

RESEARCH

Open Access



# Quantitative proteomic analysis of bronchoalveolar lavage fluid in West Highland white terriers with canine idiopathic pulmonary fibrosis

Rosemary E Maher<sup>1</sup>, Merita Määttä<sup>2\*</sup>, Robert J Beynon<sup>1</sup>, Henna P Laurila<sup>2</sup>, Paul S McNamara<sup>3</sup> and Minna M Rajamäki<sup>2</sup>

## Abstract

**Background:** Canine idiopathic pulmonary fibrosis (CIPF) is a chronic, progressive, interstitial fibrosing lung disease, manifesting as cough, exercise intolerance and ultimately, dyspnea and respiratory failure. It mainly affects West Highland white terriers (WHWTs), lacks curable treatment and has a poor prognosis. Aspiration of gastroesophageal refluxate may play a role in the development of CIPF. In the first part of this study, we completed label-free quantitative proteomic analysis of bronchoalveolar lavage fluid (BALF) from CIPF and healthy WHWTs. In the second part, we evaluated potential protein markers of reflux aspiration from canine gastric juice and vomitus and whether these were present in BALF from the two groups.

**Results:** Across all BALF samples, 417 proteins were identified, and of these, 265 proteins were identified by two or more unique tryptic peptides. Using the 265 high confidence assignments, the quantitative proteome profiles were very similar in the two cohorts, but they could be readily resolved by principal component analysis on the basis of differential protein expression. Of the proteins that were differentially abundant in the two groups, several (including inflammatory and fibrotic markers) were elevated in CIPF, and a smaller, more diverse group of proteins were diminished in CIPF. No protein markers indicative of reflux aspiration were identified.

**Conclusions:** Label-free proteomics allowed discrimination between CIPF and healthy WHWTs, consistent with fibrotic process but did not provide clear evidence for gastrointestinal aspiration. The measurement of proteins may provide a proteomics signature of CIPF that could be used to evaluate treatment options.

**Keywords:** Bronchoalveolar lavage fluid, Dog, Lung, Reflux aspiration, Canine, Reflux aspiration, Proteomics: West Highland white terrier, CIPF

## Background

Canine idiopathic pulmonary fibrosis (CIPF) is a chronic, progressive, interstitial fibrosing lung disease affecting mainly West Highland white terriers (WHWTs)

characterized by unclear etiology, lack of curable treatment and poor prognosis [1–3]. Typical clinical signs include cough, exercise intolerance and, in later stages, dyspnea and respiratory failure [1, 2]. Similarities between CIPF and human idiopathic pulmonary fibrosis (IPF) and nonspecific interstitial pneumonia exist both in pathological and clinical findings [2, 4, 5]. Since WHWTs are susceptible to CIPF, it is likely that genetics, together

\*Correspondence: merita.maatta@helsinki.fi

<sup>2</sup> Department of Equine and Small Animal Medicine, Faculty of Veterinary Medicine, University of Helsinki, Helsinki, Finland

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

with other putative risk factors such as gastroesophageal reflux and reflux aspiration, play a role in the development of CIPF [6].

Extra-esophageal reflux is thought to be common in IPF with silent and repetitive microaspiration of refluxate postulated to cause alveolar epithelial injury, dysregulated wound healing and fibrosis [7]. Detection of reflux aspiration is often based on evaluation of gastrointestinal contents in airway samples [8, 9]. In IPF, pepsin is considered an important potential marker of reflux aspiration [10, 11]. Previously, we have shown that bile acids, also considered a marker of reflux aspiration [8], were measurable in bronchoalveolar lavage fluid (BALF) in most CIPF and healthy WHWTs but not in healthy beagles [6]. Supporting a possible role of reflux aspiration in disease pathogenesis, Fastrès et al. [12] found bacteria originating from ingested food or water (*Brochothrix*, *Curvibacter*, *Pseudarcicella* and *Flavobacteriaceae* genus) to be significantly more abundant in lung microbiota of CIPF and healthy WHWTs than in healthy dogs of other breeds.

To gain a better understanding of the etiology and pathogenesis of CIPF in WHWTs, we have completed a proteomics analysis of BALF from healthy and CIPF WHWT subjects. Proteomics permits a broad survey of global protein expression and can provide relative quantitative data that permit a comparison between the two subject groups. In brief, BALF samples were digested to convert proteins into peptides that were then resolved by high resolution reverse-phase chromatography and analysed by tandem mass spectrometry, to gain the identity and a quantification measure of each protein in the mixture. From this we were able to add a functional perspective on proteins enhanced in CIPF BALF. Furthermore, using samples of canine gastric juice and vomit, we explored whether proteins from the upper gastrointestinal tract were present in BALF from CIPF and healthy WHWTs and whether they might have potential as biomarkers of reflux aspiration.

## Results

### Proteomic profiling of BALF in WHWTs

Bronchoalveolar lavage samples from CIPF ( $n=6$ ) and healthy WHWTs ( $n=6$ ) were collected and used for comparative proteomics. The total protein concentration of BALF was increased in WHWTs with CIPF (mean  $\pm$  SD,  $0.75 \pm 0.14$  mg/mL) compared to healthy WHWTs (mean  $\pm$  SD,  $0.38 \pm 0.06$  mg/mL,  $P=0.042$ , Fig. 1a). On SDS-PAGE, the protein profile of BALF was dominated by a high abundance of a protein that migrated at approximately 66 kDa (Fig. 1b), established as canine serum albumin by in gel digestion (results not shown).

Each BALF sample was digested with trypsin and subjected to proteomics analysis. Similar numbers

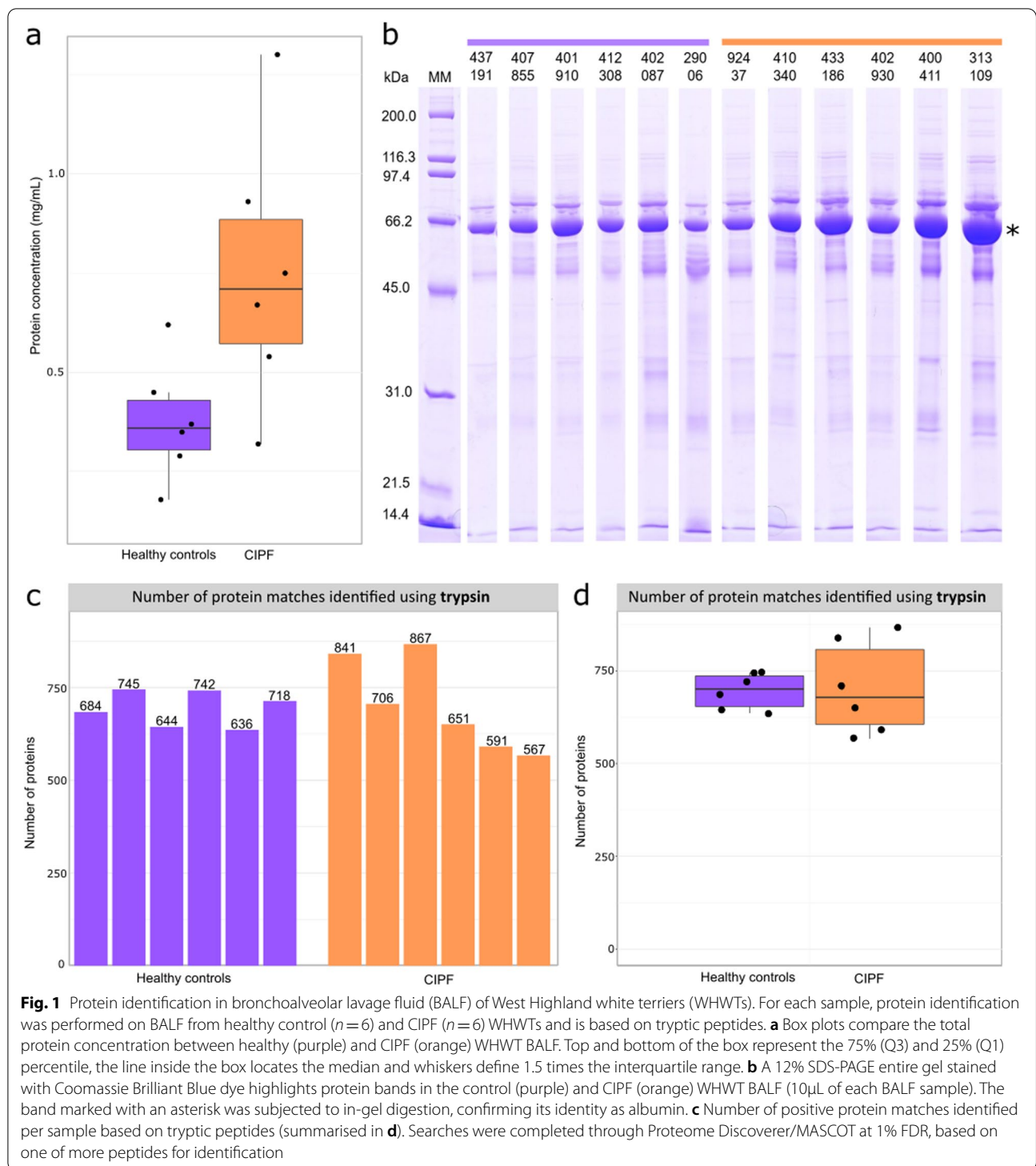
of total proteins (approximately 700, range 567–857) were identified in all samples (Fig. 1c and d). When we applied the criterion that a protein should be identified in every sample, 417 proteins were observed in BALF across all twelve WHWT samples, based on one or more peptides for identification. Overall, there was little difference between the total number of proteins identified. To increase the confidence in protein identification, the 417 proteins were further filtered to require at least two unique peptides. By applying this filter, the number of proteins in the tryptic digests was reduced to 265 high confidence assignments across all samples (we refer to this as the ‘core BALF proteome’). Functional enrichment analysis of this BALF core proteome highlighted the higher representation of secretory or exosomal proteins (Fig. 2), particularly in protease:antiprotease balance, complement and the humoral immune response.

