Plasma DNA aberrations in systemic lupus erythematosus revealed by genomic and methylomic sequencing

Rebecca W. Y. Chan[a,b], Peiyong Jiang[a,b], Xianlu Peng[a,b], Lai-Shan Tam[c], Gary J. W. Liao[a,b], Edmund K. M. Li[c], Priscilla C. H. Wong[c], Hao Sun[a,b], K. C. Allen Chan[a,b,d], Rossa W. K. Chiu[a,b], and Y. M. Dennis Lo[a,b,d,1]

[a]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China; [b]Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China; [c]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China; and [d]State Key Laboratory in Oncology in South China, Sir Y. K. Pao Centre for Cancer, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China

Contributed by Y. M. Dennis Lo, November 4, 2014 (sent for review September 18, 2014; reviewed by Chak Sing Lau and Ekkehard Schütz)

We performed a high-resolution analysis of the biological characteristics of plasma DNA in systemic lupus erythematosus (SLE) patients using massively parallel genomic and methylomic sequencing. A number of plasma DNA abnormalities were found. First, aberrations in measured genomic representations (MGRs) were identified in the plasma DNA of SLE patients. The extent of the aberrations in MGRs correlated with anti-double–stranded DNA (anti-dsDNA) antibody level. Second, the plasma DNA of active SLE patients exhibited skewed molecular size-distribution profiles with a significantly increased proportion of short DNA fragments. The extent of plasma DNA shortening in SLE patients correlated with the SLE disease activity index (SLEDAI) and anti-dsDNA antibody level. Third, the plasma DNA of active SLE patients showed decreased methylation densities. The extent of hypomethylation correlated with SLEDAI and anti-dsDNA antibody level. To explore the impact of anti-dsDNA antibody on plasma DNA in SLE, a column-based protein G capture approach was used to fractionate the IgG-bound and non–IgG-bound DNA in plasma. Compared with healthy individuals, SLE patients had higher concentrations of IgG-bound DNA in plasma. More IgG binding occurs at genomic locations showing increased MGRs. Furthermore, the IgG-bound plasma DNA was shorter in size and more hypomethylated than the non–IgG-bound plasma DNA. These observations have enhanced our understanding of the spectrum of plasma DNA aberrations in SLE and may provide new molecular markers for SLE. Our results also suggest that caution should be exercised when interpreting plasma DNA-based noninvasive prenatal testing and cancer testing conducted for SLE patients.

genomic representation | size profiling | epigenetics | massively parallel sequencing | autoimmune disease

Systemic lupus erythematosus (SLE) is a prototype autoimmune disease that has the potential of affecting multiple organ systems, including the skin, muscles, bones, lungs, kidneys, as well as the cardiovascular and central nervous systems (1, 2). SLE can cause various tissue inflammation and damages in a chronic manner. Renal complications, infections, myocardial infarctions, and central nervous system involvement are the major causes of death in SLE patients (3). The extremely variable clinical manifestations and the absence of effective tests to monitor disease activity present a challenge for clinical management (2, 3).

The etiology of SLE remains unknown and is multifactorial, involving genetic, epigenetic, environmental, hormonal, and immunologic factors (2, 4). Cell death has been regarded as an important event in the pathogenesis of SLE, as it leads to the release of antigens, such as nucleic acids, for immune complex formation, which may trigger a cascade of immune responses against the bodily tissues of the SLE patients (5, 6). Defects in the mechanism of cell death (7), impairment in the clearance of dead cells (8), and deficiency in DNase activity (9) have been implicated in SLE and suggested to be involved in the generation of autoantigens (5, 6).

In addition, epigenetic regulation is an important mechanism for maintaining the normal functioning of the immune system. Perturbation of the epigenetic regulation can disrupt the immunologic self-tolerance (10). Following the demonstration of impaired DNA methylation of T cells in SLE patients (11), an increasing amount of evidence has highlighted the contribution of epigenetic mechanisms in this disorder (12, 13). Hypomethylated apoptotic DNA from cells has been shown to be potentially pathogenic and may provoke the humoral and cellular immune responses in SLE (14).

SLE was one of the pathological conditions reported to be associated with the presence of circulating DNA nearly 50 years ago (15). Since then, studies using various detection methods have demonstrated the elevations of circulating DNA in SLE patients (16, 17). In addition, early reports have highlighted that

Significance

Through the use of massively parallel sequencing, we have demonstrated a spectrum of plasma DNA abnormalities in patients with systemic lupus erythematosus. These abnormalities include aberrant measured genomic representations, hypomethylation, and DNA fragment size shortening. The binding of anti-double–stranded DNA antibody to plasma DNA appears to be an important factor associated with these abnormalities. These findings provide valuable insights into the biology of plasma DNA in an autoimmune disease and have potential implications for the development of new molecular markers for systemic lupus erythematosus.

Author contributions: R.W.Y.C., L.-S.T., R.W.K.C., and Y.M.D.L. designed research; R.W.Y.C., P.J., and X.P. performed research; L.-S.T., E.K.M.L., and P.C.H.W. performed patient recruitment and clinical characterization; R.W.Y.C., P.J., X.P., G.J.W.L., H.S., K.C.A.C., R.W.K.C., and Y.M.D.L. analyzed data; and R.W.Y.C., G.J.W.L., R.W.K.C., and Y.M.D.L. wrote the paper.

Reviewers: C.S.L., University of Hong Kong; and E.S., Chronix Biomedical.

Conflict of interest statement: A patent application has been filed on the technology described in this paper. Y.M.D.L. and R.W.K.C. are consultants to, receive research support from, and hold equities in Sequenom. K.C.A.C. holds equities in Sequenom. Y.M.D.L., R.W.K.C., and K.C.A.C. are founders of Xcelom Limited.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in European Genome-Phenome Archive (EGA), www.ebi.ac.uk/ega, which is hosted by the European Bioinformatics Institute (EBI) (accession no. EGAS00001000962).

[1]To whom correspondence should be addressed. Email: loym@cuhk.edu.hk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1421126111/-/DCSupplemental.

the circulating DNA that form immune complexes with autoantibodies in SLE patients displays a characteristic fragmentation pattern that resembles the DNA laddering pattern of apoptosis by gel electrophoresis (18). These findings have suggested an interplay of apoptosis and circulating DNA in the pathogenesis of SLE. However, there have been very few studies reporting the detailed biological characterization of circulating DNA in SLE.

The advent of massively parallel sequencing has enabled the investigation of circulating DNA at single-base resolution on a genome-wide scale in fields such as noninvasive prenatal testing (19–21) and cancer detection (22–25). It would be of great interest to use this technology to explore the genomic and methylomic features of plasma DNA in SLE patients. In particular, the interplay of deregulated cell death, altered epigenetic regulation and production of autoimmune antibodies in SLE patients might cause abnormal patterns of circulating DNA. Hence, in this study we delineated the biological characteristics of DNA in the plasma of SLE patients using genome-wide genomic and methylomic sequencing.

