Longitudinal analysis of sarcoidosis blood transcriptomic signatures and disease outcomes

Robert Su[1], Michael M. Li[2], Nirav R. Bhakta[2], Owen D. Solberg[2], Eli P.B. Darnell[2], Joris Ramstein[2], Suresh Garudadri[2], Melissa Ho[2], Prescott G. Woodruff[2,3] and Laura L. Koth[2]

Affiliations: [1]Division of Rheumatology, University of California, San Francisco, CA, USA. [2]Division of Pulmonary, Critical Care, Sleep and Allergy, Dept of Medicine, University of California, San Francisco, CA, USA. [3]Cardiovascular Research Institute, University of California, San Francisco, CA, USA.

Correspondence: Laura L. Koth, Division of Pulmonary, Critical Care, Sleep and Allergy, Dept of Medicine, University of California, San Francisco, Box 0111, 505 Parnassus Avenue, San Francisco, CA 94143, USA. E-mail: laura.koth@ucsf.edu

ABSTRACT   Previously, we demonstrated concordance in differentially expressed genes in sarcoidosis blood and lung, implicating shared dysfunction of specific immune pathways. In the present study, we hypothesised that expression levels of candidate genes in sarcoidosis blood could predict and track with disease outcomes longitudinally.

We applied Ingenuity Pathway Analysis to a cross-sectional derivation microarray dataset (n=38) to identify canonical pathways and candidate genes associated with sarcoidosis. In a separate longitudinal sarcoidosis cohort (n=103), we serially measured 48 candidate gene transcripts, and assessed their relation to disease chronicity and severity.

In the cross-sectional derivation study, pathway analysis showed upregulation of genes related to interferon signalling and the role of pattern recognition receptors, and downregulation of T-cell receptor (TCR) signalling pathways in sarcoidosis. In the longitudinal cohort, factor analysis confirmed coregulation of genes marking these pathways and identified CXCL9 as an additional candidate pathway. CXCL9 and TCR factors discriminated between chronic *versus* nonprogressive disease, and CXCL9 predicted disease outcomes longitudinally. Interferon factor was similarly increased in both disease phenotypes. Factors associated with lung function decline included decreased TCR factor and increased CXCL9.

These findings demonstrate blood transcriptomic signatures reflecting TCR signalling and CXCL9 predict sarcoidosis chronicity and correlate with disease severity longitudinally.

@ERSpublications
Blood gene transcript measurements predict sarcoidosis chronicity and severity longitudinally http://ow.ly/zFGnP

Introduction

Sarcoidosis is a heterogeneous inflammatory disease characterised by noncaseating granulomas that can affect multiple organ systems, most commonly the lung. Although many affected individuals have a self-limited course, approximately one-third of patients develop chronic, progressive disease of the lungs that can result in fibrosis. The pathogenesis of sarcoidosis remains to be elucidated although inflammatory cytokines have been implicated, including interferon (IFN)-γ and tumour necrosis factor (TNF), which are strongly associated with localised granulomatous inflammation [1–3]. When treatment is indicated, corticosteroids remain the mainstay of therapy; however, TNF inhibitors may have benefit in refractory disease [4–6].

The development of noninvasive markers that predict clinical course or correlate with sarcoidosis disease activity remains an important unmet clinical need. Previously, in a cross-sectional study [7], we applied a genome-wide approach and identified whole-blood gene expression markers of sarcoidosis that implicated dysfunction of a set of immune pathways, including IFN signalling. Using available gene microarray expression datasets, we showed concordance of these sarcoidosis-associated gene expression markers in both blood and lung. However, our prior study, and most other studies of candidate gene expression markers in sarcoidosis, lacked longitudinal serial measurements and prospective determination of disease course [8].

Based on our prior cross-sectional study [7], we hypothesised that expression levels of specific sets of genes in sarcoidosis blood would track with disease outcomes longitudinally. To test this hypothesis, we measured serial blood quantitative (q)PCR measurements of candidate gene transcripts in a large longitudinal sarcoidosis cohort, and evaluated their relation to disease chronicity and pulmonary function.

Material and methods

Study subjects

Sarcoidosis subjects and healthy controls were recruited from the community at large, either self-referred *via* online recruitment or referred by the University of California San Francisco (UCSF) (San Francisco, CA, USA) and community providers, to participate in the UCSF Sarcoidosis Research Program. All subjects provided informed consent to participate in this study and all sarcoidosis subjects fulfilled diagnostic criteria based upon American Thoracic Society guidelines [9]. Retrospective medical records for sarcoidosis subjects were collected at enrolment. Blood samples and clinical data (*e.g.* changes in medications and clinical notes from the subject's healthcare provider) were then collected prospectively at the baseline study visit and at follow-up intervals of 6 months. Among sarcoidosis subjects, pulmonary function testing (PFT) results were used if performed within 6 months of study visits; otherwise, spirometry was performed at each visit (HDpft 1000; nSpire, Longmont, CO, USA) [10].

