บทบาทของดีเอนเอเมททิเลชั่นที่ลำดับเบสของเอชอีอาร์วีต่อการเกิดโรคเอสเอลอี

นายจีรวัฒน์ นาคขุนทด

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาจุลชีววิทยาทางการแพทย์ (สหสาขาวิชา)
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2554

ROLE OF DNA METHYLATION AT HERV SEQUENCE IN SYSTEMIC LUPUS ERYTHEMATOSUS

Mr. Jeerawat Nakkuntod

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Medical Microbiology
(Interdisciplinary Program)
Graduate School
Chulalongkorn University
Academic Year 2011

Thesis Title	ROLE OF DNA METHYLATION AT HERV SEQUENCE IN SYSTEMIC LUPUS ERYTHEMATOSUS
By	Mr Jeerawat Nakkuntod
Field of Study	Medical Microbiology
Thesis Advisor	Associate Professor Nattiya Hirankarn, M.D.,Ph.D.
Thesis Co-advisor	Professor Apiwat Mutirangura, M.D.,Ph.D.

Accepted by the Graduate School, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

…….. Dean of the Graduate School
(Associate Professor Pornpote Piumsomboon,Ph.D.)

THESIS COMMITTEE

…….. Chairman
(Associate Professor Ariya Chindamporn,Ph.D.)

…….. Thesis Advisor
(Associate Professor Nattiya Hirankarn, M.D.,Ph.D.)

…….. Thesis Co-advisor
(Professor Apiwat Mutirangura, M.D.,Ph.D.)

…….. Examiner
(Associate Professor Jongkonnee Wongpiyabovorn, M.D.,Ph.D.)

…….. Examiner
(Associate Professor Tanapat Palaga,Ph.D.)

…….. External Examiner
(Nakarin Kitkumthorn, D.D.S.,Ph.D.)

จีรวัฒน์ นาคขุนทด : บทบาทของดีเอนเอเมททิเลชั่นที่ลำดับเบสของเอชอีอาร์วีต่อการเกิดโรคเอสเอลอี. (ROLE OF DNA METHYLATION AT HERV SEQUENCE IN SYSTEMIC LUPUS ERYTHEMATOSUS) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.พญ. ดร.ณัฏฐิยา หิรัญกาญจน์,อ. ที่ปรึกษาวิทยานิพนธ์ร่วม : ศ.นพ.ดร.อภิวัฒน์ มุทิรางกูร , 175 หน้า.

การศึกษาก่อนหน้านี้พบว่าเม็ดเลือดขาวชนิดทีเซลล์ของผู้ป่วยเอสเอลอี มีระดับดีเอ็นเอเมททิลเลชั่นลดลงทั้งจีโนม และที่บริเวณโปรโมเตอร์ของบางยีน แต่กลับยังไม่มีการศึกษาระดับการเติมหมู่เมททิลบนดีเอนเอที่มีลำดับเบสซ้ำกัน ดังนั้นการศึกษานี้ต้องการเปรียบเทียบระดับเมททิลเลชั่นบริเวณตำแหน่งไลน์-1 (LINE-1) อะลู (Alu) และ HERV ในเซลล์เม็ดเลือดขาวชนิด CD4+ ทีลิมโฟไซท์, CD8+ ทีลิมโฟไซท์ และ บีลิมโฟไซท์ ของผู้ป่วยเอสเอลอีกับคนปกติ โดยพบว่าจีโนมของเซลล์ชนิด CD4+ ทีลิมโฟไซท์, CD8+ ทีลิมโฟไซท์ และ บีลิมโฟไซท์ ของผู้ป่วยเอสเอลอี มีระดับเมททิลเลชั่นลดลงที่บริเวณตำแหน่งไลน์-1 และพบความแตกต่างอย่างมีนัยสำคัญทางสถิติอย่างมาก ในเซลล์ชนิด CD4+ ทีลิมโฟไซท์ และ CD8+ ทีลิมโฟไซท์ของผู้ป่วยเอสเอลอีในระยะกำเริบของโรค และพบว่ามีระดับเมททิลเลชั่นลดลงในบริเวณตำแหน่ง HERV-E LTR2C ในเซลล์ชนิด CD4+ ทีลิมโฟไซท์ของผู้ป่วยเอสเอลอีในระยะกำเริบ โดยมีความสัมพันธ์กับภาวะ leucopenia และ lymphopenia ในขณะที่ระดับเมททิลเลชั่นที่ลดลงในบริเวณตำแหน่ง HERV-K LTR5_Hs ของเซลล์ชนิด CD4+ ทีลิมโฟไซท์มีความสัมพันธ์กับระดับของ complement activity และ ค่า SLEDAI เมื่อทำการวิเคราะห์ในระดับจีโนมพบว่า ใน CD4+ ทีลิมโฟไซท์ของผู้ป่วยเอสเอลอี และ ใน CD4+ ทีลิมโฟไซท์ที่มีระดับดีเอ็นเอเมททิลเลชั่นลดลงทั้งจีโนม มีการแสดงออกของยีนในกลุ่ม IFN-induced เพิ่มขึ้น โดยไม่ได้เกี่ยวข้องทั้งกับการลดลงของระดับดีเอ็นเอเมททิลเลชั่นที่บริเวณโปรโมเตอร์และการมีลำดับเบสของไลน์-1 อยู่ในยีน จากการศึกษานี้สรุปได้ว่าลิมโฟไซท์ของผู้ป่วยเอสเอลอี มีการลดลงของระดับเมททิลเลชั่นอย่างจำเพาะโดยขึ้นกับชนิดของลำดับเบสซ้ำ และมีหลักฐานที่แสดงว่าการลดลงของระดับเมททิลเลชั่นบริเวณตำแหน่งไลน์-1 นี้น่าจะส่งผลให้มีการแสดงออกของยีนในกลุ่ม IFN-induced เพิ่มสูงขึ้น ดังนั้นการศึกษาบทบาทของระดับเมททิลเลชั่นที่ลดลงนี้ อาจนำไปสู่การค้นพบกลไกใหม่ของการเกิดโรคเอสเอลอี

สาขาวิชา จุลชีววิทยาทางการแพทย์ ลายมือชื่อนิสิต .....

ปีการศึกษา 2554 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....

ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม .....

4989656420 : MAJOR MEDICAL MICROBIOLOGY
KEYWORDS : DNA METHYLATION / SYSTEMIC LUPUS ERYTHEMATOSUS / HERV-E / HERV-K / LINE-1 / ALU

JEERAWAT NAKKUNTOD : ROLE OF DNA METHYLATION AT HERV SEQUENCE IN SYSTEMIC LUPUS ERYTHEMATOSUS. ADVISOR : ASSOC. PROF. NATTIYA HIRANKARN, M.D.,Ph.D., CO-ADVISOR : PROF. APIWAT MUTIRANGURA, M.D.,Ph.D., 175 pp.

Previous studies reported that T cells from active SLE patients contained globally hypomethylation and demethylation at promoter of several genes which contributed to disease pathogenesis. So far there is limited information about methylation profile of retroelements in SLE. In this study, we examined and compared the methylation levels of LINE-1, Alu, HERV-E and HERV-K in normal and SLE CD4+ T lymphocytes, CD8+ T lymphocytes and B lymphocytes. Hypomethylation of LINE-1 but not Alu was found in CD4+ T lymphocytes, CD8+ T lymphocytes, and B lymphocytes of SLE patient. Strikingly, LINE-1 hypomethylation was more significantly distinguished in both CD4+ and CD8+ T lymphocytes of patients from the active SLE group. For LTR retroelements, hypomethylation of HERV-E LTR2C was observed in CD4+ T cells of active SLE. Moreover, the hypomethylation of HERV-E LTR2C was correlated with leucopenia and lymphopenia in active SLE while the hypomethylation of HERV-K LTR5_Hs in CD3+CD4+ T cells was significantly correlated with complement activity and SLEDAI score. From comprehensive analysis of genome and gene expression array, IFN-induced genes showed up-regulation in methylation sensitive gene in CD4+ T cells of SLE. However, hypomethylation at promoter region and containing of LINE-1 within gene does not explain this up-regulation. In conclusion, the hypomethylation in each lymphocyte subset of SLE was interspersed repetitive sequences (IRSs) type-specific. We also provided an evidence to propose consequence mechanism of LINE-1 hypomethylation in SLE that LINE-1 transcripts might have an effect in *trans* to up-regulate the IFN-induced gene. Further study to find out the functional role of IRSs hypomethylation might lead to the discovery of novel pathogenesis pathway in SLE.

Department : Microbiology Student's Signature .....
Field of Study : Medical microbiology Advisor's Signature .....
Academic Year : 2011 Co-advisor's Signature .....

ACKNOWLEDGEMENTS

I would like to thank everyone who encouraged and supported me to complete this dissertation. This dissertation could not have been written without the support of the Royal Golder Jubilee Ph.D. program, National Research Council of Thailand 2008-2010, Lupus Research Unit, the 90th Year Anniversary of Chulalongkorn University (Ratchadaphiseksomphot Endowment Fund), National Research University project from the Commission of Higher Education, the Thailand Research Fund (TRF) senior research scholar grant, Center of Excellence in Molecular Genetics of Cancer and Human Diseases, Department of Anatomy, Faculty of Medicine, Chulalongkorn University

CONTENTS

PAGE

PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

COBRA	Combined Bisulfite Restriction Analysis
CU-DREAM	Connection Up- or Down-Regulation Expression Analysis of Microarray
DNMTs	DNA Methyltransferases
GEO	Gene Expression Omnibus
HERVs	Human Endogenous Retroviruses
IRSs	Interspersed Repetitive Sequences
LINE-1	Long Interspersed Nuclear Element type 1
LTRs	Long Terminal Repeats
ORE	Normal Oral Epithelium
PBMCs	Peripheral Blood Mononuclear Cells
RA	Rheumatoid arthritis
SLE	Systemic Lupus Erythematosus
SLEDAI-2K	Systemic Lupus Erythematosus Disease Activity 2000

CHAPTER I

INTRODUCTION

Systemic lupus erythematosus (SLE) is an autoimmune disease that affects multiple organs, including the skin, joints, kidneys, heart, lungs, and central nervous system. The hallmark of SLE is the production of autoantibodies directed against constituents of the cell nucleus such as anti-nuclear antibodies and antibodies to double stranded DNA. These autoantibodies cause end-organ damage via inflammatory response to immune complex, which result in various clinical manifestations such as glomerulonephritis, arthritis, serositis, vasculitis and neurological disorders (1) Although the precise etiopathogenesis of SLE remains elusive, several studies have been reported that genetic and environmental factors are an important in susceptibility to SLE (2) Furthermore, there is increasing evidence that epigenetics modifications, the study of heritable changes in gene expression that occur without a change in DNA sequence,(3) contribute in the pathogenesis of SLE.

DNA methylation is a postsynthesis DNA modification, which in eukaryotes occurs at the 5 position of the cytosine ring in the context of CpG dinucleotides. There are regions of the genome termed CpG islands which are rich of CpG pairs. Most of CpG islands overlap the 5' end of gene regions, including promoter and first exon sequences. In general, the promoters of active genes and tissue-specific genes are typically hypomethylated, while inactive genes are silenced by methylation. The process of DNA methylation is carried out by DNA methyltransferase (3)

The role of DNA methylation in pathogenesis of SLE has been reported. From independent study found that CD4+ T lymphocyte of SLE patient is globally hypomethylation comparing to normal group (4-6) This hypomethylation was associated with decreasing of Dnmt1 mRNA levels (5) Moreover, CD4+ T cells treated with the DNA methylation inhibitors (5-azacytidine, procainamide or hydralazine) became autoreactive. The adoptive transfers of 5-azacytidine or procainamide-treated T cells to syngeneic

recipient mice caused a lupus-like disease (7, 8) Therefore, CD4+ T cell hypomethylation may contribute to the development of SLE. Subsequently, the CD4+ T cell autoreactivity was found to be due in part to an overexpression of the adhesion molecule lymphocyte function-associated antigen 1 (LFA-1), perforin and CD70 (9-11) The increased expression of these genes in CD4+ T cells related with demethylation at promoter sequence, as observed in SLE patients. Recently, overexpression and demethylation of *CD40LG* was found only in CD4+ T cell from female with SLE. This result was suggested that demethylation in X inactivation chromosome may be an important factor result in highly incidence of SLE in female than male (12) However, genes that are aberrant of methylation are largely unknown.

Besides the CpG-rich sequences within genes, most of the methylated cytosines can be found in the Interspersed Repetitive Sequences (IRSs) (13, 14) Human genome contains approximately 45% of IRSs which can be divided into DNA transposons and retroelements, encompass about 2.8% and 42.2% of the human genome, respective (15) Retroelements can be divided into 2 groups based on the presence or absence of long terminal repeats (LTRs). Two high copy number of non-LTR retroelement are long interspersed nuclear elements (LINEs) (20.1%) and short interspersed nuclear elements (SINEs e.g. ALU) (13.1%) (16) While, the majority of LTR retroelements is Human endogenous retroviruses (HERVs) (8.2 %) (16) IRSs have often been referred to as junk DNA sequence. However, the functional role of IRSs in human genome have suggested that they are *cis* regulatory elements to provide enhancers, repressors, alternative promoter, a new exon coding protein and polyadenylation signal to neighboring genes (17-19) Interestingly, IRSs are methylated to variable degrees depending on their location under normal conditions (20, 21) Therefore, changes in IRSs methylation should alter cellular functions. Beside global hypomethylation in CD4+ T cell of SLE, there have also been found in several condition such as cancer (22), embryogenesis (23, 24), aging (25) and congenital malformation (26) Global hypomethylation may be the one of mechanisms that promotes these conditions but the consequence is different. The methylation levels of IRSs

have been studied. In cancers, LINE-1 is the IRS element that is most frequently studied, and hypomethylation of LINE-1 is common in cancer cells (27) Therefore, hypomethylation of LINE-1 has been reported to be a prognostic marker and used as a screening tool for cancer detection (28, 29) The consequences of LINE-1 hypomethylation are genomic instability and alteration in gene expression (27) Interestingly, hypomethylation of Alu element and HERV-K, but not LINE-1, were found in aging, (30) Therefore, it is revealed that loss of genome-wide methylation is IRSs-specific.

