Clinical Epigenetics

RESEARCH

Open Access

Comparison of cell type and disease subset chromatin modifications in SLE

Katherine Beigel[1], Xiao-Min Wang[2], Li Song[2], Kelly Maurer[2], Christopher Breen[2], Deanne Taylor[1,3], Daniel Goldman[4], Michelle Petri[4] and Kathleen E. Sullivan[2*]

Abstract

Background Systemic lupus erythematosus (SLE) is an autoimmune disease with protean manifestations. There is little understanding of why some organs are specifically impacted in patients and the mechanisms of disease persistence remain unclear. While much work has been done characterizing the DNA methylation status in SLE, there is less information on histone modifications, a more dynamic epigenetic feature. This study identifies two histone marks of activation and the binding of p300 genome-wide in three cell types and three clinical subsets to better understand cell-specific effects and differences across clinical subsets.

Results We examined 20 patients with SLE and 8 controls and found that individual chromatin marks varied considerably across T cells, B cells, and monocytes. When pathways were examined, there was far more concordance with conservation of TNF, IL-2/STAT5, and KRAS pathways across multiple cell types and ChIP data sets. Patients with cutaneous lupus and lupus nephritis generally had less dramatically altered chromatin than the general SLE group. Signals also demonstrated significant overlap with GWAS signals in a manner that did not implicate one cell type more than the others.

Conclusions The pathways identified by altered histone modifications and p300 binding are pathways known to be important from RNA expression studies and recognized pathogenic mechanisms of disease. NFκB and classical inflammatory pathways were strongly associated with increased peak heights across all cell types but were the highest-ranking pathway for all three antibodies in monocytes according to *fgsea* analysis. IL-6 Jak/STAT3 signaling was the most significant pathway association in T cells marked by H3K27ac change. Therefore, each cell type experiences the disease process distinctly although in all cases there was a strong theme of classical inflammatory pathways. Importantly, this NFκB pathway, so strongly implicated in the patients with generalized SLE, was much less impacted in monocytes when cutaneous lupus was compared to the general SLE cohort and also less impacted in lupus nephritis compared to general SLE. These studies define important cell type differences and emphasize the breadth of the inflammatory effects in SLE.

Keywords Lupus, p300, H3K27ac, Rheumatoid arthritis

*Correspondence:
Kathleen E. Sullivan
sullivank@chop.edu
Full list of author information is available at the end of the article

Background

Systemic lupus erythematosus (SLE) is a chronic inflammatory disease affecting predominantly young women during their peak productive years. Typical manifestations of SLE include a photosensitive rash, glomerulonephritis, arthritis, and serositis [1]. Autoantibodies represent the best biomarker of disease subsets [2], however, these are imperfect and are not accurate for the prediction of course or severity [3, 4].

Gene expression and epigenetic studies have increasingly been utilized to better understand the drivers of disease course [5–7]. Starting with DNA methylation, studies performed on stored DNA samples revealed a globally hypomethylated genome in SLE [8]. Interferon-responsive genes have been consistently identified as demethylated in CD4 T cells [9, 10] with hypermethylation observed in B cells [11]. Mechanistically, a subset of changes have been traced back to over-expression of EZH2 in T cells in SLE [12]. DNA hypomethylation has also been observed in neutrophils from people with SLE, with particular enrichment of interferon-regulated genes. The hypomethylation was largely stable over time although there were differences related to race [13].

Less work has been done to characterize histone modifications in SLE, another important epigenetic mirror of disease state. Our data have demonstrated a clearly altered histone landscape in SLE patients with overall increased H4ac in patient monocytes with networks of NFκB and IRF1. Genes with both high H4ac and high gene expression had a signal for antiviral pathways, characteristic of interferon responses [14]. 28% of the SLE-related gene expression differences in monocytes could be replicated by treatment with α2 interferon, one of the type I interferons. In addition to altered peak height for H3K4me3 in SLE, peak shape was altered with the broadest peaks in H3K4me3 associated with the highest likelihood of increased transcription in SLE [15]. Monocyte enhancers with altered peak height in SLE were enriched in transcription factor bindings sites for IRF1, Pu.1, SP1, IRF4, IRF8, STAT1, and STAT6 [16]. When studies were performed expanding to T cells, B cells, and monocytes, only ribosome biogenesis crossed all three cell types as a conserved signature for H3K4me3 [17]. Both transcripts and H3K4me3 overlapped known GWAS signals but each cell type had largely its own histone signature in SLE. Other groups have focused on histone acetylation, identifying hypoacetylated H3 and H4 in T cells and B cells from patients with SLE [18]. This study was undertaken to examine whether histone modification changes in SLE are related to SLE phenotype and to validate our previous findings across the three cell types using an expanded analysis of chromatin.

Methods

Cell source and purification

The female healthy donor blood samples were obtained from the Human Immunology Core, under a protocol approved by the University of Pennsylvania Perelman School of Medicine Institutional Review Board (IRB) ("Controls" N = 8). They were self-declared as healthy. The female SLE patient blood samples were from the Johns Hopkins Lupus Cohort [19, 20] under a separate IRB protocol ("SLE" N = 20) and 4 patients with rheumatoid arthritis (RA) were from the same institution under the same IRB ("RA" N = 4). Control and SLE samples were handled identically with simulated shipping for the controls. Peripheral blood mononuclear cells (PBMC) and monocytes were purified as previously described [21]. Dynal beads (Dynabeads® CD3 and CD19, Invitrogen, Oslo, Norway) were used to separate CD3 T cells and CD19 B cells according to the manufacturer's protocol. Flow cytometry was used to ensure purity (> 90%) of the three cell populations (CD14, CD19, CD3).

Chromatin immunoprecipitation

Antibodies for chromatin immunoprecipitation were all validated for ChIP. Sources were H3K27ac from Abcam (4729), and H3K4me3 [39] from Active Motif, and p300 (D2X6N) from Cell Signaling Technologies. Cells purified as above were fixed, lysed and sonicated according to the ChIP-IT High Sensitivity kit instructions (Active Motif). Libraries were prepared using NEBNext Ultra II DNA Library Kit for Illumina.

Bioinformatic analyses

The workflow for ChIP-sequencing analysis, in brief, was as follows: *trimgalore* was used to trim adapters, and *bowtie2* was used to align trimmed reads, and aligned reads were sorted and indexed using *samtools*. Peaks were called using *macs2 callpeak* for narrow peaks. DiffBind was used to filter out ENCODE blacklist regions, count peaks, normalize for library size, and analyze differentially bound peaks between conditions (EdgeR method in DiffBind). Differentially bound sites identified in DiffBind were annotated using ChIPpeakAnno to find the nearest feature. For cross-referencing with GWAS SNPs, the 221 SLE-associated SNPs identified by Farh et al. [22] were examined. Sample read coverage of a 1 kb region (−500 to +500) around each SNP was calculated using *bedtools coverage*. Coverages were log-transformed and normalized by Loess method. Wilcoxon Rank Sum test was used to test significance of the difference in coverage between Lupus and Control groups. The difference in average coverage between groups was plotted for each SNP. Differentially bound genes identified by DiffBind were also cross-referenced with differentially expressed genes from

a previous bulk RNA-seq experiment of lupus and control B cells, T cells, and monocytes [17]. More details on the bioinformatics methods (including software versions) can be found in the Supplemental Methods.

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of our research.