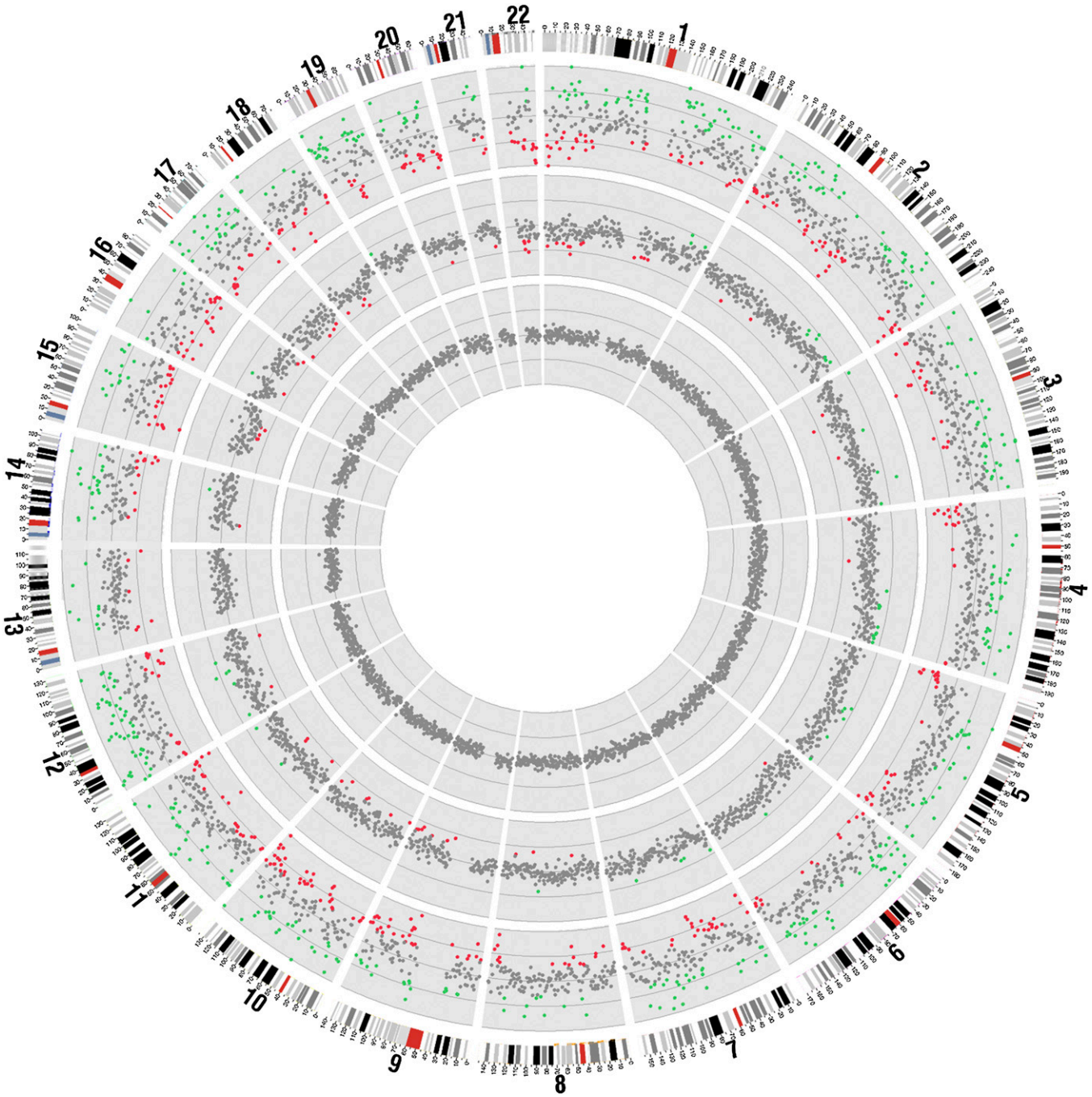

Fig. 1. Circos plots showing the genomic distributions of plasma DNA for a representative case in each of the control, inactive SLE, and active SLE groups. From inside to outside, the rings show data from a representative healthy individual (C005), inactive SLE patient (S011), active SLE patient (S112), and ideograms of the human chromosomes, respectively. Each dot represents a 1-Mb bin. The green, red, and gray dots represent bins with significant increased, decreased, and normal MGRs, respectively. The distance between intervals represents a *z*-score difference of 5.

Results

Genomic Representations of Plasma DNA in SLE Patients. First, we assessed if the plasma DNA molecules in the circulation of SLE patients were evenly distributed across the genome. Plasma DNA from 24 SLE patients and 11 healthy individuals were analyzed by paired-end massively parallel sequencing. SLE patients were divided into active and inactive groups according to their SLE disease activity index (SLEDAI), which is a clinical index for the measurement of disease activity (26). Fifteen patients with SLEDAI below or equal to 6 (median: 4, range: 0–6) were classified as the inactive group, and nine patients with SLEDAI over 6 (median: 8, range: 7–16) were classified as the active group. A median of 120 million (range: 18–208 million) alignable and nonduplicated paired-end reads were obtained for the plasma DNA per case for subsequent analyses (Dataset S1).

We assessed the genomic distribution of plasma DNA across segments (bins) of the human genome, each of 1 Mb in size. The number of sequence reads in each bin was tallied and adjusted by GC-correction as previously described (25). The control group consisted of 11 healthy individuals and this group showed even genomic distribution of the plasma DNA molecules as reported previously [Fig. 1, the innermost ring of the circos plot (27)]. To determine if the plasma DNA profiles of SLE patients showed differences in genomic representation, we compared the number of plasma DNA sequences aligned to a bin to the mean number detected among the control group for the same bin. We expressed the difference as a *z*-score, which was the number of SDs away from the mean of the control group. Bins with *z*-scores below −3 and above +3 were considered as showing significant under- and over-representation, respectively. We termed these changes aberrant measured genomic representations (MGRs).

The percentages of bins with aberrant MGRs among the healthy individuals, inactive and active SLE patients are shown in Fig. 2*A*. We tested for aberrant MGRs in the plasma of each healthy individual by comparing against the genomic distributions of the remaining control group. None of the healthy individuals exhibited any bin with aberrant MGRs in plasma (percentage of bins with aberrant MGRs = 0%). The percentages of bins with aberrant MGRs were higher among the SLE patients (active group: median 8.1%, range 1.1–52%; inactive group: median 6.5%, range 0.5–32.1%) compared with the controls ($P < 0.0001$, Kruskal–Wallis test) (Table 1). The MGR patterns of one representative case in each of the control, inactive SLE, and active SLE groups are shown in Fig. 1. Correlation analyses were performed between the percentage of bins with aberrant MGRs in SLE patients with the serum anti-double–stranded DNA (anti-dsDNA) antibody levels ($r = 0.604$, $P = 0.0018$, Spearman's correlation), and SLEDAI ($r = 0.226$, $P = 0.29$, Spearman's correlation).