Sarcoidosis subjects were classified as nonprogressive, chronic or uncertain based on the schema depicted in figure 1. In prior studies, the course of sarcoidosis disease has been characterised as chronic if disease activity persisted beyond 2 years [11–16]. As such, the 2-year threshold was incorporated into our classification schema. Use of this classification schema required the use of both retrospective and prospective clinical data to categorise patients. Nonprogressive subjects had stable PFT and no history of flare beyond 2 years of diagnosis. In contrast, chronic patients had either progressive decline in PFT during the study period or a history of disease flare beyond 2 years of diagnosis. Progressive declines in PFT were defined as a decrease in forced vital capacity (FVC) $\geqslant$10% predicted or diffusing capacity of the lung for carbon monoxide (DLCO) $\geqslant$15% predicted during the study period as per published thresholds [17]. A flare was defined as a recurrence of extrapulmonary disease or pulmonary symptoms severe enough to warrant escalation of treatment to $\geqslant$20 mg of corticosteroids daily [18] or initiation of a TNF inhibitor.

Selection of genes for longitudinal study

Using Ingenuity canonical pathway analyses (Ingenuity Systems, Redwood City, CA, USA), we identified differentially expressed genes belonging to canonical pathways within a blood gene expression microarray dataset from our previously published study of 38 sarcoidosis subjects (Gene Expression Omnibus (GEO) accession number GSE19314) [7]. Based on the results of the pathway analysis and, in part, by manual curation of genes implicated in sarcoidosis or granulomatous inflammation, 48 genes of interest were selected for Fluidigm qPCR analysis (Fluidigm San Francisco, CA, USA) (table 1). Primer and probe sequences are listed in online supplementary table E1.

RNA processing and qPCR

Blood samples were collected with PAXgene Blood RNA tubes (PreAnalytiX, Hombrechtikon, Switzerland) and RNA extracted with the automated instrument Qiacube (Qiagen, Valencia, CA, USA). RNA

```
                    ┌─────────────────────────────────────────┐
                    │ Fulfilled ATS diagnostic criteria for     │
                    │ sarcoidosis                               │
                    └─────────────────────────────────────────┘
```

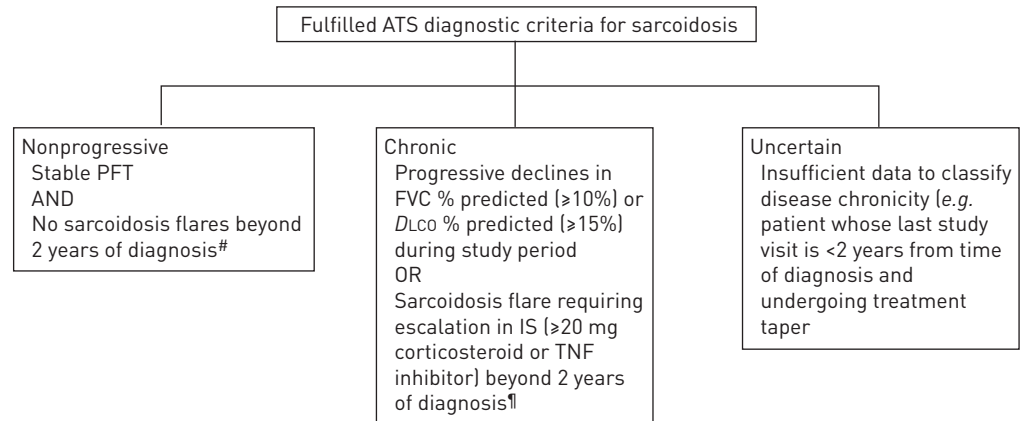| Nonprogressive Stable PFT AND No sarcoidosis flares beyond 2 years of diagnosis[#]	Chronic Progressive declines in FVC % predicted (≥10%) or D_{LCO} % predicted (≥15%) during study period OR Sarcoidosis flare requiring escalation in IS (≥20 mg corticosteroid or TNF inhibitor) beyond 2 years of diagnosis[¶]	Uncertain Insufficient data to classify disease chronicity (e.g. patient whose last study visit is <2 years from time of diagnosis and undergoing treatment taper)

FIGURE 1 Schema used for classification of disease course among recruited sarcoidosis subjects. ATS: American Thoracic Society; PFT: pulmonary function testing; FVC: forced vital capacity; D_{LCO}: diffusing capacity of the lung for carbon monoxide; IS: immunosuppression; TNF: tumour necrosis factor. [#]: 34 out of 45 nonprogressive subjects had never received IS or had IS tapered off within 2 years of diagnosis and 11 nonprogressive subjects were maintained on IS from diagnosis; [¶]: 16 out of 46 chronic subjects had progressive declines in pulmonary function testing and 30 chronic subjects developed a sarcoidosis flare requiring escalation of IS (organ manifestations of flare: nine pulmonary, seven cutaneous, five neurological, four ocular, three cardiac and two hypercalcaemia).

concentration and RNA integrity number were calculated with Agilent RNA 6000 Nano Kit and Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). High-throughput qPCR was performed using Fluidigm 96.96 Dynamic Arrays IFC with each gene run in duplicate. Sarcoidosis expression values are expressed as $\log_2$ cycle threshold (Ct) values relative to mean healthy control Ct values. See the online supplementary material for details of expression data processing and normalisation.