Data access

The data sets are available on the database of Genotypes and Phenotypes (dbGaP) under accession phs003713.v1.p1.

Results

Patient and control data

Comparisons between SLE and Control utilized samples from 8 healthy donors and 7 patients with SLE (without nephritis). The mean age for the SLE group was 49 years with mean age-of-onset of 32 years. Recruited patients were not on B cell depleting therapies or cytotoxic drugs. The autoantibody profile and racial breakdown are given in Supplemental Table 1. Library counts for each analysis are given in Supplemental Table 2. The SLEDAI at the time of sampling was 2.9. Five patients with cutaneous lupus provided samples for a subset analysis and they had a mean age of 47 years with the mean age-of-onset at 26 years of age. Eight patients with lupus nephritis were included in the subset analysis and they had a mean age of 43 years with mean age-of-onset at 23 years of age. Four subjects with RA were used as a comparator group with a different inflammatory disease and they had a mean age of 66 years of age. Differences in age between groups were not significant except for SLE vs. control where $p=0.02$. All subjects were female.

SLE pathways

We initially examined H3K4me3, H3K27ac and p300 in SLE and healthy controls using DiffBind to define differential peak height between groups. The two histone modifications and p300 binding conceptually reflect active regulatory regions with H3K4me3 being most enriched at active promoters and H3K27ac being enriched at active enhancers. P300 is enzymatically responsible for H3K27 acetylation at enhancers, among other functions [23, 24]. There were minimal genes with conserved differential peak binding across the three cell types and all three chromatin marks. Among the top 20 differentially impacted genes, there was only a single gene with concordant effects (Supplemental Fig. 1). *PRRC2A* had increased peak height across all three antibodies and all three cell types. This gene encodes a protein involved in stress granule formation and is induced by type I interferons or viral infections, consistent with the interferon signature in SLE [7]. *MIR17HG* peak height was decreased across all three antibodies in both T cells and B cells but not in monocytes. This microRNA has been implicated in cell proliferation and apoptosis. In general, there was more sharing across the lymphocytes than either T cells or B cells with monocytes. Overall, the limited concordance across cell types was expected based on recognized biology [25, 26]. We noted more concordance between H3K27ac and p300 peaks compared to other combinations, as expected based on their mechanistic relationship (Supplemental Fig. 2). Thus, concordance was highest across T cells and B cells, and p300 and H3K27ac. These findings reflect known biological attributes where p300 and enhancers are mechanistically linked. The volcano plots clearly demonstrate the magnitude of the differential between decreased peak height and increased peak height with far more decreased peak heights in SLE compared to controls, which was particularly notable in T cells and B cells (Supplemental Fig. 3).

Recognizing that distinct cell types reflect responses to their environment uniquely, we examined whether specific pathways were conserved across cell types. We identified strongly conserved pathways among genes corresponding to increased peak height in SLE using MSigDB v7.5.1 [27, 28]. MSigDB's overlap analysis identifies pathways that overlap with an input list of genes. Overlaps with MSigDB's Hallmark (H) and Canonical Pathway (C2:CP) genesets were determined for the differentially bound peaks with the greatest log2 fold change in SLE compared to control. With the exception of p300 peaks in T cells where only one pathway was enriched (GPCR signaling), all cells and chromatin marks associated with increased peak height in SLE

Table 1 *P* values of enrichment of GSEA pathways of top 500 genes with increased peak height in SLE

	Monocyte H3K4me3	Monocyte H3K27ac	Monocyte p300	B cell H3K4me3	B cell H3K27ac	B cell p300	T cell H3K4me3	T cell H3K27ac	T cell p300
TNFA signaling via NFkB	$2.2X10^{-22}$	$2.5X10^{-14}$	$1.9X10^{-9}$	$2.2X10^{-16}$	$3.1X10^{-13}$	$3.0X10^{-12}$	$2.6X10^{-9}$	$3.1X10^{-13}$	-
Inflammatory response	$2.2X10^{-22}$	$1.2X10^{-18}$	$1.9X10^{-9}$	$1.5X10^{-8}$	$1.1X10^{-7}$	$4.4X10^{-6}$	$2.4X10^{-10}$	$3.1X10^{-13}$	-