Label-free quantification, based on the summed intensities of unique peptides derived from each protein, allowed the generation of a quantitative profile of the BALF proteins. As anticipated from the SDS-PAGE analysis (Fig. 1b), protein abundances spanned more than six orders of magnitude (Fig. 3a), although the 20 most abundant proteins explained 80% of the total label-free abundance. By far the most abundant protein was serum albumin. The next five most abundant proteins (Supplementary Table 1) were transferrin, uteroglobin, also known as secretoglobulin family 1 A member 1 (SCGB1A1), alpha 2-HS glycoprotein (fetuin A, AHSG), chloride channel accessory protein 1 (CLCA1) and haptoglobin.

The confidence in protein identification varied from protein to protein. Of the 417 proteins common to CIPF and healthy WHWTs BALF, 152 (36%) were based on a single unique peptide for identification (Fig. 3b). In the lower abundance region, protein identifications were achieved with fewer unique peptides, and the scatter in the data was more pronounced. The overall quantitative proteome profiles of the two conditions (control and CIPF) were very similar, as evidenced by the overlap between the log-transformed abundance waterfall plots (Fig. 3a) and the strong correlation between most protein abundances in the two conditions (Fig. 3a and b). This analysis thus satisfied the condition of an overall proteome similarity, creating a reference comparison from which differentially abundant proteins could be extracted.

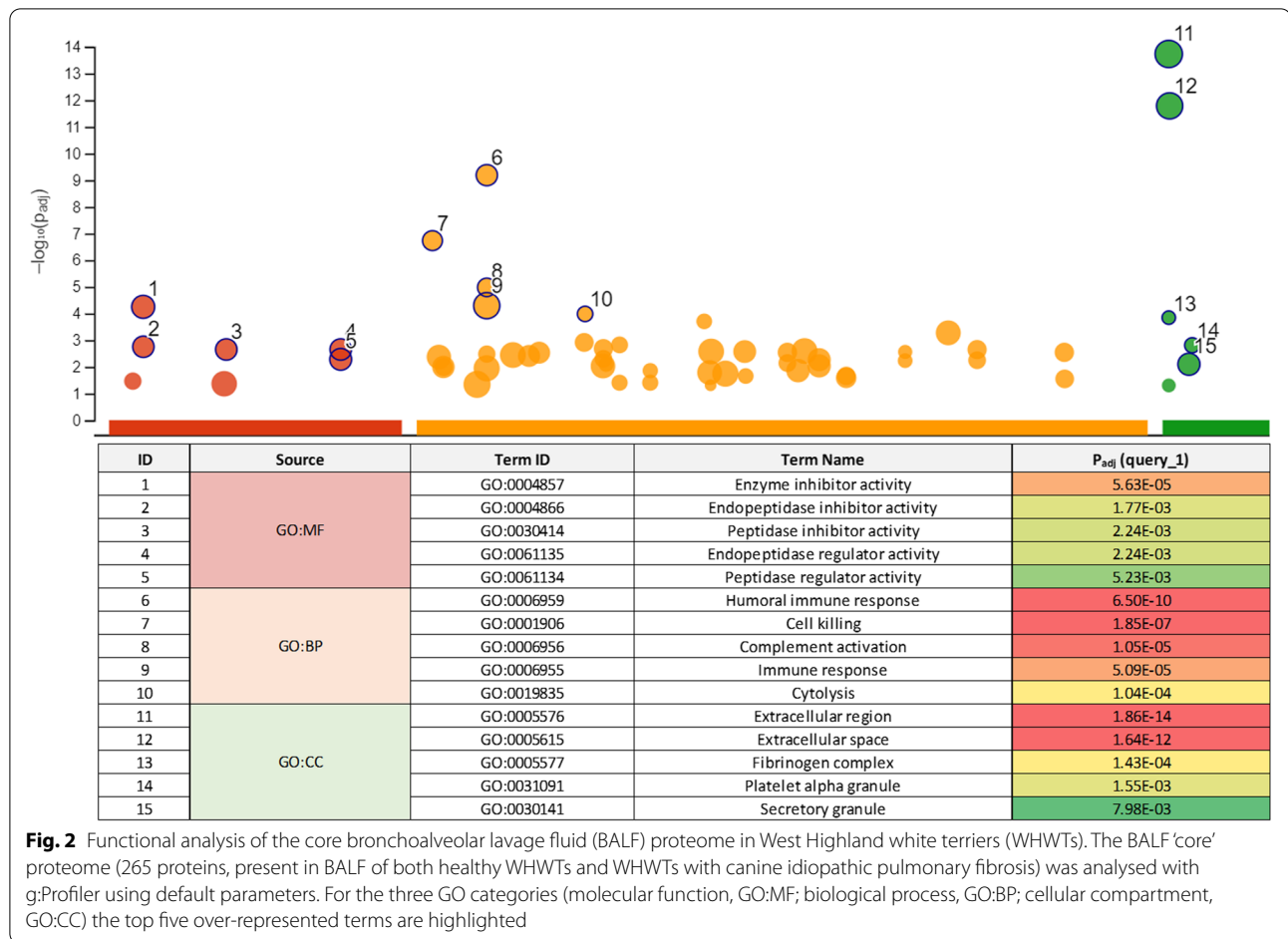
### Comparison of the BALF proteome from control and CIPF samples

The BALF proteome was compared between CIPF and healthy WHWTs. For this analysis, the normalised



abundance data from the core BALF proteome (265 proteins) were used. A volcano plot of the protein-level data revealed about 100 proteins that were significantly changed in CIPF BALF (resampled  $P < 0.05$ ) (either up or down) relative to healthy WHWTs (Fig. 4a), although

for many of these, the fold differences were small. When the top 40 most differentially expressed proteins, selected on the basis of the statistical confidence in the differential expression, were used to generate a heatmap, a clear segregation by hierarchical clustering between the two