Statistical analysis

Within Ingenuity, Fisher's exact test was used to determine the probability that the association between the genes in the sarcoidosis derivation cross-sectional microarray data set and the canonical pathway was explained by chance alone. Other statistical analyses were performed using Stata version 12 (StataCorp, College Station, TX, USA). Exploratory factor analysis was performed on sarcoidosis qPCR results. Exploratory factor analysis identifies clusters (*i.e.* factors) of intercorrelated variables (*i.e.* gene transcripts) without making any prior assumption on the potential relationships between measured variables.

Unpaired two-tailed t-tests were used for comparisons of continuous data and Fisher's exact test for categorical data. Logistic regression modelling was used to generate predicted probabilities of chronic *versus* nonprogressive disease. Receiver operating characteristic (ROC) curve analyses of the predicted probabilities and a time-to-event analysis were performed to assess the predictive utility of the factors. Mixed-effects modelling was performed with the longitudinal qPCR data using unstructured covariance for the random effects (*i.e.* subject identification number and visit number or time from diagnosis). Age, sex, lymphocyte count, time from diagnosis and visit numbers were included in mixed-effects models. Models

TABLE 1 Genes selected for quantitative PCR analysis in longitudinal cohort

Rationale for inclusion in study	Genes[#] n	Gene names
Implicated in sarcoidosis inflammation	11	*ICAM1, IFNG, IL15, IL17, IL6, CXCL9, CXCL10, TNFA, HGF, IL13, TNFSF13B*
Implicated in granulomatous inflammation	7	*IL15RA, IL18R1, IL4, STAT6, B7H1, IL25, SAA4*
Invoked in Ingenuity Pathway Analysis	16	*IRF1, ITK, JAK2, TYK2, LEF1, MYD88, TLR2, TLR4, TLR8, CTL4A, CD28, TAP1, IRF7, IRF9, STAT1, STAT2*
Differentially expressed in sarcoidosis derivation cohort but unknown role in disease	11	*LRP1, PLAUR, TBXAS1, GBP1, GBP2, SKIL, CETP, USP15, PIK3AP1, TNFAIP2, LPCAT2*
Housekeeping genes	3	*RPL13A, B2M, ACTB*

[#]: 48 genes in total.

also controlled for immunosuppression use when feasible (*i.e.* when not collinear with other covariates). p-values less than 0.05 were considered significant.

Results

Patient characteristics

Neither sex, race nor ethnicity was different between groups (table 2). The mean age of the sarcoidosis group was older than that of the healthy group and, therefore, all analyses presented in this study were adjusted for age.

At baseline, there were no differences in age, sex, race or Scadding stage in the nonprogressive, as compared to the chronic, sarcoidosis subjects (table 3). A higher proportion of chronic sarcoidosis subjects were on systemic immunosuppression, and had significantly lower FVC % predicted, FEV1 % predicted, and DLCO % predicted compared with nonprogressive subjects.

Pathway analyses of sarcoidosis blood gene expression implicates dysregulation of canonical immune pathways

We used a publically available microarray dataset that we previously generated using blood samples from an unrelated cohort of 38 sarcoidosis subjects (GEO accession number GSE19314) [7] to perform Ingenuity Pathway Analysis. This identified upregulation of genes related to IFN signalling (p=0.0024) (fig. E1) pathways and the role of pattern recognition receptors in recognition of bacteria and viruses (p=0.0006) (fig. E2), and downregulation of genes related to the T-cell receptor (TCR) signalling pathway (p=0.019) (fig. E3) in sarcoidosis blood, compared with healthy controls. Primers were selected for genes contained within these canonical pathways and, additionally, by manual curation as mentioned earlier (table 1).

Factor analysis demonstrates distinct clusters of intercorrelating genes

Factor analysis using the qPCR data derived from the longitudinal sarcoidosis cohort consisting of 103 subjects yielded eight factors, of which three (factors 1, 2, and 3; supplemental table E2) corresponded to the same three canonical biological pathways identified by Ingenuity analysis as described earlier. Factor 1 consisted of several highly intercorrelated genes associated with the IFN signalling pathway, including *STAT1*, *STAT2*, *IRF1*, *IRF7* and *IRF9*. Factor 1 also included interferon-γ-inducible genes such as *TAP1* and *GBP1*. Factor 2 genes were associated with role of pattern recognition receptors, including *TLR2*, *TLR4* and *TLR8*. Factor 3 consisted of genes involved in TCR signalling: *CD28*, *CTL4A*, *ITK* and *LEF1*. Although *LEF1* was not a member of the Ingenuity-defined canonical TCR signalling pathway, *LEF1* is an integral mediator of TCR signalling as it encodes a protein that binds the TCR-α enhancer site and confers maximal TCR-α enhancer activity. Unexpectedly, the IFN-inducible chemokine *CXCL9* gene (factor 7) was not intercorrelated with the other IFN-related genes in factor 1. We had particular interest in *CXCL9* (factor 7), as *CXCL9* is upregulated in gene network analyses of sarcoidosis organ tissues [19, 20], and we and others have previously shown that serum protein levels of this chemokine are upregulated in sarcoidosis [15, 21]. In summary, factor analysis of the qPCR results provided confirmation of the IFN, pattern recognition and TCR pathways in a separate replicate cohort, and additionally demonstrated a separate factor for CXCL9.

