In-depth serum proteomics reveals biomarkers of psoriasis severity and response to traditional Chinese medicine

Meng Xu2,3, Jingwen Deng1*, Kaikun Xu3*, Tiansheng Zhu4*, Ling Han1, Yuhong Yan1, Danni Yao1, Hao Deng1, Dan Wang3, Yaoting Sun4, Cheng Chang3, Xiaomei Zhang3, Jiyu Dai3, Liang Yue4, Qiushi Zhang4, Xue Cai4, Yi Zhu4, Hu Duan3, Yuan Liu3, Dong Li3, Yunping Zhu3, Timothy R. D. J. Radstake5, Deepak M.W. Balak6, Danke Xu2*, Tiannan Guo4*, Chuanjian Lu1*, and Xiaobo Yu3*

1. Guangdong Provincial Hospital of Chinese Medicine, Guangdong Provincial Academy of Chinese Medical Sciences, Guangdong Provincial Key Laboratory of Clinical Research on Traditional Chinese Medicine Syndrome, Guangzhou 510120, China.
2. State Key Laboratory of Analytical Chemistry for Life Science, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, China.
3. State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences, Beijing Institute of Lifeomics, Beijing, China.
4. School of Life Sciences, Westlake University; Institute for Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China.
5. Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands.
6. Department of Dermatology and Allergology, University Medical Center Utrecht, Utrecht, the Netherlands.

*Contributed equally to this manuscript

Corresponding authors: yuxiaobo@mail.ncpsb.org (X. Yu), luchuanjian88@vip.sina.com (C. Lu), xudanke@nju.edu.cn (D. Xu), guotiannan@westlake.edu.cn (T. Guo)

© Ivyspring International Publisher. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC) license (https://creativecommons.org/licenses/by-nc/4.0/). See http://ivyspring.com/terms for full terms and conditions.

Received: 2018.11.01; Accepted: 2019.02.22; Published: 2019.04.13

Abstract

Serum and plasma contain abundant biological information that reflect the body’s physiological and pathological conditions and are therefore a valuable sample type for disease biomarkers. However, comprehensive profiling of the serological proteome is challenging due to the wide range of protein concentrations in serum.

Methods: To address this challenge, we developed a novel in-depth serum proteomics platform capable of analyzing the serum proteome across ~10 orders or magnitude by combining data obtained from Data Independent Acquisition Mass Spectrometry (DIA-MS) and customizable antibody microarrays.

Results: Using psoriasis as a proof-of-concept disease model, we screened 50 serum proteomes from healthy controls and psoriasis patients before and after treatment with traditional Chinese medicine (YinXieLing) on our in-depth serum proteomics platform. We identified 106 differentially-expressed proteins in psoriasis patients involved in psoriasis-relevant biological processes, such as blood coagulation, inflammation, apoptosis and angiogenesis signaling pathways. In addition, unbiased clustering and principle component analysis revealed 58 proteins discriminating healthy volunteers from psoriasis patients and 12 proteins distinguishing responders from non-responders to YinXieLing. To further demonstrate the clinical utility of our platform, we performed correlation analyses between serum proteomes and psoriasis activity and found a positive association between the psoriasis area and severity index (PASI) score with three serum proteins (PI3, CCL22, IL-12B).

Conclusion: Taken together, these results demonstrate the clinical utility of our in-depth serum proteomics platform to identify specific diagnostic and predictive biomarkers of psoriasis and other immune-mediated diseases.

Key words: Proteomics, Data-Independent Acquisition Mass Spectrometry, Antibody Microarray, Biomarker, Psoriasis
Introduction

Over the past ten years, the rapid development of high throughput omics technologies has enabled us to quickly obtain a large number of physiological and pathological information of the human body [1-3]. This information has begun to change the way that we understand and treat disease, from reductionism to a more holistic view [4].

Human serum and plasma are valuable in clinical and biological studies because they are a reservoir of secreted proteins from organs throughout the body, reflect the patient’s physiological and pathological conditions and may contain biomarkers of disease and treatment response [4-10]. Currently, there are two main strategies for analyzing the plasma/serum proteome [4]. The first “triangular” approach compares protein expression differences between normal and disease conditions using proteomics technology, and then selects the proteins with significantly altered levels as biomarkers for downstream clinical validation [4, 11]. The second “rectangular” method builds a knowledge base for a specific disease by quickly and accurately screening the serum/plasma proteomes from a large number of clinical samples across different time points, risk and drug treatment conditions. In general, this step employs mass spectrometry (MS) [5, 12]. It is then expected that the knowledge base can be mathematically modeled to diagnose and effectively treat the disease [3, 4, 13, 14].

