Identification of CD14 and lipopolysaccharide-binding protein as novel biomarkers for sarcoidosis using proteomics of serum extracellular vesicles

Yu Futami[1,©], Yoshito Takeda[1], Taro Koba[1], Ryohei Narumi[2], Yosui Nojima[3], Mari Ito[3],
Mana Nakayama[1], Mimiko Ishida[2], Hanako Yoshimura[1], Yujiro Naito[1], Kiyoharu Fukushima[1],
Takayuki Takimoto[4], Ryuya Edahiro[1], Takanori Matsuki[1], Satoshi Nojima[5], Haruhiko Hirata[1],
Shohei Koyama[1], Kota Iwahori[1], Izumi Nagatomo[1], Yuya Shirai[1], Yasuhiko Suga[1], Shingo Satoh[1],
Shinji Futami[1], Kotaro Miyake[1], Takayuki Shiroyama[1], Yoshikazu Inoue[6], Jun Adachi[2],
Takeshi Tomonaga[2], Koji Ueda[7] and Atsushi Kumanogoh[1,8,9]

[1]Department of Respiratory Medicine and Clinical Immunology, Osaka University, Graduate School of Medicine, 2-2 Yamada-Oka, Suita, Osaka 565-0871, Japan
[2]Laboratory of Proteome Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan
[3]Laboratory of Bioinformatics, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan
[4]Department of Respiratory Internal Medicine, National Hospital Organization Kinki-Chuo Chest Medical Centre, Nagasone-Cho 1180, Kita-Ku, Sakai, Osaka 591-8555, Japan
[5]Department of Pathology, Osaka University Graduate School of Medicine, 2-2 Yamada-Oka, Suita, Osaka 565-0871, Japan
[6]Clinical Research Centre, National Hospital Organization Kinki-Chuo Chest Medical Centre, 1180 Nagasone-Cho, Kita-Ku, Sakai City, Osaka 591-8555, Japan
[7]Cancer Proteomics Group, Cancer Precision Medicine Centre, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto, Tokyo 135-8550, Japan
[8]Laboratory of Immunopathology, World Premier International Immunology Frontier Research Centre, Osaka University, 2-2 Yamada-Oka, Suita, Osaka 565-0871, Japan
[9]Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives (OTRI), Osaka University, 2-2 Yamada-Oka, Suita, Osaka 565-0871, Japan

Correspondence to: Y. Takeda; E-mail: yoshito@imed3.med.osaka-u.ac.jp

Abstract

Sarcoidosis is a complex, polygenic, inflammatory granulomatous multi-organ disease of unknown cause. The granulomatous inflammation in sarcoidosis is driven by the interplay between T cells and macrophages. Extracellular vesicles (EVs) play important roles in intercellular communication. We subjected serum EVs, isolated by size exclusion chromatography, from seven patients with sarcoidosis and five control subjects to non-targeted proteomics analysis. Non-targeted, label-free proteomics analysis detected 2292 proteins in serum EVs; 42 proteins were up-regulated in patients with sarcoidosis relative to control subjects; and 324 proteins were down-regulated. The protein signature of EVs from patients with sarcoidosis reflected disease characteristics such as antigen presentation and immunological disease. Candidate biomarkers were further verified by targeted proteomics analysis (selected reaction monitoring) in 46 patients and 10 control subjects. Notably, CD14 and lipopolysaccharide-binding protein (LBP) were validated by targeted proteomics analysis. Up-regulation of these proteins was further confirmed by immunoblotting, and their expression was strongly increased in macrophages of lung granulomatous lesions. Consistent with these findings, CD14 levels were increased in lipopolysaccharide-stimulated macrophages during multinucleation, concomitant with increased levels of CD14 and LBP in EVs. The area under the curve values of CD14 and LBP were 0.81 and 0.84, respectively, and further increased to 0.98 in combination with angiotensin-converting enzyme and soluble interleukin-2 receptor. These findings suggest that CD14 and LBP in serum EVs, which are associated with granulomatous pathogenesis, can improve the diagnostic accuracy in patients with sarcoidosis.

Graphical Abstract
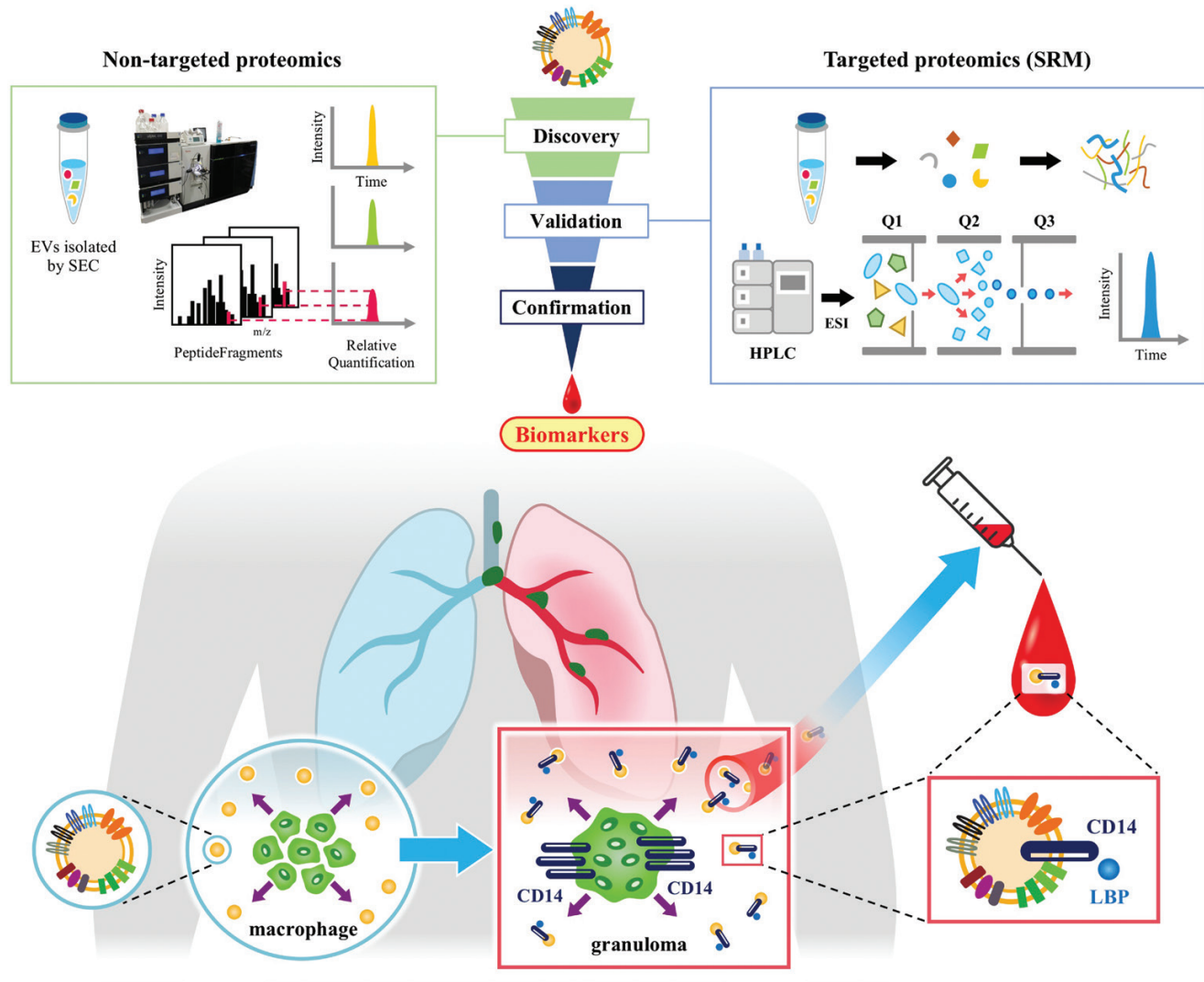
Keywords: bioinformatics, exosome, liquid biopsy, omics, precision medicine

Introduction

Sarcoidosis is a systemic granulomatous disease associated with T-lymphocyte and macrophage activation and migration into affected organs. There is heterogeneity in disease manifestation, severity and clinical course (1–3). To better understand this disease, effective biomarkers for diagnosis and prognosis are required; however, the ideal biomarker for sarcoidosis has not yet been discovered. Among the serum biomarkers identified to date, angiotensin-converting enzyme (ACE) and soluble interleukin-2 receptor (sIL-2R) are the most relevant but they lack sensitivity and specificity (4). Roughly 30–80% of patients with sarcoidosis have increased ACE levels, with a 22–86% sensitivity and 54–95% specificity (5). Similarly, although sIL-2R levels are proposed as a marker of sarcoidosis severity, elevated sIL-2R levels are not specific for sarcoidosis and can be found in other granulomatous diseases, haematological malignancies and various autoimmune

disorders (4). Thus, the use of ACE and sIL-2R as diagnostic markers for sarcoidosis remains a matter of debate.

