Published in final edited form as:

Clin Immunol. 2014 February ; 150(2): 192–200. doi:10.1016/j.clim.2013.12.002.

A T cell gene expression panel for the diagnosis and monitoring of disease activity in patients with systemic lupus erythematosus

Alexandros P. Grammatikos, Vasileios C. Kyttaris, Katalin Kis-Toth, Lisa M. Fitzgerald, Amyn Devli, Michele D. Finnell, and George C. Tsokos*

Division of Rheumatology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, 02215, USA

Alexandros P. Grammatikos: agramma1@bidmc.harvard.edu; Vasileios C. Kyttaris: vkyttari@bidmc.harvard.edu; Katalin Kis-Toth: kkistoth@bidmc.harvard.edu; Lisa M. Fitzgerald: lfitzge2@bidmc.harvard.edu; Amyn Devli: adevlin@bidmc.harvard.edu; Michele D. Finnell: mfinnell@bidmc.harvard.edu

Abstract

Systemic Lupus Erythematosus (SLE) remains a challenging disease to diagnose and follow, as no reliable biomarkers are known to date. We designed a gene expression panel with 40 genes known to play a role in SLE pathogenesis. We found that the combined expression of these genes in SLE T cells can accurately differentiate SLE from healthy individuals and patients with other autoimmune diseases. The accuracy of the test increased further (83%) when only three out of the initial genes (OAS2, CD70 and IL10) were used. A T cell score, calculated from the combined expression levels of these genes, correlated positively with various SLE activity markers in a cross-sectional cohort and in a few patients that were followed prospectively. These data showcase the usefulness of measuring mRNA levels of key molecules in diagnosing and following patients with SLE.

Keywords

OAS2; CD70; IL10; Biomarkers; Diagnostic; Flare

1. Introduction

Current tools used to diagnose systemic lupus erythematosus (SLE) and to monitor disease activity, including systemic lupus erythematosus disease activity index (SLEDAI) and the British Isles Lupus Assessment Group (BILAG) rely on clinical and laboratory markers and have acknowledged limitations given the clinical complexity of the disease and the frequently superimposed comorbid conditions [1–3].

Most of the markers used by these scoring systems reflect non-specific inflammation in the target tissues. Validated markers that reflect the immunologic aberration of SLE, including complement levels and dsDNA antibodies have not risen to be acclaimed as reliable disease

© 2013 Elsevier Inc. All rights reserved.

*Corresponding author. Fax: +1 617 735 4170. gtsokos@bidmc.harvard.edu (G.C. Tsokos).

Conflict of interest statement

The author(s) declare that there are no conflicts of interest.

Appendix A Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.clim.2013.12.002>.

biomarker. It is therefore apparent that reliable and easy to use biomarkers that can objectively diagnose SLE, reflect disease activity and/or predict flares are needed.

Novel approaches for disease diagnosis and classification rely on the use of panels of expression levels of genes related to disease pathogenesis. Recently this approach has been successfully applied to cancer prognosis and commercially available genetic-based risk tests have been produced [4,5]. Similar attempts have been performed in the field of autoimmune diseases. For example, a CD8 + T cell transcription signature was recently found to be associated with disease prognosis in anti-neutrophil cytoplasmic antibody (ANCA)-associated vasculitis and SLE patients [6]. Furthermore, interferon (IFN)-induced gene signatures have been found to correlate with disease activity in a number of SLE studies [7–9].

Several lines of evidence suggest that T cells play an important role in the pathogenesis of SLE [1,10]. T cells are key modulators of B cell activation and immunoglobulin class switching that are necessary steps for the production of high affinity autoantibodies seen in SLE. SLE T cells also fail to execute appropriately their regulatory functions while at the same time exhibit inappropriate homing to target organs. SLE T cells have an abnormally activated phenotype with pre-aggregated lipid rafts, high calcium flux upon activation, and production of pro-inflammatory cytokines such as IL-17 [10–12]. Abnormally expressed signaling molecules (CD40L, IL10) [10,13,14] have been considered as treatment targets, but their value as disease biomarkers has not been addressed formally.

We sought to evaluate the potential of developing a disease-specific biomarker in SLE based on the expression levels of genes that have been claimed to contribute to the abnormal T cell function. In a preliminary report, we showed that measuring disease specific gene expression using a PCR-based gene array can successfully differentiate SLE patients from control individuals [15]. Using an expanded study sample and prospective collection of samples we show that the expression of specific pathogenesis-linked genes can be used to assist diagnosis and monitor disease activity in individual patients.

2. Patients and methods

2.1. Ethics statement

This study was approved by the Institutional Review Board of BIDMC. Written informed consent was obtained from all participating subjects and all clinical investigation was conducted according to the principles expressed in the Declaration of Helsinki.

