UCSF
UC San Francisco Previously Published Works

Title

Genome-wide profiling identifies associations between lupus nephritis and differential methylation of genes regulating tissue hypoxia and type 1 interferon responses

Permalink

Journal

ISSN

Authors

Mok, Amanda
Solomon, Olivia
Nayak, Renuka R
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

Genome-wide profiling identifies associations between lupus nephritis and differential methylation of genes regulating tissue hypoxia and type 1 interferon responses

Amanda Mok,[1] Olivia Solomon,[1] Renuka R Nayak,[2] Patrick Coit,[3] Hong L Quach,[1] Joanne Nititham,[2] Amr H Sawalha,[3,4] Lisa F Barcellos,[1] Lindsey A Criswell,[2] Sharon A Chung[2]

ABSTRACT

Objective: Previous studies have shown that differential DNA methylation is associated with SLE susceptibility. How DNA methylation affects the clinical heterogeneity of SLE has not been fully defined. We conducted this study to identify differentially methylated CpG sites associated with nephritis among women with SLE.

Methods: The methylation status of 428 229 CpG sites across the genome was characterised for peripheral blood cells from 322 women of European descent with SLE, 80 of whom had lupus nephritis, using the Illumina HumanMethylation450 BeadChip. Multivariable linear regression adjusting for population substructure and leucocyte cell proportions was used to identify differentially methylated sites associated with lupus nephritis. The influence of genetic variation on methylation status was investigated using data from a genome-wide association study of SLE. Pathway analyses were used to identify biological processes associated with lupus nephritis.

Results: We identified differential methylation of 19 sites in 18 genomic regions that was associated with nephritis among patients with SLE (false discovery rate q<0.05). Associations for four sites in HIF3A, IFI44 and PRR4 were replicated when examining methylation data derived from CD4+ T cells collected from an independent set of patients with SLE. These associations were not driven by genetic variation within or around the genomic regions. In addition, genes associated with lupus nephritis in a prior genome-wide association study were not differentially methylated in this epigenome-wide study. Pathway analysis indicated that biological processes involving type 1 interferon responses and the development of the immune system were associated with nephritis in patients with SLE.

Conclusions: Differential methylation of genes regulating the response to tissue hypoxia and interferon-mediated signalling is associated with lupus nephritis among women with SLE. These findings have not been identified in genetic studies of lupus nephritis, suggesting that epigenome-wide association

studies can help identify the genomic differences that underlie the clinical heterogeneity of SLE.

INTRODUCTION

Lupus nephritis continues to cause substantial morbidity and increased mortality for individuals with SLE.[1] The pathogenesis of lupus nephritis involves intrarenal factors such as immune complex formation and deposition that lead to complement activation, as well as extra-renal factors including genetic variants that predispose individuals to a loss of tolerance, leading to autoantibody production against nuclear antigens.[2] In a genome-wide association study of lupus nephritis, the strongest genetic associations with lupus nephritis were observed in the PDGFRA-GSX2, SLC5A11 and ID4 regions. These results indicated that genetic variants outside of the major histocompatibility complex and not previously associated with SLE may have a greater influence in the development of nephritis among patients with SLE.[3]

More recently, DNA methylation, an epigenetic modification that influences gene expression, has been implicated in the pathogenesis of SLE. Two studies have indicated that hypomethylation of interferon-regulated genes in specific leucocyte subsets is associated with SLE susceptibility.[4] [5] Our own work has shown that among individuals with SLE, hypomethylation of interferon-regulated genes is associated with SLE-related autoantibody production.[6] More recently, Coit et al[7] also showed that, when compared with healthy controls, CD4+ T cells from patients

with SLE having nephritis showed a greater degree of hypomethylation of interferon-regulated genes than patients with SLE not having nephritis.

In this study, we build on our previous work by conducting the largest genome-wide DNA methylation study of lupus nephritis to date to identify genes and pathways that are differentially methylated between SLE cases with and without nephritis. We also examine the influence of genetic variation on the identified associations, as well as methylation status of previously identified lupus nephritis susceptibility genes among patients with SLE with and without nephritis.

METHODS
Study subjects
The primary subject group for this study comprised 80 patients with SLE having nephritis and 245 patients with SLE not having nephritis. All SLE cases were participants of the University of California, San Francisco Lupus Genetics Project and fulfilled at least four American College of Rheumatology classification criteria for SLE. To minimise confounding, all SLE cases were women of European descent who had never smoked. Lupus nephritis status was determined by medical record review. SLE cases were considered to have lupus nephritis if they fulfilled the renal classification criterion for SLE (persistent proteinuria >0.5 gm/day or >3+ on urine dipstick or presence of cellular casts)[8] or had a renal biopsy showing lupus nephritis. A SLE case was considered not to have lupus nephritis if no evidence of renal involvement (eg, lacking haematuria and proteinuria) was identified on medical record review.

