

RESEARCH ARTICLE

Plasma proteomics can discriminate isolated early from dual responses in asthmatic individuals undergoing an allergen inhalation challenge

Amrit Singh^{1,2,3}, Gabriela V. Cohen Freue^{2,3,4}, Jean L. Oosthuizen^{1,2}, Sarah H. Y. Kam^{1,2}, Jian Ruan^{1,2}, Mandeep K. Takhar^{2,3}, Gail M. Gauvreau⁵, Paul M. O'Byrne⁵, J. Mark FitzGerald^{2,6,7}, Louis-Philippe Boulet⁸, Christoph H. Borchers⁹ and Scott J. Tebbutt^{1,2,3,6}

¹ James Hogg Research Centre, St. Paul's Hospital, University of British Columbia, Vancouver, Canada

² Institute for HEART+LUNG Health, Vancouver, Canada

³ NCE CECR PROOF Centre of Excellence, Vancouver, Canada

⁴ Department of Statistics, UBC, Vancouver, Canada

⁵ Department of Medicine, McMaster University, Hamilton, Canada

⁶ Department of Medicine, Division of Respiratory Medicine, UBC, Vancouver, Canada

⁷ Vancouver Coastal Health Research Institute, Vancouver General Hospital, Vancouver, Canada

⁸ Centre de Pneumologie de L'Hôpital, Université Laval, Sainte-Foy, Canada

⁹ University of Victoria – Genome BC Proteomics Centre, University of Victoria, Victoria, Canada

Purpose: This proteomics study was designed to determine the utility of iTRAQ MALDI-TOF/TOF technology to compare plasma samples from carefully phenotyped mild, atopic asthma subjects undergoing allergen inhalation challenge.

Experimental design: Eight adult subjects with mild, allergic asthma (four early responders (ERs) and four dual responders (DRs)) participated in the allergen inhalation challenge. Blood samples were collected prior to and 2 h after the inhalation challenge. Sixteen plasma samples (two per subject), technical replicates, and pooled controls were analyzed using iTRAQ. Technical validation was performed using LC-MRM/MS. Moderated robust regression was used to determine differentially expressed proteins.

Results: Although this study did not show significant differences between pre- and post-challenge samples, discriminant analysis indicated that certain proteins responded differentially to allergen challenge with respect to responder type. At pre-challenge, fibronectin was significantly elevated in DRs compared to ERs and remained significant in the multiple reaction monitoring validation.

Conclusions and clinical relevance: This proof of principle demonstration has shown that iTRAQ can uncover differences in the human plasma proteome between two endotypes of asthma and merits further application of iTRAQ to larger cohorts of asthma and other respiratory diseases.

Keywords:

Allergen inhalation challenge / Early response / Fibronectin / Late response

Received: February 23, 2012

Revised: June 21, 2012

Accepted: July 30, 2012



Correspondence: Dr. Scott J. Tebbutt, UBC James Hogg Research Centre, Room 166, St. Paul's Hospital, 1081 Burrard Street, Vancouver, BC V6Z 1Y6, Canada

E-mail: Scott.Tebbutt@hli.ubc.ca

Fax: 604-806-9274

Abbreviations: DR, dual responder; ER, early responder; FEV₁, Forced expiratory volume in one second; FN1, Fibronectin; sPLS-DA, sparse partial least squares discriminant analysis

1 Introduction

Atopic asthmatic individuals, following inhalational exposure to allergen, develop differential early, and dual (early plus late) response profiles [1,2]. Bronchoconstriction, representing the early phase of the asthmatic response, of which the onset can be detected within ten minutes of inhalation, reaches a

Colour Online: See the article online to view Figs. 1–3 in colour.

maximum within 30 min, and resolves within 3 h. In 50–60% of atopic, asthmatic, adult subjects, the early response is followed by the late-phase asthmatic response, which usually starts between 3 and 4 h after allergen inhalation challenge [3], and is characterized by cellular inflammation of the airway, increased bronchovascular permeability, and mucus secretion. The development of the late response is associated with the hallmark features of chronic asthma. These include induction of bronchial hyperresponsiveness, long-lasting bronchoconstriction that is difficult to reverse and resolves slowly over a period of days, and inflammation of the airways. This is in contrast to the early response that is very easily reversed with bronchodilators and has minimal associated airway inflammation or subsequent airway hyperresponsiveness. Individuals that develop an isolated early asthmatic response are classified as early responders (ERs) whereas individuals that also develop a late asthmatic response in addition to the early response are classified as dual responders (DRs).

Currently, the only method used to discriminate ERs and DRs is through spirometry (forced expiratory volume in one second; FEV₁) measurements following inhalation challenge. It is very important to identify the critical pathways leading to the late (inflammatory) response, in order to find ways to inhibit it. Despite tremendous interest, the differences between the pathways leading to the dual response and those leading to the isolated early response are not completely understood. Understanding these differences is important in elucidating the underlying disease mechanisms, and improving phenotypic classification and patient care.

The use of proteomics in the field of respiratory diseases such as asthma and chronic obstructive lung disease has been limited and is attributed to the poor phenotypic selection of patients [4]. Commonly used biological samples include bronchoalveolar lavage fluid [5, 6], sputum [7], and exhaled breath condensate [8, 9]. Proteomic profiling of plasma [10], urine, nasal lavage fluid [11], saliva [12], and bronchial biopsies [13] is less common [4]. The present study uses iTRAQ 4800 MALDI-TOF/TOF proteomics technology to determine changes in the human plasma proteome in mild atopic asthmatic individuals undergoing allergen inhalation challenge. Although this study did not show significant proteomic differences between pre- and post-challenge, significant proteins were identified when comparing ERs and DRs.