2.2. SLE and control samples

Blood samples were obtained from 65 SLE patients members of the Lupus Center at the Rheumatology Division of Beth Israel Deaconess Medical Center (BIDMC). Seven of those patients were followed up longitudinally in two months interval for a period of 1 year. All participating patients fulfilled the American College of Rheumatology criteria for the diagnosis of SLE. Blood was also obtained from 27 healthy subjects and 16 rheumatoid arthritis (RA) patients.

2.3. Cell extraction and RNA isolation

Density- gradient centrifugation was used to isolate peripheral blood mononuclear cells (PBMCs) from whole blood samples. For T cell isolation, blood was incubated first for 30 min with a rosette T cell purification kit (StemCell technologies, Vancouver, Canada) that contained a tetrameric antibody mixture against CD14, CD16, CD19, CD56 and glyA that

attaches non-T cells to erythrocytes. Lymphocyte separation medium (Cellgro, Manassas, VA) was subsequently used to separate these complexes from T cells.

From 3×10^6 cells, RNA, DNA and protein were extracted using Qiagen (Valencia, CA) AllPrep extraction kit. A DNase-I treatment step (Qiagen) was added to the standard protocol to ensure exclusion of genomic DNA from the final product. OD_{260/280} measurements were used as a measure of isolated RNA quality.

Total RNA was reversely transcribed into cDNA using Promega (Madison, WI) reverse transcription system and a mixture of 1:10 oligo (dT)₂₀ to random hexamer primers. Reverse transcription was performed in a conventional thermocycler.

2.4. Design of gene array and qPCR

Quantitative real time polymerase chain reaction (qPCR) was performed to measure gene expression levels using the Universal Probe Library (UPL) system from Roche (Indianapolis, IN). Probes only bind to the amplicon of interest and do not release their signal unless the 5'/3' exonuclease activity of the Taq DNA Polymerase unquenches the reporter dye.

Forty gene targets, two reference (CD3 ϵ and GAPDH) and two control assays (human genomic contamination and no template controls) were included in the array (Supplementary Table 1).

All primers and probes were designed using the online UPL Assay Design Center tool from Roche (Indianapolis, IN). Only intron spanning assays were kept as a safeguard against genomic DNA contamination. Amplicon sizes <200 bp were chosen to ensure fast and efficient amplification. A detailed list of all included genes, along with their amplicon lengths, accession numbers, primer and probe sequences can be found in Supplementary Table 2.

For each gene, a reaction mixture was prepared using a final concentration of 200 nM for each primer, 100 nM for probes, 1 $\times$ LightCycler 480 Probes Master (containing FastStart Taq DNA polymerase), cDNA and water in a total volume of 10 μ l. HPSF (High Purity Salt Free) primers were obtained from Eurofins MWG Operon (Huntsville, AL) and all other reagents from Roche (Indianapolis, IN). Reactions were prepared in 96 well plates. All plates were sealed and centrifuged for 2 min at 1500 $\times$ g in a standard swing-bucket centrifuge prior to amplification.

Amplification was performed on a Roche LightCycler 480 PCR instrument (Roche, Indianapolis, IN). Crossing points (Ct) were calculated using the second derivative maximum method, an algorithm that is based on the kinetics of each PCR reaction. The advantage of this method is that it requires minimal user input.

To confirm that single amplification products are produced, samples were stained with ethidium bromide and electrophoresed in agarose gel before visualising under a UV light camera (Fig. 1).

2.5. Analysis and statistical methods

Ct values over 40 were excluded from the analysis (considered to be less reliable). In the resulting dataset imputation was performed using the k-nearest neighbor algorithm (KNN) applying a Euclidean metric [16].

To account for differences in amplification efficiency we converted Ct values into copy numbers. Standard curves were created for each of the 42 genes using five 10-fold dilution series of a cDNA sample. Ct values for each dilution were plotted against the log of the cDNA concentration and copy numbers (CN) calculated using the formula: $CN = 10^{\frac{Ct-b}{m}}$ (b: intercept, m: slope). Normalization towards reference genes was performed by calculating

the ratio $\frac{CN_{GOI}}{CN_{ARG}}$ (GOI: gene of interest, ARG: average of reference genes).

To dissect the most important genes in differentiating SLE from control samples comparative marker selection (CMS) was used [17]. A two-sided *t*-test that computes the standardized mean difference between two classes was used to calculate a score for each gene. First the difference in the mean expression levels of the two classes was calculated and then this was divided by the variability of expression (represented by the square root of the sum of the standard deviation for each class, divided by the number of samples in each class).

For class prediction, CART (classification and regression trees) analysis was used [18] and the leave-one-out cross-validation method applied. Leave-one-out cross-validation was chosen over other validation methods, since it allows for better generalizability of the built model. It achieves this by allowing every data point to appear at least once in the test set, in contrast to random partitioning methods, in which part of the data is always excluded.

An unsupervised clustering technique, hierarchical clustering was used to confirm classification tree results and visualize gene expression patterns [19]. Spearman's complete correlation was employed as a metric of similarity.