To replicate significant findings, we used the dataset generated by Coit et al.[7] This study assessed the methylation status of CD4+ T cells from 56 women with SLE. A total of 28 women had lupus nephritis and 28 did not have lupus nephritis. The two groups were ethnically matched (17 European-American and 11 African-American individuals in each group). Additional details regarding this study group are provided in Coit et al.[7] To assess for overlap between the two cohorts, we compared dates of birth and recruitment site for the individuals in each cohort, since genome-wide genetic data were not available for individuals in the replication cohort. Only one birth date was identical between the primary and replication cohorts and the individuals with this date of birth were recruited from different states within the USA. Thus, there is likely no overlap between the two cohorts.

DNA methylation assessment
DNA methylation profiling of genomic DNA from peripheral blood cells was performed using the Illumina HumanMethylation450 BeadChip, as previously described.[6] This high-throughput platform assesses the methylation status for ~480 000 CpG sites in ~23 000 genes and assesses methylation for sites in promoters, 5′ and 3′ regions, gene bodies, CpG islands, CpG island shores and outside of CpG islands. For each site, the primary output is the β-value, which is the ratio of the methylated probe intensity to the overall probe intensity and represents the proportion of DNA that is methylated at that site (0=completely unmethylated, 1=completely methylated).

Prior to analysis, we performed extensive normalisation and quality control measures. We removed sites which had a low detection rate (p value for detection being <0.05) in more than 20% of samples (289 CpG sites). All samples met the detection threshold (<5% missing). To normalise the data, background signal was subtracted using the 'noob' method with the R package Methylumi.[9] All sample mean normalisation was used to normalise colour channels.[10] The final normalisation step was β-mixture quantile normalisation to normalise for differences between type 1 and type 2 probes on the Illumina Infinium platform.[11] We also removed cross-reactive probes that may hybridise to more than one genomic location (30969), polymorphic probes which are known to contain a single nucleotide polymorphism (SNP) commonly found in populations of European ancestry (29 027) and non-CpG rs SNP probes (65).[12] We calculated correlations between replicate samples (38 duplicate pairs) as a quality control measure and found $r^2 > 0.98$ for all sites across the duplicate pairs. From the original dataset of 325 samples, we removed three samples that were identified as outliers through principal components analysis of methylation data. The final dataset consisted of 428 229 CpG sites assessed in 322 individuals with SLE, of whom 80 SLE cases had nephritis and 242 SLE cases did not have evidence of nephritis.

Methylation status of the 56 individuals used as the replication cohort was also assessed using the Illumina HumanMethylation450 BeadChip. Details regarding the data processing for this group are provided in Coit et al.[7]

Cell population estimation
DNA methylation patterns can vary between different blood cell populations.[13] In order to account for cell-type heterogeneity between samples, we calculated principal components using the ReFACTor method[14] to be used in multivariable regression. ReFACTor is a reference-free method that performs a principal component analysis (PCA) on the subset of CpG sites that are differentially methylated across different cell types to adjust for variance due to cell-type heterogeneity. We included the first two ReFACTor principal components in regression analyses to adjust for cell heterogeneity based on the examination of quantile-quantile plots for inflation and deflation of p values and calculation of genomic inflation factors (λ=0.90 for the first two principal components).

Population substructure
Population substructure may confound the associations between lupus nephritis and methylation since genetic ancestry may be associated with both methylation[15] and

disease status. Therefore, a genetic PCA was computed in EIGENSTRAT, as previously described,[6] using SNP data generated with the Illumina HumanHap500 and ImmunoChip genotyping platforms. The first principal component was included in multivariable linear regression analysis to adjust for confounding due to population substructure.

Statistical analyses

All regression analyses were conducted in R (V.3.2.2) (R Core Team. R: A Language and Environment for Statistical Computing [program], Vienna, Austria, 2015. http://www.R-project.org/). For each of the 428 229 CpG sites under study, multivariable linear regression was used to determine the association between lupus nephritis status and methylation. For these regression models, methylation β-values were used as the outcome and lupus nephritis status was the primary predictor. Disease duration, the first genetic principal component and the first two ReFACTor principal components were included as covariates. The Benjamini-Hochberg false discovery rate (FDR) method[16] was used to adjust for multiple comparisons, with an FDR q-value of <0.05 considered statistically significant.