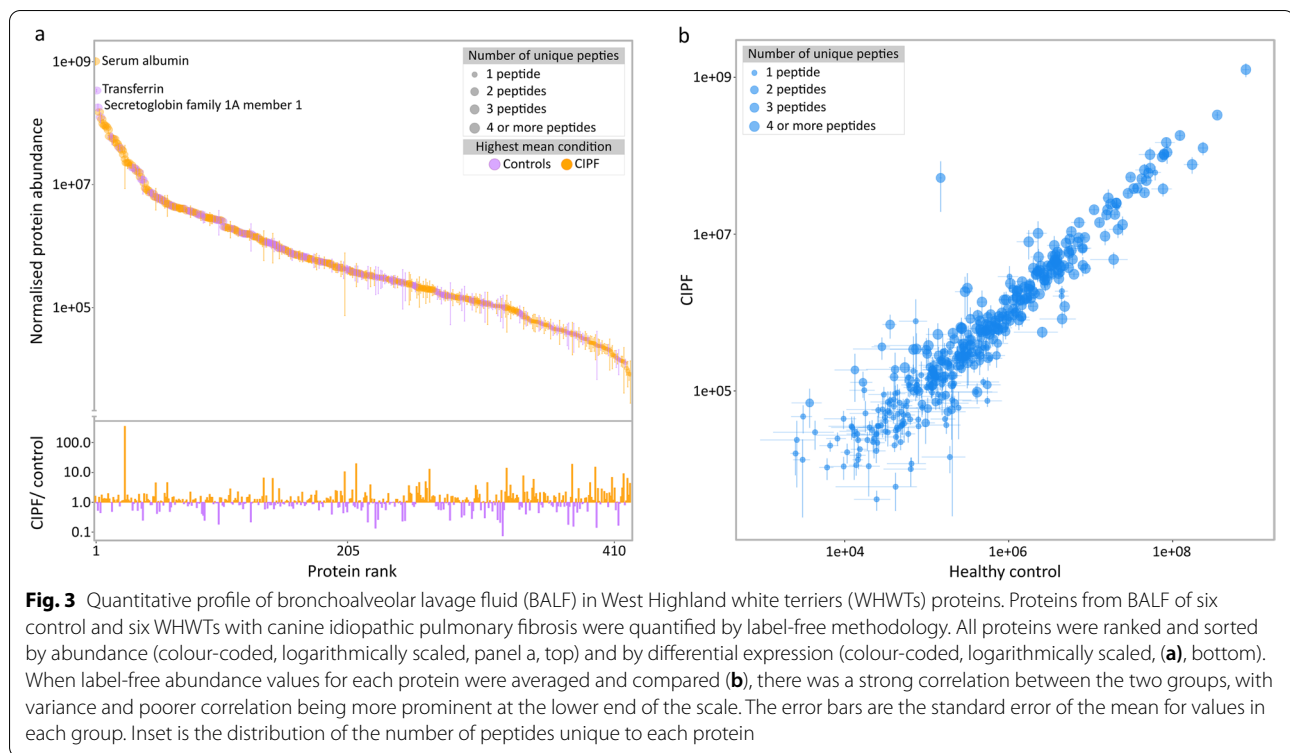
cohorts was evident, with proteins being either elevated or diminished in CIPF (Fig. 4b). Principal component analysis (PCA) of the normalized, z-scaled proteome data resolved CIPF and healthy WHWTs in the first two principal components with no overlap, explaining 34% and 19% of the variance respectively (total 53%, Fig. 4c.) There were clear proteome differences between BALF samples from healthy WHWTs and WHWTs with CIPF, although the CIPF group showed higher intragroup variability. The most obvious separation between healthy and CIPF groups was in PC2. The top 20 proteins that contributed to PC2, ranked according to the magnitude of the contribution are plotted in Supplementary Fig. 1.

The proteins with the largest magnitudes of change, either elevated or decreased in CIPF WHWTs, are presented in Fig. 5. The proteins that were most highly increased in CIPF WHWTs compared to healthy WHWTs included S100A12 (19-fold, Supplementary Table 1), cathelicidin (8.8-fold), Ly-6 like CD177 (5.6-fold), transforming growth factor beta 1 induced protein

(2.4-fold), serpin G1 (2-fold) and inter-alpha trypsin inhibitor (1.9-fold). Additionally, all three fibrinogen chains were elevated between 1.8 and 3.8-fold in CIPF WHWTs. The proteins that were most significantly decreased in CIPF compared to healthy WHWTs include carboxylic acid ester hydrolase (5-fold), non-enzymatic chitinase 3 like 1 (2-fold) and secretoglobulin (2-fold).

**Protein evidence for reflux aspiration in WHWTs with CIPF**

In the second part of this study, we used proteomics to discover potential proteins that could be indicative of reflux. First, a sample of gastric juice and one induced vomit sample from healthy dogs of other breeds were analysed to create a list of putative stomach-derived proteins. From gastric juice, 68 proteins were identified after trypsin digestion. Similarly, for the induced vomit sample, trypsin digestion revealed 63 proteins. Collectively, 126 proteins were identified in both gastric juice and induced vomit (Supplementary Table 2). Proteins of interest that might be used as potential markers of reflux



aspiration were selected based on a high level of expression in the stomach. These included the trefoil factor family (TFF), gastrokinins (GKNs), IgGfc-binding protein, pepsin A and gastric triacylglycerol lipase. Of note, these same proteins were identified in canine gastric juice in a separate study [13].

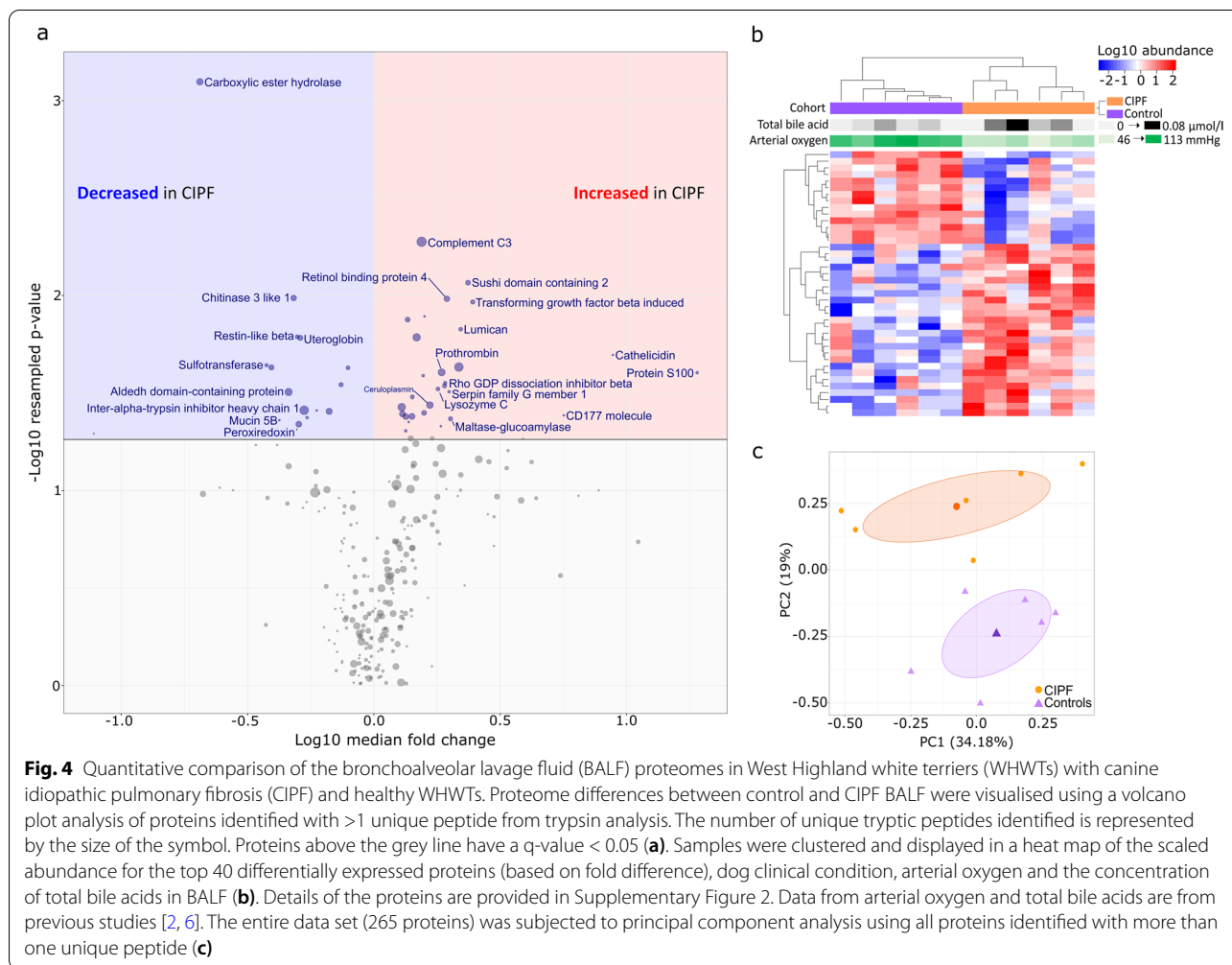
Trefoil factor 1 (TFF1) was detected with three unique peptides in digests of gastric juice and with two unique peptides in vomitus. Additionally, it was detected in trypsin digests of BALF by a single peptide and a cognate miscleavage product containing the same peptide. No significant difference was detected in TFF1 levels between CIPF and healthy WHWTs ( $P=0.08$ ). Trefoil factor 2 (TFF2) was detected with three unique peptides in gastric juice and with one unique peptide in vomitus but was not identified in BALF. Trefoil factor 3 (TFF3) was not identified in gastric juice or vomitus but was identified in tryptic digests of BALF. This protein was reduced in CIPF ( $P=0.03$ ). Gastrokine 2 (GKN2) was identified in gastric juice by three unique peptides. However, GKNs could not be identified in BALF from control or CIPF WHWTs. IgGfc-binding protein was identified in gastric juice with 11 unique peptides and vomitus (22 unique peptides). There was a broad distribution of values, but with no difference between the abundance of this protein in the two groups ( $P=0.37$ ). Pepsin A was detected in gastric juice by two peptides, and in induced vomit by a single peptide