CART, CMS, KNN imputation and hierarchical clustering were all performed on the GenePattern public server (Broad Institute, Cambridge, MA) [20]. Three-dimensional plotting was performed in JMP software (SAS Institute Inc, Middleton, MA). Comparisons of mean expression levels and correlations were performed in Prism 5 (GraphPad, La Jolla, CA) and Microsoft Excel 2010 (Redmond, WA).

3. Results

3.1. Evaluation of an SLE T cell biochemical profile for disease diagnosis

Based on several reports [11,21,22] using small number of patients, we chose 40 genes whose products are either abnormally elevated or decreased in SLE T cells in comparison to controls and have been linked to the pathogenesis of the disease (Supplementary Table 1). Genes chosen to be included in the array encode for different types of proteins. Specifically, six genes encode for cytokines (IL2, IL10, IL21, IL23A, IFN γ and BAFF), ten for cell surface receptors (CD40L, CD44, CD70, ICAM1, Notch1, CTLA4, TNFRSF4, PDCD1, ITGAM and LY9), six for transcription factors (CREB1, CREM α , FOXp3, GATA3, NFATC2, RELA) and nineteen for various cytoplasmic signaling molecules, including six protein kinases (ERK, PRKAR1B, PRKAR2B, PRKCD, PRKCQ, ROCK1) and two protein phosphatases (PPP2CA, PPP2CB) [10,11,21,23,24].

A caveat in analyzing expression levels of multiple genes at the same time is that each primer used in the qPCR assay exhibits slightly different amplification efficiency. To account for this discrepancy, standard curves were run and the amplification efficiency of each gene calculated separately. Copy numbers were then extracted and expression levels for every gene normalized to the expression of reference genes.

The initial study cohort included T cell samples from 92 individuals, 65 SLE patients (who fulfilled at least 4/11 criteria for SLE) with various disease activities and 27 control subjects. The demographic and clinical characteristics of study subjects are shown in Tables 1 and 2. To determine whether concurrent determination of the expression levels of the chosen 40 genes can identify SLE patients with sufficient fidelity we used classification trees. Classification trees help combine the predictive value of multiple variables (e.g. gene expression levels) in a single test. They are powerful tools, particularly in detecting relationships between predictors and outcomes that are non-linear and involve higher order interactions [18]. Using classification and regression trees (CART) analysis, a predicting rule was constructed based on gene expression levels. Results were validated using leave-one-out cross validation, a technique that is known [25] to be very reliable for smaller datasets.

Based solely on gene expression levels and using the classification tree algorithm SLE versus control identity of samples was predicted correctly in 75% of cases. The 40 gene panel had 91% sensitivity (59/65 SLE samples predicted correctly) and 37% specificity (10/27 control samples predicted correctly). The positive predictive value (PPV) and negative predictive value (NPV) for this test were 78% and 63% respectively.

3.2. A three-gene T cell panel correctly differentiates SLE samples from control and rheumatoid arthritis samples

The genes that were used to build this array have been shown to be important in the pathogenesis of the disease but their usefulness as disease biomarkers has not been established. We therefore asked which of the 40 genes used in the final analysis contributed the most in differentiating SLE patients from controls. To reach to the list of most important genes, we used the comparative marker selection (CMS) method [17]. CMS applies a *t*-test statistic score to determine which variables display the strongest correlation with a subgroup class. Fig. 2a shows those genes, the expression levels of which were best suited to distinguish SLE from control samples. These included, in order of decreasing importance, OAS2 (2'-5'-oligoadenylate synthetase 2), CD70, IL10 (interleukin 10), CTLA4 (cytotoxic T-lymphocyte antigen 4), PRKAR2B (protein kinase, cAMP-dependent, regulatory, type II, beta) and so on.

We then asked whether the classifier would perform better if only the genes that were better at discriminating SLE from controls were included. Interestingly, the overall prediction accuracy of the classifier remained the same when the number of genes in the array was reduced (by one each time) down to only 4 genes (Fig. 2b). However, when keeping just the top three genes, the classifier's accuracy increased to 83%. Expression levels of OAS2, CD70 and IL10 alone were sufficient to discriminate SLE from control samples successfully in most cases. This three gene classifier achieved 94% (61/65) sensitivity, 56% (15/27) specificity and a PPV and NPV of 84% and 79% respectively.

To confirm the above, an unsupervised clustering technique was used and samples were organized into groups based on their gene expression pattern. Using hierarchical clustering and plotting on expression levels of OAS2, CD70 and IL10, we were able to see that control samples tend to cluster separately from SLE samples (Fig. 3a). Relative copy numbers for all three genes were found to be low in almost all healthy subjects tested. On the contrary, expression level of at least one of the three genes was found to be high in the SLE subjects tested. Average expression levels for all three genes were significantly higher in SLE patients in comparison to control individuals (false discovery rate method: CD70, $q = 0.01$; IL10, $q = 0.02$; OAS2, $q = 0.01$).