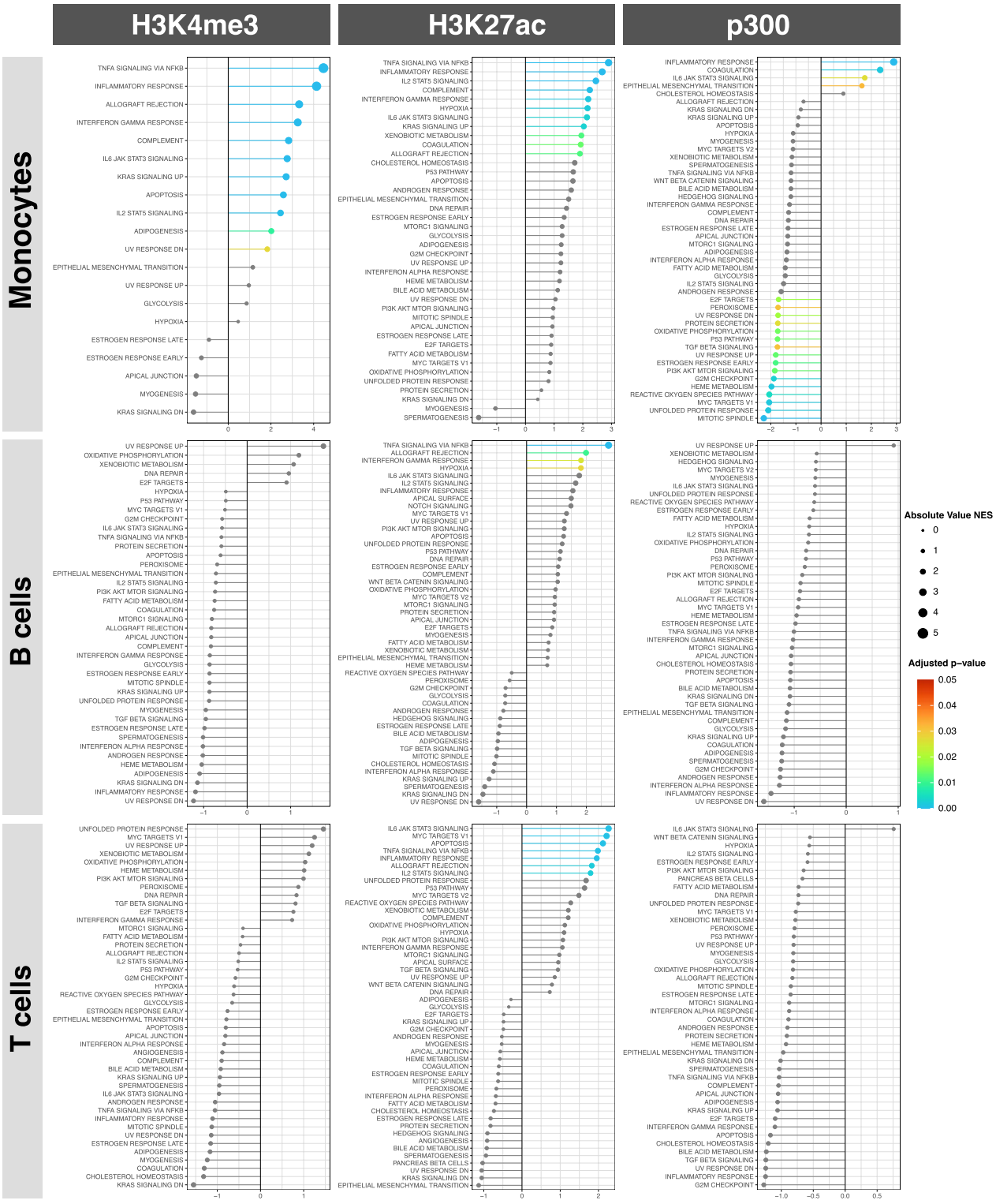

Fig. 1 Pathway analysis of SLE versus control. Differentially bound ChIP-seq peaks up and down in SLE compared to control were used for Fast Gene Enrichment Analysis (*fgsea)* analysis of Hallmark pathways. The cutoff for genes analyzed was FDR < 0.1 and the genes were ranked by log2 fold change in *fgsea*. The analyses are divided by antibody type and cell type. Pathways enriched in SLE have a positive normalized enrichment score, and pathways with decreased enrichment in SLE have a negative normalized enrichment score. The color scale indicates the significance (p-value < 0.05) and the size of the circle reflects the absolute value of the normalized enrichment score. Pathways with an adjusted p-value > 0.05 are colored in gray

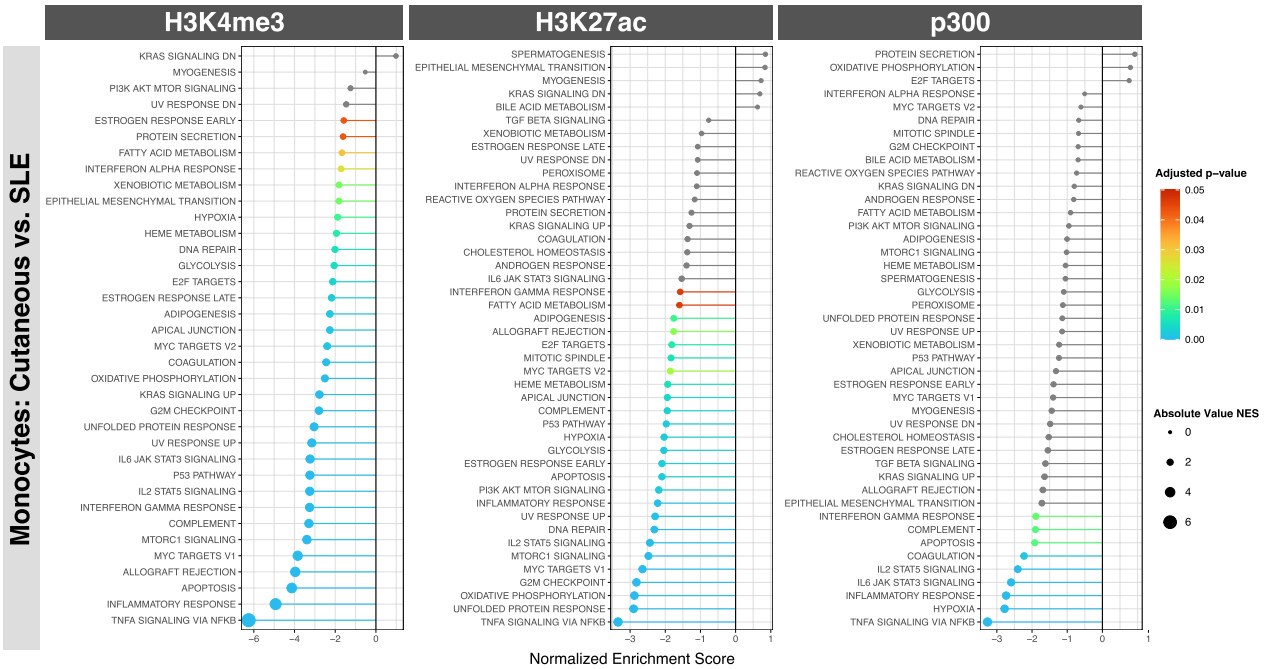

Fig. 2 Pathway analysis of cutaneous versus general SLE. Differentially bound ChIP-seq peaks up and down in cutaneous compared to general SLE were used for Fast Gene Enrichment Analysis (*fgsea)* analysis of Hallmark pathways. The cutoff for genes analyzed was FDR < 0.1 and genes were ranked by log2 fold change in *fgsea*. Only monocyte data is displayed. B cells had no significant differences and T cell had a single pathway altered in SLE (TNF). Pathways enriched in cutaneous lupus have a positive normalized enrichment score, and pathways with decreased enrichment in cutaneous lupus have a negative normalized enrichment score. The color scale indicates the significance (*p*-value < 0.05) and the size of the circle reflects the absolute value of the normalized enrichment score. Pathways with an adjusted *p*-value > 0.05 are colored in gray

shared high enrichment for "TNFA signaling via NFκB" and "Inflammatory Response" (Table 1). This finding is all the more notable given the overall decrease in peak height in SLE (Supplemental Fig. 3). Thus, individual gene signatures are varied between cells and chromatin marks, but there are similar pathways driving the changes across cell types. We then analyzed overall peak height changes (up and down) in SLE for MSigDB Hallmark pathways using *fgsea*, an implementation of GSEA [29]. The *fgsea* enrichment score (ES) represents the degree to which the query genes (here, the differentially bound genes) are present in a pathway geneset. We see a clear recapitulation of the key pathways across cells and chromatin marks (Fig. 1). TNF/NFκB (particularly prominent for H3K27ac marks) and IL-6/Jak/STAT3 (particularly prominent in monocytes) pathways all reached significance across multiple antibodies and cell types in the differential peak height analysis. These data in Table 1 and Fig. 1 suggest both myeloid and lymphoid contributions to the cellular epigenetic responses to disease and highlight the TNF/NFκB and IL-6/Jak/STAT3 pathways as being central to the disease process.

SLE clinical subsets

To understand if distinct pathways were invoked in clinical subsets, we analyzed 5 samples from patients with cutaneous lupus and 8 samples from patients with lupus nephritis. Clinical data are in Supplemental Table 1. Comparing cutaneous SLE to general SLE using *fgsea* to identify enriched pathways, nearly all of the significant pathways had a negative normalized enrichment score, which indicates that these pathways were predominantly associated with decreased peak height in monocytes in cutaneous samples compared to general SLE (Fig. 2). Across all three antibodies, the TNF pathway had the most negative enrichment score in cutaneous lupus compared to general lupus. Multiple other pathways identified as elevated in the general SLE vs. control analysis (Fig. 1) were decreased in cutaneous lupus compared to general lupus for all three antibodies, including complement, IL-2/STAT5, and γ-interferon pathways. The only T cell or B cell difference observed was increased peak height in TNF pathway genes defined by H3K27ac in T cells (not shown). Similarly, comparing nephritis to general SLE, we again found that all the significant pathways had negative enrichment scores, meaning that the genes within these pathways generally had decreased peak