using trypsin digestion. We were not able to detect pepsin A in BALF samples of CIPF or healthy WHWTs. To explore the possibility of pepsin being present in BALF but not being identified because of the paucity of peptides, we generated extracted ion chromatograms for the candidate peptides but were unable to detect peptides from pepsin A in BALF of WHWTs (data not shown). Finally, gastric triacylglycerol lipase was identified in gastric juice with three unique peptides, but this protein was not detected in any BALF sample.

## Discussion

In the first part of this study, we evaluated BALF proteomic profiles in CIPF and healthy WHWTs. As an accessible biofluid for proteomics, BALF can yield several hundred protein identifications, although it has a relatively simple SDS-PAGE profile reflecting a strong bias to serum albumin. Functional enrichment analysis of BALF core proteome in WHWTs highlighted the higher representation of core BALF proteins as secretory or exosomal proteins, which is consistent with the fluid being rich in purposefully secreted proteins rather than because of cellular damage. In particular, the strongest scoring molecular function categories were associated with endopeptidase inhibition, consistent with the demand for regulation of ectopic proteolysis and modulation of coagulation/fibrinolysis, in addition to complement

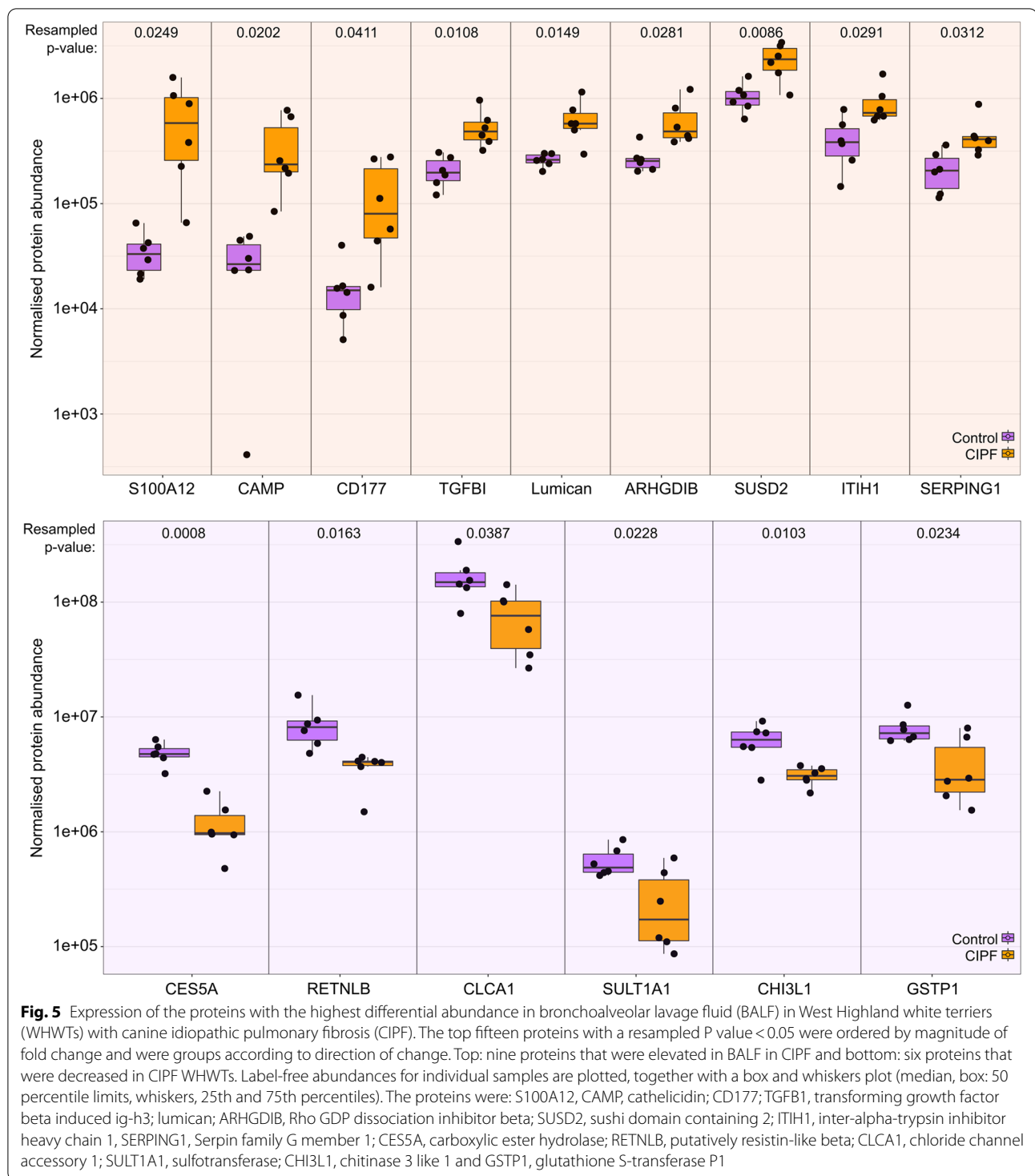




activation. As anticipated, the most abundant protein was serum albumin. The next five most abundant proteins were transferrin, AHSG, SCGB1A1, CLCA1 and haptoglobin. These high abundant proteins are a mixture of plasma proteins, presumably from plasma exudate in the respiratory/digestive system, and proteins that are elevated and secreted by lung tissue [14].