Several studies have been reported on association of HERVs and autoimmune disease. Peripheral blood mononuclear cells (PBMCs) of SLE and rheumatoid arthritis (RA) patients showed increased expression of HERV-E and HERV-K *gag* gene, respectively (31-33) Furthermore, transcription of HERV-E *gag* gene correlated with blood plasma concentrations of anti-U1 ribonucleoprotein (RNP) and anti-Sm antibodies, are components of the same spliceosome complex and participate in splicing of pre-mRNA to mRNA, of SLE patients (32) Similar to other IRSs, influences of DNA methylation to function of HERVs have been reported. HERV-E LTR-derived promoter of pleiotrophin, endothelin B receptor and midline 1 genes were different in methylation levels between placenta and blood cells. These methylation levels were associated with gene expression (21) Methylations of HERV-K LTRs in different loci were shown to have distinct pattern with inverse correlation with transcriptional activity (34) In addition, treating PBMCs and Tera-1 cell line (human testicular tumor) with 5-azacytidine showed that alteration of methylation level can increase expression of HERV-E *gag* transcript and HERV-K Gag protein, respectively (31, 35) From previous results may suggest that loss of methylation in lymphocyte influence the up-regulation of HERV-E in SLE patient. Interestingly, recent study showed that SLE patients have demethylation in HERV-E alternative promoter of CD5 gene and can lead to up-regulation of CD5-E1B isoform in SLE B cells (36) supporting that fact that the IRSs potential to regulate the adjacent cellular gene and also contribute to the pathogenesis of SLE.

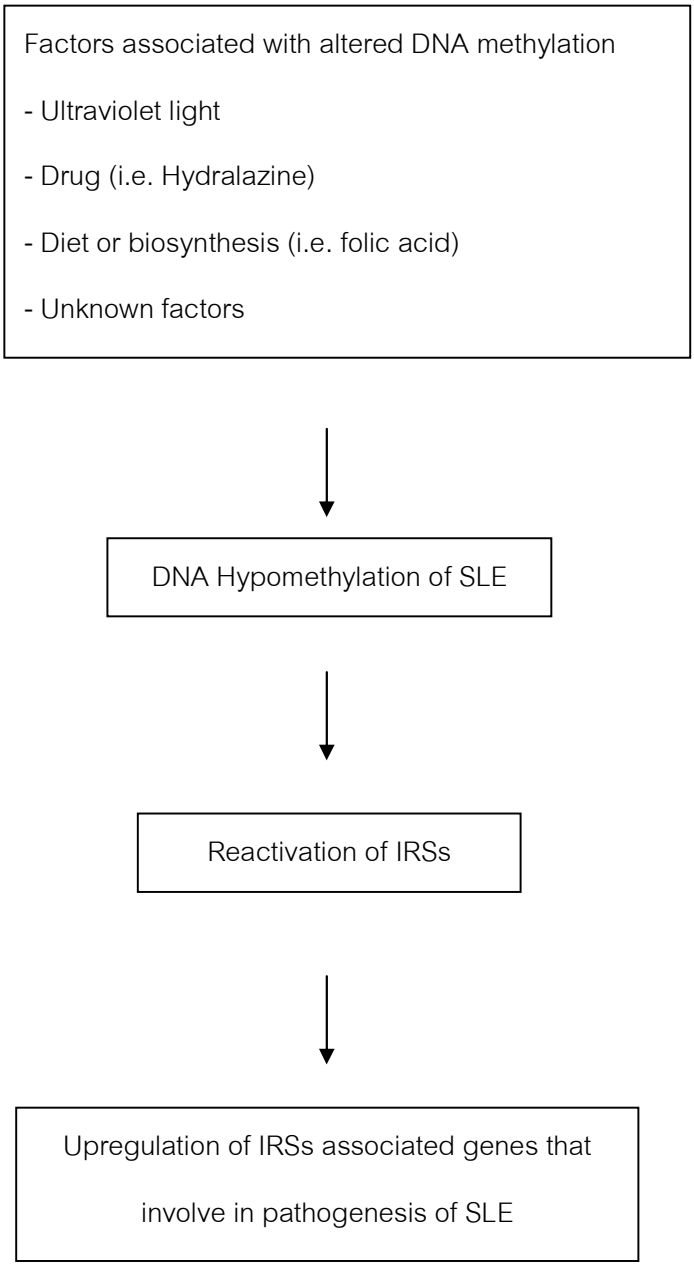
Because cellular biology of cancer and SLE are different, we hypothesized that the SLE dependent pattern of methylation losses in SINEs, LINEs and LTRs should be different. So far there is limited information about methylation profile of IRSs in SLE. In this study, we

determined and compared the methylation status of LINE-1, Alu and HERVs in normal and SLE lymphocyte subsets. We then compared methylation levels in lymphocyte subsets from inactive and active SLE with healthy subjects. As mentioned above, there is potential of IRSs sequence to regulate the adjacent cellular gene. Therefore, to study the methylation of IRSs may lead to understand how the same epigenetic characteristic is associated with a variety of cellular phenotypes.

The objectives of this thesis study were:

1. To compare the methylation levels of LINE-1, Alu, HERV-E and HERV-K in CD4+ T lymphocytes, CD8+ T lymphocytes and B lymphocytes of normal and SLE patients.
2. To investigate the role of IRSs in CD4+ T cells of SLE patients.

Conceptual framework



CHAPTER II

LITERATURE REVIEW

1. Systemic lupus erythematosus (SLE)

Systemic lupus erythematosus (SLE) is an autoimmune disease that affects multiple organs, including the skin, joints, kidneys, heart, lungs, and central nervous system. The hallmark of SLE is the production of autoantibodies directed against constituents of the cell nucleus such as anti-nuclear antibodies and antibodies to double stranded DNA. These autoantibodies cause end-organ damage via inflammatory response to immune complex, which result in various clinical manifestations such as glomerulonephritis, arthritis, serositis, vasculitis and neurological disorders (1) The disease severity can also differ between the patients. Some patients have mild symptoms, whereas others are severely clinical in chronic inflammation of multiple organs. These require aggressive treatment with high-dose corticosteroids drugs. Due to the heterogeneity of this disease, the American College of Rheumatology (ACR) has established 11 criteria for classification (Table 1) (37, 38) Fulfillment of at least four criteria is necessary for the classification of SLE.

1.1 Epidermiology

SLE is found worldwide with incedence rates range from approximately 1 to 10 per 100,000 person-years and prevalence rates range from 20 to 70 per 100,000 depending on racial and geographic background (39) The one of risk factor for SLE is clearly gender. The incidence and prevalence rates for men to women are approximately 1:10 (39) The peak age of onset is during reproductive years. Interestingly, populations of African–American, Asian and Hispanic individuals are more prevalence and severely affected than Caucasians (40)

1.2 Aetiology of SLE

Although the precise etiopathogenesis of SLE remains elusive, several studies have been reported that genetic and environmental factors are an important in susceptibility to SLE.

1.2.1 The genetic component of SLE

SLE has a complex genetic basis, as shown by an increased risk to develop the disease for relatives of SLE patients. Epidemiological studies show an estimated relative risk to siblings of affected individuals (λs) 15–20-fold over that of the general population (41) Twin studies also indicate a disease concordance of 2–5% in dizygotic compared with 24–57% in monozygotic twins (42, 43) A number of genetic risk factors for SLE have been identified by genome wide association studies (GWAS). Recently, three genome wide GWAS in Caucasian patients with SLE from international collaborations have been reported (44-46) Results from these studies confirm known candidate genes including *HLA, FCGR, PTPN22*, *STAT4* and *IRF5*. For instance, the discovery of *BLK* and *BANK1* genes emphasize the crucial role of B cell in pathogenesis of SLE. Another novel gene namely, *ITGAM* was also identified which is an adhesion molecule that regulates leukocyte adhesion to endothelial cells and may contribute to vasculitis in patients with SLE. From these reports of international collaborations found that odd ratios (OR) were rather low ranged between 1.25 and 2.36. Newly, a genome-wide association study has found the association of three independent SNPs in the *TNFAIP3* region with SLE (OR=1.3-2) (47) This data highlight the inflammatory role of TNF pathway in the pathogenesis of SLE. Furthermore, a recent genome wide association (GWA) study in Asian population demonstrated the important role of *ITPR3* (OR=3.39) (48) and *TNXB* (49) on chromosome 6p21 to be candidate gene susceptible to SLE in the Japanese population.

1.2.2 Environmental factors

The concordance rates in twins studies are still relatively low, indicating involvement of environmental factors. Various infectious agents, especially the Epstein-Barr Virus (EBV)

have been reported with the disease. Possible mechanisms are antibody cross-reactivity between viral antigens and self antigens, and immortalisation of self-reactive B cells by EBV (50-53) UV radiation can also potentially trigger disease onset (54) Approximately 73% of patients with SLE report photosensitivity (55) There are some of medications inducing lupus-like disease that is a side-effect of long-term use of medications. The drugs most frequently implicated in the development of drug-induced lupus are hydralazine and procainamide (56) These drugs could induce lupus disease by inhibit T cell DNA methylation become autoreactive after treatment (9) Other examples of environmental factors associated with risk of SLE are exposure to high levels of silica dust (57), diet (58) and smoking (59) There are also some indications that could be risk factors such as heavy metals, solvents, pollutants, pesticides and hair dyes (54)

Table 1. The American College of Rheumatology's criteria for classification of SLE, revised in 1982 and 1997.

Criterion	Definition
Malar rash	Fixed erythema, flat or raised, over the malar eminences, tending to spare the nasolabial folds
Discoid rash	Erythematosus raised patches with adherent keratotic scaling and follicular plugging; atrophic scarring may occur in older lesions
Photosensitivity	Skin rash as a result of unusual reaction to sunlight, by patient history or physician observation
Oral ulcers	Oral or nasopharyngeal ulceration, usually painless, observed by a physician
Arthritis	Nonerosive arthritis involving 2 or more peripheral joints, characterized by tenderness, swelling or effusion
Serositis	a) Pleuritis (convincing history of pleuritic pain or rub heard by a physician or evidence of pleural effusion) OR b) Pericarditis (documented by ECG or rub or evidence of pericardial effusion)
Renal disorder	a) Persistent proteinurea > 0.5 g/day, or greater than 3+ if quantification not performed OR b) Cellular casts: red cell, hemoglobin, granular, tubular or mixed
Neurologic disorder	a) Seizures (in the absence of offending drugs or known metabolic derangements) OR b) Psychosis (in the absence of offending drugs or known metabolic derangements)
Haematologic disorder	a) Haemolytic anaemia with reticulocytosis OR b) Leukopenia (<4000/mm^3 on 2 or more occasions) OR c) Lymphopenia (<1500/mm^3 on 2 or more occasions) OR d) Thrombocytopenia (<100,000/mm^3 in the absence of offending drugs)
Immunologic disorder	a) Antibody to native DNA in abnormal titre OR b) Antibody to Sm nuclear antigen OR c) Antiphospholipid antibodies - based on 1) abnormal serum level of IgG or IgM anticardiolipin antibodies, 2) positive test result for lupus anticoagulant using a standard method, or 3) a false-positive serologic test for syphilis for at least 6 months confirmed by Treponema pallidum immobilization or fluorescent treponemal antibody absorption test
Anti-nuclear antibodies	Abnormal titre of antinuclear antibody by immunofluorescence or equivalent assay, in the absence of drugs known to be associated with 'drug-induced lupus' syndrome

1.3 Pathogenesis of SLE

As mention above, SLE is multifactorial disease which genetic and environmental factors are involved in pathogenesis. The following steps have been suggested: an immune response is triggered by an environmental factor, such as an EBV infection, drug, UV. Genetic predisposition influences the immune response. Furthermore, gender is also an additional predisposing factor. Several studies indicate that SLE patients increased rate of apoptosis in circulating and abnormal of clearance an apoptotic materials. These antigens are taken up by antigen-presenting cells, then present peptides to T cells. Activated T cells stimulate B cells to produce autoantibodies that can bind to self-antigen. Defective in regulation of immune system is major cause of autoreactivity. Increasing of immune complexes and defective clearance of these complexes causes deposition in tissues. Immune complex induces an immune mediated inflammation that causes following tissue damage. This pathogenic step is shown in Figure 1.

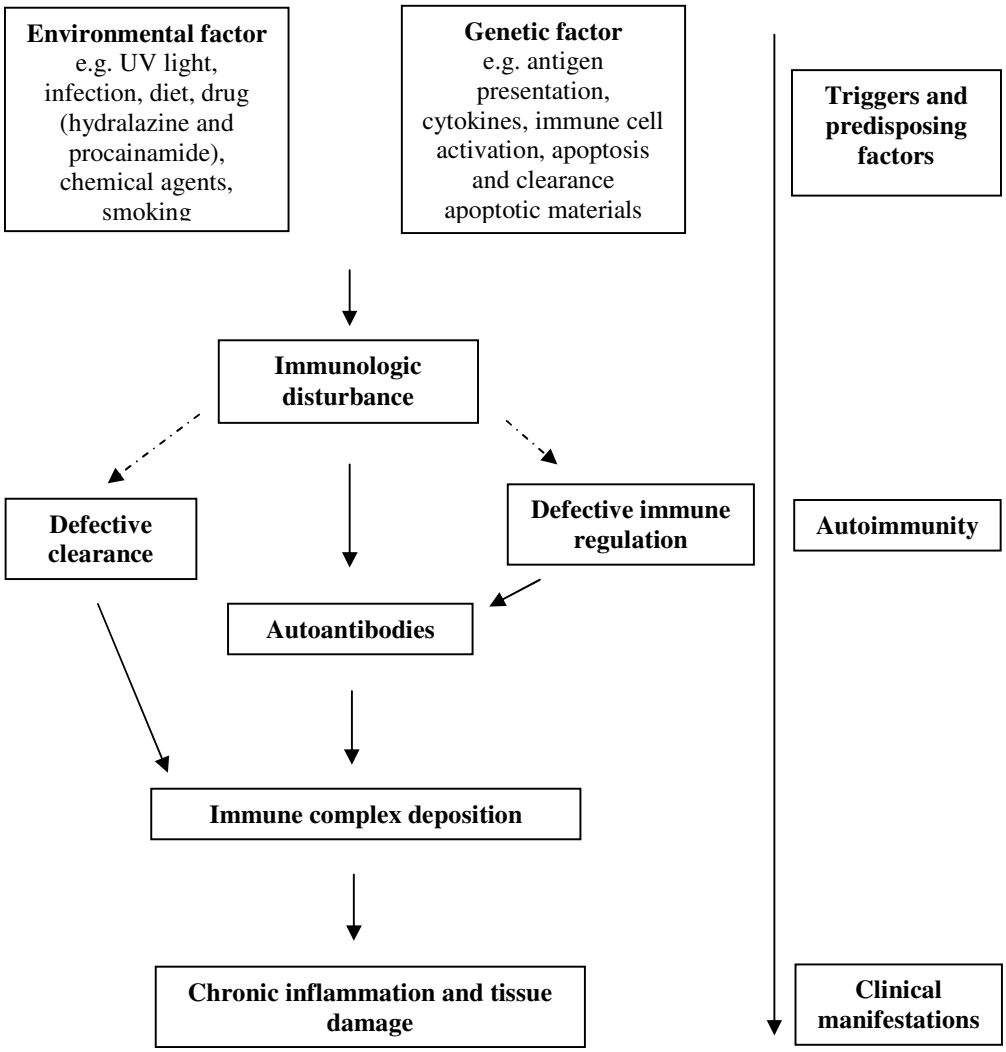


Figure 1. Pathogenic steps in SLE

2. DNA methylation and SLE

Although the role of genetic factors in SLE are well established. Several studies reported that environmental factors, including UV and drugs (hydralazine and procainamide), modulate the susceptibility of SLE in part of epigenetic changes (9, 60) Epigenetics is the study of heritable changes in gene expression that do not involve changes in DNA sequence, including histone modifications and DNA methylation (3) Growing evidence indicates that aberrant in DNA methylation plays essential roles in the pathogenesis of SLE.