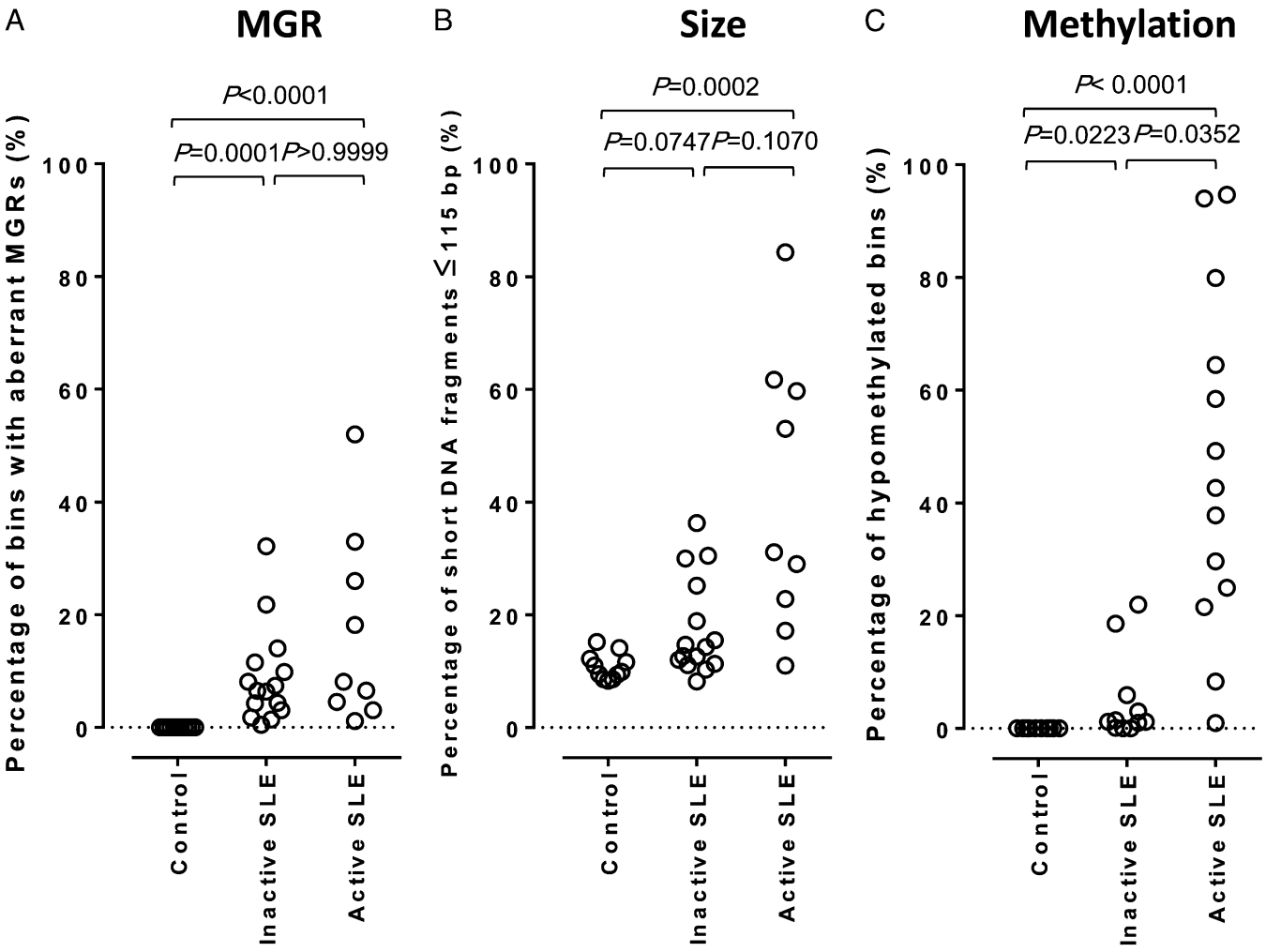

Fig. 2. Genomic and methylomic features of plasma DNA among subject groups. (*A*) Percentage of bins with aberrant MGRs. (*B*) Percentage of short DNA fragments. (*C*) Percentage of hypomethylated bins. Statistical comparisons were performed by the Kruskal–Wallis test followed by the post hoc Dunn's test.

Table 1. Percentage of bins with increased or decreased MGRs in the plasma of SLE patients

Group and case no.	SLEDAI	Anti-dsDNA antibody level (IU/mL)	Percentage of bins with increased MGR (%)	Percentage of bins with decreased MGR (%)	Percentage of bins with aberrant MGRs (increases + decreases) (%)
Inactive					
S073	0	0	1.1	0.7	1.8
S006	2	0	1.4	2.9	4.4
S017	2	107	4.3	3.0	7.4
S007	4	122	17.5	14.6	32.1
S019	4	139	6.2	5.4	11.6
S020	4	292	2.3	4.1	6.5
S002	4	312	2.5	3.9	6.4
S001	4	454	3.3	4.8	8.1
S012	4	500	0.3	0.2	0.5
S014	4	1,000	11.8	9.9	21.8
S011	5	227	1.8	2.4	4.2
S013	5	613	5.3	4.5	9.9
S008	6	0	0.8	0.6	1.4
S009	6	0	1.5	1.5	3.1
S081	6	1,000	6.0	8.0	14.0
Active					
S015	7	1,000	9.3	8.9	18.2
S016	8	218	2.0	2.5	4.5
S003	8	228	1.3	1.8	3.1
S018	8	230	0.4	0.7	1.1
S010	8	832	4.1	4.0	8.1
S004	8	1,000	3.6	2.9	6.5
S112	8	1,000	17.6	15.3	32.9
S082	10	1,000	13.9	12.1	26.0
S005	16	1,000	26.5	25.5	52.0

Size Analysis of Plasma DNA in SLE Patients. The sizes of plasma DNA molecules in the samples discussed in the previous section were deduced from the start and end coordinates of the paired-end reads (28). The size-distribution profiles of one healthy individual, one active, and one inactive SLE patient are shown in Fig. 3. The size distribution of plasma DNA in healthy individuals showed a major peak at 166 bp and a series of smaller peaks occurring at a 10-bp periodicity (Fig. 3). The size-distribution profiles of the SLE patients were different. The height of the 166-bp peak was reduced, whereas the other peaks for DNA fragments in smaller sizes, particularly those shorter than 115 bp, were elevated. These changes were more pronounced in the active SLE group than the inactive SLE group.

To systematically compare the size profiles between all samples, we defined a plasma DNA molecule ≤ 115 bp in size as a short DNA fragment and determined the percentage of short plasma DNA molecules in each sample. The data are shown in Fig. 2*B*. The median percentages of short plasma DNA fragments in the plasma of healthy individuals, inactive SLE and active SLE patients were 10% (range: 8–15%), 14% (range: 8–36%), and 31% (range: 11–84%), respectively. Positive correlations were observed between the percentage of short DNA fragments and SLEDAI ($r = 0.532$, $P = 0.0076$, Spearman's correlation) and the anti-dsDNA antibody level ($r = 0.758$, $P < 0.0001$, Spearman's correlation).

Methylation Status of DNA in Plasma of SLE Patients. Another sample set consisting of 24 SLE patients and 10 healthy individuals were subjected to methylation analysis. For the SLE patients, four inactive (S006, S013, S017, and S019) and four active cases (S004, S005, S010, and S015) had been studied in the above-mentioned MGR and size analyses, as sufficient volume of plasma could be harvested from these cases for both genomic and methylomic sequencing. For those with insufficient volume of plasma for both types of analysis, either genomic or methylomic analysis was performed on each sample. Plasma DNA was bisulfite-converted and analyzed by paired-end massively parallel sequencing, as previously described (29). A median of 125 million (range: 26–191 million) alignable and nonduplicated reads were obtained per case for subsequent analysis. Among the 24 SLE cases, 11 were in the inactive SLE group (SLEDAI median: 3, range: 0–5) and 13 were in the active SLE group (SLEDAI median: 8, range: 7–18).

The genome-wide methylation density of plasma DNA for each case refers to the proportion of CpG sites deemed to be

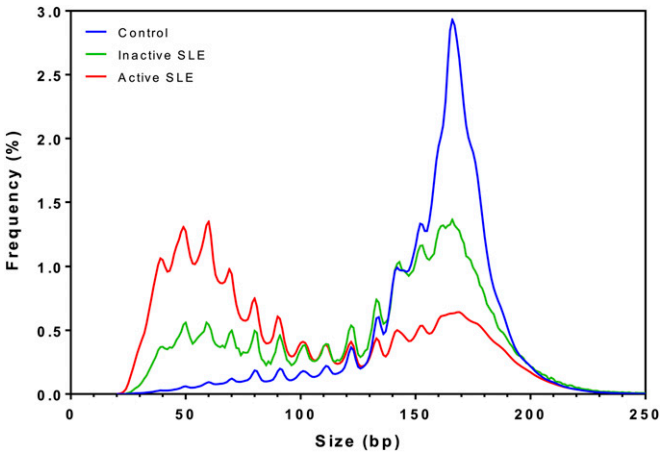

Fig. 3. Size distributions of plasma DNA molecules of representative cases for control (C005) (blue), inactive SLE (S081) (green), and active SLE (S082) (red) groups.

methylated among all of the CpG sites covered by sequence reads (29). The genome-wide methylation density of the active SLE group (70.1% ± 4.5%) was significantly reduced compared with both the healthy individuals (74.3% ± 1.4%, $P = 0.0367$, Kruskal–Wallis test, post hoc Dunn's test) and the inactive SLE group (74.4% ± 1.3%, $P = 0.0118$, Kruskal–Wallis test, post hoc Dunn's test) (Fig. S1).