In light of the progress in mass spectrometry (MS) technology, a great deal of attention has been given to omics approaches for the study of heterogeneous diseases (6). Although advances in proteomics have facilitated the discovery of protein biomarkers, the application of proteomics to measure features close to the phenotype has been limited when compared to genomics and transcriptomics (7).

Serum is an ideal source of molecules for biomarker discovery because of its high reproducibility and minimal invasiveness. However, it is difficult to detect small amounts of key proteins in blood because the dynamic range of blood proteins is extremely large (8).

To address this problem, in this study we focused on extracellular vesicles (EVs) in the serum. EVs are being

Table 1. Participants' shotgun proteomics characteristics (the discovery cohort)

	Age	Sex	Lung	Eye	Skin	Heart	Other
Control	58	F					
Control	58	F					
Control	59	F					
Control	47	M					
Control	45	M					
Sarcoidosis	47	F	I	+	–	–	Neuropahthy
Sarcoidosis	65	M	IV	+	+	+	Kidney disorder
Sarcoidosis	65	F	I	+	–	–	Muscle
Sarcoidosis	43	F	I	+	–	–	Breast
Sarcoidosis	29	M	0	+	+	–	Neurophathy
Sarcoidosis	60	F	I I	–	+	–	
Sarcoidosis	63	F	I I I	–	+	–	

Chest radiographic staging. Stage 0: normal chest radiograph, I: bilateral hilar lymphadenopathy (BHL), I: BHL plus pulmonary infiltration, I I pulmonary infiltration (without BHL), I I I: pulmonary fibrosis.

increasingly appreciated as important carriers of biologic cargo that also play key roles in intercellular communication by transferring contents such as messenger RNA, microRNA and proteins between neighbouring cells. Moreover, EV constituents play a variety of pathological roles in diseases including malignancies, inflammatory disorders and infections (7, 9, 10). Thus, an EV sandwich enzyme-linked immunosorbent assay (ELISA) system has been developed to diagnose lung cancer in clinics (11). While some methods, including ultracentrifugation, affinity-based methods and size exclusion chromatography (SEC), have been developed for the isolation of EVs, a standard method for the quantitative and qualitative analyses of EVs does not exist (12). Although ultracentrifugation is widely regarded as the gold standard to isolate EVs, SEC-based EV isolation yields much higher purity through MS analysis (9). Some comparative studies have shown that the isolation of EVs by SEC retains the biophysical properties of EVs, and results in a higher yield (13). Tandem mass tag-based non-targeted proteomic analysis of serum EVs isolated by ultracentrifugation revealed that fibulin-3 in EVs may serve as a novel diagnostic biomarker for chronic obstructive pulmonary disease (COPD) and might be closely related to its pathophysiology (14). With this background, we isolated serum EVs using SEC, which can detect less abundant serum proteins, and applied label-free proteomic strategies to discover novel biomarkers for sarcoidosis.

Methods

Study design

Participants were recruited for a discovery and a validation phase from the database of patients diagnosed with sarcoidosis at Osaka University Hospital (2013–19). All patients were diagnosed with sarcoidosis according to the international ATS/ERS/WASOG criteria (1). Serum samples were separated by centrifugation and stored at –80°C for further analyses. To sustain the sample quality, freezing and thawing of serum was avoided as much as possible. For the discovery cohort, five control subjects and seven patients with sarcoidosis were selected (Table 1). Patients with sarcoidosis were untreated when the samples were collected. The samples

Table 2. Participants' SRM proteomics characteristics (the validation cohort)

	Control subjects ($n = 10$)	Patients with sarcoidosis ($n = 46$)	*P*-value
Male/female	4/6	16/30	0.73
Average age	56.2 ± 17.6	59.0 ± 15.6	0.61
Chest X-ray stage			
0/I/II/III/IV		13/20/8/4/1	
Eye		24	
Skin		10	
Heart		3	
Other		10	
ACE (U l⁻¹)		25.7 ± 9.1	
sIL-2R (U ml⁻¹)		774.2 ± 830.4	
KL-6 (U ml⁻¹)		412.4 ± 539.9	

KL-6, Krebs von den Lungen-6.

were subjected to quantitative high-throughput proteomics using liquid chromatography-mass spectrometry (LC-MS/MS). The validation cohort included 10 control subjects and 46 patients with sarcoidosis (Table 2). For validation, candidate proteins identified in the discovery cohort were quantified by targeted proteomics using MS [selected reaction monitoring (SRM)] as described previously (15). This study was performed according to the guidelines described in the Declaration of Helsinki for medical research involving human subjects. This study protocol was approved by the Ethics Committee of Osaka University (no. 17148), and all study participants provided written informed consent.

EV isolation

Serum EVs were isolated by SEC, using an EV Second L70 SEC column (GL Science, Tokyo, Japan) (16). Briefly, after blocking with foetal bovine serum (FBS) and washing with phosphate-buffered saline (PBS), 400 µl of serum was loaded onto the column and eluted with PBS. In this study, we used equal volumes of serum to isolate EVs for each experiment. The first 300 µl of the eluate was discarded and thereafter, the eluate was collected in three fractions of 100 µl

each. The isolation of EVs was confirmed using the guide-lines delineated in minimal information for studies of extra-cellular vesicles 2018 (MISEV2018) (17). Size distributions and numbers were confirmed by NanoSight nanoparticle tracking analysis (Malvern Instruments, Malvern, UK). The processing of EV proteins for proteomic analysis was per-formed as described in Supplementary Information.

Non-targeted proteomics and targeted proteomics (SRM)

In the discovery phase, quantitative proteomics was per-formed using an Orbitrap Fusion Lumos mass spectrom-eter (Thermo Scientific) combined with UltiMate 3000 RSLC nano-flow high-performance liquid chromatography (Thermo Scientific). In the validation phase, protein abundance was measured by SRM on a TSQVantage triple quadrupole mass spectrometer (Thermo Fisher Scientific) as described pre-viously (18–20). SRM data were analysed using the Skyline software (MacCoss Lab Software, Seattle, WA, USA) (21). Peak area ratios of endogenous light (L) peptides and their heavy (H) isotope-labelled internal standards were used for accurate quantification. Peptide light/heavy ratios were log2 transformed and median-centred.

Bioinformatics analysis of the proteome with version information

To identify biologically relevant molecular networks and pathways in the proteome, up-regulated protein IDs were used as input data. The following tools were used for ana-lysis: Ingenuity Pathway Analysis (IPA) (ver 01.13, Qiagen N.V.) for enrichment analysis and upstream analysis, Clue GO/Clue Pedia plugin (ver 2.5.7) from Cytoscape (ver 3.8.0) (22) for enrichment analysis and KeyMolnet (ver 6.2, KM Data Inc., https://www.km-data.jp) for network analysis.

Statistical analysis

Pearson's chi-square test or Welch's *t*-test were used to com-pare healthy control subjects and patients with sarcoidosis. Correlation between two parameters was calculated using Spearman's rank correlation coefficients. Differences were considered statistically significant at $P < 0.05$. Receiver operating characteristic curves were constructed using the SRM results. Area under the curve (AUC) values were used to evaluate the diagnostic value of each marker. Multiple logistic regression analysis was applied to calculate the predictive probability of a multimarker for the diagnosis of sarcoidosis. Statistical analysis was conducted using JMP Pro v. 14.3.0 (SAS Institute, Cary, NC, USA).

Transmission electron microscopy

EV samples were adsorbed onto a formvar/carbon-coated nickel grid for 1 h. EVs were fixed with 2% paraformaldehyde and then incubated with the following primary antibodies: anti-CD9 (MM2/57; Thermo Fisher Scientific), anti-CD14 [EPR3653] (ab133335; Abcam, Cambridge, UK) and anti-lipopolysaccharide-binding protein (LBP) polyclonal (PA5-21642; Invitrogen, USA). Immunoreactive EVs were visualized using anti-mouse IgG(H+L) (EMGMHL10) and

anti-rabbit IgG(H+L) (EMGFAR10; BBI Solutions, UK) anti-bodies preabsorbed with 10-nm gold particles.

Cell culture and lipopolysaccharide stimulation

RAW 264.7, a murine macrophage/monocyte lineage cell line, obtained from ATCC (ATCC no.: TIB-71), was cultured in DMEM (Dulbecco's Modified Eagle Medium) containing 0.5% FBS, 100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin. For stimulation with lipopolysaccharide (LPS), these cells (1×10^6 cells per ml) were seeded into 6-well plates in DMEM with exosome-free FBS and stimulated with 10 ng ml^{-1} LPS for 6 h. EVs in the supernatant (650 µl per well) were collected using a SEC column and then concentrated by ultracentrifu-gation (100 000 × *g*, 4°C, 70 min), dissolved in RIPA buffer and evaluated by blotting. RAW 264.7 cells were fixed and stained with Diff-Quick stain (Sysmex, Kobe, Japan).