In order to determine whether this gene panel could be applied to discriminate SLE from other autoimmune diseases as well, we collected T cells from sixteen RA patients (Table 1) and tested the expression levels of OAS2, CD70 and IL10. Similarly to control subjects, expression of all three genes was found to be low in RA subjects (Fig. 3a). These samples exhibited similar expression patterns and were clustered in two separate groups by the hierarchical clustering algorithm. More importantly, these samples clustered separately from SLE samples (Fig. 3a).

To better visualize OAS2, CD70 and IL10 expression levels for each of the study participants we plotted expression levels on a 3D plot representing CN ratios of the three target genes (Fig. 3b). Again, we were able to observe that control and RA samples clustered together in the same area of the plot while SLE samples tended to spread broadly, in a separate area of the plot.

In a subgroup analysis, the three-gene panel was better in diagnosing steroid-free (77%) over steroid-treated patients (64%) and those with a longer (72%) over those with a shorter disease course (62%) (Fig. 3c). Cases exhibiting acute nephritis were the ones most accurately diagnosed (74%), followed by those exhibiting cytopenia (66%), arthritis (58%), serositis (57%) and mucocutaneous (50%) manifestations (Fig. 3c and Table 2).

3.3. The three-gene T cell panel correlates with disease activity in SLE patients

We then asked whether the expression of these three genes may be useful in monitoring disease activity in patients with SLE. To address this question we combined the expression levels of all three genes into a single score: for each sample expression of OAS2, CD70 and IL10 was summed after normalizing against the average expression and variability of control samples for each gene target.

This 'T cell score' was found to be significantly higher in SLE patients than control individuals (mean $\pm$ SEM: 3.7 ± 0.65 vs. 0 ± 0.41 , $p < 0.001$). The T cell score correlated with SLEDAI, the most widely used composite SLE disease activity index ($p < 0.001$). (Fig. 3c). It also correlated with dsDNA antibody titers ($p < 0.001$), erythrocyte sedimentation rate (ESR) ($p = 0.04$), and the Physician Global Assessment (PGA) ($p < 0.001$). A statistical trend towards negative correlation with markers known to be inversely associated with disease activity (lymphocyte count, C3 and C4 levels) was also seen.

3.4. Application of the three-gene panel in peripheral blood mononuclear cells

Although T cells were the primary target of our study, PBMCs are much easier to isolate from peripheral blood and would be more practical targets to assess. To evaluate our 3 gene panel in PBMCs we collected blood from 20 SLE patients and 9 healthy control subjects and measured expression levels of OAS2, IL10 and CD70. The three-gene panel allowed for correct classification of SLE samples in 69% of cases; achieving 78% specificity (7/9) and 65% sensitivity (13/20) (CART cross-validation). Combined expression levels of all three genes were significantly higher in SLE patients in comparison to controls (mean $\pm$ SEM: 7.86 ± 1.99 vs. 0 ± 0.55 , $p = 0.02$, Fig. 4a) and correlated significantly with PGA ($p < 0.01$, Fig. 4b). A trend towards positive correlation with SLEDAI and ESR was also seen.

3.5. Evaluation of the three-gene score prospectively

To determine whether this gene expression score correlates with disease activity in individual patients we recruited seven SLE patients and collected blood samples from them in two month-intervals for a period of 1 year (Table 3).

The T cell score (reflecting expression levels of OAS2, CD70 and IL10) correlated positively with SLEDAI in six patients, with PGA in five, with dsDNA in five, and with ESR in three (Fig. 5). On the other hand, the T cell score correlated negatively with lymphocyte counts in five patients.

4. Discussion

There is an increased interest in the use of biomarkers for disease diagnosis, monitoring and prognosis in various diseases in recent years. Such efforts have also been undertaken in the field of autoimmune disorders [7–9,26–28]. Several reasons have been cited to explain why these studies have not had the anticipated success. Analysis of peripheral blood mononuclear cells instead of specific cell populations represents a serious limitation because cell subsets vary in patients with SLE and this variation is further accentuated by use of steroids and cytotoxic drugs. In addition, previous studies evaluated the expression levels of indiscriminately put together thousands of genes. Lastly, the use of microarrays has been criticized for being less accurate than real time PCR in determining individual gene expression levels [29].

To overcome these obstacles we elected to evaluate expression levels of T cell genes using qPCR in pure T cell populations. A probe based technology was employed to avoid detection of unspecific amplification products. Real-time PCR crossing points (Cts) were identified by the second derivative maximum method, which requires minimal user input. Finally, Ct values were converted into copy numbers, in order to account for differences in amplification efficiency. T cells were particularly chosen, since multiple signaling abnormalities have been previously described in these cells in SLE patients [30,31]. A panel of 40 genes previously evaluated in a pilot study and claimed to contribute to aberrant T cell function in SLE was chosen to be studied [15].

