



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

# Novel expression signatures identified by transcriptional analysis of separated leucocyte subsets in systemic lupus erythematosus and vasculitis

### Citation for published version:

Lyons, PA, McKinney, EF, Rayner, TF, Hatton, A, Woffendin, HB, Koukoulaki, M, Freeman, TC, Jayne, DRW, Chaudhry, AN & Smith, KGC 2010, 'Novel expression signatures identified by transcriptional analysis of separated leucocyte subsets in systemic lupus erythematosus and vasculitis', *Annals of the Rheumatic Diseases*, vol. 69, no. 6, pp. 1208-1213. <https://doi.org/10.1136/ard.2009.108043>

### Digital Object Identifier (DOI):

[10.1136/ard.2009.108043](https://doi.org/10.1136/ard.2009.108043)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Publisher's PDF, also known as Version of record

### Published In:

Annals of the Rheumatic Diseases

### Publisher Rights Statement:

Available under Open Access.

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Novel expression signatures identified by transcriptional analysis of separated leucocyte subsets in systemic lupus erythematosus and vasculitis

Paul A Lyons,<sup>1</sup> Eoin F McKinney,<sup>1</sup> Tim F Rayner,<sup>1</sup> Alexander Hatton,<sup>1</sup> Hayley B Woffendin,<sup>1</sup> Maria Koukoulaki,<sup>1</sup> Thomas C Freeman,<sup>2</sup> David R W Jayne,<sup>1</sup> Afzal N Chaudhry,<sup>1</sup> Kenneth G C Smith<sup>1</sup>

► Additional data are published online only. To view these files please visit the journal online (<http://ard.bmj.com>).

<sup>1</sup>Cambridge Institute for Medical Research and Department of Medicine, Addenbrooke's Hospital, Cambridge, UK  
<sup>2</sup>Roslin Institute, University of Edinburgh, Roslin, Midlothian, UK

## Correspondence to

Dr Paul Lyons or Professor Ken Smith, Department of Medicine, Cambridge Institute for Medical Research, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0XY, UK; [pal34@cam.ac.uk](mailto:pal34@cam.ac.uk) or [kgcs2@cam.ac.uk](mailto:kgcs2@cam.ac.uk)

Accepted 25 August 2009

## ABSTRACT

**Objective** To optimise a strategy for identifying gene expression signatures differentiating systemic lupus erythematosus (SLE) and antineutrophil cytoplasmic antibody-associated vasculitis that provide insight into disease pathogenesis and identify biomarkers.

**Methods** 44 vasculitis patients, 13 SLE patients and 25 age and sex-matched controls were enrolled. CD4 and CD8 T cells, B cells, monocytes and neutrophils were isolated from each patient and, together with unseparated peripheral blood mononuclear cells (PBMC), were hybridised to spotted oligonucleotide microarrays.

**Results** Using expression data obtained from purified cells a substantial number of differentially expressed genes were identified that were not detectable in the analysis of PBMC. Analysis of purified T cells identified a SLE-associated, CD4 T-cell signature consistent with type 1 interferon signalling driving the generation and survival of tissue homing T cells and thereby contributing to disease pathogenesis. Moreover, hierarchical clustering using expression data from purified monocytes provided significantly improved discrimination between the patient groups than that obtained using PBMC data, presumably because the differentially expressed genes reflect genuine differences in processes underlying disease pathogenesis.

**Conclusion** Analysis of leucocyte subsets enabled the identification of gene signatures of both pathogenic relevance and with better disease discrimination than those identified in PBMC. This approach thus provides substantial advantages in the search for diagnostic and prognostic biomarkers in autoimmune disease.

The identification of gene expression signatures differentiating patients with systemic lupus erythematosus (SLE) and antineutrophil cytoplasmic antibody-associated small-vessel vasculitis (AAV) may well provide clinically relevant biomarkers as well as insights into disease pathogenesis.

Over the past decade the use of microarray-based expression profiling as a tool for the identification of both diagnostic and prognostic biomarkers has grown immensely. This has been most apparent in oncology, in which array-based gene signatures have been used to reveal novel clinically relevant patient subgroups,<sup>1,2</sup> predict treatment response,<sup>3</sup> identify patients at risk of metastasis,<sup>4</sup> and provide insight into the underlying molecular pathology.<sup>5</sup> Array-based diagnostic tests are starting to enter

clinical oncology practice as a direct consequence of this work.<sup>6</sup>

In autoimmune disease, unlike oncology, the choice of material to profile is not straightforward. Most studies published so far have used heterogeneous populations of blood cells, either total leucocytes or Ficoll-purified peripheral blood mononuclear cells (PBMC). As a consequence, many of the signatures that have been reported may be attributable to differences in the relative abundance of individual leucocyte populations, which has been suggested<sup>7</sup> but not yet directly addressed.

