Cell lineage-specific genome-wide DNA methylation analysis of patients with paediatric-onset systemic lupus erythematosus

Kit San Yeung, Tsz Leung Lee, Mo Yin Mok, Christopher Chun Yu Mak, Wanling Yang, Patrick Chun Yin Chong, Pamela Pui Wah Lee, Marco Hok Kung Ho, Sanaa Choufani, Chak Sing Lau, Yu Lung Lau, Rosanna Weksberg & Brian Hon Yin Chung

To cite this article: Kit San Yeung, Tsz Leung Lee, Mo Yin Mok, Christopher Chun Yu Mak, Wanling Yang, Patrick Chun Yin Chong, Pamela Pui Wah Lee, Marco Hok Kung Ho, Sanaa Choufani, Chak Sing Lau, Yu Lung Lau, Rosanna Weksberg & Brian Hon Yin Chung (2019) Cell lineage-specific genome-wide DNA methylation analysis of patients with paediatric-onset systemic lupus erythematosus, Epigenetics, 14:4, 341-351, DOI: 10.1080/15592294.2019.1585176

To link to this article: https://doi.org/10.1080/15592294.2019.1585176
Cell lineage-specific genome-wide DNA methylation analysis of patients with paediatric-onset systemic lupus erythematosus

Kit San Yeung, Tsz Leung Lee, Mo Yin Mok, Christopher Chun Yu Mak, Wanling Yang, Patrick Chun Yin Chong, Pamela Pui Wah Lee, Marco Hok Kung Ho, Sanaa Choufani, Chak Sing Lau, Yu Lung Lau, Rosanna Weksberg, and Brian Hon Yin Chung

*Department of Paediatrics and Adolescent Medicine, The University of Hong Kong, Hong Kong, China; †The Hong Kong Children’s Hospital, Hong Kong, China; ‡Department of Biomedical Sciences, The City University of Hong Kong, Hong Kong, China; §Genetics and Genome Biology Program, The Hospital for Sick Children, Toronto, Ontario, Canada; ¶Department of Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China; and #Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada.

ABSTRACT
Patients with paediatric-onset systemic lupus erythematosus (SLE) often present with more severe clinical courses than adult-onset patients. Although genome-wide DNA methylation (DNAm) profiling has been performed in adult-onset SLE patients, parallel data on paediatric-onset SLE are not available. Therefore, we undertook a genome-wide DNAm study in paediatric-onset SLE patients across multiple blood cell lineages. The DNAm profiles of four purified immune cell lineages (CD4+ T cells, CD8+ T cells, B cells and neutrophils) and whole blood were compared in 16 Chinese patients with paediatric-onset SLE and 13 healthy controls using the Illumina HumanMethylationEPIC BeadChip. Comparison of DNAm in whole blood and within each independent cell lineage identified a consistent pattern of loss of DNAm at 21 CpG sites overlapping 15 genes, which represented a robust, disease-specific DNAm signature for paediatric-onset SLE in our cohort. In addition, cell lineage-specific changes, involving both loss and gain of DNAm, were observed in both novel genes and genes with well-described roles in SLE pathogenesis. This study also highlights the importance of studying DNAm changes in different immune cell lineages rather than only whole blood, since cell type-specific DNAm changes facilitated the elucidation of the cell type-specific molecular pathophysiology of SLE.

ARTICLE HISTORY
Received 12 October 2018
Revised 10 February 2019
Accepted 12 February 2019

KEYWORDS
SLE; lupus; paediatric-onset; DNA methylation; MethylationEPIC; blood; cell lineage-specific; cell composition