Additionally, we compared the proteome of BALF in CIPF and healthy WHWTs to find potential protein biomarkers for CIPF. A previous study explored BALF proteomic profile in CIPF WHWTs, dogs with chronic bronchitis (CB) and healthy dogs with 2-D differential gel electrophoresis [15]. In this study, six proteins (complement C3,  $\alpha$ -1-antitrypsin, apolipoprotein A-1, haptoglobin,  $\beta$ -actin and transketolase) were significantly upregulated in both CIPF and CB dogs compared with healthy dogs. However, no specific CIPF biomarkers were identified. In our study, we observed proteomic differences between BALF of CIPF and healthy WHWTs. The CIPF group showed higher intragroup variability,

probably reflecting the variability in disease presentation by individual dogs. Expression of several proteins was increased in CIPF compared to healthy WHWTs. One of these, the neutrophil-derived protein S100A12, is a major modulator of the inflammatory response [16]. Elevated expression of cathelicidin and Lys-6 like CD177 in CIPF WHWTs is consistent with an activation of inflammatory response and of neutrophil function [17, 18]. Elevated TGF- $\beta$ 1 induced protein is consistent with an effect on cell proliferation and the onset of fibrotic changes [19, 20]. Increased TGF- $\beta$ 1 signaling has also been detected previously in CIPF WHWTs compared with healthy controls and TGF- $\beta$  is considered a key mediator in CIPF of WHWTs [21, 22]. Elevation of serpin G2 and inter-alpha trypsin inhibitor are consistent with a proteinase:antiproteinase reaction that might act to suppress fibrinolysis, complement activation and coagulation. Indeed, elevated levels of all three fibrinogen chains were detected in CIPF WHWTs. These proteins were consistent with an increase in plasma exudate in the



BALF from CIPF patients, although the lack of a similar increase in the most abundant plasma proteins, such as albumin and transferrin, implied a degree of selectivity.

The proteins that were most significantly decreased in CIPF compared to healthy WHWTs included carboxylic

acid ester hydrolase, non-enzymatic chitinase 3-like 1 and secretoglobulin. Carboxylic acid ester hydrolase is relevant to fatty acyl and sterol ester metabolism and capable of cleaving thioester or amide bonds. This enzyme is expressed in lung in secreted form but

its precise role there is unknown [23]. Interestingly, the non-enzymatic chitinase 3-like 1 has been implicated in the inflammatory response to respiratory virus infection where it is elevated in the inflammatory phase [24], and in the fibroproliferative phase in mouse models of pulmonary fibrosis [25]. The decrease in this protein is thus somewhat of a contradiction. Secretoglobulin, secreted by club cells, is the most abundant protein in airway fluid and the decrease in CIPF is consistent with an impairment of lung function. Indeed, a decrease in secretoglobulin production could reduce the ability of lung to resist cytokine surges [26].

To establish whether WHWTs were aspirating gastrointestinal contents, we examined BALF for the presence of proteins that were putatively derived from the upper digestive tract. In our study, 126 proteins were identified in both gastric juice and induced vomit, which concurs with the median number of proteins in canine gastric fluid in a previous study [13]. In a previous study [6], we found bile acids, a commonly used marker of reflux aspiration, in BALF from both age matched CIPF and healthy WHWTs but not in healthy Beagles.

One commonly used protein marker of reflux aspiration in humans is pepsin [8]. We were able to detect peptides derived from pepsin in gastric juice and vomitus, even though pepsin is of relatively low abundance in canine gastric fluid [13] compared to human gastric fluid [27]. However, when we examined the liquid chromatography tandem mass spectrometry (LC-MS/MS) traces for the presence of these peptides in tryptic digests of BALF, they were no different to baseline signals, suggesting that these peptides were absent or below the lower limit of detection. The identification of pepsin by tryptic digestion and LC-MS/MS is difficult. It has disproportionately few basic amino acids as it is required to function in the acidic environment of the stomach lumen; it thus possesses a low isoelectric point [28]. Because trypsin cleaves proteins at lysine and arginine residues, most peptides generated from pepsin by tryptic digestion are incompatible with an LC-MS/MS workflow meaning that only around 7% of the pepsin sequence (two peptides) could be expected to be identified using this method.

We could find no strong evidence for the presence of other secreted gastric proteins in BALF. Trefoil factor family proteins, extracellular, secreted proteins widely expressed in mucin producing cells are abundantly expressed in the gastrointestinal epithelium [29]. These are robust proteins that are resistant to acid, and thus, the acidic milieu of the gastric contents. Three TFF proteins are reported in dog (TFF 1–3, [30]). TFFs have been proposed to play a role in mucosal defence and healing [29] and TFF1 is strongly expressed in canine stomach ([https://bgee.org/?page=gene&gene\\_id=ENSCAFG000](https://bgee.org/?page=gene&gene_id=ENSCAFG000)

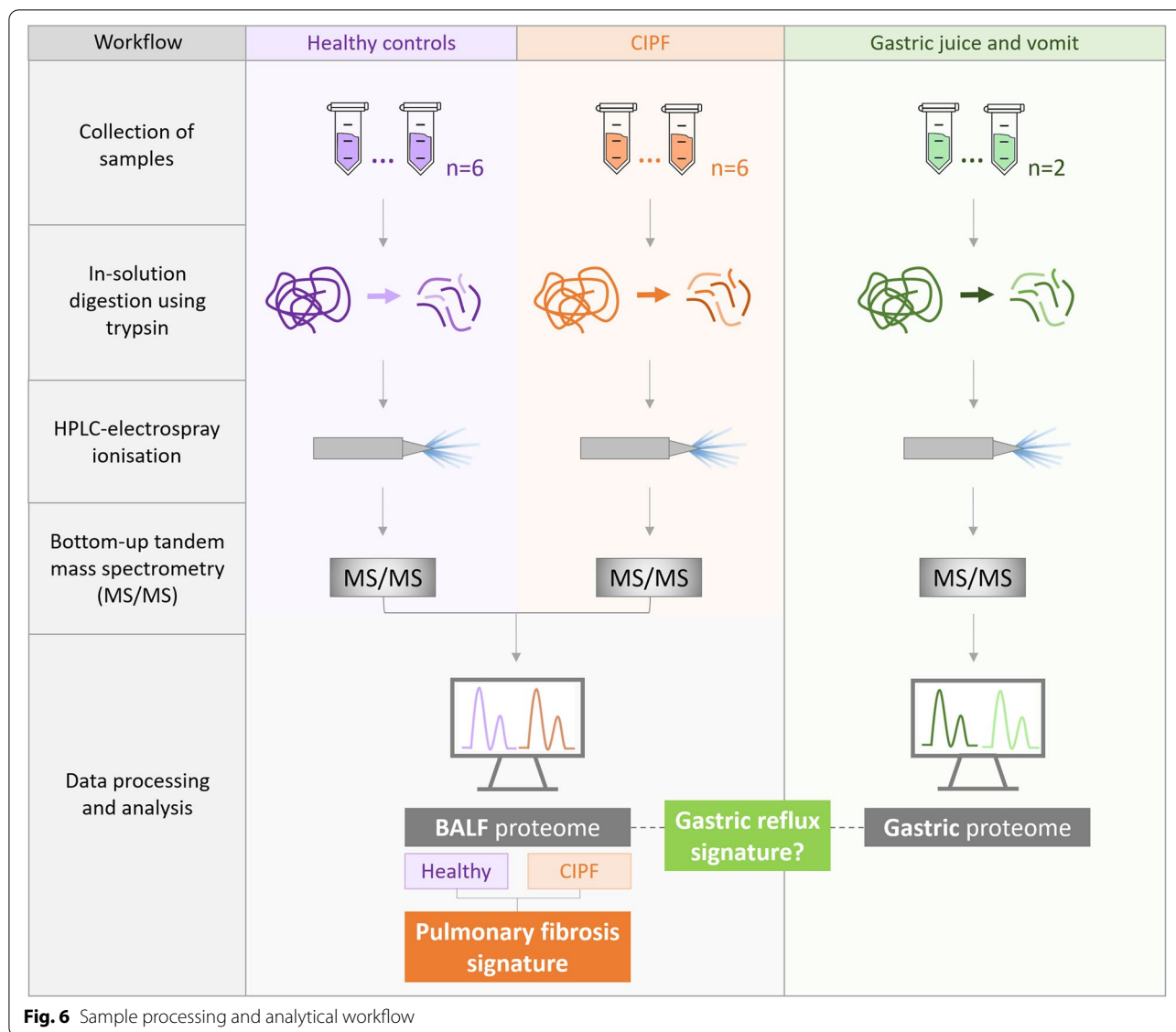
00010351). Although we were able to detect TFF1 in canine gastric juice and vomitus and in BALF samples of CIPF and healthy WHWTs, no significant differences were detected between groups. Similarly, gastrokines, stomach-specific, secreted proteins that interact with TFF proteins [31, 32], could not be identified in BALF in healthy or CIPF WHWTs. One protein that was identified through multiple peptides, with both proteases in gastric juice and vomitus, was E2RH46\_CANLF, a large, secreted, putative glycoprotein of ‘uncharacterised function’ that is flagged as being expressed in goblet cells. It might therefore meet the criterion of a suitable, stomach-derived protein to monitor reflux aspiration. This protein was detected in tryptic digests of BALF both in CIPF and healthy WHWTs but there was a broad distribution of values, and no difference in the abundance of this protein was observed between the two groups. Collectively, these data support the notion that reflux aspiration is challenging to detect in CIPF, as also noted by Grobman et al. [13]. However, group sizes were small. Also, only one gastric juice and one vomit sample were available and were not from WHWTs, although our data aligned well with an independent study [13]. Secondly, BALF samples from healthy dogs of other breeds than WHWT were not examined. Additionally, even though healthy WHWTs were thoroughly examined and considered to have healthy lungs, there is a possibility that some of them might have been in an early, subclinical stage of CIPF. Lastly, as with TFF1 and E2RH46 in this study, no significant difference in bile acid concentrations in BALF between CIPF and healthy WHWTs was detected in a previous study [6] leaving a possibility that also healthy WHWTs might aspirate refluxate. Therefore, adding two groups in comparison representing nonfibrotic lung disease and healthy dogs both of another breed might be informative.

