COVID-19 and RA share SPP1 myeloid pathway that drives PD-L1pos neutrophils and CD14pos monocytes

Lucy MacDonald, … , Mariola Kurowska-Stolarska, Elisa Gremese


Graphical abstract

Find the latest version:
https://jci.me/147413/pdf
Title: COVID-19 and RA share SPP1 myeloid pathway that drives PD-L1^{pos} neutrophils and CD14^{pos} monocytes.

Authors: 1†Lucy MacDonald & 1,2,3†Stefano Alivernini, 1,3Barbara Tolusso, 1Aziza Elmesmari, 1Domenico Somma, 2Simone Perniola, 2Annamaria Paglionicia, 2Luca Petricca, 2Silvia Laura Bosello, 4Angelo Carfi, 5,6Michela Sali, 7Egidio Stigliano, 8,9Antonella Cingolani, 8,9Rita Murri, 7Vincenzo Arena, 8,9Massimo Fantoni, 10,11Massimo Antonelli, 4Francesco Landi, 12,13Francesco Franceschi, 5,6Maurizio Sanguinetti, 1Iain B. McInnes, 14Charles McSharry, 15,16Antonio Gasbarrini, 1,^*Thomas D. Otto, 1,^*Mariola Kurowska-Stolarska & 1,2,3,^*Elisa Gremese

Affiliation:
1Research into Inflammatory Arthritis Centre Versus Arthritis (RACE), University of Glasgow, United Kingdom.
2Division of Rheumatology, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy.
3Division of Rheumatology, Università Cattolica del Sacro Cuore, Rome, Italy.
4Institute of Internal Medicine and Geriatrics, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy.
5Dipartimento di Scienze di Laboratorio e Infettivologiche, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy.
6Dipartimento di Scienze Biotecnologiche di Base, Cliniche Intensivologiche e Perioperatorie – Sezione di Microbiologia, Università Cattolica del Sacro Cuore, Rome, Italy.
7Department of Woman and Child Health and Public Health, Area of Pathology, and U.O.S.D. Coordinamento attività di Settorato Fondazione Policlinico A. Gemelli IRCCS, Università Cattolica del Sacro Cuore, Rome, Italy
8Dipartimento di Scienze di Laboratorio e Infettivologiche, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy.
9Dipartimento di Sicurezza e Bioetica, Sez. Malattie Infettive, Università Cattolica del Sacro Cuore, Rome, Italy.
10Emergency Medicine, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy.
11Emergency Medicine, Università Cattolica del Sacro Cuore, Rome, Italy.
Short Summary: Severe COVID-19 and active Rheumatoid Arthritis share tissue macrophage clusters with a hallmark pro-inflammatory mediator osteopontin (SPP1) that activates PD-L1\textsuperscript{pos} neutrophils and classical monocytes, and persist into post-COVID-19 phase.

Declaration of Interests. All authors declare no conflict of interests.
Abstract

We explored the potential link between chronic inflammatory arthritis and COVID-19 pathogenic and resolving macrophage pathways and their role in COVID-19 pathogenesis. We found that BALF macrophage clusters $FCN1^{pos}$ and $FCN1^{pos}SPP1^{pos}$ predominant in severe COVID-19 were transcriptionally related to synovial tissue macrophage (STM) clusters $CD48^{high}S100A12^{pos}$ and $CD48^{pos}SPP1^{pos}$ that drive Rheumatoid Arthritis (RA) synovitis. BALF macrophage cluster $FABP4^{pos}$ predominant in healthy lung was transcriptionally related to STM cluster $TREM2^{pos}$ that governs resolution of synovitis in RA remission. Plasma concentrations of SPP1 and S100A12 (key products of macrophage clusters shared with active RA) were high in severe COVID-19 and predicted the need for Intensive Care Unit transfer, and remained high in post-COVID-19 stage. High plasma levels of SPP1 were unique to severe COVID-19 when compared to other causes of severe pneumonia, and immunohistochemistry localized SPP1$^{pos}$ macrophages in the alveoli of COVID-19 lung. Investigation into SPP1 mechanisms of action revealed that it drives pro-inflammatory activation of CD14$^{pos}$ monocytes and development of PD-L1$^{pos}$ neutrophils, both hallmarks of severe COVID-19. In summary, COVID-19 pneumonitis appears driven by similar pathogenic myeloid cell pathways as those in RA, and their mediators such as SPP1 might be an upstream activator of the aberrant innate response in severe COVID-19 and predictive of disease trajectory including post-COVID-19 pathology.
Introduction

The severity of acute COVID-19 is attributable to genetics, immune dysregulation, abnormal blood clotting and tissue disruption, particularly implicating aberrant pro-inflammatory and anti-viral innate immunity (1-14). Rheumatoid Arthritis (RA) represents a pro-inflammatory cytokine-driven chronic articular condition often accompanied by interstitial lung disease and alveolitis (15). We (16) and others (17) have reported that SARS-CoV-2 infection is associated with emergence of polyarthritis, or flares of synovitis in RA patients in sustained disease remission, suggesting shared mechanisms of pathogenesis. RA immunopathogenesis and therapeutic targets (18) are well-understood and might be informative for COVID-19 therapeutic strategies. In RA, articular inflammation and remission are driven by distinct synovial tissue macrophage clusters (19). Emerging data suggest that similar aberrant activation of myeloid cells in the blood (7, 20-23) and lung (24, 25) contribute to acute COVID-19 severity. However, there remains a knowledge gap on myeloid cell pathways that determine the severity and resolution of acute COVID-19 pneumonitis. In addition, the immune mechanisms of long-lasting clinical sequelae (26) described in convalescent COVID-19 patients are unresolved, hindering development of effective treatments and biomarkers of disease trajectory.

Single-cell profiling and fate-mapping indicate spatial and functional macrophage heterogeneity that maintains lung homeostasis (27-31). Healthy lung alveolar macrophages (AM) expressing Fatty Acid Binding Protein 4 (FABP4) localize to the alveolar epithelial surface and recycle surfactants with type-2 alveolar epithelial cells to maintain compliance and efficient gas exchange (32, 33). This function is compromised in severe COVID-19 (34). The scRNAseq analysis (24) of bronchoalveolar lavage fluid (BALF) from severe COVID-19 patients found abnormally low numbers of resident alveolar macrophages and an increase in two macrophage clusters which share expression of ficolin-1 (FCN1) and are distinguished by
their relative expression of osteopontin (SPP1): FCN1\textsuperscript{pos} and FCN1\textsuperscript{pos}\textsuperscript{SPP1\textsuperscript{pos}} clusters. Their precise roles in the pathogenesis of severe COVID-19 have yet to be established. We recently identified similar macrophage diversity in synovial tissues (ST) of healthy donors and patients with active or remission RA (19). We delineated in RA chronic synovitis that the CD48\textsuperscript{high}S100A12\textsuperscript{pos} and CD48\textsuperscript{pos}\textsuperscript{SPP1\textsuperscript{pos}} macrophage clusters produced their distinctive pathogenic mediators (S100A12 and SPP1, respectively) in addition to hallmark pro-inflammatory cytokines IL-6, IL-8, IL-1\textbeta and TNF\alpha, and activated synovial tissue stromal cells. Early evidence suggested that this serum IL-6, IL-8, IL-1\textbeta and TNF\alpha cytokine signature also predicts the prognosis of patients with acute COVID-19 (6), and targeting the IL-6 pathway seems an effective therapeutic strategy in critically-ill COVID-19 patients (35, 36). Resolution of synovitis in RA was associated with the functions of a distinct subpopulation of synovial macrophages (TREM2\textsuperscript{pos} and FOLR2\textsuperscript{pos} clusters) that produced abundant lipid resolvins instead of inflammatory mediators and induced a repair phenotype in synovial tissue stromal cells (19). Altogether, this raised the hypothesis that macrophage clusters in the lung, functionally equivalent to those in the synovium, may govern chronic inflammation or resolution of COVID-19 pneumonitis, and that the hallmark cytokine signature related to these clusters (e.g., SPP1) might be a useful biomarker of prognosis and a therapeutic target during the unpredictable trajectory of COVID-19.
Results

COVID-19 BALF FCN1<sup>pos</sup> and FCN1<sup>pos</sup>SPP1<sup>pos</sup> macrophage clusters are transcriptionally similar to CD48<sup>high</sup>S100A12<sup>pos</sup> and CD48<sup>pos</sup>SPP1<sup>pos</sup> clusters that drive RA synovitis