Results

Proteomics for discovery of novel biomarkers for sarcoidosis

To discover novel biomarkers for sarcoidosis, we performed quantitative high-throughput proteomics using LC-MS/MS, followed by SRM verification (Fig. 1A). EVs were isolated by SEC from serum samples of control subjects and patients with sarcoidosis (Table 1) (16). The isolation of EVs was confirmed using the MISEV2018 guidelines (17). EVs from both groups expressed the EV marker protein CD9 and were similar in shape and size, less than 100 nm (Fig. 1B). Furthermore, the EVs were positive for flotillin-1, CD63 and CD9 and negative for calnexin and haptoglobin (Fig. 1C). In the nanoparticle tracking analysis, serum EVs from both groups were indistin-guishable in size and number (Fig. 1D–F).

The non-targeted proteomic analysis of EVs identified 2292 proteins. Of these, 42 proteins were significantly up-regulated in patients with sarcoidosis, while 324 were down-regulated (Fig. 2A; Table 3; Supplementary Table 1). A principal com-ponent analysis of EV protein abundance partially separ-ated control subjects from patients with sarcoidosis (Fig. 2B). The identified proteins were present in the cytoplasm (51%), plasma membrane (29%) and extracellular space (6%; Fig. 2C). Notably, IPA of the protein signature of sarcoidosis EVs re-vealed factors involved in antigen presentation, the immune re-sponse and the inflammatory response (Fig. 2D). Both tumour necrosis factor-α and transforming growth factor β1 pathways ranked highly as upstream signalling factors, suggesting that the protein fingerprints of serum EVs from patients with sarcoidosis reflect not only the disease characteristics, but also its pathogenesis (Fig. 2E). We visualized the network of functions and pathways of the 42 up-regulated proteins using the Clue GO/Clue Pedia plugin from Cytoscape. The nomin-ally significant ($P < 0.05$) pathways and associated proteins are shown in Fig. 2(F). Protein–protein interaction analyses re-vealed that the 42 up-regulated proteins were clustered in six functional groups including transfer of LPS from its carrier LBP to CD14 (37.5%), and antigen processing cross-presentation (12.5%).

Subsequently, after taking previous reports into consider-ation, we selected 25 proteins for the validation cohort using SRM (Fig. 3A; Table 2; Supplementary Table 2). Targeted
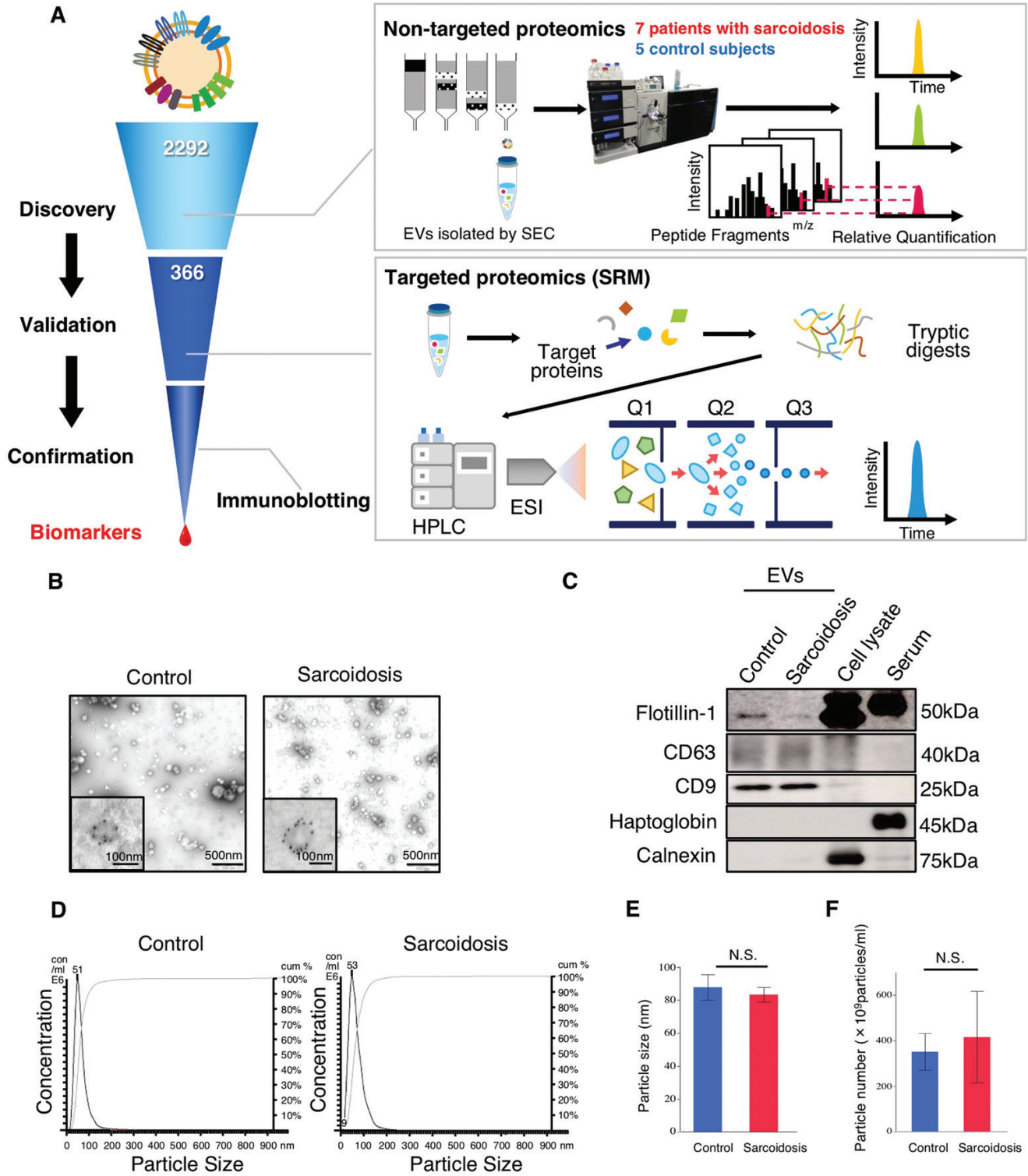
Fig. 1. Strategy for discovery of novel biomarkers for sarcoidosis. (A) In the discovery phase, serum EVs from patients with sarcoidosis and healthy control subjects were isolated using SEC and analysed. In the validation phase, biomarker candidates were quantified using SRM. Finally, the biomarkers were confirmed by immunoblotting. (B) Transmission electron microscopy images of serum EVs from a healthy control subject and a patient with sarcoidosis (CD9 immunogold labelling). (C) Immunoblots comparing flotillin-1, CD63, CD9, haptoglobin and calnexin levels in serum EVs versus serum and A549 cell lysates. (D) Representative distribution curves showing particle sizes of serum EVs from a healthy control subject and a patient with sarcoidosis, analysed using NanoSight. (E and F) The mean diameter and number of serum EVs from patients with sarcoidosis and healthy control subjects, analysed using NanoSight, are not significantly different. Error bars represent mean ± standard deviation. ESI, electrospray ionization; HPLC, high-performance liquid chromatography; N.S., not significant.
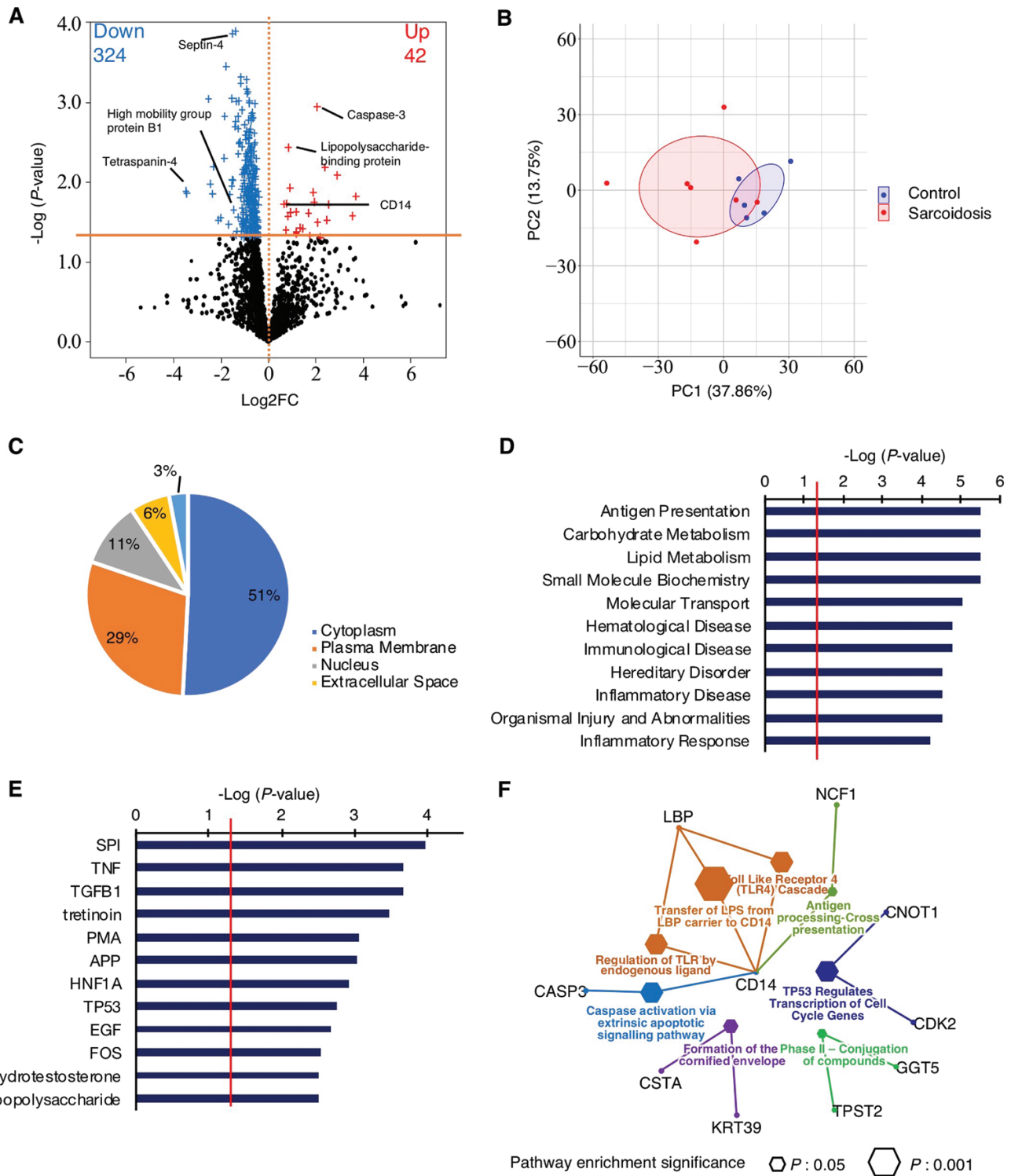
Fig. 2. The proteomic profile of serum EVs reflects sarcoidosis characteristics and pathogenesis. (A) A volcano plot of all 2292 serum EV proteins identified by non-targeted proteomic analyses in patients with sarcoidosis and healthy control subjects. A total of 42 proteins were significantly up-regulated, and 324 proteins were significantly down-regulated in sarcoidosis patients compared to controls. The horizontal line, *P*-value = 0.05. (B) Healthy control subjects and patients with sarcoidosis are separated by principal component analysis using all EV proteins identified by non-targeted proteomics analysis. (C) Localization of all identified proteins in the IPA. (D and E) Pathways (D) and upstream molecules (E) determined by the IPA to be over- (×1.5 or more) or under- (×0.8 or less) represented in the non-targeted proteomic analyses of serum EVs from patients with sarcoidosis compared to those from healthy control subjects. The vertical line, *P*-value = 0.05. (F) A protein–protein network constructed with enrichment analysis based on the Reactome database reveals interactions between the 42 up-regulated proteins.

