

RESEARCH ARTICLE

T Cell Transcriptomes Describe Patient Subtypes in Systemic Lupus Erythematosus

Sean J. Bradley^{1*}, Abel Suarez-Fueyo¹, David R. Moss², Vasileios C. Kyttaris¹, George C. Tsokos^{1*}

1 Division of Rheumatology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, United States of America, **2** Department of Anesthesiology, Tufts Medical Center, Boston, Massachusetts, United States of America

* sjbradle@bidmc.harvard.edu (SJB); gtsokos@bidmc.harvard.edu (GCT)



OPEN ACCESS

Citation: Bradley SJ, Suarez-Fueyo A, Moss DR, Kyttaris VC, Tsokos GC (2015) T Cell Transcriptomes Describe Patient Subtypes in Systemic Lupus Erythematosus. PLoS ONE 10(11): e0141171. doi:10.1371/journal.pone.0141171

Editor: Masataka Kuwana, Nippon Medical School Graduate School of Medicine, JAPAN

Received: July 2, 2015

Accepted: October 3, 2015

Published: November 6, 2015

Copyright: © 2015 Bradley et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The raw data have been deposited at the NCBI Sequence Read Archive under Bioproject Accession ID PRJNA293549.

Funding: This work was funded by the United States National Institutes for Health (<http://www.nih.gov>), with grant number R01AI042269 to GCT (Lymphocyte Signaling Defects in Patients with Lupus), R01AR060849 to VCK (The role of interleukin 23 in lupus), and SJB is funded under training grant T32 AI074549. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Background

T cells regulate the adaptive immune response and have altered function in autoimmunity. Systemic Lupus Erythematosus (SLE) has great diversity of presentation and treatment response. Peripheral blood component gene expression affords an efficient platform to investigate SLE immune dysfunction and help guide diagnostic biomarker development for patient stratification.

Methods

Gene expression in peripheral blood T cell samples for 14 SLE patients and 4 controls was analyzed by high depth sequencing. Unbiased clustering of genes and samples revealed novel patterns related to disease etiology. Functional annotation of these genes highlights pathways and protein domains involved in SLE manifestation.

Results

We found transcripts for hundreds of genes consistently altered in SLE T cell samples, for which DAVID analysis highlights induction of pathways related to mitochondria, nucleotide metabolism and DNA replication. Fewer genes had reduced mRNA expression, and these were linked to signaling, splicing and transcriptional activity. Gene signatures associated with the presence of dsDNA antibodies, low complement levels and nephritis were detected. T cell gene expression also indicates the presence of several patient subtypes, such as having only a minimal expression phenotype, male type, or severe with or without induction of genes related to membrane protein production.

Conclusions

Unbiased transcriptome analysis of a peripheral blood component provides insight on autoimmune pathophysiology and patient variability. We present an open source workflow and richly annotated dataset to support investigation of T cell biology, develop biomarkers for patient stratification and perhaps help indicate a source of SLE immune dysfunction.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: SLE, Systemic Lupus Erythematosus; PBMC, peripheral blood mononuclear cells; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; FPKM, Fragments Per Kilobase of transcript per Million mapped reads; FDR, False Discovery Rate; dsDNA, double-stranded DNA.

Background

Systemic Lupus Erythematosus (SLE) is a debilitating autoimmune disease affecting primarily women. It involves dysregulation of T and B cells resulting in excessive production of antibodies against self proteins and DNA, immune complex formation and T cell infiltration into tissues. These processes cause a variety of symptoms including arthritis, cytopenia and kidney failure. The etiologic origins of sporadic SLE are unknown, but altered regulation of T cells is well documented [1–3]. Genetic determinates of SLE severity have been elusive in part because of the heterogeneity that marks the disease [4, 5], with the majority of cases caused by genetic predisposition coupled with environmental causes. SLE T cells present a poised activation phenotype associated with lower TCR activation threshold, lipid raft aggregation, increased calcium flux upon activation, and overproduction of inflammatory cytokines. Altered gene expression usually accompanies these functional alterations [6].

Expression signatures in SLE have been addressed primarily in the peripheral blood compartment, where pioneering work by the Pascual group first described the interferon signature [7] [8]. These genes are inducible by the cytokine *in vitro* and have since been subdivided as being targets of type I or II interferon [9]. Many of these are simultaneously induced in subsets of cells including T and B cells [10] and monocytes [11] providing evidence for shared signaling abnormalities in peripheral blood mononuclear cells.

We assayed steady-state mRNA abundance by sequencing to discover molecular underpinnings of T cell dysfunction in SLE. Alterations in expression reveal patient subtypes marked by induction of genes involved in protein folding on the endoplasmic reticulum, high levels of ribosomal protein genes, or the previously identified interferon signature alone. Substantial differences in T cell expression in men and women were also found. Highlighted genes could represent biomarkers informative for disease management and may also direct investigation into other T-cell driven autoimmune conditions. This methodology is amenable to study of any disease with great variability of symptom presentation if highly relevant tissue can be obtained for transcriptome sequencing.

Materials and Methods

Sample Collection

At least 5ml of peripheral blood was collected to Lithium Heparin BD vacutainers from 14 SLE patients under treatment at the Lupus Center at the Rheumatology Division of Beth Israel Deaconess Medical Center. All participating patients fulfilled the American College of Rheumatology criteria for the diagnosis of SLE [12]. Blood was similarly obtained from 4 similarly aged healthy female controls. This study was approved by the Institutional Review Board of Beth Israel Deaconess Medical Center. 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.

Cell extraction and RNA isolation

Rosette Sep T cell Purification (StemCell technologies, Vancouver, Canada) was employed as instructed by incubation of blood for 30 min with tetrameric antibody mixture against CD14, CD19, CD20/MS4A1, CD36, CD56, CD66b, CD123, GYPA, and CD16/FCGR3A which binds non-T cells to erythrocytes. Density-gradient centrifugation with Lymphocyte Separation Medium (Cellgro, Manassas, VA) was used to isolate the unstimulated T cells. T cell purity is routinely >93% by this method as determined by CD3 APC/Cy7 HIT3a (Biolegend) staining detected on a Beckman Coulter Gallios Cytometer. RNA was then prepared by Qiagen AllPrep

Kit (Valencia, CA) from 3 million T cells with DNase-I treatment. Roughly 2ug of total RNA was submitted to sequencing, and OD260/280 ratios were approximately 2.1.

Sequencing and Analysis

Unstranded cDNA library preparation and sequencing was performed by BGI (Shenzhen, China). Illumina sequencing provided roughly 75e6 paired 90bp reads for each (~ 12GB gzipped data per sample), which were assessed by FastQC and trimmed to allow ~85% mapping to the GRCh37/hg19 assembly by TOPHAT. Expression scores in Fragments Per Kilobase of transcript per Million mapped reads (FPKM) were obtained by CUFFLINKS for the 24262 best annotated genes, including many expressed psuedogenes and noncoding RNAs (S1 Table). For comparisons between different groups of samples, CUFFDIFF2 was used primarily, alongside DESeq2 and nonparametric tests in R, to calculate expression and statistical differences. Heatmaps were generated using median-normalized expression data with Gene-e and NMF clustering was performed on the Genepattern server, both provided by the Broad Institute. Singular Value Decomposition (SVD) was performed at <http://biographserv.com/> and Venn diagrams were created with BioVenn [13]. Online supplemental files contain methods with specific program commands and R scripts which were implemented in R studio (S1 Text).

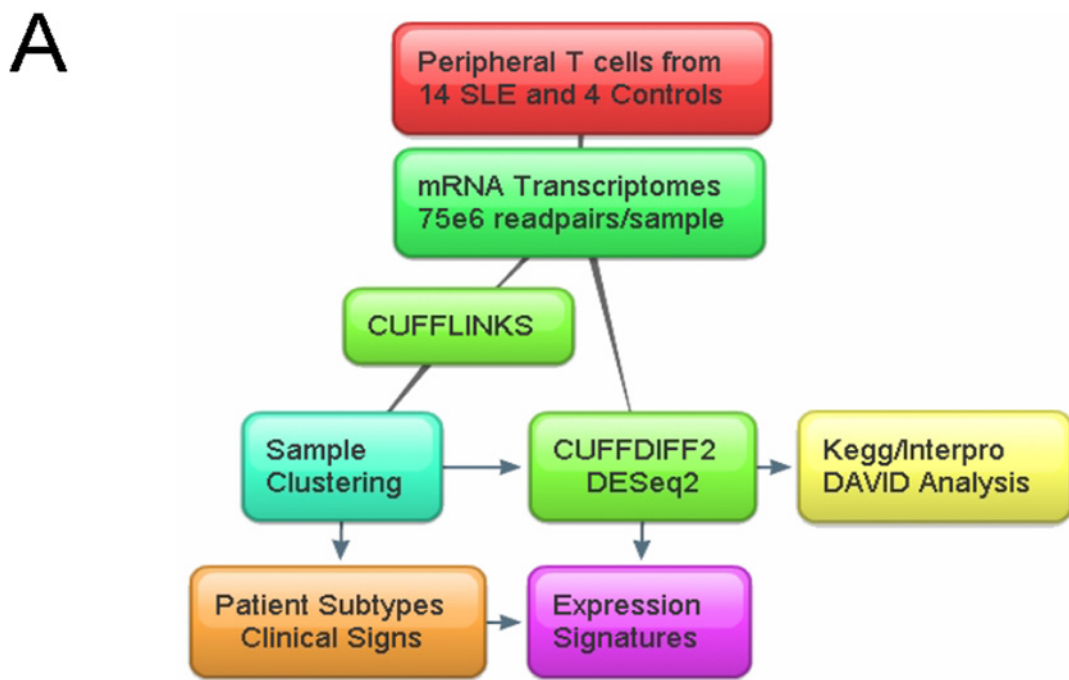
Results

We sequenced mRNA from peripheral T cells in two men and 12 women with a variety of SLE manifestations and SLEDAI disease scores (Fig 1) [14]. As controls we prepared specimens from 4 healthy women aged 25 to 37. None of the patients were receiving therapy with biological agents (Table 1). To evaluate the cell-type purity of the samples we checked non-T cell marker expression. Of the epitopes used to collect T cells by rosette negative selection, only one had expression above background. CD16/FCGR3A, a receptor on NK and T cells [15], had medium expression and was induced in some patients relative to controls. Negligible CD5 and CD19 signal indicated that the preparation was free from B cells and our T cell purity is routinely >93% based on surface CD3 detection. Genes were stratified by expression (average value of all 18 samples) into classes of high, medium, low and unexpressed (≥ 34 , ≥ 11 , ≥ 1 and less than 1 FPKM). Most analysis was carried out on the top quartile of expression in the genome (6047 genes) which included high and medium classes. Among the highest expressed were B2 microglobulin (*B2M*), several thymosins and the expected myriad ribosomal and mitochondrial proteins.

Although total T cells are readily obtained and efficiently purified, variable numbers of constituent cell types impacts transcript abundance. Our pan-T cell view provides breadth and flexibility for study but does not address alterations in cell type frequency, such as lower total lymphocyte counts or lower proportion of CD4 (fraction and absolute amounts) often found in SLE patients. Genes with significant differential expression usually had greater than 2-fold changes in transcript signal, and therefore likely reflect expression changes more so than differences in cell type frequency associated with SLE.