All sites with an FDR q<0.05 (n=19) were further examined in an independent dataset generated by Coit et al[7] for replication. For this dataset, association with lupus nephritis was determined using Wilcoxon rank-sum tests (STATA V.13.1, College Station, Texas, USA), with a stringent Bonferroni-corrected threshold of p<0.0026 (0.05/19) considered statistically significant.

To assess the role of genetic variation on the identified associations between DNA methylation sites and lupus nephritis, we used SNP data available from 264 of the study subjects from a genome-wide association study of SLE.[17] We identified SNPs on the Illumina HumanHap500 BeadChip located 500 kb upstream and downstream of the associated CpG sites and analysed SNPs with a minor allele frequency of >0.05. For each CpG, we ran multivariable regression analyses as previously described and included the SNPs in the vicinity of the CpG as a covariate, assuming an additive model. Each CpG-SNP pair was analysed separately. FDR q-values were calculated based on the number of CpG-SNP pairs for a CpG site, with FDR q<0.05 considered statistically significant.

We further examined if methylation of previously identified lupus nephritis susceptibility genes[3] was associated with nephritis among SLE cases. We examined the CpG sites on the HumanMethylation450 BeadChip that were within 5000 bp upstream and downstream of the genes. We readjusted the p values using FDR correction for this smaller subset of CpG sites from the final model in order to determine if any of these sites were differentially methylated between the two nephritis subgroups.

Last, we conducted a pathway analysis using the gometh() function in missMethyl R package to determine if differential methylation of genes in particular pathways were associated with lupus nephritis. This analysis used 1000 CpGs with the strongest evidence of association with lupus nephritis based on p values from multivariable regression models (without SNPs). The missMethyl package was developed specifically for the Illumina HumanMethylation450 array and the gometh() function tests for Gene Ontology enrichment for significant CpGs, taking into account the different number of methylation probes per gene on the array.

RESULTS

The clinical characteristics of the 322 SLE cases examined in this study are presented in table 1. To minimise confounding by sex, race/ethnicity and smoking status, all study participants in the primary dataset were women of European descent who never smoked. SLE cases with nephritis had a longer disease duration and, not unexpectedly, a higher frequency of autoantibody positivity. Of the 50 SLE cases with nephritis who underwent renal biopsy and had pathology reports available for review, 32 (64%) had class III/IV (proliferative) nephritis and 10 (20%) had class V (membranous) nephritis.[18]

We first performed a genome-wide assessment of differential DNA methylation associated with lupus nephritis. Results from the multivariable regression analysis

Table 1 Characteristics of the 322 patients with SLE in this study

Characteristic	SLE with nephritis (n=80)	SLE without nephritis (n=242)
Age at blood draw, mean years (SD)	43.20 (12.86)	46.36 (13.18)
Disease duration, mean years (SD)	16.25 (10.11)	10.69 (8.02)
ACR classification criteria for SLE, n (%)		
ANA positivity	76 (95.0)	229 (94.6)
Malar rash	43 (53.8)	110 (45.5)
Discoid rash	3 (3.8)	7 (2.9)
Oral ulcers	19 (23.8)	80 (33.1)
Photosensitivity	59 (73.8)	201 (83.1)
Arthritis	57 (71.3)	186 (76.9)
Serositis	25 (31.3)	69 (28.5)
Neurological disorder	11 (13.8)	19 (7.9)
Haematological disorder	64 (80.0)	151 (62.4)
Immunological disorder	67 (83.8)	148 (61.2)
SLE-related autoantibodies, n (%)		
Anti-dsDNA	56 (70.0)	98 (40.5)
Anti-SSA/Ro	24 (31.6*)	62 (25.6*)
Anti-SSB/La	10 (13.3*)	23 (9.5*)
Anti-Sm	14 (18.7*)	23 (9.5*)
Anti-RNP	18 (24.0*)	38 (15.7*)