Table 3. Significantly up-regulated proteins in EVs from patients with sarcoidosis compared to those from healthy subjects

Uniprot ID	Description	Fold change	*P*-value
P36222	Chitinase-3-like protein1	∞	0.032
O60704	Protein-tyrosine sulfotransferase 2	∞	0.033
Q9NZM1	Myoferlin	∞	0.004
Q7RTS3	Pancreas transcription factor 1 subunit alpha	∞	0.024
Q9Y230	RuvB-like 2	∞	0.034
Q8TAY7	Protein FAM11 D	∞	0.047
Q92522	Histone H1x	∞	0.038
P01040	Cystatin-A	∞	0.040
P06732	Creatine kinase M-type	∞	0.049
P98171	Rho GTPase-activating protein 4	∞	0.014
P36269	Glutathione hydrolase 5proenzyme	∞	0.007
A8MVU1	Putative neutrophil cytosol factor 1C	∞	0.040
P14598	Neutrophil cytosol factor 1	∞	0.040
A6NI72	Putative neutrophil cytosol factor 1B	∞	0.040
Q8NCG7	Sn1-specific diacylglycerol lipase beta	∞	0.004
Q8N5C1	ProteinFAM26E	∞	0.048
O15357	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 2	∞	0.019
Q75V66	Anoctamin-5	∞	0.045
Q14406	Chorionic somatomammotropin hormone-like 1	12.92	0.015
A5YKK6	CCR4-NOT transcription complex subunit1	11.77	0.027
Q9BYX4	Interferon-induced helicase C domain-containing protein 1	7.52	0.008
Q6A163	Keratin, type I cytoskeletal 39	5.78	0.019
P20963	T-cell surface glycoprotein CD3 zeta chain	5.53	0.030
Q15256	Receptor-type tyrosine-protein phosphatase R	5.22	0.007
O14815	Calpain-9	4.22	0.032
P42574	Caspase-3	4.15	0.001
Q10588	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 2	3.83	0.018
A0A075B6K4	Immunoglobulin lambda variable 3-1	3.74	0.013
Q92928	Putative Ras-related proteinRab-1C	3.38	0.048
Q3MII6	TBC1 domain family member 25	3.24	0.025
P24941	Cyclin-dependent kinase 2	2.72	0.039
Q8IU81	Interferon regulatory factor 2-binding protein1	2.55	0.038
Q96JQ0	Protocadherin-16	2.26	0.044
Q9Y5X9	Endothelial lipase	2.26	0.024
P58166	Inhibin beta E chain	2.26	0.042
P15291	Beta-1,4-galactosyltransferase 1	1.91	0.024
P26572	Alpha-1,3-mannosyl-glycoprotein 2-beta-*N*-acetylglucosaminyltransferase	1.89	0.012
P18428	Lipopolysaccharide-binding protein	1.79	0.004
P02766	Transthyretin	1.78	0.027
P08571	Monocyte differentiation antigenCD14	1.71	0.019
P04003	C4b-binding protein alpha chain	1.67	0.041
Q9UK55	Protein Z-dependent protease inhibitor	1.58	0.019

proteomics enables the efficient and specific verification of candidates without requiring antibodies and is, thus, a powerful tool for biomarker validation (14, 15, 20). Among the 25 biomarker candidates, the expression levels of LBP and the monocyte differentiation antigen CD14 were markedly elevated in patients with sarcoidosis (Fig. 3B). These findings corroborated our IPA upstream analysis, where LPS signalling was highly ranked (Fig. 1F), suggesting that these proteins are involved in sarcoidosis pathogenesis. Although serum CD14 levels have been previously shown to be elevated in patients with sarcoidosis (23), we did not observe this difference when we measured serum levels of both CD14 and LBP by ELISA (Fig. 3C), but only saw increased levels in the EVs. Analyses in KeyMolnet revealed that both CD14 and LBP were closely linked, both upstream and downstream, to key molecules related to granuloma formation (Fig. 3D). The presence and up-regulation of these proteins was subsequently confirmed by immunoelectron microscopy (Fig. 4A) and western blotting (Fig. 4B), respectively. Densitometric analyses of the immuno-blotting results revealed a significant increase in CD14 and

LBP levels in serum EVs (Fig. 4C and D). Furthermore, we confirmed the presence of CD14 and LBP in serum EVs isolated by both ultracentrifugation and the phosphatidylserine affinity method (data not shown).