In COVID-19 broncho-alveolar lavage fluid (BALF), Liao et al (24) identified 4 major clusters of macrophages characterized by a combination of expression of SPP1, FCN1 and FABP4. Importantly, expansion of the FCN1<sup>pos</sup> and FCN1<sup>pos</sup>SPP1<sup>pos</sup> clusters was indicative of COVID-19 severity (Fig.S1A-B). Among the 9 phenotypically distinct clusters of synovial tissue (ST) macrophages that differed in distribution between health and RA, the CD48<sup>high</sup>S100A12<sup>pos</sup> and CD48<sup>pos</sup>SPP1<sup>pos</sup> clusters were expanded in RA patients with active disease (19) (Fig.S1C-D). To test the relationship between COVID-19 BALF and RA ST macrophage clusters, we integrated the macrophage scRNAseq datasets from COVID-19 BALF (GSE145926) (24) (Fig.S1A-B) and our RA synovial tissue (E-MTAB-8322) (19) (Fig.S1C-D) using Seurat V3 anchor integration strategy (37). The dimensional reduction of integrated macrophage datasets illustrated by Uniform Manifold Approximation and Projection (UMAP) visualize overlapping and tissue-specific macrophage clusters (Fig.1A-B). To investigate the relationship between these macrophage clusters from synovitis and pneumonitis, hierarchical clustering (Fig.1C) was performed on the matrix of the average expression of each ST and BALF cluster of all 8902 genes that were common to both datasets. The hierarchical clustering dendrogram illustrating the relationship between the clusters by branch-point (split) and branch-length (distance) revealed that the macrophage clusters were separated predominantly by their precursor origin and function, i.e., embryonic-origin (homeostasis) or monocyte-derived (inflammation) rather than source tissue. The first branch-point separated 4 tissue-resident macrophage clusters (healthy lung FABP4<sup>pos</sup> and FABP4<sup>low</sup>FCN<sup>neg</sup>SPP1<sup>pos</sup>, and synovial lining TREM2<sup>high</sup> and TREM2<sup>low</sup>) from the other
BALF and ST macrophage clusters that were likely of monocytic origin (19, 24, 25). The split-points and distances in the second branch of hierarchical clustering indicate that the pro-inflammatory clusters (i.e., FCN1\(^{pos}\) and FCN1\(^{pos}\)SPP1\(^{pos}\)) which dominate in severe COVID-19 share transcriptomic profiles with pro-inflammatory clusters (i.e., CD48\(^{high}\)S100A12\(^{pos}\) and CD48\(^{pos}\)SPP1\(^{pos}\)) in active RA, respectively. Pearson correlation analysis between BALF FCN1\(^{pos}\) and ST CD48\(^{high}\)S100A12\(^{pos}\), and between BALF FCN1\(^{pos}\)SPP1\(^{pos}\) and ST CD48\(^{pos}\)SPP1\(^{pos}\) confirmed resemblance (r=0.56, p=2.2\(^{-16}\) and r=0.65, p=2.2\(^{-16}\)) of respective BALF/ST pairs. Independent differential gene expression analysis identified striking similarities in the pathogenic pathways of COVID-19 pneumonitis FCN1\(^{pos}\) and active RA synovitis CD48\(^{high}\)S100A12\(^{pos}\) clusters. They share 238 top marker-genes (Fig.1D) that include up-regulation of inflammatory and prothrombotic pathways. These consisted of IFN pathway (e.g., IFITM2, IFITM3 and ISG15), inflammation-triggering alarmins (S100A8/9/12A), B-cell activation factors (e.g., BAFF-TNFSF15B), promotors of IL-1β and TNF production (e.g., CARD16 and LITAF), prothrombotic factors (e.g., FGL2) and integrins mediating cell migration and adhesion (e.g., ITGB2 and ITGAX). They also share receptor expression profiles e.g., TNFR2 (TNFRSF1B), G-CSFR (CSF3R) and the complement receptor (C5aR) (38) that may render them increasingly susceptible to pro-inflammatory mediators to further escalate inflammation (Fig.1E). The hierarchical analysis and comparison of cluster markers also indicate a shared transcriptional profile of COVID-19 FCN1\(^{pos}\)SPP1\(^{pos}\) and synovitis CD48\(^{pos}\)SPP1\(^{pos}\) clusters consisting of 86 common marker-genes including SPP1, which have broad pro-inflammatory and fibrotic functions (39) (Fig.1F-G). These transcriptomic profiles suggest that BALF FCN\(^{pos}\) and FCN\(^{pos}\)SPP1\(^{pos}\) macrophage clusters predominant in severe COVID-19 shared pathogenic molecular pathways, including the expression of their signature mediators S100A12 and SPP1 (Fig.1H) with ST CD48\(^{high}\)S100A12\(^{pos}\) and CD48\(^{pos}\)SPP1\(^{pos}\) clusters predominant in active RA. Analysis of SPP1 and S100A12 expression in other cell
types present in lung tissue (Fig.1I and Fig.S2) demonstrated that BALF macrophage populations are the main cells expressing SPP1 and S100A12 in severe COVID-19, and due to their vast numbers in BALF of severe COVID-19 (Fig.S1), FCN\textsuperscript{pos} and FCN\textsuperscript{pos}SPP1\textsuperscript{pos} macrophage clusters are likely the main contributors to the S100A12 and SPP1 pool in the COVID lung.

Healthy lung alveolar FABP4\textsuperscript{pos} macrophages share a homeostatic regulatory transcriptomic profile, including the TAM pathway, with synovial tissue lining-layer TREM2\textsuperscript{high} macrophages

The innate immune pathways that restrain inflammation in acute COVID-19 are yet to be characterized. Recently uncovered innate mechanisms reinstating synovial tissue (ST) homeostasis in RA (19) might shed new light on potential mechanisms of resolution for COVID-19. The ST of RA patients in sustained disease remission is dominated by macrophage clusters with protective and inflammation-resolving properties (19). Of these resolving clusters, ST TREM2\textsuperscript{high} clustered tightly with the FABP4\textsuperscript{pos} alveolar macrophage cluster in unsupervised hierarchical analysis (Fig.1C). Pearson correlation of 986 TREM2\textsuperscript{pos} and FABP4\textsuperscript{pos} unique and common cluster-markers (r=0.83 p=2.2^{-16}) confirmed the similarities. Differential expression analysis of their profiles identified that BALF FABP4\textsuperscript{pos} and ST TREM2\textsuperscript{high} clusters share 170 markers indicating analogous homeostatic functions (Fig.2A-B). These include the complement pathways (\textit{e.g.} C1q) that facilitates uptake of apoptotic bodies, high expression of genes of retinoic acid production (\textit{e.g.}, ALDH1A1 and RBP4) that drives T-reg differentiation (40), and the B7-related co-inhibitory molecule VSIG4 that inhibits T-effector cells (41); together suggesting a primary role of this cluster in governing lung immunity (Fig.2B). While the functional contribution of FABP4\textsuperscript{pos} macrophages to the resolution of SARS-CoV-2 infection is yet to be established, we have shown previously that
their counterpart TREM2pos ST macrophage clusters produced inflammation-resolving lipid mediators and induce a repair phenotype in tissue stromal cells that maintain disease remission. These homeostatic responses were driven by MerTK, a member of the immunosuppressive tyrosine kinase receptor TAM family (TYRO, AXL and MerTK) (19). TAM receptors and their ligands GAS6 or the PROS1 gene product Protein S form a homeostatic brake on inflammation and autoimmunity (42-44). In addition, Protein S is an essential inhibitor of blood coagulation preventing thrombosis (45). Lung-resident FABP4 macrophages uniquely express AXL rather than MerTK. Their AXL is constitutively ligated to GAS6 (46, 47) and is key in preventing exacerbated inflammation e.g., during influenza virus infection (46, 47). Analysis of the GSE145926 dataset (24) identified profoundly altered macrophage expression of TAM receptors and their ligands in the lung (Fig.2C-D) that might explain the inadequate regulation of the tissue hyper-inflammatory and thrombotic responses in severe COVID-19. The TAM receptors and ligands pathway was mostly expressed by lung myeloid cells, with a contribution from ciliated epithelium to the lung Protein S pool (Fig.S2-3). Alveolar macrophages from healthy lungs showed high expression levels of AXL and PROS1, and these were markedly reduced in severe COVID-19 patients (Fig.2C-D). GAS6 and MerTK were not expressed by resident AMs; instead, they were increasingly expressed by infiltrating FCNpos and FCNposSPP1pos macrophage clusters, suggesting an inflammation-triggered attempt to counterbalance pathogenic responses. However, the reduced PROS1, which is the preferred activating-ligand for MerTK (44, 48), and the fewer homeostatic resident AMs in severe COVID-19 might enable unrestricted pro-inflammatory cytokine production by locally differentiated MerTK-expressing FCNpos and FCNposSPP1pos macrophages.
High plasma SPP1 is associated with a severe disease trajectory of COVID-19

To validate the computational scRNAseq findings of shared pathogenic macrophage clusters in COVID-19 and RA, we quantified plasma concentrations of their key shared functional mediators SPP1 (osteopontin) and S100A12 (calgranulin C) in a cross-sectional comparison of additional patient groups. These consisted of (i) hospitalized acute COVID-19 patients (n=92), (ii) COVID-19 patients in their post-COVID-19 phase (convalescence) at outpatient clinics (n=41), (iii) hospitalized patients with non-SARS-CoV-2 community-acquired severe acute pneumonia (n=29), and (iv) healthy donors (n=10) (Fig.3A). Demographic and clinical information, and laboratory biomarkers are presented in Table S1. Blood samples for groups (i) and (iii) were taken on hospital admission or shortly after, and before anti-inflammatory treatment (see Methods).