To interpret expression levels of multiple genes simultaneously we employed novel mathematical algorithms like classification trees (CART) and comparative marker selection. These techniques allow for the combined predictive accuracy of multiple biomarkers to be assessed. In this manner, gene signatures characteristic of patient groups, irrespective of variations in individual gene expression levels, can be detected. Both CART and comparative marker selection are methods that require very little input from the analyst and are thus considered very objective.

Using the 40 gene panel we were able to correctly differentiate SLE samples from control in 75% of the cases. To identify which genes contributed the most in differentiating SLE from control samples we used comparative marker selection. The genes that were found to be the most important have been previously confirmed to be associated with the pathogenesis of SLE [21,32] and their expression levels to vary in the same direction. An exception was CD40L, which has previously been reported to be overexpressed in SLE patients, but was found to be decreased in our study [33]. Expression levels of CD40L in inactive SLE are, nevertheless, known to be normal and our study included many patients with low disease activity. Indeed, CD40L expression was low in these patients (data not shown) and this may have skewed overall study findings.

When trying to reduce the gene panel to include only the three genes that achieved the highest score in discriminating SLE from control status, its accuracy increased to 83%. An independent clustering technique, hierarchical clustering confirmed this finding. We were also able to show that SLE samples tended to cluster separately from, not only control, but also RA samples. The sensitivity of this gene panel approached that of anti-nuclear antibodies (ANA) (94% compared to 95%), the most commonly used biomarker in SLE and

its specificity exceeded that of ANA (56% compared to 49%). In a subgroup analysis, the three-gene panel performed better in prednisone-free patients and those with a longer history of disease. Although this analysis was limited by its small sample size, it may be plausible to hypothesize that prednisone normalizes abnormal expression levels of many genes, including OAS2, CD70 and IL10.

After we combined the expression levels of OAS2, CD70 and IL10 into a single score, the “T cell score”, we were able to confirm that expression is associated with disease severity. Following seven SLE patients prospectively we were able to confirm this finding in most cases. Larger studies would be, however, necessary in order to confirm the value of the ‘T cell score’ as a valuable biomarker in disease activity monitoring.

The study of isolated T cells has many advantages over PBMCs, but the added complexity of cell isolation makes studies in PBMCs more practical. Using PBMCs to measure expression levels of OAS2, CD70 and IL10, correct identification of SLE cases from healthy controls was achieved in a lower rate than when using T cells (69% over 83%). Nevertheless, combined expression levels were found to be both significantly higher in SLE and correlated with patients’ disease activity. These results seem promising, but will require validation in larger sample sets to be confirmed.

Although we started this study by assessing the usefulness of 40 genes in diagnosing/ following SLE patients, we found through unsupervised analyses, that only three genes can in combination differentiate SLE from healthy controls.

IL-10, shown to be increased in SLE patients vs. controls [34,35], plays a role in B cell maturation and antibody production. CD70 is expressed on activated T cells, plays a role in T cell proliferation and is also known to be upregulated on SLE T cells [36]. Finally, OAS2 is an interferon-inducible gene and its expression is increased in SLE vs. controls [37]. The T cell score therefore captures three very important aspects of the immune dysregulation in SLE: B cell over-activity, T cell dysfunction and interferon production.

Our study confirms previous studies on the importance of IFN-inducible genes in SLE [8,27,28] and extends these findings to other gene pathways. The high sensitivity and good specificity of the three-gene panel make it a potential screening tool for patients with suspected SLE. However, its ability to monitor disease activity needs to be validated systematically in larger and more diverse patient cohorts. Although it would be meaningful to compare the sensitivity and specificity of the three-gene panel to that of IFN-inducible genes, this is not possible at this point because such information has not been published [8,27,28].

In conclusion, we present findings on a three-gene panel that may assist clinicians in SLE diagnosis. Future studies will help us better understand how the expression levels of only three genes that are pathophysiologically relevant to SLE can become a clinical tool.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The work reported herein was supported by RO1 AI 42269 and AI 49954

Glossary

SLE	systemic lupus erythematosus
RA	rheumatoid arthritis
SLEDAI	systemic lupus erythematosus disease activity index
BILAG	British Isles lupus assessment group
qPCR	real time polymerase chain reaction
CART	classification and regression trees
CMS	comparative marker selection
KNN	k-nearest neighbor algorithm
PPV	positive predicted value
NPV	negative predictive value
ANA	anti-nuclear antibodies
PGA	physician global assessment