SLE T cells display more genes with increased rather than decreased expression

First we sought an overall picture of mRNA abundance changes related to SLE in T cell samples. Differential expression metrics were found by CUFFDIFF2, which calculates groupwise expression in Fragments Per Kilobase of transcript per Million mapped reads (FPKM) to allow comparison across genes of varied size and applies a beta negative binomial distribution to generate a False Discovery Rate (FDR) where $q < 0.05$ is routinely considered significant [16]. A



B

Increased DNA binding	L137	L149	L140	L031	L133	L074	L078	L115	L027
Low Complement	L137	L149	L140	L031	L072	L133	L074	L115	
Nephritis	L137	L149	L140	L072	L074	L102	L115		
Hematuria	L137	L149	L140	L072	L074	L102			
Proteinuria	L137	L149	L074	L102	L115				
Arthritis	L149	L133	L078	L115	L027				
Flare	L149	L072	L133	L102					
Pyuria	L137	L074							
Urinary Casts	L102								
Leukopenia	L062								
Pleurisy	L074								
Headache	L074								
Vasulitis	L133								
Alopecia	L149								
Rash	L149								
								No Incidence	
								Organic Brain Syndrome	
								Visual Disturbance	
								Cranial Nerve Disorder	
								Psychosis	
								Seizure	
								Mucosal Ulcers	
								Myositis	
								Fever	
								Pericarditis	
								Thrombocytopenia	

Fig 1. T cell transcriptome workflow and distribution of major clinical signs in the SLE patient cohort. A) Samples were obtained from peripheral blood by negative selection and mRNA was sequenced at high depth. CUFFLINKS generated per-sample expression values and CUFFDIFF2 and DESeq2 were used for groupwise comparisons, which were repeated following discovery of novel sample subgroups. **B)** Frequency of SLE symptom presentation at blood draw for patient samples. Highlighting colors based on patient subtypes determined by downstream analysis.

doi:10.1371/journal.pone.0141171.g001

scatterplot shows the distribution of fold change by expression level for genes with a 1.5- or 2-fold difference and those detected as significantly altered (Fig 2A). Count-based methods of

Table 1. Clinical manifestations of the enrolled SLE patients.

ID	L005	L137	L101	L062	L149	L140	L031	L072	L133	L074	L078	L102	L115	L027
Subtype	0	0	0	A	A	A	B	B	B	B	B	C	C	C
SLEDAI	0	16	0	1	20	8	4	6	16	26	6	12	12	6
Race	white	white	white	black	white	asian	white	black	white	black	white	white	asian	white
Sex	F	F	F	F	M	M	F	F	F	F	F	F	F	F
Years	55	39	40	32	33	30	43	39	51	31	61	24	28	48
Low Complement		yes			yes	yes	yes	yes	yes	yes			yes	
Increased DNA binding		yes			yes	yes	yes		yes	yes	yes		yes	yes
Arthritis					yes				yes		yes		yes	
Hematuria		yes			yes	yes		yes		yes		yes		
Proteinuria		yes			yes					yes		yes	yes	
PGA	1	2	0	0	2	0	1	2	3	2	1	2	3	1
C3 (mg/dl)	137	73	99	92	128	89	89	72	91	51	91	115	73	90
C4 (mg/dl)	13	7	24	23	7	34	17	7	2	11	11	25	17	18
dsDNA	neg	(1:40)	neg	neg	(1:320)	(1:20)	(1:40)	neg	(1:80)	(1:80)	(1:80)	neg	(1:80)	(1:320)
ESR	20	12	9	15	10	49	38	39	40	120	38	17	65	5
WBC/HPF	1	57	0	0	8	0	3	0	1	6	0	2	3	0
RBC/HPF	1	125	0	0	12	8	0	175	1	28	0	23	1	0
Urine Protein	30	300	0	30	30	30	0	x	0	1566	100	100	245	0
Spot/creatinine	0.8	1.9	0.1	0.2	x	0.2	0.1	0.8	0.3	4.6	1.1	1.9	3.6	x
WBC	6	8	6	3	4	8	6	9	8	8	6	9	13	4
Lymphocyte	1298	1957	1112	1079	954	832	1224	3355	600	1771	1462	335	1454	752
Hemoglobin	12	11	13	11	14	13	13	11	13	12	13	10	10	12
Platelets	146	320	259	191	238	283	336	273	279	312	241	312	352	220
Creatinine	0.6	0.7	0.5	0.5	0.8	1	0.9	0.9	0.6	0.5	0.9	1.2	2.3	0.6
Prednisone (mg)	0	60	0	0	0	10	10	30	20	40	0	40	10	4
Hydroxychloroquine	400	0	400	400	400	200	400	0	0	0	400	400	0	400
Azathioprine	100	0	0	0	0	0	100	0	0	0	0	0	0	0
Mycophenolate	0	0	0	1000	0	1500	0	3000	0	3000	750	0	2500	1500

Increased dsDNA binding (dsDNA), low complement (C3 or C4) or nephritis (hematuria or proteinuria) were the major symptoms most prevalent. SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; PGA, Physician's global assessment; ESR, erythrocyte sedimentation rate; WBC, white blood cell; RBC, red blood cell; HPF, high-power field. Nephritis was confirmed by recent biopsy and patient L078 has minimal change disease deemed unrelated to lupus.

doi:10.1371/journal.pone.0141171.t001

differential expression are reported to be advantageous in some scenarios, so we also used DESeq2 which was usually confirmatory ($p < 0.01$). Alterations were stratified at multiple levels to provide flexibility for downstream analysis, some of which performed better with more input genes having more subtle differential expression. For high- and medium-expression genes (6047 with greatest average expression for all samples) at least twice as many displayed increased rather than reduced mRNA expression relative to controls at all thresholds. Alterations at each threshold were similarly distributed in terms of mRNA abundance. One third of the genes showing more than 2-fold expression changes were statistically significant by CUFF-DIFF2 ($q < 0.05$) but only one third of these passed a comparable threshold in DESeq2 ($p < 0.01$) (Fig 2B).

Previously reported interferon signature genes including *OAS2*, *ISG15*, *UBE2L6*, *IFI35*, *IFI44*, and *STAT1* were detected as significantly induced by both analyses [17]. Among highly-expressed and significantly upregulated genes were *IL2RG* (encoding CD132, the common

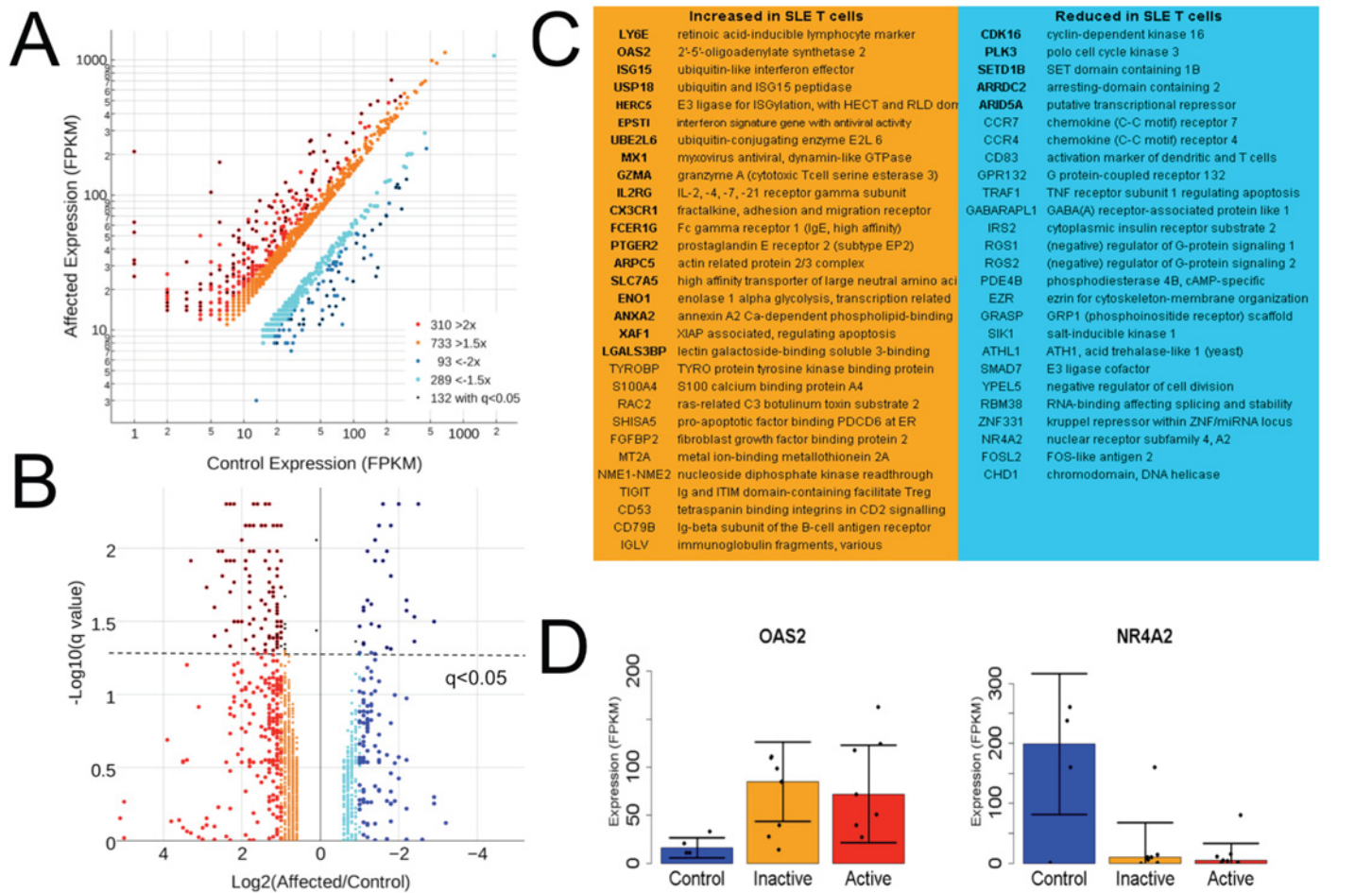


Fig 2. SLE T cells display more genes with increased rather than decreased expression. **A)** Distribution of expression stratified at the 1.5-, 2-fold and $q < 0.05$ CUFFDIFF2 significance levels. **B)** Relationship between q values and expression fold change in SLE relative to control. **C)** Select genes significantly increased or decreased as determined by sequencing and CUFFDIFF2, with those confirmed by DESeq2 in bold. **D)** Example constituent data for OAS2 and NR4A2 in control, inactive and active (SLEDAI > 6) samples, where error bars represent the median absolute deviation about the median.

doi:10.1371/journal.pone.0141171.g002

subunit of receptors for IL-2, -4, -7-15 and -21), CD53, ENO1 and many immunoglobulin fragment transcripts. Select genes with diverse functions are listed, with those in bold found by both CUFFDIFF2 and DESeq2 (Fig 2C). Genes with significantly reduced mRNA abundance included RGS1 and RGS2, which drive G-protein alpha subunits to their inactive state, EZR, which regulates cytoskeleton-membrane interactions, and several nuclear expression regulators. OAS2 and NR4A2 serve as examples of robust alterations, which usually persisted in patients both with low and high (SLEDAI > 6) disease activity (Fig 2D), but were not altered in all patients. Although CCR4 and CCR7 mRNAs were significantly reduced, their surface expression is reported to be increased on SLE T cells [18, 19], suggesting altered post-translational or membrane trafficking regulation for these receptors.