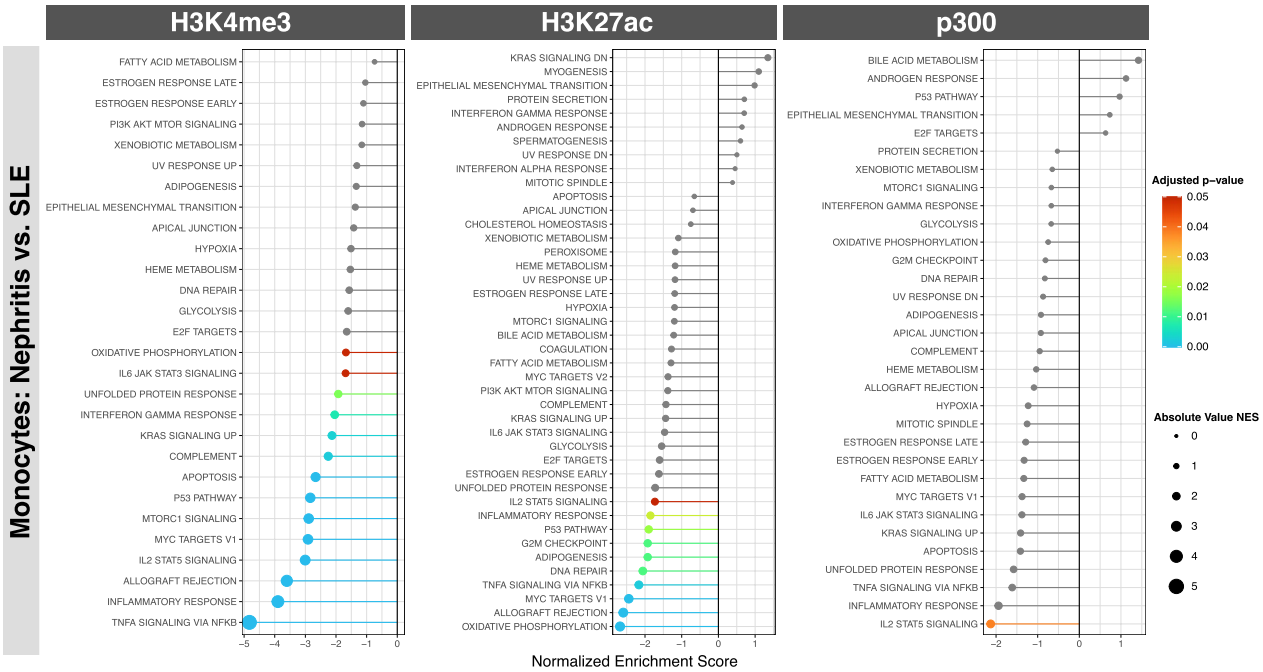

Fig. 3 Pathway analysis of lupus nephritis versus general SLE. Differentially bound ChIP-seq peaks up and down in lupus nephritis compared to general SLE were used for Fast Gene Enrichment Analysis (*fgsea)* analysis of Hallmark pathways. The cutoff for genes analyzed was FDR < 0.1 and genes were ranked by log2 fold change in *fgsea*. Only monocyte data is displayed. No significant differences were observed for T cells and B cells. Pathways enriched in lupus nephritis have a positive normalized enrichment score, and pathways with decreased enrichment in lupus nephritis have a negative normalized enrichment score. The color scale indicates the significance (p-value < 0.05) and the size of the circle reflects the absolute value of the normalized enrichment score. Pathways with an adjusted p-value > 0.05 are colored in gray

height in monocytes (Fig. 3). The IL-2/STAT5 signaling pathway was significantly lower across all three antibodies in monocytes in lupus nephritis vs general SLE. These findings support monocyte responses to the inflammatory milieu of SLE as a differentiator in clinical subsets.

SLE compared *to RA*

We analyzed 4 samples from patients with RA to understand whether distinct pathways were invoked in this clinically distinct autoimmune disease. In this analysis, we found that p300 across all three cell types was the most altered between the two cohorts and the differences were largely decreased peak height in SLE compared to RA (Fig. 4). There were no genes with significantly altered peak height shared across all three cell types.

ChIP-seq cross-reference with genome-wide association study signals

Genome-Wide Association Studies (GWAS) have been instrumental in driving pathogenic mechanism discoveries across disease states [22, 30, 31]. To define relationships between known GWAS signals and differential peak height, we used a 1 kb window for each of the 221 single nucleotide polymorphisms (SNPs) [22]. The Wilcoxon Rank sum test was used (since sequencing data showed a negative binomial distribution which necessitates a non-parametric test) to analyze the difference in mean ChIP-seq coverage between SLE and Control for each 1 kb SNP window (Fig. 5). T cells and B cells had more concordant changes at the SNPs compared to either one with monocytes. Nevertheless, there were comparable numbers of genes with significantly (Wilcoxon $p < 0.05$) altered chromatin/p300 binding across all three cell types. Of note, no gene had concordant changes across all three cell type for all three

(See figure on next page.)

Fig. 4 Differential peak height in SLE versus RA. Overall, peak height is decreased in SLE compared to RA. The 20 genes (or fewer if less than 20) with the greatest significantly diminished peak heights, i.e., negative log2 fold change (log2FC), in SLE compared to RA are shown in **A**. Peak height (annotated by nearest gene) is ranked by the log2 fold change of the decrease and the FDR indicates the significance of the difference between lupus and RA samples for each peak as calculated in DiffBind. Volcano plots of the differentially bound peaks between SLE and RA samples are shown in **B** to optimally display the overall effect on peak height

A

p300: Lupus vs. RA

Monocytes

symbol	log2FC	FDR
KYNU	-5.172	1.00e-02
NMNAT2	-4.492	8.14e-06
SPINK14	-4.391	7.72e-04
HECW2-AS1	-4.349	1.00e-02
RN7SL288P	-4.129	4.00e-03
CYTIP	-4.039	8.00e-03
ABCA13	-3.960	3.00e-03
RNU5A-2P	-3.878	1.60e-02
RNU6-960P	-3.809	1.10e-02
RPL17P5	-3.601	4.70e-02
LARP1BP2	-3.526	4.90e-02
GIRGL	-3.491	1.00e-02
LPP-AS1	-3.207	1.60e-02
FAM133B	-2.708	3.10e-02

B cells

symbol	log2FC	FDR
KANSL1L	-6.578	1.19e-05
ZNG1A	-6.517	2.43e-04
DCP1A	-6.426	1.06e-04
CDC42EP3	-6.139	3.00e-06
FYB1	-6.001	1.57e-07
TXNL1P1	-5.992	3.00e-06
MAP2	-5.982	2.59e-04
TNFAIP6	-5.946	1.23e-04
MIR4263	-5.886	3.00e-06
MIR17HG	-5.881	1.07e-05
GPC5	-5.836	2.91e-04
MTCO2P5	-5.803	3.61e-06
CXCL8	-5.693	4.26e-04
TBC1D4	-5.622	6.91e-04
ACTR6P1	-5.569	8.85e-04
RPL34P6	-5.476	3.61e-06
DGLUCY	-5.465	1.22e-05
METTL21C	-5.458	1.00e-03
MMRN1	-5.382	4.79e-04
SRSF10	-5.368	4.22e-04