One strategy to circumvent this issue and identify cell-intrinsic expression signatures would be to profile purified populations of cells. Whereas a few studies have profiled single populations of cells,<sup>8-10</sup> no studies have looked across all major leucocyte types simultaneously, or directly compared arrays of separated cells with those of PBMC. As a consequence, many subtle, cell-type-specific changes may be missed,<sup>11</sup> and signatures that are found may be attributable to differences in the relative abundance of individual leucocyte populations. We have therefore developed a cell separation protocol that allows the isolation of CD4 and CD8 T cells, CD19 B cells, CD14 monocytes and CD16 neutrophils from individual patients.<sup>12</sup> Although this protocol utilises two sequential rounds of positive selection using magnetic beads, it produces cell samples in which gene expression levels are essentially identical to those of untouched cells.<sup>12</sup>

Whereas several studies have applied expression profiling to the study of SLE,<sup>13-17</sup> only one study has examined AAV.<sup>18</sup> That study profiled unseparated leucocytes and showed some discrimination between patients with AAV and SLE, but it did not address the issue of whether the gene signatures observed could be explained by differences in the relative size of leucocyte subsets between their patient cohorts.

In this study we confirm that controls and patients with AAV and SLE cannot be accurately discriminated at a transcriptional level using unseparated PBMC, but demonstrate for the first time that arrays of purified leucocyte subsets achieve superior differentiation of disease and control populations and allow identification of novel disease-related expression signatures. The analysis of purified leucocyte subsets demonstrated that the large majority of disease specific expression



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://ard.bmj.com/info/unlocked.dtl>

differences found were not seen in the PBMC analysis. These cell intrinsic differences are likely to provide more insight into disease pathology and offer a greater chance of detecting clinically useful biomarkers.

## PATIENTS AND METHODS

Full details of patient cohorts and experimental procedures are provided online (see supplementary file).

### Patient and control cohorts

Thirteen patients with active SLE and 44 patients with active AAV attending the lupus and vasculitis clinic at Addenbrooke's Hospital, together with 25 age and sex-matched controls, were enrolled into the study (see supplementary tables 1 and 2, available online only).

### Cell separation, RNA extraction and microarray hybridisation

PBMC, CD4 and CD8 T-cell, CD19 B-cell, CD14 monocyte and CD16 neutrophil samples were isolated as previously described.<sup>12</sup> Following extraction, RNA was labelled and hybridised to microarrays as previously described.<sup>19</sup>

### Microarray data analysis

Raw image data were imported into R for normalisation and assessment of differential expression using the Limma package in bioconductor.<sup>20</sup>

## RESULTS

### Differences observed in the transcriptomes of unseparated leucocytes from normal controls and patients with SLE or AAV largely reflect changes in their cellular composition.

To determine the extent of similarity between SLE and AAV at a transcriptional level we profiled unseparated PBMC RNA from 28 patients with AAV, 12 with SLE and 21 normal controls (arrays from four AAV, three SLE and four control samples failed quality control and were not included in subsequent analyses). Data for 15 255 of the probes on the array were present in all 61 samples and were subjected to analysis of variance (ANOVA) with correction for multiple testing<sup>21</sup> to identify differentially expressed genes among the three groups. This yielded 3549 statistically significant probes when the false discovery rate was set at 5% (supplementary table 3, available online only). Unsupervised hierarchical clustering using these data broadly separated the samples into three groups; however, they bore little resemblance to the original diagnostic classes (figure 1A). Group 1 contained 17 of 21 control samples; however, 45% of group 1 samples were from patients with AAV. Groups 2 and 3 were composed predominantly of patient samples (26 of 30), with SLE and AAV being equally distributed between them.

Post hoc analysis of the differentially expressed probes identified the genes that best differentiated SLE patients from controls (n=89, supplementary table 4, available online only), AAV patients from controls (n=61, supplementary table 5, available online only), and SLE and AAV patients (n=41, supplementary table 6, available online only). The transcripts differentiating SLE from control encode type 1 interferon-inducible proteins, immunoglobulin, or proteins expressed in immature granulocytes (figure 1B,C), consistent with previous reports.<sup>13 14</sup> The 61 transcripts differentiating AAV from control mainly encoded proteins expressed in immature granulocytes (figure 1D, see also Alcorta *et al.*).<sup>18</sup> The type 1 interferon signature was completely absent in the AAV patients (figure 1E). Consistent with this, analysis

of the 41 probes discriminating SLE and AAV patients revealed that they were primarily transcripts upregulated in SLE, which encode immunoglobulin or type 1 interferon-inducible proteins. The granulopoiesis signature is thus common to both diseases, the type 1 interferon and plasmablast signatures are restricted to SLE.

The granulopoiesis signature has previously been associated with contamination of the PBMC fraction by immature granulocytes.<sup>13</sup> This was confirmed by flow cytometry (supplementary figure 1A,B, available online only), and there was a strong correlation between the expression of genes from the granulopoiesis signature and immature granulocyte numbers (supplementary figure 1C, available online only).

Interestingly, while the interferon and granulopoiesis signatures are found in almost all SLE patients, the plasmablast signature is absent in approximately one third. Patients with the plasmablast signature had significantly elevated serum IgG (figure 1F), and a trend towards higher antinuclear antibody and double-stranded DNA autoantibody levels (supplementary figure 2A, available online only). They had significantly more active disease at enrolment (figure 1G), but no difference in flare-free survival following induction therapy (supplementary figure 2B, available online only), suggesting the plasmablast signature is associated with disease activity rather than long-term prognosis.