Genes Related to Disease Symptoms

Gene expression markers of SLE symptoms could aid in diagnosis and may point to causative biology. We sought expression changes linked to the presence of increased anti-dsDNA

antibodies, low complement levels, or nephritis. Comparison to controls largely recapitulated the overall SLE analysis, showing similar gene expression related to all major symptoms. This was not surprising given that many patients exhibited more than one symptom. We therefore made comparisons among patient samples with and without each SLE manifestation. Although CUFFDIFF and DESeq analyses suggested genes with substantial fold changes and statistically significant differences in expression, the underlying data revealed great vulnerability to outlier expression. We employed the Mann-Whitney nonparametric rank-sum test, based on group-wise median rather than mean expression values, which yielded genes with statistically different expression in samples with and without each clinical sign.

Samples obtained from patients with increased dsDNA antibodies (titer greater than 1:40 at blood draw) had 579 and 44 genes detected as significant at the $p > 0.05$ and 0.01 levels, respectively (Fig 3A). Most compelling was confirmation of an association with increased expression for *LY6E* [20]. More mRNA for caspase inhibitor *CARD16* and proteasome regulator *PSMF1* was also detected. Reductions were evident for mRNA for *SEMA4D*, which has altered expression in arthritis [21], and *ITPKB*, encoding an inositol phosphate kinase involved in stem cell division [22, 23]. DESeq2 confirmed the association of all but *PSMF1* with this clinical feature ($p < 0.01$), while CUFFDIFF2 confirmed only *LY6E*. None of the genes associated with other symptoms were confirmed by either analysis.

A similar number of gene expression differences uniquely marked samples obtained from patients having low complement (C3 lower than 90 mg/dl or C4 less than 12 mg/dl). Most of the 388 and 46 genes detected at the $p < 0.05$ and < 0.01 levels had increased expression (Fig 3B). Several of these have activity at the endoplasmic reticulum. *SEC11C* encodes a subunit of microsomal signal peptidase complex while peptidyl-prolyl cis/trans isomerases cyclophilin B (*PPIB*) and *FKBP11* both support protein folding. Increases in *BOLA2*, which binds glutaredoxin 3 to regulate iron levels [24] and the poorly characterized *SCAND1* are also good candidate genes for which increased activation marks this clinical phenotype. The function of these genes hints at a disruption of normal activity rather than increased T cell activity.

Patients with nephritis (confirmed by recent biopsy and usually coincident with proteinuria or hematuria) showed fewer genes with altered expression (58 or 2 for $p < 0.05$ or 0.01) and there were no obvious links between them (Fig 3C). The coefficient of variability for altered genes ($p < 0.05$) was greater for these samples (1.2 versus 0.9 and 1.1 for DNA antibodies and low complement, respectively). Patient L078 biopsy and electron microscopy indicated minimal change disease unrelated to lupus, and exclusion from the non-nephritis group had no effect on the detected signature genes because non-parametric tests are only mildly affected by single sample values. One striking marker was *TNFRSF14/HVEM*, encoding a coreceptor for herpes virus which transduces immunosuppressive signals from BTLA [25]. Although this mRNA was marginally increased in SLE patients relative to controls, patients with nephritis had significantly lower amounts relative to those without. A similar pattern was found for many genes increased in SLE, including *OAS2* and *MX1*, which may indicate progression of disease beyond functional immune signaling within T cells and on to response to renal breakdown. Other increased mRNAs linked to the presence of nephritis were *C1orf86/FAAP20*, coding for a DNA repair factor [26], and *LINC00339*, an uncharacterized noncoding RNA. *FOS* mRNA was strikingly increased. This member of the AP1 transcription complex has numerous immune roles, and mRNA for its paralog *FOSL2* was among those significantly decreased overall in patient samples.

Although we detected mRNAs marking the presence of major clinical signs of SLE, the fold changes were less than expected and the genes involved did not suggest a clear picture of the relationship between T cell expression and cause of the symptoms. This could be due to the small size of our cohort and the fact that multiple overlapping symptoms were present in several patients.

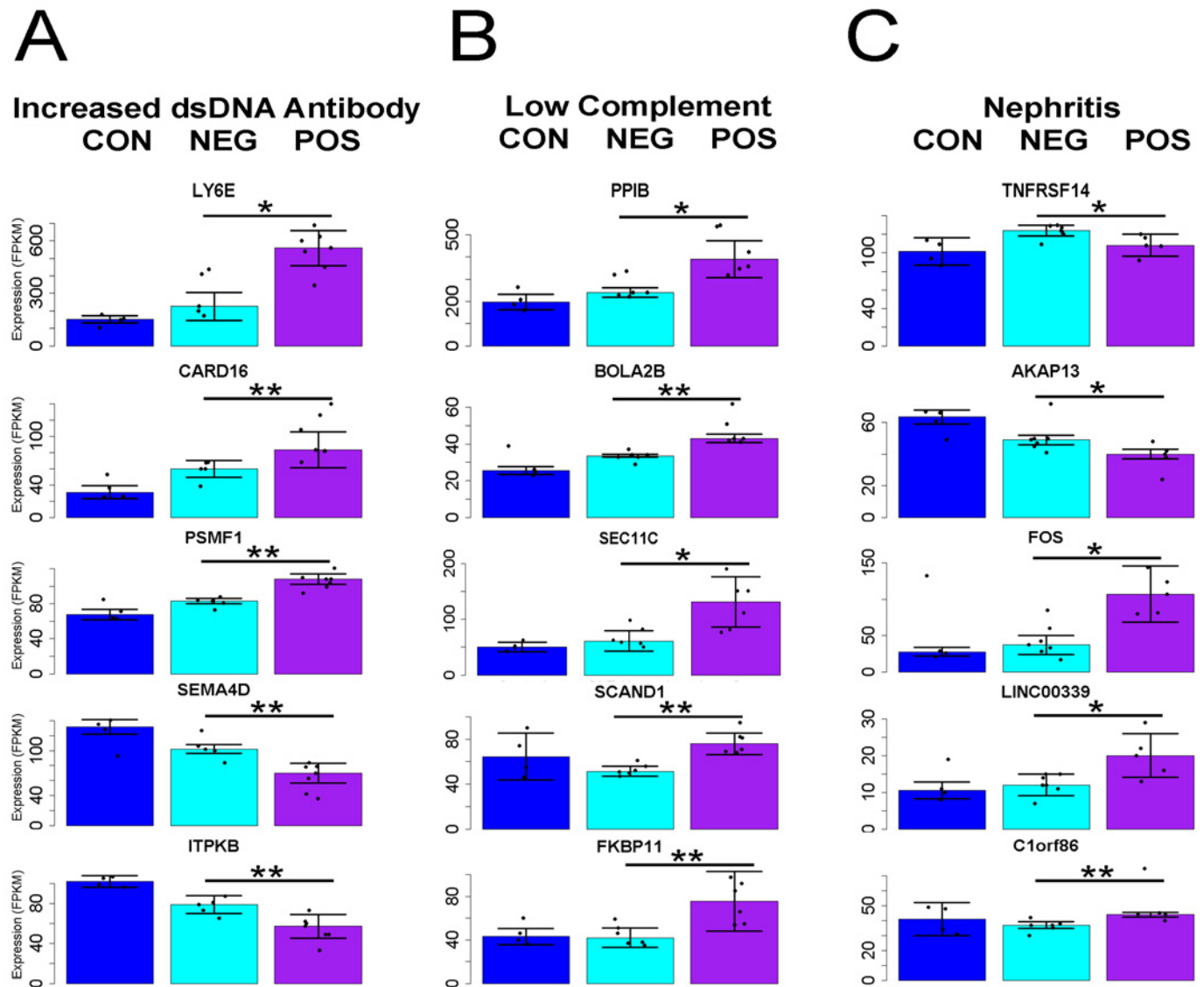


Fig 3. T cell expression of specific genes linked to major clinical manifestations of SLE. Select genes with differential expression in women largely unique to each symptom were detected by Mann-Whitney rank sum tests related to **A)** increased presence of dsDNA antibodies **B)** Low C3 or C4 complement levels or **C)** biopsy-confirmed lupus nephritis. Error bars represent median absolute deviation from the median value for each group and * $p < 0.05$, ** $p < 0.01$ for pairwise tests conducted on samples from patients with or without each symptom (healthy control data plotted only for comparison).

doi:10.1371/journal.pone.0141171.g003

Novel Patient Subtypes are Detected by T cell Expression Clustering

Next, we looked for expression patterns among all patients that might uncover subgroups. We applied unbiased clustering (Pearson) to organize samples by expression similarity in a heat-map, first using genes with at least 1.5-fold expression changes in SLE versus controls. Initial clustering was skewed by outlying expression in single samples. Though potentially interesting they disrupt visualization of groups with coherent behavior. To purify the data we filtered for genes with coefficients of variation (Standard deviation divided by average) between 2 and 0.3, which yielded roughly 1000 genes.