Introduction
Systemic lupus erythematosus (SLE; OMIM 152700) is characterized by chronic inflammation, autoantibody production and the presence of auto-reactive immune cells affecting multiple organs [1]. Paediatric-onset SLE (pSLE) is uncommon, representing only approximately 10% of the SLE population [2–4]. The prevalence of pSLE is approximately 33 to 88 per million children, with a higher prevalence in Africans and Asians [5]. The median age of diagnosis for patients with pSLE is consistently reported to be 11 to 12 y of age [5–7]. Compared with adult-onset patients, those with pSLE have a higher involvement of cardiopulmonary, musculoskeletal, ocular, renal and neuropsychiatric systems [5,8]. Patients with pSLE also exhibit higher disease activity and more frequent use of immunosuppressants [9]. Diagnosing SLE in children can be more difficult because of the broad range of disease presentations in children and SLE is rarer in children [10]. Previous genome-wide DNAm studies using the Illumina HumanMethylation450 BeadChip have been performed on patients with classical, adult-onset SLE [11–17], SLE patients with renal involvement or lupus nephritis [18–20], patients with autoantibodies [21], and patients with different SLE disease activity index (SLEDAI) [22]. However, no genome-wide DNAm profiling has been performed in patients with pSLE.

Compared with other tissues, blood is much easier to obtain and is therefore commonly used for epigenetic studies. However, blood is a heterogeneous collection of different types of immune cell populations that show considerable DNAm variation across
different lineages [23–27]. As a result, the DNAm level of a particular CpG site measured in whole blood represents the overall DNAm level across many different immune lineages. Moreover, the proportions of different cell types can vary across individuals, especially in certain disease states, particularly autoimmune diseases. As a result, cell heterogeneity can be a potentially important confounding factor for whole-blood DNAm measurements. Multiple algorithms have been developed to adjust for the contribution of DNAm variation when analysing whole blood. However, no particular algorithm has demonstrated outstanding performance [28,29]. In addition, cell type-specific DNAm cannot be identified by statistical adjustment. Cell sorting remains the only reliable method to study cell lineage-specific DNAm.

Since DNAm changes in pSLE remain unknown and because cell type-specific DNAm changes cannot be reliably identified in whole blood, we investigated the DNAm changes in pSLE in a lineage-specific manner (CD4 + T cells, CD8 + T cells, CD19 + B cells and CD16+ neutrophils). The latest version of Illumina’s DNAm microarray, the MethylationEPIC BeadChip (EPIC microarray), was used in this study and for the first time in any SLE DNAm studies. Our results show that a set of 21 CpG sites were commonly hypomethylated in the four purified lineages, representing a pSLE-specific DNAm signature. In addition, we identified some genes with lineage-specific DNAm changes, supporting the importance of studying DNAm changes in purified lineages to optimally define DNAm changes in specific disease states.

Results

Patient selection for the DNAm microarray

Sixteen patients with pSLE (14 female and 2 male) and 13 controls (all female) were selected for genome-wide DNAm analysis (Supplementary Tables 1 and 2). Four purified lineages (purity>90%) and whole blood were assessed for all participants (n = 29, a total of 145 samples) using the EPIC microarray. The purities of the four lineages ranged from 95.8±2.1% for CD4 + T cells to 99.2±0.8% for neutrophils (Supplementary Table 3). At the time of recruitment, the median age of patients with pSLE and healthy controls was 19 (range: 12–57) years and 24 (range: 17–61) years, respectively. The median age of SLE onset was 11.5 (range: 6–17) years, which was comparable to that reported in the literature.

DNAm is driven by cell composition

After filtering for non-specific probes and probes with missing values or with detection p-values>0.05, a total of 91,556 probes were removed, and the final data set contained 775,280 probes. Using all remaining probes, the PCA showed that samples were clustered according to cell type, as expected (Figure 1(a)). This result confirmed previous findings that the DNA methylome demonstrates considerable variation across different types of immune cell lineages [23–27]. Specifically, DNAm profiles of CD4 + T cells and CD8 + T cells were more similar to each other, as these are both T cell lineages, and the DNAm profiles of whole blood and CD16+ neutrophils were more similar to each other, presumably because neutrophils are the most abundant type of leukocytes in blood. The result also revealed that the effect of cell composition was greater than that of the SLE manifestation on the raw DNAm data.

