6-16-2022

Hypomethylation of miR-17-92 cluster in lupus T cells and no significant role for genetic factors in the lupus-associated DNA methylation signature

Patrick Coit

Xiavan Roopnarinesingh

Lourdes Ortiz-Fernández

Kathleen Maksimowicz-McKinnon
Henry Ford Health, kmckinn2@hfhs.org

Emily E. Lewis

See next page for additional authors

Recommended Citation

Coit P, Roopnarinesingh X, Ortiz-Fernández L, McKinnon-Maksimowicz K, Lewis EE, Merrill JT, McCune WJ, Wren JD, and Sawalha AH. Hypomethylation of miR-17-92 cluster in lupus T cells and no significant role for genetic factors in the lupus-associated DNA methylation signature. Ann Rheum Dis 2022.

Authors

Patrick Coit, Xiavan Roopnarinesingh, Lourdes Ortiz-Fernández, Kathleen Maksimowicz-McKinnon, Emily E. Lewis, Joan T. Merrill, W. Joseph McCune, Jonathan D. Wren, and Amr H. Sawalha

TRANSLATIONAL SCIENCE

Hypomethylation of miR-17-92 cluster in lupus T cells and no significant role for genetic factors in the lupus-associated DNA methylation signature

Patrick Coit,[1,2] Xiavan Roopnarinesingh,[3,4] Lourdes Ortiz-Fernández,[1] Kathleen McKinnon-Maksimowicz,[5] Emily E Lewis,[6] Joan T Merrill,[4] W Joseph McCune,[6] Jonathan D Wren,[4,7] Amr H Sawalha [1,8,9,10]

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2022-222656).

For numbered affiliations see end of article.

Correspondence to
Dr Amr H Sawalha, Rheumatology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; asawalha@pitt.edu

PC and XR contributed equally.

Received 13 April 2022
Accepted 7 June 2022

To cite: Coit P, Roopnarinesingh X, Ortiz-Fernández L, *et al*. *Ann Rheum Dis* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/annrheumdis-2022-222656

ABSTRACT

Objectives Lupus T cells demonstrate aberrant DNA methylation patterns dominated by hypomethylation of interferon-regulated genes. The objective of this study was to identify additional lupus-associated DNA methylation changes and determine the genetic contribution to epigenetic changes characteristic of lupus.

Methods Genome-wide DNA methylation was assessed in naïve CD4$^+$ T cells from 74 patients with lupus and 74 age-matched, sex-matched and race-matched healthy controls. We applied a trend deviation analysis approach, comparing methylation data in our cohort with over 16 500 samples. Methylation quantitative trait loci (meQTL) analysis was performed by integrating methylation profiles with genome-wide genotyping data.

Results In addition to the previously reported epigenetic signature in interferon-regulated genes, we observed hypomethylation in the promoter region of the miR-17-92 cluster in patients with lupus. Members of this microRNA cluster play an important role in regulating T cell proliferation and differentiation. Expression of two microRNAs in this cluster, miR-19b1 and miR-18a, showed a significant positive correlation with lupus disease activity. Among miR-18a target genes, *TNFAIP3*, which encodes a negative regulator of nuclear factor kappa B, was downregulated in lupus CD4$^+$ T cells. MeQTL identified in lupus patients showed overlap with genetic risk loci for lupus, including *CFB* and *IRF7*. The lupus risk allele in *IRF7* (rs1131665) was associated with significant *IRF7* hypomethylation. However, <1% of differentially methylated CpG sites in patients with lupus were associated with an meQTL, suggesting minimal genetic contribution to lupus-associated epigenotypes.

Conclusion The lupus defining epigenetic signature, characterised by robust hypomethylation of interferon-regulated genes, does not appear to be determined by genetic factors. Hypomethylation of the miR-17-92 cluster that plays an important role in T cell activation is a novel epigenetic locus for lupus.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Lupus is characterised by robust DNA hypomethylation in interferon-regulated genes; however, the genetic contribution to the lupus-associated epigenotype is unknown.

WHAT THIS STUDY ADDS

⇒ Our results suggest that genetic factors do not significantly contribute to the lupus-associated DNA methylation profiles.

⇒ We also report a novel epigenetic locus for lupus in a microRNA cluster involved in T cell function.

⇒ Furthermore, we provide a prototype example showing how a lupus risk genetic variant might mediate functional pathogenic effects through altering DNA methylation levels.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ This study highlights the importance of non-genetic factors in determining epigenetic changes characteristic of lupus.

INTRODUCTION

Systemic lupus erythematosus (lupus or SLE) is a heterogeneous autoimmune disease of incompletely understood aetiology. The disease is characterised by a loss of immunotolerance and the development of autoantibodies against nuclear antigens. Severe manifestations of lupus have significant impact on quality of life and can lead to organ damage and mortality in affected patients, particularly among patients of non-European genetic ancestry.[1 2] Genetic risk contributes to the development of lupus, but the estimated heritability of lupus is ~30%.[3–5] Indeed, monozygotic twin studies in lupus suggest a substantial non-genetic contribution to the aetiology of lupus.[6] Environmental exposures across the lifespan that can directly impact epigenetic regulation and cellular function are suggested to be involved in the pathogenesis of lupus.[7 8]

DNA methylation is an epigenetic mechanism that regulates gene expression through the enzyme-mediated addition of a methyl group to the cytosine bases in the genome. DNA methylation is heritable across cell generations and can promote gene silencing, making it an important component in regulating the plasticity of immune cell identity and function.[9] Early work demonstrated that adoptive transfer of CD4$^+$ T cells treated ex vivo with DNA methyltransferase (DNMT) inhibitors was sufficient to cause lupus-like disease in mice,[10] mimicking the DNA methylation inhibition in patients with drug-induced lupus.[11] Since then, other studies have

observed that CD4+ T cells of patients with lupus show a distinct shift in global DNA methylation compared with healthy individuals, potentially in part due to defective MEK/ERK signalling, suppressing DNA methyltransferase 1 (DNMT1) activity in CD4+ T cells, and leading to hypomethylation and overexpression of costimulatory genes.[12–16]

We have previously observed a robust hypomethylation signature in interferon-regulated genes defining patients with lupus.[17 18] Our initial findings in CD4+ T cells were subsequently confirmed and expanded to other cell types by our group and others.[19–21] In CD4+ T cells, we observed hypomethylation in interferon-regulated genes at the naïve CD4+ T cell stage, preceding transcriptional activity. This epigenetic 'poising' or 'priming' of interferon-regulated genes was independent of disease activity.[18] The genetic contribution to this lupus-associated epigenotype is currently unknown.

Methylation quantitative trait loci (meQTL) are genetic polymorphisms that are associated with the methylation state of CpG sites either through direct nucleotide change within the CpG dinucleotide or intermediary mechanisms. Prior studies of patients with lupus show an enrichment of meQTL associated with type I interferon genes, genetic risk loci and specific clinical manifestations in whole blood and neutrophils.[22–24] Furthermore, our previous work suggests that meQTL might at least in part explain differences in DNA methylation between African-American and European-American patients with lupus.[22]

Herein, we evaluated genome-wide DNA methylation data in naïve CD4+ T cells from a large cohort of patients with lupus compared with matched healthy controls. We integrated DNA methylation and genotyping data to better understand the influence of genetic factors on the DNA methylation changes observed in lupus.

METHODS

Study participants and demographics

Seventy-four female patients with lupus and 74 female healthy age-matched (±5 years) and race-matched controls were recruited as previously described[25 26] (online supplemental table 1). All patients fulfilled the American College of Rheumatology classification criteria for SLE.[27]

Sample collection, DNA isolation and data generation

Genomic DNA samples for this study were collected from naïve CD4+ T cells as previously described.[18] Briefly, magnetic beads and negative selection was used to isolate naïve CD4+ T cells from whole blood samples collected from patients with lupus and controls. Genomic DNA was directly isolated from collected cells and bisulfite converted using the EZ DNA Methylation Kit (Zymo Research, Irvine, California, USA). The Illumina HumanMethylation450 BeadChip (Illumina, San Diego, California, USA) was used to measure DNA methylation levels at over 485 000 methylation sites across the genome.

