOSTIC VALUES OF MARKER AND MARKER IN COMBINATIONS IN THE CLINICAL VALIDATION POPULATION:

Comparison	Marker combination	*Logistic Regression Model p value*	*Minimal false negatives*			*Minimal false positives*			AUC (%)
			Cut-off Predicted value	*Sensitivity*	*Specificity*	*Cut-off Predicted value*	*Sensitivity*	*Specificity*	
Asthma v Healthy	**Combination**								
n = 36	Ceruloplasmin Haptoglobin	0.0002	>=0.342	85.71	72.73	>=0.561	64.29	95.45	87.01
	Ceruloplasmin Hemopexin	0.0001	>=0.226	92.31	68.18	>=0.421	69.23	81.82	89.16
	Haptoglobin Hemopexin	***<0.00001***	***>=0.289***	***100.0***	***90.91***	***>=0.457***	***92.31***	***95.45***	96.85
COPD v Healthy	**Combination**								
n = 44	Ceruloplasmin Haptoglobin	0.016	>=0.460	72.73	59.09	>=0.573	63.64	81.82	75.41
	Ceruloplasmin Hemopexin	0.008	>=0.521	72.73	77.27	>=0.617	63.64	90.91	76.65
	Haptoglobin Hemopexin	***0.014***	>=0.439	72.73	63.64	***>=0.567***	***59.09***	***95.45***	73.97
	α-2-Macroglobulin Hemopexin	***0.006***	***>=0.412***	***86.36***	***68.18***	***>=0.631***	***54.55***	***95.45***	80.17
Mixed v Healthy	**Combination**								
n = 36	Ceruloplasmin Haptoglobin	0.002	>=0.321	92.86	59.09	>=0.488	64.29	81.82	80.84
	Ceruloplasmin Hemopexin	0.007	>=0.321	92.86	68.18	>=0.464	64.29	86.36	81.49
	Ceruloplasmin Hemopexin Haptoglobin	***0.0007***	>=0.277	92.86	59.09	***>=0.530***	***71.43***	***95.45***	86.04
	α-2-Macroglobulin Hemopexin Haptoglobin	***0.0002***	***>=0.295***	***100.0***	***77.27***	***>=0.435***	***92.86***	***95.45***	94.16
COPD v Asthma	**Individual**								
n = 36	α-2-Macroglobulin	0.013	>=0.585	72.73	71.43	>=0.672	54.55	92.86	71.43
	Combination								
	α-2-Macroglobulin Haptoglobin	0.008	>=0.556	72.73	64.29	>=0.621	68.18	92.86	79.55
	α-2-Macroglobulin Hemopexin	0.028	>=0.523	81.82	61.54	>=0.667	59.09	84.62	74.48
	α-2-Macroglobulin Ceruloplasmin	0.012	>=0.585	72.73	71.43	>=0.665	68.18	92.86	78.25
	α-2-Macroglobulin Haptoglobin Hemopexin	***0.003***	>=0.512	81.82	61.54	***>=0.686***	***68.18***	***100.0***	84.27
	α-2-Macroglobulin Haptoglobin Ceruloplasmin	***0.011***	***>=0.520***	***86.36***	***64.29***	>=0.641	63.64	85.71	79.87
	α-2-Macroglobulin Hemopexin Ceruloplasmin	0.024	>=0.477	81.82	61.54	>=0.684	68.18	84.62	78.67
Mixed v Asthma	**Individual**								
n = 28	α-2-Macroglobulin	0.025	>=0.392	78.57	64.29	>=0.544	57.14	71.43	70.92
	Combination								
	α-2-Macroglobulin Haptoglobin	0.069	>=0.340 or >=0.450	85.71 or 78.57	64.29 or 71.43	>=0.490	64.29	78.57	77.04
	α-2-Macroglobulin Hemopexin	0.063	>=0.380	78.57	61.54	>=0.602	57.14	84.62	75.82
	α-2-Macroglobulin Ceruloplasmin	***0.045***	***>=0.368***	***85.71***	***57.14***	***>=0.583 or >=0.509***	***57.14 or 64.29***	***92.86 or 71.43***	75.00
	α-2-Macroglobulin Haptoglobin Ceruloplasmin	0.069	>=0.390	85.71	71.43	>=0.503	71.43	78.57	78.57
	α-2-Macroglobulin Hemopexin Ceruloplasmin	0.089	>=0.428	71.43	61.54	>=0.517	64.29	76.92	73.63

COPD = chronic obstructive pulmonary disease; AUC = area under the curve.

Figure Legends

Figure 1. Representative Cy3-labelled 2D-gel of immunodepleted human plasma proteins from a healthy control. Proteins were separated on pH 4-7 IPG strips in the first dimension, and by 4-18%T SDS-PAGE in the second dimension. Numbered spots are differentially expressed ($p<0.05$; 1-ANOVA) across the 3 disease groups (healthy controls, asthmatics and COPD) and were positively identified by mass spectrometry (Table 3).

Figure 2. Validation data for **A** prothrombin, **B** ceruloplasmin, **C** haptoglobin **D** antithrombinIII in the diagnostic immunoassay validation group. Statistical significance determined using un-paired, two-tailed students t-test with multiple comparisons correction using the Bonferroni method (significant $p < 0.017$). Triplicate analyses were performed using either two-site ELISA or immunoblotting for each patient, and the mean value used for analysis. Error bars represent ± SEM.

Figure 3. Validation data for **A** ceruloplasmin, **B** haptoglobin and **C** hemopexin in the clinical validation population. Statistical significance determined using un-paired, two-tailed students t test with multiple comparisons correction using the Bonferroni method (significant $p < 0.0083$). A single analysis was performed for this population using either two-site ELISA or immunoblotting. Error bars represent ± SEM.

Figure 4. Selected ROC curves for **A** asthma versus healthy controls, **B**. COPD versus healthy controls and **C** COPD versus asthma in the clinical validation population.

Identification of Novel Diagnostic Biomarkers for Asthma and Chronic Obstructive Pulmonary Disease

Online Supplement

Nicole M. Verrills[1,4,7†], Jennifer A. Irwin[2,4,5,7†], Xiao Yan He[2,4,5,7], Lisa G. Wood[1,4,5], Heather Powell[2,4,5], Jodie L Simpson[2,4,5], Vanessa M. McDonald[2,3,4,5], Alistair Sim[1,4] and Peter G. Gibson[2,4,5,6,7].