Patients with acute COVID-19 had significantly higher plasma levels of pro-inflammatory SPP1 and S100A12 than healthy donors and patients at post-COVID-19 stage (Fig.S4A). Based on clinical criteria (see Methods) and before anti-viral/anti-inflammatory treatment administration, the acute COVID-19 patients were categorized into those who subsequently developed either mild/moderate or severe disease. We found that the S100A12 plasma levels were higher in COVID-19 patients categorized as severe compared to mild/moderate, and to those in convalescence, and to healthy donors, but comparable to the levels in patients with community acquired pneumonia (Fig.3B). Interestingly, levels of SPP1 were uniquely higher in those with severe COVID-19 compared to any other groups, including severe pneumonia induced by other pathogens (Fig.3B).

We next investigated the relationship between plasma levels of SPP1 and S100A12, with systemic inflammation measured by blood cytometry and biochemistry, and lung function measured by PaO₂/FiO₂ tested around the time of blood sampling and before administration of anti-inflammatory treatment (Table S1). Plasma concentrations of SPP1 and S100A12
correlated negatively with PaO$_2$/FiO$_2$ and positively with inflammatory biomarkers e.g., IL-6, CRP and LDH (Fig.3C and Table S1). Moreover, SPP1 and S100A12 each strongly correlated with the neutrophil/lymphocyte ratio which itself is a prognostic biomarker for COVID-19 severity (49) (Fig 3C). Stratification of patients according to levels of PaO$_2$/FiO$_2$ reflecting those with or without severe respiratory distress (PaO$_2$/FiO$_2$≤200 and PaO$_2$/FiO$_2$ >200, respectively) clearly showed that the former had significantly higher levels of SPP1 and S100A12 (Fig.3D). A refinement of this observation suggested that COVID-19 patients with plasma levels of SPP1 ≥108 ng/ml, and S100A12 ≥59 ng/ml (cut-off values based on their medians in all COVID-19 patients) were predictive of those more likely to have severe respiratory distress (PaO$_2$/FiO$_2$ ≤200) than those with lower SPP1 and S100A12 levels (Fig.3E and Fig.S4B). Of prognostic importance for COVID-19 patients, higher levels of SPP1 and S100A12 at the time of hospital admission and before anti-inflammatory treatment were predictive of urgency for subsequent ICU transfer (Fig.3F-G). The time to ICU transfer was significantly more rapid for patients with plasma SPP1 ≥108 ng/ml and S100A12 ≥59 ng/ml (Fig.3H). Altogether these data suggest that a plasma biomarker signature associated with pathogenic macrophage clusters in the lung and shared with RA synovitis (i.e., SPP1 and S100A12) might be useful for predicting the trajectory of disease severity and indicative of mechanism of pathogenesis of severe COVID-19. Analysis of the SPP1 (CD44/integrins) and S100A12 (TLR4/CD36) receptors distribution in the lungs of COVID-19 patients demonstrated that many cell types expressed receptors for SPP1, including macrophages, neutrophils, T cells and epithelial cells, whereas predominantly macrophages expressed receptors for S100A12 (Fig.S5). This suggested that the lungs were a receptive environment for the actions of pro-inflammatory SPP1 and S100A12.

We next investigated the plasma concentrations of COVID-19/RA shared regulatory TAM receptor pathway ligands GAS6 and PROS1 in the patient groups described
above. They showed a distinct plasma signature compared to pro-inflammatory SPP1 and S100A12. At the time of acute pneumonia, plasma levels of GAS6, the preferable ligand for the AXL receptor, did not differ between acute severe COVID-19 and healthy donors. However, GAS6 levels were lower in patients with milder acute COVID-19 and in post-COVID-19 subjects in whom acute inflammation had resolved, compared to healthy donors and patients with severe COVID-19 (Fig.3B). Consistent with lower levels of GAS6 in patients with mild/moderate disease, GAS6 levels in COVID-19 patients correlated negatively with PaO$_2$/FiO$_2$ ratio (Fig.3C). COVID-19 patients with PaO$_2$/FiO$_2$ >200 had lower plasma GAS6 levels than those with PaO$_2$/FiO$_2$ ≤200 (Fig 3D) and optimum cut-off plasma levels of GAS6 <24 ng/ml detected those more likely to maintain good lung function (PaO$_2$/FiO$_2$ >200) (Fig 3E). Prognostically, low levels of GAS6 at the time of hospital admission and/or before anti-inflammatory treatment were predictive of low risk for ICU transfer (Fig.3F); for example, 6.5% of COVID-19 patients with GAS6 <24 ng/ml required transfer to ICU compared to 44.4% of those with GAS6 ≥24 ng/ml (Fig.3G-H). These data suggest that lower GAS6 levels were predictive of better disease outcomes. GAS6 can be produced by many immune cells (45, 50), including lung-infiltrating inflammatory FCN1$^{pos}$ and FCN$^{pos}$SPP1$^{pos}$ macrophages and dendritic cells (Fig.2C and Fig.S2). Thus, low levels of plasma GAS6 in patients with milder disease may reflect a lower degree of lung infiltration by inflammatory macrophages.

Plasma level of PROS1, a marker of healthy lung $FABP_{4}^{pos}$ macrophages, and the preferable ligand for MerTK, did not differ between any of the COVID-19 patient categories or healthy donors (Fig.3B). However, the levels of PROS1 in COVID-19 patients with mild/moderate disease showed a wider spread of the values as illustrated by interquartile range compared to other groups, suggesting potential association between PROS1 and individual composites of the classification criteria. In contrast to SPP1, A100A12 and GAS6, PROS1 levels correlated positively with PaO$_2$/FiO$_2$ and negatively with IL-6 (Fig.3C). COVID-19
patients with PaO$_2$/FiO$_2$ >200 had higher plasma levels of PROS1 than those with PaO$_2$/FiO$_2$ ≤200 (Fig.3D). An optimum plasma PROS1 cut-off value was calculated at 15 µg/ml, above which 34.1% of COVID-19 patients had severe respiratory distress as compared to 59% of those with PROS1 <15 µg/ml (Fig.3E). This observation is consistent with the single-cell transcriptomic data (Fig.2C) showing that the expression of inflammation resolving and anti-thrombotic PROS1 in resident alveolar FABP4$^{\text{pos}}$ macrophages decreases with increasing disease severity.

Next, we investigated whether concomitant pharmacological drug-treatment, demographic factors (age and gender) and COVID-19 comorbidities (6) contributed to the plasma levels of SPP1, S100A12, GAS6 and PROS1. None of the pharmacological treatments affected the plasma levels of SPP1, S100A12, GAS6 and PROS1 in moderate and severe COVID-19 pneumonia (Fig.S6). S100A12 and PROS1 were unrelated to age in COVID-19 and in SARS-CoV-2 negative pneumonia (Fig.S4C-E). SPP1 was unaffected by age in SARS-CoV-2 negative pneumonia, but was higher in COVID-19 patients >70 (Fig.S4C-E) suggesting that SPP1 levels are related more to severity of pneumonia than age (Fig.3). GAS6 was higher in COVID-19 and non-SARS-CoV-2 pneumonia patients >70 years, and higher in COVID-19 patients with arterial hypertension, diabetes mellitus and ischemic cardiopathy (Fig.S7) thus age and comorbidities may confound interpretation of increased GAS6 in COVID-19.