The expression of CD14 and LBP in vivo and in vitro

We also assessed the expression level of these proteins *in vivo* by immunostaining of tissue samples. CD14 and LBP were weakly expressed in mononuclear cells in the lung and lymph nodes of healthy control subjects. In contrast, the expression levels of CD14 and LBP were strikingly increased in granulomatous lesions, especially in multinucleated giant cells (MGCs) and surrounding mononuclear cells (Fig. 5A and B). MGCs, the hallmarks of granuloma, are generated by monocytes in response to various stimuli, including LPS (24). To observe the dynamic changes of CD14 and LBP during MGC formation *in vitro*, we stimulated RAW 264.7 cells with LPS. Consistent with the *in vivo* results, CD14 levels were significantly elevated in EVs from cell culture supernatants as well as in LPS-stimulated macrophages
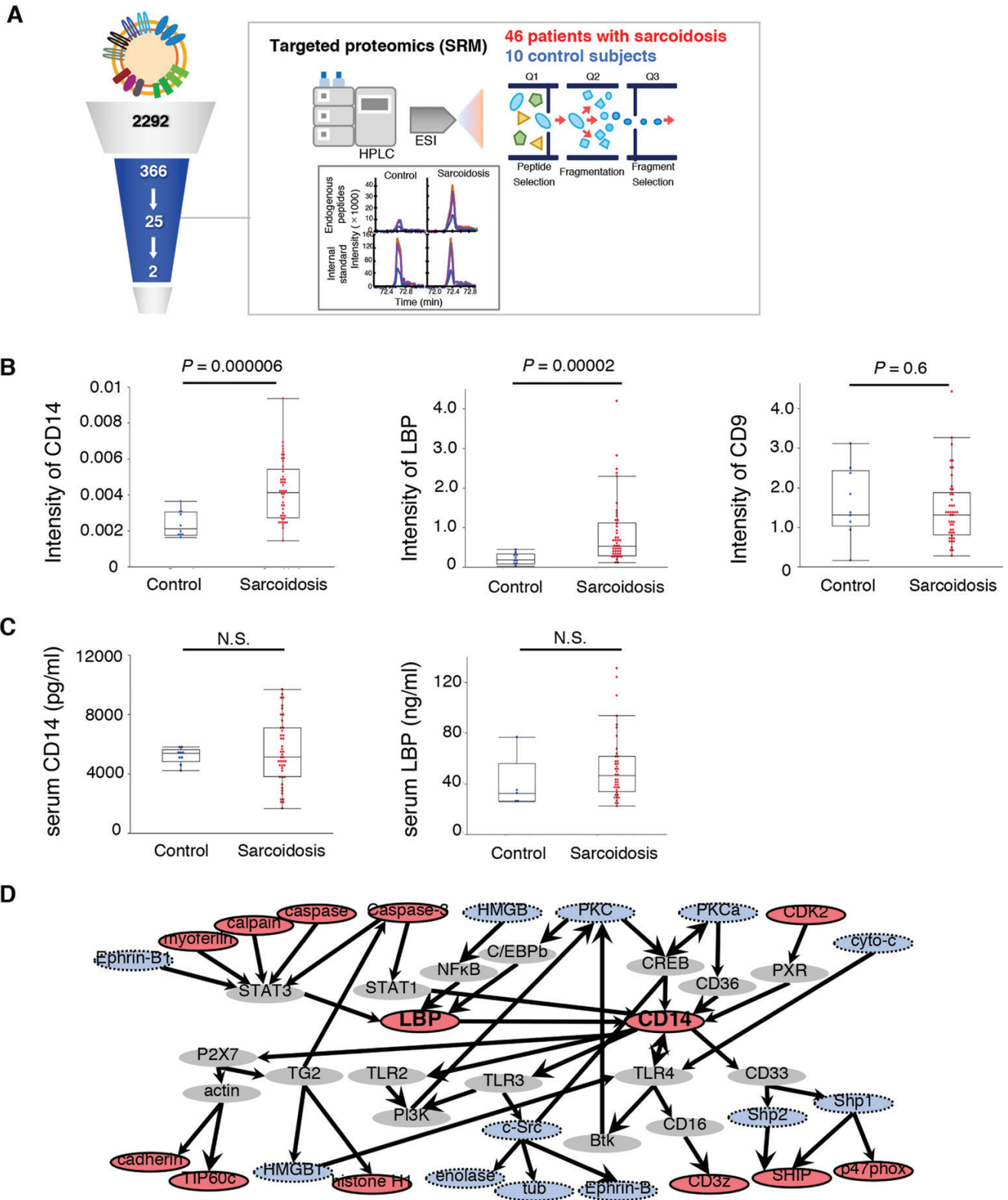
Fig. 3. Validation of the SRM results. (A) A flow chart of this study and representative SRM figure. The quantitation of each endogenous protein in serum EVs was performed by comparing the brown area of an endogenous peptide (Light) to the corresponding synthetic peptide (Heavy). For further information, please see the sub-section *Non-targeted proteomics and targeted proteomics (SRM)* in the Methods section. (B) CD14, LBP and CD9 levels in serum EVs from patients with sarcoidosis compared to healthy control subjects, as determined by SRM. (C) No significant difference in the serum levels of soluble CD14 and LBP can be detected between healthy control subjects and patients with sarcoidosis. (D) A highly complex network of targets with significant relationships, generated by KeyMolnet. Solid line ellipses: up-regulated proteins. Dotted line ellipses: down-regulated proteins. ESI, electrospray ionization; HPLC, high-performance liquid chromatography; N.S., not significant.
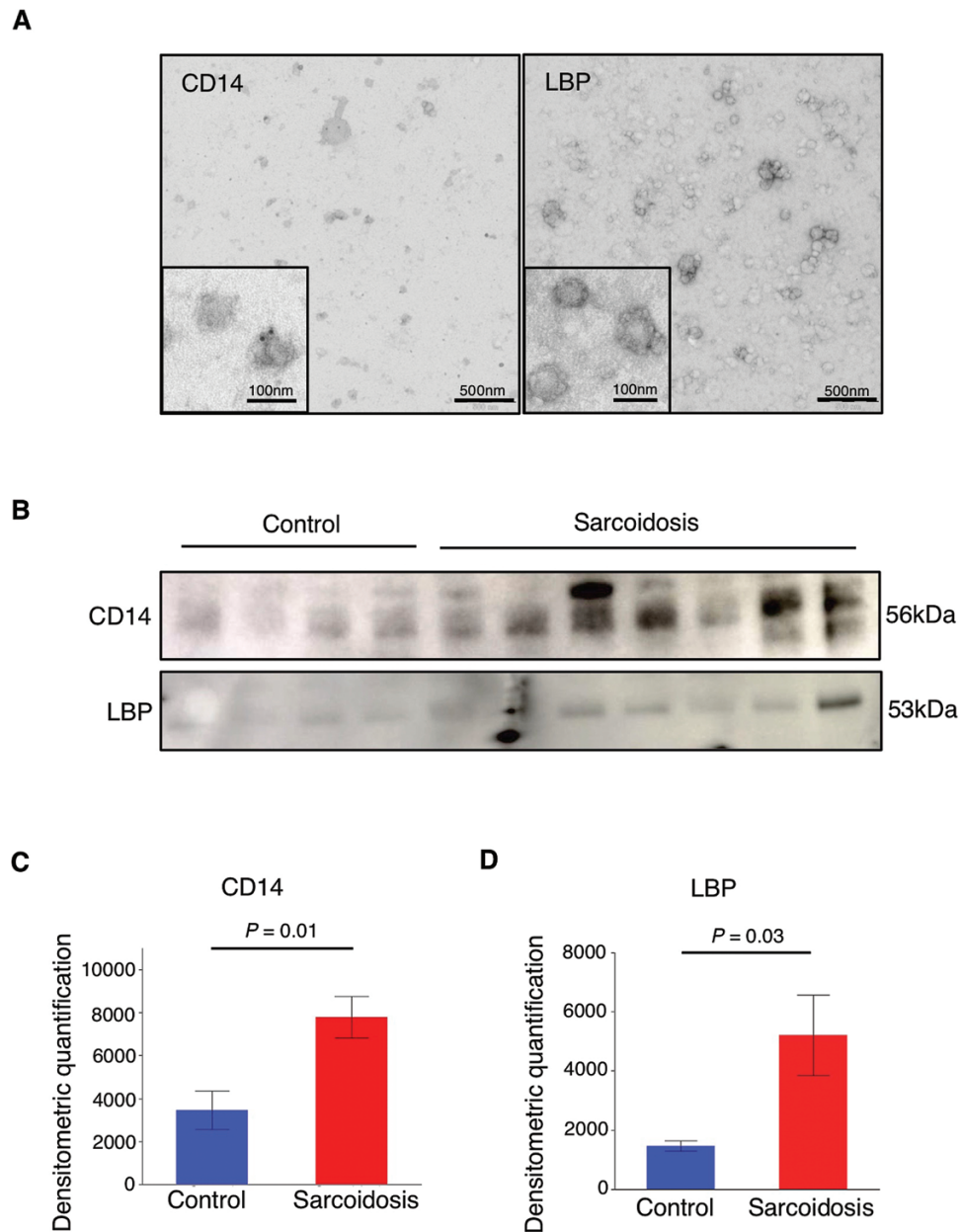
Fig. 4. Expression of CD14 and LBP in serum EVs. (A) Transmission electron microscopy images showing the CD14 and LBP expression in EVs isolated from patients with sarcoidosis, labelled with primary antibody and anti-IgG gold-labelled secondary antibody. (B) A representative immunoblot comparing CD14 and LBP in serum EVs of healthy control subjects and patients with sarcoidosis. Data are representative of three independent studies with similar results. (C and D) Densitometric analyses of the immunoblot results for CD14 (C) and LBP (D) in Fig. 4(B). Error bars indicate mean ± SEM.