*Percentages are not based on all 322 participants due to missing values.
ACR, American College of Rheumatology; ds, double-stranded; RNP, ribonucleoprotein; SSA, Sjögren's syndrome related antigen A; SSB, Sjögren's syndrome related antigen B.
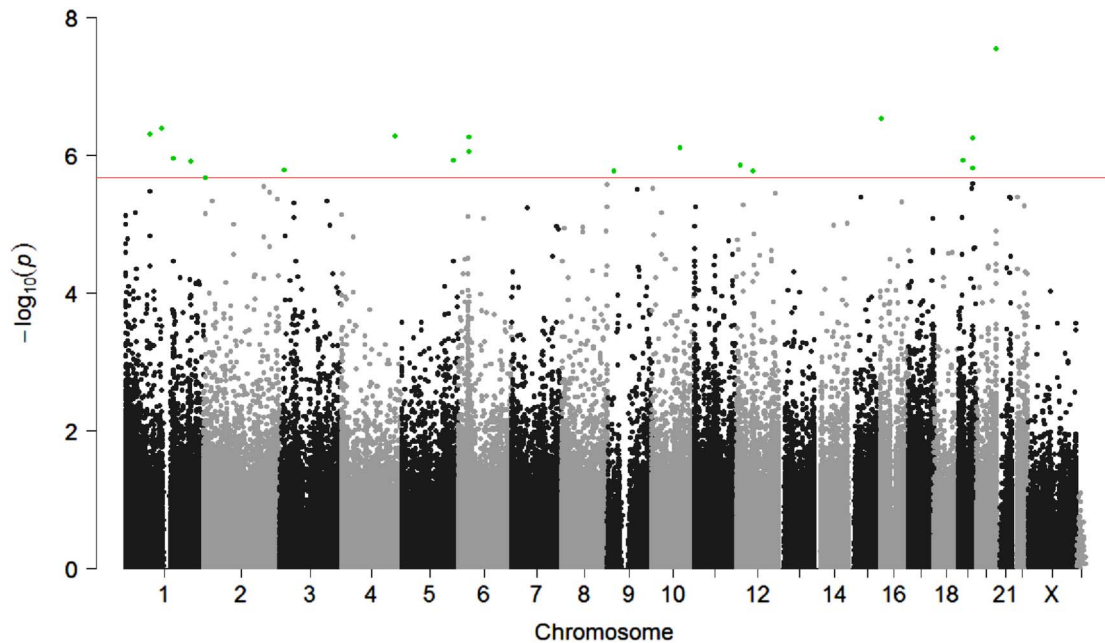
Figure 1 Manhattan plot of DNA methylation sites associated with lupus nephritis based on a multivariable linear regression. The red line shows a false discovery rate (FDR) cut-off of significance level p<2.1E-06. Significant CpG sites are highlighted in green.

are shown in figure 1. The methylation status of 19 CpG sites had statistically significant evidence of association with lupus nephritis (FDR q<0.05) after adjusting for disease duration, population substructure and peripheral blood cell heterogeneity (table 2). Eight sites were more methylated in patients with SLE having nephritis compared with those without nephritis, with the difference in methylation ranging from 1.4% to 6.4%. Among these sites, the most hypermethylated site was cg23510764 on chromosome 5. Eleven sites were less methylated in SLE cases with nephritis compared with those without nephritis, with the difference in methylation ranging from −2.3% to −12.4% between SLE cases with and without nephritis. The most hypomethylated site among these 11 sites was cg05552874 located in *IFIT1*.

We then sought replication for these findings using DNA methylation data generated with the same platform from CD4+ T cells from 56 patients with SLE.[7] Out of the 19 sites we identified in our initial analyses, methylation of four sites in or near hypoxia-inducible transcription factor *(HIF)3A*, interferon-induced protein 44 *(IFI44)* and lacrimal proline-rich protein 4 *(PRR4)* in CD4+ T cells showed evidence of association with lupus nephritis in Wilcoxon rank-sum tests (p<0.0026), as shown in table 2.

Next, we examined whether methylation of 15 previously identified lupus nephritis susceptibility genes[3] was associated with lupus nephritis, to determine if epigenetic associations are driven by genetic variation. We identified 12 794 CpG sites that were located within 5000 bp upstream or downstream of the lupus nephritis susceptibility genes for this analysis. Even with readjustment of

the FDR q-value for only these sites, no site showed evidence of association with lupus nephritis (FDR q>0.1, data not shown).

To determine if the associations between methylation status and lupus nephritis could be confounded by genetic variation (ie, that genetic variation influences both the development of lupus nephritis and extent of methylation at a site), we used genome-wide SNP data available on 264 study subjects who had been previously genotyped with the Illumina HumanHap500 BeadChip. SNPs within 500 kb of the associated CpG site were included in multivariable regression analyses. Online supplementary table S1 shows the total number of SNPs studied for each CpG site and the number of SNPs with FDR q<0.05 in multivariable regression modelling. SNPs in or near *HELZ2, HIF3A, MIR-29B-2/LOC148696* and *PRR4* showed evidence of association with methylation (FDR q<0.05) but did not substantially alter the association between methylation and lupus nephritis status (see online supplementary table S2). These findings suggest that the observed associations between methylation status and lupus nephritis do not appear to be driven by genetic variation of these genes.