2 Materials and methods

2.1 Asthma cohort

This prospective study was approved by the Institutional Review Boards of the participating health research institutes. Eight adult subjects were recruited, following written informed consent. Inclusion criteria included non-smokers with stable, mild atopic asthma, free of other lung diseases. Subjects had a FEV₁ of greater than 70% of predicted, and

baseline methacholine PC₂₀ (the provocative concentration of methacholine causing a 20% fall in FEV₁) of less than 16 mg/mL [14]. The eight subjects were selected based on their response type, and balanced for other demographics, including sex and age, in order to reduce possible confounding. All eight subjects were chosen due to their hypersensitivity to cat allergen (via skin prick test) in order to avoid additional confounding based on the use of different allergens [15].

2.2 Allergen inhalation challenge

Each subject reported to the laboratory on three consecutive days for a methacholine challenge test (MCT) on days 1 and 3, and an allergen inhalation challenge on day 2. The MCT was performed according to Cockcroft et al. and [PC₂₀] (mg/mL) was calculated [16]. Results of the MCT were used to set the initial dosage of cat allergen and to evaluate changes in airway hyperresponsiveness after allergen challenges. A step-wise approach was used to administer allergen dosage in doubling concentrations until an initial drop in FEV₁ of at least 20% was observed [17]. Subjects in which the immediate asthmatic response resolved back to baseline within 3 h of onset were classified as ERs. Subjects with a subsequent FEV₁ drop of 15% between 3 and 7 h post-challenge were classified as DRs. Subjects having an FEV₁ drop of 10% that was still falling at the end of the 7-h observation period were categorized as DRs if they also demonstrated a doubling dose shift in airway responsiveness at 24-h postallergen [2]. Ten milliliters of venous blood (EDTA tubes) was obtained from each subject just prior to the allergen challenge (pre-) and approximately 2–3 h post-challenge. Leukocyte enriched (buffy coat) and plasma fractions (in multiple aliquots) were freshly prepared from the blood and frozen at –80°C. Plasma samples for analysis were subsequently thawed on ice, and centrifuged at 1600 × g at 4°C for 10 min, before the supernatant was extracted and shipped on dry ice to the Genome BC/University of Victoria Proteomics Centre for proteomic analysis.

2.3 iTRAQ MALDI-TOF/TOF methodology

2.3.1 Plasma samples

With the exception of dual responder subject 3 (DR3; see Table 1), 200 µL of pre-challenge samples and 200 µL of post-challenge samples were used in the iTRAQ experiment. For DR3, 200 µL of the pre-challenge sample and six aliquots of 120 µL each of the post-challenge sample were used in order to assess the reproducibility of the iTRAQ technology. Lastly, three pooled reference samples of 200 µL each were prepared prior to immunodepletion by combining equal aliquots of the pre- and post-challenge samples for the eight subjects. The 24 samples were analyzed using 3 × 8-plex iTRAQ experiments (see Supporting Information Table S1 for experimental design). The samples were randomly labeled

Table 1. Demographics of subjects undergoing allergen inhalation challenge.

Patient ID	Age (year)	Sex (M:F)	Allergen	Pre [PC ₂₀] (mg/mL)	Post [PC ₂₀] (mg/mL)	Allergen-induced shift ^{b)}	% Fall in FEV1	
							Early	Late
ERs								
1	28	F	Cat pelt	12.8	ND	ND	20.3	4.8
2	34	F	Cat pelt	2.8	6.1	0.4	21	1.5
3	27	M	Cat pelt	4.5	1.8	2.5	34.4	0
4	42	F	Cat hair	5.3	8.6	0.6	42.1	11.1
Mean ± SE	32.8 ± 3.4	1:3		5.4 ^{a)}	4.6 ^{a)}	1.2 ± 0.66	29.5 ± 5.3	4.4 ± 2.5
DRs								
5	23	F	Cat hair	0.3	0.2	1.5	38.9	31.8
6	26	F	Cat hair	5.1	1.5	3.4	31.4	14.9
7	49	F	Cat hair	3.6	1.0	3.6	25.3	12.6
8	26	M	Cat hair	0.9	1.0	0.9	31.5	15.6
Mean ± SE	31.0 ± 6.0	1:3		1.5 ^{a)}	0.7 ^{a)}	2.4 ± 0.68	31.8 ± 2.8	18.7 ± 4.4*

ND, Not determined; all subjects were challenged with cat allergen.

a) Geometric mean (PC₂₀ values are measured on a log scale).

b) [PC₂₀]_{pre}/[PC₂₀]_{post}.

**p* < 0.05 versus ER group.

Note: Subjects 5–8 are also referred to as DR1, DR2, DR3, and DR4.

using iTRAQ reagents. Each run included the pooled reference plasma sample to allow comparisons between the three iTRAQ experiments.

2.3.2 Protocol

Briefly, plasma samples were depleted of the 14 most abundant proteins (Albumin, Immunoglobulins (G, A, M), α₁-Antitrypsin, Transferrin, Heptaglobin, α₁-Acid Glycoprotein, α₂-Macroglobulin, Apolipoprotein A-I, Apolipoprotein A-II, Fibrinogen, Apolipoprotein B, and Complement C3) using immune-affinity chromatography in order to enhance mass spectrometer detection sensitivity (Genway Biotech, San Diego, CA). After tryptic digestion, multiple peptides were labeled (Applied Biosystems, Foster City, CA), pooled and acidified to pH 2.5–3.0 with concentrated phosphoric acid (ACP Chemicals Inc., Montreal, Quebec, Canada). Labeled peptides were separated by 2D LC, spotted directly onto 384-spot MALDI ABI 4800 plates, and analyzed by a 4800 MALDI-TOF/TOF mass spectrometer (Applied Biosystems); acquisition time ranged from 35 to 40 h. Full details on the iTRAQ methodology can be found in [18].

2.3.3 iTRAQ data analysis

Peptide data were processed using ProteinPilot™ software v3.0 with the integrated Paragon™ search algorithm and Pro Group™ algorithm [19] (Applied Biosystems) 3.0.0.0, searching against the International Protein Index (IPI HUMAN v3.69, 87,129 entries) database [20]. All entries in the entire database were searched. The Paragon algorithm uses prob-

abilities in order to estimate parameters for modifications, substitutions, and the number of missed cleavages [21].