(Fig. 5C). In contrast, while LBP levels in EVs were also elevated, they were unaltered in LPS-stimulated macrophages (Fig. 5C). Taken together, both CD14 and LBP were up-regulated in the process of granuloma formation *in vitro* and *in vivo*. Moreover, given that in macrophages, the tetraspanin CD9 is closely colocalized with CD14, thereby inhibiting LPS-induced signalling (25), and that CD9 is strongly involved in macrophage fusion into MGCs (26), CD14 in macrophages and EVs might positively or negatively participate in the pathogenesis of granulomatous diseases (Fig. 5D).

Diagnostic potential of CD14 and LBP in serum EVs for sarcoidosis

To further evaluate the diagnostic potential of the identified biomarkers, we analysed their AUC values. The AUC values for CD14 and LBP were 0.81 and 0.84, respectively (Fig. 6A). Considering that the previously reported biomarkers appear to have some limitations in terms of sensitivity and specificity, we thought that it would be intriguing to investigate different combinations. By combining the novel biomarkers with ACE, sIL-2R and both, the AUC values increased to 0.96, 0.96 and 0.98, respectively (Fig. 6A). LBP correlated well with clinical
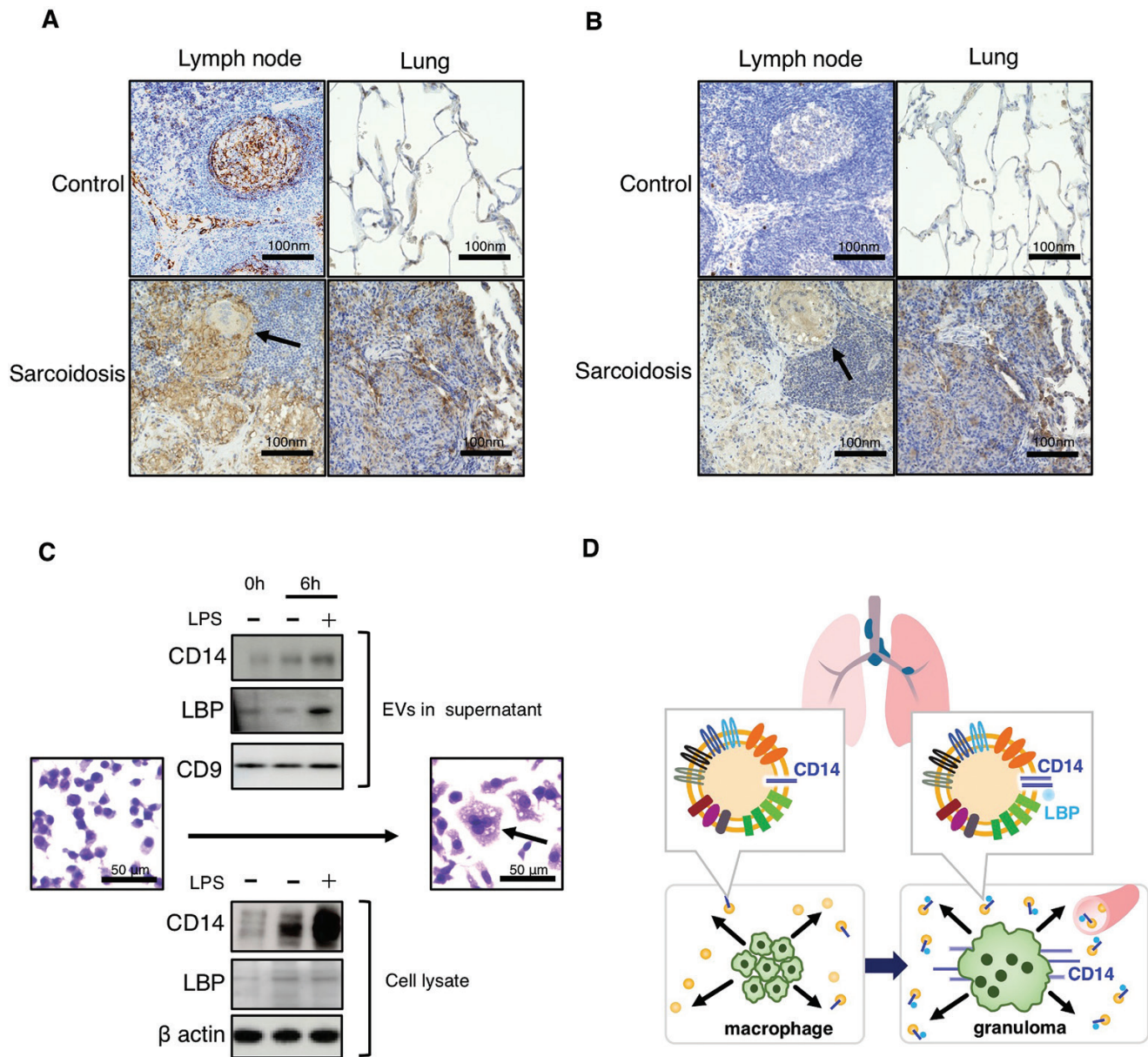
Fig. 5. Expression of CD14 and LBP *in vivo* and *in vitro*. (A and B) Immunohistochemical staining of the lung and lymph node in a patient with sarcoidosis. CD14 (A) and LBP (B) are highly expressed in the granuloma (black arrows). These markers are also expressed in macrophages present in the lung and lymph node of the healthy control subject. (C) Dynamic CD14 and LBP changes in EVs during macrophage multinucleation. For the induction of MGCs, RAW 264.7 cells (1×10^6 cells per ml) were seeded into 6-well plates in DMEM with exosome-free FBS and stimulated with 10 ng ml^{-1} LPS for 6 h. Data are representative of three independent studies with similar results. (D) The up-regulation of CD14 in serum EVs is associated with granuloma formation in patients with sarcoidosis.

parameters such as ACE (r: 0.45; $P < 0.01$), Krebs von den Lungen-6 (r: 0.35; $P < 0.05$) and C-reactive protein (CRP) (r: 0.60; $P < 0.01$), while CD14 correlated only with CRP (r: 0.29; $P < 0.05$) and not with any other parameters (Table 4). Despite 59.6% of the patients being ACE-negative, as many as 45.1% and 48.3% of these ACE-negative patients could be further diagnosed using CD14 and LBP, respectively. Similarly, while 46.8% of patients were sIL-2R-negative, 40.8% and 45.5% of these sIL-2R-negative patients could be further diagnosed using CD14 and LBP, respectively (Fig. 6C). These findings suggest that our novel biomarkers

possess distinct properties as compared to the conventional biomarkers.

Discussion

Since both ACE and sIL-2R have several limitations such as insufficient sensitivity and specificity, there is an urgent need to develop better biomarkers for sarcoidosis. Although peripheral blood may be regarded as an ideal source of biomarkers, MS-based serum proteomics is exceptionally challenging due to its broad dynamic range with abundant
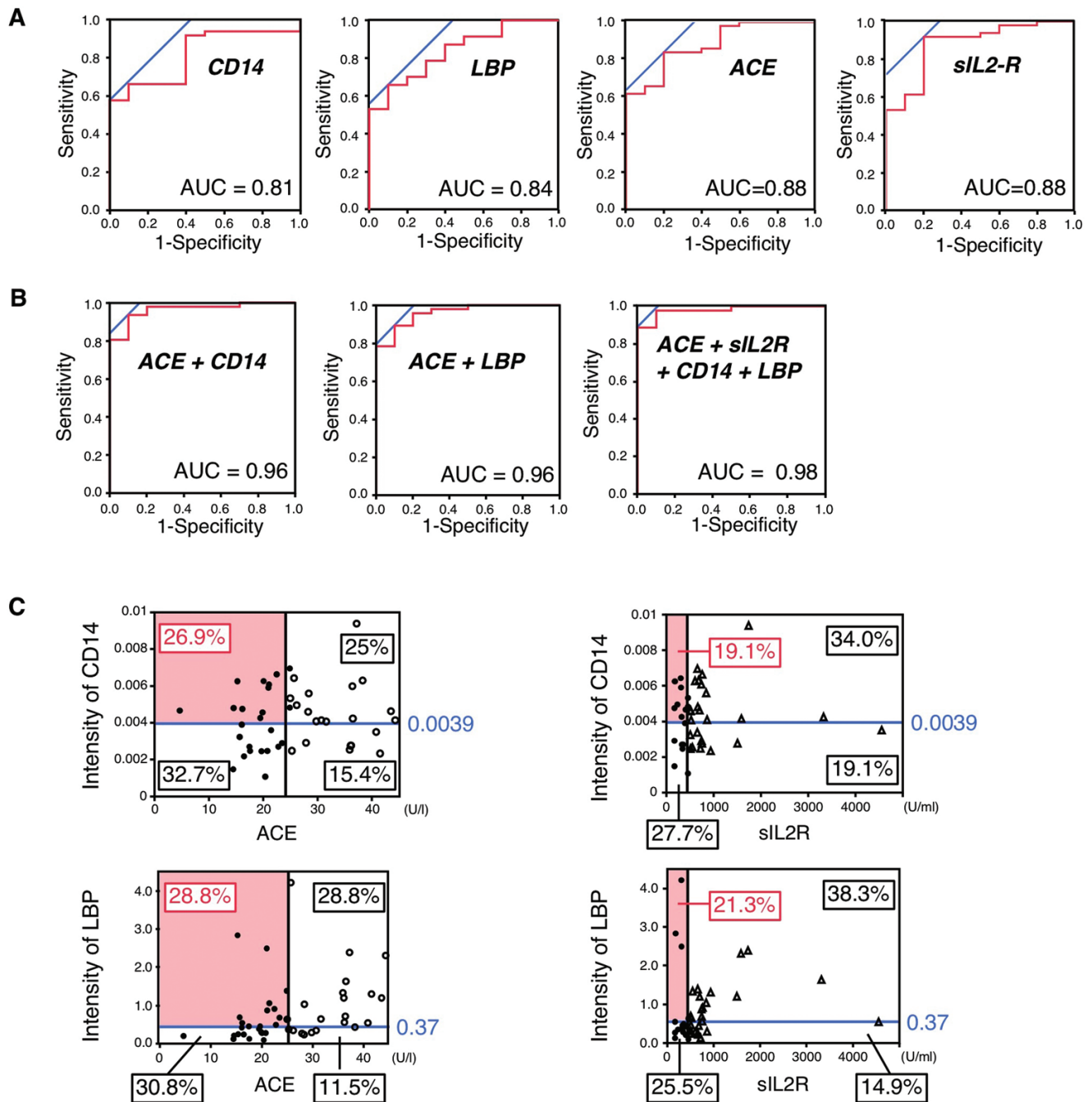
Fig. 6. Diagnostic accuracy of the novel biomarkers for sarcoidosis. (A) Receiver operating characteristics (ROC) curves of CD14 and LBP levels in serum EVs, and ACE and sIL-2R in serum to diagnose sarcoidosis. (B) ROC curves of CD14 and LBP combined with conventional biomarkers. (C) CD14 and LBP in serum EVs can partially identify sarcoidosis patients without elevation of serum ACE or sIL-2R levels. The cut-off values were determined by the Youden index.