Identification of differentially methylated CpG sites (pSLE vs. controls) in each cell lineage and whole blood

Using the Wilcoxon rank-sum test, differentially methylated CpG sites were identified in each cell type by comparing the lineage-specific DNAm data of pSLE patients and controls. The number of probes that remained after the FDR control at q-values<0.05 ranged from lowest in B cells (n = 314) to highest in whole blood (n = 66588). An additional effect-size filter was applied to select a statistically significant set of probes with a > 10 percentage point difference in average DNAm levels between pSLE patients and controls. The number of probes that remained after the FDR control at q-values<0.05 ranged from lowest in B cells (n = 314) to highest in whole blood (n = 66588). An additional effect-size filter was applied to select a statistically significant set of probes with a > 10 percentage point difference in average DNAm levels between pSLE patients and controls. After effect-size filtering, the number of probes remaining ranged from lowest in B cells (n = 46) to highest in neutrophils (n = 160), and more CpG sites showed hypomethylation than hypermethylation (Table 1, Supplementary Table 4). Analysis of the genomic locations of the differentially methylated CpG sites showed that the CpG sites were over-represented in the promoter region, i.e., 5’UTR, in all cell lineages (Supplementary Table 5).
We next examined the effect of the differentially methylated CpG sites identified in each lineage using PCA and unsupervised hierarchical clustering. In general, PCA revealed that pSLE patients clustered as a distinct group separate from controls, and two subgroups existed within the patients (Supplementary Figure 1). The two pSLE subgroups can also be demonstrated by unsupervised hierarchical clustering. These two subgroups of SLE patients were distinguished by hypomethylated CpG sites (Supplementary Figure 2). A number of parameters have been tested for the two SLE subgroups, including age of SLE onset, age at the time of patient recruitment, disease duration, and other clinical information at the time of DNA sampling, such as SLEDAI, low complement level and presence of anti-double stranded DNA. We found that the SLE subgroup with milder DNAm loss exhibited lower SLEDAI

Table 1. Summary of differentially methylated CpG sites in whole blood and different cell types.

<table>
<thead>
<tr>
<th></th>
<th>Whole blood</th>
<th>CD4 + T cells</th>
<th>CD8 + T cells</th>
<th>B cells</th>
<th>Neutrophils</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of probes with q-value &lt; 0.05</td>
<td>66588</td>
<td>25360</td>
<td>6050</td>
<td>314</td>
<td>20951</td>
</tr>
<tr>
<td>No. of differentially methylated CpG sites</td>
<td>152</td>
<td>144</td>
<td>96</td>
<td>46</td>
<td>160</td>
</tr>
<tr>
<td>No. of probes (genes) with LOM</td>
<td>127 (75)</td>
<td>86 (52)</td>
<td>72 (41)</td>
<td>39 (24)</td>
<td>133 (61)</td>
</tr>
<tr>
<td>No. of probes (genes) with GOM</td>
<td>25 (25)</td>
<td>58 (36)</td>
<td>24 (13)</td>
<td>7 (7)</td>
<td>27 (15)</td>
</tr>
</tbody>
</table>

Gene Ontology of genes with LOM
- Response to type I interferon

Gene Ontology of genes with GOM
- T cell receptor signalling
- Response to type I interferon
- Nil

For each cell type, data was obtained by comparing the differences in pSLE and controls. Please refer to supplementary Table 6 for details on the gene ontology analysis.

LOM: loss of DNAm; GOM: gain of DNAm
in both CD4 + T cells and CD8 + T cells (unpaired t-test, p = 0.044 in both T cell lineages, Supplementary Figure 1), but this association was not apparent in other specific lineages or in whole blood. Because of the small sample size, further studies should be performed in a larger cohort of patients to confirm the correlation.