Geyer et al. developed a plasma proteome analysis pipeline using label-free quantitative MS, which detected 284 ± 5 proteins containing > 40 FDA-approved biomarkers without removing high-abundance proteins [13]. Using the same approach for a cohort of 43 obese people who experienced weight loss, Geyer et al. showed that the levels of 93 plasma proteins changed, including proteins associated with insulin resistance[15]. Their study showed that the serum proteome can be used to evaluate and monitor metabolic diseases. Liu et al. developed another approach using SWATH-MS (Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra), a data independent acquisition (DIA) method with high quantitative accuracy and reproducibility. Using this method, they surveyed the changes of the plasma proteome in a longitudinal cohort containing 72 monozygotic and 44 dizygotic twins at intervals of 2-7 years. A total of 342 different proteins were quantitatively detected in 232 plasma samples. It was found that the plasma proteome was affected by genetic, environmental and longitudinal factors [5, 10].

Despite these technological advances, however, a comprehensive and in-depth determination of the serum or plasma proteome has remained a challenge due to the wide range of protein concentration [5, 8, 13, 16]. Compared to MS, antibody and protein microarrays have higher sensitivity, higher throughput, and are concentration independent, thus making them suitable for detecting low-abundance proteins in serum or plasma samples [11, 17-21].

In this work, we developed an in-depth serum proteomics technology platform by combining DIA-MS and customizable antibody microarrays. We then used this two-pronged approach to study the serological protein profile of psoriasis, a common immune-mediated inflammatory skin disease characterized by skin and joint impairment and medical comorbidities [22]. The disease is complex with a multifold pathogenesis affected by both hereditary and environmental factors. Moreover, psoriasis patients respond differently to therapy[23]. In the United States, 3.2% of the population has psoriasis with associated costs over $110 billion per year [24]. Comprehensive determination of the serum proteome of psoriasis patients may help better understand the pathogenesis of psoriasis as well as identify diagnosis and treatment biomarkers. To date, only a limited number of psoriasis serum proteins have been studied [25, 26].

In this study, differentially-expressed proteins discriminating healthy volunteers from psoriasis patients were identified, as well as proteins distinguishing responders from non-responders to an oral traditional Chinese medicine. Furthermore, a positive association between the psoriasis area and severity index (PASI) score with three serum proteins (PI3, CCL22 and IL-12B) was made. Our results demonstrate the feasibility of our platform to identify diagnostic and predictive disease biomarkers in serum.

Experimental Procedures

Clinical Samples

Serum from healthy volunteers, urticaria and psoriasis patients were obtained from Guangdong Provincial Hospital of Chinese Medicine [5, 27]. The detailed information pertaining to the psoriasis patients, including their demographics, co-morbidities, and medications, is provided in Tables 1 - 3. The study was approved by the ethical committees at the Beijing Proteome Research Center and Guangdong Provincial Hospital of Chinese Medicine, and performed according to the Declaration of Helsinki.
Fabrication of Psoriasis Antibody Arrays

Array design. All psoriasis-related protein targets for the psoriasis-specific antibody microarray were selected using text mining and manual curation as previously described (Figure S2)[9]. Briefly, the PubMed database (https://www.ncbi.nlm.nih.gov/) was searched for reported psoriasis plasma/serum proteins using “psoriasis” and “serum” or “plasma” as keywords. In parallel, the electronic PubMed abstract database was mined for psoriasis biomarkers using “psoriasis” and “biomarker” as keywords. The search resulted in a total of 544 abstracts and 364 candidate proteins. After three rounds of manual curation and de-redundancy, a total of 129 proteins associated with psoriasis remained.

Array printing. All antibodies (Bio-techne Ltd, MN, USA) were diluted to 0.2 mg/mL and then printed onto a 3D modified slide surface (Capital Biochip Corp, Beijing, China) in four replicates using an Arrayjet microarrayer (Roslin, UK). Phosphate-buffered saline (PBS) and bovine serum albumin (BSA, 100 µg/mL) (Sigma-Aldrich, MO, USA) were used as negative controls. Biotinylated proteins were dissolved in 500 µL of PBS containing 0.05% (w/v) Tween 20 (PBST).

Slides were washed three times, 10 min per wash, after centrifuging for 2 min at 1000 ×g for 2 min. The collected molecules were removed using a Bio-Spin column via centrifugation at 1000 ×g, and then alkylated with 40 mM iodoacetamide (IAA) at 25 °C for 1 h in the dark. The protein was sequentially digested with trypsin at a ratio of 1:40 (enzyme to substrate) for 16 hrs (4 h +12 h) at 33 °C. The tryptic peptides were then acidified with 1% trifluoroacetic acid (TFA, pH 2-3) prior to C18 desalting with Sep-Pak Vac 1cc (50 mg) C18 cartridges (Waters, MA, USA) according to the manufacturer’s protocol. Desalted peptides were then dried under vacuum and dissolved in 20 µL of MS buffer containing 0.1% formic acid and 2% acetonitrile in water (all HPLC grade). The peptide concentration was measured by Nanoscan (Analytik Jena AG, Jena, Germany) at an absorbance of A280 nm. 0.4 µg peptides were separated on a 30 min LC gradient using an analytical column (PepMap® RSLC, 75 µm × 250 mm, 2 µm 200 Å C18 particles) and injected into Thermo Fisher’s Q Exactive HF Hybrid Quadrupole Orbitrap™ (QE-HF) mass spectrometer. The DIA acquisition scheme consisted of 24 variable windows ranging from 400 to 1000 m/z. The sequential precursor isolation window setup was as follows: 410-430, 430-450, 450-470, 470-490, 490-510, 510-530, 530-560, 560-590, 590-610, 610-630, 630-660, 660-690, 690-710, 710-730, 730-750, 750-770, 770-790, 790-820, 820-860, 860-910, 910-970, 970-1000 m/z.