TABLE 2 Baseline cohort characteristics

	Control	Sarcoidosis	p-value
Subjects n	62	103	
Female n (%)	35 (56)	64 (62)	0.51
Age years	42.13 ± 14.55	50.88 ± 10.31	<0.01
Follow-up period months		31 ± 11.4	
Race			
African-American	6 (10)	12 (12)	0.80
White	42 (68)	77 (75)	0.37
Other	14 (22)	14 (13)	0.14
Ethnicity			
Latino	8 (13)	8 (8)	0.29
Lymphocyte count $\times 10^9$ **cells·L^{-1}**	1.81 ± 0.50	1.47 ± 1.79	0.16

Data are presented as n (%) or mean $\pm$ SD, unless otherwise stated.

TABLE 3 Sarcoidosis baseline characteristics by clinical course

	Nonprogressive	Chronic	p-value
Subjects	45	46	
Age years	49.2±10.2	52.8±10.6	0.11
Range	(33–74)	(30–76)	
Females	26 (58)	31 (67)	0.39
Time from diagnosis to enrolment years	2.87±3.9	8.01±9.06	<0.01
Race			
African-American	5	7	0.76
White	35	31	0.35
Other	5	8	0.55
Ethnicity			
Latino	2	5	0.43
Scadding stage 0/I/II/III/IV	4/6/20/6/9	5/2/22/5/12	0.64
Lymphocyte count $\times 10^9$ cells·L^{-1}	1.24±0.5	1.32±1.1	0.64
Systemic immunosuppression at enrolment	14 (31)	30 (65)	<0.01
Pulmonary function testing % predicted			
FVC	102.4±12.5	91.2±16.7	<0.01
FEV_1	97.8±16.8	86.7±18.8	<0.01
D_{LCO}	89.8±18.1[#]	73.4±20.2[¶]	<0.01

Data are presented as n, mean±SD or n (%), unless otherwise stated. FVC: forced vital capacity; FEV_1: forced expiratory volume in 1 s; D_{LCO}: diffusing capacity of the lung for carbon monoxide. [#]: n=31; [¶]: n=26.

To facilitate data reduction for factors weighted by more than one gene, a three-gene mean was calculated and applied to each subject sample to represent the mean expression values of positively intercorrelating genes within factors 1, 2 and 3. This three-gene approach has been previously applied in our genomic studies of respiratory diseases such as asthma [22–24]. Within these factors, we chose the three genes with the greatest factor loadings (*i.e.* higher loading denotes greater contribution by that gene to the factor). We calculated a three-gene mean, herein called the "IFN factor" score, which represented the mean of standardised (*i.e.* mean-centered and scaled) expression values for *STAT1*, *STAT2* and *GBP1*. Similarly, for each sample, a "pattern recognition factor" score was calculated using the mean of standardised expression values for *TLR2*, *TLR4* and *TLR8*. The mean of *CD28*, *ITK* and *LEF1* standardised expression values was taken together as the "TCR signalling factor" score.

Factors distinguish chronic from nonprogressive disease

Using blood collected at enrolment, logistic regression modelling of IFN, TCR and CXCL9 factors identified *CXCL9* (p<0.001) and *TCR* (p=0.011) as significant predictors of disease course (*i.e.* chronic *versus* nonprogressive). We generated predicted probabilities of chronic disease from these logistic regression models. Lymphopenia (measured as 1–lymphocyte count) was used for comparison as a reference predictor of chronic disease. Comparisons of the area under the ROC curve (AUC) demonstrated that the predicted probabilities based on gene factors at enrolment significantly outperformed lymphopenia as a predictor of chronic disease (AUC 0.80 *versus* 0.59, respectively; p=0.001) (fig. 2). The performance of CXCL9, TCR and IFN factors as independent measures of predicting chronic disease is shown in figure E4, and confirms that IFN, in itself, is a poor predictor of chronic disease.

To determine whether these factors demonstrated good predictive value among subjects early in their disease course, we repeated these analyses using only sarcoidosis subjects who had enrolled in the study within 2 years of their diagnosis. Again, we found that CXCL9 and TCR factor-based predicted probabilities identified chronic disease among subjects with early disease (AUC 0.84) (fig. E5). In sensitivity analysis, we also examined the predictive value of CXLC9 and TCR factors in sarcoidosis subjects not on any immunosuppression at enrolment, and found that the results were unchanged (AUC 0.84) (fig. E6).

High CXCL9 factor predicts disease course

All sarcoidosis subjects were classified as CXCL9 factor high or low at enrolment. The CXCL9 factor threshold was set at 0.4 ($\log_2$ relative expression value) based on ROC curve analysis of CXCL9 described above. We performed a time-to-event analysis, wherein an event was defined as: 1) a decrease in FVC ⩾10% predicted or D_{LCO} ⩾15% predicted during the follow-up period; or 2) a sarcoidosis flare where treatment is escalated to ⩾20 mg of corticosteroids daily [18] or there is an initiation of a disease
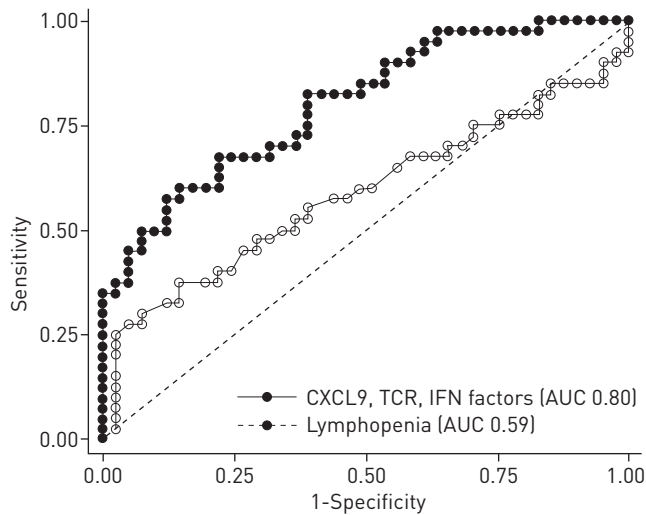
FIGURE 2 Receiver operating characteristic curve analyses of discrimination between chronic and nonprogressive disease (n=91). Logistic regression-based prediction probabilities using CXCL9, T-cell receptor (TCR) and interferon (IFN) factors outperform lymphopenia, with the difference in area under curve (AUC) being associated with p=0.001.