Next we analyzed the methylation densities of each 1-Mb bin across the genome. For every bin, the plasma DNA methylation densities of the SLE patients were compared with the mean methylation density obtained from the 10 healthy individuals of the corresponding bin. Bins with methylation densities that were more than 3 SDs lower or higher than the mean of the control group, namely with *z*-scores below −3 or above +3, were deemed as significantly hypo- and hypermethylated, respectively (24). The percentages of bins with significant hypomethylation among the healthy individuals, inactive and active SLE patients are shown in Fig. 2*C*. Patients in the active group showed more hypomethylated bins (median: 42.7%, range: 1–94.7%) than the inactive group (median: 1.2%, range: 0–22%) (Table 2). The methylation patterns of one healthy individual, one active, and one inactive SLE patients are shown in Fig. 4. The percentage of hypomethylated bins correlated with SLEDAI ($r = 0.653$, $P = 0.0005$, Spearman's correlation) and anti-dsDNA antibody levels ($r = 0.555$, $P = 0.0059$, Spearman's correlation) of the SLE patients.

We reported in a previous study that shorter plasma DNA fragments tend to be more hypomethylated (29). Here we explored if a similar relationship is present in the plasma of SLE patients. First, the genome-wide methylation density was inversely correlated with the proportion of short DNA (≤115 bp) in all individuals ($r = -0.550$, $P = 0.0007$, Spearman's correlation) (Fig. S2). Next, we determined the methylation densities of DNA fragments of different sizes ranging from 20 to 250 bp, using sequence reads that covered at least 1 CpG site (29) (Fig. S3). For fragments between 40 and 180 bp, which accounted for the majority of plasma DNA molecules, the same trend as previously reported for the plasma of pregnant women was observed (29). It is noteworthy that the active SLE group showed greater reductions in methylation densities with progressive shortening of the plasma DNA fragments compared with the healthy individuals and patients in the inactive SLE group (Fig. S3).

Effects of IgG Binding on Plasma DNA of SLE Patients. Autoantibodies have a direct contribution to the pathogenesis of SLE (2, 3) and are responsible for many of the clinical manifestations (30). One of such autoantibodies is the anti-dsDNA antibody, which can bind to the DNA in plasma (31). Studies have reported that IgG-class anti-dsDNA antibody has high avidity for dsDNA and is implicated in the pathogenesis of SLE (31, 32). We hypothesized that the binding of anti-dsDNA antibody to plasma DNA might alter the stability or clearance of DNA in plasma and might result in observable aberrations in genomic representation, size, or methylation profiles of plasma DNA. To study the effect of anti-dsDNA antibody binding on plasma DNA, two sample sets were recruited: one for genomic representation and size analysis and the other for methylation level analysis. Each sample set included two healthy individuals, two inactive SLE patients, and two active SLE patients. For each case, the plasma sample was divided into two portions. One portion was not subjected to any treatment and was termed the neat fraction. The other portion was incubated with protein G and subjected to column capture. Protein G binds human IgG, including anti-dsDNA antibody. Therefore, column-based protein G capture further allowed the plasma sample to be separated into IgG-bound and non–IgG-bound fractions. The genomic representation, molecular size, and methylation profiles were compared among the neat, IgG-bound and non–IgG-bound DNA fractions.

Table 2. Percentage of bins with significant plasma DNA hypomethylation, normal methylation, and hypermethylation in SLE patients

Group and case no.	SLEDAI	Anti-dsDNA antibody level (IU/mL)	Percentage of bins with significant hypomethylation (%)	Percentage of bins with normal methylation (%)	Percentage of bins with significant hypermethylation (%)
Inactive					
S105	0	0	1.2	98.8	0.0
S125	0	0	0.0	99.9	0.1
S006	2	0	1.2	98.8	0.0
S017	2	107	1.0	99.0	0.0
S053	2	150	0.1	99.7	0.2
S019	4	139	1.4	98.6	0.0
S124	4	378	3.0	97.0	0.0
S026	4	1,000	0.0	100.0	0.0
S059	4	1,000	5.9	94.0	0.1
S013	5	613	18.6	81.4	0.0
S132	5	758	22.0	71.9	6.1
Active					
S015	7	1,000	94.7	5.3	0.0
S203	8	793	48.7	51.3	0.0
S010	8	832	29.7	70.3	0.0
S031	8	896	79.9	20.1	0.0
S004	8	1,000	58.4	41.6	0.0
S039	8	1,000	42.7	57.3	0.0
S131	8	1,000	23.3	76.3	0.3
S147	8	1,000	49.2	50.8	0.0
S033	10	1,000	21.5	78.5	0.0
S043	10	1,000	1.0	98.9	0.1
S027	12	1,000	64.5	35.5	0.0
S005	16	1,000	94.0	6.0	0.0
S086	18	947	8.3	91.4	0.3