proteins such as albumin (8). To overcome these problems, we focused on serum EVs and identified new biomarkers for sarcoidosis by employing both non-targeted, label-free proteomics and targeted proteomics. Although the importance of EVs, especially exosomes, is being increasingly demonstrated in cancer and immune diseases (11, 27–29), proteomics-based discovery of biomarkers for inflammatory diseases is still in its infancy. Importantly, serum EVs reflected the pathophysiology of sarcoidosis in our study, as shown by the IPA in Fig. 2(D). Using data from previous reports that compared gene expression changes in the lungs and peripheral blood mononuclear cells (PBMCs) of patients with sarcoidosis with those of control subjects (30–32), we examined the significance of each pathway

Table 4. Spearman's rank correlation coefficients between novel biomarkers and clinical parameters

	CD14 (r)	LBP (r)
ACE	0.12	0.45*
sIL-2R	0.046	0.29
KL-6	0.13	0.35**
SP-D	0.011	0.12
WBC	0.16	0.027
Neut	0.043	0.0012
Ly	0.15	−0.067
Mo	0.062	0.25
Eo	0.014	0.16
Baso	0.15	0.11
Creatinine	−0.1	−0.029
Calcium	−0.23	0.13
CRP	0.29**	0.6*
BNP	−0.16	−0.06
FVC	0.2	−0.3
FEV$_1$	0.077	−0.32
FEV$_1$%	−0.19	−0.053
%VC	0.012	−0.18

Baso, basophils; BNP, brain natriuretic peptide; Eo, eosinophils; FEV$_1$, forced expiratory volume in 1 s; FVC, forced vital capacity; KL-6, Krebs von den Lungen-6; Ly, lymphocyte; Mo, monocytes; Neut, neutrophils; SP-D, surfactant protein D; VC, vital capacity; WBC, white blood cells.
*$P < 0.01$, **$P < 0.05$.

in sarcoidosis on the basis of the overlap with our top-enriched pathways in Fig. 2(D). We found that pathways such as Immunological Disease, Organismal Injury and Abnormalities, Inflammatory Disease, Hereditary Disorder and Carbohydrate Metabolism were enriched by differentially expressed genes in the lungs of sarcoidosis patients. In addition, Immunological Disease, Inflammatory Disease, Organismal Injury and Abnormalities, Hereditary Disorder, Small molecule Biochemistry and Lipid Metabolism were enriched by differentially expressed genes in the PBMCs of sarcoidosis patients. Typical roles of these pathways include inflammatory cell infiltration and tissue repair, which are related to the pathophysiology of sarcoidosis.

Considering that isolating EVs can reduce the complexity of serum protein analysis and that EVs reflect the pathophysiology of the disease, serum EVs could become an ideal source of novel biomarkers, providing liquid biopsy samples for personalized medicine. While EVs have advantages such as stability and accessibility, they also have certain disadvantages such as difficulties in both isolation and quantitation (17). Although ultracentrifugation is widely regarded as the gold standard, SEC-based EV isolation yields much higher purity through MS analysis than ultracentrifugation (9). For this reason, we isolated EVs using SEC instead of ultracentrifugation. Previous studies leveraging MS-based proteomics to examine bronchoalveolar lavage fluid, alveolar macrophages, plasma and EVs in sarcoidosis identified several differentially expressed proteins, including pulmonary surfactant A2, vitamin D-binding protein and amyloid P (7). Regarding serum EVs, the vitamin D-binding protein levels in EVs derived from plasma samples or bronchoalveolar lavage fluid are elevated in patients with sarcoidosis (33). Although we detected a higher number of proteins, we could not confirm

this finding, which was presumably due to the differences in isolation and MS methods.

The verification of many biomarker candidates by conventional immunoblotting would have been time-consuming and less specific; the targeted proteomics approach allowed us to verify biomarker candidates specifically and efficiently without using any antibody. Combining non-targeted and targeted proteomics approaches, we identified CD14 and LBP as potential biomarkers of sarcoidosis, and further confirmed them by western blots and immunohistochemical tissue staining.

CD14 is a myeloid differentiation antigen mainly expressed on monocytes and macrophages. Both CD14 and LBP are required for recognition of LPS by Toll-like receptor (TLR) 4 (34). LPS is a major component of the gram-negative bacterial cell wall and is a typical example of a pathogen-associated molecular pattern. To facilitate microbial recognition and to amplify cellular responses, certain TLRs require additional proteins, such as LBP, CD14, CD36 and high mobility group box 1. These additional proteins are required since several infectious organisms like viruses, *Mycobacterium* spp., and *Propionibacterium acnes* have been implicated in the pathogenesis of sarcoidosis (1, 35, 36). Thus, the newly identified biomarkers CD14 and LBP may support an infectious aetiology of sarcoidosis. Previous studies have also shown that the Janus kinase/signal transducer and activator of transcription (JAK/STAT) signalling pathway was more activated in patients with sarcoidosis, and particularly STAT1 and STAT3 played a role in granuloma formation (37–39). In this study, some EV proteins detected in the discovery phase were correlated with CD14 and LBP through TLRs and the STAT pathway (Fig. 3D). Given that CD14 and LBP were up-regulated in the process of granuloma formation *in vitro* and *in vivo* (Fig. 5) and that CD14-positive monocytes promote MGC formation (40), both CD14 and LBP in serum EVs could be involved in the pathophysiology of sarcoidosis. Considering the limited sensitivity and specificity of the conventional biomarkers ACE and sIL-2R, as well as the positive correlations between various biomarkers, different combinations might be warranted. Our novel biomarkers possess both distinct properties and higher diagnostic potential in comparison to conventional biomarkers. Several studies have demonstrated that a combination of different biomarkers increases both the sensitivity and specificity (41). Hence, a combination of our novel biomarkers with either ACE or sIL-2R may strikingly improve the diagnostic potential in patients with sarcoidosis.

Despite the great advantages of SRM to verify many biomarkers efficiently, our approach has a couple of limitations. First, we examined our new biomarkers regarding organ specificity and disease severity but observed no significant correlation (Supplementary Figure 1). This could be caused by the relatively small sample size, and thus further studies are warranted. Moreover, the expression levels of the novel biomarkers were not increased in inflammatory lung disease, including COPD, bronchial asthma and lung fibrosis (data not shown), demonstrating their disease specificity. Next, given that genomic, transcriptomic and metabolomic networks are also important for precision medicine, it would be intriguing to integrate our data with multiple omics approaches (6, 42). A system biology platform using clinical data, omics and

bioinformatics may lead to a better understanding of sarcoidosis (6).