2.1 DNA methylation

DNA methylation is a postsynthesis DNA modification, which methylation in human cells mainly affects the cytosine residues that are located at 5' to a guanosine. CpGs are enriched in short stretches of genome known as CpG islands, and the majority of them are located 5' end of gene regions, including promoters and first exon (3) CpG island is considered a contiguous window of DNA of at least 500 bp in length, with a C+G content more than 55% and an observed CpG over expected CpG ratio in excess of 0.65 (61) In mammalian cells, methylation of cytosines in CpG dinucleotides is catalyzed by DNA methyltransferases (DNMTs). This enzyme transfers the methyl groups from S-adenosylmethionine to cytosine residues in DNA. Functional DNMTs are encoded by at least three different genes called *DNMT1*, *DNMT3A* and *DNMT3L*. Maintenance of DNA methylation is known to be carried out by DNMT1. This enzyme preferentially methylates hemimethylated DNA to create the methylation pattern for newly replicated daughter strands during mitosis, based on that of the parent strand (62) DNMT3A and DNMT3L mainly add a methyl group to unmethylated CpG (*De novo* methylation), resulting in the creation of new hemimethylated and following fully methylated CpG (63) (figure 2). In general, the promoters of active genes and tissue-specific genes are typically hypomethylated, while inactive genes are silenced by methylation. There are 2 mechanisms of transcriptional repression mediated by cytosine methylation. First, methylated CpG can directly interfere with of transcription factors binding. This binding can

also prevented by indirectly mechanism from recruiting methyl binding domains and co-repressors proteins to methylated CpG (64)

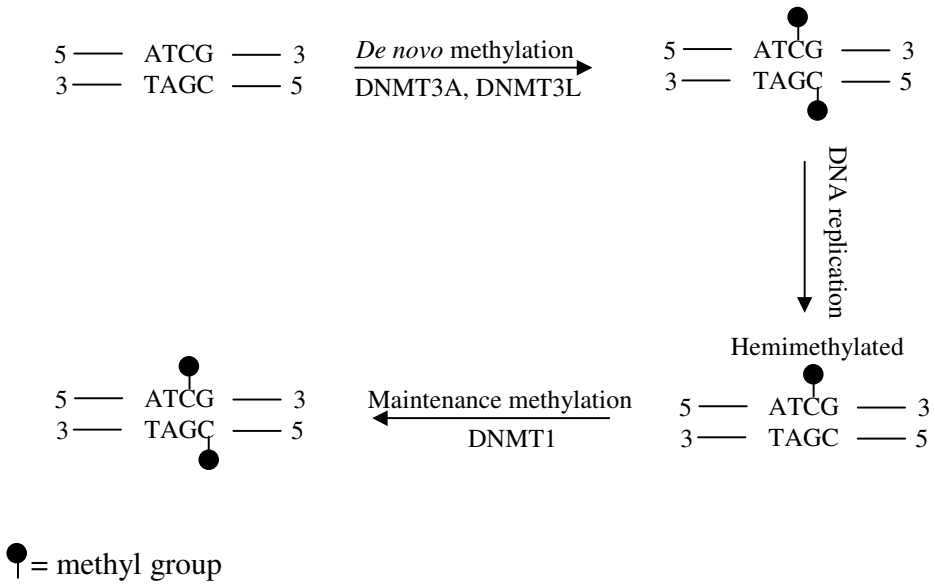


Figure 2. Mechanism of DNA methylation

2.2 Role of DNA methylation in SLE

From independent studies it was found that leukocytes (65), PBMCs (60), T cell (66), CD4+ T lymphocyte (4-6) of SLE patient is globally hypomethylation comparing to normal group. This hypomethylation was associated with decreasing of DNMT1 mRNA levels (5) The evidence of the role of demethylation in development of SLE comes from studies with demethylating agents. One of the most common demethylating drugs is 5-azacytidine, a cytosine analog that contains a nitrogen atom at the 5' position of the pyrimidine ring and is incorporated into newly synthesized DNA. Treatment with 5-azacytidine causes global hypomethylation. The mechanisms by which hypomethylated T cells induce SLE have been performed by Richard et al. They showed that treating T cell with demethylating agents (procainamide, hydralazine, or 5-azacytidine) induces major histocompatibility complex–specific T cell autoreactivity (8, 67-70) 5-azacytidine caused autoreactivity by demethylation at upstream of the ITGAL promoter leading to overexpression of leukocyte function-associated antigen-1 (LFA-1; CD11a/CD18). Furthermore, T cells transfected with LFA-1 caused autoreactivity in vitro, and injecting into syngeneic mice lead to lupus-like autoimmunity (9) Demethylation and overexpression of LFA-1 were also observed in SLE patients (71) Subsequently, demethylation and overexpression of perforin (*PRF1*) and CD70 (*TNFSF7*) were found in 5-azacytidine treated T cell as in SLE patients (10, 11) Perforin could contribute to macrophage killing by SLE T cells, generating a source of antigenic apoptotic nucleosomes (10) While CD70 overexpression contributes to excessive B cell stimulation in lupus (11) Recent study reported that overexpression and demethylation of CD40 ligand (*TNFSF5*) was found only in CD4+ T cell from female with SLE. This result was suggested that demethylation in X inactivation chromosome may be an important factor result in highly incidence of SLE in female than male (12) Moreover, the adoptive transfers of 5-azacytidine or procainamide-treated T cells to syngeneic recipient mice caused a lupus-like disease (7, 8) Therefore, CD4+ T cell hypomethylation may contribute the development of SLE.

Inhibiting ERK pathway signaling by using hydralazine caused decreasing of Dnmt1 level, demethylate and overexpression of methylation-sensitive genes (11, 72, 73) This

observation indicated that T cell DNA methylation in SLE may be due to defect in ERK signaling pathway. T cells of active SLE patients also decreased in ERK signaling pathway and Dnmt1 levels. Moreover, the degree of impairment was negative correlation to disease activity (74) Abnormality of ERK signaling in CD4+ T cells from SLE patients and in hydralazine-treated cells localized to protein kinase C (PKC) δ. Inhibiting of PKCδ or transfection with a dominant negative PKCδ caused demethylation of *TNFSF7* promoter and overexpression of CD70 similar to SLE and hydralazine-treated T cells (75) In animal model, murine CD4+ T cells showed demetylate and overexpression of LFA-1 after treating with hydralazine or the MEK inhibitor U0126 (73) Adoptive transfers of inhibited ERK pathway T cells to syngeneic recipient mice was sufficient to induce a lupus-like disease (73) Interestingly, a dominant negative MEK transgenic mice model decreased expression of Dnmt1 and overexpresson of CD11a and CD70, as in T cell of human SLE. This transgenic mice also develops anti-dsDNA antibodies. Moreover, interferon-regulated genes were upregulated in the spleen similar to PBMCs of SLE patients (76)

3. Retroelements

Human genome contains approximately 45% of transposable elements which can be divided into DNA transposons and retroelements, encompass about 2.8% and 42.2% of the human genome, respective (15) DNA transposons are generally excised themselves as DNA from one genome site and inserted into new position by a “cut and paste” mechanism. Meanwhile, retroelements duplicate by “copy and paste” mechanism. Retroelements are transcripted into RNA and reversed by using reverse transcriptase and inserted at new location (77)

Retroelements can be divided into 2 groups based on the presence or absence of long terminal repeats (LTRs). Two high copy number of non-LTR retroelement are long interspersed nuclear elements (LINEs) (20.1%) and short interspersed nuclear elements (SINEs e.g. ALU) (13.1%) (16) While, the majority of LTR retroelements is Human endogenous retroviruses (HERVs) (8.2 %) (16)

4. Retrovirus

Retroviruses are enveloped viruses containing two single-stranded RNA genomes. A typical size of retroviral provirus is about 6-11 kb. The structure of a provirus consists of *gag, pro, pol* and *env* genes which are flanked by two LTRs (Figure 3). LTRs are composed of three domains from the 5' to the 3' end namely unique 3' (U3), repeat (R) and unique 5' (U5). The U3 region contains the promoter and regulatory regions (78), which influence the expression of neighboring cellular genes (79) The primer binding site (PBS) for a specific tRNA molecule used to initiate reverse transcription. The tRNA varies among different retroviruses and was used to name and group endogenous retroviruses in the human genome. For example, members of the HERV-E and HERV-K have a PBS with glutamic acid-tRNA and lysine-tRNA, repectively (80) Four important genes of provirus, the *gag* gene is encoded for three structural proteins in the virion including matrix (MA), capsid (CA) and nucleocapsid (NC). The *pro* gene encodes a protease to cleave the initial protein into functional proteins. The *pol* gene is coded for reverse transcriptase, RNase H and integrase. Reverse transcriptase is used to transcript RNA into DNA. RNase H is used to cleave a DNA/RNA hybrid during DNA synthesis. Integrase is used to insert the viral DNA into the host chromosomal DNA. Finally, the *env* gene encodes the viral membrane protein to bind with host receptor and penetrate the cell membrane, which is importance of intercellular transmission.

After infecting of host cell, the reverse transcriptase is used to make a viral DNA from viral RNA. Then, a complementary viral DNA strand is made. These double strand copies of viral DNA are migrated to nucleus and inserted into the host chromosome as provirus, that can be transmitted vertically in a Mendelian fashion (81) The proviral DNA is also transcribed to virus-related RNA by using the host-cell RNA polymerase. These RNA strands are used as templates for making new copies of the viral chromosomal RNA and can be translated into viral proteins that are used to make the virus envelope. Then, new viral particles are assembled, bud from the plasma membrane, and are released. This virus is now prepared to initiate a new round of replication (82)

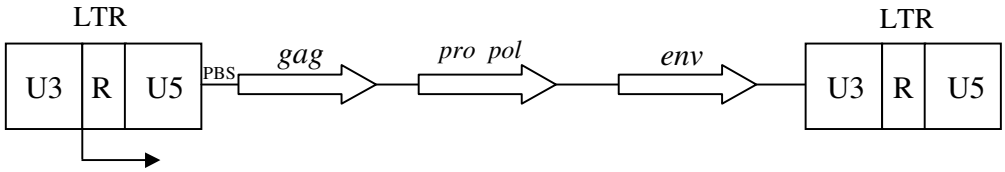


Figure 3. Structure of the retroviral genome. LTR: long terminal repeat, U3: unique 3' sequence, R: repeat sequence; U5: unique 5' sequence; PBS: primer binding site; *gag*: group-specific antigen; *pro*: protease; *pol*: RNA-dependent DNA polymerase, *env*: envelope.

5. Human endogenous retrovirus (HERVs)

HERVs are sequences that resulted from ancestral infections by provirus genomes into the germ-line DNA (15, 83) The majority of HERVs are also detected in apes and Old World monkeys genomes. This can indicate that HERVs originated in our ancestors at least 25 million years ago (81) Human genome contains approximately 8% of HERV sequences (84) A typical size is about 6-7 kb. The structure of intact HERVs consists of four internal universal genes, including *gag, pro, pol* and *env*, which are flanked by two LTRs as exogenous retrovirus (Figure 3). Because most infectious retroviruses are harmful, HERVs are maintained by mutations involving numerous deleterious mutations such as smaller deletions, frame shifts or stop codons and hypermethylation of promoters to prevent the expression of functional proteins from proviral genes (85) In most cases, HERVs contain in the human genome as solitary LTRs due to recombination of the two LTRs (86)

5.1 Classification of HERVs

According to their similarities to mammalian retroviruses, HERVs have been divided into three classes (Table 2). Where Class I consists of gamma-like retroviruses formerly known as type C, this class is subdivided into six groups include HERV-H, HERV-I, HERV-W and HERV-E. Class II includes beta-like retroviruses known as type A, B and D which is subdivided into 10 groups include HERV-K. Finally, Class III consists of foamy virus related HERVs, member include HERV-L (80) However, the classification of ERVs is difficult due to broad sequence variation and numerous mutations. Therefore a suitable classification is needed (87)

Table 2. Classification of human endogenous retrovirus (HERVs) (80)

HERV family	Representative accession number
Class I HERVs (type C related HERVs)	
Group 1, HERV-HF	
HERV-H (RTVL-H, RGH)	AF108842
HERV-F	AF070684
Group 2, HERV-RW	
HERV-W	AF072506
HERV-R (ERV9)	X57147
HERV-P (HuERS-P, HuRRS-P)	X06279
Group 3, HERV-ERI	
HERV-E (4-1, ERVA, NP-2*)	S46403
51-1	J00273
HERV-R (ERV3)	M12140
RRHERV-I	M64936
Group 4, HERV-T	
HERV T (S71, CRTK1, CRTK6)	M32788
Group 5, HERV-IP	
HERV-I (RTVL-I)	X14953
HERV-IP-T47D (ERV-FTD)	U27241
Group 6, ERV-FRD	
ERV-FRD	U27240
Class II HERVs (type A, B, and D related HERVs)	
Group 1, HERV-K (HML-1)	
HERV-K (HML-1.1)	U35102
Group 2, HERV-K (HML-2)	
HERV-K10	M14123
HERV-K-HTDV	X8227
Group 3, HERV-K (HML-3)	
HERV K (HML3.1)	U35153
Group 4, HERV-K (HML-4)	
HERV-K-T47D	AF020092
Group 5, HERV-K (HML-5)	
HERV -K-NMWV2	AF015995
Group 6, HERV-K (HML-6)	
HERV K (HML-6p)	U86698
Group 7, HERV-K (HML-7)	
HERV-K-NMWV7	AF016000
Group 8, HERV-K (HML-8)	
HERV-K-NMWV3	AF015996
Group 9, HERV-K (HML-9)	
HERV-K-NMWV9	AF016001
Group 10, HERV-K (HML-10)	
HERV-KC4	U07856
Class III Foamy virus related HERVs	
HERV-L	X89211

5.2 Functional of HERVs in human transcriptome

The functional role of HERV elements in human gene expression have been shown to provide enhancers, repressors, alternative promoter, a new exon coding protein and polyadenylation signal to neighboring genes (79, 88-91)

5.2.1 Exonization and alternative splicing

Retrotransposon sequences are through insertion into an intron, and subsequent are sometimes integrated into genes as a new exon in a process termed exonization (Figure 4A). Recent study based on computational analysis showed 50 protein coding exons completely derived from LTR of retrotransposons (89) For instance, LTR derived alternatively spliced exon of the interleukin 22 receptor, alpha 2 gene (IL22RA2) contributes additional 32 amino acids at exon 3/4 in splicing variant 1 (NM_052962) (89) However, the role of additional exon in this isofome is uncertain.