We also examined whether these results could be confounded by medication exposure. While we were unable to adjust for medication dose at sample collection (data not available), we ascertained previous exposure to prednisone, hydroxychloroquine, azathioprine, methotrexate, mycophenolate mofetil and cyclophosphamide by medical record review and patient questionnaire. While medication exposure did differ between those with and without lupus nephritis as expected (eg, higher prevalence of mycophenolate mofetil and cyclophosphamide

Table 2 Differentially methylated CpG sites associated with nephritis in women with SLE

CpG	Chr	Position (hg19)	Gene(s)	Primary dataset from whole blood (n=80 SLE with nephritis, n=242 SLE without nephritis)			Replication dataset from CD4+T cells (n=28 SLE with nephritis, n=28 SLE without nephritis)			
				β*	Regression p value	FDR q-value	Median SLE with nephritis	Median SLE without nephritis	Median difference	Wilcoxon p value
cg15262954	20	62198871	*HELZ2*	0.014	2.84E-08	0.012	0.307	0.213	0.094	0.25
cg19168338	16	4465730	*CORO7*	0.016	2.89E-07	0.035	0.323	0.318	0.005	0.31
cg00356916	1	116256617	*CASQ2*	0.015	4.01E-07	0.035	0.619	0.604	0.015	0.67
cg01079652	**1**	**79118190**	***IFI44***	**−0.086**	**4.96E-07**	**0.035**	**0.687**	**0.814**	**−0.127**	**0.0013**
cg05883128	4	169239130	*DDX60*	−0.093	5.30E-07	0.035	0.124	0.151	−0.027	0.16
cg25114611	6	35696869	*FKBP5; LOC285847*	−0.026	5.46E-07	0.035	0.381	0.403	−0.022	0.024
cg22891070	**19**	**46801641**	***HIF3A***	**−0.094**	**5.69E-07**	**0.035**	**0.700**	**0.807**	**−0.106**	**0.0001**
cg05552874	10	91153142	*IFIT1*	−0.124	7.87E-07	0.04	0.516	0.691	−0.175	0.0069
cg18468088	6	35490817	No gene within 5 kb	−0.026	8.91E-07	0.04	0.104	0.13	−0.026	0.33
cg00791854	1	154392069	*IL6R*	0.023	1.11E-06	0.04	0.482	0.474	0.008	0.78
cg23510764	5	166419954	No gene within 5 kb	0.064	1.17E-06	0.04	0.773	0.789	−0.016	0.0096
cg02812482	19	13974981	No gene within 5 kb	0.023	1.19E-06	0.04	0.4	0.396	0.004	0.95
cg04845871	1	207996318	*MIR-29B-2†, LOC148696†*	0.019	1.22E-06	0.04	0.363	0.353	0.010	0.074
cg23256579	**12**	**11002402**	***PRR4***	**−0.05**	**1.41E-06**	**0.041**	**0.661**	**0.712**	**−0.051**	**0.0015**
cg16672562	**19**	**46801671**	***HIF3A***	**−0.092**	**1.55E-06**	**0.041**	**0.725**	**0.840**	**−0.114**	**0.0001**
cg16032408	3	14466410	*SLC6A6*	−0.025	1.61E-06	0.041	0.342	0.337	0.005	0.62
cg13437525	12	51318486	*METTL7A‡*	−0.023	1.67E-06	0.041	0.256	0.284	−0.028	0.12
cg14519664	9	20356458	*MLLT3*	0.037	1.71E-06	0.041	0.243	0.236	0.007	0.34
cg10959651	2	7018019	*RSAD2*	−0.054	2.11E-06	0.048	0.105	0.132	−0.027	0.069

Associations with bolded sites were replicated in CD4+ T cells.
*β represents the difference in methylation between SLE cases with nephritis and SLE cases without nephritis adjusted for disease duration, population substructure and mixed cell populations. Positive values indicate that the site is more methylated in SLE cases with nephritis, while negative values indicate the site in less methylated in SLE cases with nephritis.
†Found within 5 kb downstream of the site.
‡Found within 5 kb upstream of the site.
FDR, false discovery rate.

Genetics

Table 3 The top 10 Gene Ontology groups represented by the 1000 CpG sites with the strongest evidence of association with lupus nephritis

Gene Ontology accession #	Group/term	No. of genes in group	No. of genes represented in the top 1000 CpG sites	p Value based on the number of differentially expressed genes	FDR
GO:0034340	Response to type 1 interferon	76	17	1.77E-10	3.68E-06
GO:0060337	Type 1 interferon signalling pathway	72	16	6.80E-10	4.72E-06
GO:0071357	Cellular response to type 1 interferon	72	16	6.80E-10	4.72E-06
GO:0002376	Immune system process	2544	138	2.01E-08	1.05E-04
GO:0051607	Defence response to virus	315	30	1.04E-07	4.32E-04
GO:0098542	Defence response to other organisms	521	38	2.09E-07	7.25E-04
GO:0009615	Response to virus	396	34	5.09E-07	0.00151
GO:0065007	Biological regulation	10877	443	1.09E-06	0.00283
GO:0045071	Negative regulation of viral genome replication	48	10	1.75E-06	0.00405
GO:0006952	Defence response	1758	94	2.19E-06	0.00455

FDR, false discovery rate; GO, gene ontology.

use in patients with SLE having lupus nephritis compared with patients with SLE not having nephritis), adjusting for previous medication exposure did not significantly alter the associations between methylation of 19 sites in table 2 and lupus nephritis (data not shown).