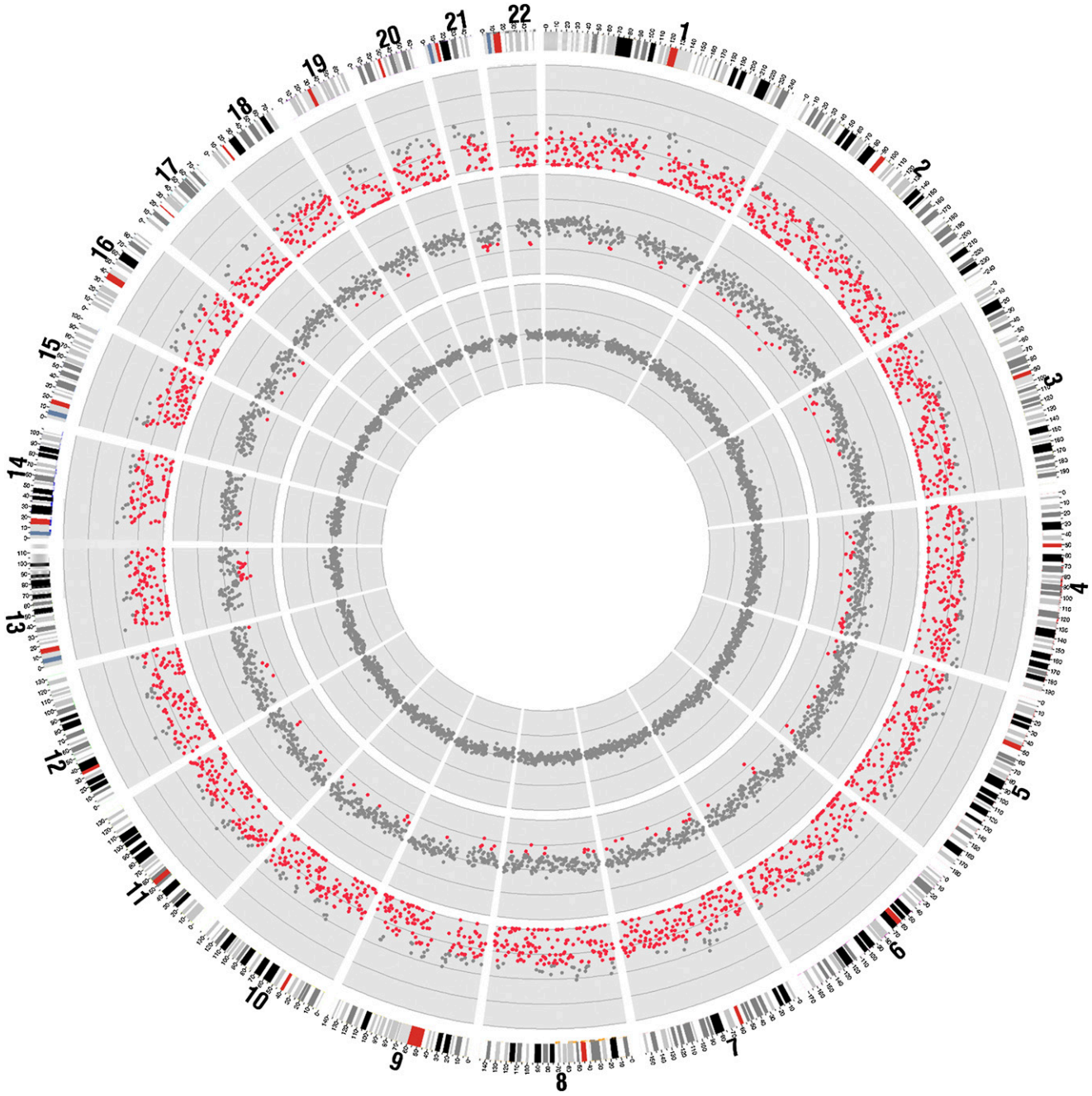

Fig. 4. Plasma methylation profile analysis for a representative case in each of the control, inactive SLE, and active SLE groups. From inside to outside, the rings showed data from a representative healthy individual (C040), inactive SLE patient (S124), active SLE patient (S027), and ideograms of the human chromosomes, respectively. Each dot shows the methylation density for a 1-Mb bin. The green, red, and gray dots represent bins with significant hypermethylation, hypomethylation, and normal methylation, respectively. The distance between intervals represents a *z*-score difference of 5.

Higher concentrations of IgG-bound DNA were found in the plasma of SLE patients (median: 494.5 copies/mL, range: 92–983 copies/mL), compared with the healthy individuals (median: 68.5 copies/mL, range: 41–139 copies/mL) (Table 3).

For genomic representation analysis, the *z*-score for each 1-Mb bin was calculated in the neat, IgG-bound, and non–IgG-bound fractions. Next, we calculated the *z*-score difference between the IgG-bound and non–IgG-bound fractions for each bin, expressed as the "IgG binding index." The IgG binding index of each bin was then compared with the *z*-score of MGR in the neat fraction for the corresponding bin (Fig. 5). We hypothesized that a higher proportion of plasma DNA fragments originating from regions showing increased MGRs were bound by anti-dsDNA antibody and would be found in the IgG-bound fraction. Hence, the IgG binding index for such locations should be higher. For regions exhibiting decreased MGRs, the reverse would be true and the IgG binding index should be more negative. Indeed, for SLE cases with high anti-dsDNA antibody levels (S081, S082, S112), the *z*-scores of bins with aberrant MGRs in the neat fractions showed a positive relationship with the corresponding IgG binding index ($r = 0.31$, $P < 0.0001$, Pearson's correlation) (Fig. 5).

For the size analysis, the SLE case (S073) with SLEDAI 0 and negative anti-dsDNA antibody showed similar plasma DNA size-distribution profiles in the neat, IgG-bound, and non–IgG-bound

Table 3. Absolute DNA concentration of neat, IgG-bound, and non–IgG-bound fractions

Group and case no.	SLEDAI	Anti-dsDNA antibody level (IU/mL)	Concentration of plasma DNA (copies/mL)		
			Neat	IgG-bound	Non–IgG-bound
Control					
C020	—	0	1,350	96	1,112
C021	—	0	1,714	139	1,549
C073	—	0	932	30	713
C074	—	0	860	41	771
Inactive					
S073	0	0	1,544	92	1,344
S125	0	0	2,379	901	1,600
S124	4	378	1,349	983	419
S081	6	1,000	1,376	634	703
Active					
S203	8	793	515	316	168
S112	8	1,000	368	355	39
S147	8	1,000	6,560	794	5,086
S082	10	1,000	1,042	277	459

fractions to the two healthy individuals (Fig. S4 *A*–*C*). In contrast, for three SLE patients (S081, S082, and S112) with high anti-dsDNA antibody levels and SLEDAI ≥ 6, the size-distribution profiles demonstrated a shortening of plasma DNA in each of the neat, IgG-bound, and non–IgG-bound fractions (Fig. 6 and Fig. S4 *D* and *E*). The IgG-bound fractions of these three SLE patients showed an enrichment of short DNA fragments ≤115 bp compared with the non–IgG-bound fractions (Fig. S5).

For the methylation analysis, compared with the two healthy individuals, SLE patients had lower genome-wide methylation densities in each of the neat, IgG-bound, and non–IgG-bound fractions (Table 4). For the two active SLE patients with high anti-dsDNA antibody levels (S147 and S203), the IgG-bound fractions showed the lowest genome-wide methylation densities and the highest number of hypomethylated bins. Correspondingly, the non–IgG-bound fractions had the highest genome-wide methylation densities and lowest number of hypomethylated bins (Table 4). Among these two SLE cases, the percentages of hypomethylated bins decreased in the following order: IgG-bound, neat, and non–IgG-bound fractions (Fig. 7). On the other hand, for the two SLE cases with low anti-dsDNA antibody levels (S124, 125), the genome-wide methylation densities and number of hypomethylated bins were similar between the IgG-bound and non–IgG-bound fractions (Table 4).

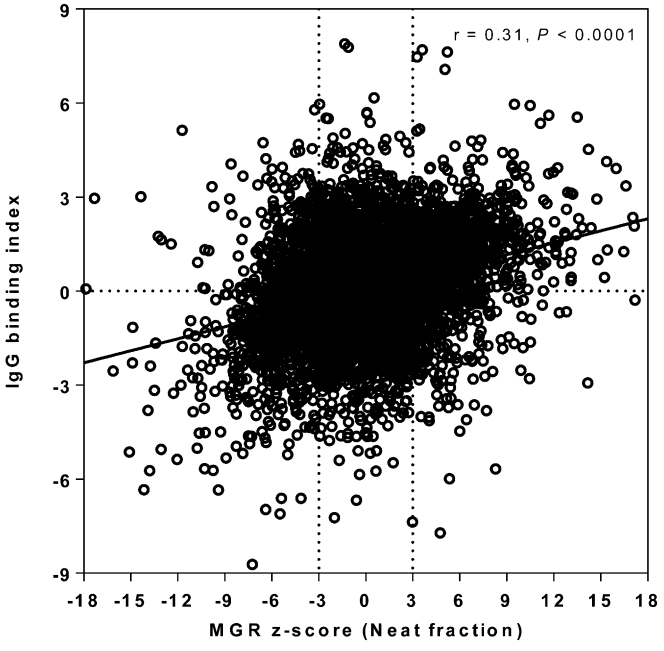

Fig. 5. Relationship between IgG binding index and MGR *z*-score of the SLE cases with high anti-dsDNA antibody levels (S081, S082, and S112). The correlation coefficient was calculated by Pearson's correlation analysis.