Epigenome-wide association study

Epigenome-wide association study (EWAS) for identifying associations between specific CpG sites and disease status was performed using GLINT.[28 29] Covariates for age, race and technical batch were included for the analysis prior to other preprocessing. No outliers beyond 4 SD were detected in the first two components of the principal component analysis (PCA) space, all 148 samples were included in the analysis. Reference-less cell type composition correction was performed using *ReFACTor*, with six components used in the downstream analysis to account for any cell-type heterogeneity in the samples. An additional covariate was included to account for effects of genetic admixture using the EPISTRUCTURE algorithm included in GLINT. Cell-type composition covariate components generated by *ReFACTor* were included at this step to reduce bias from potential cell-type heterogeneity, and polymorphic CpG sites were excluded from this step and the EWAS. Using the age, race and technical batch covariates, along with six *ReFACTor* components and one *EPISTRUCTURE* component, logistic regression for disease status was performed across all CpG sites, excluding the polymorphic and unreliable cross-reactive probes previously described in the literature, as well as CpG sites with low variance (SD <0.01).[30 31]

Differential DNA methylation analysis of gene promoters

Raw .idat files were used to generate methylation beta value profiles across all samples using GenomeStudio (Illumina) after background subtraction and normalising to internal control probes. Missing probe values were imputed using *sklearn.impute.KNNImputer*, and *ComBat* was used to correct for batch effects associated with sample preparation.[32–34] Ensembl gene loci for hg19 were downloaded using *PyEnsembl*.[35] For each gene, loci for 1500 base pairs upstream of the transcription start site[36] to the transcription start site (TSS) were mapped to the overlapping CpG probes using *PyBedtools*, and the mean of the associated probes for each gene was used as the representative methylation value for the resulting 20 437 mapped genes.[37] Differential methylation analysis comparing patients and controls was performed on the mean TSS1500 methylation using *limma*, and false discovery rate adjustment using the Benjamini-Hochberg method was used to correct p values for multiple testing. Gene Ontology Enrichment for Biological Process terms was performed on the differentially methylated gene list using *Enrichr* with the mapped promoter gene list used as the background.[38 39]

Trend deviation analysis

DNA methylation data derived using the Illumina 450k methylation array from 23 415 samples were downloaded from Gene Expression Omnibus (GEO).[40] To reduce batch effects, samples from experiments with fewer than 50 samples were omitted, and the remaining samples were quantile normalised.[41] A matrix of pairwise Pearson's correlation values for DNA methylation levels was computed across TSS1500 gene promoters in 16 541 samples across 37 tissues to create a multitissue correlation network (online supplemental figure 1). The differentially methylated genes in lupus-naïve CD4+ T cells were clustered by their correlation in global signature created from the GEO data. Hierarchical clustering was performed using *Scipy*'s hierarchical linkage. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed using *Enrichr*,[42] and p values were reported after false discovery rate (FDR) adjustment.

The goal of a trend deviation analysis is to detect correlation patterns among differentially methylated genes in large DNA methylation datasets. A correlation in methylation among a set of differentially methylated genes between patients and controls suggests a trend is being observed, reinforcing the significance and robustness of the differential DNA methylation detected between patients and controls.

Genotyping

Genomic DNA isolated from naïve CD4+ T cells was used as input for the Infinium Global Screening Array-24 V.2.0 (Illumina). Single nucleotide polymorphisms (SNPs) with a genotyping call

rate <98%, minor allele frequencies (MAF) <5% and deviating from Hardy-Weinberg equilibrium (p<1E-3) were filtered out. Samples were removed if they had a genotyping call rate <95%. Sex chromosomes were not analysed. About 100 000 independent SNPs were pruned and used to perform PCA with *Eigensoft* (V.6.1.4) software.[43] Genotyping data were analysed using *PLINK* (V.1.9).[44] Genotype profiles were generated for n=63 patients and n=68 controls.

Methylation quantitative trait loci analysis

Raw .idat files were used to generate methylation profiles using *minfi* (V.1.32.0)[45 46] and to check median intensity values and reported sex in the R statistical computing environment (V.3.6.3).[47] Probes with less than three beads and zero intensity values across all samples were removed using the *DNAmArray* package (V.0.1.1).[48] Background signal and dye bias were corrected, followed by normalisation of signal intensities using functional normalisation in the *preprocessFunnorm.DNAmArray* function[48 49] using the first three principal component values calculated from signal intensities of control probes present on all array spots to correct for technical variation. Any probe with a detection p<0.01 or returned signal intensities in fewer than 98% of samples was removed. This resulted in a total of 476 767 probes used for further analysis. Signal intensities were then converted to M values with a maximum bound of ±16. M values were used for meQTL analysis and converted to beta values (0%–100% methylation scale) using *minfi* for reporting.

We removed any probe for meeting any of the following technical criteria: a unique probe sequence of <30 bp, mapping to multiple sites in the genome, polymorphisms that cause a colour channel switching in type I probes, inconsistencies in specified reporter colour channel and extension base, mapping to the Y chromosome and/or having a polymorphism within 5 bp of the 3' end of the probe with a MAF >1% with the exception of CpG-SNPs with C>T polymorphisms which were retained.[50] Batch correction for chip ID was performed using the *ComBat* function in the *sva* (V.3.34.0) package.[51] After technical filtering, there were a total of 421 214 probes used for meQTL analysis.

We implemented a mixed correspondence analysis with the *PCAmixdata* package (V.3.1)[52] to calculate eigenvalues using patient medication data for prednisone, hydroxychloroquine, azathioprine, mycophenolate mofetil and cyclophosphamide. The top four components accounted for a cumulative 89.3% of variability in the medication data. Each component value was used as an independent variable in regression analysis to adjust for medication usage across individuals. MeQTL association analysis was performed in patients and controls separately using methylation M value profiles and corresponding sample genotypes. Age, the top 4 medication components and top 10 genotype principal components were included as covariates to build a linear model for detecting meQTL using *MatrixEQTL* (V.2.3).[53] *Cis*-meQTL were defined as CpG sites with methylation values associated with an SNP within a conservative 1000 bp of the CpG dinucleotide. We used a Benjamini-Hochberg FDR-adjusted p value cut-off of <0.05 as a threshold for significant associations. The above EWAS results were compared with the meQTL results to determine overlap of lupus-associated differentially methylated CpG sites and those CpG sites in an meQTL.

Functional enrichment analysis

ToppGene Suite was used for functional enrichment analysis[54] of Molecular Function and Biological Process Gene Ontologies and KEGG pathways in meQTL. P values were derived using a hypergeometric probability mass function, and a Benjamini-Hochberg FDR-adjusted p value cut-off of <0.05 was used as a threshold of significance. A minimum membership of 3 genes and maximum of 2000 genes in each term was used as a threshold for inclusion. Interferon-regulated genes were identified using the set of genes associated with the CpG site in each meQTL as input using Interferome (V.2.01).[55] The type I interferon response genes were defined as genes with an expression fold change of 1.5 or greater between type I interferon-treated and untreated samples using gene expression datasets from all available $CD4^+$ T cell experiments in the Interferome database.

For the analysis of miR-18a-regulated genes, literature-based network association analysis was performed using IRIDESCENT to create a weighted network of published relationships as previously described.[56]

MicroRNA expression microarray

MicroRNA (miRNA) expression was measured in naïve $CD4^+$ T cells from a subset of patients with lupus and healthy matched controls (n=16). Cells were immediately lysed with TRIzol Reagent (ThermoFisher Scientific, New York, USA) followed by storage at −80°C. Total RNA was isolated using the Direct-zol RNA MiniPrep Kit (Zymo Research, California, USA) following the manufacturer's directions. The Affymetrix miRNA V.4.1 Array Strip (Affymetrix, California, USA) was used to measure expression of over 2000 premature and 2500 mature human miRNA sequences. RNA sequences were polyadenylated and ligated to a biotin-labelled oligomer using the FlashTag Biotin HSR RNA Labeling Kit (Affymetrix). Biotin-labelled sequences were hybridised to array probes and washed then stained with streptavidin-phycoerythrin (PE). The Affymetrix Expression Console & Transcriptome Analysis Console V.2.0 software (Affymetrix) was used to analyse biotin/streptavidin-PE fluorescence measurements. All samples passed signal intensity, polyadenylation and ligation quality controls. Signal intensities were background adjusted and normalised. Log2-transformed expression values for each probeset was calculated using a robust multi-array average model.[23] The Pearson's r correlation coefficient for median expression values of probes for miR-17, miR-18a, miR-19a, miR-19b1 and miR-20a and Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) scores were calculated using GraphPad Prism (V.9.3.0) (GraphPad Software, California, USA).