ADDITIONAL METHODS

Chemicals & Reagents

Cy2, Cy3 and Cy5 (minimal dye), Immobiline Drystrips (pH 4-7, 24 cm) PlusOne™ drystrip cover fluid, Pharmalyte™ 3-10 for IEF, Bind-silane solution, 2D quant kit, ECL anti-mouse IgG HRP-linked whole antibody (from sheep), ECL anti-rabbit IgG HRP-linked whole antibody (from sheep), Hybond® C-extra nitrocellulose and ECL Plus western blotting detection system were purchased from GE Healthcare Bio-Sciences AB, Uppsala, Sweden. SDS, 40% acrylamide solution, 30% acrylamide/Bis solution (37.5:1, 2.6%C), SyproRuby™ protein stain and SDS PAGE standards (broad range, unstained) were all electrophoresis purity and purchased from BioRad, NSW, Australia. Agarose (Type I-A: Low EEO), tributyl phosphine solution (TBP), N,N,N',N'-tetramethylethylenediamine (TEMED, >99%), CF_3COOH (TFA, Fluka Biochimika, >99.5%) NH_4HCO_3 (Reagent Plus®, ≥ 99%), CCl_3COOH (TCA, 99%), *L*-lysine monochloride (> 98%), thiourea (ACS reagent), CHAPS (> 98%), bovine serum albumin (BSA) Fraction V, protease inhibitor cocktail, and anti-goat/sheep-HRP were purchased from Sigma-Aldrich, NSW, Australia. $(NH_4)_2S_2O_8$ (>98 %), urea (>98 %), glycine (>98.5%) were sourced from Chem. Supply, Gillman, SA, Australia. Tris(hydroxymethyl)-aminomethane (ACS reagent), glycerol (BDH AnalR, > 99.5%), CH_3CN (BDH HiPer Solv for HPLC, >99.9 %), NaOH (BDH AnalR, 99%) and CH_3COOH (BDH, 100%) were purchased from Merck Australia, VIC, Australia. Methanol (CH_3OH, > 99.8%) was purchased from Ajax Finechem., NSW, Australia. Dithiothreitol (DTT, >99.5%) was purchased from Applichem, Darmstadt, Germany. Bromophenol blue (BPB) was purchased as the sodium salt from Research Organics, Cleveland, OH, USA. $(CH_3)_2NCOH$ (DMF, ≥99.5%) was purchased from USB Corporation, Cleveland, OH, USA and was stored in the dark

under a N_2 atmosphere. α-cyano-4-hydoxycinnamic acid (α-CHCA, recrystallised) and peptide calibration mix 1 (1000-2500 Da) were purchased from Laser BioLabs, Sophia-Antipolis Cedex, France. Sequencing-grade modified porcine trypsin was purchased from Promega (Madison, WI, USA). Ziptip® μ-C18 pipette tips and 0.45 μm white nylon filters were purchased from Millipore. MA, USA. Human IgA ELISA quantitation kit and ELISA starter accessory package were purchased from Bethyl Laboratories, Inc., TX, USA. Human haptoglobin, α-2 macroglobulin, and hemopexin ELISA kits were purchased from Genway Biotech San Diego, CA, USA. Human ceruloplasmin and antithrombin III ELISA kits were purchased from Assay Pro, St. Charles, MO, USA. Human complement factor H ELISA kit was purchased from Hycult biotechnology, Uden, The Netherlands. Anti-prothrombin (ab48627) was purchased from Abcam, Cambridge, UK. Anti-fibrinogen γ chain monoclonal antibody (M01), clone 1F2 was purchased from Abnova Corporation, Taipei, Taiwan. Anti-inter-α-trypsin inhibitor heavy chain H4 (45A12) monoclonal antibody was purchased from Ab Frontier, Seoul, Korea. All reagents were used as received without further purification.

Clinical Assessment

The recruitment criteria for the proteomic discovery study were established to clinically characterise the subject groups of interest with the aim of establishing mutually exclusive clinical diagnostic groups of asthma, COPD, and controls without airway disease. None of the subjects were current smokers. The selection criteria ensured there was control for smoking and that there was control for age and sex effects by matching subjects within defined age ranges. Stable airway disease was defined as no increase in bronchodilator use, no use of oral corticosteroids, no limitation in activities, no doctor's visit and no hospitalisation due to asthma in the

past 4 weeks. Subjects with recent respiratory tract infection (past 4 weeks) were excluded.

Asthma was defined as a compatible history of episodic wheeze, cough and dyspnea that were responsive to asthma treatment, together with airway hyperresponsiveness to hypertonic (4.5%) saline. Asthmatics had fully reversible airflow obstruction and no evidence of COPD as FEV_1/FVC ratio after bronchodilator was normal (>70%) and DLCO was >70% predicted.

The COPD group were selected to have COPD (GOLD stage II or greater) with predominant airway disease with minimal emphysema (D_{LCO}>60% predicted) and no asthma. They were ex-smokers with a significant smoking history, compatible symptoms and a doctors diagnosis of COPD, and incompletely reversible airflow obstruction (FEV_1<80% predicted, FEV_1/FVC < 0.7 postbronchodilator). In addition, there was a negative bronchodilator response (BDR).

Healthy controls had no respiratory symptoms, nor a diagnosis of respiratory disease, together with normal measures of airway function.

Sample Collection and Clinical Measurements

Subjects were assessed following an overnight fast (minimum 12 hours). Peripheral blood was collected from a vein in the forearm. Sputum was induced during hypertonic saline challenge as described (E1, E2). Atopy was assessed using skin prick testing to common allergen extracts (Dome/Hollister-Steir; Bayer Pharmaceuticals, Sydney, Australia) for house dust mites (*Dermatophagoides pteronyssinus*), mold mix (*Alternaria, Tenuis, Aspergillus mix*), mixed grasses, and cockroach, together with positive (histamine) and negative (glycerine) controls. Participants were asked to withhold antihistamine 5 days prior to testing. A skin-prick test was defined as positive if the wheal diameter was 3 mm or greater at 15

min. Participants were considered atopic if a positive skin-prick test was recorded for any allergen, with no reaction to the negative control.

Airway responsiveness was assessed from spirometry (KoKo spirometer, PDS instrumentation, Louisville, Co, 80027, USA) with hypertonic saline (4.5%) provocation challenge as described (1). Airflow obstruction was assessed in each participant using spirometry (KoKo K313100 PDS Instrumentation, Louisville, CO, USA) to measure pre and post bronchodilator FEV_1, FVC and FEV1/FVC%, Participants withheld bronchodilators for their duration of action before testing. Three reproducible measurements of FEV1 and FVC were obtained before and after inhalation of 200 mg albuterol via a metered dose inhaler with valved holding chamber (Volumatic, Allen and Hanbury's, Melbourne Victoria, Australia) using predicted values according to Knudson *et al*. Airflow obstruction was defined as an FEV_1 < 80% predicted and an FEV_1/FVC% as <70% and performed using the American Thoracic Society/European Respiratory Society standards of lung function testing.

The carbon monoxide transfer co-efficient (KCO) was determined according to ATS guidelines (Med- Graphics Elite DX Pulmonary function testing system Medical Graphics Corporation, Minnesota, MN, USA) (4).

Induced Sputum Processing

All sputum samples were processed within 2 hours of collection. Mucus portions were selected from saliva and divided into two portions; the first portion was processed for RNA extraction and the second portion was mixed with a dispersing agent (dithiothreitol, DTT, 0.1%, Calbiochem, La Jolla Ca USA). The tube was capped and rocked for 30 minutes at room temperature, after which PBS was added and the dispersed suspension was filtered (60µm, Millipore, Australia). A leukocyte total cell

count (TCC) and cell viability (trypan blue exclusion) was performed using a haemocytometer. After centrifugation at 400 × *g* for 10 minutes, the cell pellet was resuspended to 1 × 10^6 cells/mL using phosphate buffered saline (PBS). The supernatant was aspirated and stored at -80°C. Cytospins were prepared from the resuspended cell pellet.