**Increased SPP1 levels persist post-COVID-19**

We investigated the persistence of increased plasma SPP1, S100A12, GAS6 and PROS1 into the SARS-CoV-2-negative post-COVID-19 phase; often characterized by complex pathologies (26). We compared plasma concentrations in 41 post-COVID-19 patients attending outpatient clinic at (mean±SEM) 68.60 ± 4.36 days after discharge (Fig.4A), 26 of whom plasma samples were available from the peak of acute COVID-19. Longitudinal
comparison showed that plasma levels of SPP1, S100A12 and GAS6 were significantly lower in convalescence compared peak disease, but PROS1 was unchanged (Fig. 4B). However, although reduced, the levels of SPP1 and S100A12 remained significantly higher than in healthy control donors, irrespective of whether the prior disease trajectory of COVID-19 was mild/moderate or severe (Fig 4C). This was in contrast to the levels of sensitive markers of inflammation including IL-6 and LDH, which normalized in post-COVID-19 to levels of those in healthy donors, consistent with resolved acute inflammation (Fig.4D and Table S1). GAS6, lower than normal control levels in mild/moderate COVID-19 on hospital admission (Fig.3B), remained low in post-COVID-19 (Fig.4C).

Most (36/41) convalescent patients reported persistence of at least one post-COVID-19 symptom (fatigue, musculoskeletal or respiratory). Increased levels of SPP1 and S100A12 persisted in all, although extremely high concentrations SPP1 were restricted to symptomatic patients (Fig.4E). Stratification of patients by symptoms showed that SPP1 and S100A12 levels remained increased and GAS6 decreased in convalescent COVID-19 patients irrespective of symptom category (Fig.S8).

Together, these data suggest that the SPP1 and S100A12 myeloid cell inflammatory signature persists after resolution of SARS-CoV-2 infection and may contribute to the pathogenesis of long-COVID-19 syndrome.

**SPP1 protein is expressed by COVID-19 BALF macrophages, drives pro-inflammatory activation of classical monocytes, and the differentiation of neutrophils towards a pro-inflammatory CD274 (PD-L1)pos phenotype.**

High plasma SPP1 selective for severe COVID-19 (Fig.3B) suggested a role for SPP1 in pathogenesis. SPP1 and CD68 staining of post-mortem lung tissue confirmed abundant clusters of SPP1pos macrophages in alveoli of COVID-19 patients (n=2), rare in normal lung (n=3)
(Fig.5A-B and Table S2) and sparse in alveoli of bacterial (n=3) and H1N1 (n=3) pneumonia (Fig.S9A-B). To investigate the biological effects of SPP1, we stimulated healthy whole blood cells with SPP1 at concentrations equivalent to those in severe and post-COVID-19 (200 and 50 ng/ml, respectively Fig.3B). SPP1 receptors (e.g., CD44/integrins) are expressed by many immune cell types in lung and blood (Fig.S5 and Fig.S10); therefore, to capture the effects of SPP1 on all cell types we used scRNAseq of the whole blood culture (n=3 healthy donors) using the immune gene panel (399 genes) with the BD_Rhapsody system after 16h stimulation with SPP1. We sequenced 13,580 cells and identified 14 distinct immune clusters (Fig.6A-B) including neutrophil, eosinophil, monocyte, dendritic cell, NK cell and lymphocyte cell-clusters. The relative proportions of cell clusters revealed that SPP1 stimulation at 200 ng/ml, equivalent to plasma concentration of severe COVID-19, was associated with a significant increased neutrophil proportion (Fig.6B-C) with a shift from the normally dominant CD10\(^\text{MME}\) neutrophil phenotype to a CD274 (PD-L1)\(^\text{pos}\) neutrophil phenotype. The top 20 marker-genes of the CD10\(^\text{pos}\) and SPP1-differentiated CD274 (PD-L1)\(^\text{pos}\) neutrophil clusters (Fig.5D and Table S3) revealed that the CD274 (PD-L1)\(^\text{pos}\) neutrophils represent an aberrantly-activated phenotype characterized by high expression of PD-L1 (CD274) which mediates suppression of adaptive immunity, high levels of pro-inflammatory cytokines/chemokines (e.g., IL1B, TNF, CCL3 and CCL4), c-type lectins (CLEC4E/D) that facilitate neutrophils migration, and the SPP1-receptor CD44. Similar aberrant neutrophil activation is a key pathogenic characteristic of severe COVID-19 (51, 52). To investigate if the in vitro SPP1-driven neutrophil cluster transcriptionally replicated those in severe COVID-19, we performed unsupervised mapping of the SPP1-driven CD274 (PD-L1)\(^\text{pos}\) neutrophil gene module (37 genes, Table S3) into a whole-blood scRNAseq transcriptomic signature of healthy, mild and severe COVID-19 (22). We identified that enrichment of transcriptional signature of SPP1-driven CD274 (PD-L1)\(^\text{pos}\) neutrophils was significantly higher in Neutrophil_2 phenotype than any other cell population.
(Fig.5E-F) in COVID-19 patients, and that this phenotype was uniquely increased in severe disease, suggesting that SPP1 might be responsible for the pathogenic activation of neutrophils in severe COVID-19. This link between SPP1 and neutrophils is supported by the correlation between plasma SPP1 and the blood neutrophils/lymphocytes ratio in our COVID-19 cohort (Fig.3C).

Our analysis also revealed that SPP1 induced the activation of CD14pos classical monocytes. At 200 ng/ml, SPP1 strongly increased the expression of alarmins (S100A12, S100A9), IL-1β and chemokines (CXCL8, CCL2-4) with a commensurate reduction of MHCII and anti-inflammatory ENTPD1, (encoding CD39) (Fig.6G) which resembles the pro-inflammatory changes in blood CD14pos monocytes subset in severe COVID-19 (22, 23). SPP1 at post-COVID-19 convalescent plasma concentration (50 ng/ml) did not affect neutrophil activation, but increased expression levels of alarmins S100A12 and S100A9, and decreased the expression of MHCII and ENTPD1 of CD14pos monocytes (Fig.6G), suggesting that persistent levels of SPP1 may contribute to long-COVID-19 pathologies by skewing monocytes towards a pro-inflammatory phenotype. The effect of SPP1 on other cell types was minimal (Table S3). In summary, SPP1 produced by pro-inflammatory lung-infiltrating macrophages might be an upstream activator of the aberrant innate response in COVID-19 by driving pro-inflammatory activity of CD14pos monocytes, and differentiation of CD274 (PDL1)pos neutrophils that suppress adaptive immunity and support inflammation.
Discussion

Our comparative single-cell transcriptomic analysis integrating myeloid cell clusters from COVID-19 pneumonitis and RA synovitis suggested that COVID-19 and RA pathogenesis and resolution might be driven by similar myeloid cell clusters and their signature functional pathways (Fig.S1). The shared key pathogenic macrophages were SPP1\textsuperscript{pos} and S100A12\textsuperscript{pos} clusters and their signature SPP1 and S100A12 mediators were confirmed in COVID-19 cross-sectional and longitudinal patient plasma samples. COVID-19 patients with severe respiratory distress were characterized by an aberrant raised SPP1 and S100A12 cytokine signature that could predict the urgency for ICU transfer and persisted into the post-COVID-19 phase after hospital discharge. An abundance of SPP1\textsuperscript{pos} macrophages in the alveolar spaces, and high SPP1 plasma levels were unique to severe COVID-19 pneumonia as compared to pneumonias induced by other pathogens. Investigation into SPP1 mechanisms of action revealed that it drives pro-inflammatory activation of CD14\textsuperscript{pos} monocytes and development of PD-L1\textsuperscript{pos} pathogenic neutrophils, both hallmarks of severe COVID-19 (1-14). Thus, COVID-19 pneumonitis appears driven by similar pathogenic myeloid cell pathways as those in RA, and their mediators such as SPP1 may be an upstream activator of the aberrant innate response in severe COVID-19 and predictive of disease trajectory including post-COVID-19 monitoring.

In contrast, COVID-19 patients with milder respiratory distress were characterised by increased plasma levels of the inflammation-resolving, anti-thrombotic cytokine Protein S (PROS1), which characterised the resident alveolar macrophage FABP4\textsuperscript{pos} cluster. This cluster shared transcriptomic similarities with the inflammation-resolving synovial tissue macrophage (STM) TREM2\textsuperscript{pos} cluster, suggesting a potential role in controlling lung inflammation.