Finally, gene ontology analysis was performed using GeneMANIA. Genes that overlapped the differentially hypomethylated CpG sites were enriched in the interferon signalling pathway (Table 1), consistent with previous studies of SLE [11–16,20]. However, unlike previous studies using the 450k microarray that were unable to identify significant pathways in hypermethylated genes, we identified enrichment of the T cell receptor signalling pathway in CD4 + T cells, as well as the type I interferon signalling pathway in whole blood, CD8 + T cells and neutrophils (Table 1, Supplementary Table 6). The novel identification of pathway enrichments in hypermethylated genes may be specific to pSLE patients or to the utilization of the latest EPIC microarray.

**A pSLE-specific DNAm signature**

We then investigated the common differential DNAm changes in the four cell lineages, by examining the degree of overlap between the differentially methylated CpG sites in each cell lineage (Supplementary Table 5). We identified 21 CpG sites that showed differential loss of DNAm across the four purified lineages, and these differential changes can also be found in whole blood (Figure 2, Table 2). The 21 differentially hypomethylated CpG sites overlapped 15 genes, and clearly distinguished pSLE patients from healthy individuals. Therefore, we refer these 21 CpG sites to as a pSLE-specific DNAm signature. PCA and unsupervised hierarchical clustering of the pSLE-specific DNAm signature on all 145 samples revealed that samples can be separated based on case-control and cell lineages (Figures 1(b) and 3). To determine the significance of the cell composition effect on the pSLE-specific DNAm signature, the Friedman test was used for multiple group comparisons using pSLE patients DNAm data, and p-value<0.002 was considered as statistically significant (Bonferroni-adjusted for 21 tests and α = 0.05). Twenty of the 21 CpG sites showed significant differences across lineages and whole blood (the mean rank of each cell type and test statistics can be found in Supplementary Table 7). The Friedman test indicated that the pSLE-specific DNAm signature is also partially affected by cell composition.

**Lineage-specific differentially methylated genes**

We also observed a subset of probes that were uniquely present in pSLE in a cell lineage-specific manner (Figures 2 and 4). The genes that found to
be differentially methylated in just one cell lineage and are referred to as lineage-specific differentially methylated genes. The number of lineage-specific differentially methylated genes ranged from lowest in B cells (n = 1) to highest in CD4 + T cells (n = 38). Some of these lineage-specific genes have previously been reported in association with SLE (Supplementary Table 8). For example, the hypermethylated gene \( LCK \), identified in CD4 + T cells, demonstrated reduced expression in T cells of SLE patients, and the level of reduction was associated with disease activity \([30, 31]\). The hypomethylated genes \( OAS1 \) and \( OAS2 \), identified in CD8 + T cells and neutrophils, respectively, were upregulated in the peripheral blood of SLE patients \([32]\). On the other hand, some of the cell lineage-specific genes have never been reported in SLE. For example, the hypermethylated gene \( TART1 \) has never been reported in SLE, yet interestingly, it is involved in T cell receptor signalling \([33]\). These cell lineage-specific changes could not be reflected in whole blood, 

![Figure 3](image-url). Heatmap of the pSLE-specific DNAm signature.

Unsupervised hierarchical clustering using the pSLE-specific DNAm signature showed that samples were first separated according to disease manifestation, followed by cell lineage. Hierarchical clustering was performed using Euclidian distance metrics.
indicating the importance of performing DNAm studies on purified lineages.

Discussion

In this study, DNAm changes in patients with pSLE were compared to those in healthy controls using the EPIC microarray. In contrast to previous genome-wide DNAm studies on SLE, the experimental design of the current study differed in three important ways. First, we investigated DNAm changes in four major purified immune cell lineages and whole blood in the same cohort of patients and controls. This approach allows direct comparison of differentially methylated genes across different lineages and the identification of lineage-specific differentially methylated genes. Second, DNAm profiling was performed in patients with pSLE, which has never been performed in a genome-wide manner. Patients with pSLE are rare and, compared with adult-onset SLE, have a more severe clinical presentation and different patterns of organ involvement [5,8]. Therefore, the results of this study provide insights into the pathogenesis of pSLE. Third, we used the latest Illumina microarray to measure DNAm levels, allowing us to include an exploration of other regulatory regions such as enhancers and open chromatin. Consistent with previous studies using the 450k microarray, most of the differentially methylated CpG sites we identified were located either in the promoter or the body of the genes.