Sputum Differential

Cytospins were fixed in methanol and stained with May and Grunwald stain and subsequently with Giemsa stain. A total of 400 non-squamous cells were counted, with the squamous cells proportion recorded separately. Cells were identified by their morphology and the differential cell count was expressed as a percentage of non-squamous cells. Cytospin quality was determined using a scale to evaluate squamous contamination, cell and nuclear integrity, presence of airway macrophages and number of cells present on the slide.

Sample collection and immunodepletion

Blood samples were collected from a peripheral vein in the forearm. Serum samples were collected in tubes in the absence of anti-coagulant and plasma with anti-coagulant. For serum isolation, blood was allowed to clot and serum collected via ultrafiltration at 1400g. For plasma collection, samples were centrifuged at 2000g for 10 minutes at room temperature. Plasma was removed and centrifuged at 2500g for 15 minutes at room temperature to deplete platelets. Protease inhibitor cocktail (Sigma-Aldrich, Missouri, USA) (1% v/v) was added to the platelet depleted plasma prior to storage at -80^{o}C. Each plasma sample was subjected to immuno-depletion using an IgY-12 LC-2 Proteome Partitioning Kit (Beckman Coulter™, CA, USA) and an AKTA P920 FPLC system (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Using this methodology, the twelve most abundant plasma proteins (albumin, IgG,

α1-antitrypsin, IgA, IgM, transferrin, haptoglobin, α1-acid glycoprotein, α2-macroglobulin, HDL (apolipoproteins A-I and A-II) and fibrinogen) were depleted from samples to yield a sub-sample for analysis that was enriched with respect to the other proteins present (E3). Briefly, human plasma (50 µL) was diluted in dilution buffer (200 µL, 10 mM Tris-HCl pH 7.4, 150 mM NaCl), centrifuged at 8300 rpm for 1 min and loaded into an injection loop. The IgY column was pre-equilibrated (dilution buffer, 1 column volume (CV), 0.35 mL min^{-1}) prior to automated sample injection. The unbound protein fraction was collected (fraction 1, 5 mL) using a Frac 900 fraction collector. Dilution buffer was passed through the column (3.5 CV, 0.35 mL min^{-1}) prior to elution of the bound (high abundance) proteins with stripping buffer (50 mM glycine-HCl, pH 2.5, 8 CV, 0.35 mL min^{-1}). The eluate was subsequently neutralised with 10 × neutralisation buffer (1M Tris-HCl, pH 8.0) and stored at -80 °C until required. The column was then immediately re-equilibrated with dilution buffer (32 CV, 0.35 mL min^{-1}) to neutralise the column pH. As 100µg of protein was required per sample hence multiple depletion runs (three to four) were performed and pooled in order to obtain the required quantity of protein. Eluted proteins were stored immediately at -80 ºC until required. The unbound column fractions were thawed, replicates pooled and subsequently concentrated by TCA precipitation. Briefly, an aqueous stock solution of TCA (90% $^{w}/v$) was added to each of the pooled replicates to yield a final concentration of 9% and the samples were left overnight at 4 °C to facilitate protein precipitation. Each tube was then centrifuged at 15,000 × g (15 min, 4 °C) and the supernatant discarded. The protein pellet was washed sequentially in acidified ethanol (40 mM CH_3COOH in CH_3OH) and ice-cold acetone (precipitate pelleted via centrifugation at 15,000 ×g (15 min, 4 °C)), dried in a rotary vacuum concentrator. Proteins were allowed to precipitate overnight at 4 ºC,

centrifuged for 15 min at 15000 g (4 ºC) and the resultant pellet resuspended in DIGE lysis buffer (30 mM Tris pH 8.5, 2M thiourea, 7M urea, 4%(w/v) CHAPS, pH 8.5). Samples were vortexed briefly and pulse sonicated (2 × 2 second pulse at an amplitude of ≤ 2 microns whilst on ice) prior to analysis. Sample pH was adjusted to 8.5 with NaOH solution (50 mM) and the quantity of protein recovered estimated using a 2-D Quant Kit (GE Biosciences) according to the manufacturer's recommendations.

2D-DIGE & Image analysis

Reproducibility study: An experimental design was formulated in order to test the reproducibility of the sample preparation and 2D-DIGE technique. Plasma was collected from one healthy control (#9173) and divided into 3 separate aliquots (labelled A, B, C). Each sample was immunodepleted, TCA precipitated, and quantitated as described previously. The reproducibility of the FPLC separation is shown in Figure E1. Separate aliquots were then labelled with Cy2, Cy3 or Cy5 dyes. Equal protein from the Cy2 samples were pooled for the internal control to be run on every gel. The remaining samples were separated by 2D-DIGE in two independent gel runs, as shown in Table E1. The variability between dye labelling, sample preparation, and gel runs was determined across all spots using the DeCyder generated volume ratios (Table E2).

Discovery study: An experimental design (Table E3) was formulated with the 43 samples assigned to a particular gel and CyDye™ label. Within each clinical group, some samples were labelled with Cy3 and others the Cy5. An internal pooled control was produced from equal quantities of ten of the samples (two per clinical group). Proteins were CyDye labelled and separated on 24cm pH 4-7 IPG strips in the first dimension and 4-18% gradient acrylamide gels in the second dimension as previously

described. Protein separations were visualised on a Typhoon 9410 Variable Mode Imager with Ettan DALT alignment guides using excitation/emission wavelengths specific for Cy2 (488/520 nm), Cy3 (532/580 nm) and Cy5 (633/670 nm) The photomultiplier tube voltage (PMT) was adjusted to preclude spot intensity saturation. The resultant image files were cropped using the program ImageQuant™ Tools 2.1 and saved using the DIGE file naming format.

Image analysis was performed using the Batch Processor and Biological Variation Analysis (BVA) modules of the DeCyder 2D software version 6.5 (GE Healthcare, Australia). This analysis initially normalizes each sample to its respective in-gel Cy2 internal standard, and then matches all controls and samples between different gels. A total of 1918 spots were detected in the master gel (automatically assigned by DeCyder) and matched across the gel images. Comparing each disease group in the BVA module generated average expression ratios and Student's *t* tests of individual protein spots. One way analysis of variance (1-ANOVA) was also used to identify protein spots that showed a significant change across the three groups (healthy control, asthmatic and COPD). Proteins with an expression ratio of +/-1.2 (assigned following reproducibility analysis – see results) and $p<0.05$ by 1-ANOVA or individual group comparison Student's *t*-test, were assigned as a protein of interest (POI). Each of these proteins spots were manually inspected for densitometric Gaussian distribution and match quality.