Discussion

In this study, we investigated the characteristics of plasma DNA of SLE patients in a high-resolution and genome-wide manner with the use of paired-end massively parallel sequencing. In general, the higher the disease activity, the greater the extent and the wider the range of plasma DNA aberrations were observed. Plasma DNA of SLE patients showed aberrant MGRs, size shortening, and hypomethylation. We further obtained evidence that suggested that those plasma DNA aberrations were at least in part related to DNA binding by IgG class antibodies, for example anti-dsDNA antibodies. These observations are particularly interesting because they suggest that a blood constituent, namely anti-dsDNA antibody, could alter the molecular characteristics and profile of plasma DNA.

We showed that the genomic representations of plasma DNA in SLE patients were different from those of healthy controls. The percentages of bins showing aberrant MGRs correlated with the anti-dsDNA antibody level but not the SLEDAI. We therefore hypothesized that antibody binding of plasma DNA may be related to the observed aberrant MGRs. In the IgG binding experiments, we indeed showed that more plasma DNA molecules originating from the regions with increased MGRs were associated with the IgG binding. Perhaps the binding of anti-dsDNA antibody to plasma DNA would protect the bound DNA from enzymatic degradation, or would impair the clearance mechanism, such as the phagocytic system (33–35). Studies have also reported the preferential binding of anti-DNA antibody to particular DNA sequences, for example, DNA fragments containing certain CpG motifs (36, 37). The retention of antibody-bound DNA might therefore enhance the genomic representation in regions with preference for antibody binding, whereas regions with less antibody binding preference would be underrepresented. These changes might then be detected as increased MGRs and decreased MGRs, respectively.

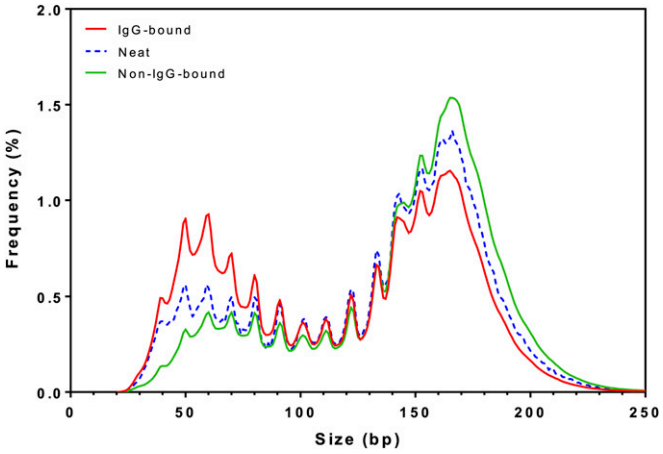

Fig. 6. Size distributions of DNA molecules with and without IgG binding in plasma of a representative SLE patient (S081). Blue dash, red solid, and green solid lines show data for the neat, IgG-bound, and non–IgG-bound fractions, respectively.

Currently, a number of noninvasive prenatal tests for fetal aneuploidy screening are based on the detection of copy number aberrations in maternal plasma DNA (19, 21, 38). Therefore, caution should be taken when applying such tests to pregnancies with underlying SLE, as the aberrant MGRs might reflect the activity of SLE, rather than the genomic aberrations originating from the fetus. Similarly, as plasma DNA sequencing for copy-number aberrations (23, 25) and hypomethylation (24) has been reported for cancer detection, one should also be wary of the possibility of false-positive results for patients with SLE.

In terms of size analysis, plasma of SLE patients showed an increased proportion of short DNA fragments. The proportion of short DNA fragments (≤115 bp) in the plasma of the active SLE patients was threefold higher than that of healthy individuals and could contribute up to 84% of the total DNA in plasma. Hence, when plasma DNA size analysis is used for noninvasive prenatal testing (39) in pregnant subjects with SLE, the possibility of false-positive results because of the underlying SLE should be borne in mind.

The shortening of plasma DNA in SLE patients may be a result of the increased production or decreased clearance of short DNA fragments. Our data show that the shortening of plasma DNA was positively correlated with the disease activity of the SLE patients and the anti-dsDNA antibody level. Furthermore, the IgG-binding data demonstrated that short DNA fragments (≤115 bp) were enriched in the IgG-bound fraction. Therefore, the data seem to suggest that there is preferential binding of anti-DNA antibody to short DNA fragments hindering the clearance of short DNA from the circulation.

Our previous data showed that plasma DNA of healthy individuals and pregnant women have a characteristic size profile with a prominent peak at 166 bp and a series of smaller peaks that are 10-bp apart (20). This characteristic pattern was reminiscent of the length of DNA in a nucleosomal unit. We have also previously shown that the shorter plasma DNA molecules tend to be more hypomethylated than the longer molecules (29). Hypomethylated DNA tended to be less densely packed with histones (40) and might be more susceptible to enzymatic degradation and provide more access to antibody binding. We therefore studied the methylation profile of DNA in the plasma of SLE patients. We found that plasma DNA of active SLE patients was generally hypomethylated compared with that of healthy individuals. The degree of hypomethylation correlated with the disease activity and the anti-dsDNA antibody level. The IgG binding experiment showed that the IgG-bound plasma DNA molecules were more hypomethylated (Fig. 7). This observation suggests that some hypomethylated DNA molecules are retained in plasma of SLE patients because of antibody-binding. On the other hand, our data show that antibody-binding results in an enrichment of short plasma DNA fragments. These observations were therefore consistent with our finding that there was a relationship between the genome-wide methylation density and the size of plasma DNA in SLE patients, in which the shorter DNA fragments were more hypomethylated. This finding was also consistent with our previous study on circulating fetal DNA (29).

Early studies have reported the phenomenon of DNA release from lymphocytes into serum (41, 42). Interestingly, a number of recent studies have reported the hypomethylation of T cells in SLE patients (43, 44). Hence, the release of DNA from hypomethylated cells, such as T cells, might be another mechanism that could contribute toward the hypomethylation of plasma DNA of SLE patients.

In summary, plasma DNA of SLE patients could exhibit aberrant MGRs, size shortening, and hypomethylation. These features might potentially be useful as biomarkers for SLE. Further studies with larger sample size, longitudinal specimen collections, and subgroup analysis of cases with specific clinical manifestations (e.g., renal involvement) may provide a more in-depth understanding of plasma DNA in SLE patients and its potential value for clinical practice. Our study also highlights the possibility that the study of plasma nucleic acids would be a valuable venue for research for other autoimmune diseases.

Table 4. Methylation levels of plasma DNA in neat, IgG-bound, and non–IgG-bound fractions

Group and case no.	SLEDAI	Anti-dsDNA antibody level (IU/mL)	Neat		IgG-bound		Non–IgG-bound	
			Genome-wide MD (%)	No. of bins with significant hypomethylation	Genome-wide MD (%)	No. of bins with significant hypomethylation	Genome-wide MD (%)	No. of bins with significant hypomethylation
Control								
C073	—	0	75.8	0	76.2	11	75.6	3
C074	—	0	77.1	0	78.0	0	77.2	0
Inactive								
S125	0	0	74.9	1	74.1	1	73.2	3
S124	4	378	74.7	82	73.5	142	74.3	161
Active								
S203	8	793	70.5	1,033	68.6	1,646	71.0	766
S147	8	1,000	70.7	1,345	70.1	1,624	72.1	896

MD, methylation density.