The results of the current study have four important implications. First, we identified a pSLE-specific DNAm signature comprising 21 CpG sites that show differential loss of DNAm across the major immune cell lineages of pSLE patients. Both myeloid lineages (i.e., neutrophils) and lymphoid lineages (i.e., T cells and B cells) showed the same pSLE-specific DNAm signature, suggesting that epigenetic reprogramming of type I interferon-related genes is a widespread phenomenon in blood cells of pSLE patients, which can occur in haematopoietic stem cells. While the time point at which the reprogramming occurs is unknown, hypomethylation of type I interferon-
related genes may predispose the gene to increased responsiveness to stimulation (i.e., increase their readiness for expression). Another possibility is that such hypomethylation increases the baseline expression of certain type I interferon-related genes, since meta-analysis revealed that some of these genes (MX1, IFIT1, PLSCR1 and IFI44) had higher expression in SLE patients than in healthy individuals [32].

Second, while the pSLE-specific DNAm signature showed significant differences between pSLE patients and healthy controls, it also showed milder yet significant differences across immune cell lineages. This implies that the disease-specific DNAm signature can still be confounded by cell composition. Since blood is commonly used as a source of DNA for epigenome-wide association studies, our results imply that the differentially methylated CpG sites identified should be examined carefully using our approach to simultaneously evaluate whole blood and different lineages in order to ensure that cell composition is not the primary source of variation for the differential changes. This strategy is particularly important in studying autoimmune diseases or when the differentially methylated genes identified are immune-related.

Third, we identified a subset of lineage-specific differentially methylated genes, indicating that some of the dysregulation is specific to a particular cell lineage and hence affects normal function for that specific cell type. For example, hypermethylated genes in helper T cells were found to be enriched in pathways such as T cell receptor signalling, suggesting the dysregulation of this pathway in pSLE patients, which may affect helper T cell function in adaptive immunity. In this study, we also identified some cell lineage-specific DNAm changes that potentially pinpoint the cell type that contributes to gene expressions changes in pSLE. Meta-analysis showed that interferon-regulated genes such as MX1, IFI44L, IFIT1, PLSCR1, LY6E, LGALS9, OAS1 and OAS2 were upregulated in the peripheral blood of SLE patients [32]. The DNAm data showed that MX1, IFI44L, IFIT1 and PLSCR1 were hypomethylated in the four purified lineages, suggesting that each lineage contributed to the increased expression of these genes. However, some genes showed hypomethylation only in specific lineages. For example, LY6E was hypomethylated in helper T cells, cytotoxic T cells and neutrophils but not in B cells; LGALS9 was T cell-specific, with hypomethylation only detected in helper T cells and cytotoxic T cells. Although OASI and OAS2 belong to the same family of the enzyme synthase, OASI was hypomethylated only in cytotoxic T cells, whereas OAS2 hypomethylation was only detected in neutrophils. DNAm studies should therefore be performed in specific lineages rather than in whole blood, especially for diseases that affect the immune system.

Lastly, significant overlap was observed between hypomethylated genes in patients with pSLE and other SLE subgroups, such as adult-onset SLE or patients with lupus nephritis. A number of studies have been performed to investigate the DNAm changes in SLE [11–22]. However, the experimental designs are different in terms of the ethnic group, gender, cell types and comparison groups (Supplementary Table 9). It is therefore difficult to examine the DNAm differences across SLE subgroups or cell lineages. However, numerous studies, including one of ours, all identified type I interferon-related genes (e.g., MX1, IFI44L, PARP9 and DTX3L) as the most significant hypomethylated genes that distinguish SLE patients and controls [11–14,16,17,20]. These studies suggest that the hypomethylation of interferon-related genes is highly robust in SLE and is common in SLE regardless of age of disease onset, ethnicity or gender. As a result, DNAm may potentially be used as one of the diagnostic biomarkers for SLE. In fact, Zhao et al. showed that IFI44L promoter DNAm in whole blood has a high specificity and sensitivity, and can distinguish SLE from healthy controls and other autoimmune diseases [34]. Nevertheless, it is necessary to carry out additional studies in newly diagnosed SLE patients, so as to confirm the possibility of using DNAm as one of the diagnostic biomarkers for SLE.