The probes representing the interferon, plasmablast and granulopoiesis signatures represent only a small fraction of the differentially expressed genes (vertical blue bar, figure 1A). The majority of the remaining genes could be broken down into two broad groups, those either upregulated or downregulated in controls compared with patients irrespective of diagnosis (figure 1A, vertical orange or yellow bar, respectively). Whereas these genes contained no obvious functional signatures, closer examination revealed that those upregulated in controls tended to be expressed more highly in cells of the lymphoid lineage (supplementary figure 3A, available online only), and those downregulated were expressed more highly in the myeloid lineage (supplementary figure 3B, available online only). Analysis of PBMC composition in AAV and SLE patients and controls revealed a significant relative T-cell lymphopaenia affecting both the CD4 and CD8 populations in AAV patients, but just the CD4 population in SLE patients, together with a significant relative monocytosis in both groups compared with controls (figure 1H). These differences in composition are likely to explain the majority of the expression differences observed earlier.

The ANOVA results also pointed to a CD8 T-cell signature differentiating AAV patients and controls (vertical purple bar, figure 1A), due to a number of genes primarily expressed in CD8 T cells being downregulated in AAV patients (supplementary table 7, available online only). This most likely reflects the specific relative CD8 T-cell lymphopaenia seen in AAV patients (figure 1H-I).

### Microarray analysis of purified leucocyte subsets from AAV and SLE patients identifies cell-intrinsic expression differences.

To look for cell-specific expression changes that might provide clinically useful biomarkers we carried out a microarray analysis on purified CD4 and CD8 T cells, B cells, monocytes and neutrophils isolated at the same time as the PBMC from SLE or AAV patients described earlier.

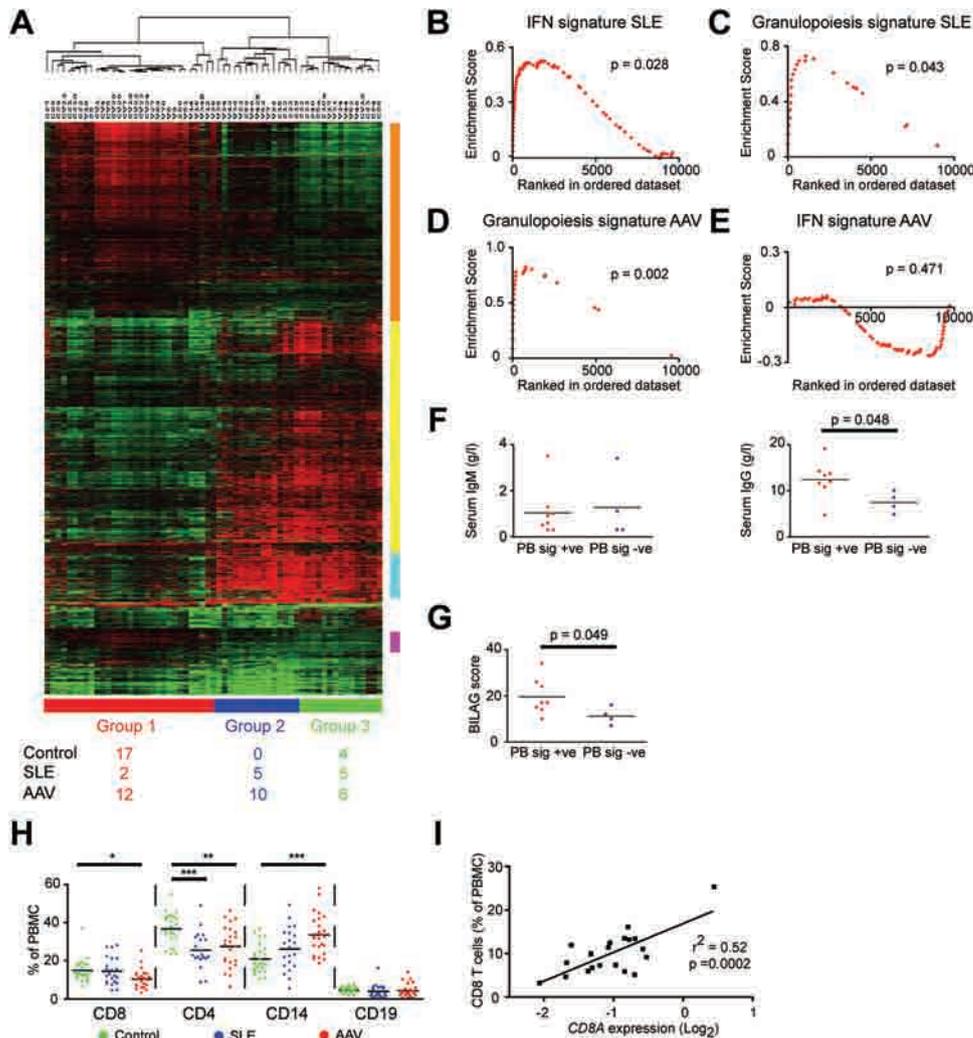
Imposition of a 1.5-fold change cut-off identified 229 probes from the PBMC data that showed significant differential expression between SLE patients and controls (figure 2A). Of