In a conclusion, label-free proteomics allowed discrimination between CIPF and healthy WHWTs, consistent with inflammatory and fibrotic processes in lungs, but could not find clear evidence for reflux aspiration. Targeted measurement of key proteins identified herein may provide a proteomics signature of CIPF that could be used to evaluate treatment options and progress. However, evidence for reflux aspiration remains rather conjectural at this stage.

## Methods

The overall sample workflow is presented in Fig. 6. In the first part of this study, we evaluated proteomic profile of BALF samples from six CIPF and six healthy WHWTs. In the second part of this study, we conducted a proteome analysis of gastric juice and induced vomit from single individuals, with the goal





of building a list of abundant, secreted gastrointestinal proteins that might be of value as markers of reflux aspiration. In this first pass, we allowed identifications that were based on single, high-quality peptide matches to maximise the list of candidates. Systematic exploration of the literature relating to these proteins and exploring the homologues in the Human Protein Atlas (<https://www.proteinatlas.org/>) revealed that the majority of these would not provide the tissue specificity to allow them to be used to assess reflux aspiration. Putative candidates were then used to focus a search in BALF samples in CIPF and healthy WHWTs, ( $n = 6$ , both groups).

**Study subjects**

A total of 6 CIPF WHWTs (median age 11.6 years) and 6 healthy WHWTs (median age 8.3 years) were retrospectively selected for this study (Supplementary Table 4). All WHWTs were thoroughly examined; clinical workup included physical examination, hematology, serum biochemistry, arterial blood gas analysis, fecal sample analysis, cervicothoracic radiographs, bronchoscopy with BALF sampling and high-resolution computed tomography. CIPF diagnosis was verified with postmortem lung histopathology in all dogs. Owners reported no signs indicating gastrointestinal disease in any of the dogs at the time of sample collection. Additionally, two healthy

dogs of other breeds were included in this study (gastric juice sample: Belgian shepherd, 7.5 years, female and vomit sample: Australian shepherd, 10.9 years, female). All dogs were client-owned pet dogs and were examined in the Veterinary Teaching Hospital of Helsinki University.

#### Collection of samples

Bronchoscopy and BALF sampling were performed under general anesthesia. BALF samples were collected from left and right caudal lung lobes using sterile saline (2 ml/kg/lobe). Supernatant was separated by centrifugation and stored at  $-80^{\circ}\text{C}$  [33]. Gastric fluid sample was collected during gastroscopy by flushing with saline and vomit sample by inducing vomiting with apomorphine (0.2 mg/kg). Both were stored similarly as BALF samples.

#### Rapigest in-solution digestion

BALF, gastric juice and vomit samples were diluted in 25 mM ammonium bicarbonate to a final volume of 160  $\mu\text{l}$  to yield a total protein concentration between 0.11 and 0.13 mg/ml. Samples were treated with a 1% (w/v) solution of Rapigest SF Surfactant (Waters) to a final concentration of 0.05% (w/v) before incubation at  $80^{\circ}\text{C}$  for 10 min. Disulfide bonds were reduced by the addition of dithiothreitol (DTT) to a final concentration of 3 mM and incubated for 10 min at  $60^{\circ}\text{C}$ . Iodoacetamide (IA) was added to a final concentration of 9 mM to alkylate the free sulfhydryl groups on the cysteine residues. After incubation for 30 min at room temperature in the dark, trypsin was added to a 50:1 protein:protease ratio and incubated overnight at  $37^{\circ}\text{C}$ . Digestion was stopped by the addition of trifluoroacetic acid (TFA) to a final concentration of 0.5% (v/v) and incubated for 45 min at  $37^{\circ}\text{C}$ . The samples were centrifuged to remove particulates.

#### Liquid chromatography-mass spectrometry

Peptides were analysed by LC-MS/MS using an Ultimate 3000 nano system (Dionex/Thermo Fisher Scientific, Hemel Hempstead, UK) coupled to a Q-Exactive-HF mass spectrometer (Thermo Fisher Scientific, Hemel Hempstead, UK) to acquire the masses of peptides derived from proteins extracted from the BALF samples. Peptides were loaded onto a trap column (50 cm) (Acclaim Pepmap 100) at 12  $\mu\text{l}/\text{min}$  over 7 min with an aqueous solution containing 0.1% (v/34v) TFA and 2% (v/v) acetonitrile. The trap column was set in-line with an analytical column (Easy-Spray Pepmap<sup>®</sup> RSLC) (Dionex). Peptides were eluted by applying a linear gradient of 3.8–50% solvent B (acetonitrile 80% (v/v) with 0.1% (v/v) formic acid FA

over 35 min at 300 nl/min. Solvent A was HPLC grade water with 0.1% (v/v) formic acid. The mass spectrometer was operated in data dependent positive (ESI+) mode and full scan MS spectra (350–2000 m/z) were acquired in the Orbitrap. The 16 most intense multiply charged ions ( $z \geq 2$ ) were sequentially isolated and fragmented in the octopole collision cell by high energy collisional dissociation (HCD) and detected in the Orbitrap.

#### Protein identification and analysis

Samples were searched against a database comprised of all canine protein sequences in the Uniprot Uni\_Canine database (45,397 proteins) using trypsin as the specified enzyme, carbamidomethylation of cysteine as fixed modification, methionine oxidation as variable modification and one trypsin missed cleavage, precursor and fragment ion error tolerances were set to 10 ppm and 0.01 Da, respectively. The false discovery rate (FDR) was calculated using the decoy database tool in MASCOT. Only those proteins identified with an FDR < 1% were accepted. Label-free quantification, including run alignment and peak picking, was carried out in Progenesis QI for Proteomics v4 (Nonlinear Dynamics), while database searching was carried out by MASCOT search engine (Matrix Science version 2.6.0.). The data exported from Progenesis was imported into Simplifi (simplifi.protefi.com) for rapid visualisation and calculation of non-parametric statistics and the output table from Simplifi is presented as Supplementary Table 3. The Simplifi visualisation of these data is available at <https://simplifi.protefi.com/#/p/f1d8cd90-2416-11eb-8cd5-fb853c2621cb>. Functional enrichment analysis of BALF core proteome were performed using g:Profiler (<https://biit.cs.ut.ee/gprofiler/gost>). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository [34] with the dataset identifier PXD028916.