The BALF FCN1\textsuperscript{pos} and FCN1\textsuperscript{pos}SPP1\textsuperscript{pos} macrophage clusters from severe COVID-19 patients shared transcriptomic profiles with STM CD48\textsuperscript{high}S100A12\textsuperscript{pos} and CD48\textsuperscript{pos}SPP1\textsuperscript{pos}
clusters from active RA, respectively. The functional biology of the synovial tissue macrophage counterparts (i.e., CD48\textsuperscript{high}S100A12\textsuperscript{pos} and CD48\textsuperscript{pos}SPP1\textsuperscript{pos}) showed that they were the main producers of pathogenic TNF, IL-6, IL-1\(\beta\) and chemokines in the synovium of RA and this resembles the lung hyper cytokine environment characterizing severe COVID-19 respiratory distress syndrome (53, 54), suggesting that SPP1\textsuperscript{pos} and S100A12\textsuperscript{pos} BALF macrophage clusters might be the key source of these mediators in the COVID-19 lung. These shared macrophage clusters of RA/COVID-19 are also characterized by their cluster-unique mediators, SPP1 (CD48\textsuperscript{pos}SPP1\textsuperscript{pos} cluster) and S100A12 (CD48\textsuperscript{high}S100A12\textsuperscript{pos}cluster) (19). These mediators were upregulated in severe compared to mild/moderate COVID-19, and their high plasma concentrations correlated with respiratory insufficiency, and was predictive for urgency of ICU transfer. While high plasma levels of the alarmin S100A12 characterised both severe COVID-19 and SARS-CoV-2-negative community-acquired pneumonia, the presence of SPP1\textsuperscript{pos} macrophages and high plasma SPP1 levels were selectively upregulated in severe COVID-19. Many biological functions for SPP1, mostly in tissue remodelling, have been described (39, 55) however, its contribution to COVID-19 pathologies is unclear. We revealed here that SPP1 at the level detected in patients with severe COVID-19 is a potent driver of pathogenic PD-L1\textsuperscript{pos} neutrophils that are characteristic of severe COVID-19 (22). Thus, SPP1 might be an upstream regulator of neutrophil-mediated tissue damage and immune-thrombosis that are key pathogenic features of severe COVID-19 (51, 52). In addition, SPP1 induced strong pro-inflammatory activation of CD14\textsuperscript{pos} monocytes, including the expression of S100A alarmins, suggesting that it may contribute to the development of S100A12\textsuperscript{pos} macrophage clusters in the lung.

There is an increasing recognition of a syndrome of persistent debilitating symptoms after the COVID-19 patients become SARS2-CoV-2 negative (post-COVID-19). The symptom patterns are complex but dyspnoea and fatigue are prominent (26). The pathogenic
mechanisms of this syndrome are unknown. Emerging transcriptomic evidence shows that changes in blood myeloid cells elicited during acute infection can normalize within 14 days of becoming SARS2-CoV-2 negative (56). However, some changes can persist for at least 12 weeks, including high levels of surface molecules that regulate monocyte migration into tissue e.g., CXCR6 and VLA-4 (57). We found that in contrast for example to IL-6, plasma SPP1 and S100A12 remained significantly higher than normal for at least 10 weeks after infection clearance in the post-COVID-19 convalescent phase. Our study revealed that SPP1, at the concentration detected in post-COVID-19 stage, induced some of the features of pro-inflammatory CD14pos monocytes, for example an increase in alarmins (S100A12 and S100A9) and decrease in MHCII molecules. This suggests that sustained raised level of SPP1 may contribute to post-COVID-19 pathologies and may influence future responses against pathogens by skewing monocytes towards pro-inflammatory phenotype. At this stage of understanding, it remains unclear whether the plasma SPP1 and S100A12 exclusively originate from the pathogenic BALF clusters persisting in the lung of these patients. However, regardless of SPP1 and S100A12 source their wide range of pro-inflammatory and fibrotic functions may be responsible for the respiratory and musculoskeletal symptoms that persist in convalescence.

One of the aims of this study was to apply our recent discovery of inflammation-resolving macrophage pathways in RA to investigate potentially similar inflammation-resolving pathways for COVID-19. We found that tissue resident lung alveolar FABP4pos macrophages and homeostatic synovial lining layer TREM2pos macrophages share transcriptomic profiles. This may reflect their homeostatic functions in their respective tissues. In the synovium, the TREM2pos macrophages form a lining-layer producing and recycling lubricant synovial fluid which facilitates joint movement (58) while in lung alveoli, the FABP4pos macrophages contribute to alveolar integrity; recycling surfactants to maintain patency and facilitate gas exchange (33). One of the pathways they shared was inflammation-
resolving TAM (TYRO, AXL, MerTK) receptor pathways. TAM pathways are widely expressed by resident macrophages in many tissues (59, 60) and their deregulation has pro-inflammatory/autoimmune consequences in pre-clinical animal models (61, 62) and in patients with RA (19). The scRNAseq data showed that in healthy human BALF the FABP4pos macrophages express the TAM receptor AXL, and the ligand for the TAM receptor MerTK (PROS1); and that these were profoundly repressed in severe COVID-19. Consistent with scRNAseq data, plasma Protein S levels were higher in COVID-19 patients who maintained lung functions (i.e., sustained PaO2/FiO2 >200), suggesting a potential role in counterbalancing the severity of inflammation of SARS-Cov-2 infection.

Limitations of the study included the younger age and fewer numbers of the healthy control group used as a comparator of some of the COVID-19 analysis. These limitations were mitigated by findings that SPP1, S100A12 and PROS1 were unaffected by age. In addition, the control group for the key mediator SPP1 in severe COVID-19 was the non-SARS-CoV-2 pneumonia group that matched COVID-19 for age, and showed lower levels of SPP1 compared to COVID-19, supporting the link between SPP1 and COVID-19 pathogenesis.

In summary, this study suggests that the pathogenesis of acute severe COVID-19 pneumonitis and RA synovitis might be driven by similar pathogenic myeloid cell clusters/pathways, producing SPP1 that persist into post-COVID19 syndrome. Further functional studies on identification of common tissue factors driving the differentiation of these pathogenetic macrophage clusters and better understanding the functional interaction between them and lung environment could clarify mechanisms of pneumonitis and provide evidence for potential repurposing of current anti-inflammatory/anti-fibrotic treatments for COVID-19 (63, 64). Promising data from current COVID-19 clinical trials of drugs already used for the treatment of RA (e.g., dexamethasone, tocilizumab and baricitinib) (35, 65-69) further support
the concept of common pathogenic and resolution mechanisms that could be capitalized upon for COVID-19 therapeutic exploitation.
Materials and Methods:

Data Acquisition and Analysis: comparison and integration of 10X Genomics scRNAseq data (BALF myeloid cells (24) and from synovial tissue (19)). BALF data was acquired as CellRanger output (GEO: GSE145926) and ST data from EMBL-EBI :E-MTAB-8322. COVID-19 blood and PBMC data (22) were from https://beta.fastgenomics.org/p/schulte-schrepping_covid19.

Quality checks, filtering, clustering of 10X Genomics datasets. Seurat (3.1.2) in R created an object (CreateSeuratObject). Cell-filtering involved removal of cells with <500 expressed genes (subset, subset=nFeatures_RNA >500), with set-thresholds for gene-expression level, including mitochondrial genes (percent.mt) to exclude doublets and dying cells. Data normalized using Seurat’s NormalizeData function. Myeloid cells were filtered for expression of CD14, MARCO, LYZ with the subset function. The top 2000 variable genes were identified using the FindVariableFeatures function. Sample integration of each dataset followed the Seurat vignette, integrating all genes common between samples of each condition, using FindIntegrationAnchors and IntegrateData functions (features.to.integrate to normalize all common genes). Individual Clustering and Dimensional Reduction: UMAP based on PCA cell-embeddings were generated by Seurat for each dataset using RNA assay. To visualize ST data, the first 12 principal components were used (RunUMAP). These PCs determined the k-nearest neighbors for each cell during SNN graph construction before clustering at a resolution of 0.5 (FindNeighbors, FindClusters) and identifying the populations as described (19). In comparison, the first 50 principal components (0.8 resolution) were used for visualization and clustering of BALF data (24), and populations were identified by merging clusters based on expression of FCN1, SPP1 and FABP4. The relative proportion of clusters between conditions was analysed using Kruskal-Wallis test with Dunn’s correction (Prism v8.4.2). Differential
Expression Analysis. The Seurat FindAllMarkers function identified cluster markers in individual datasets. The “test.use” function determined genes differentially expressed between clusters within each dataset using MAST. As recommended, for DE comparison the non-batch normalized counts were used. Cluster markers must be expressed by ≥40% cells in the cluster (‘min.pct’ parameter 0.4). Default values were used for all other parameters. Genes are considered significantly DE if p<0.05 (Bonferroni Correction, multiple by number of tests/clusters). Ligand-receptor expression analysis. CellTalkDB (70) identified putative receptors of mediators of interest (S100A12, SPP1, GAS6 and PROS1). An object containing all sequenced BALF cell-types was prepared for visualization of expression of mediators and identified receptors by distinct BALF immune and epithelial clusters. The same QC and integrative clustering procedure were followed as described above for analysis of myeloid cells, without filtering cells for expression of CD14, MARCO, LYZ. The first 30 principal components (0.3 resolution) were used for classification of BALF immune and epithelial cell types (24). Clusters were annotated by marker-gene expression, and expression of genes-of-interest were visualized (FeaturePlot, RNA assay). Pseudobulk expression values of genes-of-interest were generated per sample, per cell-type, and exported for analysis (GraphPad Prism 9.1.0). A similar strategy was used to analyse SPP1 receptor-expression in COVID-19 blood (22), and data was obtained as a Seurat object with clustering and annotation pre-performed.