The resolution of MS1 was 60,000, and MS2 was 30,000. OpenSWATH (version 2.0.0 Sep 26 2017) was performed against a plasma library containing 325 proteins as described previously [30]. Finally, Pyrophet limited the peptide and protein identification to a 1% false discovery rate (FDR).

Measurement of Psoriasis Serum Proteome using DIA-MS

2 µL of serum samples were added in a 96-well plate and further diluted with lysis buffer containing 0.1 M NH4HCO3 (Sigma, MO, USA), 6 M Urea (Sigma, MO, USA), and 2 M thiourea (Sigma, MO, USA). Disulfide reduction was performed for 45 min at 33 °C, 600 rpm, with 10 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and then alkylated with 40 mM iodoacetamide (IAA) at 25 °C for 1 h in the dark. The protein was vacuum dried, then subjected to 1 μL of NHS-PEG4-Biotin (20 g/L in DMSO) (Thermo Fisher Scientific, MA, USA). After incubating for 1 h at room temperature, excess biotin molecules were removed using a Bio-Spin column via centrifugation at 1000 ×g for 2 min. The collected biotinylated proteins were dissolved in 500 µL of PBS containing 5% milk (w/v) and stored at 4 °C.

Sera screening. Antibody microarrays were assembled into an incubation tray (PEPperPRINT, Heidelberg, Germany) and blocked with 600 µL 5% milk (w/v) for 1 h at room temperature. After removing the milk, the arrays were incubated with pre-labeled serum proteins at 4 °C overnight. The slides were washed three times, 10 min per wash, with PBS containing 0.05% (w/v) Tween 20 (PBST). For detection, the arrays were incubated with 2 µg/mL streptavidin-PE for 1 h at room temperature in the dark and then washed three times with PBST. After centrifuging for 2 min at 1000×g, the slide was scanned using the GenePix 4000A microarray scanner (Molecular Devices, CA, USA). The fluorescent images were analyzed and the signal intensity was extracted using the GenePix Pro image analysis software (Molecular Devices, CA, USA).

Validation of biomarker candidates by ELISA

Human CCL22, PI3 and CD14 ELISA kits were obtained from R&D Systems (MN, USA). The ELISA experiments were executed according to the manufacture’s protocol. Briefly, 100 µL of assay diluent, human recombinant protein standards and diluted serum samples were added to appropriate wells sequentially. The resulting plate was incubated

http://www.thno.org
for 2 hrs at 4 °C, followed by four times of washing with washing buffer. After removing washing buffer by aspirating, 200 µL of HRP-conjugated detection antibody was added to each well and incubated for 2 hrs at 4 °C. After washing the wells to remove unbound detection antibody, 200 µL of TMB substrate solution was added and incubated for 30 min at room temperature. Lastly, the reaction was stopped with addition of 50 µL stop solution (2 M H2SO4) and the signal was read immediately at 450 nm.

**Bioinformatics analysis**

Functional annotation of serological proteins employed the PANTHER database (http://pantherdb.org/) [31]. The interaction network analysis of biological processes and signaling pathways used Cytoscape and ClueGO with a p-value cut-off ≤ 0.01 [32]. Pathway analysis employed the KEGG database (https://www.genome.jp/kegg/) [33].

**Statistical analysis**

Missing values in SWATH-MS data and non-signal values in microarray data were replaced with the minimum of each sample such that the signals would not be zero. Inter-sample data were normalized using quantile and log10 normalization and then subjected to Satterthwaite t-test analysis (p-value = 0.05) to identify the serological proteins related to psoriasis via the Python scipy.stats package (v1.1.0) [34].

After normalization, correlation analysis was performed with the Python scipy.stats package (v1.1.0). The Pearson correlation coefficient is used to represent correlations between continuous variables, while the Kendall rank correlation coefficient plays the same role between continuous variables and categorical variables (i.e., Drug, Sex, Smoking, Drinking and Family in the clinical data). The circos plot was made using circos (http://circos.ca/) [35]. The nonbiased hierarchical clustering analysis and biased hierarchical clustering analysis were performed using R package pheatmap (v1.0.10). The values of principal components in principal components analysis (PCA) were computed using sklearn.decomposition (v0.19.1) package in Python and result charts were drawn using matplotlib (v2.2.2) and seaborn (v 0.8.1) packages.