modifying antirheumatic agent or TNF inhibitor after enrolment. As shown in figure 3, CXCL9-high sarcoidosis subjects developed a significant decline in pulmonary function testing or flare during the follow-up period more rapidly than CXCL9-low sarcoidosis subjects (p=0.009 by log-rank test).

Comparisons of IFN, TCR and CXCL9 factors longitudinally

Using mixed-effects modelling of longitudinal qPCR data from healthy and sarcoidosis subjects, both nonprogressive (p=0.001) and chronic sarcoidosis groups (p<0.001) demonstrated a higher IFN factor score compared with healthy subjects (fig. 4a). There was no significant difference in IFN factor scores between chronic and nonprogressive sarcoidosis (p=0.167). Both nonprogressive and chronic sarcoidosis groups showed lower TCR factor scores than healthy controls (p<0.001) (fig. 4b). Additionally, TCR factor was significantly lower in chronic sarcoidosis compared with nonprogressive sarcoidosis (p=0.005). CXCL9 was persistently and significantly higher in chronic compared with nonprogressive sarcoidosis and healthy subjects (p<0.001) (fig. 4c). There was no difference in CXCL9 between healthy and nonprogressive sarcoidosis subjects (p=0.826).

In subset analysis, subjects who met criteria for chronic disease based on declining PFT did not differ from those who met criteria for chronic disease based on flares in regards to IFN (p=0.31), TCR (p=0.11) and CXCL9 (p=0.95) factor scores longitudinally. This subset analysis suggests that our phenotypic group definitions yielded groups with relatively consistent patterns of gene expression.
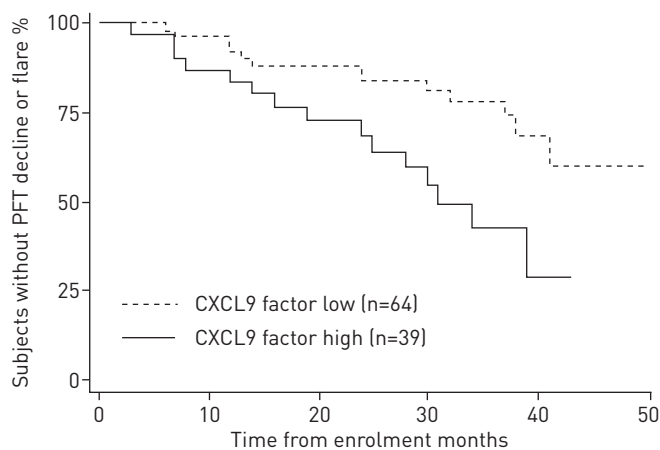

FIGURE 3 Sarcoidosis subjects with high CXCL9 factor at enrolment more rapidly develop significant declines in pulmonary function or a sarcoidosis flare requiring escalation in immunosuppression during follow-up (p=0.009 by log-rank test). PFT: pulmonary function testing.
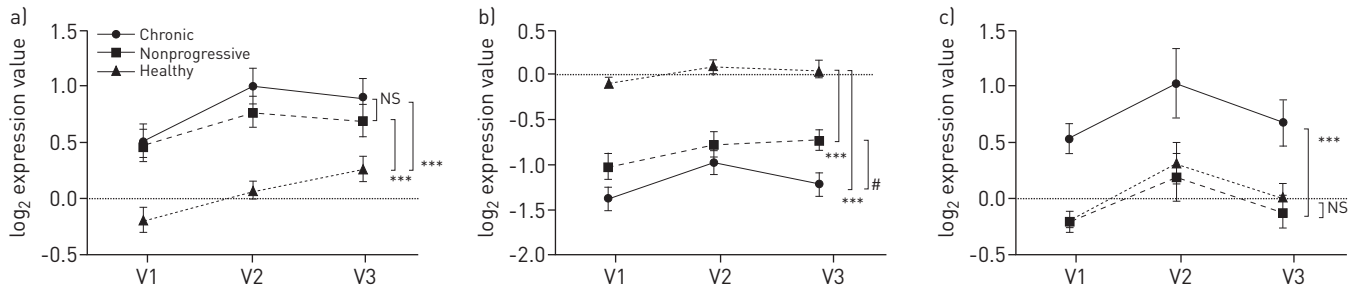
FIGURE 4 a) Longitudinal analysis demonstrates a persistent elevation of interferon factor (three-gene mean of *STAT1, STAT2* and *GBP1*) in nonprogressive and chronic sarcoidosis compared with healthy controls. b) Longitudinal analysis reveals a suppression of T-cell receptor (TCR) factor (three-gene mean of *CD28, ITK* and *LEF1*) in nonprogressive and chronic sarcoidosis compared with healthy controls, and lower TCR factor in chronic subjects compared with nonprogressive subjects. c) CXCL9 expression remains persistently high among subjects with chronic sarcoidosis compared with both healthy control and nonprogressive sarcoidosis subjects. Mean values are presented; errors bars represent the standard error of the mean. Numbers of PAXgene samples processed at time of quantitative PCR analysis (V1: baseline visit; V2: 6-month visit; V3: 12-month visit) to generate data were as follows. Nonprogressive: V1, n=45; V2, n=27; V3, n=27; chronic: V1, n=46; V2, n=31; V3, n=27; healthy: V1, n=62; V2, n=40; V3, n=36. NS: nonsignificant. #: p=0.005. ***: p<0.001.