5.2.2 Alternative promoter

LTR of HERV has been shown to contribute promoter sequences that can control the expression of nearby genes (92) (Figure 4B). Initiation of transcription by LTR results in the production of alternative isoforms that are tissue specific. For instance, HERV-E LTR was shown to act as an alternative promoter of amylase gene (93), placenta-specific gene include endothelin-B receptor (94), midline-1 (95), pleiotrophin (96) Recent study reported B cell-specific alternative isoform of CD5 gene (CD5-E1B) that also derived from HERV-E LTR promoter (97, 98) The bidirectional LTR promoters activity were reported for the human Down Dyndrome critical region 4 (DSCR4) and DSCR8 and also examples in transcriptomic sequencing data (99, 100) Furthermore, it were recently showed that an antisense-oriented HERV-LTR act an alternative promoter for genes *GSDML* (gasdermin-like protein), *SLC4A8* (sodium bicarbonate cotranspoter) and *IFT172* (intraflagellar transport protein 172) (101, 102) They also observed that antisense transcript of *SLC4A8* and *IFT172* genes are shown to decrease the mRNA level of the corresponding genes (102)

Genome-wide analysis, van de Lagemaat *et al*., found many LTR-derived promoters including aromatase (*CYP19*) gene in placenta, bile acid Coenzyme A (*BAAT*) gene in liver and the erythroid-specific promoter of the carbonic anhydrase *(CA1)* gene (103) Recent study by using systematic computational analysis method revealed that many human genes produced LTR-chimeric transcripts (79) The transcriptomic sequence data also demonstrated that LTR promoters have an impact on tissue-specific and specie-specific of transcription (100)

5.2.3 Polyadenylation signal

HERV-LTR sequences provide polyadenylation signals that induce the termination of gene transcripts (104) (Figure 4C). Mager *et al*., reported that LTR of HERV-H provide the polyadenylation signal for HHLA2 and HHLA3 genes (91) They also analyzed of RNA from baboon, showed that the baboon HHLA2 and 3 genes use other polyadenylation signals. This study demonstrates that retroviral insertion have potential to shape the human transcriptome during the course of evolution (91)

5.2.4 Epigenetic regulation

Epigenetic is the one mechanism to prevent retrotransposon activity for cell. Similar to other transposable elements, influences of DNA methylation to function of HERVs have been reported. HERV-E LTR-derived promoter of pleiotrophin, endothelin B receptor and midline 1 genes were different in methylation levels between placenta and blood cells. These methylation levels were associated with gene expression (21) Methylations of HERV-K LTRs in different loci were shown to have distinct pattern with inverse correlation with transcriptional activity (34) In addition, treating PBMCs and Tera-1 cell line (human testicular tumor) with 5-azacytidine showed that alteration of methylation level can increase expression of HERV-E *gag* transcript and HERV-K Gag protein, respectively (31, 35) Because retroelements are frequently found near genes so heterochromatin formed at retroelements could spread and repress the transcription of nearby genes (19) (Figure 4D).

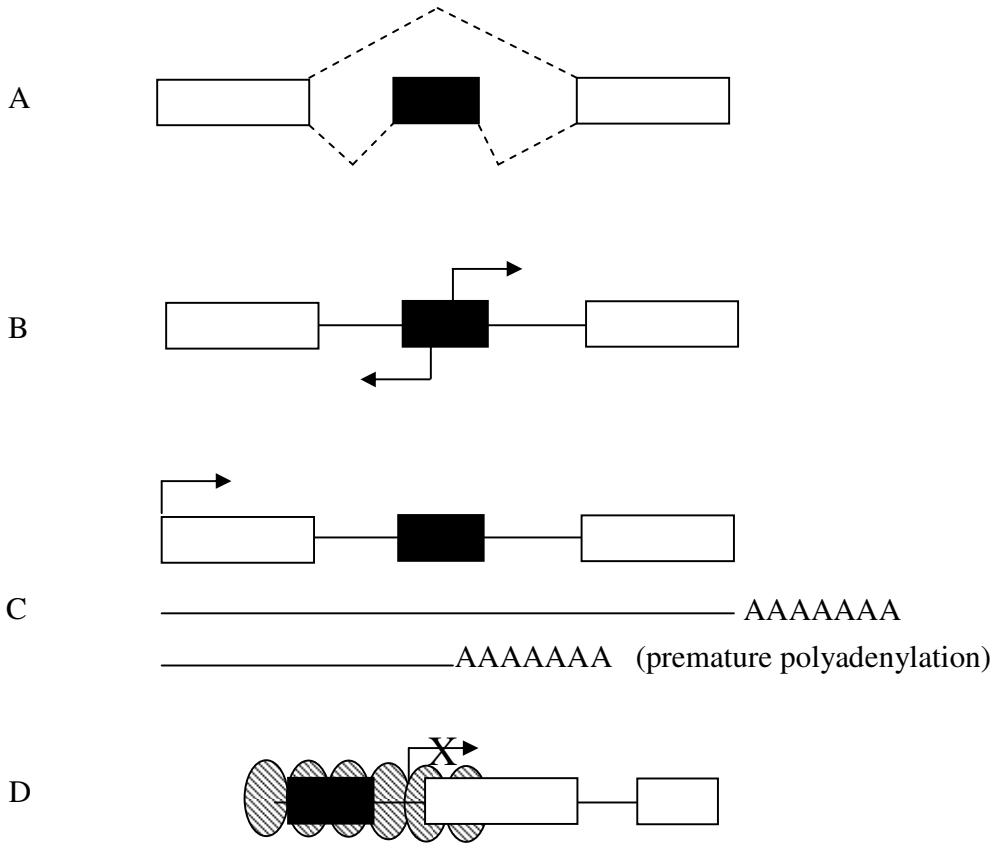


Figure 4. Impact of HERVs on human gene expression (edited from (19)) (A) HERVs sequence can be recruited as a new exon in a process termed exonization and associated with alternative splicing (dashed line). (B) LTR of HERV carried sense and antisense promoter activity (horizontal arrow) that can control the expression of nearby genes. (C) The presence of the HERV sequencing in the intron of a gene can cause of transcription elongation defects such as premature polyadenylation. (D) HERV sequence can be methylated. So, formation of heterochromatin (striped ovals) at HERVs sequence could spread and repress the transcription of nearby genes. Dark box represent a HERV sequence.

6. Role of HERVs in SLE

Several studies have been reported on association of HERV and autoimmune disease. Peripheral blood mononuclear cells (PBMCs) of SLE patients showed increased expression of HERV-E *gag* mRNA (31, 32, 105), whereas significantly up-regulation of HERV-K10 *gag* mRNA was found in rheumatoid arthritis patients when compared to normal controls (33) Furthermore, transcription of HERV-E *gag* gene correlated with blood plasma concentrations of anti-U1 ribonucleoprotein (RNP) and anti-Sm antibodies, are components of the same spliceosome complex and participate in splicing of pre-mRNA to mRNA, of SLE patients (32) Similar to other transposable elements, influences of DNA methylation to function of HERVs have been reported. HERV-E LTR-derived promoter of pleiotrophin, endothelin B receptor and midline 1 genes were different in methylation levels between placenta and blood cells. These methylation levels were associated with gene expression (21) Methylations of HERV-K LTRs in different loci were shown to have distinct pattern with inverse correlation with transcriptional activity (34) Furthermore, treating PBMCs and Tera-1 cell line (human testicular tumor) with 5-azacytidine showed that alteration of methylation level can increase expression of HERV-E *gag* transcript and HERV-K Gag protein, respectively (31, 35) These results suggest that loss of methylation in lymphocyte might result in the up-regulation of HERV expression in autoimmune patients. Interestingly, recent study showed that SLE patients have demethylation in HERV-E alternative promoter of CD5 gene and can lead to up-regulation of CD5-E1B isoform in SLE B cells (36) So far there is limited information about methylation profile of repetitive sequences especially the LTR retroelements in SLE. However, the hypomethylation of HERVs were observed in cancer cells (106-108) and in PBMC of aging process (109) In this study, we determined and compared the methylation status of HERV-E and HERV-K in normal and SLE lymphocyte subsets.

CHAPTER III

MATERIALS AND METHODS

Subjects

Twenty-three patients with SLE were recruited from King Chulalongkorn Memorial Hospital (Table 1). SLE patients with at least 4 of the American College of Rheumatology (ACR) criteria for the classification of SLE were enrolled into the study. Disease activity was assessed by the Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) (110) Inactive disease was defined as SLEDAI-2K score ≤ 6 whereas active disease had to have SLEDAI-2K score > 6 (111) Ten healthy female volunteers with no history of any autoimmune disease served as normal controls. Informed consent was obtained from each patient before entering the study. The trial was approved by the the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand.

Cell isolation

The peripheral blood mononuclear cells (PBMCs) were prepared by Ficoll density gradient centrifugation. In brief, heparinized blood was collected and diluted 1:1 in RPMI. The diluted was layered onto Ficoll-Paque (Invitrogen) and the gradients were centrifuged to obtain the mononuclear cell fraction, which was collected and washed twice with RPMI. PBMCs were counted on a hemocytometer. Then, B and T cell subsets were isolated from collected PBMCs by magnetic beads. First, B cells were isolated from PBMCs by positive selection using Dynabeads CD19 (Pan B) (Dynal Biotech; Invitrogen). After B cells were collected, negative flow through was used to isolate T cells by negative selection using Dynal T Cell Negative Isolation Kit Ver II (Dynal Biotech; Invitrogen). Cells from the second negative flow through were used to isolate CD4+ cells by positive selection using Dynal CD4 Positive Isolation Kit (Dynal Biotech; Invitrogen). Cells from the third negative flow through, unbound cells or non-CD4 T cells, were collected. The purity of the different cell

types was evaluated by flow cytometry. The purity of CD3+CD4+ T cells was more than 98%. Approximately 80% of the cells from the non-CD4 T cell fraction were CD3+CD8+ T cells. These cells were referred to as CD8+ T cells.

Normal oral epithelium (ORE) samples from 13 individuals were collected as previously described (20) Briefly, each subject used 20 ml of sterile 0.9% NaCl solution to rinse and gargle for 15 seconds then spit the solution into a sterile 50 ml tube. Cells from the oral rinses were collected by centrifugation at 2500g for 15 minutes. The supernatant was discarded and the pellets were washed with sterile 1X phosphate buffered saline (PBS).

Cell lines

Ten cell lines of carcinomas and lymphomas were used including cervical (HeLa, SiHa and CaSki) colorectal (RKO), hepatocellular (HepG2), larynx (HEp-2) and leukemia (K-562, Daudi, Jurkat and MOLT-4). Cells were cultured in Dulbecco's modified Eagle's medium or RPMI1640 (Gibco BRL, Life Technologies, Pairly, UK), supplemented with 10% heat-inactivated fetal bovine serum (Sigma, St Louis, MO, USA) and incubated at 37^{o}C in 5% CO_2.

DNA extraction

DNA was isolated from buffy coat collected with ethylenediaminetetraacetic acid (EDTA) as anticoagulant, using a salting out method (112) In brief, 1 ml of red cell lysis buffer (RCLB) was added to 0.5 ml of buffy coat, vortex for 30 seconds. This solution was centrifuged at 10,000-12,000 rpm for 30 seconds and the supernatant was discarded to obtain the pellet. The pellet remaining should be white to pink. This step may be repeated if necessary. To this pellet, 200 µl nuclei lysis buffer (NLB) and 50 µl 10% SDS were added. Pellet was broken up with pipette tip and vortex to get powdery, tiny flakes. The solution, 150 µl of NLB and 10 µl of proteinase K (10 mg/ml in H_2O stored frozen) were added, followed by incubation at 65 ºC for 2 hours. Precipitation of proteins was obtained by adding 175 µl of 5.3 M NaCl. This solution was centrifuged at 10,000-12,000 rpm for 15

minutes in micro centrifuge. After centrifugation, the DNA in the supernatant was precipitated in 1 ml of cold absolute ethanol. Invert 6-10 times to precipitate DNA, it will appear as a white to translucent stringy mass. This solution was centrifuged at 10,000-12,000 rpm for 10 minutes and the supernatant was discarded to obtain the pellet. This pellet was resuspend in 1 ml of cold 70% ethanol (break pellet by tapping), followed by centrifugation 1-2 minutes at 10,000-12,000 rpm and the supernatant was discarded to obtain the pellet. After removal of the ethanol, the pellet was dried at 37 ºC with the cap open to evaporate the ethanol, this pellet was dissolved in 200 µl of sterile distilled water, followed by incubation at 65 ºC for 15 minutes. Use gentle vortexing to resuspend. If clums of undissolved DNA are present, it will be incubated at 65 ºC until completely resuspended.

Combined bisulfite restriction analysis (COBRA)

Bisulfite conversion of DNA was performed using the EZ DNA methylation kit (Zymo Research), according to the manufacturer's protocol. Briefly, 20 µl of 500 ng of DNA was dissolved in 130 µl of the bisulfite solution and then was denatured in thermal cycler under the following conditions: 98 ºC for 10 min, 64 ºC for 2.5 hours and 4 ºC storage up to 20 hours. Next, the bisulfite-treated DNA was isolated using the Zymo-Spin™ IC Column. The DNA was eluted with 10 µl of M-Elution Buffer. Bisulfite-treated DNA was stored at -20 ºC until ready for use.