The Psoriasis Area and Severity Index (PASI) scores is frequently used in the clinic to assess the severity of psoriasis and the treatment response [36]. Therefore, the remission of psoriasis disease after therapy, or clinical improvement rate, was calculated using the equation as previously described [5, 27, 37]:

\[
\text{Improvement rate (％)} = \frac{\text{PASI}(0 \text{ week}) - \text{PASI}(12 \text{ week})}{\text{PASI}(0 \text{ week})} \times 100\%
\]

**Data deposition**

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the iProX partner repository [77] with the dataset identifier PXD013089.

**Results**

**Design and Fabrication of Psoriasis Antibody Microarrays**

Using psoriasis as a disease model, we collected serum samples from 16 healthy controls and 23 psoriasis patients before and after 12 weeks of treatment with a well-established oral traditional Chinese medicine (i.e., YinXieLing) for psoriasis treatment (Table 1) [27, 38-41]. The serum proteomes from these clinical samples were measured with our proteomics platform using DIA-MS and customizable antibody microarrays. All of the obtained data were analyzed with the following aims: 1) to elucidate the biology of psoriatic pathogenesis; 2) to find correlations between the clinical data, serum proteome and psoriasis severity index; and 3) to identity proteomic biomarkers for the diagnosis and treatment of psoriasis. The schematic illustration of this translational study using our in-depth serum proteomics approach is shown in Figure 1.

Table 1. Clinical serum samples for the identification of psoriasis-associated proteins.

<table>
<thead>
<tr>
<th>Age</th>
<th>Healthy controls (N=16) Mean±SD</th>
<th>Psoriatic patients (N=23) Mean±SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex N (%)</td>
<td>40.94 ± 12.61</td>
<td>47.09 ± 11.65</td>
</tr>
<tr>
<td>Male</td>
<td>14 (87.50)</td>
<td>20 (86.96)</td>
</tr>
<tr>
<td>Female</td>
<td>2 (12.50)</td>
<td>3 (13.04)</td>
</tr>
</tbody>
</table>

To select target proteins for antibody microarray detection, we first text-mined and manually curated the PubMed database as previous described (Figure S1A) [9]. Using psoriasis and plasma or serum as keywords, we identified 898 articles in which 120 protein candidates were obtained after three rounds of curation. In parallel, we identified 113 articles using psoriasis and biomarker as keywords, in which 33 candidates were obtained. After removing redundant biomarkers-of-interest, 129 plasma/serum proteins and biomarkers related to psoriasis remained (Table S1, Figure S1B). ClueGO analysis revealed that these
plasma/serum proteins are indeed involved in multi-signaling pathways associated with psoriasis, including cell proliferation, cell differentiation, cell migration, programmed cell death, chemotaxis and endocytosis (Figure S2). Of note, two pivotal psoriasis pathways, the MAPK and STAT pathways, were also identified in this study using our proteomics platform (Figure S2, Table S2) [42, 43].

Based on this 129-protein dataset, we designed and prepared an antibody microarray to detect all of these psoriasis plasma/serum proteins reported in the literature (Table S1). By creating a customized antibody microarrays focused on the disease of interest, the reagents, cost and time for the serum screening and data analysis are significantly reduced compared to using high-density antibody arrays [11]. Almost all (98.4%, 127/129) of the antibodies used to prepare the microarray were verified by the antibody vendors to detect the conformational or partially conformational epitopes of target proteins via ELISA, immunoprecipitation (IP), immunofluorescence (IF), immunohistochemistry (IHC) or flow cytometry. Only 1.6% (2/129) antibodies were exclusively validated by western blotting (WB), which recognizes linear epitopes (Figure S3). All 129 antibodies were printed on the surface of a 3D-modified slide with appropriate negative controls (PBS and BSA) and positive controls (biotinylated BSA and Alexa Fluor 555 goat anti-human IgG).

Using custom-made antibody microarrays, we screened 50 serum proteomes of healthy controls and psoriasis patients (Figure 2A, Table 1 and Table 2) [27, 38]. The r correlations of array-to-array and slide-to-slide were 0.99 respectively (Figure 2B). Using serum from a healthy control, the intra- and inter-array coefficient of variation (CV) were calculated as 2.09% and 5.72%, respectively. Furthermore, the CV of antibody microarray measurements in all 50 serum samples ranged from 2.13% to 8.2%, indicating the high reproducibility of antibody microarrays in serum screening (Figure S4). The majority (86.82%; 112/129) of the proteins were detected in these serum samples (Figure S5).