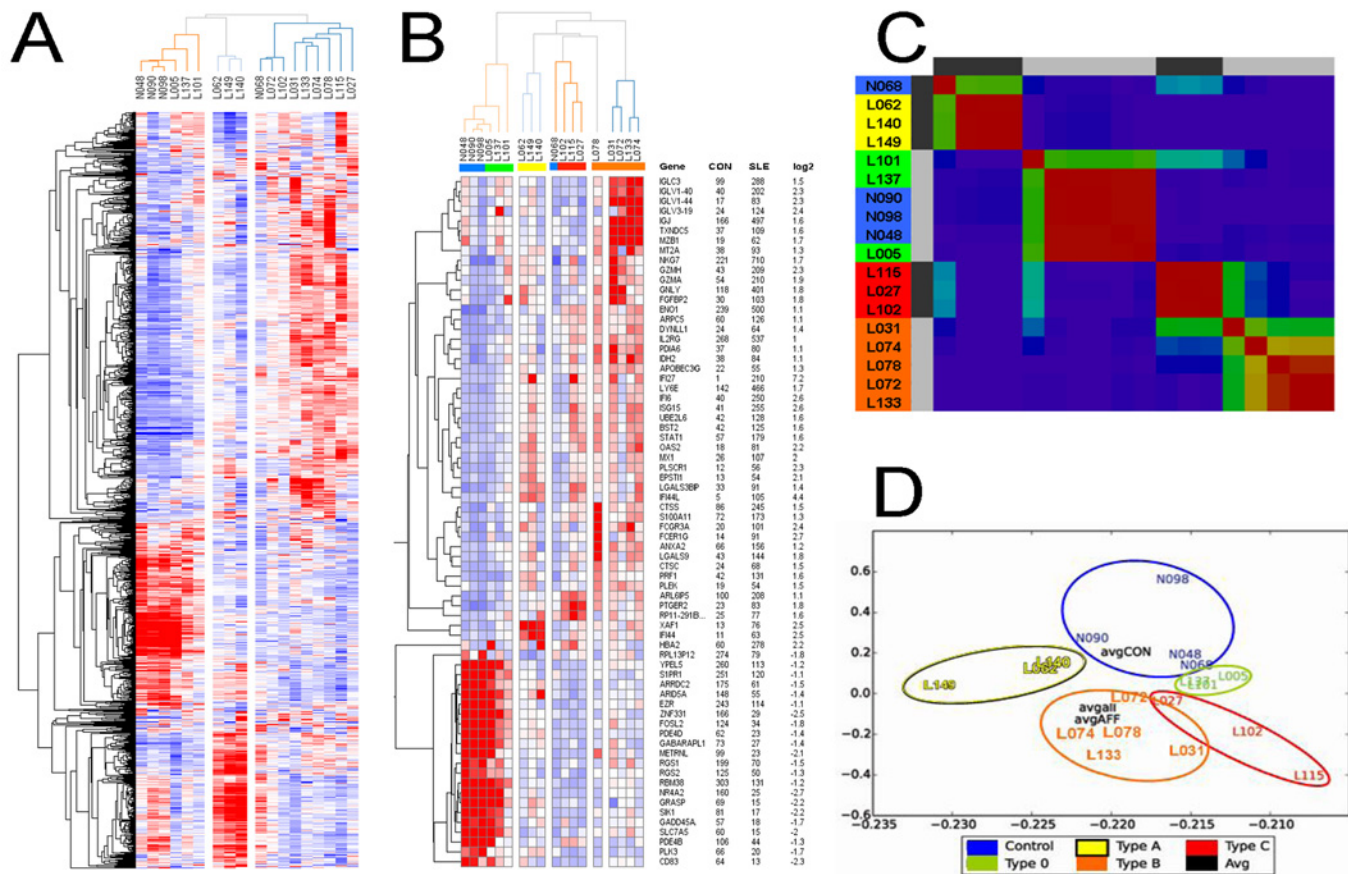


Fig 4. Unbiased clustering identifies patient subtypes by T cell gene expression. **A)** Genes with altered expression in lupus T cells relative to controls, which also showed moderate variation across all samples, were median normalized and Pearson clustered with average linkage, where red and blue denote high and low expression. **B)** Removal of genes with expression outside of the top quartile permits identification of subtype signature genes. **C)** Unbiased NMF clustering using the same input genes yields similar patient subgroups. **D)** SVD clustering of samples provides an approximate metric of sample similarity, where average values for control, affected and all samples are centrally located.

doi:10.1371/journal.pone.0141171.g004

Immediately obvious was that three SLE samples cluster with the controls, indicating similar T cell gene expression (Fig 4A). These samples exhibit only a minor T cell expression phenotype we term Type 0. Two of these patients had low disease activity (SLEDAI 0), but L137 had a higher disease score so its control-like expression pattern is surprising, and may be explained by high dose prednisone treatment. Two other sample groups were delineated by high expression for two different sets of genes (red in the bottom middle for three sample columns, or more diffusely in the upper right for nine samples). The middle group contains samples from two men with active disease and a woman (L062) with SLEDAI 1, which we denote Type A. This unbiased organization of samples is striking because active and inactive samples cluster together, uncoupling disease score from T cell expression in some cases.

To view modules with common patterns we repeated clustering following removal of entries with lower than top quartile expression. This made additional sample types apparent among those from SLE patients (Fig 4B). At the top of this heatmap is a module of genes induced in four patients with increased immunoglobulin fragment mRNAs as well as two genes whose products act at the endoplasmic reticulum, peptide isomerase chaperone *TXNDC5* [27] and *MZB1*, which promotes IgM assembly. Another module specific to three patients was marked

by induction of *ARL6IP5/JWA*, an ROS-sensitive ER protein involved in DNA repair, prostaglandin receptor *PTGER2* and an expressed pseudogene of ribosomal protein *RP11*. We denote these sample groups as Type B and C, and they were similar both in the cohort of genes and extent of expression change outside of these striking modules. They present a more severe expression phenotype than the other sample groups, but most of the alterations are detectable to a minor degree in Type 0 patient samples. In this view, outlier expression for L078 is readily identified. The induction of several genes unique to individual samples was pervasive in our cohort, and may hint at a common disease mechanism linked to transcriptional regulatory failure.

Membership in each group was somewhat dependent on the algorithm and expression thresholds employed, but the four types of patient samples persisted across clustering schemes. We used other methods to verify the tendency to form these groups. Non-negative matrix factorization (NMF) largely recapitulated the Pearson clustering (Fig 4C) and helped confirm that L078 was most similar to Type B, although it represents an edge case. Singular Value Decomposition (SVD) organization of the samples provided additional evidence for structured similarity, where Type 0 grouped with the controls, Type A is quite separate, and Types B and C are closer together (Fig 4D). In each case Type 0 samples were positioned between the controls and Types B and C, indicating an intermediary or perhaps transitional expression phenotype. We repeated CUFFDIFF analysis to find genes altered in each sample subtype with regard to the controls (S1A Fig) and found more alterations linked to Type B and C, consistent with the segregation resulting from unbiased clustering. Most of the altered expression found in Type 0 was also detected in Type B and C samples. Type B had a greater number of genes that were different from all other sample types (S1B Fig). The clinical data associated with these sample groups did not show striking patterns, other than the fact that Type B and C samples were obtained from patients with more symptoms and higher disease scores (orange and red highlighting, Table 1).

We next sought expression markers capable of partitioning samples into the patient subgroups. CUFFDIFF and DESeq comparisons yielded largely overlapping lists of genes, similar to our findings related to potential mRNA biomarkers of SLE clinical manifestation. Although many genes have significant expression alterations in SLE T cells, most are driven by differences present in less than half of the patients in our cohort due to the relatively mild expression phenotype samples in subgroup 0 and A. We applied the nonparametric Kruskal–Wallis and Dunn tests, which allows for comparison between multiple groups. Putative markers were then prioritized by specificity and high expression (Fig 5). An exception was Type 0 samples, for which *LY6E* and *NME1-NME2* expression was chosen on account of the induction present in all subtypes relative to controls. Type A samples exhibited high levels of *DDX17* and *ZAP70* mRNA, while type B samples show increased *MZB1* and *TXNDC5* are expected to best differentiate them from type C samples. Type C samples were less obviously marked but displayed higher levels of *PTGER2* and *ARL6IP5* mRNA compared to Type A.

Immunosuppression is a serious confounder of human autoimmunity studies, so we also looked for expression differences related to prednisone use. CUFFDIFF and DESeq detected relatively few significant alterations (88 and 4 respectively at $q < 0.05$ and $p < 0.01$), for which only *TXNDC5* and *MZB1* had been highlighted as related to patient status or subtype. We next looked among the 324 genes with at least 1.5-fold expression change in patients with or without prednisone administration for genes of interest in other comparisons. This more liberal view revealed *FOS* and *LINC00339* (induction linked to nephritis), as well as *MZB1* and *TXNDC5* (induced in Type B patient samples), as increased in patients under steroid treatment. Mann-Whitney analysis detected 139 genes at the $p < 0.05$ threshold, of which only *FOS*, *LINC00339*, and *C1orf86* had been previously noted (all increased in nephritis). These results indicate that

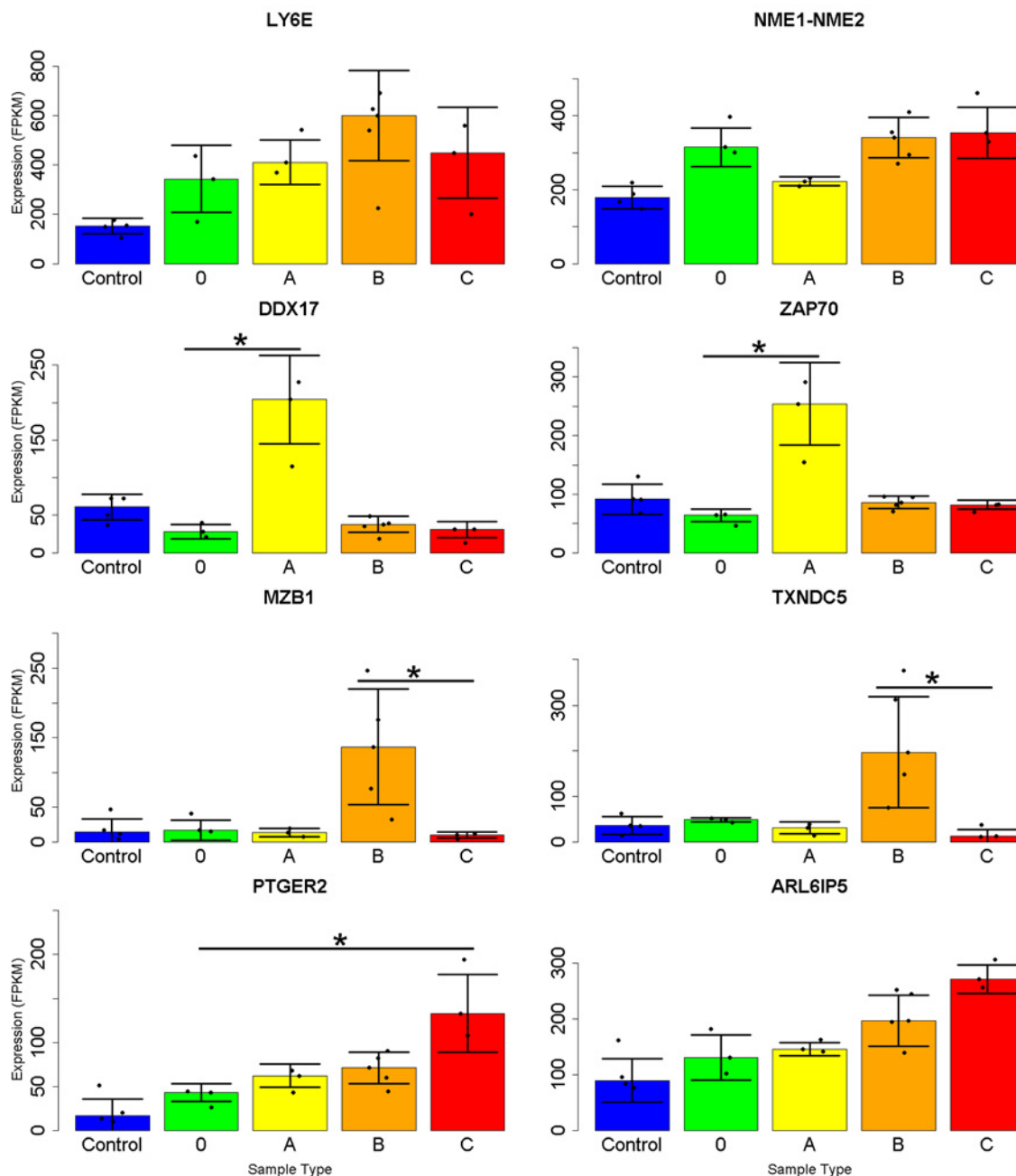


Fig 5. Marker gene expression suggested to identify SLE patient subtypes. Genes were selected based on their ability to differentiate first all SLE samples from controls (top) and then subtypes A or B from the others. Although not all of the genes selected had statistically different expression from all other groups, their use in concert is expected to be sufficient for stratification. Error bars represent median absolute deviation from the median value for each group, and * signifies $p < 0.05$ by Kruskal-Wallis rank-sum followed by Bonferroni-corrected Dunn post test.

doi:10.1371/journal.pone.0141171.g005

prednisone may underlie expression that we found related to nephritis or Type B status. There is considerable overlap for these attributes in our cohort, especially for nephritis and prednisone use. However, based on relatively high expression of Type B markers even for patient L078, who was steroid-free, we expect confounding effects to be ruled out in future studies. We

suspect patient L137 to be present in the mild phenotype subgroup on account of high dose prednisone treatment.