Dataset Comparison. Dataset integration was performed as described above for 8993 common genes between ST and BALF. These “integrated” batch-corrected values were set as the default assay, and gene expression values are scaled before principal component analysis. Clustering and Dimensional Reduction. To prevent bias in clustering, we matched BALF and ST cell numbers at 32,139 random cells. Data were re-scaled (ScaleData) and the first 30 principal components were visualised by UMAP. Cells coloured by original identity illustrated
how clusters of each dataset overlay. **Comparative analyses.** A count matrix with the average expression of all common genes by each cluster was generated (AverageExpression) before hierarchical clustering (dist, hclust) and plotting a dendrogram. Principal component analysis (PCA) was performed on this pseudobulk expression matrix (prcomp). These analyses identified similar clusters between BALF and ST. Shared marker-genes between similar clusters were identified by Venny (2.1). To visualize pseudobulk expression of overlapping marker-genes, the pheatmap (1.0.12) package visualized shared genes as a heatmap using a custom script. Between-dataset comparison of pseudobulk of expression marker genes (unique and common to all clusters) was performed by Pearson correlation (ggscatter, add = "reg.line", conf.int = TRUE, cor.coef = TRUE, cor.method = "pearson") performed on the 916, 721 and 986 unique and common genes of BALF FCN1\textsuperscript{pos}/ST CD48\textsuperscript{high}S100A12\textsuperscript{pos}, BALF FCN\textsuperscript{pos}SPP1\textsuperscript{pos}/ST CD48\textsuperscript{pos}SPP1\textsuperscript{pos} and BALF FABP4\textsuperscript{pos}/ST TREM2\textsuperscript{high} clusters.

**Patients and clinical assessment:**

**COVID-19 pneumonia patients.** RT-PCR-positive for SARS-CoV-2 by NPS were enrolled (COVID-Hospital, Fondazione Policlinico Universitario A.Gemelli, Rome). Each provided blood samples and COVID-19 severity was classified on hospital admission and/or before anti-inflammatory treatment. **Mild/moderate** COVID-19 pneumonia was defined based on (mild) normal O\textsubscript{2} saturation (>94%), or (moderate) abnormal O\textsubscript{2} saturation (<94%); or pneumonia on imaging (chest X-Ray or CT scan) and arterial oxygen partial pressure (PaO\textsubscript{2} in mmHg) to fractional inspired oxygen (FiO\textsubscript{2}) PaO\textsubscript{2}/FiO\textsubscript{2} (>200). **Severe** COVID-19 pneumonia was defined based on abnormal O\textsubscript{2} saturation (<94%), presence of pneumonia on imaging (X-Ray or CT scan) and PaO\textsubscript{2}/FiO\textsubscript{2} \leq 200, or the use of high-flow nasal cannula (HFNC), non-rebreather mask (NRB), bilevel positive airway pressure (BIPAP) or mechanical ventilation and vasopressor drugs use, and on Creatinine Clearance greater than 30, and alanine
aminotransferase (ALT) less than 5X the upper limit of normal. Concomitant comorbidities including arterial hypertension, diabetes mellitus, ischemic cardiopathy, COPD, dyslipidaemia and cancer were recorded. The subsequent need and date of ICU admission, intubation status and O₂ therapy for each patient was recorded. Patients was treated following the internal protocol defined by the Gemelli Against COVID-19 taskforce as described (35).

**Convalescent post-COVID-19 patients.** Follow-up outpatient attendance (Fondazione Policlinico Universitario A.Gemelli) was offered to all discharged COVID-19 patients. Those recruited attended 68.6±4.4 days (mean±SEM) after discharge, were RT-PCR negative for SARS-CoV-2. They were assessed by a multidisciplinary team who collected demographic and clinical data, blood samples, and performed physical examination. Persistence of symptoms was recorded by questionnaire.

**Patients with community-acquired pneumonia.** Patients with pneumonia, RT-PCR negative for SARS-CoV-2 by NPS were enrolled at the Emergency Room (Fondazione Policlinico Universitario A.Gemelli). Demographic and clinical parameters were recorded, and blood sample collected. These formed a comparison pathological control group for COVID-19 pneumonia. All patients and health controls demographics and clinical characteristics are summarized in Supplementary Table 1.

**Lung tissue immunohistochemistry for SPP1⁺ macrophages.** Post-mortem lung tissue was obtained from COVID-19 pneumonia (n=2), bacterial pneumoniae (n=3), H1N1 pneumonia (n=3); and normal donors (n=3). Demographic and clinical data are summarized in Supplementary Table 2. Autopsies were performed at the Department of Woman and Child Health and Public Health, Institute of Pathology of The Fondazione Policlinico Universitario A.Gemelli in accordance with the guidelines of the Royal College of Pathologists (www.rcpath.org). Tissue collection was in accordance with appropriate protocols (71). Lung
tissue was fixed in 10% neutral-buffered formalin, embedded in paraffin, sectioned at 4 mm, and stained as described (19). Briefly, after blocking (2h, 10% human serum, 10% of goat serum, 1% BSA in TBS; Thermo Fisher Scientific) at room temperature, sections were incubated overnight at 4°C with the primary rabbit anti-human osteopontin (SPP1) antibody, (Ab8448/Abcam, at 1:100) and mouse anti-human CD68 (Dako, clone PG-M1 at 1:40) in blocking buffer above. Sections were washed twice (5 minutes, TBS plus 0.025% Triton X-100). Secondary antibody goat anti-rabbit IgG Alexa Fluor 488 (1:100), and goat anti-mouse IgG Alexa Fluor 660 (1:100) (both Invitrogen); in dilution buffer (TBS with 1% BSA) and incubated for 1 h at room temperature. Sections were washed x3 for five minutes in TBS and counterstained with DAPI (H-1800-2/ VECTASHIELD® Vibrance™ Antifade Mounting Medium with DAPI) and visualized with confocal microscopy with Zeiss Zen Black software. Negative controls were performed throughout using isotype matched antibodies.

**SPP1, S100A12, GAS6, PROS1, IL-6, IL-8, IL-1β and TNFα ELISA assay.** Blood samples were centrifuged (3500rpm/15 minutes) and plasma aliquots stored at -80°C until analysis by ELISA for SPP1(BMS2066), S100A12(DY1052), GAS6(DY885B), and PROS1(NBP2-60585) [R&D Systems, UK, Invitrogen/Thermo Fisher Scientific, NOVUS Biological] IL-6, IL-8, TNFα and IL-1β plasma levels were quantified by single-plex ELISA (Multi-cytokine test for ELLA, Bio-Techne, Minneapolis).