References

1. Tsokos GC. Systemic lupus erythematosus. *N. Engl. J. Med.* 2011; 365:2110–2121. [PubMed: 22129255]
2. Sturfelt AG, Nived O. Instruments to measure disease activity in systemic lupus erythematosus – is our net tight enough to catch the hydra. *J. Rheumatol.* 2011; 38:3–4. [PubMed: 21196581]
3. The American College of Rheumatology response criteria for systemic lupus erythematosus clinical trials: measures of overall disease activity. *Arthritis Rheum.* 2004; 50:3418–3426. [PubMed: 15529383]
4. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* 2002; 415:530–536. [PubMed: 11823860]
5. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D, Foekens JA. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet.* 2005; 365:671–679. [PubMed: 15721472]
6. McKinney EF, Lyons PA, Carr EJ, Hollis JL, Jayne DR, Willcocks LC, Koukoulaki M, Brazma A, Jovanovic V, Kemeny DM, Pollard AJ, Macary PA, Chaudhry AN, Smith KG. A CD8+ T cell transcription signature predicts prognosis in autoimmune disease. *Nat. Med.* 2010; 16:586–591. (581p following 591). [PubMed: 20400961]
7. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, Pascual V. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J. Exp. Med.* 2003; 197:711–723. [PubMed: 12642603]
8. Feng X, Wu H, Grossman JM, Hanvivadhanakul P, FitzGerald JD, Park GS, Dong X, Chen W, Kim MH, Weng HH, Furst DE, Gorn A, McMahon M, Taylor M, Brahn E, Hahn BH, Tsao BP. Association of increased interferon-inducible gene expression with disease activity and lupus nephritis in patients with systemic lupus erythematosus. *Arthritis Rheum.* 2006; 54:2951–2962. [PubMed: 16947629]
9. Nikpour M, Dempsey AA, Urowitz MB, Gladman DD, Barnes DA. Association of a gene expression profile from whole blood with disease activity in systemic lupus erythematosis. *Ann. Rheum. Dis.* 2008; 67:1069–1075. [PubMed: 18063674]
10. Kyttaris VC, Tsokos GC. Targeting lymphocyte signaling pathways as a therapeutic approach to systemic lupus erythematosus. *Curr. Opin. Rheumatol.* 2011; 23:449–453. [PubMed: 21720246]

11. Grammatikos AP, Tsokos CG. T cell abnormalities in patients with SLE (systemic lupus erythematosus). *World Hell. Biomed. News.* 2011; 2:7–11.
12. Crispin JC, Kyttaris VC, Terhorst C, Tsokos GC. T cells as therapeutic targets in SLE. *Nat. Rev. Rheumatol.* 2010; 6:317–325. [PubMed: 20458333]
13. Lo MS, Tsokos GC. Treatment of systemic lupus erythematosus: new advances in targeted therapy. *Ann. N. Y. Acad. Sci.* 2012; 1247:138–152. [PubMed: 22236448]
14. Grammatikos AP, Tsokos GC. Immunodeficiency and autoimmunity: lessons from systemic lupus erythematosus. *Trends Mol. Med.* 2012; 18:101–108. [PubMed: 22177735]
15. Juang YT, Peoples C, Kafri R, Kyttaris VC, Sunahori K, Kis-Toth K, Fitzgerald L, Ergin S, Finnell M, Tsokos GC. A systemic lupus erythematosus gene expression array in disease diagnosis and classification: a preliminary report. *Lupus.* 2011; 20:243–249. [PubMed: 21138984]
16. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB. Missing value estimation methods for DNA microarrays. *Bioinformatics.* 2001; 17:520–525. [PubMed: 11395428]
17. Gould J, Getz G, Monti S, Reich M, Mesirov JP. Comparative gene marker selection suite. *Bioinformatics.* 2006; 22:1924–1925. [PubMed: 16709585]
18. Speybroeck N. Classification and regression trees. *Int. J. Public Health.* 2012; 57:243–246. [PubMed: 22015650]
19. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A.* 1998; 95:14863–14868. [PubMed: 9843981]
20. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nat. Genet.* 2006; 38:500–501. [PubMed: 16642009]
21. Crispin JC, Liossis SN, Kis-Toth K, Lieberman LA, Kyttaris VC, Juang YT, Tsokos GC. Pathogenesis of human systemic lupus erythematosus: recent advances. *Trends Mol. Med.* 2010; 16:47–57. [PubMed: 20138006]
22. Fernandez D, Bonilla E, Phillips P, Perl A. Signaling abnormalities in systemic lupus erythematosus as potential drug targets. *Endocr. Metab. Immune Disord. Drug Targets.* 2006; 6:305–311. [PubMed: 17214576]
23. Crispin JC, Kyttaris VC, Juang YT, Tsokos GC. How signaling and gene transcription aberrations dictate the systemic lupus erythematosus T cell phenotype. *Trends Immunol.* 2008; 29:110–115. [PubMed: 18249583]
24. Moulton VR, Tsokos GC. Alternative splicing factor/splicing factor 2 regulates the expression of the zeta subunit of the human T cell receptor-associated CD3 complex. *J. Biol. Chem.* 2010; 285:12490–12496. [PubMed: 20118245]
25. Breiman, L.; Friedman, J.; Stone, CJ.; Olshen, RA. *Classification and Regression Trees.* Washington DC: Chapman and Hall/CRC; 1984.
26. Landolt-Marticorena C, Bonventi G, Lubovich A, Ferguson C, Unnithan T, Su J, Gladman DD, Urowitz M, Fortin PR, Wither J. Lack of association between the interferon-alpha signature and longitudinal changes in disease activity in systemic lupus erythematosus. *Ann. Rheum. Dis.* 2009; 68:1440–1446. [PubMed: 18772188]
27. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, Shark KB, Grande WJ, Hughes KM, Kapur V, Gregersen PK, Behrens TW. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc. Natl. Acad. Sci. U. S. A.* 2003; 100:2610–2615. [PubMed: 12604793]
28. Kirou KA, Lee C, George S, Louca K, Papagiannis IG, Peterson MG, Ly N, Woodward RN, Fry KE, Lau AY, Prentice JG, Wohlgemuth JG, Crow MK. Coordinate overexpression of interferon-alpha-induced genes in systemic lupus erythematosus. *Arthritis Rheum.* 2004; 50:3958–3967. [PubMed: 15593221]
29. Valasek MA, Repa JJ. The power of real-time PCR. *Adv. Physiol. Educ.* 2005; 29:151–159. [PubMed: 16109794]
30. Moulton VR, Tsokos GC. Abnormalities of T cell signaling in systemic lupus erythematosus. *Arthritis Res. Ther.* 2011; 13:207. [PubMed: 21457530]