these, 112 belonged to the interferon, granulopoiesis or plasmablast signatures, with the remaining 117 falling into the unassigned category. Analysis of the purified leucocyte subsets identified many more differentially expressed genes, with a remarkably small overlap with those found in PBMC (6.5% in neutrophils to 39% in monocytes; figure 2A and specific examples in supplementary figure 4, available online only). Of the 1607 cell-specific genes differentiating disease from controls seen across the five cell types, 86% were missed on PBMC analysis. The plasmablast signature was not seen in the B-cell samples, nor the granulopoiesis signature in the neutrophils, further supporting the conclusion that these arise from differences in PBMC composition between SLE patients and controls. The only expression signature seen consistently across all cell sub-

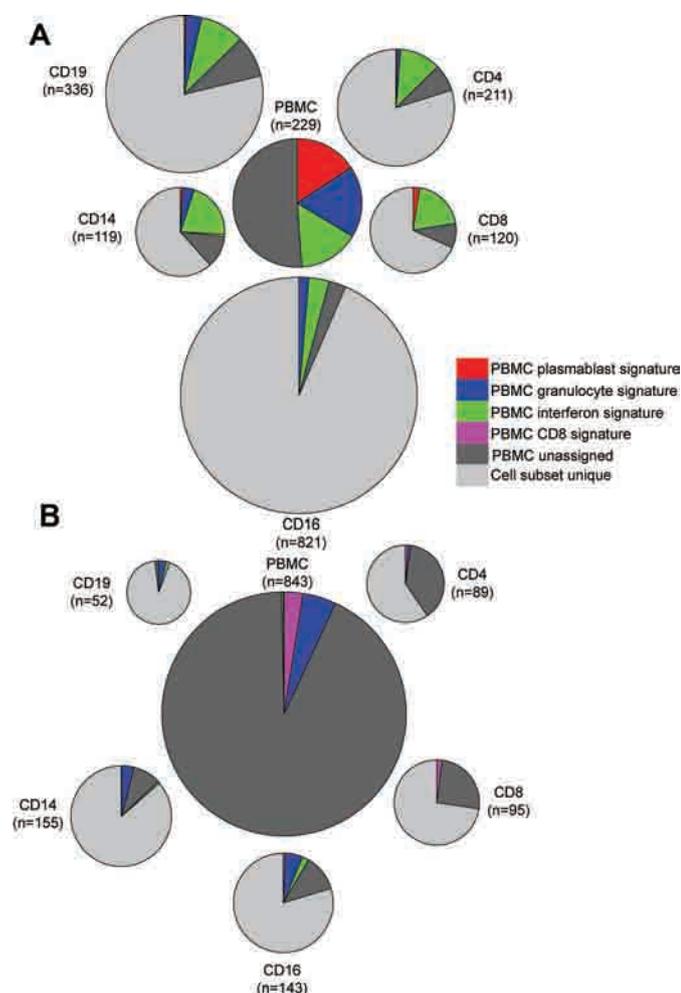
sets was the interferon signature (figure 2A and supplementary figure 4, available online only).

In a similar analysis of the AAV samples, 843 probes were found to be differentially expressed in PBMC samples between patients and controls (figure 2B), most falling into the 'unassigned' category. As in the SLE patients, there was little overlap between differentially expressed genes identified in the purified cell subsets and PBMC (specific examples in supplementary figure 4, available online only); from 8% in B cells to 43% in CD4 T cells (figure 2B).

We performed quantitative PCR, which confirmed the significant correlation between expression measured by microarray or quantitative PCR (supplementary figure 5, available online only).



**Figure 1** Differences in cellular composition largely drive gene expression signatures seen in comparisons of peripheral blood mononuclear cell (PBMC) samples. (A) Hierarchical clustering of PBMC samples from antineutrophil cytoplasmic antibody-associated small-vessel vasculitis (AAV) and systemic lupus erythematosus (SLE) patients and controls broadly separates the samples into three groups. The numbers below the heat map show the distribution of AAV, SLE and control samples between the three groups. Vertical bars identify the following gene signatures; interferon, plasmablast and granulopoiesis (blue), CD8 T cell (purple), lymphoid (orange) and myeloid (yellow). (B–E) Gene set enrichment analysis was performed to look for enrichment of the interferon and granulopoiesis signatures described by Bennett *et al.*<sup>13</sup> (B) Interferon-inducible genes are significantly enriched in SLE patients. (C and D) Granulopoiesis-associated genes are significantly enriched in both (C) SLE and (D) AAV patients and (E) interferon-inducible genes are not enriched in AAV patients. (F) Serum IgG but not IgM levels correlate with the presence of the plasmablast signature. PB sig +ve and –ve indicates patients with or without the plasmablast signature, respectively. (G) The plasmablast signature is associated with increased disease activity at enrolment. Horizontal bars denote the mean, statistical significance was determined using a t test. (H) AAV and SLE patients have reduced CD4 T-cell and increased CD14 monocyte levels and in addition, AAV patients have significantly reduced CD8 T-cell levels compared with controls. Horizontal bars denote the mean, statistical significance was determined using analysis of variance followed by post hoc testing. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . (I) CD8A messenger RNA levels correlate with the relative abundance of CD8 T cells in the PBMC samples arrayed.