COBRA-HERV-E and COBRA-HERV-K

To compare the methylation of global HERV-E and HERV-K, bisulfited DNA were amplified by using primers sequence as follows: HERV-E forward 5'- TTT TGT TAG TTG ATG T(A/G)(G/T) GTA- 3', HERV-E reverse 5'- CCC CAA AAA AAA AAT TC(C/T) TAA CC-3' and HERV-K forward 5'-ATA TTA AGG GAA TTT AGA GGT TGG-3', HERV-K reverse 5'-CCC CTA CAC ACC TAT AAA TAT TTC-3'. We selected PCR primers that located at LTR of HERV-E and HERV-K sequences. PCR was performed under the following conditions: 95 ºC for 5 min, followed by 35 cycles of 1 min at 95 ºC for, 1 min at 58 ºC of HERV-E primer or 60 ºC of HERV-K primer, 1 min at 72 ºC, and finally 7 min at 72 ºC. HERV-E (126-bp) and

HERV-K (156-bp) PCR products were digested in 10 µl reaction volumes with 2 U of *Tai*I (Fermentas International Inc., Burlington, Canada). Each reaction was incubated overnight at 65 ºC and then electrophoresed on 8% non-denaturing polyacrylamide gels. The gel was stained with SYBR Green (Invitrogen), and the intensities of DNA fragments were measured on a PhosphoImager, using the ImageQuant software (Molecular Dynamics, GE Healthcare, Slough, UK). HERV-E positive methylated amplicons yielded 91, 82, 72 and 53bp fragments, whereas HERV-K positive methylated yielded 112 and 44 bp fragments. The methylation level was calculated as a percentage (the intensity of the digested methylated fragment divided by the sum of the undigested and digested-amplicons). To reduce inter-assay variation, we used bisulfite-treated DNA of cell line as a control in HERV-E assay (HeLa, Daudi and Jurkat) and HERV-K assay (K562, HepG2 and CaSki). Methylation levels of each cell line was standardized. Therefore, all experiments were adjusted with the same control methylation levels.

For COBRA for unique HERV-E LTR2C, locus-specific HERV were amplified by using one primer bind at gene specific sequences and another at LTR2C sequence. Primers for HERV-E LTR2C-APC2 were 5'-TAA AAG TAA GAT AAA TAG ATA A-3' and 5'-CCT TTT TAA AAA CTT ACA ACT-3'. Primers for HERV-E LTR2C-VAV2 were 5'-GGT TTT AGG AGT TAA AAA AAA G-3' and 5'-CCT TTT TAA AAA CTT ATA ACT-3'. PCR products of LTR2C-APC2 and LTR2C-VAV2 were digested with *Tai*I or *Taq*I (New England Biolabs, Beverly, MA), respectively.

COBRA-LINE-1 and COBRA-ALU

Global assessment of LINE-1 and Alu was performed by the combined bisulfite restriction analysis-interspersed repetitive sequences (COBRA-IRS) which is designed to detect methylation of thousands of CpG loci using one set of conserved primers in each IRS. The detection of methylation of IRSs in LINE-1 and Alu used in this study has been validated previously with pyrosequencing to be an accurate and reliable technique (22, 109) The primer sequences that correspond to the nucleotides in the regulatory region of the LINE-1 sequence (GenBank: M80343) are LINE-1 forward 5'-CCG TAA GGG CTT AGG

GAG TTT TT -3' and LINE-1 reverse 5'-(AG)TA AAA CCC TCC (AG)AA CCA AAT ATA AA 3'. The PCR reactions consisted of 35 cycles of 95 ºC for 1 min, 53 ºC for 1 min and 72 ºC for 1 min. The PCR products were subsequently digested with 2 U of *TaqI* (Fermentas International Inc., Burlington, Canada) and 2 U of *TasI* (Fermentas International Inc., Burlington, Canada) at 65 ºC overnight and were then run on an 8% nondenaturing polyacrylamide gel. The gel was stained with SyBr Green, and band intensities were measured by PhosphoImager using Image Quant software (Molecular Dynamics). The Alu primer sequences, which correspond to the nucleotides of the Alu Sx subfamily sequence (109), are Alu forward 5'-GG(T/C) G(C/T)G GTG GTT TA(C/T) GTT TGT AA-3' and Alu reverse 5'-CAC CAT ATT AACCAA ACT AAT CCC GA3'. Alu PCR conditions and restriction digestion conditions were similar to those of LINE-1. Gel electrophoresis, staining and band quantitative were measured as described in COBRA-HERV-E and HERV-K. COBRA LINE-1 and Alu amplicons were 160 and 99 bp, respectively. After digestion, the LINE-1 and Alu methylated bands were 80 and 57 bp, respectively. Unmethylated bands of LINE-1 were 97 and 63 bp, and the unmethylated band of Alu was 78 bp. Methylation levels were calculated as the intensity of methylated bands divided by sum of the methylated and unmethylated bands. We used Daudi, Jurkat and HeLa cell lines as a control to validate the inter-assay variation.

Cloning and Sequencing

Bisulfite DNA sequencing of HERV-E LTR2C-APC2, bisulfite-treated DNA was amplify by using primers 5'-AAG TAA GAT AAA TAG ATA ATT TTG G-3' and 5'-AAA ACT CCT AAC ACT AAA TTA AAA-3'. The PCR products were cloned into the pGEM-T easy vector (Promega, Santhan, UK). Ten independent clones were sequenced for each of the amplified fragments.

Pyrosequencing

Bisulfited DNA were amplified by using primers sequence as follows: HERV-E forward primer 5'-TTT TGT TAG TTG ATG T(A/G)(G/T) GTA-3', reverse biotinylated primer

5'-CCC CAA AAA AAA AAT TC(C/T)T AAC C-3' and HERV-K forward primer 5'-ATA TTA AGG GAA TTT AGA GGT TGG-3' reverse biotinylated primer 5'-CCC CTA CAC ACC TAT AAA TAT TTC-3'. Total 20 µl of PCR was performed under the following conditions: 95 ºC for 5 min, followed by 35 cycles of 1 min at 95 ºC for, 1 min at 58 ºC of HERV-E primer or 60 ºC of HERV-K primer, 1 min at 72 ºC, and finally 7 min at 72 ºC. Biotinylated PCR products were immobilized onto Streptavidin Sepharose HP beads (GE Healthcare). Biotinylated strands were selected by Vacuum Prep Workstation. Beads were released in the plate filled with 0.4 µM sequencing primer that is forward primer. The HERV-E and HERV-K methylation levels were quantitated by using the PSQ HS96 Pyrosequencing System. The ratio of C to T nucleotides was evaluated for HERV-E and HERV-K methylation. All values were represented by multiply the peak height methylated by 100 and dividing by sum peak height methylated and peak height non-methylated.

Real-time quantitative PCR

Total RNA was extracted from PBMCs by using the Trizol reagent (Life technologies, Inc.). DNA contamination was removed by RQ1 RNase-Free DNase (Promega, Southampton, UK) following the manufacturer's instructions. Then, all treated-DNase total RNA were reverse transcribed into cDNA as previously described (113) HERV-E clone 4-1 and HERV-K10 *gag* were amplified by using primers and protocols as previously reported (32, 114) The *gag* mRNA levels was quantitated relative to 18S rRNA transcripts (113) Real-time RT-PCR was performed in a Light Cycler machine (Roche Molecular Biochemicals, Indianapolis, IN, USA) using QuantiTect SYBR Green I (Qiagen, Hilden, Germany), according to the manufacturer's instructions. To ensure specific amplicons, PCR product melt curves were analyzed and visualized after electrophoresis in 2% agarose gels by ethidium bromide staining.

RNA sequencing (RNA-Seq)

PBMCs of healthy female donor were isolated by Ficoll density gradient centrifugation. Then, CD4+ T cells were isolated from collected PBMCs by negative CD4+

T cell isolation kits (EasySep, STEMCELL technologies). Where indicated, CD4+ T cells were stimulated for 20 hr with 10 μg/ml phytohemagglutinin (PHA), then untreated or treated with 1 μM 5-azacytidine (Sigma, St. Louis, MO) for an additional 72 hours as described (115, 116) These cells were used for the isolation of total RNA by RNeasy mini kits (Qiagen, Hilden, Germany). mRNA was purified by Dynabeads® mRNA Purification Kits (Dynal Biotech; Invitrogen) and fragmented by RNase III. SOLiD whole transcriptome libraries were performed according to manufactures protocols. Then, samples were purified and size selected on an agarose gel and subsequently extracted (Qiaquick gel extraction, Qiagen). cDNA libraries were amplified (15 cycles) by using SOLiD RNA barcoding kit before assess the yield and size distribution. Amplified libraries were quantified to estimate input requirements for emulsion PCR (emPCR). Titration points 0.5 ng was used to set-up the emPCR. Then, amplified beads were enriched and modified the 3' ends. Selected RNA libraries were then loaded onto quadrant scale slide on the AB SOLiD and sequenced according to manufactures protocols.

RNA-sequencing reads were analyzed by the BioScope software, whole transcriptome analysis pipeline. RPKM (Reads Per Kilobase of exon model per Million mapped reads) was used for normalization of expression and quantify transcript, which was fallowed the formula (117)

$$\text{RPKM} = \frac{10^9 \times \text{exon read counts}}{\text{Total read counts} \times \text{exon length}}$$

Differential expression of methylation sensitive genes in CD4+ T cells of SLE patients

The average RPKM per exon for each transcript to produce a mean RPKM for the gene were used to identify methylation sensitive CD4+ T cell genes. Then, up- or down regulated genes were identified by fold change between the two libraries. Connection Up- or Down- Regulation Expression Analysis of Microarrays (CU-DREAM) was used to identify methylation sensitive genes that differently expressed in CD4+ T cells of SLE patients (18, 118) Differential expression genes from RNA-seq were classified as up- or down-regulated

and not up- or not down-regulated to compare with expression microarray data available from Gene Expression Omnibus (GEO), a public repository that archives and freely distributes microarray data submitted by the scientific community (119, 120) The numbers of genes in each subset were compared using a chi-square test (figure 5). The significant *P*-value indicated the specific association between the two independent studies than random chance. The odds ratios > 1 indicate the regulated genes in one study were also regulated in another study. This result confirmed that the two studies involved a common gene regulation mechanism. On the contrary, odds ratios < 1 indicate that the event of interest (up or down) inhibit the mechanism (s) that altered gene expression in another experiment (118)

	Up- or Down-regulated genes of microarray experiment	**Not up- or Not down-regulated genes of microarray experiment**
Up- or Down-regulated genes of RNA-Seq experiment	Numbers of genes that regulated in both experiments	Numbers of genes that not regulated only in microarray experiment
Not up- or Not down-regulated genes of RNA-Seq experiment	Numbers of genes that not regulated only in RNA-Seq experiment	Numbers of genes not regulated in both experiments

Figure 5. 2 x 2 table of chi-square test.

Functional analysis

Gene Functional classification analysis was performed using DAVID Bioinformatics Resources (http://david.abcc.ncifcrf.gov/home.jsp) (121, 122) Gene list was performed with DAVID functional annotation chart to get EASE scores, a modified Fisher Exact *P*-Value, for each enriched annotation terms. Following, calculate geometric mean of EASE scores of those terms involved in this gene group. Minus log transformation is applied on the geometric mean to emphasize that the geometric mean is a relative score instead of an absolute *P* value. For example, enrichment scores 0.05 is equivalent to 1.3 in minus log scale. Ranks the biological significant of gene groups are based on overall EASE scores (enrichment score) of all enriched annotation terms. Therefore, a higher enrichment score for a group indicates that the group members are involved in more important roles (121) Furthermore, pathway analysis was performed by Ingenuity Pathway Analysis (IPA) system (http://www.ingenuity.com/index.html).

Database searches and sequence analyses

Analysis of HERV-E and HERV-K sequences were performed by using an *in silico* PCR of BiSearch with bisulfite converted-human genome DNA (123, 124), UCSC Genome Browser (125), blast 2 sequences alignment, and RepeatMasker (126)

Statistical analyses

Statistical significance was determined by applying an independent or dependent sample t-test. The Mann-Whitney U test was used to make non-parametric comparisons of HERV-E clone 4-1 *gag* and HERV-K10 *gag* expression between normal and SLE group. Pearson's correlation coefficient was used to examine the relationship between two continuous variables. All analyses were performed using SPSS, version 11.5 (SPSS Inc., Chicago, IL).

Table 3. Patient demographics and medications

Patient	Group	Age	Sex	SLEDAI-2K	Medications
1	active	35	F	12	Oral Prednisolone 20 mg/day, Cellcept 1.5 g/day
2	inactive	35	F	4	Oral Prednisolone 5 mg/day, Cellcept 500 g/day
3	active	33	F	8	Oral Prednisolone 25 mg/day, Endoxan 50 mg/day
4	active	33	F	14	Oral Prednisolone 15 mg/day
5	active	33	F	10	Imuran 75 mg/day
6	active	38	F	17	Oral Prednisolone 5 mg/day
7	active	30	F	20	Oral Prednisolone 10 mg/day, Cellcept 1.5 g/day
8	active	24	F	20	Oral Prednisolone 10 mg/day
9	active	25	F	10	Oral Prednisolone 10 mg/day
10	active	36	F	8	Oral Prednisolone 7.5 mg/day, Imuran 50 mg/day
11	inactive	30	F	0	Oral Prednisolone 10 mg/day
12	inactive	37	F	0	Oral Prednisolone 2.5 mg/day
13	inactive	16	F	6	Oral Prednisolone 10 mg/day
14	inactive	27	F	5	Oral Prednisolone 10 mg/day
15	active	41	F	9	Oral Prednisolone 10 mg/day
16	active	37	M	12	Oral Prednisolone 15 mg/day, Endoxan 1400 mg IV x 6 monthly
17	active	15	F	16	Endoxan 1000 mg IV x 6 monthly
18	active	42	F	16	Oral Prednisolone 5 mg/day
19	inactive	27	F	4	Oral Prednisolone 10 mg/day
20	inactive	37	F	0	None
21	active	32	F	9	None
22	active	27	F	16	None
23	active	20	F	18	Oral Prednisolone 15 mg/day

CHAPTER IV

RESULTS

Analysis of HERV-E clone 4-1 and HERV-K10 *gag* expression in PBMCs of SLE patients and healthy individuals

This study was performed to investigate whether expression of HERV-E clone 4-1 and HERV-K10 sequences differ among PBMCs of SLE and healthy groups. All primers were located in *gag* of HERV-E clone 4-1 and HERV-K10 sequences as published by Piotrowski et al. and Depil et al., respectively (32, 114) Expression was measured as of HERV-*gag* mRNA relative to 18S rRNA (Figure 6). As previously reported (31, 32), HERV-E clone 4-1 *gag* transcripts in PBMCs of SLE patients (n = 10) were significantly increased when compared with normal individuals (n = 11) (Figure 6A, P = 0.02). However, there was no significant difference in the HERV-K10 *gag* expression between the two groups (Figure 6B, P = 0.38). No significant correlations between expression levels of each HERV *gag* and with disease activity were observed (Figure 7 and 8).