### Table 2. Clinical serum samples for the identification of biomarkers associated with psoriasis treatment.

<table>
<thead>
<tr>
<th></th>
<th>0 week (n=11)</th>
<th>12 weeks (n=11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>44.72 ± 11.1</td>
<td>44.72 ± 11.1</td>
</tr>
<tr>
<td><strong>Sex N (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>9 (81.82)</td>
<td>9 (81.82)</td>
</tr>
<tr>
<td>Female</td>
<td>2 (18.18)</td>
<td>2 (18.18)</td>
</tr>
<tr>
<td><strong>PASI</strong></td>
<td>7.73 ± 3.12</td>
<td>7.09 ± 4.89</td>
</tr>
</tbody>
</table>

Figure 1. Study design using in-depth serum proteomics. In this study, we developed an in-depth serum proteomics platform by combination of DIA-MS and customizable antibody microarray microarrays, which allows the measurement of serum proteome in unprecedented depth (~10 orders). Furthermore, the comprehensive determination of serum proteome for psoriasis disease enable the integrative analysis of correlations between serum protein expression and clinical data as well as the identification of biological information and potential biomarkers for psoriasis disease.
To assess the relationship between proteins reported in the literature with psoriasis, we compared the protein expression between psoriasis patients (n=23) and healthy controls (n=16) using the Satterthwaite t-test analysis. In addition to the well-known psoriasis-related proteins (e.g., VEGFA, IFNG, SELE, CXCL8, IL4, APOB, CCL22, EGFR, FAS, CD40LG), we identified 9 additional proteins (i.e., PI3, TNFRSF8, PFN1, KRT16, TNFSF8, KLK1, APOC3, CXCR3, CCL4) associated with psoriasis with statistical significance (p ≤ 0.05) (Figure 2C and 2D) (Table S3).

Surprisingly, two well-known serum cytokines that play an important role in the pathogenesis of psoriasis (i.e., TNF, IL17A) were not identified after statistical analysis. We therefore tested the binding specificity of the TNF and IL17A antibodies to their target proteins with the anti-TNF and anti-IL17A antibodies, respectively, by probing a protein array containing 7 different proteins (TNF, IL2, IL17A, IL1A, CCL2, CXCL10, CXCL9 and CCL4) (Figure S6) [44]. The antibodies bound to their targets in both cases, thus suggesting that the reason TNA and IL17A were not identified in this study may be due to the pathophysiologic heterogeneity of psoriasis disease in the population [45-48].

**Profiling of Psoriasis Serum Proteomes using DIA-MS and Antibody Microarray Technology**

The workflow for plasma/serum screening using SWATH-MS has been described before [5, 30]. In this study, we extended the SWATH methodology from TripleTOF (Sciex) to the DIA method using the QE-HF instrument (Thermo Fisher) (Figure 3A). To enable screening in high-throughput and reduce technical variation, all serum samples were added to a 96-well plate (2 μL/well), and then lysed, reduced, alkylated and digested in the same batch. DIA analysis led to the identification of 283 proteotypic peptides that uniquely belong to a single protein sequence to ensure high degree of quantitative accuracy.

Next, we analyzed the reproducibility of our DIA-MS assay, in which the Pearson correlation of 50 duplicated samples was 0.85 (0.69 - 0.95) (Figure 3B) and the average variation was 4.17% (1.75% - 12.6%) (Figure 3C). The Satterthwaite t-test further identified 58 proteins differentially expressed in the serum of psoriatic patients (p ≤ 0.05) when compared to the healthy controls (Figure 3D, Table S4). Some of these proteins are known to be involved in inflammation (PF4V1, PF4, PLCH1) and blood coagulation (F11, SERPINF2 and PLG) (Figure 3E).

Lastly, we analyzed the distribution of serological protein concentrations detected by our DIA-MS and microarray platform using the reference concentrations from the human plasma proteome database (http://www.plasmaproteomedatabase.org/) [49] (Figure 4A, Table S5). The proteins detected by DIA-MS and antibody microarrays range from 10⁶ to 10¹⁰ pg/mL and from 10⁰ to 10⁶ pg/mL.
respectively, spanning a total of ~10 orders of magnitude (Figure 4A).

Bioinformatics analysis of the serum proteins’ functions identified by each proteomics method (Figure 4B) showed that most of proteins identified by DIA-MS and antibody microarrays are linked to the same cellular components and biological processes (Figure S7 and S8). However, while proteins detected by DIA-MS are enriched with enzyme modulator, defense immunity and the blood coagulation pathway, the proteins detected by antibody microarrays are enriched with signaling molecules and the CCKR, P53, interleukin and hormone receptor pathways (Figure 4B and 4C). For both DIA-MS and antibody microarrays, differentially expressed proteins in the serum of psoriasis patients were enriched in critical pathways involved in the regulation of inflammation and immune system (i.e., cytokine-chemokine, coagulation, complement) (Figure S9 and S10). Taken together, these data suggest that our in-depth proteomics analysis using DIA-MS and antibody microarrays comprehensively reveals the functional landscape of the psoriasis serum proteome, including molecular functions, protein classes, cellular components, biological processes and signaling pathways (Figure S11).