Protein Identification

For protein identification, five preparative 2D-gels were prepared, one from each disease group. Immuno-depleted plasma (600 µg) was separated by 2D-PAGE as described above. Gels were stained with SyproRuby™ protein stain or Colloidal Coomassie G250 and protein spots of interest excised and subjected to tryptic

digestion. Tryptic digests were analysed by matrix assisted laser desorption ionisation time of flight (MALDI-TOF) mass spectrometry. Peptide mass fingerprint data was acquired using an Ettan MALDI-ToF/Pro mass spectrometer and analysed using the NCBI FASTA human database provided with the system. The instrument was operated in positive reflectron mode acquiring duplicate spectra (700 – 4000 *m/z*), each comprising signal from 400 laser shots (fixed laser power, 337 nm nitrogen laser). Positive ions were extracted into the mass analyser at 20 kV using pulsed extraction. The instrument was calibrated using Pepmix 1 (Laser Biolabs, Ang II (1046.542 Da) and hACTH 18-39 (2465.199 Da)) according to the manufacturers directions. Each sample spectrum was further internally calibrated using Trypsin I (842.508 Da) and trypsin III (2211.108 Da) autolytic peaks prior to database searching. Tryptic digestion was specified with only 1 missed cleavage allowed. The partial amino acid modifications oxidation (M) and propionamide (C) were considered and a fragment mass tolerance of 0.2 *m/z* allowed. PMF data was searched against Homo Sapien entries in the NCBI FASTA database (20060322) with the protein database update tool v1.30 using the PROFOUND peptide mass fingerprinting search engine licensed from ProteoMetrics (NY, USA). Bayesian statistics were used to rank the protein sequences in the database according to their probability of generating the experimental data (expt < 0.010 significant with $p<0.05$). Proteins were assigned as positive identifications if they showed an expectation value <0.05. Samples that were unable to be identified by MALDI-TOF were analysed by MALDI-TOF/TOF at the Australian Proteome Analysis Facility on an Applied Biosystems 4700 Proteomics Analyser with TOF/TOF optics in MS mode. Data was submitted to the database search program Mascot (Matrix Science Ltd, London, UK) and searched against *Homo Sapien* entries in the Swisprot database. Significant Mascot scores in

the database search indicated a likely match, and were confirmed or qualified by operator inspection.

ELISA analysis

Non-depleted serum from the primary validation population was tested for IgA, complement factor H, haptoglobin, α-2 macroglobulin, ceruloplasmin, hemopexin and antithrombin III concentrations using commercially available two-site ELISA kits following the manufacturer's instructions. Both samples and standards were analysed in duplicate wells for each determination and the quantification was performed in triplicate (three separate runs) for the primary validation population and duplicate wells singly for the secondary validation population. Standard curves were fit using a four parameter logistic regression and individual sample concentrations calculated via interpolation. Triplicate determinations (primary validation population) for each sample were averaged and the mean data analysed for differentiation between known subject groupings using GraphPad Prism® 4.02 for Windows. Unpaired t-test analysis was used to test for statistical significance of the observed experimental trends.

Western Blotting

Blood serum was analysed via Western blotting for the candidate markers prothrombin, inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4) and fibrinogen γ chain . Proteins were separated on 10 % or 4-18% acrylamide gels and transferred to nitrocellulose. Membranes were blocked with 2 % BSA in Tris-buffered-saline with Tween (TBST, 20 mM Tris-Cl, 140 mM NaCl, pH 7.5, 0.05 % Tween-20) for one hour at room temperature or overnight at 4 °C. Primary antibodies anti-prothrombin (1:200), anti-ITIH4 (1/1000) and anti-fibrinogen gamma chain (1:1000) in TBST were incubated with the appropriate blocked membranes with agitation for 60 min at

room temperature, the membranes washed in TBST (triplicate, 20 min) and then incubated with secondary antibody (rabbit anti-mouse IgG, HRP-linked whole antibody, 1:10000 in TBST, 60 min) with agitation. Reactive bands were visualised using ECL chemiluminescent reagents and a Fujifilm Luminescent Image Analyser LAS-300 and quantified using Multi Gauge V3.0 software. The chemiluminescent intensity was normalised to a healthy control serum sample run on every gel. Blots were run in triplicate (primary validation group) or singly (secondary validation group). Unpaired Students *t*-test was used to test for statistical significance between the clinical groups.

Statistical Analysis for quantitative assessment of multiple biomarkers

In this programme logistic regression was used to calculate the predicted value of an individual having the disease based on their level of a particular marker (simple logistic regression) or combination of markers (multiple logistic regression). The regression equation utilised was:

$$\ln(p/(1-p)) = \beta_0 + \beta_i x_i$$

$$p = \exp(\beta_0 + \beta_i x_i)/(1 + \exp(\beta_0 + \beta_i x_i)$$

The predicted value is the probability that the person has the disease given the level of the markers. Receiver-Operating Characteristic (ROC) Curves were generated from the predicted values from the regression analyses. The ROC curves were then used to determine the sensitivity and specificity of both individual or combinations of markers at specified cut points of the predicted values. An individuals' protein measurements can be entered into the logistic regression equation to determine their predicted value, a value above or below the cut point would determine their likelihood of disease with the defined sensitivity and specificity calculated from the

ROC curve. The area under the curve (AUC) was also calculated. The AUC value gives a measure of the ability of the test to correctly classify people.

The significance of the regression models were used to assist in determining the best combination of markers. The combinations with the greatest significance and greatest AUC are highlighted in red in Table E4 (primary validation population) and Table E5 (secondary validation population). Sensitivity and specificity were used to describe the discriminatory power of the protein combinations. We selected cut-points in order to maximise the discriminatory power in any given diagnostic situation. Different batches of ELISA kits were used in the in the two validations and this is why there were some differences in the cut offs for the two datasets. We are now able to source kits from the same company on a consistent basis.

Statistically significant combinations are presented in the two tables in red. The combinations highlighted (bold + italics) display highest sensitivity (rule out) and specificity (rule in). High AUC was also considered.

SUPPLEMENTARY RESULTS

Reproducibility analysis of immuno-depletion and 2D-DIGE

The reproducibility of the sample preparation procedure and 2D-DIGE separation was examined. This involved evaluation of the reproducibility of 2D DIGE separations from one healthy control sample in order to test the variability of three parameters: (i) dye labelling; (ii) sample preparation; (iii) independent gel runs (temporal changes). The densitometric volume relative to the internal control (volume ratio) was determined for each protein using the program DeCyder, and the variation between parameters determined by linear regression analysis (Figure. E1). There is a strong correlation between the spot volume ratio of samples labelled with Cy3 and Cy5, slope = 1.07 ± 0.01, r^2 = 0.9744, $p < 0.0001$. (Figure E1A). Similarly, a good

correlation is observed between two samples that have been independently prepared (IgY column, TCA precipitation, protein quantitation, Cy dye labelling), slope = 0.786 ± 0.02, $r^2 = 0.8155$, $p < 0.0001$ (Figure E1B) and between the same sample Cy-labelled and separated by 2D-DIGE in independent experimental gel runs one week apart, slope = 091 ± 0.02, $r^2 = 0.8495$, $p < 0.0001$ (Figure E1C). Additionally, DeCyder was used to analyse the differences between all possible comparator combinations across the 12 gel images, and the analysis of variance determined (Table E2). For downstream analyses, we arbitrarily set proteins of interest as statistically significant changes that show a fold change of at least the mean ± 1 standard deviation. From the reproducibility analysis (Table E2) the mean ± 1SD correlates to a 1.2 fold up- or down-regulation. Thus only protein changes greater than or equal to ±1.2 fold were considered in the subsequent analyses.

The clinical details of the diagnostic immunoassay validation group are shown in Table E6. These results are similar to the discovery population.