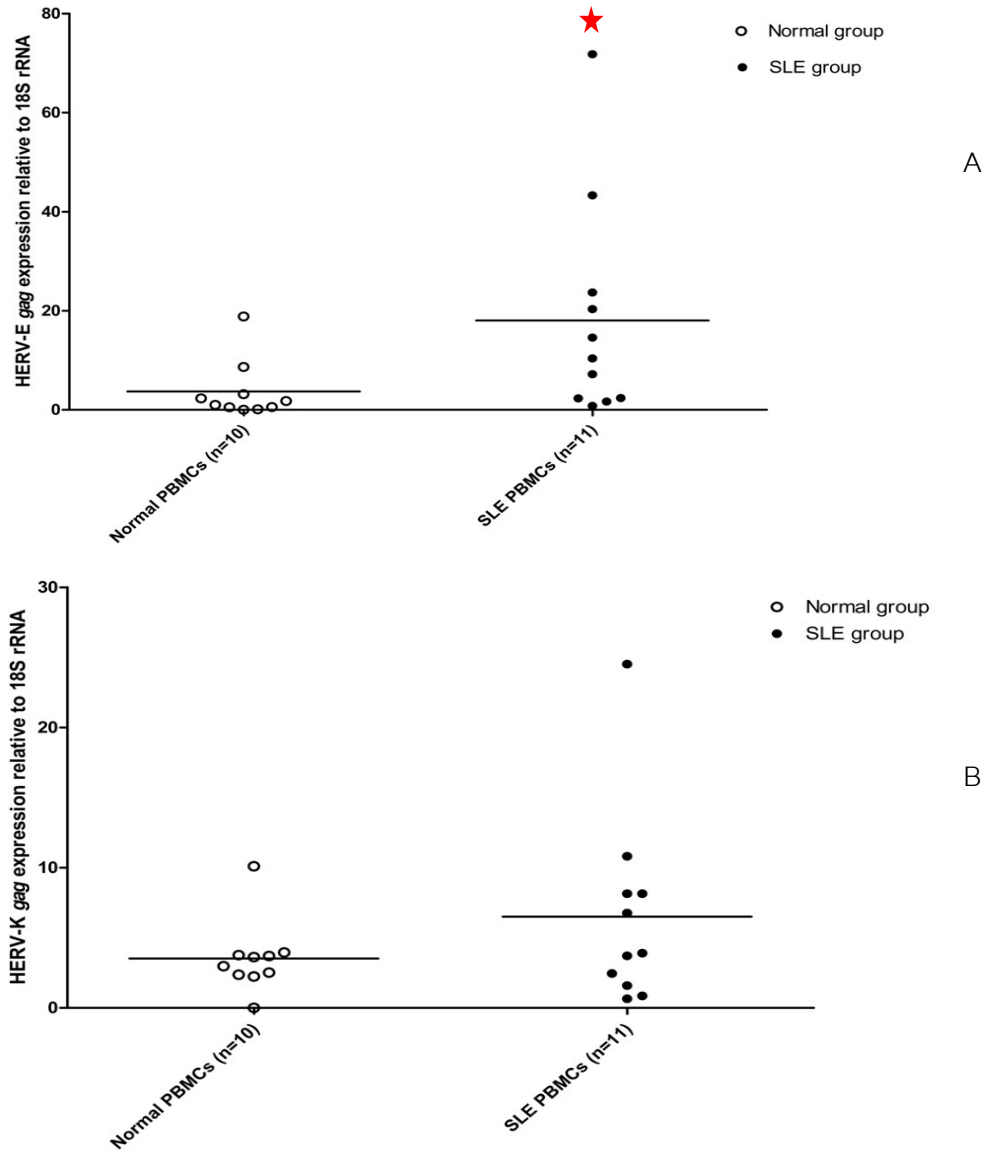


Figure 6. Expression of HERV-E and HERV-K *gag* sequences in normal and SLE PBMCs. The *gag* mRNA levels was quantitated relative to 18S rRNA transcripts. (A) HERV-E 4-1 *gag* transcripts was significant at $P = 0.02$ and (B) HERV-K10 *gag* transcripts was not significant ($P = 0.38$). Numbers in the parentheses indicate the number of samples. Each symbol represents an individual sample and horizontal lines show mean values. *P*-value calculated by Mann–Whitney test.

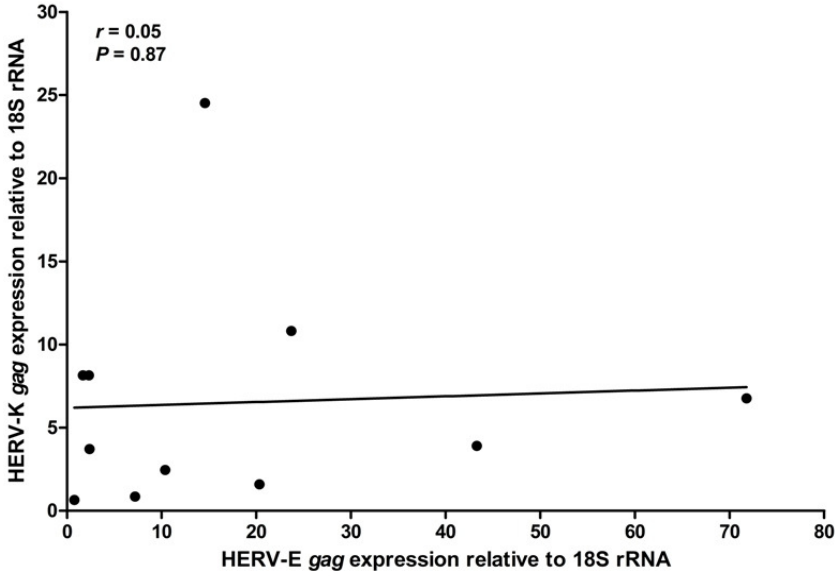


Figure 7. Correlation between expression levels of HERV-E *gag* and HERV-K *gag*. Each dot represents an individual patient. *P*-value calculated by Pearson's correlation coefficient test.

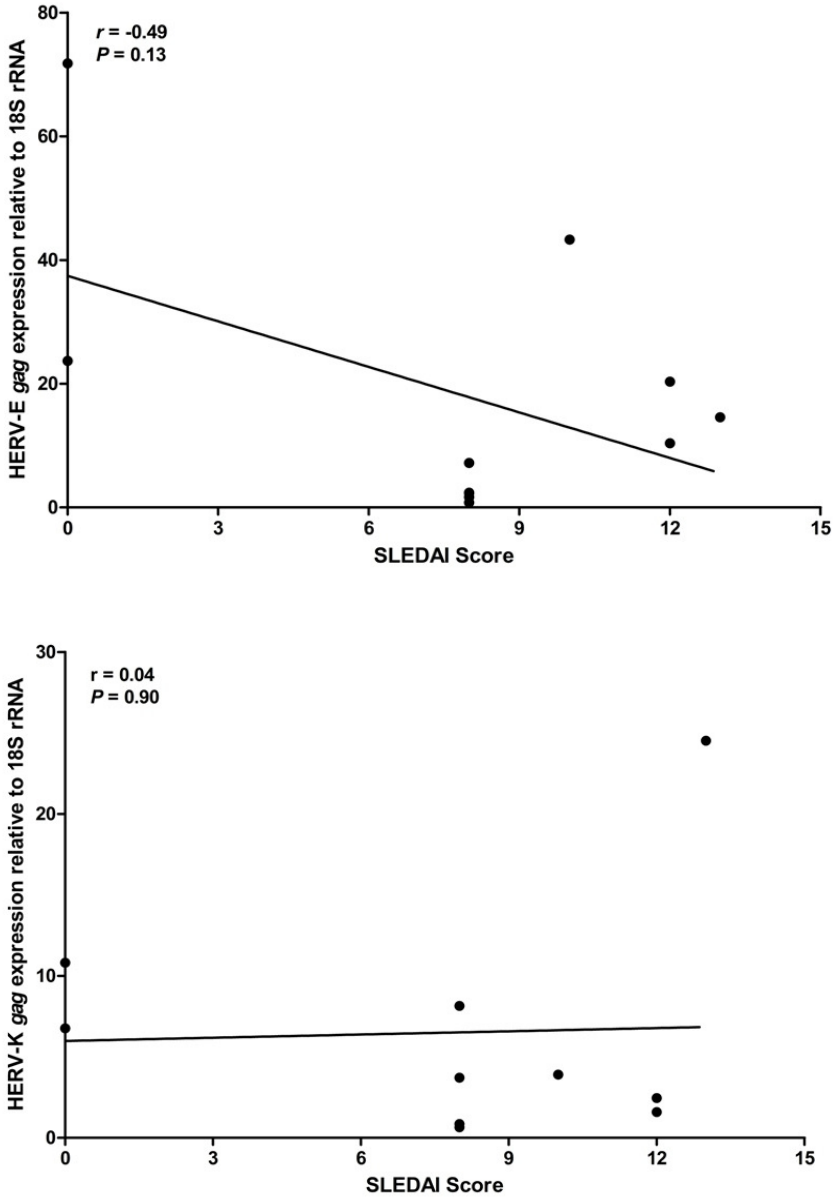


Figure 8. Correlations between expression levels of each HERV *gag* and with disease activity. The relationship of HERV-E *gag* (A) and HERV-K *gag* (B) expression levels in PBMCs with SLEDAI score of SLE patients were shown.

COBRA- long terminal repeats (LTRs) Design and Set Up

The primer design for HERV in this study is based on the HERV-E clone 4-1 and HERV-K10 *gag* transcripts aligned in Genbank accession no. M10976 and no.M14123 sequence, respectively. To identify LTR subclass, HERV sequences were analyzed by Repbase nomenclature (126) Subclass of HERV-E clone 4-1(M10976) and HERV-K10 (M14123) are LTR2C and LTR5_Hs, respectively. To study the global DNA methylation of LTR2C and LTR5_Hs, we selected PCR primers that located at LTR2C and LTR5_Hs sequences for COBRA (Figure 9). A total of 265 (LTR2C) and 618 (LTR5_Hs) were mapped in human genome by using Table Browser of UCSC Browser (125) Primers were predicted to amplify in bisulfite converted-human genome DNA by using an *in silico* PCR of BiSearch (123, 124) A total of 77 (HERV-E LTR2C) and 283 (HERV-K LTR5_Hs) were predicted approximately 29% and 45% of the total, respectively. From LTR2C-COBRA, multiple bands result from variation of restriction site in amplicons presumably due to low identity (80%) between HERV-E copies (21, 127) While, low levels of sequence divergence (1-2%) in HERV-K sequence was reported (21, 128) To assess the reproducibility of each primer pair, we analyzed the methylation levels of HERV-LTRs in triplicate samples using different lymphocyte subsets and cancer cell lines. All cell lines were cultured at least three flasks in order to replicate in COBRA assays and compared among same cell type. Our results showed limited result deviations (Figure 10).

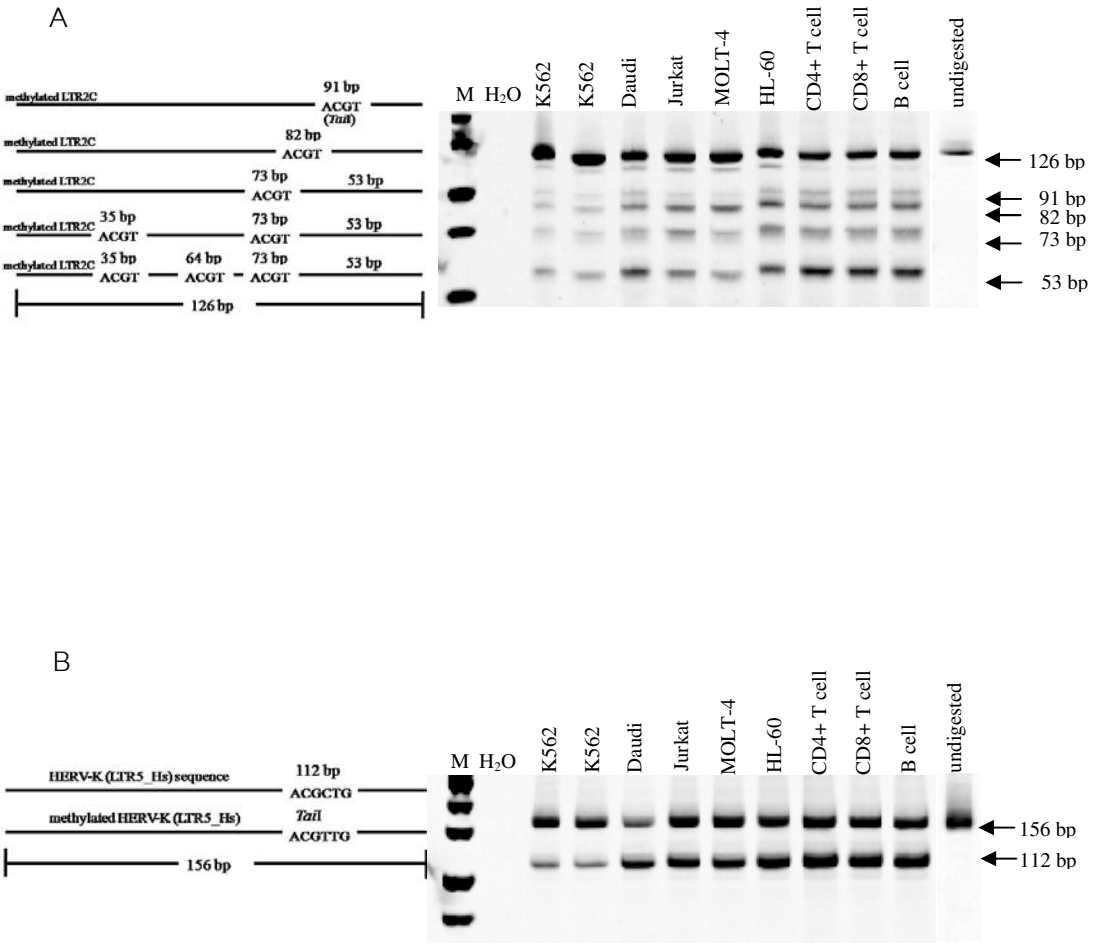
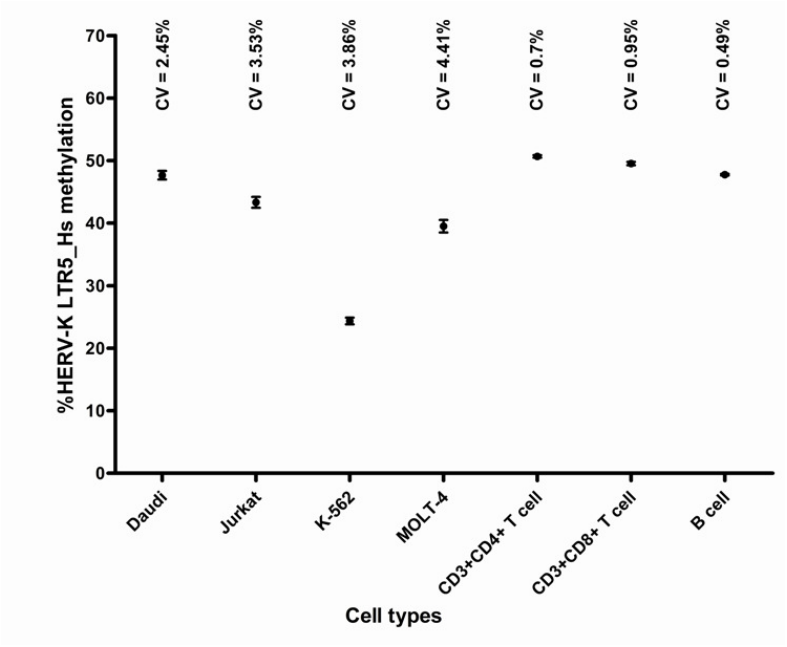
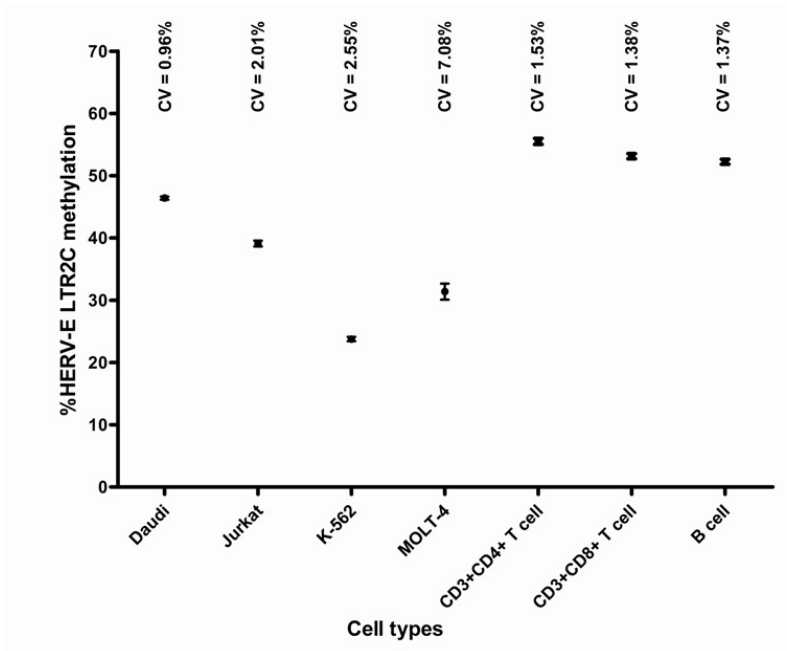