Correlation Analysis between Clinical Data, Serum Proteome and Psoriasis Severity

To further demonstrate the clinical utility of our in-depth serum proteomics platform, we analyzed the correlations between clinical data, serum proteome and psoriasis severity. Patients’ visual analog scale (VAS) is a method sometimes used to measure itch intensity to assess psoriasis severity and treatment benefit [50, 51]. Our results indicate that the number of neutrophils (NEUT) is positively correlated with the expression levels of four serum proteins (SERPINE1, PI3, IL4, CX3CL1) and negatively correlated with the expression levels of six serum proteins (CCL4, EGFR, CD8A, IFNG, IGLC2, MTRF1) (Figure S12). Notably, our results show that NEUT and sex correlated with the Psoriasis Area and Severity Index (PASI) score, which is consistent with previous reports [52, 53]. In addition, we found that the PASI score is associated with three serum proteins (PI3, CCL22, IL-12B) (Figure 5C) that were uniquely different from the two serum proteins (TNFRSF8 and CD14) associated with the VAS scale (Figure S13). Among these proteins, PI3 and CCL22 were validated using ELISA (Figure S14). The positive and negative correlation data are represented by Figure 5A and Figure 5B, respectively (Table S6).

Figure 3. Detection of psoriasis serum proteins using DIA-MS. (A) High throughput screening of serum from psoriasis patients using DIA-MS; (B and C) Reproducibility of DIA-MS detection of serum proteins; (D) Identification of psoriasis-associated proteins in serum using volcano plot analysis; (E) Box plot analysis of newly identified serological psoriasis-associated proteins in inflammation and blood coagulation pathways. The selection of differentially expressed proteins was performed using the Satterthwaite t-test analysis (p-value = 0.05). N and P represent healthy controls and psoriatic patients, respectively. *,** and *** represent the p-value less than 0.05, 0.005 and 0.0005, respectively.
Identification of Psoriasis Biomarkers using In-depth Serum Proteomics

We performed unbiased hierarchical clustering on our collected serum proteome data and clinical status. The analyses showed that 87.3% (18/23) of psoriasis patients could be distinguished from healthy controls with a specificity of 100% (16/16) (Figure S15). After using the Satterthwaite t-test analysis with a p-value of 0.01, 58 proteins in the nonbiased cluster analysis could discriminate healthy controls from psoriasis patients (Figure 6A). These results were confirmed by principle component analysis, indicating that these proteins could be used as diagnostic biomarkers for psoriasis (Figure 6B).

Among these proteins, PFN1, a protein involved in actin polymerization, was cross-validated by DIA-MS and antibody microarrays (Figure 6C).

The potential biomarkers of psoriasis identified in our study, PI3, CCL22 and CD14, were further validated using quantitative ELISA using the same cohort of samples. We found that the expression of PI3, CCL22 and CD14 were higher in the psoriasis group (Figure 6D, Figure S14 and S16), which is consistent with our microarray results. We also validated PI3 using an independent cohort of 20 healthy controls, 10 urticaria patients, and 50 psoriasis patients (Table 3). The results show that PI3 expression is consistently higher in psoriasis patients compared to the other groups (Figure 6D).
Table 3. Clinical serum samples used for validating the biomarkers of psoriasis identified in this study.

<table>
<thead>
<tr>
<th></th>
<th>Healthy controls (N=20)</th>
<th>Urticaria Patients (N=10)</th>
<th>Psoriatic patients (N=50)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>38.80 ± 15.20</td>
<td>41.30 ± 8.84</td>
<td>45.34 ± 12.91</td>
</tr>
<tr>
<td><strong>Sex N (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>15 (75)</td>
<td>5 (50)</td>
<td>20 (78)</td>
</tr>
<tr>
<td>Female</td>
<td>5 (25)</td>
<td>5 (50)</td>
<td>11 (22)</td>
</tr>
<tr>
<td><strong>PASI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12.82 ± 3.09</td>
</tr>
</tbody>
</table>

There are different drugs on the market for the treatment of psoriasis, including therapeutic monoclonal antibodies targeting specific cytokines (e.g., Adalimumab, Infliximab). Although highly effective, the use of these treatments can be limited by side effects, primary or secondary loss of efficacy, and high treatment costs [54]. Therefore, finding biomarkers to predict treatment outcomes will reduce treatment costs, avoid side-effects and improve treatment effectiveness. In this study, we selected 11 psoriasis patients treated with an oral Chinese herbal medicine, YinXieLing, for 12 weeks [27, 38] (Table 2). Treatment results showed YinXieLing was effective in 6 patients and ineffective in 5 patients (Figure 7A).

To identify predictive biomarkers of treatment response, we compared the serum protein expression in responders (n=6) and nonresponders (n=5) prior to the drug treatment and identified 12 differentially-expressed proteins with a p-value ≤ 0.05 (Table S7). These proteins distinguished responders and nonresponders using unbiased clustering and principle component analysis (Figure 7B and Figure S17). The proteins were then validated by measuring their expression changes after treatment with YinXieLing for 12 weeks (Figure 7C). After 12 weeks of treatment, the expression levels of FCN2, MIF and MMP1 in the serum of responders were consistently higher than nonresponders, indicating their potential as predictive biomarkers of YinXieLing effectiveness in treating psoriasis.