**Figure 2** Expression profiling purified cells identify large numbers of novel expression differences. (A,B) Differential gene expression was measured in purified cell subsets from both (A) systemic lupus erythematosus (SLE) and (B) antineutrophil cytoplasmic antibody-associated small-vessel vasculitis (AAV) patients. Pie charts represent individual cell types and their size is proportional to the number of differentially expressed genes observed in each. Differentially expressed genes were defined as expression differences greater than 1.5-fold found to be statistically significant following correction for multiple testing by setting the false discovery rate to 5%.

#### Purified CD4 T-cell analysis reveals an SLE-specific activation signature.

Among the 211 genes that show significant differential expression in CD4 T cells between SLE patients and controls (supplementary table 8, available online only), the biggest cluster represents interferon-inducible transcripts (figure 2A). In agreement with other studies,<sup>13 14</sup> the upregulation of interferon-inducible genes was not seen in all patients, and was independent of treatment at enrolment (supplementary figure 6, available online only).

To look for coordinated dysregulation of other pathways in the CD4 expressed genes differentially expressed in SLE patients we searched the KEGG ([www.genome.ad.jp](http://www.genome.ad.jp)), BioCarta ([www.biocarta.com](http://www.biocarta.com)) and GenMAPP ([www.genmapp.org](http://www.genmapp.org)) databases using the Pathway Miner tool.<sup>22</sup> This analysis revealed a network of differentially expressed genes broadly related to T-cell activation (figure 3A). An identical analysis of genes differentially expressed in PBMC samples from SLE patients failed to reveal these expression differences (supplementary

figure 7, available online only, and data not shown), as did a specific search for them using GSEA (figure 3B).

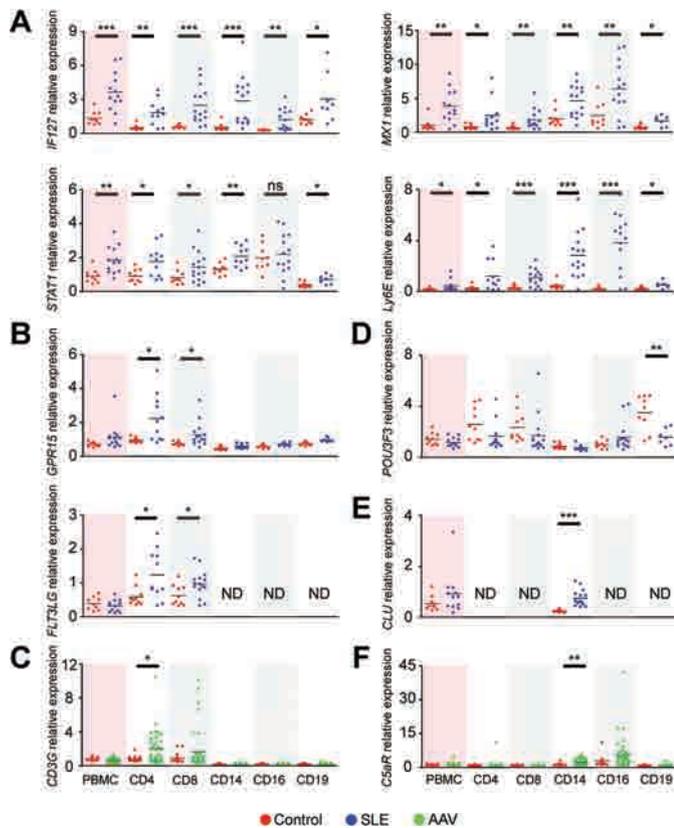
Interestingly, the expression of many genes in the network was strongly correlated with *STAT1* expression (figure 3C), suggesting that their expression may be regulated by signalling downstream of the type 1 interferon receptor. In keeping with this, analysis of a 2 kb segment of DNA immediately upstream of the start site of each of the 28 genes in the signature revealed the presence of consensus interferon response factor (IRF) 3 and/or seven binding sites<sup>23</sup> in the majority of cases (supplementary table 9, available online only). GSEA did not identify this signature in CD4 T cells from AAV patients (peak enrichment score 0.35,  $p=0.81$ ), indicating that it may be SLE specific.

This finding is in agreement with the observation that type 1 interferon signalling keeps activated murine T cells alive, independent of a measurable effect on *BCL2* or *BCL2L1* (*Bcl-X*) levels.<sup>24</sup> Consistent with this, we observed no difference in the expression level of these genes between SLE patients and controls (figure 3D); however, expression of the pro-apoptotic molecule *BNIP3L* is reduced and that of the anti-apoptotic molecule *CFLAR* elevated in SLE patients (figure 3D). Cell-specific array analysis has thus identified a novel interferon-associated, CD4 T-cell signature, which immediately suggests a mechanism for driving SLE pathogenesis.