Functional Annotation detects Pathways and Protein Domains Linked to SLE Expression Changes

DAVID analysis provides a literature-based overview of biological functions related to differential expression [28], where KEGG pathways and INTERPRO protein domains offer concise and non-redundant terminology. We compared SLE samples to controls, and also pooled controls with mild expression phenotype samples (Type 0 and A females) for comparison with grouped Type B and C samples. Patient subtype information strengthened this analysis because the comparison of samples with weak or strong expression phenotypes detected more significant terms, and more genes associated with each, than did the initial control versus SLE analysis (Fig 6).

For genes induced in SLE, significantly related KEGG pathways are readily associated with activated and proliferating immune cells, and included Oxidative Phosphorylation (37 genes), Lysosome (26), Proteasome (19) Antigen Processing (16), Glycolysis (14), N-Glycan biosynthesis (11) and Fatty Acid Elongation (5). Significant INTERPRO terms included Immunoglobulin (33) and NAD(P) Binding (27) domains. Down-regulated genes were associated with signaling and nuclear pathways including Spliceosome (13 genes), FC receptor RI Signaling (8), Erb Signaling (9) and Apoptosis (9) and Circadian Rhythms (4). Genes with reduced expression in SLE were enriched for domains related to signaling and gene expression, including Kinase (31), Zinc-Finger (13), Basic Leucine Zipper (9) and Jumonji Transcription Factor (6) motifs. The genes related to each term are listed in the supplement (S2 Table). Analysis of patient subtype samples did not offer compelling differences from these ontology terms, presumably because fewer genes were specific to each.

Induced Transcription Factors Are Suggested to Regulate Induced Genes

We next looked at ENCODE immunoprecipitation data for factors detected in chromatin that might share responsibility for any observed expression changes. Examination of the body and 3kb flanking regions of genes induced at least 2-fold in SLE T cells revealed thousands of binding events. As the consortium data is derived from various cell lines, we looked at the expression of these binding ChIP targets in our data. Among those with signal at induced loci, nine had greater than 1.5-fold increased mRNA in SLE (Fig 7A), most of which are known to impact lymphocyte development or activity. Minimally characterized in T cells was WRNIP1 (Werner helicase interacting protein 1), an ATP-dependent DNA-binding protein related to DNA repair [29, 30]. As they show increased mRNA and are found at induced loci, these factors likely act as transcriptional activators.

A greater number of transcription/chromatin factors were reduced in expression (Fig 7B), a trend detected by DAVID analysis. Among the 14 ChIP targets with mRNAs reduced at least 1.5-fold, several have little described role in T cells. These include ZNF274, which recruits repressive factors SETDB1 and TRIM28/KAP1 [31], chromatin modifiers CHD1 and 2 [32] and ZBTB7A, which represses glycolytic genes [33].

Because mRNA levels are frequently unchanged for transcription factors directing an expression program, we checked ChIP targets unaltered in SLE and depict those with at least 100 sites at induced loci (Fig 7C). Several top hits are well known to influence T cell biology. Runx3 is critical for thymocyte development [34] and YY1 influences Th2 cytokine production [35]. SMC3 and RAD21 interact with MXI1 (found among ChIP targets with decreased

	CON v SLE (1020)			CON v SLE(F)(1256)			C+0+A(F) v B+C(1278)		
	Count	%	Benjamini	Count	%	Benjamini	Count	%	Benjamini
Increased Expression									
UP-Regulated KEGG TERMS									
Oxidative phosphorylation	29	3.1	2.4E-06	42	3.7	4.1E-13	37	3.2	1.5E-08
Purine metabolism	18	1.9	2.7E-01	20	1.7	3.5E-01	27	2.3	1.1E-02
Lysosome	23	2.5	2.0E-04	25	2.2	4.9E-04	26	2.2	3.8E-04
Proteasome	11	1.2	1.9E-02	16	1.4	8.3E-05	19	1.6	1.2E-06
Natural killer cell mediated cytotoxicity	19	2.1	5.5E-02	19	1.7	2.2E-01	19	1.6	2.3E-01
Antigen processing and presentation	12	1.3	2.7E-01	13	1.1	3.2E-01	16	1.4	4.5E-02
Glycolysis / Gluconeogenesis	9	1.0	3.9E-01	12	1.0	1.1E-01	14	1.2	2.7E-02
DNA replication	8	0.9	1.4E-01	9	0.8	1.1E-01	13	1.1	6.4E-04
Glutathione metabolism	12	1.3	9.2E-03	13	1.1	1.3E-02	12	1.0	4.1E-02
N-Glycan biosynthesis				8	0.7	5.0E-01	11	0.9	4.8E-02
Aminoacyl-tRNA biosynthesis							11	0.9	3.4E-02
Pyruvate metabolism	7	0.8	3.9E-01	9	0.8	1.7E-01	9	0.8	1.6E-01
Fatty acid metabolism							9	0.8	1.6E-01
Citrate cycle (TCA cycle)				12	1.0	3.8E-01	9	0.8	4.6E-02
Pentose phosphate pathway	6	0.6	2.7E-01	8	0.7	5.1E-02	8	0.7	4.6E-02
Mismatch repair							8	0.7	3.8E-02
Fatty acid elongation in mitochondria	4	0.4	1.9E-01	5	0.4	5.0E-02	5	0.4	3.8E-02
UP-Regulated INTERPRO DOMAINS									
Immunoglobulin V-set	33	3.6	2.0E-04	34	3.0	5.2E-03	33	2.8	1.9E-02
NAD(P)-binding domain	20	2.2	2.9E-02	27	2.4	3.4E-04	27	2.3	6.3E-04
EF-Hand type	19	2.1	8.5E-01	25	2.2	4.8E-01	23	2.0	8.4E-01
Thioredoxin fold	21	2.3	4.1E-05	26	2.3	1.0E-06	22	1.9	4.9E-04
Small GTP-binding protein	14	1.5	9.0E-01				20	1.7	5.7E-01
Nucleic acid-binding, OB-fold	10	1.1	1.2E-01	11	1.0	1.3E-01	12	1.0	6.2E-02
Short-chain dehydrogenase/reductase SDR	8	0.9	7.3E-01	9	0.8	6.8E-01	9	0.8	8.1E-01
Proteasome, subunit alpha/beta	4	0.4	8.8E-01	7	0.6	9.2E-02	8	0.7	1.9E-02
Reduced Expression									
DOWN-Regulated KEGG TERMS									
MAPK signaling pathway	11	3.3	2.2E-01	12	3.0	1.9E-01	17	3.0	3.2E-02
Neurotrophin signaling pathway	10	3.0	5.2E-02	9	2.3	1.1E-01	15	2.7	7.1E-04
T cell receptor signaling pathway	8	2.4	6.0E-02	9	2.3	9.8E-02	13	2.3	2.0E-03
Spliceosome	9	2.7	5.1E-02	11	2.8	3.2E-02	13	2.3	6.1E-03
Cell cycle	10	3.0	2.8E-02	9	2.3	8.2E-02	11	2.0	3.1E-02
Chemokine signaling pathway	8	2.4	2.7E-01	9	2.3	2.3E-01	11	2.0	1.5E-01
Chronic myeloid leukemia	7	2.1	4.8E-02	6	1.5	1.8E-01	10	1.8	5.5E-03
ErbB signaling pathway	8	2.4	4.1E-02	7	1.8	1.2E-01	10	1.8	1.1E-02
Apoptosis	5	1.5	2.9E-01				9	1.6	2.9E-02
Fc epsilon RI signaling pathway	6	1.8	1.5E-01	7	1.8	8.6E-02	8	1.4	3.8E-02
Toll-like receptor signaling pathway	6	1.8	2.6E-01	6	1.5	2.5E-01	8	1.4	1.1E-01
Acute myeloid leukemia	4	1.2	3.2E-01				7	1.2	3.7E-02
Phosphatidylinositol signaling system				7	1.8	9.3E-02	7	1.2	9.7E-02
B cell receptor signaling pathway	5	1.5	2.5E-01	5	1.3	2.7E-01	7	1.2	9.3E-02
Fc gamma R-mediated phagocytosis				6	1.5	2.4E-01	7	1.2	1.9E-01
Circadian rhythm	4	1.2	4.3E-02	4	1.0	7.8E-02	5	0.9	7.6E-03
DOWN-Regulated INTERPRO DOMAINS									
Protein kinase, core	15	4.5	6.2E-01	16	4.0	8.1E-01	31	5.5	8.5E-03
Nucleotide-binding, alpha-beta plait	11	3.3	2.3E-01	21	5.3	1.8E-05	19	3.4	8.0E-03
RNA recognition motif, RNP-1	11	3.3	2.6E-01	20	5.1	4.0E-05	18	3.2	1.2E-02
Zinc finger, PHD-type	8	2.4	1.2E-01	8	2.0	2.1E-01	13	2.3	2.8E-03
Helicase, superfamily 1 and 2, ATP-binding	7	2.1	3.2E-01	7	1.8	5.6E-01	13	2.3	8.0E-03
DNA/RNA helicase, C-terminal	7	2.1	3.2E-01	7	1.8	5.6E-01	13	2.3	8.0E-03
Basic-leucine zipper (bZIP) TF	7	2.1	9.3E-02	7	1.8	1.0E-01	9	1.6	1.1E-02
Nuclear hormone receptor, ligand-binding	5	1.5	3.2E-01	5	1.3	5.3E-01	7	1.2	1.1E-01
Cyclin-like F-box							7	1.2	4.0E-01
Transcription factor jumonji	5	1.5	1.4E-01	5	1.3	1.5E-01	6	1.1	3.2E-02
Transcription regulator cyclin	3	0.9	3.2E-01	3	0.8	4.6E-01	5	0.9	6.6E-03

Fig 6. Distinct biological pathways and protein domains are identified following sample clustering. The number of genes with altered expression used for each query is in parentheses for each comparison. Significant terms with Benjamini q values <0.05 are in bold. More terms were detected for Control versus SLE samples by removal of male samples (CON v SLE(F)). Further refinement was obtained by grouping minor expression phenotype samples instead with controls (C+0+A(F) v B+C). Redundant and nonspecific terms were discarded and the remaining were ranked by the number of genes associated.