ProteinPilot assembles peptide data into protein groups (local) for each iTRAQ experiment containing related redundant proteins, and estimates the protein ratios: levels of subject sample labels relative to the reference sample label. These ratios are computed using the weighted geometric means of individual peptides contributing to the identification of a protein group and have been corrected using the Auto Bias option in the Pro Group algorithm. Proteins with an Unused Protein Score greater than 2 (99% confidence level) were considered for downstream analyses. An Unused Protein Score indicates the number of unique peptide spectra contributing only to the corresponding identified protein.

The protein group code algorithm (PGCA) assembles proteins groups (global) by connecting overlapping local protein groups across multiple iTRAQ experiments [22]. Using PGCA as a preprocessing tool to analyze proteomics data ensures the comparison of related proteins, including homologous proteins, redundant proteins, and proteins from the same families, if they could not be distinguished based on the observed peptide data. Further details on the iTRAQ data processing methodology and PGCA algorithm can be found in the Supporting Information Material [22].

2.4 Technical validation using LC-MRM/MS

Multiple reaction monitoring (MRM) assays consisting of one peptide per protein were used to perform technical validation of significant proteins identified using iTRAQ. Additional information regarding sample preparation, and data acquisition and analysis can be found in the Supporting Information Material.

2.4.1 Sample preparation and LC-MRM/MS analysis

Technical validation of selected proteins was performed using mass spectrometry based MRM [23]. The plasma samples underwent tryptic digestion, after which acidified stable-isotope-labeled peptide mixture with formic acid was added. An AB/MDS Sciex 4000 QTRAP with a nano-electrospray ionization source controlled by Analyst 1.5 software (Applied Biosystems) was used for all LC-MRM/MS analyses.

2.4.2 MRM data analysis

All MRM data were processed using MultiQuant 1.2 (Applied Biosystems) with the MQL algorithm for peak integration. A *p*-value cut-off of 0.05 was used to determine significant proteins.

2.5 Statistical analyses

Protein groups (PGs) that were detectable across all samples were used for all statistical analyses.

2.5.1 Differential expression

Moderated robust regression in the Linear Models for Microarrays (limma) package [24] was used to determine statistically significant PGs and MRM proteins using a Benjamini–Hochberg false discovery rate (BH-FDR) of 5% [25]. Limma uses a moderation factor [24] that shrinks the variance of each PG toward a common value in order to attenuate the effect of small sample sizes. The robust version of this test uses an M-estimator [26] to reduce the effects of outliers on the parameters of interest. Both methods are ideal for cohorts with small sample sizes where the risk of false positives is propagated. Since the sample size was very limited in this study with three variables of interest; time (pre-challenge versus post-challenge), response (ER versus DR) and the interaction (time:response), each comparison was tested in a separate linear model, one variable at a time (a total of three linear models) in order to avoid over fitting. All statistical analyses were implemented in R version 2.14.2 [27].

2.5.2 Sparse partial least squares discriminant analysis (sPLS-DA)

Classification of subjects into their respective phenotypic class using a subset of the total number of PGs was performed using sPLS-DA [28] using the mixOmics library. sPLS-DA is a supervised classification method in which the predictor is a high dimensional $n \times p$ dataset, X ; n = number of samples, p = number of features (e.g. PGs) and the response is a $n \times 1$ dummy vector, Y indicating the class labels of the individuals. Variable selection is achieved using lasso penalizations result-

ing in sparse loading vectors. The number of sparse loading vectors is equal to the number of retained dimensions (specified by the user). The length of each loading vector is equal to p features in the X dataset. Loading vectors contain non-zero weights corresponding to discriminative features and zero weights for non-discriminative features. The number of features with non-zero weights on each loading vector is equal to the number of variables specified for each dimension. In addition, the features with non-zero weights on each loading vector may or may not overlap. Therefore the two variables to tune for this algorithm are (i) the number of dimensions and (ii) the number of variables to select on each dimension. The average misclassification error rate using the maximum distance method was used in order to determine the optimal number of dimensions and features to retain.

3 Results

3.1 Subject demographics

Subject demographics are presented in Table 1. There were no significant differences in terms of age, sex, and initial (early) drop in FEV₁ between ERs and DRs. DRs had a maximum% fall in FEV₁ of $18.7 \pm 4.4\%$ during the late phase versus $4.4 \pm 2.5\%$ drop in ERs ($p < 0.05$). DRs also demonstrated a >2 doubling dose shift in methacholine PC₂₀ at 24-h postallergen, as compared to a <1 doubling dose shift in ERs.

3.2 Quality of data

In order to limit the number of false protein identifications from the list of observed peptides, a protein confidence interval of 99% was used in this initial discovery study. A total of 43 protein groups (PGs), corresponding to 14 depleted proteins (see Supporting Information Table S2) were removed. PGs corresponding to immunoglobulins were also removed. A total of 142 PGs were identified across the 21 samples (Supporting Information Table S3). Eighty-four of the 142 PGs were detectable across all samples and retained for downstream analyses. The dictionary file used to map local groups (within an iTRAQ experiment) to global groups (between independent iTRAQ experiments) is provided in Supporting Information Table S4.