Figure 5. Correlation network of serum proteome, clinical data and psoriasis index. (A) Positive and (B) negative correlations between the serum proteome (gray), clinical data (blue) and psoriasis indices (red) using circos, respectively. The statistical methods employed are explained in the Materials and Methods section; (C) Psoriasis serum proteins associated with the PASI score.
Figure 6. Identification of serum proteins for the detection of psoriasis. (A) Classification of healthy and psoriatic patient groups based on differentially expressed proteins and unbiased clustering analysis; (B) Classification of healthy and psoriatic patient groups based on differentially expressed proteins and principle component analysis; (C) Cross-validation of PFN1 protein expression using DIA-MS and antibody microarrays. The statistical analysis was performed using the Satterthwaite t-test with a p-value of 0.01. (D) Validation of PI3 as a psoriasis biomarker using ELISA. *, ** and *** represent the p-value less than 0.05, 0.005 and 0.0005, respectively.

Figure 7. Identification of serum proteins to predict the efficacy of psoriasis treatment with YinXieLing. (A) Different responses of psoriasis patients to Chinese medicine treatment; (B) Classification of responders and non-responders based on differentially expressed proteins and unbiased clustering analysis; (C) Changes in protein expression levels before treatment and 12 weeks after treatment. The statistical analysis was performed using the Satterthwaite t-test with a p-value ≤ 0.05.
Discussion

In-depth proteome profiling of serum and plasma has been a major challenge due to the wide range of protein concentrations and the detection capability of current technologies. To help address this, we utilized data from DIA-MS and customizable antibody microarrays to compare the serological proteome of healthy controls and psoriasis patients. The combination of DIA-MS and antibody microarrays enabled the unprecedented in-depth analysis of the psoriasis serum proteome spanning 10 orders of magnitude in protein concentration (Figure 4A), thus allowing the exploration of protein changes during psoriasis and their relationship to clinical data (Figure 5). We found that the high proteome coverage of DIA-MS and the high sensitivity of antibody microarrays complemented each other, such that more informative data were collected on high and low abundance proteins from DIA-MS data and antibody microarrays, respectively (Figure 4B) (Figure S8-S10). This is not surprising since it is well known that the presence of highly abundant proteins can mask the detection of lower abundant proteins during mass spectrometry analysis [55].

We found that four proteins (SERPINE1, PI3, IL4, CX3CL1) and six proteins (IFNG, CCL4, EGFR, CD8A, IGLC2, MTRF1) were positively and negatively correlated with neutrophils, respectively (Figure S12). In our previous studies, we found that the oral traditional Chinese medicine, YinXieLing, exerts its therapeutic effect on psoriasis by inhibiting cell proliferation and pro-inflammation while increasing CD4+ Foxp3+ regulatory T cell generation [39-41]. These results align well with previous studies that show that PI3 downregulates neutrophil-mediated inflammation and IFNG inhibits neutrophil accumulation [56-59].

Three serum proteins (PI3, CCL22 and IL12B) that correlated with the PASI score (i.e., disease severity) were identified in this study (Figure 6D and Figure S16). PI3 (Elafin) is an epithelial host-defense protein that is absent in normal skin but highly expressed in keratinocytes of inflamed skin, which is a hallmark of psoriasis [60]. The gene expression of PI3 has been found to be up-regulated in psoriatic skin samples (n=58) compared to controls (n=63) [61]. More recently, Elgharib et al. found serological PI3 was significantly elevated in 26 psoriasis patients compared to 26 healthy controls [62], which appears to also confirm our results as well. CCL22 is a chemoattractant for monocytes, dendritic cells, natural killer cells and activated T lymphocytes, and likely plays a role in trafficking these cell types to inflammatory sites. Increased CCL22 expression in psoriatic skin has been associated with the positive response to infliximab therapy, a chimeric monoclonal antibody that targets TNF alpha [23]. IL12B mutation is associated with the susceptibility to psoriasis and psoriatic lesions are particularly rich in IL12B proteins [63]. Our study, which employed serum samples, suggests that PI3, CCL22 and IL12B are secreted into the circulatory system. In addition, we found that two serum proteins (TNFRSF8, CD14) correlated with the VAS score (i.e., another score system to measure disease severity) (Figure S13). CD14 is a marker on monocytes and macrophages, whose expression increased in psoriasis patients, but did not correlate with the PASI score [64].

Unbiased clustering analysis and principle component analysis identified 58 proteins that distinguished healthy controls from psoriasis patients (Figure 6A and 6B). A protein that interacts with actin and regulates the structure of the cytoskeleton, PFN1, was identified as a disease biomarker of psoriasis by both DIA-MS and the antibody array (Figure 6C). Interestingly, elevated PFN1 protein levels was previously observed in the synovial fluid of psoriatic arthritis patients using liquid chromatography coupled to MS (LC-MS/MS) and multiplexed selected reaction monitoring (MRM) assays[65].