This study has several limitations. First, the median age of pSLE patients is 19, which implies that more than half of the patients are in their age of adulthood at the time of recruitment. In addition, patients participated in this study have received immunosuppressive treatments of variable doses for a significant period of time. Most
of the drugs used in our patients are not known DNA methylation modifiers (except for methotrexate, which is used by one patient); however, the effect can yet to be discovered. Therefore, medications can also be a confounding factor in this study. Ideally, patients recruited should be children who are newly diagnosed with SLE to avoid the confounding effect of age and medications. Second, DNA methylation (DNAm) analysis was limited to whole blood and four purified lineages only, which included CD4+ T cells, CD8+ Tcells, B cells, and neutrophils. Although we included the major immune cell lineages in our DNAm analysis, these cell types can be further divided into sub-classes, while some cell types were completely missing. For example, monocytes and natural killer cells were not included in this study, and T cells can be further divided into memory T cells, regulatory T cells, natural killer T cells and gamma delta T cells. In addition, although all nucleated cells can secrete type I interferon, plasmacytoid dendritic cells are professional type I interferon-secreting cells. Therefore, characterization of the genome-wide DNAm of the plasmacytoid dendritic cells may also be relevant [35]. With advances in single-cell sequencing-based technologies, single-cell genome-wide DNAm analysis may soon be possible, thereby providing a better understanding of the molecular pathophysiology of SLE in specific cell lineages.

In conclusion, we identified a pediatric systemic lupus erythematosus (pSLE)-specific DNAm signature that was commonly found in the major immune cell lineages of patients with pSLE. This pSLE-specific DNAm signature has potential as a diagnostic biomarker for pSLE. This finding is important for pediatric patients because diagnosing SLE in children can be challenging. We also highlighted the importance of studying DNAm changes in a lineage-specific manner, because the disease-specific DNAm signature can be cofounded by cell composition. In addition, studying lineage-specific DNAm changes can facilitate the elucidation of cell type-specific molecular pathophysiology of the disease. Finally, we have deposited the DNAm dataset that consists of whole blood and the four purified lineages in the GEO repository, which can be used for cell-type adjustments for other epigenome-wide association studies involving the use of the EPIC microarray.

Patients and methods

Patient recruitment

In this study, patients with pSLE were recruited regardless of their current age. All patients were younger than 18 y of age at the time of SLE diagnosis. They were all diagnosed using the 1997 American College of Rheumatology Criteria for Classification of SLE. Written consent was obtained from each patient or from their parents if the patient was less than 18 y of age. A total of 30 ml of blood was obtained from each patient. For healthy controls, blood was obtained from the Hong Kong Red Cross. All participants were self-reported to be of Chinese ethnicity. This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (reference number: UW 14–327).

Positive selection of immune cell populations

Four immune cell lineages (CD4+ T cells, CD8+ T cells, CD19+ B cells and CD16+ neutrophils) were sequentially isolated from each participant. Sample processing was initiated within 30 min of the blood draw. Peripheral blood mononuclear cells (PBMCs) were separated from red blood cells and polynuclear cells using Ficoll-Paque PLUS (GE Healthcare, Chicago, USA). Isolated PBMCs were sequentially incubated with anti-CD19, anti-CD8, and anti-CD4 microbeads (Miltenyi Biotec, Bergisch, Germany). Red blood cells and polynuclear cells were separated using 3% dextran (Fisher Scientific, Hampton, USA), as previously reported [36], and neutrophils were subsequently isolated using anti-CD16 microbeads (Miltenyi Biotec, Bergisch, Germany). The purity of each cell lineage was evaluated by flow cytometry. Specific cell lineages were stained with anti-CD4-PE, anti-CD8-FITC, anti-CD20-FITC and anti-CD16-FITC primary antibodies (BioLegend, San Diego, USA), respectively. Immune cells were considered to be a purified lineage when the purity was >90%. DNA extraction and bisulfite conversion were performed as described previously [14].