References

1   Hunninghake, G. W., Costabel, U., Ando, M. *et al.* 1999. Statement on sarcoidosis. *Am. J. Respir. Crit. Care Med.* 160:736.
2   Spagnolo, P., Rossi, G., Trisolini, R. *et al.* 2018. Pulmonary sarcoidosis. *Lancet Respir. Med.* 6:389.
3   Crouser, E. D., Maier, L. A., Wilson, K. C. *et al.* 2020. Diagnosis and detection of sarcoidosis. An Official American Thoracic Society Clinical Practice Guideline. *Am. J. Respir. Crit. Care Med.* 201:e26.
4   Chopra, A., Kalkanis, A. and Judson, M. A. 2016. Biomarkers in sarcoidosis. *Expert Rev. Clin. Immunol.* 12:1191.
5   Ramos-Casals, M., Retamozo, S., Sisó-Almirall, A. *et al.* 2019. Clinically-useful serum biomarkers for diagnosis and prognosis of sarcoidosis. *Expert Rev. Clin. Immunol.* 15:391.
6   Crouser, E. D., Fingerlin, T. E., Yang, I. V. *et al.* 2017. Application of 'omics' and systems biology to sarcoidosis research. *Ann. Am. Thorac. Soc.* 14:S445.
7   Guerrero, C. R., Maier, L. A., Griffin, T. J. *et al.* 2020. Application of proteomics in sarcoidosis. *Am. J. Respir. Cell Mol. Biol.* 63:727.
8   Tirumalai, R. S., Chan, K. C., Prieto, D. R. A. *et al.* 2003. Characterization of the low molecular weight human serum proteome. *Mol. Cell. Proteomics* 2:1096.
9   Wu, A. Y. T., Ueda, K. and Lai, C. P. K. 2019. Proteomic analysis of extracellular vesicles for cancer diagnostics. *Proteomics* 19:e1800162.
10 El Andaloussi, S., Mäger, I., Breakefield, X. O. *et al.* 2013. Extracellular vesicles: biology and emerging therapeutic opportunities. *Nat. Rev. Drug Discov.* 12:347.
11 Ueda, K., Ishikawa, N., Tatsuguchi, A. *et al.* 2014. Antibody-coupled monolithic silica microtips for highthroughput molecular profiling of circulating exosomes. *Sci. Rep.* 4:6232.
12 Baranyai, T., Herczeg, K., Onódi, Z. *et al.* 2015. Isolation of exosomes from blood plasma: qualitative and quantitative comparison of ultracentrifugation and size exclusion chromatography methods. *PLoS One* 10:e0145686.
13 Hong, C. S., Funk, S., Muller, L. *et al.* 2016. Isolation of biologically active and morphologically intact exosomes from plasma of patients with cancer. *J. Extracell. Vesicles* 5:29289.

14 Koba, T., Takeda, Y., Narumi, R. *et al.* 2021. Proteomics of serum extracellular vesicles identifies a novel COPD biomarker, fibulin-3 from elastic fibres. *ERJ Open Res.* 7:00658-2020.
15 Kume, H., Muraoka, S., Kuga, T. *et al.* 2014. Discovery of colorectal cancer biomarker candidates by membrane proteomic analysis and subsequent verification using selected reaction monitoring (SRM) and tissue microarray (TMA) analysis. *Mol. Cell. Proteomics* 13:1471.
16 Sakha, S., Muramatsu, T., Ueda, K. *et al.* 2016. Exosomal microRNA miR-1246 induces cell motility and invasion through the regulation of DENND2D in oral squamous cell carcinoma. *Sci. Rep.* 6:38750.
17 Théry, C., Witwer, K. W., Aikawa, E. *et al.* 2018. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J. Extracell. Vesicles* 7:1535750.
18 Narumi, R., Murakami, T., Kuga, T. *et al.* 2012. A strategy for large-scale phosphoproteomics and SRM-based validation of human breast cancer tissue samples. *J. Proteome Res.* 11:5311.
19 Narumi, R., Shimizu, Y., Ukai-Tadenuma, M. *et al.* 2016. Mass spectrometry-based absolute quantification reveals rhythmic variation of mouse circadian clock proteins. *Proc. Natl Acad. Sci. USA* 113:E3461.
20 Shiromizu, T., Kume, H., Ishida, M. *et al.* 2017. Quantitation of putative colorectal cancer biomarker candidates in serum extracellular vesicles by targeted proteomics. *Sci. Rep.* 7:12782.
21 MacLean, B., Tomazela, D. M., Shulman, N. *et al.* 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26:966.
22 Mlecnik, B., Galon, J. and Bindea, G. 2018. Comprehensive functional analysis of large lists of genes and proteins. *J. Proteomics* 171:2.
23 Pforte, A., Schiessler, A., Gais, P. *et al.* 1994. Expression of CD14 correlates with lung function impairment in pulmonary sarcoidosis. *Chest* 105:349.
24 Okamoto, H., Mizuno, K. and Horio, T. 2003. Monocyte-derived multinucleated giant cells and sarcoidosis. *J. Dermatol. Sci.* 31:119.
25 Suzuki, M., Tachibana, I., Takeda, Y. *et al.* 2009. Tetraspanin CD9 negatively regulates lipopolysaccharide-induced macrophage activation and lung inflammation. *J. Immunol.* 182:6485.
26 Takeda, Y., Tachibana, I., Miyado, K. *et al.* 2003. Tetraspanins CD9 and CD81 function to prevent the fusion of mononuclear phagocytes. *J. Cell Biol.* 161:945.
27 Genschmer, K. R., Russell, D. W., Lal, C. *et al.* 2019. Activated PMN exosomes: pathogenic entities causing matrix destruction and disease in the lung. *Cell* 176:113.
28 Melo, S. A., Luecke, L. B., Kahlert, C. *et al.* 2015. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature* 523:177.
29 Hoshino, A., Costa-Silva, B., Shen, T. L. *et al.* 2015. Tumour exosome integrins determine organotropic metastasis. *Nature* 527:329.
30 Crouser, E. D., Culver, D. A., Knox, K. S. *et al.* 2009. Gene expression profiling identifies MMP-12 and ADAMDEC1 as potential pathogenic mediators of pulmonary sarcoidosis. *Am. J. Respir. Crit. Care Med.* 179:929.
31 Koth, L. L., Solberg, O. D., Peng, J. C. *et al.* 2011. Sarcoidosis blood transcriptome reflects lung inflammation and overlaps with tuberculosis. *Am. J. Respir. Crit. Care Med.* 184:1153.
32 Su, R., Li, M. M., Bhakta, N. R. *et al.* 2014. Longitudinal analysis of sarcoidosis blood transcriptomic signatures and disease outcomes. *Eur. Respir. J.* 44:985.
33 Martinez-Bravo, M. J., Wahlund, C. J. E., Qazi, K. R. *et al.* 2017. Pulmonary sarcoidosis is associated with exosomal vitamin D-binding protein and inflammatory molecules. *J. Allergy Clin. Immunol.* 139:1186.
34 Anas, A., Poll, T. and de Vos, A. F. 2010. Role of CD14 in lung inflammation and infection. *Crit. Care* 14:209.
35 Brownell, I., Ramírez-Valle, F., Sanchez, M. *et al.* 2011. Evidence for mycobacteria in sarcoidosis. *Am. J. Respir. Cell Mol. Biol.* 45:899.

36 Patterson, K. C. and Chen, E. S. 2018. The pathogenesis of pulmonary sarcoidosis and implications for treatment. *Chest* 153:1432.

37 Zhou, T., Casanova, N., Pouladi, N. *et al.* 2017. Identification of Jak-STAT signaling involvement in sarcoidosis severity via a novel microRNA-regulated peripheral blood mononuclear cell gene signature. *Sci. Rep.* 7:4237.

38 Rosenbaum, J. T., Pasadhika, S., Crouser, E. D. *et al.* 2009. Hypothesis: sarcoidosis is a STAT1-mediated disease. *Clin. Immunol.* 132:174.

39 Damsky, W., Thakral, D., McGeary, M. K. *et al.* 2020. Janus kinase inhibition induces disease remission in cutaneous sarcoidosis and granuloma annulare. *J. Am. Acad. Dermatol.* 82:612.

40 Yuasa, K., Avnet, S., Salerno, M. *et al.* 2010. Spontaneous formation of multinucleated giant cells with bone resorbing activity by long culture of human peripheral blood CD14-positive monocytes in vitro. *Cell Commun. Adhes.* 17:13.

41 Kraaijvanger, R., Bonás, M. J., Vorselaars, A. D. M. *et al.* 2020. Biomarkers in the diagnosis and prognosis of sarcoidosis: current use and future prospects. *Front. Immunol.* 11:1443.

42 Landi, C., Bargagli, E., Carleo, A. *et al.* 2015. A functional proteomics approach to the comprehension of sarcoidosis. *J. Proteomics* 128:375.