Last, we performed a pathway analysis to determine if genes with evidence of differential methylation associations with lupus nephritis represent particular biological pathways. Using Gene Ontology as the basis for this analysis, the top 1000 CpG sites with the strongest evidence of association by regression p value were studied. Thirty-three Gene Ontology groups were found to be over-represented in the top 1000 CpG sites. The 10 most significantly associated Gene Ontology groups as determined by FDR q-value are shown in table 3. These results indicate that differential methylation of genes involved in interferon signalling and host response pathways show the strongest evidence of association with lupus nephritis among individuals with SLE.

DISCUSSION

In this study, we present the largest genome-wide assessment of DNA methylation in lupus nephritis to date. In our initial samples obtained from whole blood, we identified evidence for differential methylation of 19 sites in 18 genomic regions associated with lupus nephritis. Association with four sites in *HIF3A*, *IFI44* and *PRR4* were subsequently replicated when we examined DNA methylation of CD4+ T cells from an independent set of women with SLE. Observed methylation associations are not driven by DNA sequence variation surrounding the associated CpG site. In pathway analyses, differential methylation of interferon signalling and host defence genes had the strongest evidence of association with lupus nephritis.

Two of the replicated associations, cg22891070 and cg16672562, are located approximately 200 bp upstream from the transcription start site or within the gene body of *HIF3A*, depending on the gene transcript. *HIF3A* encodes the third isoform of the α subunit of the HIF, which can activate the transcription of over 100 genes involved with the cell's response to hypoxia, including genes involved in angiogenesis, glucose metabolism and cell proliferation. The role of HIF3A is less well delineated compared with the two other isoforms, in part because *HIF3A* mRNA undergoes alternative splicing, with at least seven known variants. Specific variants of *HIF3A* have been shown to act as a negative regulator of HIF1 and HIF2, while overexpression of other variants upregulate specific HIF target genes.[19] Hypermethylation (instead of the hypomethylation observed in this study) of these CpG sites has been associated with increased body mass index.[20] Previous epidemiological studies have suggested that increased body mass index is associated with chronic kidney disease,[21] even among patients with SLE.[22] Thus, it is not evident that the associations between lupus nephritis and hypomethylation of *HIF3A* is due to confounding by body mass index. Chronic hypoxia in the tubulointerstitium of the kidney is proposed to be a common pathway that leads to end-stage renal disease and thus HIF plays a critical pathophysiological role in this process. HIF is also being studied as a target for therapies to prevent progression of chronic kidney disease.[23] Our findings for *HIF3A* are novel, since neither genetic variation nor methylation of HIF genes has been associated with SLE or specific SLE-related subphenotypes in prior studies. Hypomethylation of HIF-related genes may contribute to the development of lupus nephritis or occur as the result of organ damage from nephritis in patients with SLE. In either scenario,

therapy targeting the HIF pathway may provide a novel strategy to decrease the progression of kidney damage for patients with lupus nephritis.

We also replicated an association between lupus nephritis and hypomethylation of cg01079652, which is located within the gene body of *IFI44*. While the role of IFI44 is not completely defined, some evidence suggests that this interferon-inducible gene acts as a suppressor of viral replication.[24] In the study by Coit *et al*,[7] hypomethylation of this CpG site was identified when comparing SLE cases with nephritis with healthy controls and was not identified when SLE cases without nephritis were compared with healthy controls. In a study by Absher *et al*,[5] this site was significantly hypomethylated in CD4+ T cells, CD19+ B cells and CD14+ monocytes when comparing patients with SLE with healthy controls. Of note, hypomethylation of *IFI44L*, a paralog of *IFI44*, was found to be associated with SLE,[4] SLE-related autoantibody production[6] by our group and more recently was shown to be a biomarker for SLE when compared with individuals with other autoimmune diseases and healthy controls.[25]

The third replicated association was with cg23256579, which occurs ~1500 bases upstream or within the body of *PRR4*, depending on the transcript variant. *PRR4* has been identified as a potential biomarker of the dry-eye syndrome.[26] Read-through transcription also exists between this gene and the upstream gene *PRH1* (proline-rich protein HaeIII subfamily 1). This transcript is a candidate for nonsense-mediated mRNA decay and is unlikely to produce a protein product. No previous associations between methylation of this site and SLE-related phenotypes have been reported.