3.3 Reproducibility of iTRAQ

Six technical replicates from dual responder subject 3 (DR3; see Table 1) at post-challenge were included in the iTRAQ study design in order to assess the reproducibility of the iTRAQ technology. Figure 1 shows the scatter plot matrices for six samples [six DR3s at post-challenge (two per iTRAQ run)]. A weak correlation was seen between experimental replicates (replicates from different iTRAQ experiments). However, a strong correlation was seen

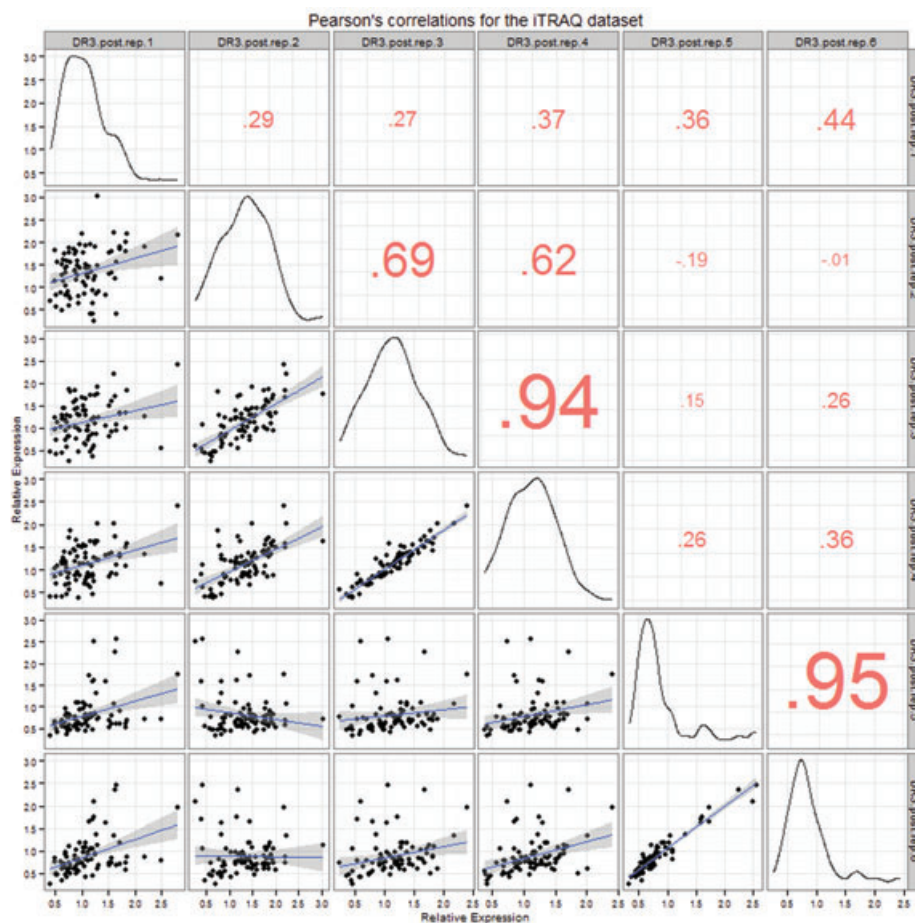


Figure 1. Intraassay variability of iTRAQ. Pearson correlations between six technical replicates in the iTRAQ dataset (84 protein groups $\times$ six samples). The within-run correlations are greater than between-run correlations, with the exception of run 1 (DR3.post.rep.1 and DR3.post.rep.2).

between within-run replicates, with the exception of run 1. The median value of PGs across all iTRAQ experiments was used for the DR3 post-challenge sample in subsequent analyses.

3.4 Statistical analyses

Limma revealed no significant changes in the levels of any of the 84 PGs, at a BH-FDR of 5% when comparing the eight pre- and eight post-challenge samples (Supporting Information Table S5).

A separate independent linear model considering the interaction between time and response [change (post minus pre) in protein expression in ERs versus change in protein expression in DRs (Δ ERs versus Δ DRs)] revealed no significant PGs at a BH-FDR of 5%, although 16 PGs had p -values less than 0.05 (Supporting Information Table S6). Instead, sPLS-DA was used to determine a subset of the 84 PGs that could discriminate ERs from DRs. In order to choose the optimal number of dimensions and variables to retain, the average classification error rate was estimated using a tenfold cross-validation averaged over 50 permutations (Supporting Information Fig. S1). Supporting Information Fig. S1 indi-

cated that the lowest misclassification rate corresponded to four PGs for each of the two retained dimensions. Therefore, sPLS-DA was able to classify the 16 samples (eight subjects) into their respective response-class, using eight PGs (four PGs on each of the two loading vectors; see Section 2.5.2). PCA applied to the dataset consisting of the 8 PGs resulted in the clear segregation of ERs and DRs (Fig. 2A). The correlation circle in Fig. 2B displays the eight PGs into two inversely correlated groups (Group 1 and Group 2). PGs in each group are positively correlated with PGs in the same group but negatively correlated with PGs in the other group. This is shown in Fig. 2C where Group 1 PGs are all up-regulated in DRs and down-regulated in ERs. This is opposite to the pattern found for PGs in Group 2, where PGs are down-regulated in DRs and up-regulated in ERs after challenge.

Four significant (BH-FDR of 5%) PGs were identified when comparing ERs and DRs at pre-challenge (Table 2). Elevated levels of plasma proteins identified in DRs relative to ERs included Alpha-1B-glycoprotein (A1BG), Fibronectin (FN1), and Transthyretin (TTR). Inter-alpha-inhibitor H4 (ITIH4) was reduced in DRs relative to ERs at pre-challenge. The complete list of differential expressed proteins between ERs and DRs at pre-challenge can be found in Supporting Information Table S7).