Figure 9. COBRA for HERV-E LTR2C and HERV-K LTR5_Hs. (A) LTR2C and (B) LTR5_Hs amplicon sizes are 126 and 156 bp, respectively. The ACGT sequence were presented after sodium bisulfite treatment and then were digested by *Tai*I, LTR2C methylated sequences yielded 91, 82, 73 and 53 bp fragments, whereas LTR5_Hs methylated yielded 112 and 44 bp (not detected) fragments. The methylation level was calculated as a percentage of the intensity of the digested methylated fragment divided by the sum of the undigested and digested-amplicons. M is a standard size marker. The negative control is water (H_2O).



A



B

Figure 10. Mean (±SEM) methylation levels of (A) HERV-E LTR2C and (B) HERV-K LTR5_Hs in different cell types. The coefficients of variance (%CV) of the measurement from triplicates of each cell type were determined.

The correlation between COBRA methylation levels and mean methylation levels of all CpG by pyrosequencing

We have validated the COBRA method by comparing it to pyrosequencing technique. Pyrosequencing analysis was performed in 48 samples that were chosen by sampling from lymphocyte subsets of 5 normal controls and 10 SLE patients and 3 control-cell lines (Figure 11 and 12).

Methylation levels of 5 CpG sites of HERV-E and 3 CpG sites of HERV-K were measured by Pyrosequencing. The methylation levels of each CpG nucleotides, including COBRA *Tai*I representative nucleotides of HERV-E and HERV-K, were commonly in direct correlation with the mean methylation levels. For HERV-E, there are positive correlations between methylation levels of means and each CpG dinucleotides ($r = 0.92$, $P = 1.56 \times 10^{-19}$ for CpG1, $r = 0.77$, $P = 1.48 \times 10^{-10}$ for CpG2, $r = 0.56$ $P = 3.5 \times 10^{-5}$ for CpG3, $r = 0.55$, $P = 6.7 \times 10^{-5}$ for CpG4 and $r = 0.49$, $P = 4.1 \times 10^{-4}$ for CpG5) (Figure 13A-E). CpG1, CpG2, CpG3 and CpG5 are *Tai*I CpG sites of COBRA HERV-E LTR2C. Also in HERV-K, positive correlations of methylation levels of means and each CpG dinucleotides were observed ($r = 0.79$, $P = 1.57 \times 10^{-11}$ for CpG1, $r = 0.67$, $P = 2.06 \times 10^{-7}$ for CpG2 and $r = 0.64$ $P = 1.07 \times 10^{-6}$ for CpG3 which is *Tai*I CpG site of COBRA HERV-K LTR5_Hs) (Figure 14 A-C). Moreover, there are positive correlations between methylation levels of means from pyrosequencing with COBRA HERV-E LTR2C ($r = 0.45$, $P = 0.0015$) (Figure 13F) and HERV-K LTR5-Hs ($r = 0.45$, $P = 0.0003$) (Figure 14D).

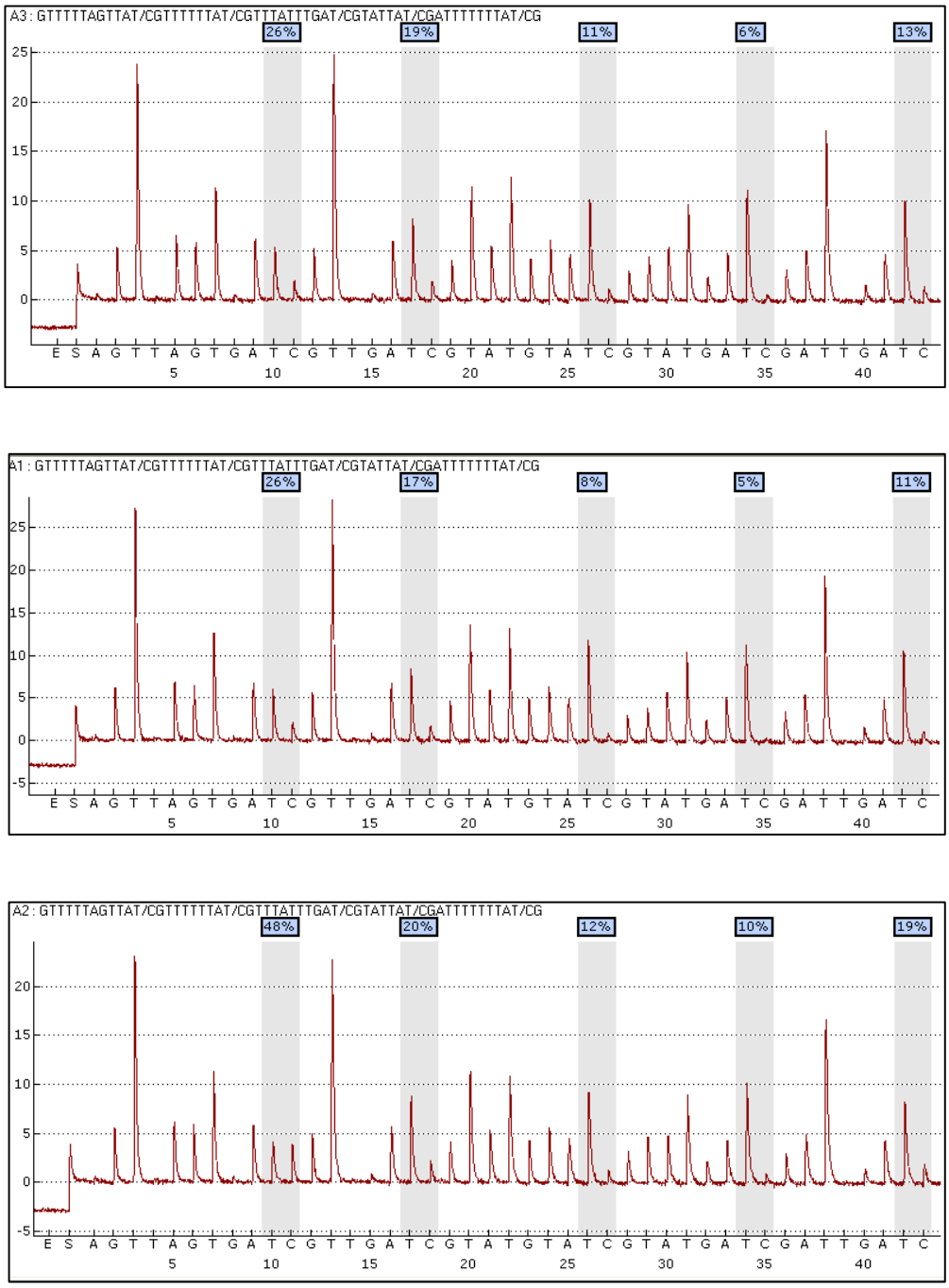


Figure 11. The representative of pyrogram of HERV-E LTR2C. A) Jurkat (15%) B) HeLa (13%) C) Daudi (22%) D) Normal CD3+CD4+ T cell (19%) and E) Normal CD3+CD4+ T cell (20%). Mean of methylation was shown in parentheses.

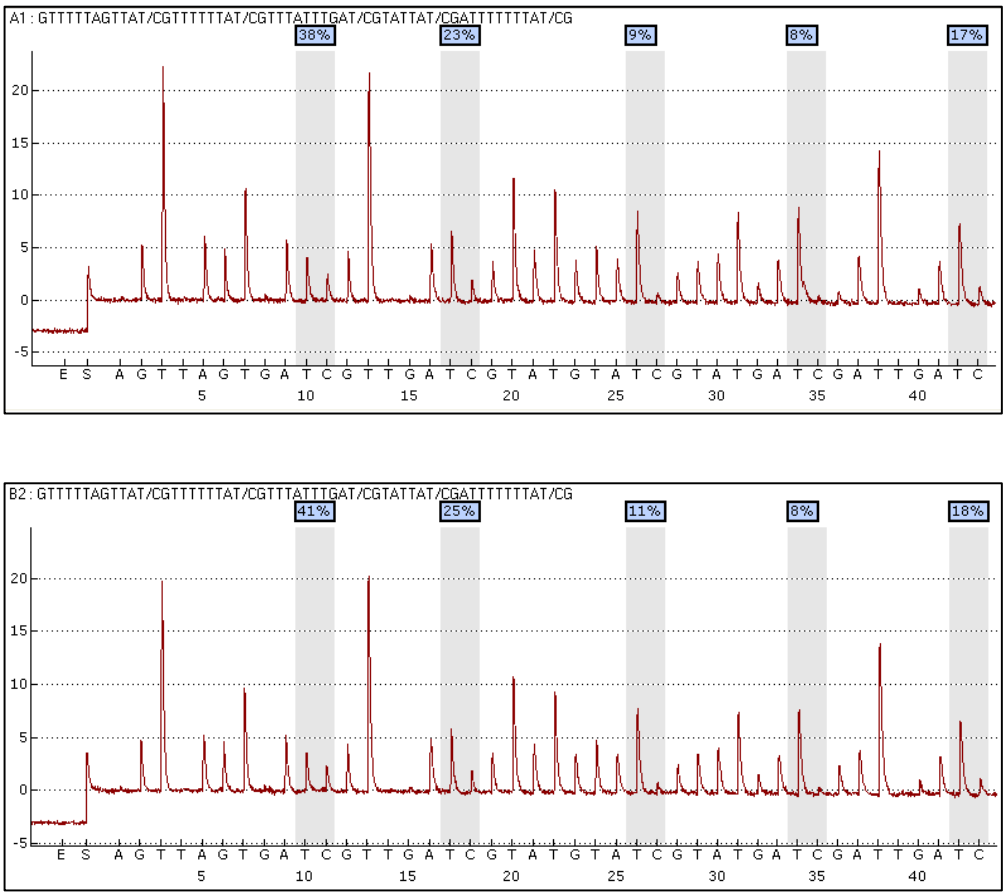


Figure 11. The representative of HERV-E LTR2C pyrogram. (Continue)