#### Gene expression in monocytes differentiates patients with AAV and SLE from each other and from controls more robustly than PBMC data.

Gene expression data from unseparated PBMC samples is unable to discriminate SLE and AAV patients efficiently from each other or from controls (figure 1A and supplementary figure 8, available online only). To determine whether expression data from purified cells result in improved clustering, we analysed monocyte-derived microarray data from 13 SLE patients, 44 AAV patients and 25 normal controls (two SLE samples failed quality control). Data from 20 495 probes were subjected to ANOVA with correction for multiple testing to detect differentially expressed probes among the three groups. Post-hoc analysis identified 15 probes that best differentiated AAV patients from controls, 38 probes that differentiated SLE patients from controls and 41 probes that separated AAV and SLE patients. Hierarchical clustering of the 82 monocyte samples with the non-redundant list of probes generated from this analysis (supplementary table 10, available online only) clearly separated the samples into three groups (figure 4). Of the 14 samples in group 1, 12 are SLE patients, in group 2, 20 of the 21 samples are controls, and in group 3, 42 of the 47 samples are AAV patients (figure 4). Therefore, in comparison with the clustering obtained using PBMC data (figure 1A and supplementary figure 8, available online only), the monocyte data gave a much cleaner separation between diagnostic groups.

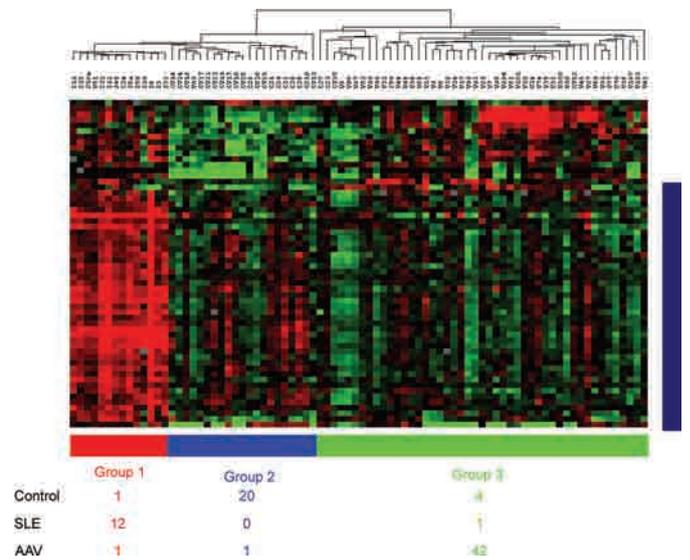
The most prominent feature of the ANOVA-derived gene list is a cluster of 42 interferon-inducible genes (vertical blue bar, figure 4 and supplementary table 10, available online only), high expression of which differentiates the SLE patients in group 1 from the rest. The remaining 16 probes predominantly represent genes upregulated in both disease groups compared with controls (supplementary table 10, available online only), and are enriched for genes encoding components of the acute phase response, including haptoglobin and orosomucoid (supplementary table 10, available online only).



**Figure 3** A gene signature consistent with T-cell activation is found in CD4 samples from systemic lupus erythematosus (SLE), but not antineutrophil cytoplasmic antibody-associated small-vessel vasculitis (AAV), patients. (A) Analysis of the differentially expressed genes in CD4 samples from SLE, but not AAV, patients (supplementary table 8, available online only) using Pathway Miner ([www.biorag.org](http://www.biorag.org)) revealed a gene association network consistent with T-cell activation. Nodes represent differentially expressed genes, orange and green nodes indicate upregulated and downregulated genes, respectively, in SLE patients compared with controls. Edges indicate annotated associations between genes, with the thickness of the edge indicating the number of annotated associations. (B) Gene set enrichment analysis found no evidence for this network in peripheral blood mononuclear cell (PBMC) samples from SLE patients. (C) Microarray-based expression levels of LAT and STAT1 are highly correlated. (D) BNIP3L and CFLAR, but not BCL2 or BCL2L1, are differentially expressed in CD4 samples from SLE patients compared with controls. Horizontal bars denote means, statistical significance was determined using a t test.

**DISCUSSION**

While the analysis of PBMC array data may identify expression differences of pathological importance, for example, the interferon signature in SLE, it is clear that many of the differences between disease and controls are due to variation in the proportion of cell subsets and are thus not seen when purified cell subsets are arrayed. Perhaps the best examples of such signatures are the plasmablast and granulocyte signatures seen in SLE patients; both reflect plasmablasts or immature granulocytes present in the PBMC. Most differentially expressed genes found in purified subsets were not found in PBMC. This was true for genes differentially expressed in rarer cell types, for example, B cells, but also for those expressed in both CD4 and CD8 cells, which together make up the biggest proportion of PBMC. Surprisingly, PBMC analysis even failed to detect expression differences unique to genes expressed in only a single cell type. However, once a gene signature has



**Figure 4** Monocyte gene expression data clearly differentiates patients with systemic lupus erythematosus (SLE) and antineutrophil cytoplasmic antibody-associated small-vessel vasculitis (AAV) from each other and from normal controls. Hierarchical clustering using data from genes expressed in CD14 monocytes separates the samples into the three diagnostic groups (compared with Figure 1A). Hierarchical clustering was performed using the Pearson correlation coefficient as distance metric and average-linkage clustering. The vertical blue bar denotes interferon-inducible genes.

been identified in purified cells, it may then be feasible in follow-up studies to test for the signature in unseparated cells using quantitative PCR of selected genes less influenced by population variation or by bioinformatics normalisation for cell proportions.