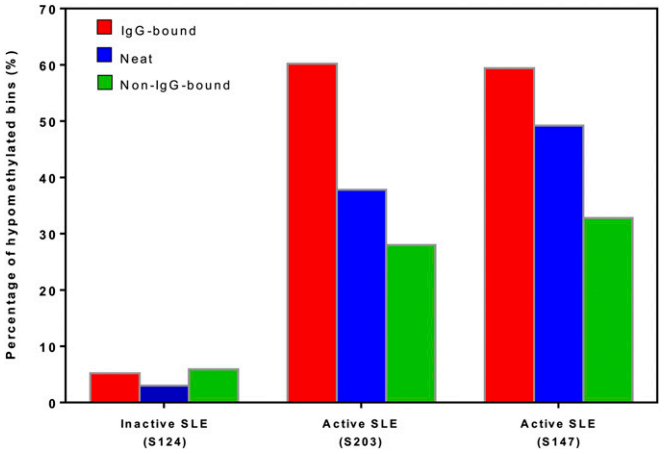

Fig. 7. Percentage of significant DNA hypomethylation in plasma of SLE patients with high anti-dsDNA antibody levels (S124, S203, and S147). Red, blue, and green bars indicate the IgG-bound, neat, and non–IgG-bound fractions, respectively.

Materials and Methods

Case Recruitment and Sample Processing. SLE patients attending the rheumatology clinic at the Department of Medicine and Therapeutics, Prince of Wales Hospital, Hong Kong, were recruited with written informed consent. The study was approved by the Joint Chinese University of Hong Kong–Hospital Authority New Territories East Cluster Clinical Research Ethics Committee. All patients fulfilled the American College of Rheumatology diagnostic criteria and their lupus disease activities were assessed by the SLEDAI (26). Clinical information on the SLE patients is summarized in Dataset S2. Peripheral blood was collected in EDTA-containing tubes. The blood samples were first centrifuged at 1,600 × *g* for 10 min at 4 °C and the plasma portion was further subjected to centrifugation at 16,000 × *g* for 10 min at 4 °C to pellet the residual cells.

Separation of IgG-Bound and Unbound DNA in Plasma. The plasma samples were separated into the IgG-bound and non–IgG-bound fractions using the NAb Spin Column (Thermo Fisher Scientific), which contained an immobilized protein G resin for IgG protein capturing. The separation was performed according to the manufacturer's instructions, except for the following modifications for purifying the IgG-bound fraction. After the elution of the non–IgG-bound fraction, the NAb Spin Column was washed with PBS for six times. To ensure a complete removal of the non–IgG-bound fraction before elution of the IgG-bound fraction, the final PBS wash was confirmed to have undetectable DNA level by using a real-time PCR quantification of the leptin (*LEP*) gene (45). Then, the IgG-bound DNA remaining in the resin of the column was eluted with six washes of freshly prepared buffer containing 1% SDS and 0.1 M sodium bicarbonate.

DNA Extraction and Preparation of DNA Libraries. DNA was extracted from 4 to 10 mL of plasma using the DSP Blood Mini Kit (Qiagen) with modifications of the manufacturer's protocol as previously reported (38). The concentrations of plasma DNA were determined by the real-time PCR assay of the *LEP* gene (45). Short DNA fragments of less than 62 bp can be underestimated with the use of this assay, as the amplicon size was 62 bp. For DNA sequencing, plasma DNA libraries were prepared using the Paired-End Sequencing Sample Preparation Kit (Illumina), as previously described (28). For bisulfite sequencing, bisulfite-treated plasma DNA libraries were prepared using the Paired-End Sequencing Sample Preparation Kit (Illumina) and the EpiTect Plus DNA Bisulfite Kit (Qiagen) (29).

DNA Sequencing and Alignment. The bisulfite-treated or untreated DNA libraries were sequenced for 75 bp of each end in a paired-end format on HiSeq2000 instruments (Illumina). DNA clusters were generated with a Paired-End Cluster Generation Kit v3 on a cBot instrument (Illumina). Real-time image analysis and base calling were performed using the HiSeq Control Software v1.4 and Real Time Analysis Software v1.13 (Illumina), by which the automated matrix and phasing calculations were based on the spiked-in PhiX control v3 sequenced with the libraries. After base calling, adapter sequences and low-quality bases (i.e., quality score < 5) on the fragment ends were removed.

For the analysis of sequencing data, the sequenced reads were aligned to the nonrepeat-masked human reference genome (NCBI build 37/hg19) using the Short Oligonucleotide Alignment Program 2 (46). Alignment was performed with the following settings: (*i*) two mismatches were allowed; (*ii*) alignment to multiple regions was not allowed; (*iii*) insertions or deletions compared with the reference genome were not analyzed; and (*iv*) DNA fragment size ranged from 20 to 600 bp.

Molecular Size Determination of Plasma DNA. Paired-end sequencing, where sequencing was performed for both ends of each DNA molecule, was used to analyze each sample. By aligning the pair of end sequences of each DNA molecule to the reference human genome and noting the genome coordinates of the extreme ends of the sequenced reads, the sizes of the sequenced DNA molecules were determined. Plasma DNA molecules are naturally fragmented and hence the sequencing libraries for plasma DNA are typically prepared without any fragmentation steps (28). Hence, the lengths deduced by the sequencing represented the sizes of the original plasma DNA molecules.

Methylation Analysis of Plasma DNA. For the analysis of bisulfite-converted DNA sequencing data, an additional step for identification of methylated cytosines was performed. The trimmed reads were processed by a methylation data analysis pipeline called Methy-Pipe (47). To align the bisulfite-converted sequencing reads, we first performed in silico conversion of all cytosine residues to thymines, on the Watson and Crick strands separately, using the reference human genome (NCBI build 37/hg19). We then performed in silico conversion of each cytosine to thymine in all of the processed reads and kept the positional information of each converted residue. After that, Methy-Pipe was used to align the converted reads to the two preconverted reference human genomes, with a maximum of two mismatches allowed for each aligned read. The cytosines originally present on the sequenced reads were recovered based on the positional information kept during the in silico conversion. The recovered cytosines among the CpG dinucleotides were scored as methylated. Thymines among the CpG dinucleotides were scored as unmethylated. We determined the methylation density of the whole human genome or any particular regions in the genome by determining the total number of unconverted cytosines at CpG sites as a proportion of all CpG sites covered by sequence reads mapped to the genome or the particular regions in the genome.

Statistical Analysis. Analysis was performed by using in-house bioinformatics programs, which were written in Perl and R languages, and SigmaStat v3.5 software (Systat Software). A *P* value of less than 0.05 was considered as statistically significant and all probabilities were two-tailed.

ACKNOWLEDGMENTS. We thank Andy Guo, Patty Tse, Emily Leung, and Michelle Pui for technical assistance. This work was supported by the Hong Kong Research Grants Council Theme-Based Research Scheme (T12-404/11) and the S. K. Yee Foundation. Y.M.D.L. is supported by an endowed chair from the Li Ka Shing Foundation.