TCR and CXCL9 factor scores correlate with pulmonary disease severity

In models that incorporate all time points, lower D_{LCO} % predicted was associated with decreased TCR factor (p=0.011, coefficient 6.54), increased CXCL9 (p=0.036, coefficient -2.90) and longer disease duration (p=0.006, coefficient -0.002). Lower FVC % predicted was also associated with lower TCR factor (p=0.016, coefficient 2.49). Furthermore, higher Scadding stage was associated with lower TCR factor (p=0.011) by ordinal logistic regression.

Effect of immunosuppression on factor scores

In mixed-effects modelling incorporating data from all sarcoidosis subjects, immunosuppression had no significant effect on IFN (p=0.12) or CXCL9 (p=0.13) factors. In secondary subset analyses, immunosuppression use was associated with lower IFN factor score among patients with nonprogressive disease (p=0.04). However, in the chronic disease group, immunosuppression use had no significant effect on IFN factor score (p=0.23). These subset analyses of the effects of immunosuppression are limited by multiple hypothesis testing. Pattern recognition factor scores strongly positively correlated with immunosuppression use in both nonprogressive and chronic groups (p<0.001 for both groups). Because we could not distinguish whether pattern recognition receptor scores related to disease activity *versus* immunosuppression use [25] with the available data, we could not pursue further analysis of the pattern recognition factor at this time.

Discussion

Previously, in a cross-sectional study of 38 subjects with sarcoidosis, we identified gene expression markers of specific inflammatory pathways that were associated with sarcoidosis, and were concordant in blood and affected lung tissue biopsies. In the current study of a separate and larger, longitudinal cohort, we demonstrate the robustness of these gene expression markers and show that they predict disease severity and course. Whereas the majority of published sarcoidosis studies are cross-sectional, our study design allowed us to investigate longitudinal measurements of biological signals over time in relation to clinical course. We identified differential expression patterns relating to three distinct gene signatures, of which TCR and CXCL9 corresponded with disease course and severity. The results support our hypothesis that measurement of whole-blood gene transcripts in sarcoidosis may afford minimally invasive means of assessing sarcoidosis disease course and could be relevant to the underlying immunopathogenesis.

We demonstrated that peripheral-blood gene transcripts for factors TCR and CXCL9 have predictive value in discriminating between chronic and nonprogressive sarcoidosis subjects. Although currently, there are no reliable means for predicting sarcoidosis disease course, some studies have suggested that lymphopenia may correlate with chronically active sarcoidosis [26–28]. In ROC curve analyses, a logistic-based prediction variable incorporating TCR and CXCL9 factor significantly outperformed lymphopenia in identifying chronic disease. This result was reproduced when the analysis was restricted to sarcoidosis subjects who enrolled early in their disease course or were not on immunosuppression. In our longitudinal cohort, sarcoidosis subjects high in CXCL9 factor at enrolment experienced a shorter time to developing a significant decline in PFT or a clinical flare during the follow-up period. Our findings suggest that increased CXCL9 factor may identify those at risk for chronic disease, and warrant further study in validation cohorts.

It is notable that our IFN signature was predominantly driven by *STAT1* and *STAT2* expression in sarcoidosis peripheral blood, as this finding supports prior studies showing that *STAT1* may play an important role in disease pathogenesis [19, 20]. Using microarray gene network analyses and immunohistochemical staining of granulomas, ROSENBAUM *et al.* [20] previously demonstrated that STAT1 signalling is strongly upregulated in peripheral blood (n=12), and lung and lymph node tissues. CROUSER *et al.* [19] similarly confirmed an increase in STAT1-related networks in sarcoidosis lung biopsies. Our study lends evidence that augmentation of STAT1 signalling is not limited to the affected granulomatous tissue but, importantly, is robustly upregulated in circulating immune cells. The finding of shared transcriptional signature of IFN pathways in the blood and diseased tissues identifies molecular targets that are not completely "compartmentalised" in the diseased organ [7, 29]. Furthermore, whereas the aforementioned studies focused on subjects at a cross-sectional level, our longitudinal study demonstrates that even nonprogressing sarcoidosis subjects, whose disease course is often considered to be benign, may exhibit persistent elevations in IFN signalling.