**SPP1 stimulation of whole blood for scRNAseq (BDRhapsody) analysis.** Leukocytes were collected by centrifugation from heparinised blood samples (30 ml from 3 healthy donors) after RBC lysis (ACK Lysing Buffer, A1049201, ThermoFischer Scientific). Leukocytes were plated at 5×10^6/well in a 24-well cell-culture plate in 0.5 ml of RPMI 1640 compete medium containing SPP1 (PeproTech, no.120-035) at concentrations of 0, 50 and 200 ng/ml. After 16
hours, cells were de-attached (Accutase solution; no. A6964, Merck), transferred to U-bottom 96-well plates, and harvested by centrifugation at 200g/4min/4°C. Cells from each donor/condition were labelled with unique Tags (below) using a Single-Cell Multiplexing Kit (BD, no. 633781) for 20 min at room temperature according to manufacturer protocol. Tag 1—ATTCAAGGGCAGCCGCGTCACGATTGGATACGACTGTTGGACCGG Tag 2—TGGATGGGATAAAGTCGATGTGGGACCAGGGACCTCCTTGCGCCCG Tag 3—CGGCTTCGTGCTGCCTCTCAAGTCCAGAAACTCCGTGTATCCT Tag 4—ATTGGGAGGCTTTCGTACCGCTGCCGCCACCAGGTGATACCCGCT Tag 5—CTCCCTGTTGTTCAATACCCCGATGTGGTGGGCGAATGTTGGCTGG Tag 6—TTACCCGCAGGAAGCAGTATACCCCTCGTGCCAGGCAGCAATGC Tag 7—TGCTACGTTCGAGCGCCAAGAAGTGAGTCAGAGGTGCACGTGT Tag 8—CCCCACCAGTTGCTTTGTCAGGAGCGAGCCGCACACAGCGACTAGGAT Tag 9—GTGATCCCGCCGACACACATACCCGACTCAGATGGGGTTGTCACCG Cell were washed x3 with PBS, with centrifugation (200g/4min/4°C), after which the tagged culture variants were pooled and loaded onto the scRNA-seq BD Rhapsody Cartridge using the BD Rhapsody Cartridge Reagent Kit (no. 633731) according to the manufacturer’s protocol. Single-cell cDNA was prepared using the BD Rhapsody cDNA Kit (no. 633773). This was followed by single-cell mRNA and tag-library preparation using BD Rhapsody Targeted mRNA and the Tag Amplification Kit (no. 633774) and primers for the BD Rhapsody Immune Response Panel (399 genes; no. 633750). Libraries were sequenced at a depth of 1,083,775 ± 236,302 (Mean ± S.E.M.) reads per tag using Illumina NextSeq 500 (Glasgow Polyomics). Then, 1535±383 cells (Mean±S.E.M.) per tag were successfully sequenced. For analysis, the sequencing-reads were processed with BD Genomics Rhapsody Analysis Pipeline CWL v.1.0 using the Seven Bridges platform. The Seurat package (3.1.5) in R was used to create an object from the RSEC_MolsPerCell.csv output file for each sample tag
Following the standard analysis protocol (19), normalization and data scaling were performed, followed by principal component analysis of the top 2000 variable genes (RunPCA). A UMAP plot (RunUMAP) was generated from the first 30 principal components. The same PCs were used to determine $k$-nearest neighbors for each cell during SNN graph construction, before clustering at a chosen resolution of 1.2 (FindNeighbors, FindClusters). Differential expression was performed (FindAllMarkers, test.use=MAST) to identify cluster-markers and variable genes between SPP1 doses. Genes are considered significantly DE if $P<0.05$ with Bonferroni correction (Table S3). To visualize heatmaps the pheatmap (1.0.12) package was adapted. The normalized expression values were used to perform pseudobulk expression analysis of each sample (AverageExpression). Raw data are accessible at EMBL-EBI (E-MTAB-10430). Using Seurat (AddModuleScore), we scored the clusters of BDRhapsody COVID-19 whole-blood dataset (22) based on average expression of 37 marker-genes that characterized the SPP1-driven CD274\textsuperscript{pos} neutrophil phenotype. A positive score suggests that this module of genes is expressed more highly in a particular cell than expected across the general population, described in more detail in (72). The module score was used to create a new assay for visualization and was illustrated using UMAP expression and pseudobulk sample expression heatmap.

**Study Approval.** The study was approved by the Committee of the Fondazione Policlinico Universitario A. Gemelli IRCCS - Università Cattolica del Sacro Cuore (prot. N. 12401/20 and prot. N.0024184/20) and University of Glasgow, MVLS College Ethical Committee (prot. N.2012073).

**Statistical analysis:** The scRNAseq data comparison is described in *Data Comparison* and *Pathway Analysis.* Statistical analysis of patient clinical, demographic and plasma cytokine
levels were performed using SPSS V.20.0 (SPSS. Chicago, Illinois, USA) and GraphPad Prism software version 9.0.0 (San Diego, California, USA). Categorical and quantitative variables were described as frequencies, percentage, median with the interquartile range or Mean ± Standard Error of the Mean (S.E.M). Data on demographic and clinical parameters were compared between patients by Mann-Whitney u-test or χ² test, as appropriate. Cytokine concentrations between multiple patient groups and healthy donors were compared using One-way ANOVA (or Kruskal-Wallis) with Dunn’s or Tukey’s correction for multiple comparisons, or Two-sided Mann-Whitney for two groups. Cytokine categories at cut-off level 108 ng/ml, 59 ng/ml, 24 ng/ml and 15 µg/ml for SPP1, S100A12, GAS6 and PROS1 respectively were selected based on the median levels in all COVID-19 patients. Linear correlations between cytokines and continuous parameters were performed using the Spearman’s-rank test. Kaplan-Meyer analysis estimated the probability of “no need to be transferred to ICU during hospitalization” for COVID-19 patients with acute pneumonia based on the previously determined cut-off plasma values of SPP1, S100A12, GAS6 and PROS1 at the time of hospital admission. P value <0.05 was considered as statistically significant. Exact p-values are provided on graphs.

**Author Contribution.** M.K.-S. oversaw the project, with S.A. interpreted clinical results and with L.M interpreted computational and experimental results. M.K-S, L.M and S.A wrote the manuscript with feedback from all authors. L.M. analyzed and integrated 10X Genomics scRNAseq data and interpreted all computational data. S.A. and B.T. evaluated cytokine levels in patients’ samples and performed statistical analysis with clinical data integration. L.M and M.K.-S performed SPP1 stimulation of whole blood and subsequent scRNAseq and data analysis. D.S. performed BD_Rhapsody sequencing library preparation and pathway analysis on scRNAseq data. L.M analyzed BD_Rhapsody SPP1 data, and with D.S mapped to COVID-
19 data sets. **T.D.O.** and **M.K.-S.** supervised all computational analysis in the study. **B. T., M.S. and Ma.S.**, organized COVID-19 samples collection and handling and performed ELISA with ELLA technology. **S.P., A.P., L.P., A.L.F., A.M.M., A.C., R.M., M.F., M.A., F.L., F.F., A.G. and E.G.**, provided clinical management of patient cohorts during acute phase hospitalization and at the post-acute outpatient clinic and provided longitudinal clinical data. **E.S. and V.A.** provided lung tissues from cadavers and clinical data. **A.E.**, assisted with ELISAs and preparation of tissue samples. **I.B.M.** and **E.G.**, assisted with running the project. **E.G.** conceived the COVID-19 translational research activity for the “Immunology Core facility” of the Fondazione Policlinico Universitario A. Gemelli IRCCS. Our study is a collaboration between two centres. The first two, and last two authors contributed equally. With agreement between all authors, and reflection on each author’s contribution, we alternated the first and last authors as listed to accurately reflect the collaborative effort of the two centres.

**Acknowledgments.** We thank all the healthcare workers who are fighting SARS-CoV-2 in the front line and all patients who agreed to take part in the study. This work was supported by UKRI-MRC (grant no. MR/V036998/1 to S.A., M.K-S, E.G., T.D.O., C.M. and I.B.M), the Research into Inflammatory Arthritis Centre Versus Arthritis UK (grant no. 22072) to L.M., T.D.O., I.B.M. and M.K-S, Versus Arthritis UK project grant (no.22253) to M.K-S and A.E, Linea D1 (Università Cattolica del Sacro Cuore, no. R4124500654) to S.A., Ricerca Finalizzata Ministero della Salute (no. GR-2018-12366992) to S.A., Versus Arthritis UK Program grant (no. 21802) to M.K-S; Versus Arthritis UK project grant (no. 22272) to M.K-S.; Wellcome Trust (no. 204820/Z/16/Z) to T.D.O. and M.K-S., and BPF-Medical Research Trust to C.M.