Effect of Co-morbid Conditions in the Clinical Validation Population

A literature review identified 5 medical conditions as potential confounders of the blood based marker diagnosis panel: hepatobiliary disease, ischemic coronary disease, obesity , endocrine and metabolic disorders as well as psychiatric conditions. Table E7 summarises a co-morbidity analysis of the clinical validation population and indicates no significant effect of these co-morbid conditions on the marker levels, apart from an effect of psychiatric disease on levels of α-2 macroglobulin.

REFERENCES

E1. Gibson PG, Wlodarczyk JK, Hensley MJ, Gleeson M, Henry RL, Cripps AW, Clancy RL. Epidemiological association of airway inflammation with asthma

symptoms and airway hyperresponsiveness in childhood. *Am J Respir Crit Care Med* 1998; 158:36-41.

E2. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology* 2006; 11:54-61.

E3. Liu T, Qian W-J, Mottaz HM, Gritsenko MA, Norbeck AD, Moore RJ, Purvine SO, Camp II DG, Smith RD. Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol Cell Proteomics* 2006; 5(11):2167-2174.

E4. American Thoracic Society. Single-breath carbon monoxide diffusing capacity (transfer factor). *Am J Respir Crit Care Med* 1995;152:2185-98.

TABLE E1. 2D-DIGE EXPERIMENTAL DESIGN FOR REPRODUCIBILITY STUDY

Gel #	Cy3	Cy5	Run #
1	Sample A	Sample A	1
2	Sample B	Sample B	1
3	Sample C	Sample C	1
4	Sample A	Sample A	2
5	Sample B	Sample B	2
6	Sample C	Sample C	2

TABLE E2. STATISTICS OF SPOT VOLUME RATIOS ACROSS ALL REPRODUCIBILITY ANALYSES

Number of values	3126
Minimum	0.2128
25% Percentile	0.9259
Median	1
75% Percentile	1.08
Maximum	3.68
Mean	1.008
Std. Deviation	0.1745
Std. Error	0.003121
Lower 95% CI of mean	1.002
Upper 95% CI of mean	1.014

TABLE E3. EXPERIMENTAL DESIGN FOR BIOMARKER DISCOVERY STUDY

Gel #	Cy3 (group #)	Cy5 (group #)	Cy2
1	9028 (1)	9787 (4)	Pooled Control
2	9295 (3)	10336 (4)	Pooled Control
3	9404 (5)	9428 (3)	Pooled Control
4	9629 (4)	9620 (3)	Pooled Control
5	9197 (2)	9750 (5)	Pooled Control
6	9210 (2)	9738 (5)	Pooled Control
7	9559 (4)	9311 (2)	Pooled Control
8	9239(1)	9651 (2)	Pooled Control
9	9694 (5)	9249 (1)	Pooled Control
10	9371 (3)	9305 (1)	Pooled Control
11	9444 (1)	10359 (2)	Pooled Control
12	9476 (3)	10404 (4)	Pooled Control
13	10928 (4)	9506 (3)	Pooled Control
14	9612 (3)	10662 (2)	Pooled Control
15	10911 (2)	9666 (3)	Pooled Control
16	9770 (3)	9450 (1)	Pooled Control
17	11173 (4)	9903 (3)	Pooled Control
18	10855 (2)	10540 (4)	Pooled Control
19	9525 (4)	9494 (1)	Pooled Control
20	11165 (2)	10570 (4)	Pooled Control
21*	9197 (2)	9750 (5)	Pooled Control
22*	11108 (3)	10336 (4)	Pooled Control

* represents repeat gels

TABLE E4. ANALYSIS OF DIAGNOSTIC VALUES OF MARKER AND MARKER IN COMBINATIONS IN PRIMARY VALIDATION POPULATION.

Comparison	Marker combination	Logistic regression: Constant	Logistic regression: Coefficient	Logistic regression: p value	Logistic regression: Model p value	Minimal false negatives/rule disease out: Cut off measurement	Minimal false negatives/rule disease out: Cut-off, Predicted value	Minimal false negatives/rule disease out: Sensitivity	Minimal false negatives/rule disease out: Specificity	Minimal false positives/ rule disease in: Cut off measurement	Minimal false positives/ rule disease in: Cut-off, Predicted value	Minimal false positives/ rule disease in: Sensitivity	Minimal false positives/ rule disease in: Specificity	AUC (%)
Asthma v Healthy														
	Individual													
n = 44	Ceruloplasmin	-4.018952	0.005213	0.002	0.0002	>=761.22mcg/ml	>=0.487	85.71	62.5	>=1017.89mcg/ml	>=0.784	64.29	93.75	81.47
	Haptoglobin	-3.276754	1.109178	0.003	0.0001	>=3.37mg/ml	>=0.613	85.71	75.0	>=3.84mg/ml	>=0.727	67.86	87.50	82.37
	Hemopexin	0.307188	0.0002305	0.904	0.904									50.89
	α-2-Macroglobulin	1.891332	0.0005327	0.190	0.181									65.63
	Combination													
	Ceruloplasmin	***-6.880227***	***0.0045298***	***0.015***	***<0.00001***		***>=0.448***	***92.86***	***75.0***		***>=0.785***	***67.86***	***93.75***	***89.29***
	Haptoglobin		***1.036785***	***0.014***										
	Ceruloplasmin	-2.27678	0.005567	0.002	0.0006		>=0.442	89.29	62.5		>=0.753	67.86	93.75	83.04
	Hemopexin		-0.0018857	0.436										
		-2.215251	1.144461	0.003	0.0006		>=0.590	85.71	68.75		>=0.676	78.57	81.25	82.59
	Haptoglobin		-0.0010982	0.626										
	Hemopexin													
COPD v Healthy														
	Individual													
n = 30	Ceruloplasmin	-7.004728	0.0079928	0.005	0.0003	>=794.79mcg/ml	>=0.343	85.71	75.0	>=972.71mcg/ml	>=0.684	64.29	93.75	86.61
	Haptoglobin	-3.684739	1.104089	0.034	0.013	>=2.96mg/ml	>=0.398	76.92	56.25	>=3.15mg/ml	>=0.448	61.54	68.75	73.56
	Hemopexin	4.024576	-0.004047	0.093	0.073									
	α-2-Macroglobulin	-1.084965	0.0003458	0.535	0.531									
	Combination													
	Ceruloplasmin	-9.850005	0.0073615	0.013	0.0004		>=0.342	84.62	68.75		>=0.590	69.23	87.50	88.94
	Haptoglobin		1.063352	0.072										
	Ceruloplasmin	***-0.8243742***	***0.0118777***	***0.007***	***<0.00001***		***>=0.426***	***92.86***	***75.0***		***>=0.480***	***85.71***	***87.50***	***91.07***
	Hemopexin		***-0.0092461***	***0.024***										
		0.9626632	1.420984	0.035	0.009		>=0.444	76.92	62.50		>=0.547	61.54	87.50	81.25
	Haptoglobin		-0.0055647	0.104										
	Hemopexin													
	α-2-Macroglobulin	3.912215	0.0000273	0.964	0.199									
	Hemopexin		-0.0040107	0.114										
COPD v Asthma														
	Individual													