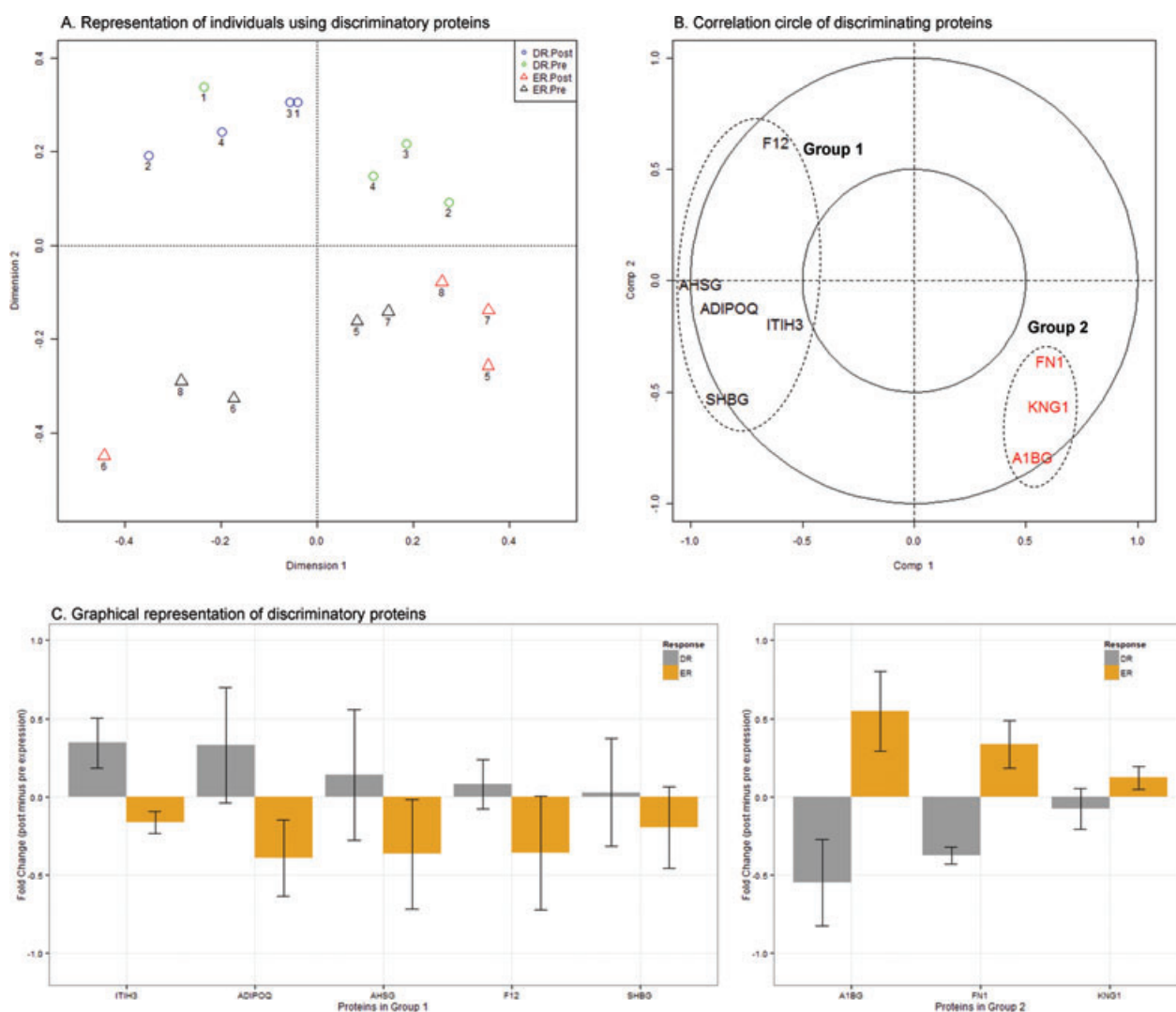


Figure 2. PCA on a dataset consisting of eight discriminatory proteins for 16 samples. (A) Individual plot showing clear separation of early and DRs. (B) A correlation circle displaying two clusters (group 1 and group 2) of protein groups. (C) Graphical representation of discriminatory proteins in group 1 and group 2.

3.5 Technical validation

Technical validation of A1BG, FN1, TTR, and ITIH was performed using MRM. Calculated concentrations (pmol/μL) of protein targets for selected proteins can be found in Supporting Information Table S8 (note; the tube number is identical to that found in Supporting Information Table S1). FN1 was the only protein that successfully validated. Figure 3 shows elevated levels of FN1 in the plasma of DRs compared to ERs prior to allergen inhalation challenge.

4 Discussion

Asthma is a complex disease encompassing many phenotypic clusters such that proteomic tools must be applied indepen-

dently to various endotypes of asthma [4]. For the present study, a cohort of ERs and DRs was considered. Classification of these individuals into their respective phenotypic classes was made using FEV₁ as well as PC₂₀ measurements (Table 1). Table 1 also indicates that the cohort contained equal numbers of ERs and DRs, matched for age and sex, and were of Caucasian descent. In addition, given the allergen specificity of asthmatic responses [15], only cat allergen was administered during these challenges. Due to the large number of potential confounding variables, a limited sample size of eight asthmatic individuals was used in this study. For each individual, a blood sample was collected prior to allergen challenge and 2-h post allergen inhalation resulting in a greater than 20% drop in FEV₁. This was in line with our previous study that showed significant changes in the peripheral blood transcriptome of individuals with mild asthma

Clinical Relevance

Individuals with atopic (allergic) asthma display characteristic physiological response profiles when they inhale an allergen to which they are sensitized. The nature of these responses has a profound influence on subsequent clinical course. Some individuals develop an isolated early asthmatic response while others also go on to develop a late asthmatic response (dual responders, DRs). In any individual, the pattern of response is generally consistent. Since it is the consequences of the late response that are be-

lieved to contribute to chronic airway inflammation and uncontrolled disease it is important to determine the molecular mechanisms responsible for this fundamentally different response. In the study reported here, we have used an unbiased proteomics analytical approach on blood samples donated by well-phenotyped atopic asthmatic individuals who are undergoing experimentally controlled allergen inhalation challenges.

2 h after allergen inhalation challenge [29]. This exploratory study used iTRAQ 4800 MALDI-TOF/TOF proteomics technology to determine changes in the human plasma proteome in eight asthmatic individuals undergoing allergen inhalation challenge.