Our finding that both nonprogressive and chronic progressive sarcoidosis subjects share upregulation of IFN factor confirms that IFN plays an important role in the aetiopathogenesis of sarcoidosis. However, our results suggest that additional biological signals aside from IFN contribute to disease severity and progression. Some studies have suggested higher TNF signalling predicts disease progression [30, 31].

Notably, we found that peripheral *CXCL9* transcript levels, which are known to be synergistically enhanced by TNF [32], predicted disease progression in terms of pulmonary decline and relapse. CXCL9 is an IFN-γ-inducible chemokine that, upon binding to its receptor CXCR3, facilitates recruitment of T-cells to inflamed tissues in human autoimmune diseases. We and others have demonstrated elevations in CXCL9 protein levels in sarcoidosis sera [15, 21], bronchoalveolar lavage (BAL) fluid [15, 21, 33] and granulomas [34]. We further show that *CXCL9* gene expression inversely correlated with D_{LCO} % predicted, independent of time from diagnosis. These data suggest that *CXCL9* is associated with chronic sarcoidosis.

Additionally, we demonstrate that expression of genes directly involved in the TCR signalling pathway are suppressed in sarcoidosis compared with healthy controls, including *CD28*, *ITK* and *LEF1*. Our analyses were adjusted for lymphocyte counts, as lymphopenia may be seen in sarcoidosis [2]. In future studies, we propose to perform cell sorting analyses of T-cell subsets to ascertain whether these findings stemmed from peripheral changes in differential subsets of lymphocytes in sarcoidosis; however, a recent study by OSWALD-RICHTER *et al.* [35] showed that sarcoidosis peripheral and BAL sorted CD4+ T-cells exhibited decreases in the TCR-responsive genes *LCK* and *PKCQ*, both of which were also shown to be downregulated in our Ingenuity TCR signalling pathway analyses. Our results add to their study by demonstrating that the degree of *TCR* downregulation has important clinical correlations longitudinally. Specifically, decreased TCR scores were associated with chronic disease and lower FVC and $D_{L,CO}$ % predicted. Taken together, these findings demonstrate that the TCR signalling complex and its downstream mediators are downregulated in sarcoidosis, which may, in part, explain suppressed responses to TCR polyclonal stimulation and peripheral anergy in sarcoidosis patients.

There are limitations to our study. Although we recruited from a diverse, multiethnic population, our cohort largely consisted of Caucasian subjects. Subjects within the cohort were followed longitudinally but were not necessarily enrolled at the time of diagnosis. Consequently, we cannot be certain that, at the time of diagnosis, nonprogressive patients did not exhibit high levels of *CXCL9* initially; however, in a subset analysis of all sarcoidosis patients whose baseline visits were within 2 years of diagnosis, the chronic group still exhibited significantly higher *CXCL9* expression levels compared with those with nonprogressive disease. Finally, we could not apply the clinical categorisations of acute and nonacute sarcoidosis disease onset reported by PRASSE *et al.* [36], as our cohort consisted of only nonacute disease. As many published studies in sarcoidosis have used a 2-year threshold to define chronic sarcoidosis, we chose this schema as well.

In summary, we provide evidence for longitudinal measurements of gene expression patterns in sarcoidosis peripheral blood that are associated with dysregulated immune signalling pathways, and predict disease chronicity and correlate with pulmonary severity. These findings have important implications for investigations into noninvasive markers of disease activity that may inform clinical management of sarcoidosis.

References
1   Baughman RP, Culver DA, Judson MA. A concise review of pulmonary sarcoidosis. *Am J Respir Crit Care Med* 2011; 183: 573–581.
2   Chen ES, Moller DR. Sarcoidosis – scientific progress and clinical challenges. *Nat Rev Rheumatol* 2011; 7: 457–467.

3   Valeyre D, Prasse A, Nunes H, *et al.* Sarcoidosis. *Lancet* 2014; 383: 1155–1167.

4   Baughman RP, Drent M, Kavuru M, *et al.* Infliximab therapy in patients with chronic sarcoidosis and pulmonary involvement. *Am J Respir Crit Care Med* 2006; 174: 795–802.

5   Judson MA, Baughman RP, Costabel U, *et al.* Efficacy of infliximab in extrapulmonary sarcoidosis: results from a randomised trial. *Eur Respir J* 2008; 31: 1189–1196.

6   Baughman RP, Lower EE, Drent M. Inhibitors of tumor necrosis factor (TNF) in sarcoidosis: who, what, and how to use them. *Sarcoidosis Vasc Diffuse Lung Dis* 2008; 25: 76–89.

7   Koth LL, Solberg OD, Peng JC, *et al.* Sarcoidosis blood transcriptome reflects lung inflammation and overlaps with tuberculosis. *Am J Respir Crit Care Med* 2011; 184: 1153–1163.

8   Tzouvelekis A, Kouliatsis G, Anevlavis S, *et al.* Serum biomarkers in interstitial lung diseases. *Respir Res* 2005; 6: 78.

9   Statement on sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the *European Respiratory Society* (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee, February 1999. *Am J Respir Crit Care Med* 1999; 160: 736–755.

10   Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.