RESULTS

Differential methylation of gene promoters in naïve $CD4^+$ T cells isolated from patients with lupus

A comparison of DNA methylation profiles from circulating naïve $CD4^+$ T cells isolated from 74 patients with lupus and 74 age-matched, sex-matched and race-matched healthy controls revealed a total of 2627 CpGs, out of 334 337 total CpG sites included in the EWAS, with a significant difference in average methylation. Significant hypomethylation in interferon-regulated genes was observed, consistent with previous reports (online supplemental table 2). Average promoter methylation for each gene was calculated by including all CpG sites on the array within 1500 bp of the associated gene's TSS. A total of 51 genes showed a significant difference in average promoter methylation between patients with lupus and controls (17 hypomethylated and 34 hypermethylated in patients compared with controls) (table 1) (figure 1). Biological Process Gene Ontology enrichment analysis of differentially methylated promoter regions did not show significant enrichment compared with the background

Table 1 Genes with differentially methylated promoter regions in naive CD4$^+$ T cells of patients with lupus compared with healthy controls

Gene	Δβ	$-\log_{10}$ (FDR-adjusted p value)	t-statistic
IFI44L	−0.177	Infinity	−10.757
DTX3L	−0.130	Infinity	−11.566
BST2	−0.089	11.323	−9.285
RABGAP1L	−0.088	9.165	−8.421
BCL2L14	−0.086	5.520	−6.908
MIR19B1	−0.059	3.169	−5.846
IFI44	−0.059	2.057	−5.304
MIR20A	−0.055	3.088	−5.807
MIR17	−0.054	6.882	−7.487
MIR18A	−0.051	6.537	−7.342
MIR19A	−0.049	4.771	−6.579
IKZF4	−0.048	3.289	−5.902
MX1	−0.046	10.624	−9.004
TRIM34	−0.045	2.184	−5.367
ODF3B	−0.034	1.712	−5.128
GNG2	−0.033	2.138	−5.344
FAM177B	−0.025	1.897	−5.223
MZF1	0.008	1.493	5.014
SSBP4	0.015	1.344	4.934
ATP6V0D1	0.018	2.594	5.569
DCUN1D1	0.025	2.068	5.309
C14orf93	0.025	1.922	5.236
TIPARP	0.026	2.069	5.310
LMBRD1	0.027	2.211	5.381
HAVCR2	0.027	2.574	5.560
KIAA1949	0.030	3.158	5.841
GPD2	0.032	1.953	5.251
CNTF	0.033	1.705	5.124
CD47	0.034	4.259	6.350
ARHGAP9	0.036	3.339	5.926
IL27RA	0.036	1.367	4.946
RAP1A	0.036	2.573	5.559
LAMA3	0.037	1.445	4.988
ABI3	0.037	1.436	4.983
FAM102A	0.038	3.161	5.842
CXCR5	0.039	1.439	4.985
DPEP2	0.040	1.889	5.219
DYRK2	0.041	3.924	6.197
TMEM71	0.044	2.757	5.649
ADORA2A	0.046	2.234	5.392
SEPT9	0.047	2.036	5.293
PSMB4	0.052	2.935	5.734
TOM1	0.055	5.415	6.862
PRIC285	0.057	9.934	8.729
LTB	0.062	2.036	5.293
MIR1205	0.067	1.698	5.121
ACER3	0.073	2.612	5.578
BCL9L	0.079	4.034	6.248
MDS2	0.080	3.149	5.836
SNORA5B	0.083	1.712	5.128
PTPRCAP	0.091	3.620	6.057

FDR correction was performed using the Benjamini-Hochberg method with an FDR-adjusted p value threshold of <0.05. Δβ: methylation difference in median methylation value of CpG sites within 1500 bp upstream of the associated gene's transcription start site (TSS1500) between patients with lupus and healthy controls. FDR, false discovery rate; TSS, transcription start site.

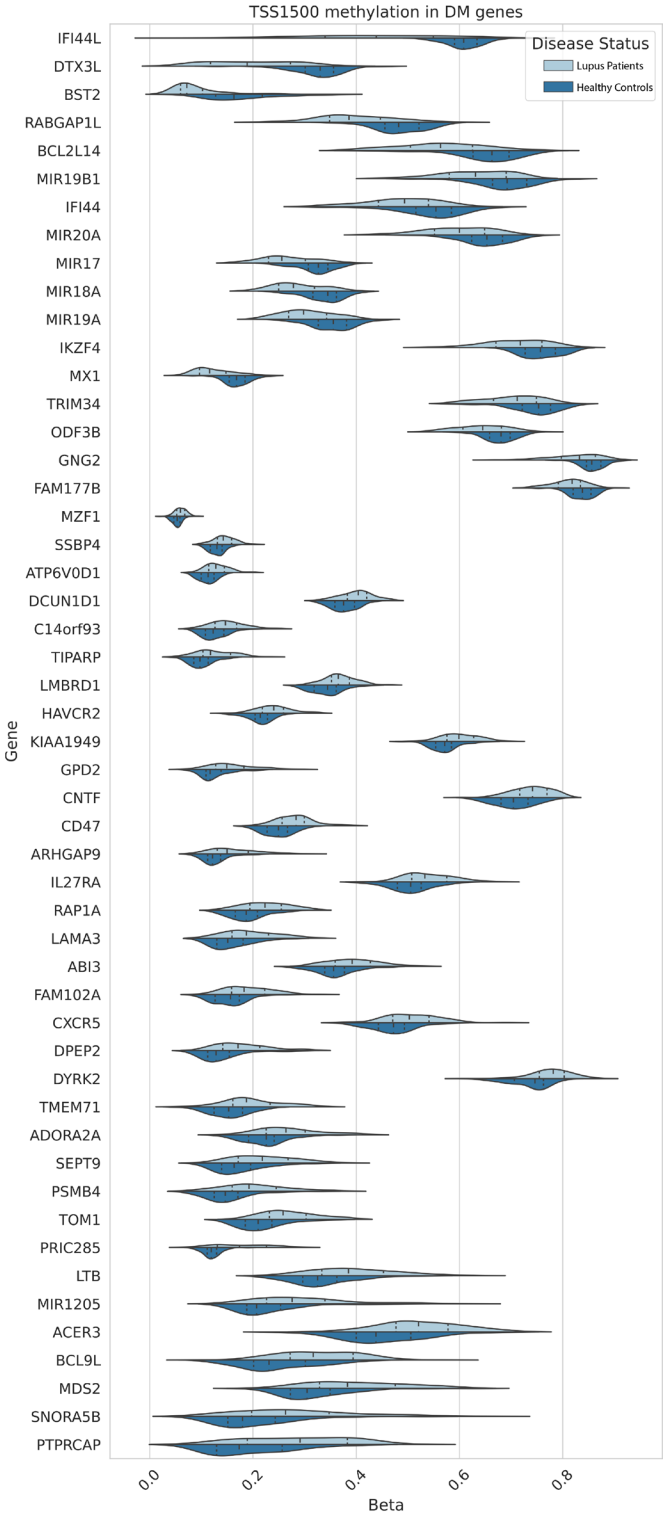

Figure 1 Distribution of average CpG methylation levels within 1500 bp of the transcription start site (TSS1500) for the respective genes differentially methylated (DM) in naïve CD4$^+$ T cells of patients with lupus compared with healthy controls.

of all gene promoters after adjusting for multiple testing (online supplemental table 3).