T cells

symbol	log2FC	FDR
PDC-AS1	-6.137	0.000543
RNY3P3	-5.845	0.001000
RPS18P8	-5.676	0.003000
C4orf19	-5.659	0.003000
ME1	-5.600	0.003000
CCSER1	-5.587	0.009000
RPL21P75	-5.520	0.003000
NDUFB2-AS1	-5.491	0.003000
RSBN1L	-5.449	0.004000
MED28P5	-5.427	0.005000
PLA2G4A	-5.391	0.004000
BMP5	-5.358	0.007000
DAOA-AS1	-5.295	0.009000
PTMAP5	-5.229	0.000933
RNU6-1001P	-5.206	0.007000
INSC	-5.197	0.009000
RNU6-620P	-5.140	0.011000
RAB32	-5.088	0.008000
SNRPGP18	-5.071	0.010000
ANKRD7	-5.054	0.015000

B

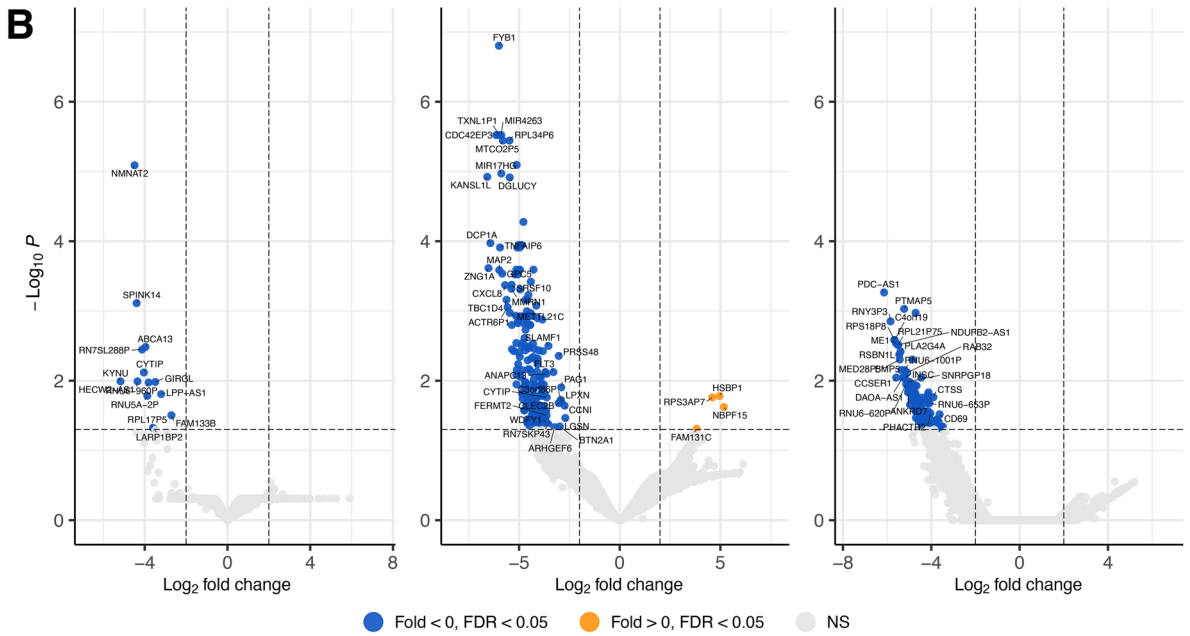

Fig. 4 (See legend on previous page.)

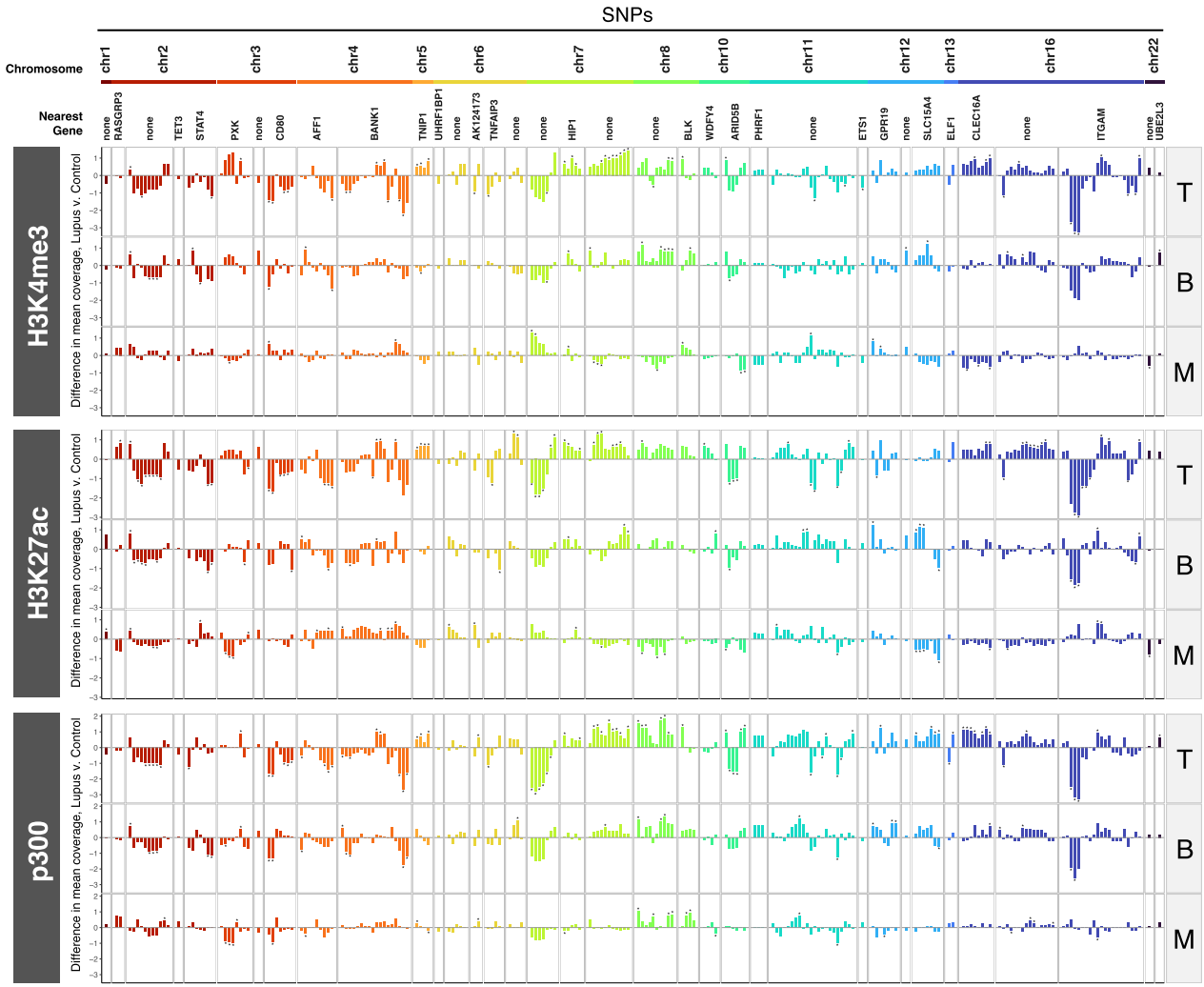

Fig. 5 Cross-referencing of ChIP-seq signals with known single nucleotide polymorphisms (SNP) associated with SLE. Stacked bar plots of the difference in mean coverage for each chromatin mark for each SNP window, shown along the x-axis. For each SNP window, the difference in mean coverage between SLE and Control is shown (y-axis). The three graphs per antibody are T cells on the top ("T"), B cells in the middle ("B"), and monocytes ("M") on the bottom, indicated by corresponding letters on the right of the graphs. Labeling at the top of the stacked bar plot shows the chromosome and gene annotation for the SNPs. Genes are ordered (left to right) by their position in the chromosome. Within each gene, SNPs are sorted by position. An asterisk ("*") above a bar indicates that the difference between SLE and Control coverage for the SNP was significant according to the Wilcoxon Rank Sum Test (p-value < 0.05)

marks. *CD80*, *STAT4*, and *HIP1* had concordant changes across all three marks for both T cells and B cells. *HIP1* and *CD80* changes were significant in both lymphocytes and two of three data sets in monocytes. Several other loci deserve mention. *ITGAM* has been highlighted as functionally significant in the evolution of SLE [32]. While not uniformly altered, differences between SLE and control did reach significance across multiple marks and cell types. Finally, chromosomes 7 and 8 had broad regions that were not clearly associated with a single gene with widely altered chromatin in SLE. These data serve as both validation of the biological consistency across different approaches and also serves as a demonstration of the complexity of this disease.