Finally, this study discovered 12 proteins that appear to be related to the effectiveness of YinXieLing treatment, distinguishing responders from non-responders with a p-value ≤ 0.05 (Figure 7B). Responders had sustained high expression levels of three proteins (FCN2, MMP1, MIF) after 12 weeks of treatment (Figure 7C). FCN2 encodes a ficolin-2 protein that activates the lectin complement signaling pathway in innate immunity [66]. MMP1 (Interstitial collagenase), which cleaves I, II, III, VII and X collagens, is elevated in the serum of patients with psoriatic arthritis, rheumatoid arthritis, osteoarthritis and systemic lupus erythematosus [67]. MIF (Macrophage migration inhibitory factor) is a pleiotropic pro-inflammatory factor involved in inflammation and immunological processes that is elevated in the serum and peripheral blood mononuclear cells of psoriasis patients [68, 69].

There are several limitations in this study. First, only proteins associated with psoriasis in abstracts within the PubMed database were selected as targets for antibody microarrays. A larger antibody array could be constructed based on full-length text mining. Second, the sensitivity of antibody microarrays is still limited to pg/mL to ng/mL, thus no low abundance proteins below the detection limit can be analyzed. Signal amplification technology such as near-infrared fluorescence-enhanced detection [70], proximity ligation assay [71] and single molecule array [72] may increase array sensitivity for improved protein
detection. Third, proteins may be quantified differently with DIA-MS and antibody microarrays due to different assay principles and detected protein isoforms. DIA-MS is an untargeted data acquisition approach with high proteomics coverage and accurate quantification of digested peptides [73]. Antibody microarrays, on the other hand, detects target proteins using pre-immobilized antibodies on the slide matrix that recognize specific epitopes [11]. The accurate quantification of multiple isoforms and splice variants, in which the amino acid sequences are highly homologous, is challenging for mass spectrometry and antibody arrays since their concentrations may be dramatically different in plasma and serum [74, 75]. As such, biomarkers identified by DIA-MS or antibody microarrays must be validated with a different technology, such as ELISA [76]. Notably, the number of clinical samples used in this study was limited, and therefore the biomarkers will need to be validated using a large independent clinical cohort before their use in clinical practice would be feasible.

The potential biomarkers revealed by this study distinguishing healthy controls from psoriasis patients and YinXieLing responders from non-responders, may help to better diagnose and treat psoriasis patients. These data also indicate that our in-depth serum proteomics approach using DIA-MS and antibody arrays has great potential in translational studies for other diseases.

Conclusion

We developed an in-depth serum proteomics technology platform that enable the measurement of serological proteome in unprecedented depth (~10 orders). Furthermore, we demonstrated that our methodology detected a large number of serum proteins associated with psoriasis disease and validated PI3 as a diagnostic biomarker for psoriasis in an independent clinical cohort. These results demonstrate the clinical utility of our novel proteomic assays and its usefulness to further advance our understanding, diagnosis, and treatment of psoriasis. Altogether, our findings suggest that the DIA-MS combined with antibody microarrays has great potential for application in translation studies and precision medicine in the future.

Supplementary Material

Supplementary table S1.
http://www.thno.org/v09p2475s2.xlsx
Supplementary table S2.
http://www.thno.org/v09p2475s3.xls

Acknowledgements

This project was financially supported by the National Key Basic Research Project (2018ZX09730003, 2017YFC0906703 and 2018YFA0507503), the National Natural Science Foundation of China (81673040 and 31870823), the Program on the Joint Proteomics Center for Chinese Medicine between PHOENIX Center and Guangdong Provincial Academy of Chinese Medical Sciences, the State Key Laboratory of Proteomics (SKLP-O201504, SKLP-O201703 and SKLP-K201505) and Capital’s Funds for Health Improvement and Research (2018-2-4034) to X.Y. We also thank Dr. Brianne Petritis for her critical review and editing of this manuscript.

Author Contributions

D. L. M. X., H. D. and Y. L. executed the text mining and manual curation. M.X., X. Z. and J. D. executed the array experiments; T. Z., Y. S., L. Y., Q. Z., X. C. and Y. Z., executed the sample preparation and DIA-MS experiments; J. D., L. H., Y. Y., D. Y., H. D. and C. L. collected the serum samples and clinical information; K. X., D. W., Y. Z., R. T., B. D. and C. C. executed the data analysis; X. Y., T. G., D. X. and C. L. designed and supervised all experiments and wrote the paper. All authors approved the final manuscript.

Competing Interests

The research group of T.G. is supported by Thermo Fisher, which provided access to mass spectrometry instrumentation.

References


73. Martins-de-Souza D, Faca VM, Gozzo FC. DIA is not a new mass spectrometry acquisition method. Proteomics. 2017; 17: 1700017.