We also found evidence of differential methylation of 15 CpG sites near or within *HELZ2, COR07, CASQ2, DDX60, FKBP5, IFIT1, IL6R, MIR-29B-2b, SLC6A6, METTL7A, MLLT3* and *RSAD2* in whole blood, which did not meet our stringent threshold for replication in the CD4+ T-cell dataset. Of the 15 sites, 13 showed concordant direction of differential methylation between the initial and replication cohorts. Two sites, located in/near in *IFIT1* and *FKBP5* had suggestive statistical evidence of association in the replication cohort (p=0.0069 and p=0.024, respectively). While a lack of replication may indicate false-positive results, these findings may not have replicated due to insufficient statistical power given the small sample size of the replication dataset, the differences in race/ethnic compositions between the two sample groups or because the differential methylation observed for these sites are not driven by CD4+ T cells. Differential methylation of these genes in other cell populations may be more relevant for the development of lupus nephritis. Differential methylation of sites near or within *IFIT1* and *RSAD2* have been previously associated with SLE-related autoantibody production,[6] while differential methylation of the associated sites in *HELZ2* (also known as *PRIC285*), *DDX60, IFIT1, CASQ2, IFI44* and *FKBP5* has been observed in specific leucocyte

subsets among patients with SLE compared with healthy controls.[4][5] Thus, while the p value for association with lupus nephritis for these sites did not meet our threshold for replication, methylation of many of these sites is likely associated with SLE-related outcomes given their previously identified association with SLE susceptibility.

Previously, we showed that methylation of CpG sites in 11 genes was associated with anti-double-stranded (ds) DNA and SLE-related autoantibody production and 7 out of the 11 genes either induced or regulated type 1 interferon signalling. Given the strong correlation between anti-dsDNA autoantibodies and lupus nephritis, we expected to see interferon-related genes dominate the list of statistically significant results. However, among the 18 genomic regions with the strongest statistical evidence of association with lupus nephritis in this study, only four are known to be induced by type 1 interferon (*IFI44, IFIT1, DDX60* and *RSAD2*). Thus, in contrast to SLE-related autoantibodies, the majority of the most significantly associated methylation sites with lupus nephritis are not in interferon-related genes. This finding may represent a key pathophysiological difference between autoantibody production and lupus nephritis. However, interferon-related genes do likely influence the development of lupus nephritis to some extent, given the results of the pathway analyses.

Associations with DNA methylation and lupus nephritis were also examined by Coit *et al*.[7] However, their results are not directly comparable with ours given differences in study design—instead of performing a within-SLE case analysis as in this study, SLE cases with and without nephritis were compared with healthy controls. However, similar to our findings, Coit *et al* observed that interferon-related genes were even less methylated in SLE cases with lupus nephritis than SLE cases without nephritis when both groups were compared with healthy controls.

Strengths of our study include the large sample size and selection of SLE cases to minimise confounding due to sex, ethnicity and smoking status. Unlike many studies of DNA methylation, we were able to adjust for differences in genetic ancestry among the SLE cases using genome-wide SNP data. Our analyses also accounted for disease duration to mitigate the effect of the disease itself on DNA methylation status and the potential misclassification of lupus nephritis status. Last, we employed a reference-free method to estimate and adjust for potential differences in the leucocyte composition between study subjects.

Limitations of our study include its cross-sectional nature. Thus, we are unable to determine if these changes in methylation occur prior to or after the development of lupus nephritis. However, the identification of novel disease-relevant pathways provides insights into the pathogenesis and progression of lupus nephritis, even if these pathways are not the biological processes that initiate disease. In addition, the body mass index of the participants was not available to include in our analyses. Last, we examined DNA from peripheral blood

cells since it can be easily collected in a clinical setting for the development of future clinical tests and is readily available from our study collection. Therefore, we could not assess which specific leucocyte population harbours these DNA methylation changes. However, our findings for CpG sites within or near *HIF3A*, *PRR4* and *IFI44* were replicated in a dataset derived from CD4+ T cells, which suggests that methylation changes in CD4+ T cells contribute to the differences we identified in our study.

In summary, we have shown that hypomethylation of CpG sites in or near tissue hypoxia and interferon signalling response genes is associated with lupus nephritis. Our results support the premise that genetic and epigenetic variations can be independently associated with disease manifestations. Thus, epigenome-wide association studies can identify novel pathways associated with disease manifestations not identified by genome-wide association studies and can provide complementary insight into the pathogenesis and progression of genetically complex autoimmune diseases.