1. Gladman DD, Urowitz MB (2007) Prognosis, mortality and morbidity in systemic lupus erythematosus. *Dubois' Lupus Erythematosus*, eds Wallace DJ, Hahn BH (Lippincott Williams & Wilkins, Philadelphia), 7th Ed, pp 1333–1353.
2. Tsokos GC (2011) Systemic lupus erythematosus. *N Engl J Med* 365(22):2110–2121.
3. Cervera R, et al.; European Working Party on Systemic Lupus Erythematosus (2003) Morbidity and mortality in systemic lupus erythematosus during a 10-year period: A comparison of early and late manifestations in a cohort of 1,000 patients. *Medicine (Baltimore)* 82(5):299–308.
4. Rullo OJ, Tsao BP (2013) Recent insights into the genetic basis of systemic lupus erythematosus. *Ann Rheum Dis* 72(Suppl 2):ii56–ii61.
5. Pisetsky DS (2004) The immune response to cell death in SLE. *Autoimmun Rev* 3(7-8): 500–504.
6. Su KY, Pisetsky DS (2009) The role of extracellular DNA in autoimmunity in SLE. *Scand J Immunol* 70(3):175–183.
7. Emlen W, Niebur J, Kadera R (1994) Accelerated in vitro apoptosis of lymphocytes from patients with systemic lupus erythematosus. *J Immunol* 152(7):3685–3692.
8. Munoz LE, et al. (2005) SLE—A disease of clearance deficiency? *Rheumatology (Oxford)* 44(9):1101–1107.
9. Sallai K, Nagy E, Derfalvy B, Müzes G, Gergely P (2005) Antinucleosome antibodies and decreased deoxyribonuclease activity in sera of patients with systemic lupus erythematosus. *Clin Diagn Lab Immunol* 12(1):56–59.
10. Strickland FM, Richardson BC (2008) Epigenetics in human autoimmunity. Epigenetics in autoimmunity—DNA methylation in systemic lupus erythematosus and beyond. *Autoimmunity* 41(4):278–286.
11. Richardson B, et al. (1990) Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis. *Arthritis Rheum* 33(11): 1665–1673.

12. Ballestar E, Esteller M, Richardson BC (2006) The epigenetic face of systemic lupus erythematosus. *J Immunol* 176(12):7143–7147.
13. Yang ML, et al. (2013) Lupus autoimmunity altered by cellular methylation metabolism. *Autoimmunity* 46(1):21–31.
14. Wen ZK, et al. (2007) DNA hypomethylation is crucial for apoptotic DNA to induce systemic lupus erythematosus-like autoimmune disease in SLE-non-susceptible mice. *Rheumatology (Oxford)* 46(12):1796–1803.
15. Tan EM, Schur PH, Carr RI, Kunkel HG (1966) Deoxybonucleic acid (DNA) and antibodies to DNA in the serum of patients with systemic lupus erythematosus. *J Clin Invest* 45(11):1732–1740.
16. Raptis L, Menard HA (1980) Quantitation and characterization of plasma DNA in normals and patients with systemic lupus erythematosus. *J Clin Invest* 66(6):1391–1399.
17. Chen JA, et al. (2007) Sensitive detection of plasma/serum DNA in patients with systemic lupus erythematosus. *Autoimmunity* 40(4):307–310.
18. Rumore PM, Steinman CR (1990) Endogenous circulating DNA in systemic lupus erythematosus. Occurrence as multimeric complexes bound to histone. *J Clin Invest* 86(1):69–74.
19. Chiu RWK, et al. (2008) Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc Natl Acad Sci USA* 105(51):20458–20463.
20. Lo YMD, et al. (2010) Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci Transl Med* 2(61):61ra91.
21. Palomaki GE, et al. (2011) DNA sequencing of maternal plasma to detect Down syndrome: An international clinical validation study. *Genet Med* 13(11):913–920.
22. Beck J, Urnovitz HB, Mitchell WM, Schütz E (2010) Next generation sequencing of serum circulating nucleic acids from patients with invasive ductal breast cancer reveals differences to healthy and nonmalignant controls. *Mol Cancer Res* 8(3):335–342.
23. Leary RJ, et al. (2012) Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. *Sci Transl Med* 4(162):162ra154.
24. Chan KCA, et al. (2013) Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. *Proc Natl Acad Sci USA* 110(47):18761–18768.
25. Chan KCA, et al. (2013) Cancer genome scanning in plasma: Detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin Chem* 59(1):211–224.
26. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH; The Committee on Prognosis Studies in SLE (1992) Derivation of the SLEDAI. A disease activity index for lupus patients. *Arthritis Rheum* 35(6):630–640.
27. Krzywinski M, et al. (2009) Circos: An information aesthetic for comparative genomics. *Genome Res* 19(9):1639–1645.
28. Zheng YWL, et al. (2012) Nonhematopoietically derived DNA is shorter than hematopoietically derived DNA in plasma: A transplantation model. *Clin Chem* 58(3):549–558.
29. Lun FMF, et al. (2013) Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. *Clin Chem* 59(11):1583–1594.
30. Arbuckle MR, et al. (2003) Development of autoantibodies before the clinical onset of systemic lupus erythematosus. *N Engl J Med* 349(16):1526–1533.
31. Isenberg DA, Manson JJ, Ehrenstein MR, Rahman A (2007) Fifty years of anti-ds DNA antibodies: Are we approaching journey's end? *Rheumatology (Oxford)* 46(7):1052–1056.
32. Winfield JB, Faiferman I, Koffler D (1977) Avidity of anti-DNA antibodies in serum and IgG glomerular eluates from patients with systemic lupus erythematosus. Association of high avidity antinative DNA antibody with glomerulonephritis. *J Clin Invest* 59(1): 90–96.
33. Haakenstad AO, Mannik M (1974) Saturation of the reticuloendothelial system with soluble immune complexes. *J Immunol* 112(5):1939–1948.
34. Frank MM, Hamburger MI, Lawley TJ, Kimberly RP, Plotz PH (1979) Defective reticuloendothelial system Fc-receptor function in systemic lupus erythematosus. *N Engl J Med* 300(10):518–523.
35. Mir A, Porteu F, Levy M, Lesavre P, Halbwachs-Mecarelli L (1988) C3b receptor (CR1) on phagocytic cells from SLE patients: Analysis of the defect and familial study. *Clin Exp Immunol* 73(3):461–466.
36. Uccellini MB, et al. (2008) Autoreactive B cells discriminate CpG-rich and CpG-poor DNA and this response is modulated by IFN-alpha. *J Immunol* 181(9):5875–5884.
37. Uccellini MB, Busto P, Debatis M, Marshak-Rothstein A, Viglianti GA (2012) Selective binding of anti-DNA antibodies to native dsDNA fragments of differing sequence. *Immunol Lett* 143(1):85–91.
38. Chiu RWK, et al. (2011) Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: Large scale validity study. *BMJ* 342:c7401.
39. Yu SCY, et al. (2014) Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing. *Proc Natl Acad Sci USA* 111(23):8583–8588.
40. Kelly TK, et al. (2012) Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res* 22(12):2497–2506.
41. Anker P, Stroun M, Maurice PA (1976) Spontaneous extracellular synthesis of DNA released by human blood lymphocytes. *Cancer Res* 36(8):2832–2839.
42. Rogers JC, Boldt D, Kornfeld S, Skinner A, Valeri CR (1972) Excretion of deoxyribonucleic acid by lymphocytes stimulated with phytohemagglutinin or antigen. *Proc Natl Acad Sci USA* 69(7):1685–1689.
43. Lei W, et al. (2009) Abnormal DNA methylation in CD4+ T cells from patients with systemic lupus erythematosus, systemic sclerosis, and dermatomyositis. *Scand J Rheumatol* 38(5):369–374.
44. Zhang Y, Zhao M, Sawalha AH, Richardson B, Lu Q (2013) Impaired DNA methylation and its mechanisms in CD4(+)T cells of systemic lupus erythematosus. *J Autoimmun* 41:92–99.
45. Tsui NBY, et al. (2012) High resolution size analysis of fetal DNA in the urine of pregnant women by paired-end massively parallel sequencing. *PLoS ONE* 7(10):e48319.
46. Li R, et al. (2009) SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics* 25(15):1966–1967.
47. Jiang P, et al. (2014) Methy-Pipe: An integrated bioinformatics pipeline for whole genome bisulfite sequencing data analysis. *PLoS ONE* 9(6):e100360.