n = 42	Ceruloplasmin	0.0630528	-0.0007368	0.576	0.575									
	Haptoglobin	1.20993	-0.5029125	0.102	0.078									
	Hemopexin	4.72965	-0.0052496	0.032	0.02	<=1090.89mg/ml	>=0.270	85.71	57.14	<=1009.75mg/ml	>=0.361	64.29	75.0	71.68
	α-2-Macroglobulin	-2.927178	0.0008705	0.066	0.050	>=2518.45mg/ml	>=0.324	85.71	75.0	>=2689.38mg/ml	>=0.358	64.29	82.14	76.28
	Combination													
	α-2-Macroglobulin	-0.9782056	0.000794	0.092	0.045		>=0.289	84.62	64.29		>=0.382	61.54	78.57	74.73
	Haptoglobin		-0.4664416	0.137										
	α-2-Macroglobulin	2.392418	0.0007116	0.133	0.019		>=0.314	78.57	67.86		>=0.435	57.14	78.57	75.51
	Hemopexin		-0.0047623	0.058										
	α-2-Macroglobulin	***2.530823***	***0.0006869***	***0.147***	***0.046***		***>=0.285***	***92.31***	***64.29***		>=0.355	76.92	75.0	***76.10***
	Haptoglobin		***-0.3750563***	***0.251***										
	Hemopexin		***-0.0034578***	***0.193***										
	α-2-Macroglobulin	***2.668874***	***0.0007084***	***0.134***	***0.048***		>=0.338	78.57	71.43		***>=0.493***	***50.0***	***89.29***	***75.77***
	Hemopexin		***-0.0047018***	***0.063***										
	Ceruloplasmin		***-0.0003201***	***0.829***										

COPD = chronic obstructive pulmonary disease; AUC = area under the curve.

TABLE E5. ANALYSIS OF DIAGNOSTIC VALUES OF MARKER AND MARKER IN COMBINATIONS IN SECONDARY VALIDATION POPULATION.

Comparison	Marker combination	Logistic regression				*Minimal false negatives/rule disease out*				*Minimal false positives/ rule disease in*				AUC (%)
		Constant	*Coefficient*	*p value*	*Model p value*	*Cut off measurement*	*Cut-off, Predicted value*	*Sensitivity*	*Specificity*	*Cut off measurement*	*Cut-off, Predicted value*	*Sensitivity*	*Specificity*	
Asthma v Healthy														
	Individual													
n = 36	Ceruloplasmin	-7.266969	0.0085751	0.008	0.0007	>=744.93 mcg/ml	>=0.293	85.71	68.18	>=898.31 mcg/ml	>=0.607	50.0	95.45	82.47
	Haptoglobin	-3.976472	0.5013821	0.014	0.002	>=6.36 mg/ml	>=0.313	85.71	54.55	>=8.70 mg/ml	>=0.596	50.0	86.36	79.87
	Hemopexin	-9.184255	0.0187517	0.004	0.0002	>=402.35 mg/ml	>=0.162	100.00	50.0	>=479.65 mg/ml	>=0.453	69.23	86.36	82.17
	α-2-Macroglobulin	0.1041666	-0.1956642	0.677	0.672	<=2.89 mg/ml	>=0.386	64.29	65.22	<=2.85 mg/ml	>=0.388	57.14	69.57	53.73
	Combination													
	Ceruloplasmin	-10.4399	0.0087462	0.022	0.0002		>=0.342	85.71	72.73		>=0.561	64.29	95.45	87.01
	Haptoglobin		0.434046	0.046										
	Ceruloplasmin	-11.97446	0.0052252	0.074	0.0001		>=0.226	92.31	68.18		>=0.421	69.23	81.82	89.16
	Hemopexin		0.015727	0.024										
		-22.76312	***1.248297***	***0.014***	***<0.00001***		***>=0.289***	***100.0***	***90.91***		***>=0.457***	***92.31***	***95.45***	96.85
	Haptoglobin		***0.0276208***	***0.015***										
	Hemopexin													
COPD v Healthy														
	Individual													
n = 44	Ceruloplasmin	-3.5269	0.0045915	0.032	0.014	>=692.70 mcg/ml	>=0.414	81.82	50.0	>=800.27 mcg/ml	>=0.537	54.55	72.73	71.07
	Haptoglobin	-1.58826	0.245888	0.047	0.034	>=5.17 mg/ml	>=0.422	77.27	40.91	>=7.81 mg/ml	>=0.582	50.0	81.82	65.91
	Hemopexin	-4.232852	0.0093937	0.019	0.009	>=462.65 mg/ml	>=0.523	72.73	86.36	>=517.64 mg/ml	>=0.652	50.0	95.45	73.76
	α-2-Macroglobulin	-2.46742	0.6973598	0.038	0.016	>=3.29 mg/ml	>=0.456	72.73	69.57	>=3.79 mg/ml	>=0.544	50.0	91.3	71.15
	Combination													
	Ceruloplasmin	-4.246169	0.0039499	0.079	0.016		>=0.460	72.73	59.09		>=0.573	63.64	81.82	75.41
	Haptoglobin		0.1850704	0.150										
	Ceruloplasmin	-6.018214	0.0034743	0.110	0.008		>=0.521	72.73	77.27		>=0.617	63.64	90.91	76.65
	Hemopexin		0.0074654	0.071										
		-4.623185	***0.1674288***	***0.200***	***0.014***		>=0.439	72.73	63.64		***>=0.567***	***59.09***	***95.45***	73.97
	Haptoglobin		***0.007912***	***0.061***										
	Hemopexin													
	α-2-Macroglobulin	***-6.03456***	***0.6182299***	***0.090***	***0.006***		***>=0.412***	***86.36***	***68.18***		***>=0.631***	***54.55***	***95.45***	80.17
	Hemopexin		***0.0086636***	***0.039***										