doi:10.1371/journal.pone.0141171.g006

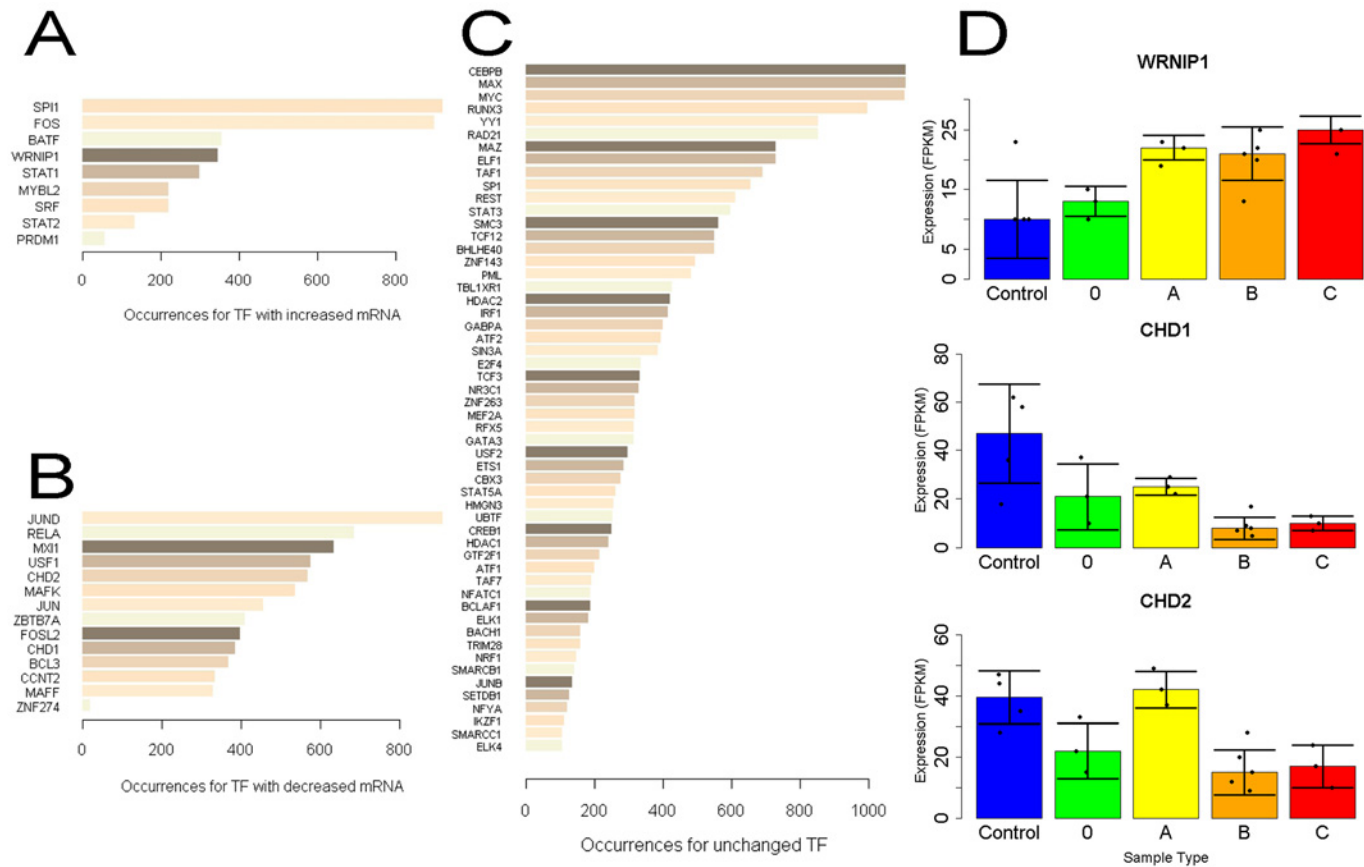


Fig 7. ENCODE ChIP analysis identifies factors which account for induced expression in SLE T cells. The number of chromatin immunoprecipitation binding sites within a 3kb window about significantly induced genes were counted **A**) Factors binding near induced genes which themselves are induced. **B**) Factors binding near these genes having reduced mRNA. **C**) ChIP factors with unchanged mRNA and more than 100 binding sites near induced loci. **D**) Expression of select chromatin factors by sample type, where error bars represent the median absolute deviation from the median.

doi:10.1371/journal.pone.0141171.g007

mRNA) to function in the cohesin complex, and the former is associated with atopic asthma [36]. Expression for these ChIP factors was usually similar in Type B and C samples (Fig 7D).

Conclusions

Patient variability challenges diagnosis and treatment of many diseases, and peripheral blood provides a window into health status capable of reporting on tissues throughout the body. The transcriptome of peripheral blood components show variation with circadian [37] and seasonal [38] periodicity and are growing in descriptive utility in autoimmunity [39] and other clinical settings including transplantation [40] heart failure [41] amyotrophic lateral sclerosis [42] and several cancers [43, 44] where patient subtypes can be identified based on tumor immune cell expression [45]. Expression analysis in disease-relevant tissue is also useful for prioritization of genomic variants [46].

SLE patients present great clinical heterogeneity as a result of genetic diversity and epigenetic changes related to immunological memory. Robust molecular diagnostics have the potential to guide treatment and describe the causes of the disease. Our mRNA analysis identified new genes related to T cell dysfunction and confirmed induction of interferon signature genes (ISG), including *OAS2* which we previously showed is specific to SLE autoimmunity [47].

Patient stratification by ISG expression, however, remains poorly correlated to disease activity [48]. While many of the pathways and domains we found are unsurprising, their notation will aid study of lymphocyte function. Induction of small groups of genes unique to single patients was unexpected, and may prove to be a source of pathological variability related to more common failures of transcriptional repression. Study of nuclear regulators may be most fruitful, in light of the persistent hypomethylation and expression activation displayed SLE T cells [49–51].

We expected patients with various clinical signs to show more distinctive expression patterns, as has been shown for rheumatoid arthritis [52]. Our data indicate that the extent of expression alteration in T cells correlates more with the severity of disease rather than which major symptoms were apparent. Both Type B and C sample groups had more genes with significant expression differences, and higher average SLEDAI scores, than did Type 0 in comparison to controls. We were encouraged that expression analysis detects subtypes of patients. Although these expression phenotypes do not correlate with specific symptoms, they will support patient stratification for study of SLE T cell function. We expect that patient subtypes, at least with regard to sexual dimorphism, will increase discriminatory power and reveal common symptoms or treatment response subsequent analysis of a larger cohort.

Several genes primarily associated with B cells, such as *CD38* and *MZB1*, had striking induction in SLE T cells. This may mark an aberrant or immature cell type that is resistant to negative selection T cell purification. Signals related to B cell activation may mistakenly be received on or within T cells, driving trans- or dedifferentiation to a close lineage member. Expression of immunoglobulin transcript fragments is perplexing, and their profound induction in some patients indicates a potential mechanism of disease, again, perhaps related to a transcriptional or chromatin regulatory lesion. While we saw no evidence of Ig protein, the number of loci and degree of induction for these and several genes with products located on the endoplasmic reticulum offer strong evidence that regulation of antibody production deserves further examination.

We conclude that study of diseases hampered by patient heterogeneity can be supported by high coverage transcriptomics of purified tissues, even in small cohorts. Unbiased detection of patient subtypes and biomarkers associated with symptoms may both be revealed if expression variability is carefully considered. The genes and domains suggested herein will hopefully aid in the study of SLE lymphocyte biology and eventually provide aid to clinical decision making.

Supporting Information

S1 Data. This file contains expression and comparison information extracted from [S1 Table](#) for use as input for analysis in R. (CSV)

S1 Fig. T cell sample types B and C present the most alterations relative to controls. **A)** Gene counts for 1.5- and 2-fold expression changes (FPKM) apparent in various comparisons show that sample Types B and C have the most extreme expression phenotypes relative to controls. The number of samples used as input is listed in parentheses for each comparison. **B)** Overlaps of mRNAs increased or reduced at least 1.5-fold in abundance in three sets of three comparisons. Left, refinement effect on the overall control v SLE analysis. Middle, patient Types B and C show most of the altered genes found in Type 0 in addition to many others. Right, comparisons of clinical signs show greater similarity between increased dsDNA antibody and low complement samples, and that nephritis is accompanied by reductions in many mRNAs. (TIF)

S1 Table. This file contains the gene expression data resulting from the Cufflinks and Cuffdiff analysis, including FPKM and descriptions for the most annotated 24,263 human genes along with metrics from the comparisons performed. The raw sequence data has been deposited at the Sequence Read Archive under Bioproject Accession ID PRJNA293549. (XLS)

S2 Table. This file includes lists of ENSG IDs for all CUFFDIFF comparisons yielding greater than 1.5x mRNA expression changes and the David Analysis of Kegg pathway and Interpro domain terms for the refined SLE versus control comparison. (XLS)

S3 Table. This file includes count-based expression data (matrix_counts) and the comparison matrix inputs for DESeq analysis and an overview of the results. (XLS)

S1 Text. This document contains supplemental methods information including commands and scripts employed for informatics analysis. (DOC)

Acknowledgments

This work was funded by R01AI042269 to GCT, R01AR060849 to VK, and SB is funded under training grant T32 AI074549. We thank members of the Tsokos lab for comments during manuscript development. We also thank James Kozubek, Nikolaus Obholzer and Nathalie Pochet for assistance with initial sequencing analysis, as well as Kenneth Westerman and Kristina Holton at Harvard Research Computing for help with downstream analysis. We are also extremely grateful to all patients participating in our ongoing SLE studies.

Author Contributions

Conceived and designed the experiments: SJB VCK GCT. Performed the experiments: SJB DRM VCK. Analyzed the data: SJB ASF DRM. Contributed reagents/materials/analysis tools: SJB DRM VCK. Wrote the paper: SJB ASF VCK GCT.

References

1. Koga T, Hedrich CM, Mizui M, Yoshida N, Otomo K, Lieberman LA, et al. CaMK4-dependent activation of AKT/mTOR and CREM-alpha underlies autoimmunity-associated Th17 imbalance. *J Clin Invest*. 2014; 124(5):2234–45. Epub 2014/03/29. doi: [10.1172/JCI73411](https://doi.org/10.1172/JCI73411) PMID: [24667640](https://pubmed.ncbi.nlm.nih.gov/24667640/); PubMed Central PMCID: PMC4001553.
2. Kis-Toth K, Tsokos GC. Engagement of SLAMF2/CD48 prolongs the time frame of effective T cell activation by supporting mature dendritic cell survival. *J Immunol*. 2014; 192(9):4436–42. Epub 2014/03/29. doi: [10.4049/jimmunol.1302909](https://doi.org/10.4049/jimmunol.1302909) PMID: [24670806](https://pubmed.ncbi.nlm.nih.gov/24670806/); PubMed Central PMCID: PMC4017928.
3. Mizui M, Koga T, Lieberman LA, Beltran J, Yoshida N, Johnson MC, et al. IL-2 protects lupus-prone mice from multiple end-organ damage by limiting CD4-CD8- IL-17-producing T cells. *J Immunol*. 2014; 193(5):2168–77. Epub 2014/07/27. doi: [10.4049/jimmunol.1400977](https://doi.org/10.4049/jimmunol.1400977) PMID: [25063876](https://pubmed.ncbi.nlm.nih.gov/25063876/); PubMed Central PMCID: PMC4135016.
4. Dai C, Deng Y, Quinlan A, Gaskin F, Tsao BP, Fu SM. Genetics of systemic lupus erythematosus: immune responses and end organ resistance to damage. *Curr Opin Immunol*. 2014; 31:87–96. Epub 2014/12/03. doi: [10.1016/j.coi.2014.10.004](https://doi.org/10.1016/j.coi.2014.10.004) PMID: [25458999](https://pubmed.ncbi.nlm.nih.gov/25458999/); PubMed Central PMCID: PMC4274270.
5. Kariuki SN, Ghodke-Puranik Y, Dorschner JM, Chrobot BS, Kelly JA, Tsao BP, et al. Genetic analysis of the pathogenic molecular sub-phenotype interferon-alpha identifies multiple novel loci involved in systemic lupus erythematosus. *Genes Immun*. 2015; 16(1):15–23. Epub 2014/10/24. doi: [10.1038/gene.2014.57](https://doi.org/10.1038/gene.2014.57) PMID: [25338677](https://pubmed.ncbi.nlm.nih.gov/25338677/); PubMed Central PMCID: PMC4305028.
6. Moulton VR, Tsokos GC. T cell signaling abnormalities contribute to aberrant immune cell function and autoimmunity. *J Clin Invest*. 2015:1–8. Epub 2015/05/12. doi: [10.1172/JCI78087](https://doi.org/10.1172/JCI78087) PMID: [25961450](https://pubmed.ncbi.nlm.nih.gov/25961450/).

7. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med*. 2003; 197(6):711–23. Epub 2003/03/19. doi: [10.1084/jem.20021553](https://doi.org/10.1084/jem.20021553) PMID: [12642603](https://pubmed.ncbi.nlm.nih.gov/12642603/); PubMed Central PMCID: PMC2193846.
8. Chaussabel D, Quinn C, Shen J, Patel P, Glaser C, Baldwin N, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity*. 2008; 29(1):150–64. Epub 2008/07/18. doi: [10.1016/j.immuni.2008.05.012](https://doi.org/10.1016/j.immuni.2008.05.012) PMID: [18631455](https://pubmed.ncbi.nlm.nih.gov/18631455/); PubMed Central PMCID: PMC2727981.
9. Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. *Arthritis Rheumatol*. 2014; 66(6):1583–95. Epub 2014/03/20. doi: [10.1002/art.38628](https://doi.org/10.1002/art.38628) PMID: [24644022](https://pubmed.ncbi.nlm.nih.gov/24644022/); PubMed Central PMCID: PMC4157826.
10. Becker AM, Dao KH, Han BK, Kornu R, Lakhpanpal S, Mobley AB, et al. SLE peripheral blood B cell, T cell and myeloid cell transcriptomes display unique profiles and each subset contributes to the interferon signature. *PLoS One*. 2013; 8(6):e67003. Epub 2013/07/05. doi: [10.1371/journal.pone.0067003](https://doi.org/10.1371/journal.pone.0067003) PMID: [23826184](https://pubmed.ncbi.nlm.nih.gov/23826184/); PubMed Central PMCID: PMC3691135.
11. Shi L, Zhang Z, Yu AM, Wang W, Wei Z, Akhter E, et al. The SLE transcriptome exhibits evidence of chronic endotoxin exposure and has widespread dysregulation of non-coding and coding RNAs. *PLoS One*. 2014; 9(5):e93846. Epub 2014/05/07. doi: [10.1371/journal.pone.0093846](https://doi.org/10.1371/journal.pone.0093846) PMID: [24796678](https://pubmed.ncbi.nlm.nih.gov/24796678/); PubMed Central PMCID: PMC4010412.
12. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum*. 1997; 40(9):1725. Epub 1997/10/27. doi: [10.1002/1529-0131\(199709\)40:9<1725::AID-ART29>3.0.CO;2-Y](https://doi.org/10.1002/1529-0131(199709)40:9<1725::AID-ART29>3.0.CO;2-Y) PMID: [9324032](https://pubmed.ncbi.nlm.nih.gov/9324032/).
13. Hulsen T, de Vlieg J, Alkema W. BioVenn—a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics*. 2008; 9:488. Epub 2008/10/18. doi: [10.1186/1471-2164-9-488](https://doi.org/10.1186/1471-2164-9-488) PMID: [18925949](https://pubmed.ncbi.nlm.nih.gov/18925949/); PubMed Central PMCID: PMC2584113.
14. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. *Arthritis Rheum*. 1992; 35(6):630–40. Epub 1992/06/01. PMID: [1599520](https://pubmed.ncbi.nlm.nih.gov/1599520/).
15. Clemenceau B, Vivien R, Berthome M, Robillard N, Garand R, Gallot G, et al. Effector memory alpha-beta T lymphocytes can express FcγRIIIa and mediate antibody-dependent cellular cytotoxicity. *J Immunol*. 2008; 180(8):5327–34. Epub 2008/04/09. PMID: [18390714](https://pubmed.ncbi.nlm.nih.gov/18390714/).
16. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol*. 2013; 31(1):46–53. Epub 2012/12/12. doi: [10.1038/nbt.2450](https://doi.org/10.1038/nbt.2450) PMID: [23222703](https://pubmed.ncbi.nlm.nih.gov/23222703/); PubMed Central PMCID: PMC3869392.
17. Li QZ, Zhou J, Lian Y, Zhang B, Branch VK, Carr-Johnson F, et al. Interferon signature gene expression is correlated with autoantibody profiles in patients with incomplete lupus syndromes. *Clin Exp Immunol*. 2010; 159(3):281–91. Epub 2009/12/09. doi: [10.1111/j.1365-2249.2009.04057.x](https://doi.org/10.1111/j.1365-2249.2009.04057.x) PMID: [19968664](https://pubmed.ncbi.nlm.nih.gov/19968664/); PubMed Central PMCID: PMC2819494.
18. Yang PT, Kasai H, Zhao LJ, Xiao WG, Tanabe F, Ito M. Increased CCR4 expression on circulating CD4 (+) T cells in ankylosing spondylitis, rheumatoid arthritis and systemic lupus erythematosus. *Clin Exp Immunol*. 2004; 138(2):342–7. Epub 2004/10/23. doi: [10.1111/j.1365-2249.2004.02617.x](https://doi.org/10.1111/j.1365-2249.2004.02617.x) PMID: [15498047](https://pubmed.ncbi.nlm.nih.gov/15498047/); PubMed Central PMCID: PMC1809206.
19. Sen Y, Chunsong H, Baojun H, Linjie Z, Qun L, San J, et al. Aberration of CCR7 CD8 memory T cells from patients with systemic lupus erythematosus: an inducer of T helper type 2 bias of CD4 T cells. *Immunology*. 2004; 112(2):274–89. Epub 2004/05/19. doi: [10.1111/j.1365-2567.2004.01862.x](https://doi.org/10.1111/j.1365-2567.2004.01862.x) PMID: [15147571](https://pubmed.ncbi.nlm.nih.gov/15147571/); PubMed Central PMCID: PMC1782491.
20. Feng X, Wu H, Grossman JM, Hanvivadhanakul P, FitzGerald JD, Park GS, et al. Association of increased interferon-inducible gene expression with disease activity and lupus nephritis in patients with systemic lupus erythematosus. *Arthritis Rheum*. 2006; 54(9):2951–62. Epub 2006/09/02. doi: [10.1002/art.22044](https://doi.org/10.1002/art.22044) PMID: [16947629](https://pubmed.ncbi.nlm.nih.gov/16947629/).
21. Yoshida Y, Ogata A, Kang S, Ebina K, Shi K, Nojima S, et al. Semaphorin 4D Contributes to Rheumatoid Arthritis by Inducing Inflammatory Cytokine Production: Pathogenic and Therapeutic Implications. *Arthritis Rheumatol*. 2015; 67(6):1481–90. Epub 2015/02/25. doi: [10.1002/art.39086](https://doi.org/10.1002/art.39086) PMID: [25707877](https://pubmed.ncbi.nlm.nih.gov/25707877/).
22. Siegemund S, Rigaud S, Conche C, Broaten B, Schaffer L, Westernberg L, et al. IP3 3-kinase B controls hematopoietic stem cell homeostasis and prevents lethal hematopoietic failure in mice. *Blood*. 2015; 125(18):2786–97. Epub 2015/03/20. doi: [10.1182/blood-2014-06-583187](https://doi.org/10.1182/blood-2014-06-583187) PMID: [25788703](https://pubmed.ncbi.nlm.nih.gov/25788703/); PubMed Central PMCID: PMC4416530.
23. Hoofd C, Devreker F, Deneubourg L, Deleu S, Nguyen TM, Sermon K, et al. A specific increase in inositol 1,4,5-trisphosphate 3-kinase B expression upon differentiation of human embryonic stem cells. *Cell Signal*. 2012; 24(7):1461–70. Epub 2012/03/27. doi: [10.1016/j.cellsig.2012.03.006](https://doi.org/10.1016/j.cellsig.2012.03.006) PMID: [22446005](https://pubmed.ncbi.nlm.nih.gov/22446005/).