Author affiliations
[1]Genetic Epidemiology and Genomics Laboratory, School of Public Health, University of California, Berkeley, Berkeley, California, USA
[2]Russell/Engleman Rheumatology Research Center, University of California, San Francisco, San Francisco, California, USA
[3]Division of Rheumatology, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA
[4]Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA

Competing interests None declared.

Ethics approval Institutional Review Board, University of California, San Francisco, USA.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Methylation data for the study participants are available at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000947.v1.p1. Additional phenotype data is available to academic investigators by contacting SAC, the corresponding author.

REFERENCES

1. Ortega LM, Schultz DR, Lenz O, *et al.* Review: Lupus nephritis: pathologic features, epidemiology and a guide to therapeutic decisions. *Lupus* 2010;19:557–74.
2. Lech M, Anders H-J. The pathogenesis of lupus nephritis. *J Am Soc Nephrol* 2013;24:1357–66.
3. Chung SA, Brown EE, Williams AH, *et al.* Lupus nephritis susceptibility loci in women with systemic lupus erythematosus. *J Am Soc Nephrol* 2014;25:2859–70.
4. 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.
5. Absher DM, Li X, Waite LL, *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:e1003678.
6. Chung SA, Nititham J, Elboudwarej E, *et al.* Genome-wide assessment of differential DNA methylation associated with autoantibody production in systemic lupus erythematosus. *PLoS ONE* 2015;10:e0129813.
7. Coit P, Renauer P, Jeffries MA, *et al.* Renal involvement in lupus is characterized by unique DNA methylation changes in naïve CD4+ T cells. *J Autoimmun* 2015;61:29–35.
8. Tan EM, Cohen AS, Fries JF, *et al.* The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271–7.
9. Dedeurwaerder S, Defrance M, Bizet M, *et al.* A comprehensive overview of Infinium HumanMethylation450 data processing. *Brief Bioinformatics* 2014;15:929–41.
10. Yousefi P, Huen K, Aguilar Schall R, *et al.* Considerations for normalization of DNA methylation data by Illumina 450K BeadChip assay in population studies. *Epigenetics* 2014;8:1141–52.
11. Teschendorff AE, Marabita F, Lechner M, *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 2013;29:189–96.
12. Chen YA, Lemire M, Choufani S, *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 2014;8:203–9.
13. Reinius LE, Acevedo N, Joerink M, *et al.* Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS ONE* 2012;7:e41361.
14. Rahmani E, Zaitlen N, Baran Y, *et al.* Sparse PCA corrects for cell type heterogeneity in epigenome-wide association studies. *Nat Methods* 2016;13:443–5.
15. 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.
16. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Statist Soc Ser B (Methodological)* 1995;57:289–300.
17. Hom G, Graham RR, Modrek B, *et al.* Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N Engl J Med* 2008;358:900–9.
18. Weening JJ, D'Agati VD, Schwartz MM, *et al.* The classification of glomerulonephritis in systemic lupus erythematosus revisited. *J Am Soc Nephrol* 2004;15:241–50.
19. Heikkilä M, Pasanen A, Kivirikko KI, *et al.* Roles of the human hypoxia-inducible factor (HIF)-3alpha variants in the hypoxia response. *Cell Mol Life Sci* 2011;68:3885–901.
20. Dick KJ, Nelson CP, Tsaprouni L, *et al.* DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 2014;383:1990-8.
21. Chertow GM, Hsu CY, Johansen KL. The enlarging body of evidence: obesity and chronic kidney disease. *J Am Soc Nephrol* 2006;17:1501–2.
22. Rizk A, Gheita TA, Nassef S, *et al.* The impact of obesity in systemic lupus erythematosus on disease parameters, quality of life, functional capacity and the risk of atherosclerosis. *Int J Rheum Dis* 2012;15:261–7.
23. Shoji K, Tanaka T, Nangaku M. Role of hypoxia in progressive chronic kidney disease and implications for therapy. *Curr Opin Nephrol Hypertens* 2014;23:161–8.
24. Power D, Santoso N, Dieringer M, *et al.* IFI44 suppresses HIV-1 LTR promoter activity and facilitates its latency. *Virology* 2015;481:142–50.
25. Zhao M, Zhou Y, Zhu B, *et al.* IFI44L promoter methylation as a blood biomarker for systemic lupus erythematosus. *Ann Rheum Dis* 2016;75:1998–2006.
26. Aluru SV, Agarwal S, Srinivasan B, *et al.* Lacrimal Proline Rich 4 (LPRR4) protein in the tear fluid is a potential biomarker of dry eye syndrome. *PLoS ONE* 2012;7:e51979.