The pairwise correlation of the 51 gene promoters identified above was calculated across a collection of 16 541 samples from 37 tissues available in GEO. Hierarchical clustering of correlations showed that 21 out of the 51 gene promoters were highly

Table 2 KEGG pathway gene enrichment of 21 gene promoters highly correlated with each other in multitissue DNA methylation data constructed from 16 541 samples available through Gene Expression Omnibus

Pathway (KEGG_2019_Human)	P value	FDR-adjusted, p value	OR	Genes
MicroRNAs in cancer	1.21E-05	0.00039	20.92	*MIR19B1;MIR20A;MIR17;MIR18A;MIR19A*
Cytokine-cytokine receptor interaction	0.0034	0.043	11.28	*CNTF;CXCR5;LTB*
Rheumatoid arthritis	0.0041	0.043	23.52	*LTB;ATP6V0D1*

FDR, false discovery rate; KEGG, Kyoto Encyclopedia of Genes and Genomes; OR, odds ratio.

correlated. KEGG pathway enrichment analysis showed a significant enrichment for three pathways among the 21 correlated gene promoters: 'microRNAs in cancer' (p=3.86E-04), 'cytokine-cytokine receptor interaction' (p=4.34E-02) and 'rheumatoid arthritis' (p=4.34E-02) (table 2) (figure 2). The 'microRNAs in cancer' pathway included genes encoding miR-17, miR-18a, miR-19a, miR-19b1 and miR-20a. Four of seven CpG sites used to calculate the average promoter methylation (TSS1500) in this locus showed a significant reduction in median methylation in patients with lupus compared with healthy controls (figure 3A). These sites: cg17799287 (Δβ=−5.5%; p=2.05E-03), cg07641807 (Δβ=−4.4%; p=1.71E-02), cg23665802 (Δβ=−5.8%; p=1.19E-02) and cg02297838 (Δβ=−4.9%; p=3.48E-02) were all hypomethylated in patients with lupus compared with healthy controls, and overlapped with enhancers and regions flanking TSS in peripheral naïve CD4$^+$ T cells using data collected from the Epigenome Roadmap[57] and visualised using the WashU Epigenome Browser.[58] We examined expression levels of the miRNAs included in the 'microRNAs in cancer' pathway (miR-17, miR-18a, miR-19a, miR-19b1 and miR-20a) in naïve CD4$^+$ T cells of a subset of our patients with lupus (n=16) and healthy matched controls (n=16). We did not observe a difference in expression between patients and controls. However, two miRNAs, miR-18a-5p and miR-19b1-5p, showed a significant positive correlation (hsa-miR-18a-5a p=0.038 and hsa-miR-19b1-5p p=0.042) between median expression levels and SLEDAI scores in patients with lupus (figure 3B) (online supplemental table 4).

Examining publicly available miRNA expression data from total CD4$^+$ T cells revealed overexpression of miR-18a in patients with lupus compared with healthy control individuals.[59]

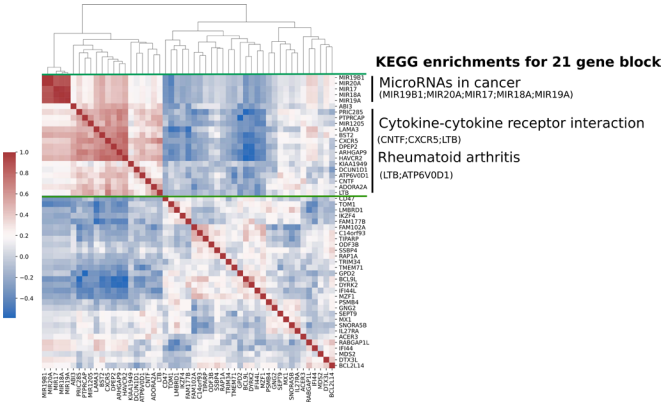

Figure 2 Heatmap of hierarchical clustering of pairwise Pearson's correlation coefficient values of 51 differentially methylated gene promoters (transcription start site (TSS)1500) in global tissue signature derived from 16 541 samples. Range from +1 (red) to −1 (blue), represent a greater to lower correlation in global tissue, respectively. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways are significantly enriched (false discovery rate-adjusted p<0.05) in a block of 21 genes (green bars).

In these same samples, a total of 74 miR-18a-target genes were downregulated in patients with lupus compared with controls. Using a literature-based network association analysis, we identified 15 of these 74 genes with relatedness to lupus (online supplemental figure 2). *TNFAIP3*, which encodes a negative regulator of nuclear factor kappa B (NF-κB) targeted by miR-18a, was downregulated in lupus CD4$^+$ T cells compared with controls.

We examined the expression of *MIR17HG*, which is the host gene that encodes the miR-17-92 cluster, in single cell RNA-sequencing data from lupus nephritis tissue samples generated by the Accelerating Medicines Partnership (AMP) project.[60] We show evidence for MIR17HG mRNA expression in multiple immune cells infiltrating the kidneys of patients with lupus nephritis, including multiple T cell subsets, although in a small percentage of kidney infiltrating cells. While over 8% of tissue-resident macrophages in lupus nephritis tissues express

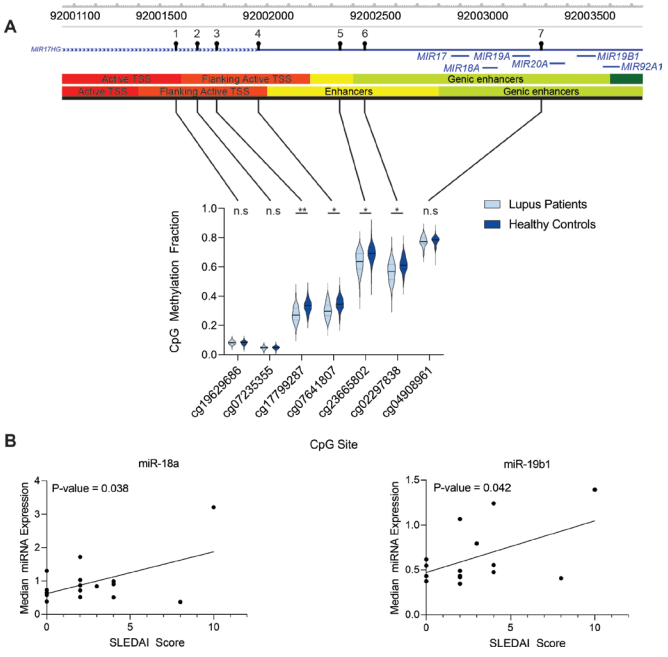

Figure 3 (A) Violin plots of the seven CpG sites in patients with lupus and healthy controls used to calculate the average promoter methylation (transcription start site (TSS)1500) for the miR-17-92 cluster. The solid black bar represents the median value and the dashed lines the first and third quartiles. Genomic visualisation and annotation are from WashU Epigenome Browser using AuxillaryHMM tracks from peripheral naïve CD4$^+$ T cells (E038 and E039, top and bottom tracks, respectively). n.s., not significant. *P<0.05, **p<0.01. (B) Correlation of median microRNA (miRNA) expression in naïve CD4$^+$ T cells of a subset (n=16) of patients with lupus with Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) score. Hsa-miR-18a-5p and hsa-miR-19b1-5p had a Pearson's correlation (r) of 0.52 (p=0.038) and 0.51 (p=0.042), respectively.

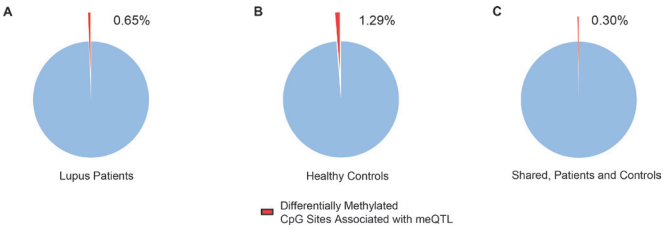

Figure 4 Proportion of differentially methylated CpG sites in naïve CD4$^+$ T cells of patients with lupus compared with healthy controls associated with a methylation quantitative trait loci (meQTL) in (A) patients with lupus, (B) healthy controls and (C) the subset of meQTL shared between patients with lupus and healthy controls.

MIR17HG mRNA, the highest levels of expression were observed in T cell subsets (online supplemental figure 3).