#### Statistical analysis

Data were analysed and visualised in the R environment, using RStudio version 4.0.3 (updated 2020-10-10). Heatmap with hierarchical clustering and principal component analysis (PCA) was completed in RStudio using heatmap3 and prcomp functions, respectively. Protein abundance data were log transformed for both analyses. The PCA data was additionally univariate scaled.

#### Abbreviations

BALF: Bronchoalveolar lavage fluid; CIPF: canine idiopathic pulmonary fibrosis; IPF: idiopathic pulmonary fibrosis; LC-MS/MS: liquid chromatography tandem mass spectrometry; WHWT: West Highland white terrier.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12917-022-03202-x>.

**Additional file 1: Supplementary Figure 1.** Abundance profiles for the 20 proteins making the largest contributions to PC2 in the principal components analysis (Figure 4, main text).

**Additional file 2: Supplementary Figure 2.** Abundance profiles for the 40 proteins with the largest fold differences (Figure 4b, main text).

**Additional file 3: Supplementary Table 1.** The core BALF proteome.

**Additional file 4: Supplementary Table 2.** Gastric juice and induced vomit proteins.

**Additional file 5: Supplementary Table 3.** Output table from Simplifi.

**Additional file 6: Supplementary Table 4.** Patient data.

### Acknowledgements

We are grateful to Dr Philip Brownridge for exceptional instrument support and to Victoria Harman who assisted with training and methods development for RM. We also wish to acknowledge Thomas Spillmann and, Marcus Candido for collection of canine gastric fluid and Sanna Viitanen for canine vomit sample. Additionally, we would like to thank Ninna Koho for research assistance.

### Authors' contributions

MR, PM, RB and MM conceived and designed the study. HL collected BALF samples from WHWTs. All proteomics analysis, including data analysis, was conducted by RM. All authors contributed to writing the manuscript and approved the final version.

### Authors' information

None.

### Funding

This work was funded by The Finnish Foundation of Veterinary Research, The Finnish Veterinary Foundation (Finland) and the University of Liverpool, who had no part in the design of the study and collection, analysis, and interpretation of data, nor writing of the manuscript. Open access is funded by Helsinki University Library.

### Availability of data and materials

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository [34] with the dataset identifier PXD028916. Username: [reviewer\\_pxd028916@ebi.ac.uk](mailto:reviewer_pxd028916@ebi.ac.uk). Password: fGhtVAcA

### Declarations

#### Ethics approval and consent to participate

The study protocol was approved by the Ethics Committee for Animal Experimentation at Helsinki University, Finland (statements 5B/2008, 4/2014 and 5/2015) and by the Committee for Experimental Animals of Southern Finland (ESAVI/1005/04.10.03/2011, ESAVI/9116/04.10.07/2014, ESAVI/7383/04.10.07/2013, and ESLH-2008-05403/Ym-23). Dog owners provided an informed consent to allow the use of the samples. We confirm that all methods were performed in accordance with the relevant guidelines and regulations. The study is reported in accordance with ARRIVE guidelines.

#### Consent for publication

Not applicable.

#### Competing interests

None.

#### Author details

<sup>1</sup>Centre for Proteome Research, Institute of Systems and Integrative Biology, University of Liverpool, Liverpool, United Kingdom. <sup>2</sup>Department of Equine

and Small Animal Medicine, Faculty of Veterinary Medicine, University of Helsinki, Helsinki, Finland. <sup>3</sup>Department of Child Health (University of Liverpool), Institute in the Park, Alder Hey Children's Hospital, Eaton Rd, L12 2AP Liverpool, United Kingdom.

Received: 16 October 2021 Accepted: 3 March 2022

Published online: 30 March 2022

### References

- Corcoran, B. M.; Cobb, M.; Martin, M. W.; Dukes-McEwan, J.; French, A.; Fuentes, V. L.; Boswood, A., Rhind, S.: Chronic pulmonary disease in West Highland white terriers. *Vet Rec* 1999, *144*:611–616, doi: <https://doi.org/10.1136/vr.144.22.611>
- Heikkilä, H. P.; Lappalainen, A. K.; Day, M. J.; Clercx, C.; Rajamäki, M. M.: Clinical, bronchoscopic, histopathologic, diagnostic imaging, and arterial oxygenation findings in West Highland White Terriers with idiopathic pulmonary fibrosis. *J Vet Intern Med* 2011, *25*:433–439, doi: <https://doi.org/10.1111/j.1939-1676.2011.0694.x>
- Lilja-Maula, L. I.; Laurila, H. P.; Syrjä, P.; Lappalainen, A. K.; Krafft, E.; Clercx, C.; Rajamäki, M. M.: Long-term outcome and use of 6-minute walk test in West Highland White Terriers with idiopathic pulmonary fibrosis. *J Vet Intern Med* 2014, *28*:379–385, doi: <https://doi.org/10.1111/jvim.12281>
- Raghu, G.; Richeldi, L.; Jagerschmidt, A.; Martin, V.; Subramaniam, A.; Ozoux, M. L.; Esperet, C. A.; Soubrane, C.: Idiopathic Pulmonary Fibrosis: Prospective, Case-Controlled Study of Natural History and Circulating Biomarkers. *Chest* 2018, *154*:1359–1370, doi: <https://doi.org/10.1016/j.chest.2018.08.1083>
- Syrjä, P.; Heikkilä, H. P.; Lilja-Maula, L.; Krafft, E.; Clercx, C.; Day, M. J.; Rönty, M.; Myllärniemi, M.; Rajamäki, M. M.: The histopathology of idiopathic pulmonary fibrosis in West Highland white terriers shares features of both non-specific interstitial pneumonia and usual interstitial pneumonia in man. *J Comp Pathol* 2013, *149*:303–313, doi: <https://doi.org/10.1016/j.jcpa.2013.03.006>
- Määttä, O. L. M.; Laurila, H. P.; Holopainen, S.; Lilja-Maula, L.; Melamies, M.; Viitanen, S. J.; Johnson, L. R.; Koho, N.; Neuvonen, M.; Niemi, M.; Rajamäki, M. M.: Reflux aspiration in lungs of dogs with respiratory disease and in healthy West Highland White Terriers. *J Vet Intern Med* 2018, *32*:2074–2081, doi: <https://doi.org/10.1111/jvim.15321>
- Lee, J. S.: The Role of Gastroesophageal Reflux and Microaspiration in Idiopathic Pulmonary Fibrosis. *Clin Pulm Med* 2014, *21*:81–85, doi: <https://doi.org/10.1097/cpm.0000000000000031>
- Houghton, L. A.; Lee, A. S.; Badri, H.; Devault, K. R.; Smith, J. A.: Respiratory disease and the oesophagus: reflux, reflexes and microaspiration. *Nat Rev Gastroenterol Hepatol* 2016, *13*:445–460, doi: <https://doi.org/10.1038/nrgastro.2016.91>
- Trinick, R.; Southern, K. W.; Mcnamara, P. S.: Assessing the Liverpool Respiratory Symptom Questionnaire in children with cystic fibrosis. *Eur Respir J* 2012, *39*:899–905, doi: <https://doi.org/10.1183/09031936.00070311>
- Lee, J. S.; Song, J. W.; Wolters, P. J.; Elicker, B. M.; King, T. E.; Kim, D. S.; Collard, H. R.: Bronchoalveolar lavage pepsin in acute exacerbation of idiopathic pulmonary fibrosis. *Eur Respir J* 2012, *39*:352–358, doi: <https://doi.org/10.1183/09031936.00050911>
- Savarino, E.; Carbone, R.; Marabotto, E.; Furnari, M.; Sconfienza, L.; Ghio, M.; Zentilin, P.; Savarino, V.: Gastro-oesophageal reflux and gastric aspiration in idiopathic pulmonary fibrosis patients. *Eur Respir J* 2013, *42*:1322–1331, doi: <https://doi.org/10.1183/09031936.00101212>
- Fastrès, A.; Roels, E.; Vangrinsven, E.; Taminiau, B.; Jabri, H.; Bolen, G.; Merveille, A. C.; Tutunaru, A. C.; Moysse, E.; Daube, G.; Clercx, C.: Assessment of the lung microbiota in dogs: influence of the type of breed, living conditions and canine idiopathic pulmonary fibrosis. *BMC Microbiol* 2020, *20*:84, doi: <https://doi.org/10.1186/s12866-020-01784-w>
- Grobman, M.; Rindt, H.; Reiner, C. R.: Proteomic Characterization of Canine Gastric Fluid by Liquid Chromatography-Mass Spectrometry for Development of Protein Biomarkers in Regurgitation, Vomiting, and Cough. *Front Vet Sci* 2021, *8*:670007, doi: <https://doi.org/10.3389/fvets.2021.670007>
- Wattiez, R.; Falmagne, P.: Proteomics of bronchoalveolar lavage fluid. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, *815*:169–178, doi: <https://doi.org/10.1016/j.jchromb.2004.10.029>