In this proof of principle study, we restricted the analysis to only those proteins that were identified with a high confidence (Unused Protein Score ≥ 2 equal to a 99% confidence level) to reduce the variation due to protein identification. In addition, to compensate for the small sample size, moderated test statistics in the Linear Models of Microarrays (limma) [24]

library were used. Limma uses shrinkage methods such as “Empirical Bayes” that borrow information across features (protein groups) making the analyses stable for a small number of samples. The initial analysis using limma, comparing eight pre-challenge samples with eight post-challenge samples yielded no differential discovery at a BH-FDR of 5%. Similarly no PGs were found differentially expressed in the interaction between time and response (Δ ERs versus Δ DRs), which may be due to the small number of samples used in this study. However using sPLS-DA, a supervised learning method with variable selection, a panel of eight PGs was

Table 2. Differentially expressed proteins identified using iTRAQ at pre-challenge using a BH-FDR of 5%

PG	Accession	Gene symbol	Protein name	p-value	Relative fold change ^{a)}
46	IPI:IPI00022895.7	A1BG	Alpha-1B-glycoprotein	$p < 0.001$	+2.90
16	IPI:IPI00845263.1	FN1	Fibronectin 1 isoform 2 preproprotein	$p < 0.001$	+2.19
	IPI:IPI00339225.1		Isoform 5 of fibronectin		
	IPI:IPI00339224.2		Fibronectin 1 isoform 4 preproprotein		
	IPI:IPI00022418.1		Isoform 1 of fibronectin		
	IPI:IPI00339223.1		Isoform 3 of fibronectin		
	IPI:IPI00339227.4		Isoform 7 of fibronectin		
	IPI:IPI00339228.1		Isoform 8 of fibronectin		
	IPI:IPI00414283.6		Isoform 9 of fibronectin		
	IPI:IPI00479723.4		FN1 protein		
	IPI:IPI00855777.1		Isoform 14 of fibronectin		
	IPI:IPI00855785.1		Isoform 15 of fibronectin		
	IPI:IPI00867588.1		Isoform 13 of fibronectin		
5	IPI:IPI00896413.1	ITIH4	Inter-alpha (globulin) inhibitor H4 isoform 2 precursor	0.001	−0.53
	IPI:IPI00218192.3		Isoform 2 of Inter-alpha-trypsin inhibitor heavy chain H4c		
	IPI:IPI00896419.3		Isoform 1 of Inter-alpha-trypsin inhibitor heavy chain H4		
	IPI:IPI00922043.2		cDNA FLJ51742, highly similar to Inter-alpha-trypsin inhibitor heavy chain H4		
	IPI:IPI00944960.1		ITIH4 protein		
33	IPI:IPI00855916.1	TTR	Transthyretin	0.002	+1.89
	IPI:IPI00022432.1		Transthyretin		
	IPI:IPI00940791.1		20 kDa protein		

a) average expression in DR/average expression in ER.

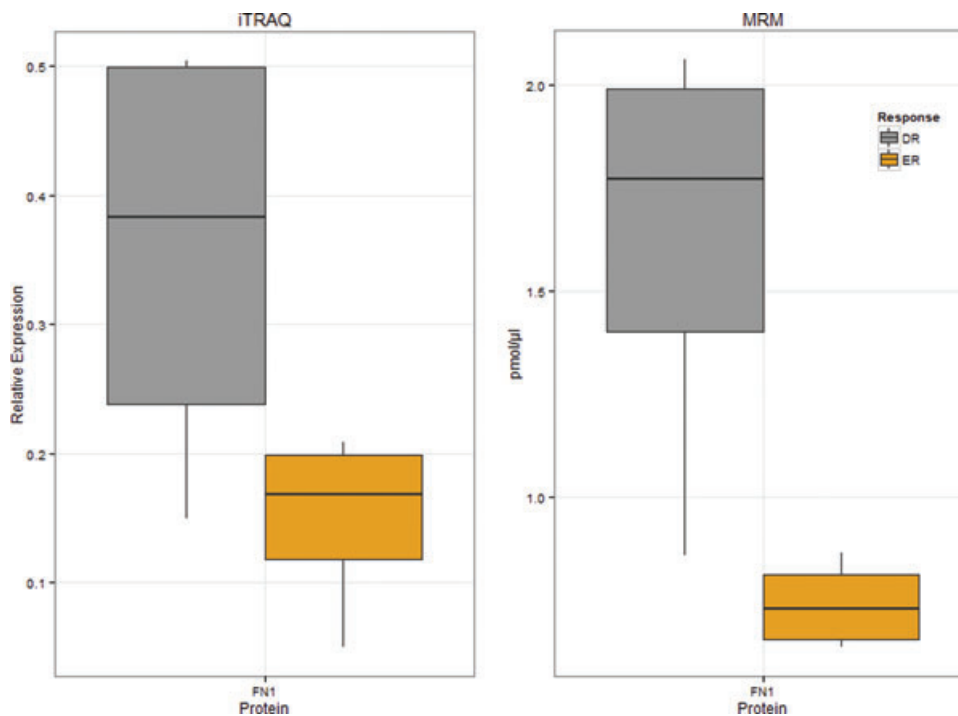


Figure 3. Technical validation of fibronectin using LC-MRM/MS. Fibronectin was significantly elevated in DRs compared to ERs in both the iTRAQ and MRM datasets, respectively.

identified that could classify subjects into ERs and DRs (Figs. 2A and B). These PGs consisted of two groups that responded differentially to the allergen exposure with respect to responder-type (Fig. 2C). These results indicate that panels of PGs may be more efficient at discriminating ERs from DRs than individual PGs, but require validation using independent test sets.