Figure 1

A. Integrated

B. BALF clusters

C. Distance between clusters

D. Synovial Tissue

E. CD48^{pos} S100A12^{pos}

F. Balloon热

G. FCN1^{pos} SPP1^{pos}

H. Healthy

I. BALF

J. Mild COVID-19

K. Severe COVID-19

L. Healthy Synovium

M. Active RA Synovium

N. Remission RA Synovium

O. S100A12

P. SPP1

Q. Normalized Expression

R. Normalized Expression
Figure 1. COVID-19 BALF FCN1pos and FCN1posSPP1pos macrophage clusters are transcriptionally related to rheumatoid arthritis synovial CD48highS100A12pos and CD48posSPP1pos macrophage clusters. (A) UMAP of integrated an equal number of synovial tissue and BALF macrophages (32,000 from each data set). (B) coloured according to their ST and BALF cluster identification. (C) Dendrogram of hierarchical clustering analysis of integrated pseudo-bulk gene expression (average expression in each cluster) of ST and BALF clusters. (D) Venn diagram illustrating the numbers of unique and shared marker-genes of ST CD48highS100A12pos and BALF FCN1pos clusters. p-value <0.05 after Bonferroni correction for multiple comparisons. (E) Heatmap illustrating scaled, pseudo-bulk gene expression of shared upregulated marker-genes (highlighted in D) by BALF and ST clusters. (F) Venn diagram illustrating numbers of unique and shared marker-genes of ST CD48posSPP1pos and BALF FCN1posSPP1pos clusters generated as in (E). (G) Heatmap illustrating scaled, pseudo-bulk gene expression of shared upregulated marker-genes (highlighted in F) by ST and BALF clusters. (H) Split UMAP plots comparing S100A12 and SPP1 expression in BALF and ST macrophage clusters across different conditions. Intensity of purple indicates expression level. HC, healthy control. (I) Dot plots illustrating normalized (Mean±S.E.M) expression values of S100A12 and SPP1 per cell across all immune and epithelial cell clusters in severe COVID-19 BALF (n=6). Framed populations showed the highest expression of these markers.
Figure 2. COVID-19 BALF FABP4$^{\text{pos}}$ and RA synovial TREM2$^{\text{pos}}$ macrophages share transcriptomic profiles and regulatory TAM receptor pathways. (A) Venn diagram illustrating numbers of unique and shared marker-genes of ST TREM2$^{\text{high}}$ and BALF FABP4$^{\text{pos}}$ macrophage clusters as described in Figure 1. Marker-genes were identified prior to integration of datasets (Liao et al. 2020, ref 24 and Alivernini et al. 2020, ref 19) and were calculated using MAST, setting minimum percentage of cells in clusters expressing each marker to 40%. Genes considered differentially expressed at p<0.05 after Bonferroni correction. (B) Heatmap illustrating scaled, pseudo-bulk expression of shared upregulated marker-genes from ST and BALF clusters indicated in (A). (C) Split UMAP plots comparing BALF macrophage clusters in health, and in mild and severe COVID-19 illustrating changes in expression of the TAM receptors AXL and MerTK, with their respective preferred ligands GAS6 and PROS1. Intensity of purple indicates expression level. (D) Heatmap illustrating scaled, pseudo-bulk expression of TAM receptors and associated ligands by each BALF cluster, across patient groups. TAM receptors and their ligands were significantly differentially expressed in severe COVID-19 versus healthy tissues (p-value≤0.005), with Bonferroni correction for multiple comparison, as confirmed by MAST.
Figure 3. High levels of SPP1 and S100A12 are associated with a severe COVID-19 disease trajectory. (A) Patients and Healthy donors: n=121 patients with acute pneumonia (n=29 community acquired SARS-CoV-2 negative pneumonia, n=29 mild/moderate COVID-19, n=63 severe COVID-19), convalescent COVID-19 (n=41) and healthy controls (n=10). Representative images of lung CT-scans. (B) Plasma levels of SPP1, S100A12, GAS6 and PROS1 in groups as in (A). (C) Spearman’s rank correlations between SPP1, S100A12, GAS6 and PROS1 plasma levels in patients with acute COVID-19 pneumonia (n=92) with demographic and clinical parameters. Each box displays the r-value, and * indicates statistical significance p< 0.05. (D) Plasma levels of SPP1, S100A12, GAS6 and PROS1 in patients with acute COVID-19 pneumonia (n=92) stratified based on lung functions measured by PaO2/FiO2 at the time of hospital admission. Severe respiratory failure was defined by PaO2/FiO2 ≤200. (E) % of acute COVID-19 pneumonia patients (n=92) with PaO2/FiO2 ≤200 based on high plasma levels of SPP1 (≥108 ng/ml), S100A12 (≥59 ng/ml), GAS6 (≥24 ng/ml), and PROS1 (≥15 µg/ml). (F) COVID-19 patients plasma levels of SPP1, S100A12, GAS6 and PROS1 at the time of hospital admission (n=92) stratified based on patients’ subsequent need to be transferred to ICU. (G) % of patients with acute COVID-19 pneumonia (n=92) transferred to ICU during the hospitalization based on having high levels of SPP1 (≥108 ng/ml), S100A12 (≥59 ng/ml), GAS6 (≥24 ng/ml) and PROS1 (≥15 µg/ml) at the time of hospital admission. (B, D, F). Data are presented as violin plots with median and interquartile range. *One-way ANOVA (Kruskal-Wallis test) with Dunn’s correction for multiple comparisons if more than two groups were compared or two-sided Mann-Whitney was used when two groups were compared. (H) Kaplan-Meier analysis of the rate of transfer of COVID-19 patients to ICU based on their cut-off values for SPP1, S100A12, GAS6 and PROS1 at the time of hospital admission.
**Figure 4. Increased SPP1 and S100A12 levels persist in post-COVID-19 phase.** (A) Representative images of lung CT-scans (transversal and sagittal view, respectively) of a COVID-19 patient taken during acute pneumonia and during convalescence (at 164 days post hospital discharge. (B) Plasma levels of SPP1, S100A12, GAS6 and PROS1 in paired plasma samples from COVID-19 patients at the time of acute pneumonia and at the convalescent phase (n=26). (C) Plasma levels of SPP1, S100A12, GAS6 and PROS1 in convalescent COVID-19 patients (n=41) stratified based on the severity of prior acute pneumonia and compared to the levels of healthy donors (n=10). D) Plasma levels of IL-6 in acute pneumonias and post-COVID-19. (E) SPP1, S100A12, GAS6 and PROS1 in convalescent COVID-19 patients (n=41) stratified based on suffering (n=36) or not (n=5) at least one of the symptoms (fatigue, musculoskeletal or respiratory symptoms). (B) Data is presented as before and after plot. Wilcoxon test on paired samples was used and exact p-values are provided on the graphs. (C-E) Data are presented as violin plots with median and interquartile range. *One-way ANOVA with correction for multiple comparisons if more than two groups were compared or two-sided Mann-Whitney was used when two groups were compared. Exact p-values are provided on the graphs.
Figure 5. SPP1 protein is expressed in COVID-19 lung by macrophages but not healthy alveolar macrophages (A) Representative Immunofluorescence staining of normal lung (n=3) showing SPP1 negative alveolar macrophage (CD68pos). (B) Representative immunofluorescence staining of COVID-19 lung (n=2) showing SPP1 positive macrophages (CD68posSPP1pos) in alveoli. Solid white arrows indicate macrophages double positive for SPP1 and CD68; hollow arrows indicate CD68 positive and SPP1 negative macrophages.
Figure 6. SPP1 stimulation drives pro-inflammatory CD14<sup>pos</sup> monocyte and CD274(PD-L1)<sup>pos</sup> neutrophil phenotypes. Whole blood cells from healthy donors (n=3) were stimulated with SPP1 (50 or 200 ng/ml) or were unstimulated, for 16h. (A) UMAP of 13,580 integrated control and SPP1-stimulated blood cells coloured by cluster identity. (B) Stacked bar plot illustrating cluster proportion of total whole blood cells per condition/dose of SPP1. * P < 0.05, one-way ANOVA with correction for multiple comparisons. (C) UMAPs illustrating change in expression of MME (CD10) and CD274 (PD-L1) by neutrophil clusters following SPP1 stimulation. (D) Heatmap showing expression of the top 20 marker-genes of each neutrophil phenotype, illustrated as average expression of each gene per sample per condition. Adjusted P < 0.05, MAST with Bonferroni correction. (E) Average of SPP1-driven CD274<sup>pos</sup> neutrophil gene-module score by neutrophil clusters of COVID-19 whole blood cell dataset (Schulte-Schrepping et al. 2020; ref 22). (F) UMAP expression of the SPP1-driven CD274<sup>pos</sup> neutrophil gene module in COVID-19 whole blood dataset. (G) Heatmap showing expression of genes significantly differentially expressed in CD14<sup>pos</sup> monocytes comparing SPP1 (200 ng/ml) stimulation with unstimulated control, illustrated as average expression of gene per condition. Adjusted P < 0.05, MAST with Bonferroni correction. (-), unstimulated control; red and orange arrows in (D and G), genes regulated by 200 ng/ml only, or by 200 and 50 ng/ml SPP1, respectively.