15. Lilja-Maula, L. I.; Palviainen, M. J.; Heikkilä, H. P.; Raekallio, M. R.; Rajamäki, M. M.: Proteomic analysis of bronchoalveolar lavage fluid samples obtained from West Highland White Terriers with idiopathic pulmonary fibrosis, dogs with chronic bronchitis, and healthy dogs. *Am J Vet Res* 2013, *74*:148–154, doi: <https://doi.org/10.2460/ajvr.74.1.148>
16. Xia, C.; Braunstein, Z.; Toomey, A. C.; Zhong, J.; Rao, X.: S100 Proteins As an Important Regulator of Macrophage Inflammation. *Front Immunol* 2017, *8*:1908, doi: <https://doi.org/10.3389/fimmu.2017.01908>
17. Herr, C.; Shaykhiev, R.; Bals, R.: The role of cathelicidin and defensins in pulmonary inflammatory diseases. *Expert Opin Biol Ther* 2007, *7*:1449–1461, doi: <https://doi.org/10.1517/14712598.7.9.1449>
18. Xie, R. F.; Hu, P.; Wang, Z. C.; Yang, J.; Yang, Y. M.; Gao, L.; Fan, H. H.; Zhu, Y. M.: Platelet-derived microparticles induce polymorphonuclear leukocyte-mediated damage of human pulmonary microvascular endothelial cells. *Transfusion* 2015, *55*:1051–1057, doi: <https://doi.org/10.1111/trf.12952>
19. Frangiannis, N.: Transforming growth factor- $\beta$  in tissue fibrosis. *J Exp Med* 2020, *217*:e20190103, doi: <https://doi.org/10.1084/jem.20190103>
20. Saito, A.; Horie, M.; Nagase, T.: TGF- $\beta$  Signaling in Lung Health and Disease. *Int J Mol Sci* 2018, *19*:E2460, doi: <https://doi.org/10.3390/ijms19082460>
21. Krafft, E.; Lybaert, P.; Roels, E.; Laurila, H. P.; Rajamäki, M. M.; Farnir, F.; Myllärniemi, M.; Day, M. J.; Mc Entee, K.; Clercx, C.: Transforming growth factor beta 1 activation, storage, and signaling pathways in idiopathic pulmonary fibrosis in dogs. *J Vet Intern Med* 2014, *28*:1666–1675, doi: <https://doi.org/10.1111/jvim.12432>
22. Lilja-Maula, L.; Syrjä, P.; Laurila, H. P.; Sutinen, E.; Rönty, M.; Koli, K.; Rajamäki, M. M.; Myllärniemi, M.: Comparative study of transforming growth factor- $\beta$  signalling and regulatory molecules in human and canine idiopathic pulmonary fibrosis. *J Comp Pathol* 2014, *150*:399–407, doi: <https://doi.org/10.1016/j.jcpa.2013.12.001>
23. Gabriele, M.; Puccini, P.; Lucchi, M.; Vizziello, A.; Gervasi, P. G.; Longo, V.: Presence and inter-individual variability of carboxylesterases (CES1 and CES2) in human lung. *Biochem Pharmacol* 2018, *150*:64–71, doi: <https://doi.org/10.1016/j.bcp.2018.01.028>
24. Kim, M. J.; Shim, D. H.; Cha, H. R.; Moon, K. Y.; Yang, C. M.; Hwang, S. J.; Kim, K. W.; Park, J. H.; Lee, C. G.; Elias, J. A.; Sohn, M. H.; Lee, J. M.: Chitinase 3-like 1 protein plays a critical role in respiratory syncytial virus-induced airway inflammation. *Allergy* 2019, *74*:685–697, doi: <https://doi.org/10.1111/all.13661>
25. Zhou, Y.; Peng, H.; Sun, H.; Peng, X.; Tang, C.; Gan, Y.; Chen, X.; Mathur, A.; Hu, B.; Slade, M. D.; Montgomery, R. R.; Shaw, A. C.; Homer, R. J.; White, E. S.; Lee, C. M.; Moore, M. W.; Gulati, M.; Lee, C. G.; Elias, J. A.; Herzog, E. L.: Chitinase 3-like 1 suppresses injury and promotes fibroproliferative responses in Mammalian lung fibrosis. *Sci Transl Med* 2014, *6*:240ra76, doi: <https://doi.org/10.1126/scitranslmed.3007096>
26. Xu, M.; Yang, W.; Wang, X.; Nayak, D. K.: Lung Secretoglobin Scgb1a1 Influences Alveolar Macrophage-Mediated Inflammation and Immunity. *Front Immunol* 2020, *11*:584310, doi: <https://doi.org/10.3389/fimmu.2020.584310>
27. Paulo, J. A.; Lee, L. S.; Wu, B.; Repas, K.; Banks, P. A.; Conwell, D. L.; Steen, H.: Proteomic analysis of endoscopically (endoscopic pancreatic function test) collected gastroduodenal fluid using in-gel tryptic digestion followed by LC-MS/MS. *Proteomics Clin Appl* 2010, *4*:715–725, doi: <https://doi.org/10.1002/prca.201000018>
28. Andreeva, N. S.; James, M. N.: Why does pepsin have a negative charge at very low pH? An analysis of conserved charged residues in aspartic proteinases. *Adv Exp Med Biol* 1991, *306*:39–45, doi: [https://doi.org/10.1007/978-1-4684-6012-4\\_4](https://doi.org/10.1007/978-1-4684-6012-4_4)
29. Aihara, E.; Engevik, K. A.; Montrose, M. H.: Trefoil Factor Peptides and Gastrointestinal Annurev. *Annu Rev Physiol* 2017, *79*:357–380, doi: <https://doi.org/10.1146/annurev-physiol-021115-105447>
30. Campbell, B. G.; Jabbes, M.: Canine and feline trefoil factor family peptides: highly conserved molecules with some unique characteristics. *Res Vet Sci* 2008, *85*:68–73, doi: <https://doi.org/10.1016/j.rvsc.2007.08.008>
31. Sharman, M. J.; Bacci, B.; Santos, L.; Mansfield, C. S.: Gastrokine mRNA expression in gastric tissue from dogs with helicobacter colonisation but without inflammatory change during treatment. *Vet Immunol Immunopathol* 2017, *187*:28–34, doi: <https://doi.org/10.1016/j.vetimm.2017.03.004>
32. Menheniott, T. R.; Kurklu, B.; Giraud, A. S.: Gastrokines: stomach-specific proteins with putative homeostatic and tumor suppressor roles. *Am J Physiol Gastrointest Liver Physiol* 2013, *304*:G109–21, doi: <https://doi.org/10.1152/ajpgi.00374.2012>
33. Rajamäki, M. M.; Järvinen, A. K.; Saari, S. A.; Maisi, P. S.: Effect of repetitive bronchoalveolar lavage on cytologic findings in healthy dogs. *Am J Vet Res* 2001, *62*:13–16, doi: <https://doi.org/10.2460/ajvr.2001.62.13>
34. Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.: The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic acids research* 2019, *47*:D442–D450, doi: <https://doi.org/10.1093/nar/dkz115>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