Mixed v Healthy	**Individual**													
n = 36	Ceruloplasmin	-3.238282	0.0037112	0.119	0.092	>=694.91 mcg/ml	>=0.341	71.43	50.0	>=785.22 mcg/ml	>=0.420	50.0	72.73	64.61
	Haptoglobin	-5.139459	0.6587639	0.009	0.0005	>=6.64 mg/ml	>=0.318	85.71	68.18	>=7.74 mg/ml	>=0.490	71.43	81.82	81.49
	Hemopexin	-8.831491	0.0187215	0.010	0.002	>=435.08 mg/ml	>=0.335	85.71	63.64	>=454.53 mg/ml	>=0.420	64.29	81.82	79.87
	α-2-Macroglobulin	-2.941042	0.7262588	0.089	0.071	>=3.07 mg/ml	>=0.330	71.43	56.52	>=3.50 mg/ml	>=0.401	50.0	82.61	68.63
	Combination													
	Ceruloplasmin	-6.705737	0.0024152	0.468	0.002		>=0.321	92.86	59.09		>=0.488	64.29	81.82	80.84
	Haptoglobin		0.6268258	0.016										
	Ceruloplasmin	-9.419774	0.0017056	0.552	0.007		>=0.321	92.86	68.18		>=0.464	64.29	86.36	81.49
	Hemopexin		0.017201	0.022										
		-12.00865	***0.0006532***	***0.866***	***0.0007***		>=0.277	92.86	59.09		***>=0.530***	***71.43***	***95.45***	86.04
	Ceruloplasmin		***0.0156691***	***0.069***										
			0.5881633	***0.030***										
	Hemopexin													
	Haptoglobin													
	α-2-macroglobulin	***-15.55849***	***1.120276***	***0.106***	***0.0002***		***>=0.295***	***100.0***	***77.27***		***>=0.435***	***92.86***	***95.45***	94.16
	Hemopexin		***0.0153898***	***0.078***										
	Haptoglobin		***0.63758***	***0.029***										
COPD v Asthma	**Individual**													
n = 36	Ceruloplasmin	1.67445	-0.0014039	0.414	0.408									
	Haptoglobin	1.521738	-0.1376352	0.323	0.310									
	Hemopexin	2.222555	-0.0033549	0.346	0.334									
	α-2-Macroglobulin	-2.734693	0.9372688	0.038	0.013	>=3.29 mg/ml	>=0.585	72.73	71.43	>=3.68 mg/ml	>=0.672	54.55	92.86	71.43
	Combination													
	Ceruloplasmin	2.242203	-0.0010095	0.575	0.510									
	Haptoglobin		-0.1169128	0.422										
	Ceruloplasmin	3.0697	-0.0011649	0.506	0.502									
	Hemopexin		-0.0030206	0.404										
		3.373483	-0.2144594	0.193	0.247									
	Haptoglobin		-0.0022269	0.557										
	Hemopexin													
	α-2-Macroglobulin	-1.485353	1.281106	0.024	0.008		>=0.556	72.73	64.29		>=0.621	68.18	92.86	79.55
	Haptoglobin		-0.304815	0.088										
	α-2-Macroglobulin	-1.175074	0.9805878	0.044	0.028		>=0.523	81.82	61.54		>=0.667	59.09	84.62	74.48
	Hemopexin		-0.0031468	0.417										
	α-2-Macroglobulin	-0.9186874	1.211829	0.026	0.012		>=0.585	72.73	71.43		>=0.665	68.18	92.86	78.25
	Ceruloplasmin		-0.0031181	0.118										
	α-2-Macroglobulin	***-0.1642794***	***1.898634***	***0.019***	***0.003***		>=0.512	81.82	61.54		***>=0.686***	***68.18***	***100.0***	84.27
	Haptoglobin		***-0.5863676***	***0.034***										
	Hemopexin		***-0.0015585***	***0.749***										
	α-2-Macroglobulin	***-0.0511131***	***1.439518***	***0.019***	***0.011***		***>=0.520***	***86.36***	***64.29***		>=0.641	63.64	85.71	79.87
	Haptoglobin		***-0.2743187***	***0.153***										
	Ceruloplasmin		***-0.002518***	***0.234***										
	α-2-Macroglobulin	0.0666635	1.279262	0.030	0.024		>=0.477	81.82	61.54		>=0.684	68.18	84.62	78.67

	Hemopexin		-0.0023254	0.576										
	Ceruloplasmin		-0.0030108	0.141										
Mixed v Asthma	**Individual**													
n = 28	Ceruloplasmin	2.635702	-0.0031127	0.183	0.154									
	Haptoglobin	-0.225885	0.0273526	0.888	0.888									
	Hemopexin	3.613586	-0.0070479	0.193	0.177									
	α-2-Macroglobulin	-4.068517	1.232145	0.047	0.025	>=2.94 mg/ml	>=0.392	78.57	64.29	>=3.45 mg/ml	>=0.544	57.14	71.43	70.92
	Combination													
	Ceruloplasmin	1.894628	-0.0035794	0.144	0.294									
	Haptoglobin		0.1383806	0.528										
	Ceruloplasmin	5.38602	-0.0025878	0.260	0.202									
	Hemopexin		-0.0062152	0.271										
		4.881399	-0.141323	0.567	0.341									
			-0.0071849	0.186										
	Haptoglobin													
	Hemopexin													
	α-2-Macroglobulin	-5.300792	1.308788	0.046	0.069		>=0.340	85.71	64.29		>=0.490	64.29	78.57	77.04
	Haptoglobin		0.1190857	0.586			or	or	or					
							>=0.450	78.57	71.43					
	α-2-Macroglobulin	-3.045245	1.231603	0.083	0.063		>=0.380	78.57	61.54		>=0.602	57.14	84.62	75.82
	Hemopexin		-0.0018068	0.773										
	α-2-Macroglobulin	***-1.67291***	***1.150547***	***0.068***	***0.045***		***>=0.368***	***85.71***	***57.14***		***>=0.583***	***57.14***	***92.86***	75.00
	Ceruloplasmin		***-0.0024947***	***0.302***							***or***	***or***	***or***	
											>=0.509	***64.29***	***71.43***	
	α-2-Macroglobulin	-2.289593	1.202923	0.094	0.133									
	Haptoglobin		-0.067749	0.803										
	Hemopexin		-0.0019742	0.755										
	α-2-Macroglobulin	-3.341737	1.276604	0.063	0.069		>=0.390	85.71	71.43		>=0.503	71.43	78.57	78.57
	Haptoglobin		0.2225828	0.371										
	Ceruloplasmin		-0.0031672	0.206										
	α-2-Macroglobulin	-1.267479	1.18816	0.100	0.089		>=0.428	71.43	61.54		>=0.517	64.29	76.92	73.63
	Hemopexin		-0.0011687	0.858										
	Ceruloplasmin		-0.0022912	0.337										
Mixed v COPD	**Individual**													
n = 36	Ceruloplasmin	0.3701332	-0.0009998	0.567	0.558									
	Haptoglobin	-1.829842	0.1760592	0.243	0.227									
	Hemopexin	-0.1913662	-0.0005365	0.884	0.884									
	α-2-Macroglobulin	0.3487275	-0.2102742	0.474	0.458									
	Combination													
	Ceruloplasmin	-0.6409984	-0.0019024	0.326	0.285									
	Haptoglobin		0.225127	0.157										
	Ceruloplasmin	0.3984816	-0.0009929	0.578	0.843									
	Hemopexin		-0.0000701	0.985										
		-0.5461736	0.2702018	0.162	0.316									
			-0.0041748	0.370										
	Haptoglobin													
	Hemopexin													
	α-2-Macroglobulin	-0.9540765	-0.2721244	0.366	0.308									
	Haptoglobin		0.1971568	0.196										

α-2-Macroglobulin	0.6977482	-0.2143809	0.471	0.747
Hemopexin		-0.0006869	0.853	
α-2-Macroglobulin	0.8839842	-0.1852103	0.538	0.689
Ceruloplasmin		-0.0007668	0.664	
α-2-Macroglobulin	0.794288	-0.3306576	0.309	0.322
Haptoglobin		0.3095113	0.119	
Hemopexin		-0.0049681	0.302	
α-2-Macroglobulin	-0.0225753	-0.2367757	0.441	0.369
Haptoglobin		0.238172	0.136	
Ceruloplasmin		-0.0016801	0.388	
α-2-Macroglobulin	1.022907	-0.1882617	0.535	0.861
Hemopexin		-0.0003211	0.933	
Ceruloplasmin		-0.0007321	0.686	

COPD = chronic obstructive pulmonary disease; AUC = area under the curve.