24. Li H, Mapolelo DT, Randeniya S, Johnson MK, Outten CE. Human glutaredoxin 3 forms [2Fe-2S]-bridged complexes with human BOLA2. *Biochemistry*. 2012; 51(8):1687–96. Epub 2012/02/09. doi: [10.1021/bi2019089](https://doi.org/10.1021/bi2019089) PMID: [22309771](https://pubmed.ncbi.nlm.nih.gov/22309771/); PubMed Central PMCID: [PMC3331715](https://pubmed.ncbi.nlm.nih.gov/PMC3331715/).
25. Derre L, Rivals JP, Jandus C, Pastor S, Rimoldi D, Romero P, et al. BTLA mediates inhibition of human tumor-specific CD8+ T cells that can be partially reversed by vaccination. *J Clin Invest*. 2010; 120(1):157–67. Epub 2009/12/30. doi: [10.1172/JCI40070](https://doi.org/10.1172/JCI40070) PMID: [20038811](https://pubmed.ncbi.nlm.nih.gov/20038811/); PubMed Central PMCID: [PMC2799219](https://pubmed.ncbi.nlm.nih.gov/PMC2799219/).
26. Yan Z, Guo R, Paramasivam M, Shen W, Ling C, Fox D 3rd, et al. A ubiquitin-binding protein, FAAP20, links RNF8-mediated ubiquitination to the Fanconi anemia DNA repair network. *Mol Cell*. 2012; 47(1):61–75. Epub 2012/06/19. doi: [10.1016/j.molcel.2012.05.026](https://doi.org/10.1016/j.molcel.2012.05.026) PMID: [22705371](https://pubmed.ncbi.nlm.nih.gov/22705371/); PubMed Central PMCID: [PMC3398238](https://pubmed.ncbi.nlm.nih.gov/PMC3398238/).
27. Jin Y, Sharma A, Bai S, Davis C, Liu H, Hopkins D, et al. Risk of type 1 diabetes progression in islet autoantibody-positive children can be further stratified using expression patterns of multiple genes implicated in peripheral blood lymphocyte activation and function. *Diabetes*. 2014; 63(7):2506–15. Epub 2014/03/07. doi: [10.2337/db13-1716](https://doi.org/10.2337/db13-1716) PMID: [24595351](https://pubmed.ncbi.nlm.nih.gov/24595351/); PubMed Central PMCID: [PMC4066338](https://pubmed.ncbi.nlm.nih.gov/PMC4066338/).
28. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009; 4(1):44–57. Epub 2009/01/10. doi: [10.1038/nprot.2008.211](https://doi.org/10.1038/nprot.2008.211) PMID: [19131956](https://pubmed.ncbi.nlm.nih.gov/19131956/).
29. Crosetto N, Bienko M, Hibbert RG, Perica T, Ambrogio C, Kensche T, et al. Human Wrip1 is localized in replication factories in a ubiquitin-binding zinc finger-dependent manner. *J Biol Chem*. 2008; 283(50):35173–85. Epub 2008/10/10. doi: [10.1074/jbc.M803219200](https://doi.org/10.1074/jbc.M803219200) PMID: [18842586](https://pubmed.ncbi.nlm.nih.gov/18842586/); PubMed Central PMCID: [PMC3259897](https://pubmed.ncbi.nlm.nih.gov/PMC3259897/).
30. Ramakrishnan R, Liu H, Donahue H, Malovannaya A, Qin J, Rice AP. Identification of novel CDK9 and Cyclin T1-associated protein complexes (CCAPs) whose siRNA depletion enhances HIV-1 Tat function. *Retrovirology*. 2012; 9:90. Epub 2012/11/01. doi: [10.1186/1742-4690-9-90](https://doi.org/10.1186/1742-4690-9-90) PMID: [23110726](https://pubmed.ncbi.nlm.nih.gov/23110726/); PubMed Central PMCID: [PMC3494656](https://pubmed.ncbi.nlm.nih.gov/PMC3494656/).
31. Cruvinel E, Budinetz T, Germain N, Chamberlain S, Lalande M, Martins-Taylor K. Reactivation of maternal SNORD116 cluster via SETDB1 knockdown in Prader-Willi syndrome iPSCs. *Hum Mol Genet*. 2014; 23(17):4674–85. Epub 2014/04/25. doi: [10.1093/hmg/ddu187](https://doi.org/10.1093/hmg/ddu187) PMID: [24760766](https://pubmed.ncbi.nlm.nih.gov/24760766/).
32. Siggins L, Cordeddu L, Ronnerblad M, Lennartsson A, Ekwall K. Transcription-coupled recruitment of human CHD1 and CHD2 influences chromatin accessibility and histone H3 and H3.3 occupancy at active chromatin regions. *Epigenetics Chromatin*. 2015; 8(1):4. Epub 2015/01/27. doi: [10.1186/1756-8935-8-4](https://doi.org/10.1186/1756-8935-8-4) PMID: [25621013](https://pubmed.ncbi.nlm.nih.gov/25621013/); PubMed Central PMCID: [PMC4305392](https://pubmed.ncbi.nlm.nih.gov/PMC4305392/).
33. Liu XS, Haines JE, Mehanna EK, Genet MD, Ben-Sahra I, Asara JM, et al. ZBTB7A acts as a tumor suppressor through the transcriptional repression of glycolysis. *Genes Dev*. 2014; 28(17):1917–28. Epub 2014/09/04. doi: [10.1101/gad.245910.114](https://doi.org/10.1101/gad.245910.114) PMID: [25184678](https://pubmed.ncbi.nlm.nih.gov/25184678/); PubMed Central PMCID: [PMC4197949](https://pubmed.ncbi.nlm.nih.gov/PMC4197949/).
34. Egawa T, Tillman RE, Naoe Y, Taniuchi I, Littman DR. The role of the Runx transcription factors in thymocyte differentiation and in homeostasis of naive T cells. *J Exp Med*. 2007; 204(8):1945–57. Epub 2007/07/25. doi: [10.1084/jem.20070133](https://doi.org/10.1084/jem.20070133) PMID: [17646406](https://pubmed.ncbi.nlm.nih.gov/17646406/); PubMed Central PMCID: [PMC2118679](https://pubmed.ncbi.nlm.nih.gov/PMC2118679/).
35. Hwang SS, Kim YU, Lee S, Jang SW, Kim MK, Koh BH, et al. Transcription factor YY1 is essential for regulation of the Th2 cytokine locus and for Th2 cell differentiation. *Proc Natl Acad Sci U S A*. 2013; 110(1):276–81. Epub 2012/12/19. doi: [10.1073/pnas.1214682110](https://doi.org/10.1073/pnas.1214682110) PMID: [23248301](https://pubmed.ncbi.nlm.nih.gov/23248301/); PubMed Central PMCID: [PMC3538243](https://pubmed.ncbi.nlm.nih.gov/PMC3538243/).
36. Cheng Q, Huang W, Chen N, Shang Y, Zhang H. SMC3 may play an important role in atopic asthma development. *Clin Respir J*. 2014. Epub 2014/12/18. doi: [10.1111/crj.12247](https://doi.org/10.1111/crj.12247) PMID: [25515564](https://pubmed.ncbi.nlm.nih.gov/25515564/).
37. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA, et al. Individuality and variation in gene expression patterns in human blood. *Proc Natl Acad Sci U S A*. 2003; 100(4):1896–901. Epub 2003/02/13. doi: [10.1073/pnas.252784499](https://doi.org/10.1073/pnas.252784499) PMID: [12578971](https://pubmed.ncbi.nlm.nih.gov/12578971/); PubMed Central PMCID: [PMC149930](https://pubmed.ncbi.nlm.nih.gov/PMC149930/).
38. Dopico XC, Evangelou M, Ferreira RC, Guo H, Pekalski ML, Smyth DJ, et al. Widespread seasonal gene expression reveals annual differences in human immunity and physiology. *Nat Commun*. 2015; 6:7000. Epub 2015/05/13. doi: [10.1038/ncomms8000](https://doi.org/10.1038/ncomms8000) PMID: [25965853](https://pubmed.ncbi.nlm.nih.gov/25965853/); PubMed Central PMCID: [PMC4432600](https://pubmed.ncbi.nlm.nih.gov/PMC4432600/).
39. McKinney EF, Lee JC, Jayne DR, Lyons PA, Smith KG. T-cell exhaustion, co-stimulation and clinical outcome in autoimmunity and infection. *Nature*. 2015; 523(7562):612–6. Epub 2015/07/01. doi: [10.1038/nature14468](https://doi.org/10.1038/nature14468) PMID: [26123020](https://pubmed.ncbi.nlm.nih.gov/26123020/).
40. Roedder S, Li L, Alonso MN, Hsieh SC, Vu MT, Dai H, et al. A Three-Gene Assay for Monitoring Immune Quiescence in Kidney Transplantation. *J Am Soc Nephrol*. 2014. Epub 2014/11/28. doi: [10.1681/ASN.2013111239](https://doi.org/10.1681/ASN.2013111239) PMID: [25429124](https://pubmed.ncbi.nlm.nih.gov/25429124/).

41. Maciejak A, Kiliszek M, Michalak M, Tulacz D, Opolski G, Matlak K, et al. Gene expression profiling reveals potential prognostic biomarkers associated with the progression of heart failure. *Genome Med.* 2015; 7(1):26. Epub 2015/05/20. doi: [10.1186/s13073-015-0149-z](https://doi.org/10.1186/s13073-015-0149-z) PMID: [25984239](https://pubmed.ncbi.nlm.nih.gov/25984239/); PubMed Central PMCID: PMC4432772.
42. Ladd AC, Keeney PM, Govind MM, Bennett JP Jr. Mitochondrial oxidative phosphorylation transcriptome alterations in human amyotrophic lateral sclerosis spinal cord and blood. *Neuromolecular Med.* 2014; 16(4):714–26. Epub 2014/08/02. doi: [10.1007/s12017-014-8321-y](https://doi.org/10.1007/s12017-014-8321-y) PMID: [25081190](https://pubmed.ncbi.nlm.nih.gov/25081190/).
43. Liong ML, Lim CR, Yang H, Chao S, Bong CW, Leong WS, et al. Blood-based biomarkers of aggressive prostate cancer. *PLoS One.* 2012; 7(9):e45802. Epub 2012/10/17. doi: [10.1371/journal.pone.0045802](https://doi.org/10.1371/journal.pone.0045802) PMID: [23071848](https://pubmed.ncbi.nlm.nih.gov/23071848/); PubMed Central PMCID: PMC3461021.
44. Shi M, Chen MS, Sekar K, Tan CK, Ooi LL, Hui KM. A blood-based three-gene signature for the non-invasive detection of early human hepatocellular carcinoma. *Eur J Cancer.* 2014; 50(5):928–36. Epub 2013/12/18. doi: [10.1016/j.ejca.2013.11.026](https://doi.org/10.1016/j.ejca.2013.11.026) PMID: [24332572](https://pubmed.ncbi.nlm.nih.gov/24332572/).
45. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity.* 2013; 39(4):782–95. Epub 2013/10/22. doi: [10.1016/j.immuni.2013.10.003](https://doi.org/10.1016/j.immuni.2013.10.003) PMID: [24138885](https://pubmed.ncbi.nlm.nih.gov/24138885/).
46. Tsalik EL, Langley RJ, Dinwiddie DL, Miller NA, Yoo B, van Velkinburgh JC, et al. An integrated transcriptome and expressed variant analysis of sepsis survival and death. *Genome Med.* 2014; 6(11):111. Epub 2014/12/30. doi: [10.1186/s13073-014-0111-5](https://doi.org/10.1186/s13073-014-0111-5) PMID: [25538794](https://pubmed.ncbi.nlm.nih.gov/25538794/); PubMed Central PMCID: PMC4274761.
47. Grammatikos AP, Kyttaris VC, Kis-Toth K, Fitzgerald LM, Devlin A, Finnell MD, et al. A T cell gene expression panel for the diagnosis and monitoring of disease activity in patients with systemic lupus erythematosus. *Clin Immunol.* 2014; 150(2):192–200. Epub 2014/01/18. doi: [10.1016/j.clim.2013.12.002](https://doi.org/10.1016/j.clim.2013.12.002) PMID: [24434273](https://pubmed.ncbi.nlm.nih.gov/24434273/); PubMed Central PMCID: PMC3932542.
48. Kennedy WP, Maciuga R, Wolslegel K, Tew W, Abbas AR, Chaivorapol C, et al. Association of the interferon signature metric with serological disease manifestations but not global activity scores in multiple cohorts of patients with SLE. *Lupus Sci Med.* 2015; 2(1):e000080. Epub 2015/04/11. doi: [10.1136/lupus-2014-000080](https://doi.org/10.1136/lupus-2014-000080) PMID: [25861459](https://pubmed.ncbi.nlm.nih.gov/25861459/); PubMed Central PMCID: PMC4379884.
49. Absher DM, Li X, Waite LL, Gibson A, Roberts K, Edberg J, et al. Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. *PLoS Genet.* 2013; 9(8):e1003678. Epub 2013/08/21. doi: [10.1371/journal.pgen.1003678](https://doi.org/10.1371/journal.pgen.1003678) PMID: [23950730](https://pubmed.ncbi.nlm.nih.gov/23950730/); PubMed Central PMCID: PMC3738443.
50. Hedrich CM, Rauen T, Apostolidis SA, Grammatikos AP, Rodriguez Rodriguez N, Ioannidis C, et al. Stat3 promotes IL-10 expression in lupus T cells through trans-activation and chromatin remodeling. *Proc Natl Acad Sci U S A.* 2014; 111(37):13457–62. Epub 2014/09/05. doi: [10.1073/pnas.1408023111](https://doi.org/10.1073/pnas.1408023111) PMID: [25187566](https://pubmed.ncbi.nlm.nih.gov/25187566/); PubMed Central PMCID: PMC4169908.
51. Coit P, Renauer P, Jeffries MA, Merrill JT, McCune WJ, Maksimowicz-McKinnon K, et al. Renal involvement in lupus is characterized by unique DNA methylation changes in naive CD4+ T cells. *J Autoimmun.* 2015. Epub 2015/05/26. doi: [10.1016/j.jaut.2015.05.003](https://doi.org/10.1016/j.jaut.2015.05.003) PMID: [26005050](https://pubmed.ncbi.nlm.nih.gov/26005050/).
52. Anderson AE, Pratt AG, Sedhom MA, Doran JP, Routledge C, Hargreaves B, et al. IL-6-driven STAT signalling in circulating CD4+ lymphocytes is a marker for early anticitrullinated peptide antibody-negative rheumatoid arthritis. *Ann Rheum Dis.* 2015. Epub 2015/02/05. doi: [10.1136/annrheumdis-2014-205850](https://doi.org/10.1136/annrheumdis-2014-205850) PMID: [25649145](https://pubmed.ncbi.nlm.nih.gov/25649145/).