ChIP-seq cross-reference with RNA expression

Changes in histone modifications represent a control step for transcription and combinations of histone modifications serve to define both the set-point and transcriptional state of a gene. Therefore, it would be anticipated that there would be significant sharing across chromatin and transcriptional signatures. To test this, we compared our ChIP-seq signatures with those from a previously published RNA-seq study of the same three cell types

[17]. Genes were much more likely to have concomitantly increased H3K27ac and mRNA than H3K4me3 or p300 with upregulated RNAs (Fig. 6). All three cell types were very consistent. Among downregulated RNA species in SLE, there was less consistency across the cell types compared to upregulated RNAs, however, among downregulated RNAs, H3K27ac was less likely to be downregulated than either H3K4me3 or p300. Therefore, there is concordance for RNA and histone changes in SLE, but the associations are dependent on cell type and directionality of change.

Conclusions

Efforts to understand the epigenetic alterations in SLE are driven by two goals. Pre-disease changes could provide clues to environmental factors that drive susceptibility. Our study focused on disease-related changes in an effort to enhance our understanding of disease mechanisms. Therapy for SLE has advanced dramatically since the advent of unbiased approaches to understanding disease with emphasis on B cells and the interferon signature as therapeutic targets. Nevertheless, there is an unmet need for new therapeutics in SLE.

Efforts to discern pathways from GWAS, DNA methylation, RNA expression and combinations of these data sets have generally agreed on interferon pathways as being central. A strategy that used cross-referencing GWAS hits with ENCODE-defined cell type-specific H3K4me3 peaks implicated regulatory T cells, B cells, and hematopoietic stem cells as key mediators in SLE [33]. The EBV transcription factor, EBNA, along with NFκB family members comprised the most significant intersection of GWAS signals and a known transcription factor [34]. This study builds on our prior work by validating in broad strokes our previous findings of interferon and inflammatory signals and expands the recognition of the TNF/NFκκB pathway. We were also able to address the organ specificity of the SLE signature by using two disease subsets and a small group of patients with RA. The current study again underscored

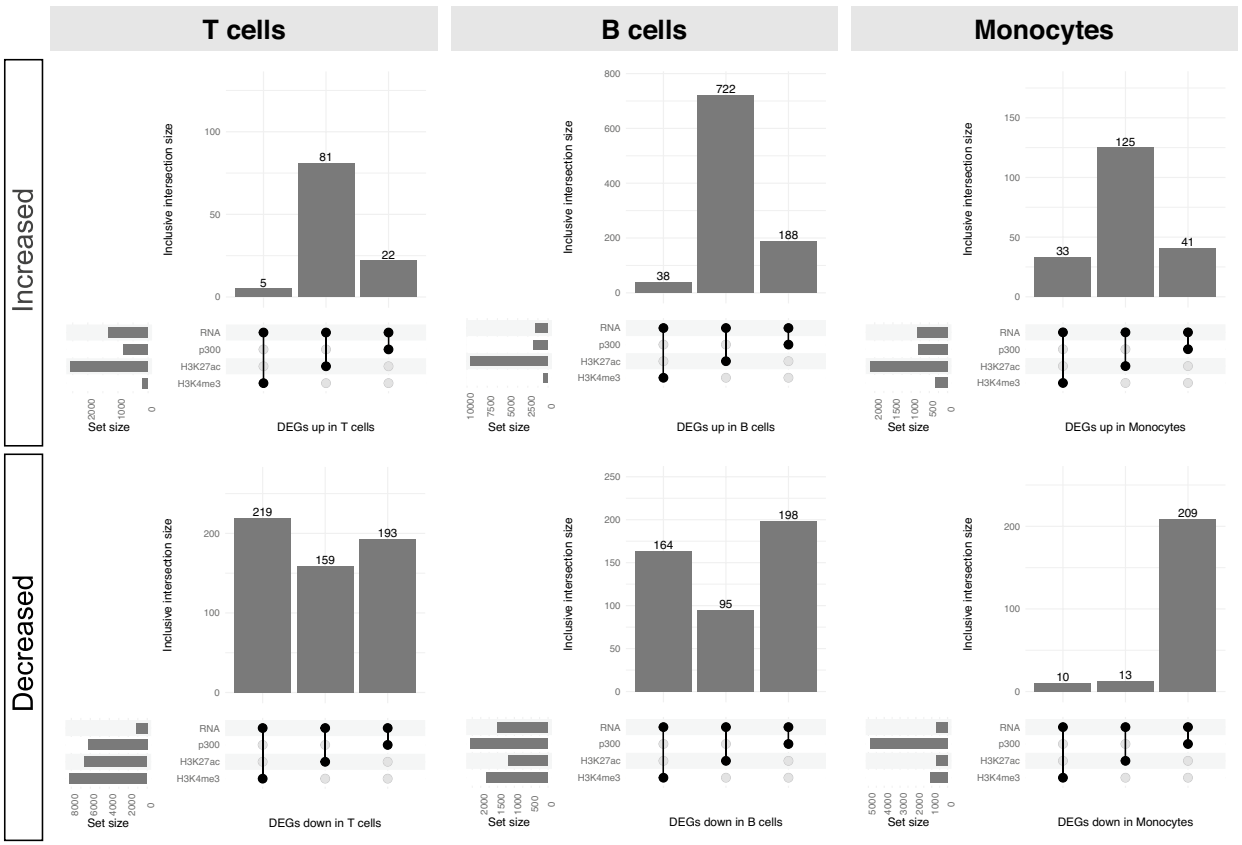

Fig. 6 Cross-referencing of RNA with ChIP-seq signals. Utilizing a previously published data set of RNA-seq changes in SLE in three cells isolated identically to the cell types reported in the current study, the changes in ChIP-seq signals are reported as a number of co-regulated signals in each set of vertical bar graphs. The top row displays increased ChIP-seq signals that overlapped genes with increased RNA-seq signals. The lower row displays decreased ChIP-seq signal overlapping genes with decreased expression. The horizontal bar graph insert displays the gene set size for each cross-referencing

the differences in cell responses to the same disease state and importantly highlighted the monocyte as a central cell type that distinguished lupus nephritis and cutaneous lupus from generalized SLE. The lower signatures seen in cutaneous lupus and lupus nephritis compared to SLE might reflect the overall impact of the disease on peripheral blood cells. This study has several strengths including the incorporation of disease subsets and overall sample size. Nevertheless, there are limitations. We utilized total T cells, total B cells and total monocytes to ensure sufficient signal. Within these populations, altered distributions of subsets could impact our findings. The sample size for RA limited conclusions regarding distinctions between RA and SLE. This study also analyzes largely the reflections of the disease state not primary etiologic events. We specifically recruited participants with relatively low disease activity to try to limit heterogeneity related to activity. Whether the same signatures woudl be amplified during a flare or whether new peaks would appear is not known. Lastly, this study evaluated two histone marks and one histone acetylase, p300. Analysis of transcription factors or overall chromatin structure might lead to different conclusions. Nevertheless, this is the largest ChIP-seq data set reported for SLE to date and significant findings were identified that impact our understanding of the disease.