31. Kyttaris VC, Krishnan S, Tsokos GC. Systems biology in systemic lupus erythematosus: integrating genes, biology and immune function. *Autoimmunity*. 2006; 39:705–709. [PubMed: 17178567]
32. Rhodes B, Vyse TJ. The genetics of SLE: an update in the light of genome-wide association studies. *Rheumatology*. 2008; 47:1603–1611. [PubMed: 18611920]
33. Desai-Mehta A, Lu L, Ramsey-Goldman R, Datta SK. Hyperexpression of CD40 ligand by B and T cells in human lupus and its role in pathogenic autoantibody production. *J. Clin. Invest.* 1996; 97:2063–2073. [PubMed: 8621796]
34. Csiszar A, Nagy G, Gergely P, Pozsonyi T, Pocsik E. Increased interferon-gamma (IFN-gamma) IL-10 and decreased IL-4 mRNA expression in peripheral blood mononuclear cells (PBMC) from patients with systemic lupus erythematosus (SLE). *Clin. Exp. Immunol.* 2000; 122:464–470. [PubMed: 11122256]
35. Nguyen H, Weng NP. IL-21 preferentially enhances IL-15-mediated homeostatic proliferation of human CD28+ CD8 memory T cells throughout the adult age span. *J. Leukoc. Biol.* 2010; 87:43–49. [PubMed: 19797296]
36. Oelke K, Lu Q, Richardson D, Wu A, Deng C, Hanash S, Richardson B. Overexpression of CD70 and overstimulation of IgG synthesis by lupus T cells and T cells treated with DNA methylation inhibitors. *Arthritis Rheum.* 2004; 50:1850–1860. [PubMed: 15188362]
37. Tang J, Gu Y, Zhang M, Ye S, Chen X, Guo Q, Qian J, Bao C, Chen S, Shen N. Increased expression of the type I interferon-inducible gene lymphocyte antigen 6 complex locus E in peripheral blood cells is predictive of lupus activity in a large cohort of Chinese lupus patients. *Lupus*. 2008; 17:805–813. [PubMed: 18755862]

Figure 1.

Real time PCR amplification products visualised in agarose gel. Real-time PCR amplification products for all genes included in the array were electrophoresed in agarose gel and visualised under UV light. In all cases single band products were seen and their size agreed with the predicted one (shown in brackets, bp: base pairs).

Figure 2.

Accuracy of the genetic panel in predicting SLE from control status. a) Genes were ranked according to their ability to differentiate SLE status from control using comparative marker selection method. For each gene a test statistic score was calculated: the higher the score, the higher the potential of the gene to be used as a biomarker. Gene targets with a positive score were those upregulated in SLE, while those with a negative, in healthy controls. OAS2, CD70 and IL10 were the genes that ranked the highest scores. b) Reducing the gene panel to only include the top differentially expressed genes improves its overall classification accuracy. Measurement of only three genes, OAS2, CD70 and IL10 correctly differentiates SLE samples from control in 83% of cases. Classification accuracy was assessed using CART (classification and regression trees).

Figure 3.