At pre-challenge, TTR, ITIH4, A1BG, and FN1 were differentially expressed (BH-FDR of 5%) comparing ERs and DRs (Table 2). Previous proteomic studies using human plasma samples have indicated that TTR increases after house dust mite challenge in asthmatic individuals with an immediate phase response (early response) compared to asthmatic non-responders [30]. On the other hand, ITIH4 was under expressed in the plasma of asthmatic individuals relative to controls [31]. However, both these proteins were not successful in the technical validation using LC-MRM/MS. It is also important to note that ITIH4 was not available as an MRM assay, and therefore a general peptide mapping to inter alpha trypsin inhibitor (ITIH) was used. A1BG, also not validated ($p = 0.08$), has been previously identified in bronchoalveolar lavage fluid but did not significantly change following segmental allergen challenge [32]. FN1 was the only significant ($p < 0.05$) protein in the technical validation using LC-MRM/MS (Fig. 3).

FN1 is a 460 kDa [33] glycoprotein that functions in cell adhesion, cell migration, phagocytosis, and as a chemo-attractant of inflammatory cells [34]. The concentration of FN1 is dependent on both age and sex [35]. Plasma FN1 levels in DRs were double that found in ERs prior to undergoing the allergen inhalation challenge. The protein group correspond-

ing to FN1 remained significant ($p = 0.02$) in an ordinary least-squares linear model, using age and sex as covariates. Previous studies have shown that airway hyperresponsiveness seen in asthma can be attributed to the cell-mediated contraction of collagen fibrils due to increased FN1 deposition in the extracellular matrix [36]. In addition, increased deposition of FN1 in the subepithelium leads to the thickening of the airway walls of patients with mild atopic asthma [37]. Based on these findings, one can speculate that increased FN1 deposition in DRs can explain the increased airway hyperresponsiveness of these subjects. However, in a previous study using bronchial biopsies from patients with asthma, fibronectin expression measured by densitometry was not significantly correlated with lung function measures [38]. A follow-up study using endobronchial biopsies of atopic mild asthmatic patients and nonatopic healthy controls stated that fibronectin expression in the airway smooth muscle layer was not different between asthmatics and controls [39]. This suggests that our finding should be taken with caution given the small size ($n = 8$) used for discovery. In addition, the MRM validation was made using one peptide mapping to FN1, whereas current studies are using up to five peptides in MRM-based multiplexed quantitation of FN1 (Table 2 in [40]). Biological validation in larger independent cohorts is required.

Using an unbiased approach to proteomic profiling in the human plasma proteome, we have shown that careful phenotyping of subjects is critical if discovery is to be made. Although this study did not show significant differences between pre- and post-challenge samples, significant proteins were identified with differential levels between carefully se-

lected ERs and DRs. The ability of FN1 to discriminate ERs from DRs is novel in asthma, but warrants additional validation in larger cohorts. In addition, given the poor intraassay variability of iTRAQ (separate tryptic digestions on the same plasma sample) in this study, a larger study is required to validate these results (Fig. 1). Collection of blood samples is ongoing and larger sample sizes will be available to us in the future with additional sources of funding.

Despite the limitations inherent to this exploratory study, we have demonstrated that iTRAQ MALDI-TOF/TOF proteomics technology can uncover differences in the human plasma proteome between two endotypes of asthma and merits further application of iTRAQ to larger cohorts of asthma and other respiratory diseases.

The authors wish to thank the eight research participants for taking part in these studies, as well as the research coordinators and study nurses as part of the AllerGen Clinical Investigator Collaborative. We would also like to thank the University of Victoria – Genome BC Proteomics Centre in Victoria, BC for performing the iTRAQ MALDI-TOF/TOF and MRM experiments. The Centre is supported by Genome Canada and Genome BC. This research was also supported by grant funding from the BC Proteomics Network (MSFHR), AllerGen NCE Inc. (Allergy, Genes and Environment Network), and the Canadian Institutes of Health Research.

The authors have declared no conflict of interest.