We cross-referenced our results with two previously published datasets. We examined the association of our altered ChIP-seq signals with disease-associated SNPs and with altered RNA. Most disease-associated SNPs are in regulatory regions [35–37] and therefore can be expected to have an altered state. The themes from this analysis recapitulated other themes in this study, showing more concordance for T cells and B cells than other cell combinations. *STAT4* consistently had significantly altered chromatin which is a locus previously highlighted as driving severity in SLE [38]. Less consistent signals, but still significant in individual cell types, were *ETS1* implicated, in SLE by murine studies and a gene association study [39, 40]. *CLEC16A* appears to have a broad role in autoimmunity [41, 42] and it exhibited significantly altered histone marks in T cells. Cross-referencing with RNA demonstrated strong alignment of increased H3K27ac and RNA. H3K27ac is a mark of an enhancer and increased H3K27ac in enabling for transcription acting as a crucial set-point and with activation of transcription driven often by acute stimuli. H3K4me3 can be increased in promoters with active transcription but if the nucleosomes are remodeled off the promoter altogether, then H3K4me3 can have a decreased signal, making it less likely to find high concordance. P300 exhibited an interesting pattern with limited enrichment among upregulated RNAs but fairly strong association of decreased p300 signal with decreased RNAs. P300 plays a unique role as a histone acetyltransferase, requiring the binding of multiple transcription factors for chromatin association [43]. Thus, loss of a single transcription factor could lead to p300 dissociation but not loss of H3K27ac. The stringent requirement of p300 for multiple transcription factors may underlie its stronger association with decreased RNA. Overall, the results of our cross-referencing studies are consistent with known biological pathways and these data strongly aligned with recognized pathways.

One of the main findings was that among highly increased peak heights, the TNF/NFκB pathway was enriched across all three cell types. This is not wholly surprising since NFκB has been identified as central in many cell models of SLE, some RNA expression analyses, and by pharmaceutical companies as a potential drug target. In a study examining transcription factor binding sites at disease-associated SNPs, the NFκB pathway was enriched across multiple inflammatory diseases, including SLE [44]. Nevertheless, it has been overshadowed by the interferon pathway which has been detected by multiple groups as a strong signature in peripheral blood transcript analyses [5, 7, 45]. Indeed, this critical finding was responsible for the most recent therapeutic advance for SLE, the licensing of anifrolumab. Yet, embedded within those early array studies, there was an inflammatory signature in these mixed cell studies. A recent study that attempted to dissect the individual cell contributions to the inflammatory landscape in SLE used RNA-seq on 27 sorted cell types [46]. This study demonstrated that the interferon signature was particularly strong in neutrophils and monocytes. B cells and CD4 T cells exhibited altered ribosome transcription related to disease activity and altered oxidative phosphorylation that was independent of disease activity. Our previous study of three cells also identified ribosome biogenesis as a strong signal [17]. The study of 27 sorted cell types also found that cytokine and chemokine upregulation was highest in the monocyte lineages and this could be tied to the TNF/NFκB pathway that we also observed as central. A central observation of the RNA-seq study was the signals associated with disease activity and disease subsets were cell type-specific and monocytes had a significant contribution to both disease activity and organ involvement. They were the highest-ranking source of gene expression variance for musculoskeletal disease in SLE. Thus, our data represent an orthogonal approach to understand disease subsets. We found monocytes had the most peak height variance for lupus nephritis and cutaneous lupus compared to generalized SLE, supporting a key role for monocytes in molding disease phenotype. Collectively, these types of studies are important

in driving the precision medicine approaches to the care of patients with SLE by illuminating the drivers of organ effects. The landscape is evolving from the conceptualization of SLE as a T cell-driven disease process with T cell-directed induction of autoantibodies to a far more complex disease, suggesting that narrow targeting of cells or cytokines may have a limited therapeutic effect.

This study made several critical contributions to the understanding of the lupus disease state. Examining three cell types in an unbiased manner led to a fuller picture of the pathways implicated in each cell type. Strong signals for the IL-2 pathway, largely related to heightened T cell responses, and the TNF/NFκB pathway lead to a more nuanced and granular perspective. This study, therefore, identifies pathways with potential for therapeutic targeting in this chronic disease for which new therapeutics are critically necessary. Our study also directly identified chromatin changes at known SNPs, a beginning toward implication of these variants as central to multiple cells' behavior in SLE.

Availability of data and material

No datasets were generated or analyzed during the current study. The sequence data are available on the database of Genotypes and Phenotypes (dbGaP) under accession phs003713.v1.p1.

Abbreviations

ENCODE	Encyclopedia of DNA elements
GWAS	Genome-wide association study
RA	Rheumatoid arthritis
SLE	Systemic lupus erythematosus

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13148-024-01754-3.

Additional file1

Acknowledgements

The authors wish to acknowledge the patients and families and the study staff.

Author contributions

KB performed the bioinformatic analyses with contributions from DT. XW, LS, and CB performed the experiments. DG, MP, and KES performed statistical analyses and oversaw the activities of the work. KES wrote the first draft, and the final draft was approved by all authors.

Funding

This study was funded by the Department of Defense grant W81XWH-19–1-0562 and the Wallace Chair of Pediatrics.

Declarations

Ethical approval and consent to participate

The Johns Hopkins University School of Medicine Institutional Review Board approved this study. IRB00216558. UPENN Institutional Review Board for healthy controls was 705906.

Patient consent

All patients consented for this study.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Biomedical and Health Informatics, Abramson Research Center, Children's Hospital of Philadelphia Research Institute, Philadelphia, PA 19104, USA. [2]Division of Allergy Immunology, Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. [3]Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA. [4]Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA.