Naïve CD4$^+$ T cell methylation quantitative trait loci in patients with lupus

Global genotype profiles were generated in a subset of patients and controls and compared with global DNA methylation profiles to identify CpG sites with allele-specific methylation associations. There was no significant difference in the average age (years) between the patient (n=63) and control (n=68) subsets (patient average age=41.6; patient age SD=12.8; control average age=40.8; control age SD=12.5; t-test statistic=0.381; two-tailed p=0.704). Allele-specific DNA methylation associations were measured as meQTL, where the CpG site was within 1000 bp of the measured SNP separately in patients and controls. After adjusting for age, genetic background and medication use in patients, we identified 5785 meQTL in naïve CD4$^+$ T cells of patients with lupus with an FDR-adjusted p<0.05 (online supplemental table 5). These meQTL include 4649 unique CpG sites and 4120 unique polymorphisms.

A linear model adjusting for age and genetic background was fit to healthy controls separately. We identified a total of 7331 meQTL with an FDR-adjusted p<0.05 in controls (online supplemental table 6). These meQTL include 5885 unique CpG sites and 5138 unique polymorphisms.

Of 2627 CpG sites differentially methylated between patients and controls, we identified 17 (0.65%) and 34 (1.29%) CpG sites that overlapped with CpG sites included in meQTL in patients and controls, respectively (figure 4A,B). We examined the overlap of meQTL in patients with lupus and healthy controls and identified a total of 3957 meQTL (68.4% of lupus patient meQTL and 54.0% of healthy control meQTL) shared between both patients and controls (online supplemental table 7). This shared set of meQTL contained 8 (0.3%) CpG sites that we identified as differentially methylated between patients with lupus and controls (figure 4C).

Functional enrichment analysis was performed using genes associated with CpG sites in our meQTL shared between patients and controls. Functional enrichment analysis revealed multiple ontologies and pathways for cell adhesion ('cell-cell adhesion'; p=1.04E-12, 'biological adhesion'; p=6.80E-12, 'cell adhesion'; p=8.25E-12, 'cell adhesion molecules'; p=2.25E-06), transporter associated with antigen processing (TAP) proteins and antigen presentation ('TAP binding'; p=1.59E-7, 'peptide antigen binding'; p=4.40E-5) and immune disorder pathways ('type I diabetes mellitus'; p=1.92E-8, 'graft-versus-host disease (GVHD)'; p=4.38E-7) (online supplemental table 8).

There were 1828 meQTL detected only in patients with lupus but not in controls. These were enriched in gene ontologies and pathways related to tissue growth and development ('animal organ morphogenesis'; p=8.44E-10, 'urogenital system development'; p=1.05E-07) and gene silencing ('negative regulation of gene silencing by miRNA'; p=2.54E-6, 'negative regulation of post-transcriptional gene silencing'; p=5.41E-6) (online supplemental table 9).

We compared our list of meQTL in patients with lupus with previously identified lupus susceptibility loci from genome-wide association studies.[4 61–64] We found 41 meQTL with CpG site-associated genes that overlapped with 20 unique lupus risk loci genes (online supplemental table 10). This included interferon regulatory factor genes *IRF5* and *IRF7*. We found three meQTL in naïve CD4$^+$ T cells that included, or were in high linkage disequilibrium (LD) ($r^2 \geq 0.80$) with, a known lupus genetic risk variant (table 3).[65] We also performed a similar analysis using data previously collected from granulocytes of patients with lupus to determine if these effects were present across tissues.[22] We found meQTL associated with lupus risk variants in *CFB* (rs170942) and *IRF7* (rs1131665) in both naïve CD4$^+$ T cells and granulocytes isolated from patients with lupus. In addition, an meQTL associated with the *TMEM86B-PTPRH* locus was observed in naïve CD4$^+$ T cells. When we compared the lupus risk alleles with DNA methylation levels, we found that the presence of the risk allele at rs1270942 (*CFB*) is associated with increased DNA methylation of cg16505946. The presence of the risk allele at rs1131665 (*IRF7*) (figure 5) and rs56154925 (*TMEM86B-PTPRH*) was associated with decreased DNA methylation of cg16486109 and cg01414877, respectively. The direction of the risk allele-DNA methylation association in the *CFB* and *IRF7* meQTL was the same in both naïve CD4$^+$ T cells and granulocytes.

We examined the overlap between genes associated with CpG sites in meQTLin lupuspatients and genes that respond to type I interferon treatment in CD4$^+$ T cells, to better understand the association between genetics and type I interferon-response gene methylation differences in lupus. A total of 101 unique type I interferon-response genes were identified as meQTL in our data (online supplemental table 11).

Because IRF7 is a master regulator of type I interferon response,[66] and the lupus-associated epigenotype is dominated by hypomethylation in interferon-regulated genes, we examined if rs1131665 (*IRF7*) had an effect on the methylation levels of the 2627 CpGs differentially methylated in naïve CD4$^+$ T cells between patients with lupus and healthy controls. This *trans*-meQTL analysis revealed no significant difference in methylation levels across these CpG sites based on rs1131665 genotypes, among patients with lupus (analysis of variance, data not shown).

DISCUSSION

We generated genome-wide DNA methylation data in naïve CD4$^+$ T cells from a large cohort of patients with lupus and matched healthy controls. Implementing an innovative trend deviation analysis, we identified a cluster of miRNAs (miR-17, miR-18a, miR-19a, miR-19b1, miR-20a) among differentially methylated loci in patients with lupus. Promoter methylation analysis revealed significant hypomethylation in this miRNA cluster in patients with lupus compared with controls. Trend deviation analysis suggested a coordinated, disease-associated change in promoter methylation for these miRNAs. Indeed, the expression of miR-18a and miR-19b1 included within this cluster positively correlated with disease activity, as measured using SLEDAI score, in our patients with lupus.

Table 3 MeQTL in naive CD4+ T cells and granulocytes of patients with lupus that include a known lupus risk variant

Lupus-naïve CD4+ T cell meQTL					
CpG site	**meQTL SNP**	**Lupus risk SNP***	**Risk SNP-associated gene**	**Lupus risk allele**	**Direction of CpG methylation associated with risk allele**
cg16505946	rs558702	rs1270942	*CFB*	C	↑
cg16486109	rs1131665	rs1131665	*IRF7*	A	↓
cg01414877	rs56154925	rs56154925	*TMEM86B-PTPRH*	C	↓
Lupus granulocyte meQTL					
CpG site	**meQTL SNP**	**Lupus risk SNP***	**Risk SNP-associated gene**	**Lupus risk allele**	**Direction of CpG methylation associated with risk allele**
cg16505946	rs558702	rs1270942	*CFB*	C	↑
cg16486109	rs1131665	rs1131665	*IRF7*	A	↓

*rs558702 and rs1270942 have an LD r^2 ≥0.80.
meQTL, methylation quantitative trait loci; SNP, single nucleotide polymorphism.

MiRNAs play an important role in post-transcriptional gene regulation by targeting specific complementary gene transcripts for degradation.[67] Peripheral blood cells in patients with lupus show altered expression of miRNAs.[68] Some dysregulated miRNAs in lupus target DNMT1, and as a result, contribute to altered DNA methylation patterns in lupus CD4+ T cells.[69–71] miR-17, miR-18a and miR-20a form the 'miR-17 family' while miR-19a and miR-19b1 form the 'miR-19 family'. These miRNAs are grouped by sequence homology and encoded in a single polycistronic miRNA gene as the 'miR-17-92 cluster'. This cluster has been well-studied as an oncogene and an immune regulator.[72] Average promoter methylation of miR-17, miR-18a, miR-19a, miR-19b1 and miR-20a was reduced by ~5% in patients with lupus compared with controls, which has not been previously described in immune cells of patients with lupus. Enterovirus 71 infection has been observed to suppress miR-17-92 cluster expression by increasing DNMT-mediated promoter methylation,[73] and chemical inhibition of DNMT1 activity in bleomycin-induced lung fibrosis mouse model increases miR-17-92 cluster expression in lung fibroblasts.[74] This suggests that miR-17-92 cluster promoter methylation plays an important role in regulating the expression of its members.