TABLE E6. CLINICAL DATA FOR THE DIAGNOSTIC IMMUNOASSAY VALIDATION (DIVA) GROUP

Group Description	1 Healthy Controls	2 Stable Asthma	3 COPD	P value
Number of patients	16	28	14	
Age (Years) [a]	45.2 ± 14.3	50.4 ± 13.2	§*65.5 ± 7.3	0.0001
Sex (Male/ Female) [c]	7/9	13/15	8/6	0.737
Smoking, n(%)[c]:			*	0.012
Never	7 (43.8%)	14 (50.0%)	1 (7.1%)	
Ex	9 (56.3%)	14 (50.0%)	11 (78.6%)	
Current	0	0	2 (14.3%)	
Pack years [a]	20.3 ± 17.4	15.9 ± 14.5	§*67.7 ± 31.4	<0.0001
Atopy, n(%)[c]	6 (37.5%)	*24 (85.7%)	§5 (35.7%)	0.001
%predicted FEV_1 [a,e]	97.2 ± 9.0	*76.9 ± 18.3	*71.2 ± 14.7	<0.0001
%predicted FEV_1 [a,f]	100.4 ± 8.6	*85.9 ± 14.4[g]	*72.7 ± 14.4	<0.0001
%predicted FVC [a,e]	101.4 ± 11.6	91.8 ± 15.5	*85.2 ± 14.8	0.011
FEV_1/FVC % [a,e]	79.0± 8.0	*68.2 ± 9.3	*62.5 ± 7.6	<0.0001
PD_{15} (mL) [b, d]	N/A	5.1 (2.7, 8.8)	4.8 (2.2, 8.4)	0.941
D_{LCO} [a]	N/A	85.5 ± 14.6	65.3 ± 13.2	0.01
ICS use, n(%)[c]	0	27 (96.4%)	§0	0.0003
ICS (µg beclomethasone equivalents /day) [b]	N/A	1406 ± 1121	N/A	
ACQ score	N/A	1.13 ± 0.82	N/A	
GOLD Classification, n (%)				
I	N/A	N/A	4 (28%)	
II	N/A	N/A	10 (71,4%)	
III	N/A	N/A	0 (0%)	
IV	N/A	N/A	0 (0%)	
Sputum Cell Counts				
Quality [b]	18.5 (17, 20)	19 (17, 20)	19 (17, 20)	0.926
Total cell count x 10^6/ml [b]	3.5 (2.4, 6.7)	4.1 (2.3, 9.7)	5.7 (2.9, 10.4)	0.728
Neutrophils% [b]	28.7 (17.8, 54.8)	48.9 (35.3, 74.5)	57.0 (53.5, 64.5)	0.065
Eosinophils% [b]	0 (0, 0.25)	*1.5 (0.5, 2.75)	§0.25 (0, 0.75)	0.0001
Macrophages% [b]	62.0 (38.3, 79.87)	42.8 (22.8, 54.8)	40.5 (31.8, 44.5)	0.066
Lymphocytes% [b]	0.75 (0.25, 1.25)	0.5 (0.25, 0.75)	1.0 (0.75, 1.5)	0.083
Columnar epithelial% [b]	2.25 (0.5, 5.5)	2.0 (0.5, 4.0)	1.75 (0.75, 2.5)	0.815
Squamous% [b]	2.3 (1.7, 5.7)	3.6 (0.99, 7.2)	2.2 (0.25, 3.9)	0.190

[a]Values are Mean ± SD, oneway ANOVA; [b]Values are median (interquartile range), Kruskall-Wallis test; [c]Chi Square or Fisher's exact test; [d] PD_{15} is provocation dose resulting in 15% drop in baseline FEV_1 expressed as geometric mean (log SD); FEV_1 is forced expiratory volume in 1 second either pre[e]- or post-[f] bronchodilator ([g] data only available for 20/28 individuals); FVC is forced vital capacity; D_{LCO} is carbon monoxide diffusing capacity; ICS is inhaled corticosteroids. Bonferroni post hoc test significant compared to: *Healthy Controls; [§]Stable Asthma. Asthma: symptoms with fully reversible variable airflow obstruction[airway hyperresponsiveness and/or increased bronchodilator reversibility]. COPD: symptoms, incomplete reversibility of airflow obstruction [postbronchodilator FEV<80%predicted].

TABLE E7 BIOMARKER LEVELS BY COMORBIDITIES

Biomarker	No Diabetes N= 66	Diabetes N=6	P value
Ceruloplasmin, mcg/ml [b]	785(668, 906)	717 (695, 749)	0.328
Haptoglobin, mg/ml [a]	7.0 ± 2.6	8.3 ± 3.0	0.255
Hemopexin, mcg/ml [a]	469 ± 95.5	493 ± 39.6	0.549
α-2-macroglobulin, mg/ml [b]	3.2 (2.8, 3.8)	3.8 (2.4, 4.1)	0.962
	BMI<=30 N=54	BMI>30 N=19	
Ceruloplasmin mcg/ml [b]	800 (668, 898)	732 (680, 875)	0.754
Haptoglobin, mg/ml [a]	6.9 ± 2.8	7.6 ± 2.1	0.354
Hemopexin, mcg/ml [a]	470 ± 94	476 ± 90	0.806
α-2-macroglobulin, mg/ml [b]	3.2 (2.8, 3.7)	3.0 (2.7, 4.1)	0.880
	No Cardiac disease N=60	Cardiac disease n=12	
Ceruloplasmin, mcg/ml [b]	793 (666, 910)	721 (681, 783)	0.257
Haptoglobin, mg/ml [a]	6.9 ± 2.7	8.1 ± 2.3	0.158
Hemopexin, mcg/ml [a]	474 ± 95	460 ± 83	0.632
α-2-macroglobulin, mg/ml [b]	3.1 (2.8, 3.7)	3.7 (3.2, 4.1)	0.181
	No Liver disease N=67	Liver disease N=5	
Ceruloplasmin, mcg/ml [b]	766 (668, 898)	800 (708, 801)	0.938
Haptoglobin, mg/ml [a]	7.2 ± 2.7	7.0 ± 3.0	0.867
Hemopexin, mcg/ml [a]	477 ± 85	400 ± 158	0.072
α-2-macroglobulin, mg/ml [b]	3.2 (2.8, 3.8)	3.7 (2.7, 5.0)	0.600
	No Psychiatric disorder, n=55	Psychiatric disorder, n=17	
Ceruloplasmin, mcg/ml [b]	785 (680, 913)	708 (652, 820)	0.210
Haptoglobin, mg/ml [a]	7.2 ± 2.5	6.9 ± 3.2	0.626
Hemopexin, mcg/ml [a]	464 ± 88	493 ± 103	0.262
α-2-macroglobulin, mg/ml [b]	3.1 (2.7, 3.6)	3.7 (3.5, 4.5)	0.031

[a]Values are Mean ± SD, Student's tests; [b]Values are median (interquartile range), Wicoxon rank sum test;

FIGURE LEGENDS

Figure E1. Reproducibility of plasma sample preparation and 2D-DIGE. Correlation of protein spot ratios between (A) the same sample labelled with Cy3 or Cy5; (B) the same plasma sample independently immunodepleted and TCA precipitated; and (C) the same sample run on a different gel one week apart.