Received: 6 May 2024 Accepted: 26 September 2024
Published online: 14 November 2024

References

1. Murimi-Worstell IB, Lin DH, Nab H, Kan HJ, Onasanya O, Tierce JC, et al. Association between organ damage and mortality in systemic lupus erythematosus: a systematic review and meta-analysis. BMJ Open. 2020;10(5):e031850.
2. Choi MY, Chen I, Clarke AE, Fritzler MJ, Buhler KA, Urowitz M, et al. Machine learning identifies clusters of longitudinal autoantibody profiles predictive of systemic lupus erythematosus disease outcomes. Ann Rheum Dis. 2023;82(7):927–36.
3. Choi MY, Clarke AE, Urowitz M, Hanly J, St-Pierre Y, Gordon C, et al. Longitudinal analysis of ANA in the Systemic lupus international collaborating clinics (SLICC) inception cohort. Ann Rheum Dis. 2022;81(8):1143–50.
4. Giannakou I, Chatzidionysiou K, Magder LS, Gyori N, van Vollenhoven R, Petri MA. Predictors of persistent disease activity and long quiescence in systemic lupus erythematosus: results from the Hopkins Lupus Cohort. Lupus Sci Med. 2018;5(1):e000287.
5. 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.
6. Banchereau J, Pascual V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. Immunity. 2006;25(3):383–92.
7. Kirou KA, Lee C, George S, Louca K, Papagiannis IG, Peterson MG, et al. Coordinate overexpression of interferon-alpha-induced genes in systemic lupus erythematosus. Arthritis Rheum. 2004;50(12):3958–67.
8. Hedrich CM, Crispin JC, Tsokos GC. Epigenetic regulation of cytokine expression in systemic lupus erythematosus with special focus on T cells. Autoimmunity. 2014;47(4):234–41.
9. 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.
10. Coit P, Jeffries M, Altorok N, Dozmorov MG, Koelsch KA, Wren JD, et al. Genome-wide DNA methylation study suggests epigenetic accessibility and transcriptional poising of interferon-regulated genes in naive CD4+ T cells from lupus patients. J Autoimmun. 2013;43:78–84.
11. Ulff-Moller CJ, Asmar F, Liu Y, Svendsen AJ, Busato F, Gronbaek K, et al. Twin DNA methylation profiling reveals flare-dependent interferon signature and B cell promoter hypermethylation in systemic lupus erythematosus. Arthritis Rheumatol. 2018;70(6):878–90.
12. Yang Y, Liu K, Liu M, Zhang H, Guo M. EZH2: Its regulation and roles in immune disturbance of SLE. Front Pharmacol. 2022;13:1002741.
13. Coit P, Ortiz-Fernandez L, Lewis EE, McCune WJ, Maksimowicz-McKinnon K, Sawalha AH. A longitudinal and transancestral analysis of DNA methylation patterns and disease activity in lupus patients. JCI Insight. 2020. https://doi.org/10.1172/jci.insight.143654.
14. Zhang Z, Song L, Maurer K, Petri MA, Sullivan KE. Global H4 acetylation analysis by ChIP-chip in systemic lupus erythematosus monocytes. Genes Immun. 2010;11(2):124–33.

15. Zhang Z, Shi L, Dawany N, Kelsen J, Petri MA, Sullivan KE. H3K4 trimethylation breadth at transcription start sites impacts the transcriptome of systemic lupus erythematosus. Clin Epigenetics. 2016;8:14.
16. Shi L, Zhang Z, Song L, Leung YT, Petri MA, Sullivan KE. Monocyte enhancers are highly altered in systemic lupus erythematosus. Epigenomics. 2015;7(6):921–35.
17. Zhang Z, Shi L, Song L, Maurer K, Petri MA, Sullivan KE. Overall downregulation of mRNAs and enrichment of H3K4me3 change near genome-wide association study signals in systemic lupus erythematosus: cell-specific effects. Front Immunol. 2018;9:497.
18. Gautam P, Sharma A, Bhatnagar A. Global histone modification analysis reveals hypoacetylated H3 and H4 histones in B Cells from systemic lupus erythematosus patients. Immunol Lett. 2021;240:41–5.
19. Fangtham M, Petri M. 2013 update: Hopkins lupus cohort. Curr Rheumatol Rep. 2013;15(9):360.
20. Petri M. Hopkins lupus cohort. 1999 update. Rheum Dis Clin North Am. 2000;26:199-213.
21. 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.
22. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015;518(7539):337–43.
23. Ortega E, Rengachari S, Ibrahim Z, Hoghoughi N, Gaucher J, Holehouse AS, et al. Transcription factor dimerization activates the p300 acetyltransferase. Nature. 2018;562(7728):538–44.
24. Narita T, Ito S, Higashijima Y, Chu WK, Neumann K, Walter J, et al. Enhancers are activated by p300/CBP activity-dependent PIC assembly, RNAPII recruitment, and pause release. Mol Cell. 2021;81(10):2166.
25. Dong M, Wang B, Wei J, de Fonseca O, Perry CJ, Frey A, et al. Causal identification of single-cell experimental perturbation effects with CINEMA-OT. Nat Methods. 2023;20(11):1769–79.
26. Cui A, Huang T, Li S, Ma A, Perez JL, Sander C, et al. Dictionary of immune responses to cytokines at single-cell resolution. Nature. 2024;625(7994):377–84.
27. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27(12):1739–40.
28. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417–25.
29. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545–50.
30. Boix CA, James BT, Park YP, Meuleman W, Kellis M. Regulatory genomic circuitry of human disease loci by integrative epigenomics. Nature. 2021;590(7845):300–7.
31. Wu Y, Qi T, Wray NR, Visscher PM, Zeng J, Yang J. Joint analysis of GWAS and multi-omics QTL summary statistics reveals a large fraction of GWAS signals shared with molecular phenotypes. Cell Genom. 2023;3(8):100344.
32. Fagerholm SC, MacPherson M, James MJ, Sevier-Guy C, Lau CS. The CD11b-integrin (ITGAM) and systemic lupus erythematosus. Lupus. 2013;22(7):657–63.
33. Lim J, Kim K. Genetic variants differentially associated with rheumatoid arthritis and systemic lupus erythematosus reveal the disease-specific biology. Sci Rep. 2019;9(1):2739.
34. Harley JB, Chen X, Pujato M, Miller D, Maddox A, Forney C, et al. Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. Nat Genet. 2018;50(5):699–707.
35. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, et al. Population genomics of human gene expression. Nat Genet. 2007;39(10):1217–24.
36. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012;337(6099):1190–5.
37. Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome. Genome Res. 2012;22(9):1748–59.
38. Taylor KE, Remmers EF, Lee AT, Ortmann WA, Plenge RM, Tian C, et al. Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus. PLoS Genet. 2008;4(5):e1000084.
39. Garrett-Sinha LA, Kearly A, Satterthwaite AB. The role of the transcription factor Ets1 in lupus and other autoimmune diseases. Crit Rev Immunol. 2016;36(6):485–510.
40. Sullivan KE, Piliero LM, Dharia T, Goldman D, Petri MA. 3′ polymorphisms of ETS1 are associated with different clinical phenotypes in SLE. Hum Mutat. 2000;16(1):49–53.
41. Pandey R, Bakay M, Hakonarson H. CLEC16A-an emerging master regulator of autoimmunity and neurodegeneration. Int J Mol Sci. 2023;24(9):8224.
42. Tam RC, Lee AL, Yang W, Lau CS, Chan VS. Systemic lupus erythematosus patients exhibit reduced expression of CLEC16A isoforms in peripheral leukocytes. Int J Mol Sci. 2015;16(7):14428–40.
43. Ferrie JJ, Karr JP, Graham TGW, Dailey GM, Zhang G, Tjian R, et al. p300 is an obligate integrator of combinatorial transcription factor inputs. Mol Cell. 2024;84(2):234.
44. Karczewski KJ, Dudley JT, Kukurba KR, Chen R, Butte AJ, Montgomery SB, et al. Systematic functional regulatory assessment of disease-associated variants. Proc Natl Acad Sci USA. 2013;110(23):9607–12.
45. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci USA. 2003;100(5):2610–5.
46. Nakano M, Ota M, Takeshima Y, Iwasaki Y, Hatano H, Nagafuchi Y, et al. Distinct transcriptome architectures underlying lupus establishment and exacerbation. Cell. 2022;185(18):3375.

Publisher's Note