5 References

- [1] Bentley, A. M., Kay, A. B., Durham, S. R., Human late asthmatic reactions. *Clin. Exp. Allergy* 1997, 27 (Suppl 1), 71–86.
- [2] Gauvreau, G. M., Lee, J. M., Watson, R. M., Irani A-M. A. et al. Increased numbers of both airway basophils and mast cells in sputum after allergen inhalation challenge of atopic asthmatics. *Am. J. Respir. Crit. Care Med.* 2000, 161, 1473–1478.
- [3] Robertson, D. G., Kerigan, A. T., Hargreave, F. E., Chalmers, R. et al., Late asthmatic responses induced by ragweed pollen allergen. *J. Allergy Clin. Immunol.* 1974, 54, 244–254.
- [4] O’Neil, S. E., Lundback, B., Latvall, J., Proteomics in asthma and COPD phenotypes and endotypes for biomarker discovery and improved understanding of disease entities. *J. Proteomics* 2011, 75, 192–201.
- [5] Magi, B., Bargagli, E., Bini, L., Rottoli, P., Proteome analysis of bronchoalveolar lavage in lung diseases. *Proteomics* 2006, 6, 6354–6369.
- [6] Wattiez, R., Falmagne, P., Proteomics of bronchoalveolar lavage fluid. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2005, 815, 169–178.
- [7] Nicholas, B., Skipp, P., Mould, R., Rennard, S. et al. Shotgun proteomic analysis of human-induced sputum. *Proteomics* 2006, 6, 4390–4401.
- [8] Bloemen, K., Van Den, H. R., Govarts, E., Hooyberghs, J. et al. A new approach to study exhaled proteins as potential biomarkers for asthma. *Clin. Exp. Allergy* 2011, 41, 346–356.
- [9] Fumagalli, M., Dolcini, L., Sala, A., Stolk, J. et al. Proteomic analysis of exhaled breath condensate from single patients with pulmonary emphysema associated to alpha1-antitrypsin deficiency. *J. Proteomics* 2008, 71, 211–221.
- [10] Rhim, T., Choi, Y.-S., Nam, S. T. Uh., Park, J. S. et al. Plasma protein profiles in early asthmatic responses to inhalation allergen challenge. *Allergy* 2008, 64, 47–54.
- [11] Benson, L. M., Mason, C. J., Friedman, O., Kita, H. et al. Extensive fractionation and identification of proteins within nasal lavage fluids from allergic rhinitis and asthmatic chronic rhinosinusitis patients. *J. Sep. Sci.* 2009, 32, 44–56.
- [12] Ghafouri, B., Tagesson, C., Lindahl, M., Mapping of proteins in human saliva using two-dimensional gel electrophoresis and peptide mass fingerprinting. *Proteomics* 2003, 3, 1003–1015.
- [13] O’Neil, S. E., Sitkauskienė, B., Babusyte, A., Krisiukienė, A. et al. Network analysis of quantitative proteomics on asthmatic bronchi: effects of inhaled glucocorticoid treatment. *Respir. Res.* 2011, 12, 124.
- [14] Gauvreau, G. M., Hessel, E. M., Boulet, L. P., Coffman, R. L., O’Byrne, P. M., Immunostimulatory sequences regulate interferon-inducible genes but not allergic airway responses. *Am. J. Respir. Crit. Care Med.* 2006, 174, 15–20.
- [15] Hatzivlassiou, M., Grainge, C., Kehagia, V., Lau, L. et al., The allergen specificity of the late asthmatic reaction. *Allergy* 2010, 65, 355–358.
- [16] Cockcroft, D. W., Berscheid, B. A., Murdock, K. Y., Unimodal distribution of bronchial responsiveness to inhaled histamine in a random human population. *Chest* 1983, 83, 751–754.
- [17] O’Byrne, P. M., Dolovich, J., Hargreave, F. E., Late asthmatic responses. *Am. Rev. Respir. Dis.* 1987, 136, 740–751.
- [18] Kuzyk, M., Ohlund, B., Elliot, M., Smith, D. et al., A comparison of MS/MS-based, stable-isotope-labeled, quantitation performance on ESI-quadrupole TOF and MALDI-TOF/TOF mass spectrometers. *Proteomics* 2009, 9, 3323–3340.
- [19] Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A. et al. The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* 2007, 6, 1638–1655.
- [20] Kersey, P. J., Duarte, J., Williams, A., Karavidopoulou, Y. et al., The international protein index: an integrated database for proteomics experiments. *Proteomics* 2004, 4, 1985–1988.
- [21] Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A. et al., The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* 2007, 6, 1638–1655.
- [22] Freue, G. V., Sasaki, M., Meredith, A., Gunther, O. P. et al., Proteomic signatures in plasma during early acute renal allograft rejection. *Mol. Cell. Proteomics* 2010, 9, 1954–1967.

- [23] Kuzyk, M., Smith, D., Cross, T., Yang, J. et al., MRM based, multiplexed, absolute quantitation of 45 proteins in human plasma. *Mol. Cell. Proteomics* 2009, 8, 1860–1877.
- [24] Smyth, G. K., Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 2004, 3, Article 3.
- [25] Benjamini, Y., Hochberg, Y., Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Series B (Methodological)* 1995, 57, 289–300.
- [26] Huber, P. J., Robust estimation of a location parameter. *Annals Mathematical Stat.* 1964, 35, 73–101.
- [27] Team, R. D. C., *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria 2009.
- [28] Le Cao, K.-A., Boitard, S., Besse, P., Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* 2011, 12, 253.
- [29] Kam, S. H., Singh, A., He, J. Q., Ruan, J. et al., Peripheral blood gene expression changes during allergen inhalation challenge in atopic asthmatic individuals. *J. Asthma* 2012, 49, 219–226.
- [30] Rhim, T., Choi, Y. S., Nam, B. Y., Uh, S. T. et al., Plasma protein profiles in early asthmatic responses to inhalation allergen challenge. *Allergy* 2009, 64, 47–54.
- [31] Verrills, N. M., Irwin, J. A., He, X. Y., Wood, L. G. et al., Identification of novel diagnostic biomarkers for asthma and chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 2011, 183, 1633–1643.
- [32] Wu, J., Kobayashi, M., Sousa, E. A., Liu, W. et al., Differential proteomic analysis of bronchoalveolar lavage fluid in asthmatics following segmental antigen challenge. *Mol. Cell. Proteomics* 2005, 4, 1251–1264.
- [33] Mosher, D. F., Plasma fibronectin concentration: a risk factor for arterial thrombosis? *Arterioscler Thromb. Vasc. Biol.* 2006, 26, 1193–1195.
- [34] Yamada, K. M., Hasegawa, T., Hasegawa, E., Kennedy, D. W. et al. Fibronectin and interactions at the cell surface. *Prog. Clin. Biol. Res.* 1984, 151, 1–15.
- [35] Eriksen, H. O., Clemmensen, I., Hansen, M. S., Ibsen, K. K., Plasma fibronectin concentration in normal subjects. *Scand. J. Clin. Lab. Invest.* 1982, 42, 291–295.
- [36] Hocking, D. C., Fibronectin matrix deposition and cell contractility: implications for airway remodeling in asthma. *Chest* 2002, 122, 275S–278S.
- [37] Roche, W. R., Beasley, R., Williams, J. H., Holgate, S. T., Subepithelial fibrosis in the bronchi of asthmatics. *Lancet* 1989, 1, 520–524.
- [38] Slats, A. M., Janssen, K., van Schadewijk, A., van der Plas, D. T. et al., Expression of smooth muscle and extracellular matrix proteins in relation to airway function in asthma. *J. Allergy Clin. Immunol.* 2008, 121, 1196–1202.
- [39] Yick, C. Y., Ferreira, D. S., Annoni, R., von der Thusen, J. H. et al., Extracellular matrix in airway smooth muscle is associated with dynamics of airway function in asthma. *Allergy* 2012, 67, 552–559.
- [40] Domanski, D., Percy, A. J., Yang, J., Chambers, A. G. et al., MRM-based multiplexed quantitation of 67 putative cardiovascular disease biomarkers in human plasma. *Proteomics* 2012, 12, 1222–1243.