MiR-17-92 cluster genes play a vital role in regulating T cell activities including proliferation and differentiation. Overexpression of miR-17-92 cluster genes promotes lymphoproliferative disease and autoimmunity in mice by targeting critical immunotolerance regulators Bim and PTEN.[75] Conditional knockout of miR-17-92 cluster in a murine model of chronic GVHD (cGVHD) reduced disease-associated T cell infiltration and IgG deposition in the skin.[76] In cGVHD mice, miR-17-92 cluster expression in CD4+ T cells supports T helper (Th)1, Th17 and T follicular helper (Tfh) cell differentiation. Loss of miR-17-92 cluster expression leads to a corresponding reduction in Tfh-dependent germinal centre formation and plasma cell differentiation.[76] MiR-17, miR-18a, miR-19a and miR-20a are overexpressed in splenic T cells of MRL/*lpr* mice.[77] Similarly, miR-17, miR-17a, miR-18a, miR-19a, miR-19b1 and miR-20a are overexpressed in circulating CD4+ T cells of patients with lupus.[78] MiR-19b1 expression, specifically, has a significant positive correlation with disease activity as measured by SLEDAI score.[78] MiR-17 and miR-20 are downregulated in circulating peripheral blood mononuclear cells,[79] B cells[80] and as circulating free miRNAs[81] in patients with lupus compared with healthy controls, suggesting tissue-specific and miRNA-specific expression patterns. Of the miR-17-92 cluster miRNAs identified as differentially methylated in our analysis, only miR-18a and miR-19b1 showed a significant positive correlation between median expression in naïve CD4+ T cells and disease activity in patients with lupus, consistent with these prior observations. MiR-19b1 promotes proliferation of mature CD4+ T cells, Th1 differentiation and interferon-γ production, and suppresses inducible Treg differentiation.[82] MiR-18a expression increases rapidly early on in CD4+ T cell activation,[83 84] and suppresses Th17 cell differentiation through direct targeting of critical Th17 transcription factor transcripts including SMAD4, HIF1A and RORA in human CD4+ T cells in vitro and in vivo murine airway inflammation models.[83] We did not observe a difference in the expression of members in the miR-17-92 cluster between patients with lupus and controls in naïve CD4+ T cells, likely because these miRNAs are upregulated upon T cell activation. Evidence for hypomethylation in lupus in naïve CD4+ T cells suggests epigenetic priming of this locus, similar to what we previously observed in interferon-regulated gene loci in lupus.[18]

Consistent with our DNA methylation data and the epigenetic priming concept in naïve CD4+ T cells discussed above, gene expression data in total CD4+ T cells isolated from patients with lupus compared with normal healthy controls revealed upregulation of miR-18a in lupus and concomitant downregulation of several genes known to be targeted by miR-18a.[59] Of 74 miR-18a target genes downregulated in lupus CD4+ T cells, our literature-based analysis highlighted 15 genes, including *HIF1A* which is involved in T cell differentiation as discussed above. The most robustly lupus-related gene was *TNFAIP3*, which

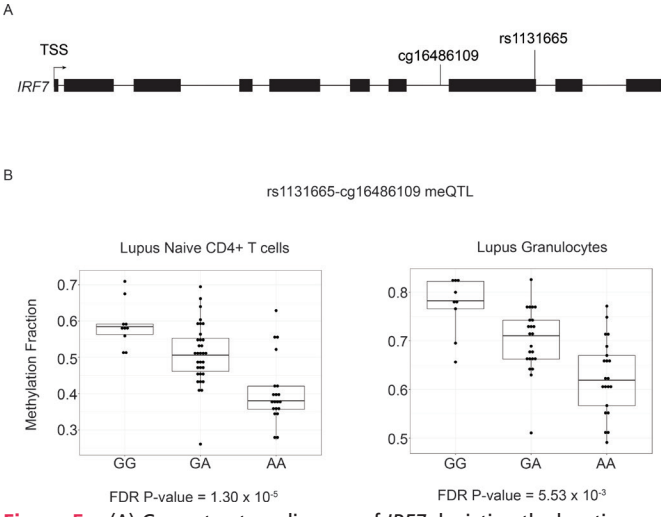

Figure 5 (A) Gene structure diagram of *IRF7* depicting the location of rs1131665 and cg16486109. (B) The presence of the lupus risk allele at rs1131665 (allele A) is associated with significantly lower DNA methylation levels of cg16486109 located in *IRF7*. FDR, false discovery rate; TSS, transcription start site.

encodes the NF-κB negative regulator A20. Indeed, the genetic association between *TNFAIP3* loss-of-function polymorphisms and lupus has been repeatedly confirmed.[85]

Single cell RNA sequencing data from lupus nephritis kidney tissues revealed evidence for expression of *MIR17HG*, the host gene encoding the miR-17-92 cluster, in kidney-infiltrating immune cells, including multiple T cell subsets. Further studies are needed to determine if altered DNA methylation at the miR-17-92 cluster promoter is associated with expression changes with a causal role in the development of lupus, and to determine if methylation levels at this locus can be used as biomarker for monitoring disease activity.

We used analysis of meQTL to identify allele-specific DNA methylation associations across the genome of naïve $CD4^+$ T cells from patients with lupus and healthy controls. Our primary objective was to understand to what extent are DNA methylation changes associated with lupus (the lupus-defining epigenetic profile), explained by genetic factors. We found that <1% of differentially methylated sites in patients with lupus compared with healthy controls were associated with a *cis*-meQTL. This suggests that almost all of the DNA methylation alterations observed in lupus are not associated with local allelic differences in the genome, suggesting a greater contribution from non-genetic and possibly environmental factors to epigenetic dysregulation in lupus. A previous study of meQTL in whole blood of patients with lupus found that a majority of meQTLs were shared between patients and controls.[24] We observed that about 68% of meQTL in patients with lupus and 54% of meQTL in healthy controls were shared by both groups, supporting this observation.

Our prior analysis of granulocytes from a cohort of patients with lupus identified overlap in meQTL genes and lupus genetic risk loci.[22] MeQTL pairs including *ARID5B* (cg13344587-rs10821936), *HLA-DQB1* (cg13047157-rs9274477), and *IRF7* (cg16486109-rs1131665) were found in both granulocytes and naïve $CD4^+$ T cells from patients with lupus. Risk loci genes unique to naïve $CD4^+$ T cell meQTLs included *CD80* (cg06300880-rs3915166), *TYK2* (cg06622468-rs280501), *IKBKE* (cg22577136-rs17020312) and *CTLA4* (cg05092371-rs16840252, cg05092371-rs4553808). Naïve $CD4^+$ T cell-specific meQTL risk loci genes are related to signal response and activation in $CD4^+$ T cells compared with the more general DNA repair and type I interferon signalling seen in the shared meQTL risk loci genes. Disease-relevant meQTL show tissue-specific patterns which should be considered when teasing apart their potential impact.

We identified three meQTL that include SNPs previously identified as lupus genetic risk variants. One meQTL is in the complement factor B gene *CFB* (cg16505946-rs558702), where the risk allele is associated with increased DNA methylation of the nearby CpG site. Complement factor B (CFB) combines with C3 to form the C3 convertase after cleavage by complement factor D as part of the alternative complement pathway. Complement pathway defects have long been studied as a model of monogenic lupus and contribute to increased risk of polygenic lupus.[65] We identified an additional meQTL that included a known lupus risk variant in *IRF7* (cg16486109-rs1131665). Rs1131665 is a missense variant in the inhibitory domain of IRF7 (Q412R). This lupus-associated amino acid change was demonstrated to enhance IRF7-induced expression response in a luciferase reporter assay.[86] This same risk allele is also associated with decreased DNA methylation of cg16486109. Although the relative DNA methylation fractions are different between naïve $CD4^+$ T cells and granulocytes of patients with lupus, the direction of the allele-specific DNA methylation is the same. This suggests that the observed meQTL effect may be present in other lymphoid and myeloid tissues, potentially including plasmacytoid dendritic cells, which are major producers of type I interferons. We describe a direct association between a lupus risk allele and local hypomethylation of a CpG site in *IRF7* in lupus. This observation provides new insights regarding possible biological mechanisms underlying pathogenic consequences of lupus-associated genetic polymorphisms.