Analysis of purified T cells revealed an SLE-associated T-cell signature not seen in PBMC analyses,<sup>13,14</sup> which may play a role in pathogenesis. T cells infiltrate inflamed organs and provide help for autoreactive B cells in SLE,<sup>25</sup> and a number of T-cell defects have been reported in SLE patients.<sup>26,27</sup> The majority of the genes making up the T-cell signature has promoters that contain consensus IRF3 and/or IRF7 binding sites and their expression levels correlate strongly with that of *STAT1*, suggesting that the signature might be driven by type 1 interferon signalling.

One component of this signature is increased expression of *CXCR6*, the receptor for *CXCL16*,<sup>28</sup> which is expressed on a proportion of memory T cells and CD16 natural killer cells and is upregulated on activation.<sup>29</sup> *CXCR6* defines a subset of memory/effector T cells with tissue-homing potential,<sup>30</sup> and consistent with this *CXCR6*<sup>+</sup> T cells are enriched in inflamed tissues from a number of inflammatory conditions, including rheumatoid arthritis,<sup>30</sup> Graves' disease<sup>31</sup> and sarcoidosis,<sup>32</sup> and treatment with an anti-*CXCL16* monoclonal antibody reduces the severity of collagen-induced arthritis in mice.<sup>33</sup> This first association of *CXCR6* on CD4 T cells with SLE might thus be associated with tissue inflammation.

Type 1 interferon signalling promotes activated T-cell survival.<sup>24</sup> The decreased expression of the pro-apoptotic molecule *BNIP3L* in SLE patients is particularly interesting as it is a functional homologue of *BNIP3*, which has been shown to play a role in activation-induced T-cell death.<sup>34</sup> Similarly, *CFLAR*, also upregulated by type 1 interferon,<sup>35</sup> protects T cells from T-cell receptor-mediated apoptosis.<sup>36</sup> This is an example of

a disease-related signature providing insight into pathogenesis that is only seen in purified cells.

Fundamental to the generation of expression-based diagnostic biomarkers is accurate clustering of patients by diagnosis. Expression data from unseparated cells is very poor at differentiating SLE and AAV. However, data from purified cells result in significantly improved discrimination, presumably because the differentially expressed genes reflect genuine differences in processes underlying disease pathogenesis rather than simply differences in relative cell proportion. It has long been hoped that microarrays would yield diagnostic tests in autoimmunity; our data suggest that the analysis of purified cell populations is a superior way to achieve this.

While array analysis of PBMC provides some useful information, the use of purified cell subsets identifies many more differentially expressed disease-specific genes. Such analysis promises to provide substantial advantages in the search for diagnostic and prognostic biomarkers in autoimmune disease.

**Acknowledgments** The authors would like to thank the Wellcome Trust Clinical Research Facility at Addenbrooke's Hospital, and all the patients and controls who have contributed to this study. KGCS is a Lister Prize Fellow.

**Funding** The study received financial support from the NIHR Cambridge Biomedical Research Centre, the Wellcome Trust, Kidney Research UK, the Medical Research Council and the Evelyn Trust. The Cambridge Institute for Medical Research is in receipt of a Wellcome Trust Strategic Award (079895).

**Competing interests** None.

**Ethics approval** This study was conducted with the approval of the Cambridge Local Research Ethics Committee.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

1. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;**286**:531–7.
2. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;**403**:503–11.
3. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;**415**:530–6.
4. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;**365**:671–9.
5. Lossos IS. Molecular pathogenesis of diffuse large B-cell lymphoma. *J Clin Oncol* 2005;**23**:6351–7.
6. van't Veer LJ, Bernards R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* 2008;**452**:564–70.
7. Häupl T, Østensen M, Grützka A, et al. Interaction between rheumatoid arthritis and pregnancy: correlation of molecular data with clinical disease activity measures. *Rheumatology (Oxford)* 2008;**47**(Suppl 3):iii19–22.
8. Satoh J, Nakanishi M, Koike F, et al. Microarray analysis identifies an aberrant expression of apoptosis and DNA damage-regulatory genes in multiple sclerosis. *Neurobiol Dis* 2005;**18**:537–50.
9. Biesen R, Demir C, Barkhudarova F, et al. Sialic acid-binding Ig-like lectin 1 expression in inflammatory and resident monocytes is a potential biomarker for monitoring disease activity and success of therapy in systemic lupus erythematosus. *Arthritis Rheum* 2008;**58**:1136–45.
10. Corvol JC, Pelletier D, Henry RG, et al. Abrogation of T cell quiescence characterizes patients at high risk for multiple sclerosis after the initial neurological event. *Proc Natl Acad Sci USA* 2008;**105**:11839–44.
11. Bryant PA, Smyth GK, Robins-Browne R, et al. Detection of gene expression in an individual cell type within a cell mixture using microarray analysis. *PLoS ONE* 2009;**4**:e4427.
12. Lyons PA, Koukoulaki M, Hatton A, et al. Microarray analysis of human leucocyte subsets: the advantages of positive selection and rapid purification. *BMC Genomics* 2007;**8**:64.
13. Bennett L, Palucka AK, Arce E, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med* 2003;**197**:711–23.
14. Baechler EC, Batliwalla FM, Karypis G, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci USA* 2003;**100**:2610–15.
15. Crow MK, Wohlgemuth J. Microarray analysis of gene expression in lupus. *Arthritis Res Ther* 2003;**5**:279–87.
16. Han GM, Chen SL, Shen N, et al. Analysis of gene expression profiles in human systemic lupus erythematosus using oligonucleotide microarray. *Genes Immun* 2003;**4**:177–86.
17. Rus V, Chen H, Zernetkina V, et al. Gene expression profiling in peripheral blood mononuclear cells from lupus patients with active and inactive disease. *Clin Immunol* 2004;**112**:231–4.
18. Alcorn DA, Barnes DA, Dooley MA, et al. Leukocyte gene expression signatures in antineutrophil cytoplasmic autoantibody and lupus glomerulonephritis. *Kidney Int* 2007;**72**:853–64.
19. Willcocks LC, Lyons PA, Clatworthy MR, et al. Copy number of FCGR3B, which is associated with systemic lupus erythematosus, correlates with protein expression and immune complex uptake. *J Exp Med* 2008;**205**:1573–82.
20. Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, et al., eds. *Bioinformatics and computational biology solutions using R and bioconductor*. New York: Springer, 2005:397–420.
21. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Statist Soc B* 1995;**57**:289–300.
22. Pandey R, Guru RK, Mount DW. Pathway Miner: extracting gene association networks from molecular pathways for predicting the biological significance of gene expression microarray data. *Bioinformatics* 2004;**20**:2156–8.
23. Lin R, Génin P, Mamane Y, et al. Selective DNA binding and association with the CREB binding protein coactivator contribute to differential activation of alpha/beta interferon genes by interferon regulatory factors 3 and 7. *Mol Cell Biol* 2000;**20**:6342–53.
24. Marrack P, Kappler J, Mitchell T. Type I interferons keep activated T cells alive. *J Exp Med* 1999;**189**:521–30.
25. Hoffman RW. T cells in the pathogenesis of systemic lupus erythematosus. *Clin Immunol* 2004;**113**:4–13.
26. Crispin JC, Kyttaris VC, Juang YT, et al. How signaling and gene transcription aberrations dictate the systemic lupus erythematosus T cell phenotype. *Trends Immunol* 2008;**29**:110–15.
27. Fritsch RD, Shen X, Illei GG, et al. Abnormal differentiation of memory T cells in systemic lupus erythematosus. *Arthritis Rheum* 2006;**54**:2184–97.
28. Matloubian M, David A, Engel S, et al. A transmembrane CXC chemokine is a ligand for HIV-coreceptor Bonzo. *Nat Immunol* 2000;**1**:298–304.
29. Unutmaz D, Xiang W, Sunshine MJ, et al. The primate lentiviral receptor Bonzo/STRL33 is coordinately regulated with CCR5 and its expression pattern is conserved between human and mouse. *J Immunol* 2000;**165**:3284–92.
30. Kim CH, Kunkel EJ, Boisvert J, et al. Bonzo/CXCR6 expression defines type 1-polarized T-cell subsets with extralymphoid tissue homing potential. *J Clin Invest* 2001;**107**:595–601.
31. Aust G, Kamprad M, Lamesch P, et al. CXCR6 within T-helper (Th) and T-cytotoxic (Tc) type 1 lymphocytes in Graves' disease (GD). *Eur J Endocrinol* 2005;**152**:635–43.
32. Agostini C, Cabrelle A, Calabrese F, et al. Role of CXCR6 and its ligand CXCL16 in the pathogenesis of T-cell alveolitis in sarcoidosis. *Am J Respir Crit Care Med* 2005;**172**:1290–8.
33. Nanki T, Shimaoka T, Hayashida K, et al. Pathogenic role of the CXCL16-CXCR6 pathway in rheumatoid arthritis. *Arthritis Rheum* 2005;**52**:3004–14.
34. Wan J, Martinvalet D, Ji X, et al. The Bcl-2 family pro-apoptotic molecule, BNIP3 regulates activation-induced cell death of effector cytotoxic T lymphocytes. *Immunology* 2003;**110**:10–17.
35. Kanetaka Y, Hayashida M, Hoshika A, et al. Interferon-alpha induces transient upregulation of c-FLIP through NF-kappaB activation. *Exp Cell Res* 2008;**314**:246–54.
36. Zhang N, Hopkins K, He YW. c-FLIP protects mature T lymphocytes from TCR-mediated killing. *J Immunol* 2008;**181**:5368–73.