DNAm microarray

Sixteen pSLE patients and 13 controls were selected (Supplementary Tables 1 and 2), and
genome-wide DNAm analysis was performed on whole blood and four purified immune cell populations for each individual, generating a DNAm dataset of 145 samples. In five pSLE patients, the purity of one of the cell lineages was not measured because of an inadequate number of cells (B cells, n = 4; CD8 + T cells, n = 1, Supplementary Table 3). DNAm microarray using the EPIC microarray (Illumina, San Diego, USA) was performed at the Centre for Applied Genomics at The Hospital for Sick Children (Toronto, Canada). To minimize batch effects, all 145 samples were scanned in the same batch. In addition, the distributions of samples on the arrays were randomized for both case and controls, as well as for all purified lineages. In other words, all 145 samples were randomized such that different cell lineages from different individuals were mixed together on the same array [37]. Sample randomization served to minimize intra-chip differences. Raw intensity files were normalized using Illumina’s internal normalization probes and algorithms, and background noise was removed by subtracting the background intensities from negative probes [14]. The relative DNAm level, represented as beta values (β), was derived from the intensities.

Statistical analysis was performed according to the methodology described in our previous study on DNAm changes in Chinese patients with adult-onset SLE [14]. Briefly, raw data were exported into the R statistical environment and tested by a minfi package [38]. A number of problematic probes were removed from further analysis, including cross-reactive probes, probes with SNPs at target sites [39,40], probes from sex chromosomes, and probes with a missing β value or with median detection p-values > 0.01 in 25% of all samples. Statistical analysis was performed independently on whole blood and for each cell type by comparing the lineage-specific DNAm data of pSLE patients and controls. For each probe, the Wilcoxon rank-sum test was used to compare differences in pSLE and controls, and a false discovery rate control was used for multiple testing corrections. Probes were considered statistically significant when their FDR-corrected p-value (or q-value)<0.05. To ensure robust results, probes with statistical significance were also filtered for effect size, that is, only probes with β differences>0.10 between SLE patients and controls (i.e., 10 percentage point difference) were selected for further analysis. As a result, CpG sites with an adjusted p-value<0.05 and a mean β change>0.10 were considered to be differentially methylated, as described previously [14].

Principal component analysis (PCA) and unsupervised hierarchical clustering were performed using Qlucore Omics Explorer 3.2 (Qlucore, Lund, Sweden). Gene ontology analysis was performed using GeneMANIA (https://genemania.org), as described previously [14,41], and pathways with q-values (Benjamini-Hochberg-corrected p-value)<0.05 were considered significant. For the pSLE-specific DNA methylation signature, the Friedman test was performed in SPSS Statistics version 24 (IBM, Armonk, USA) to compare differences across multiple cell lineages in pSLE patients.

Acknowledgments
The authors gratefully acknowledge the participation of our pSLE patients and their families.

Availability of data and materials
The DNAm data generated in this study is available at the Gene Expression Omnibus (GEO) under the accession code, GSE118144, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE118144

Disclosure statement
No potential conflict of interest was reported by the authors.

Ethics approval and consent to participate
This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW 14-327). All patients, or their parents if the patient was less than 18 years of age, provided written informed consent.

Funding
This work was supported by the Research Grant Council’s General Research Fund (Ref: HKU 765513M).

ORCID
Mo Yin Mok http://orcid.org/0000-0003-4898-2114
References