In summary, we investigated genome-wide DNA methylation changes in naïve $CD4^+$ T cells from an extended cohort of patients with lupus and controls, and using a methylation trend deviation analysis method, we showed promoter hypomethylation of the miR-17-92 cluster that has a significant regulatory role in T cell growth, function and differentiation. Combining genome-wide DNA methylation and genotyping data, we were able to determine genetic contribution to the lupus-defining epigenotype. Our data indicate that epigenetic changes characteristic of lupus are not under direct genetic influence. This suggests a more important role for non-genetic factors in the epigenetic dysregulation observed in patients with lupus, including the robust demethylation of interferon-regulated genes.

Author affiliations
[1]Division of Rheumatology, Department of Pediatrics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[2]Graduate Program in Immunology, University of Michigan, Ann Arbor, Michigan, USA
[3]Graduate Program, Department of Biochemistry and Molecular Biology, The University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA
[4]Arthritis and Clinical Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA
[5]Division of Rheumatology, Henry Ford Health System, Detroit, Michigan, USA
[6]Division of Rheumatology, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA
[7]Department of Biochemistry and Molecular Biology, The University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA
[8]Division of Rheumatology and Clinical Immunology, Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[9]Lupus Center of Excellence, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA
[10]Department of Immunology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Correction notice This article has been corrected since it published Online First. The wording in the 'Naïve CD4+ T cell methylation quantitative trait loci in patients with lupus' section has been amended.

Contributors All authors fulfilled the following criteria: substantial contributions to the conception or design of the work, or the acquisition, analysis or interpretation of data. Drafting the work or revising it critically for important intellectual content. Final approval of the version published. Dr. Sawalha accepts full responsibility for the work and the conduct of the study, had access to the data, and controlled the decision to publish.

Funding This work was supported by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) grant number R01 AI097134 to AHS. JDW is supported by NIH grants number P30 AG050911 and P20 GM103636.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s)

Ethics approval This study was approved by Institutional Review Board of the University of Pittsburgh (STUDY19020379). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. All relevant data used to generate the results presented, are included in supplementary tables associated with this submission.

ORCID iD

Amr H Sawalha http://orcid.org/0000-0002-3884-962X

REFERENCES

1 Olesińska M, Saletra A. Quality of life in systemic lupus erythematosus and its measurement. *Reumatologia* 2018;56:45–54.
2 Lewis MJ, Jawad AS. The effect of ethnicity and genetic ancestry on the epidemiology, clinical features and outcome of systemic lupus erythematosus. *Rheumatology* 2017;56:i67–77.
3 Kwon Y-C, Chun S, Kim K, *et al*. Update on the genetics of systemic lupus erythematosus: genome-wide association studies and beyond. *Cells* 2019;8:1180.
4 Morris DL, Sheng Y, Zhang Y, *et al*. Genome-Wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. *Nat Genet* 2016;48:940–6.
5 Sun C, Molineros JE, Looger LL, *et al*. High-Density genotyping of immune-related loci identifies new SLE risk variants in individuals with Asian ancestry. *Nat Genet* 2016;48:323–30.
6 Generali E, Ceribelli A, Stazi MA, *et al*. Lessons learned from twins in autoimmune and chronic inflammatory diseases. *J Autoimmun* 2017;83:51–61.
7 Javierre BM, Fernandez AF, Richter J, *et al*. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res* 2010;20:170–9.
8 Fraga MF, Ballestar E, Paz MF, *et al*. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 2005;102:10604–9.
9 Lee PP, Fitzpatrick DR, Beard C, *et al*. A critical role for DNMT1 and DNA methylation in T cell development, function, and survival. *Immunity* 2001;15:763–74.
10 Quddus J, Johnson KJ, Gavalchin J, *et al*. Treating activated CD4+ T cells with either of two distinct DNA methyltransferase inhibitors, 5-azacytidine or procainamide, is sufficient to cause a lupus-like disease in syngeneic mice. *J Clin Invest* 1993;92:38–53.
11 Cornacchia E, Golbus J, Maybaum J, *et al*. Hydralazine and procainamide inhibit T cell DNA methylation and induce autoreactivity. *J Immunol* 1988;140:2197–200.
12 Richardson B, Scheinbart L, Strahler J, *et al*. Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis. *Arthritis Rheum* 1990;33:1665–73.
13 Lu Q, Wu A, Tesmer L, *et al*. Demethylation of CD40LG on the inactive X in T cells from women with lupus. *J Immunol* 2007;179:6352–8.
14 Lu Q, Wu A, Richardson BC. Demethylation of the same promoter sequence increases CD70 expression in lupus T cells and T cells treated with lupus-inducing drugs. *J Immunol* 2005;174:6212–9.
15 Lu Q, Kaplan M, Ray D, *et al*. Demethylation of ITGAL (CD11a) regulatory sequences in systemic lupus erythematosus. *Arthritis Rheum* 2002;46:1282–91.
16 Sawalha AH, Jeffries M, Webb R, *et al*. Defective T-cell ERK signaling induces interferon-regulated gene expression and overexpression of methylation-sensitive genes similar to lupus patients. *Genes Immun* 2008;9:368–78.
17 Jeffries MA, Dozmorov M, Tang Y, *et al*. Genome-Wide DNA methylation patterns in CD4+ T cells from patients with systemic lupus erythematosus. *Epigenetics* 2011;6:593–601.
18 Coit P, Jeffries M, Altorok N, *et al*. Genome-wide DNA methylation study suggests epigenetic accessibility and transcriptional poising of interferon-regulated genes in naïve CD4+ T cells from lupus patients. *J Autoimmun* 2013;43:78–84.
19 Coit P, Yalavarthi S, Ognenovski M, *et al*. Epigenome profiling reveals significant DNA demethylation of interferon signature genes in lupus neutrophils. *J Autoimmun* 2015;58:59–66.
20 Hedrich CM, Mäbert K, Rauen T, *et al*. DNA methylation in systemic lupus erythematosus. *Epigenomics* 2017;9:505–25.
21 Ballestar E, Sawalha AH, Lu Q. Clinical value of DNA methylation markers in autoimmune rheumatic diseases. *Nat Rev Rheumatol* 2020;16:514–24.
22 Coit P, Ortiz-Fernandez L, Lewis EE, *et al*. A longitudinal and transancestral analysis of DNA methylation patterns and disease activity in lupus patients. *JCI Insight* 2020;5:e143654.
23 Lanata CM, Paranjpe I, Nititham J, *et al*. A phenotypic and genomics approach in a multi-ethnic cohort to subtype systemic lupus erythematosus. *Nat Commun* 2019;10:3902.
24 Imgenberg-Kreuz J, Carlsson Almlöf J, Leonard D, *et al*. Dna methylation mapping identifies gene regulatory effects in patients with systemic lupus erythematosus. *Ann Rheum Dis* 2018;77:736–43.
25 Coit P, Ognenovski M, Gensterblum E, *et al*. Ethnicity-specific epigenetic variation in naïve CD4+ T cells and the susceptibility to autoimmunity. *Epigenetics Chromatin* 2015;8:49.
26 Mok A, Solomon O, Nayak RR, *et al*. Genome-Wide profiling identifies associations between lupus nephritis and differential methylation of genes regulating tissue hypoxia and type 1 interferon responses. *Lupus Sci Med* 2016;3:e000183.
27 Hochberg MC. Updating the American College of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
28 Flanagan JM. Epigenome-Wide association studies (EWAS): past, present, and future. *Methods Mol Biol* 2015;1238:51–63.
29 Rahmani E, Yedidim R, Shenhav L, *et al*. GLINT: a user-friendly toolset for the analysis of high-throughput DNA-methylation array data. *Bioinformatics* 2017;33:1870–2.
30 Chen Y-an, Lemire M, Choufani S, *et al*. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 2013;8:203–9.
31 Rahmani E, Shenhav L, Schweiger R, *et al*. Genome-Wide methylation data mirror ancestry information. *Epigenetics Chromatin* 2017;10:1.
32 Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet* 2007;3:1724–35.
33 Pedregosa F, Varoquaux G, Gramfort A. Scikit-learn: machine learning in python. *the Journal of machine Learning research* 2011;12:2825–30.
34 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;8:118–27.
35 Hubbard T, Barker D, Birney E, *et al*. The Ensembl genome database project. *Nucleic Acids Res* 2002;30:38–41.
36 FANTOM Consortium and the RIKEN PMI and CLST (DGT), Forrest ARR, Kawaji H, *et al*. A promoter-level mammalian expression atlas. *Nature* 2014;507:462–70.
37 Dale RK, Pedersen BS, Quinlan AR. Pybedtools: a flexible python library for manipulating genomic datasets and annotations. *Bioinformatics* 2011;27:3423–4.
38 Kuleshov MV, Jones MR, Rouillard AD, *et al*. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016;44:W90–7.
39 Ritchie ME, Phipson B, Wu D, *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.
40 Barrett T, Wilhite SE, Ledoux P, *et al*. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* 2013;41:D991–5.
41 Wang T, Guan W, Lin J, *et al*. A systematic study of normalization methods for Infinium 450K methylation data using whole-genome bisulfite sequencing data. *Epigenetics* 2015;10:662–9.
42 Virtanen P, Gommers R, Oliphant TE, *et al*. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat Methods* 2020;17:261–72.
43 Price AL, Patterson NJ, Plenge RM, *et al*. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904–9.
44 Chang CC, Chow CC, Tellier LC, *et al*. Second-Generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015;4:7.
45 Aryee MJ, Jaffe AE, Corrada-Bravo H, *et al*. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 2014;30:1363–9.
46 Fortin J-P, Triche TJ, Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* 2017;33:558–60.
47 Team RC. *R: a language and environment for statistical computing*, 2020.
48 Sinke L, van Iterson M, Cats D. DNAmArray: streamlined workflow for the quality control, normalization, and analysis of illumina methylation array data. *Zenodoa* 2019.
49 Fortin J-P, Labbe A, Lemire M, *et al*. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol* 2014;15:503.
50 Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res* 2017;45:e22.
51 Leek JT, Johnson WE, Parker HS, *et al*. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012;28:882–3.
52 Chavent M, Kuentz-Simonet V, Labenne A. Multivariate analysis of mixed data: the R package PCAmixdata. *arXiv preprint arXiv* 2014:14114911.
53 Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 2012;28:1353–8.
54 Chen J, Bardes EE, Aronow BJ, *et al*. ToppGene suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009;37:W305–11.
55 Rusinova I, Forster S, Yu S, *et al*. Interferome v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Res* 2013;41:D1040–6.
56 Knight JS, Meng H, Coit P, *et al*. Activated signature of antiphospholipid syndrome neutrophils reveals potential therapeutic target. *JCI Insight* 2017;2:e93897.
57 Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, *et al*. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;518:317–30.
58 Li D, Hsu S, Purushotham D, *et al*. WashU epigenome Browser update 2019. *Nucleic Acids Res* 2019;47:W158–65.
59 Zhao M, Liu S, Luo S, *et al*. Dna methylation and mRNA and microRNA expression of SLE CD4+ T cells correlate with disease phenotype. *J Autoimmun* 2014;54:127–36.
60 Arazi A, Rao DA, Berthier CC, *et al*. The immune cell landscape in kidneys of patients with lupus nephritis. *Nat Immunol* 2019;20:902–14.