The three-gene panel differentiates SLE samples from healthy control and rheumatoid arthritis (RA) samples. **a**) Heatmap of expression levels of OAS2, CD70 and IL10 in 65 SLE, 27 control and 16 RA samples. Control and RA samples (blue and green bars respectively, top of the heatmap) grouped together, and separately from SLE samples (red bars) in hierarchical clustering. **b**) Three dimensional plotting of SLE, RA and control samples based on their OAS2, CD70 and IL10 expression levels (SLE samples in red, control samples in blue and RA samples in green). Plotting on expression levels, control and RA samples were seen to cluster separately from SLE samples. **c**) Accuracy of the three-gene panel was higher in steroid-free (77%), compared to steroid-treated (64%) patients, and in those with a longer (>1 year) history of disease (72%). Patients having nephritis were the ones most accurately diagnosed (74%), followed by those with cytopenia (66%), arthritis (58%), serositis (57%) and mucocutaneous manifestations (50%). **d**) Expression levels of

OAS2, CD70 and IL10 were combined to create a 'T cell score'. T cell score levels were significantly higher in the SLE than the healthy control group and correlated significantly with dsDNA, PGA, SLEDAI and ESR. Platelet and creatinine levels also correlated positively while C3, C4 levels and lymphocyte count negatively with the T cell score.

Figure 4.

Evaluation of the three-gene panel in PBMCs. a) RNA extracted from PBMCs was used to measure expression levels of OAS2, CD70 and IL10 in 20 SLE and 9 control samples. After converting into copy numbers, expression levels of these 3 genes were combined into a single score. Cumulative expression was significantly higher in SLE patients than in controls and correlated significantly with PGA (b). A positive correlation was also noted for SLEDAI and ESR; and a negative for lymphocyte count and C3 levels.

Figure 5.

T cell score in longitudinally followed up patients. Seven SLE patients were monitored longitudinally on 2 month intervals. Blood was drawn on each visit to determine T cell expression levels of OAS2, CD70 and IL10 and measure standard laboratory markers. Fifty two time points were recorded in total over a period of two and a half years. The T cell score correlated positively with SLEDAI in six patients. Three examples are shown (left panel: $r = 0.71$ (1) 12.6, 16 (2) -1.9, 12 (3) -2.8, 4 (4) -1.5, 4 (5) -2.6, 4 (6) 1.8, 2 (7) -1.4, 2 (8) -2, 4; middle panel: $r = 0.35$ (1) 7.2, 5 (2) -0.02, 4 (3) 0.12, 4 (4) -1.2, 4 (5) -1.3, 2 (6) 0.02, 2 (7) -2.4, 2 (8) -1, 8 (9) -1.7, 2 (10) -2, 2; right panel: $r = 0.64$ (1) 1.87, 16 (2) -0.8, 2 (3) 5.4, 8 (4) -0.3, 0 (5) -2.7, 0 (6) -0.5, 0; correlation coefficient, T cell and SLEDAI scores in each patient visit).

Table 1

Demographic characteristics of study participants.

	SLE patients	controls	RA patients	p values ^a
Samples	65	27	16	
Age	mean	41	54	NS
	SD	12	17	NS
Gender (%)	female	94	100	NS
	male	6	0	NS
Race (%)	white	49	88	NS
	black	35	12	NS
	asian	11	0	NS
	mixed	5	0	NS

^a comparison of SLE patients to controls.

Table 2

Clinical information of cross-sectional study participants.

Total SLE patients	65
SLEDAI , mean (range)	4.9 (0–36)
PGA , mean (range)	0.9 (0–3)
Steroid-free	29 (45%)
Steroid-treated	36 (55%)
Duration of disease, years mean (range)	8.3 (0–43)
Recent-onset (≤ 1 year)	13 (20%)
Past-onset (>1 year)	52 (80%)
Nephritis	16 (25%)
Cytopenia	11 (17%)
Arthritis	9 (14%)
Serositis	8 (12%)
Mucocutaneous	23 (35%)

Table 3

Clinical information of prospectively studied SLE patients.

	Recorded visits	SLEDAI ^a	PGA ^a	Lymphocytes ^a (k/mL)	Platelets ^a (k/mL)	Creatinine ^a (mg/dL)	ESR ^a (mm/h)
Pt1	9	1.9 (0–6)	0.1 (0–1)	0.5 (0.4–0.7)	184 (138–220)	1.2 (0.8–3.8)	70 (60–101)
Pt2	10	1.1 (0–9)	0.3 (0–2)	3.7 (2.7–4.8)	170 (117–196)	0.7 (0.5–0.9)	56 (33–100)
Pt3	10	3.5 (2–8)	0.2 (0–1)	1 (0.5–1.5)	219 (133–260)	0.8 (0.8–0.9)	11 (3–25)
Pt4	8	6 (2–16)	0.6 (0–3)	1.3 (0.6–1.8)	225 (3–312)	0.7 (0.6–0.8)	15 (3–40)
Pt5	8	2.6 (0–11)	0.4 (0–3)	12 (2–40)	347 (273–612)	1.1 (0.8–1.4)	12 (2–40)
Pt6	6	4.3 (0–16)	0.4 (0–2)	1.5 (1.1–2)	370 (320–426)	0.7 (0.6–0.8)	19 (12–34)
Pt7	6	3.7 (2–6)	0.5 (0–3)	0.8 (0.02–1.5)	215 (186–239)	1 (0.8–1.2)	13 (2–29)

^a mean (range).