11   Neville E, Walker AN, James DG. Prognostic factors predicting the outcome of sarcoidosis: an analysis of 818 patients. *Q J Med* 1983; 52: 525–533.

12   Pabst S, Baumgarten G, Stremmel A, *et al.* Toll-like receptor (TLR) 4 polymorphisms are associated with a chronic course of sarcoidosis. *Clin Exp Immunol* 2006; 143: 420–426.

13   Judson MA. An approach to the treatment of pulmonary sarcoidosis with corticosteroids: the six phases of treatment. *Chest* 1999; 115: 1158–1165.

14   Baughman RP, Lower EE. A clinical approach to the use of methotrexate for sarcoidosis. *Thorax* 1999; 54: 742–746.

15   Su R, Nguyen ML, Agarwal MR, *et al.* Interferon-inducible chemokines reflect severity and progression in sarcoidosis. *Respir Res* 2013; 14: 121.

16   Grunewald J, Berlin M, Olerup O, *et al.* Lung T-helper cells expressing T-cell receptor AV2S3 associate with clinical features of pulmonary sarcoidosis. *Am J Respir Crit Care Med* 2000; 161: 814–818.

17   Mitchell DN, Scadding JG. Sarcoidosis. London, Hodder Arnold, 2012.

18   Gottlieb JE, Israel HL, Steiner RM, *et al.* Outcome in sarcoidosis. The relationship of relapse to corticosteroid therapy. *Chest* 1997; 111: 623–631.

19   Crouser ED, Culver DA, Knox KS, *et al.* Gene expression profiling identifies MMP-12 and ADAMDEC1 as potential pathogenic mediators of pulmonary sarcoidosis. *Am J Respir Crit Care Med* 2009; 179: 929–938.

20   Rosenbaum JT, Pasadhika S, Crouser ED, *et al.* Hypothesis: sarcoidosis is a STAT1-mediated disease. *Clin Immunol* 2009; 132: 174–183.

21   Takeuchi M, Oh IK, Suzuki J, *et al.* Elevated serum levels of CXCL9/monokine induced by interferon-γ and CXCL10/interferon-γ-inducible protein-10 in ocular sarcoidosis. *Invest Ophthalmol Vis Sci* 2006; 47: 1063–1068.

22   Woodruff PG, Modrek B, Choy DF, *et al.* T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* 2009; 180: 388–395.

23   Woodruff PG, Boushey HA, Dolganov GM, *et al.* Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci USA* 2007; 104: 15858–15863.

24   Bhakta NR, Solberg OD, Nguyen CP, *et al.* A qPCR-based metric of Th2 airway inflammation in asthma. *Clin Transl Allergy* 2013; 3: 24.

25   Galon J, Franchimont D, Hiroi N, *et al.* Gene profiling reveals unknown enhancing and suppressive actions of glucocorticoids on immune cells. *FASEB J* 2002; 16: 61–71.

26   Crouser ED, Lozanski G, Fox CC, *et al.* The CD4[+] lymphopenic sarcoidosis phenotype is highly responsive to anti-tumor necrosis factor-α therapy. *Chest* 2010; 137: 1432–1435.

27   Morell F, Levy G, Orriols R, *et al.* Delayed cutaneous hypersensitivity tests and lymphopenia as activity markers in sarcoidosis. *Chest* 2002; 121: 1239–1244.

28   Sweiss NJ, Salloum R, Gandhi S, *et al.* Significant CD4, CD8, and CD19 lymphopenia in peripheral blood of sarcoidosis patients correlates with severe disease manifestations. *PLoS One* 2010; 5: e9088.

29   Maertzdorf J, Weiner J 3rd., Mollenkopf HJ., et al. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. *Proc Natl Acad Sci USA* 2012; 109: 7853–7858.

30   Ziegenhagen MW, Benner UK, Zissel G, *et al.* Sarcoidosis: TNF-α release from alveolar macrophages and serum level of sIL-2R are prognostic markers. *Am J Respir Crit Care Med* 1997; 156: 1586–1592.

31   Moller DR, Chen ES. Genetic basis of remitting sarcoidosis: triumph of the trimolecular complex? *Am J Respir Cell Mol Biol* 2002; 27: 391–395.

32   Ohmori Y, Schreiber RD, Hamilton TA. Synergy between interferon-γ and tumor necrosis factor-α in transcriptional activation is mediated by cooperation between signal transducer and activator of transcription 1 and nuclear factor κB. *J Biol Chem* 1997; 272: 14899–14907.

33   Nishioka Y, Manabe K, Kishi J, *et al.* CXCL9 and 11 in patients with pulmonary sarcoidosis: a role of alveolar macrophages. *Clin Exp Immunol* 2007; 149: 317–326.

34   Agostini C, Cassatella M, Zambello R, *et al.* Involvement of the IP-10 chemokine in sarcoid granulomatous reactions. *J Immunol* 1998; 161: 6413–6420.

35   Oswald-Richter KA, Richmond BW, Braun NA, *et al.* Reversal of global CD4[+] subset dysfunction is associated with spontaneous clinical resolution of pulmonary sarcoidosis. *J Immunol* 2013; 190: 5446–5453.

36   Prasse A, Katic C, Germann M, *et al.* Phenotyping sarcoidosis from a pulmonary perspective. *Am J Respir Crit Care Med* 2008; 177: 330–336.