61 Bentham J, Morris DL, Graham DSC, *et al*. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat Genet* 2015;47:1457–64.
62 Han J-W, Zheng H-F, Cui Y, *et al*. Genome-Wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet* 2009;41:1234–7.
63 Alarcón-Riquelme ME, Ziegler JT, Molineros J, *et al*. Genome-Wide association study in an Amerindian ancestry population reveals novel systemic lupus erythematosus risk loci and the role of European admixture. *Arthritis Rheumatol* 2016;68:932–43.
64 Lee YH, Bae S-C, Choi SJ, *et al*. Genome-Wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis. *Mol Biol Rep* 2012;39:10627–35.
65 Harley ITW, Sawalha AH. Systemic lupus erythematosus as a genetic disease. *Clin Immunol* 2022;236:108953.
66 Honda K, Yanai H, Negishi H, *et al*. IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature* 2005;434:772–7.
67 Gebert LFR, MacRae IJ. Regulation of microRNA function in animals. *Nat Rev Mol Cell Biol* 2019;20:21–37.
68 Shen N, Liang D, Tang Y, *et al*. MicroRNAs--novel regulators of systemic lupus erythematosus pathogenesis. *Nat Rev Rheumatol* 2012;8:701–9.
69 Pan W, Zhu S, Yuan M, *et al*. Microrna-21 and microRNA-148a contribute to DNA hypomethylation in lupus CD4+ T cells by directly and indirectly targeting DNA methyltransferase 1. *J Immunol* 2010;184:6773–81.
70 Zhao S, Wang Y, Liang Y, *et al*. Microrna-126 regulates DNA methylation in CD4+ T cells and contributes to systemic lupus erythematosus by targeting DNA methyltransferase 1. *Arthritis Rheum* 2011;63:1376–86.
71 Qin H, Zhu X, Liang J, *et al*. Microrna-29B contributes to DNA hypomethylation of CD4+ T cells in systemic lupus erythematosus by indirectly targeting DNA methyltransferase 1. *J Dermatol Sci* 2013;69:61–7.
72 Mogilyansky E, Rigoutsos I. The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death Differ* 2013;20:1603–14.
73 Fu Y, Zhang L, Zhang R, *et al*. Enterovirus 71 suppresses miR-17-92 cluster through up-regulating methylation of the miRNA promoter. *Front Microbiol* 2019;10:625.
74 Dakhlallah D, Batte K, Wang Y, *et al*. Epigenetic regulation of miR-17~92 contributes to the pathogenesis of pulmonary fibrosis. *Am J Respir Crit Care Med* 2013;187:397–405.
75 Xiao C, Srinivasan L, Calado DP, *et al*. Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. *Nat Immunol* 2008;9:405–14.
76 Wu Y, Schutt S, Paz K, *et al*. MicroRNA-17-92 is required for T-cell and B-cell pathogenicity in chronic graft-versus-host disease in mice. *Blood* 2018;131:1974–86.
77 Dai R, Zhang Y, Khan D, *et al*. Identification of a common lupus disease-associated microRNA expression pattern in three different murine models of lupus. *PLoS One* 2010;5:e14302.
78 Qin HH, Zhu XH, Liang J, *et al*. The expression and significance of miR-17-92 cluster miRs in CD4+ T cells from patients with systemic lupus erythematosus. *Clin Exp Rheumatol* 2013;31:472–3.
79 Dai Y, Huang Y-S, Tang M, *et al*. Microarray analysis of microRNA expression in peripheral blood cells of systemic lupus erythematosus patients. *Lupus* 2007;16:939–46.
80 Te JL, Dozmorov IM, Guthridge JM, *et al*. Identification of unique microRNA signature associated with lupus nephritis. *PLoS One* 2010;5:e10344.
81 Carlsen AL, Schetter AJ, Nielsen CT, *et al*. Circulating microRNA expression profiles associated with systemic lupus erythematosus. *Arthritis Rheum* 2013;65:1324–34.
82 Jiang S, Li C, Olive V, *et al*. Molecular dissection of the miR-17-92 cluster's critical dual roles in promoting Th1 responses and preventing inducible Treg differentiation. *Blood* 2011;118:5487–97.
83 Montoya MM, Maul J, Singh PB, *et al*. A distinct inhibitory function for miR-18a in Th17 cell differentiation. *J Immunol* 2017;199:559–69.
84 Teteloshvili N, Smigielska-Czepiel K, Kroesen B-J, *et al*. T-Cell activation induces dynamic changes in miRNA expression patterns in CD4 and CD8 T-cell subsets. *Microrna* 2015;4:117–22.
85 Adrianto I, Wen F, Templeton A, *et al*. Association of a functional variant downstream of TNFAIP3 with systemic lupus erythematosus. *Nat Genet* 2011;43:253–8.
86 Fu Q, Zhao J, Qian X, *et al*. Association of a functional IRF7 variant with systemic lupus erythematosus. *Arthritis Rheum* 2011;63:749–54.