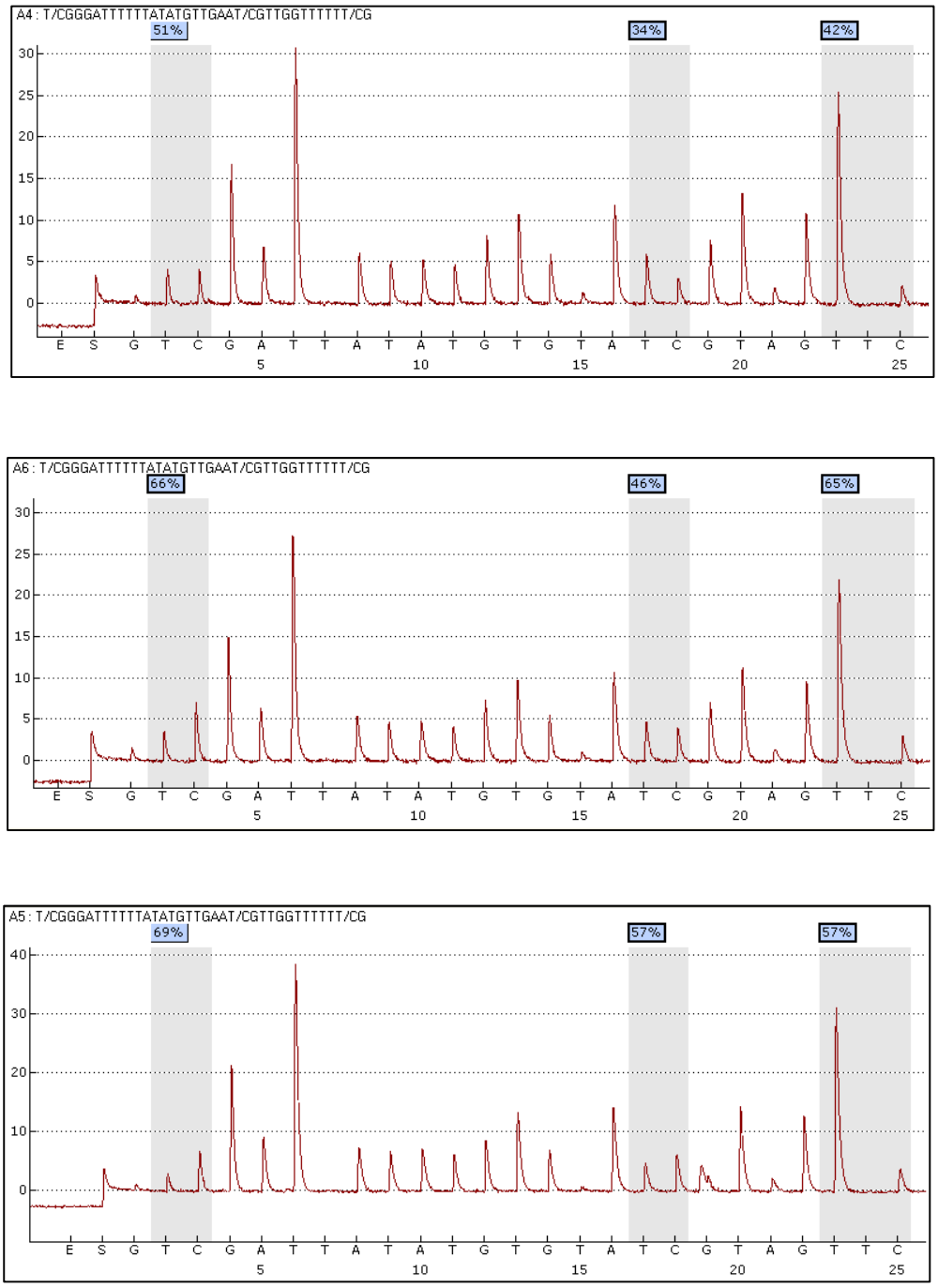
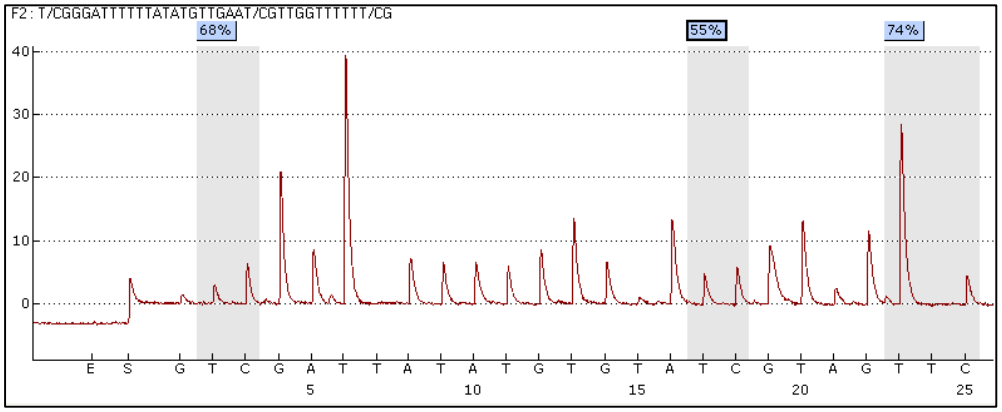
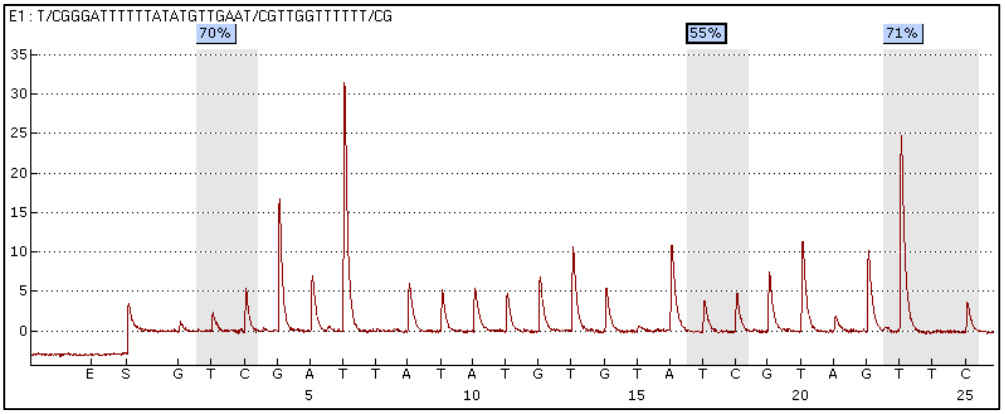


Figure 12. The representative of pyrogram of HERV-K LTR5_Hs. A) K562 (42%) B) HepG2 (59%) C) CaSki (61%) D) Jurkat (66%) and E) Normal CD3+CD4+ T cell (65%). Mean of methylation was shown in parentheses.

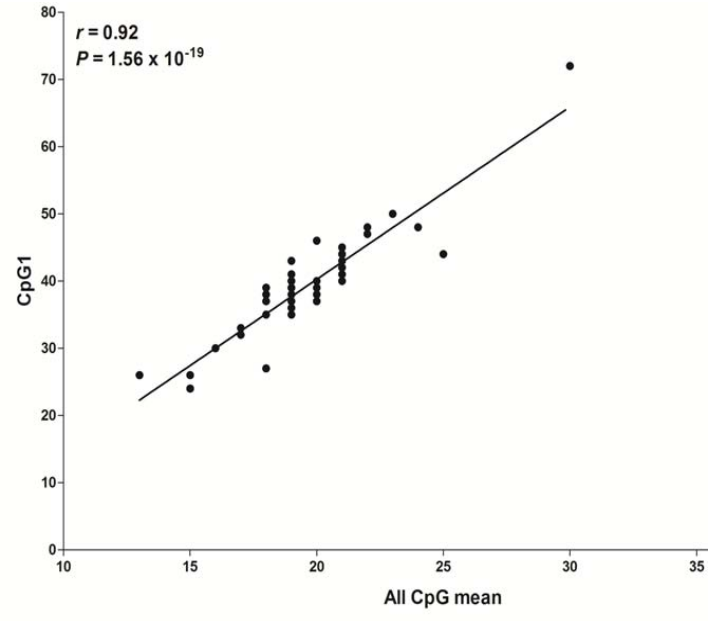


D

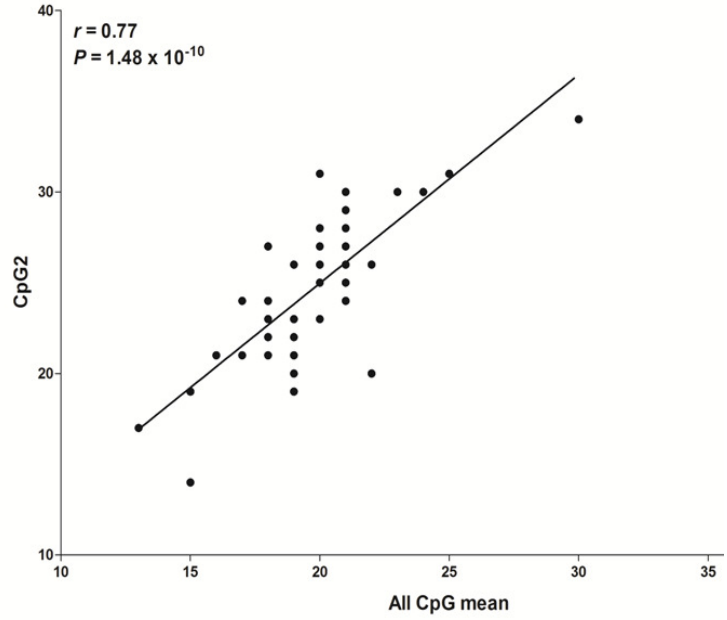


E

Figure 12. The representative of HERV-K LTR5_Hs pyrogram. (Continue)



A



B

Figure 13. The correlation between mean methylation levels of all CpG by pyrosequencing with methylation levels of each CpG nucleotide (A-E) and COBRA HERV-E LTR2C (F). A, B, C and E are *Tai*I COBRA HERV-E LTR2C sites and means methylation levels. Each dot represents an individual sample. *P*-value calculated by Pearson's correlation coefficient test.

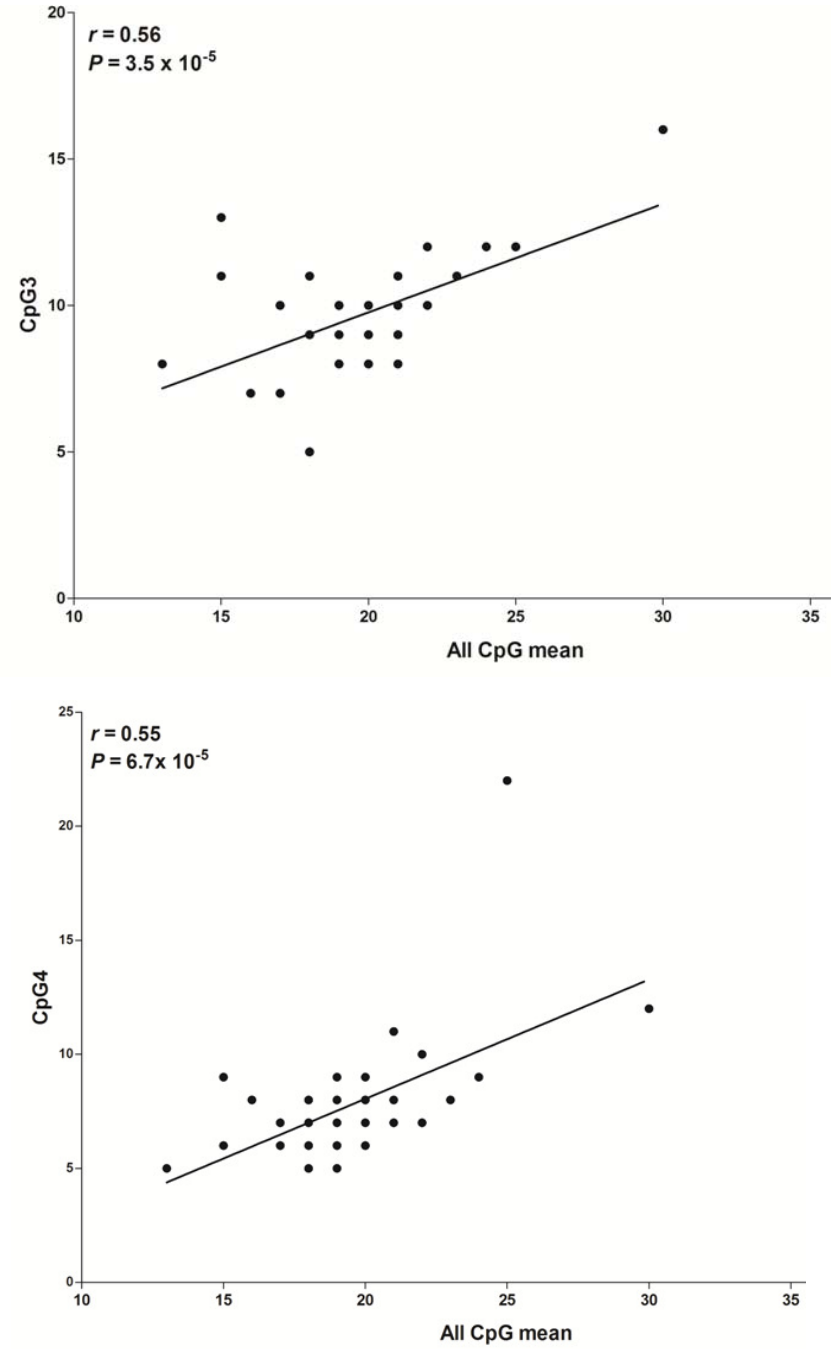
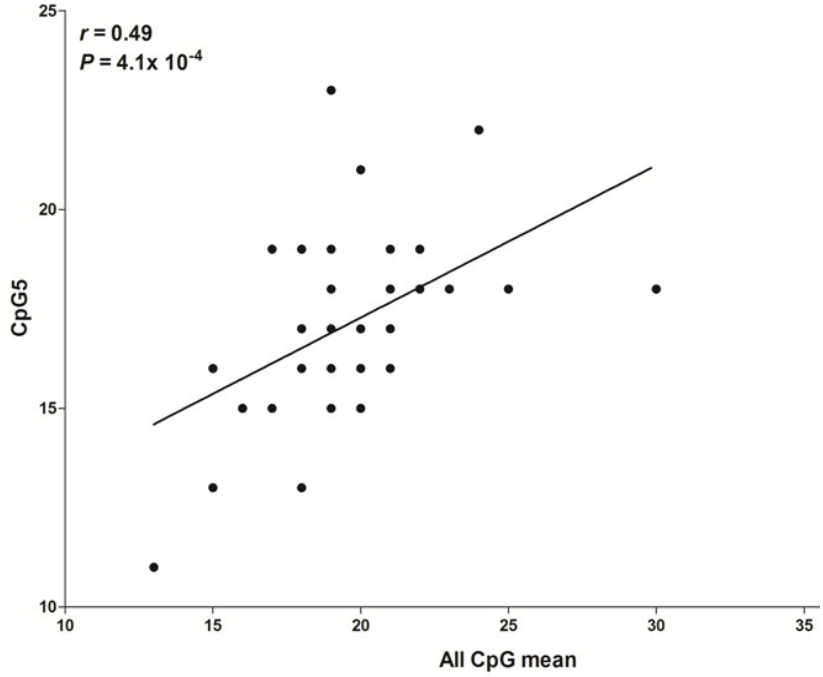
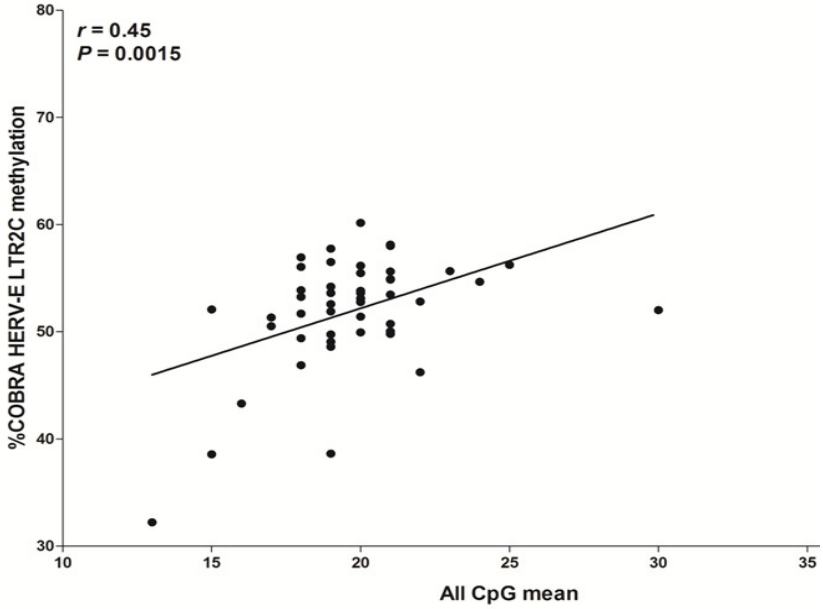


Figure 13. The correlation between mean methylation levels of all CpG by pyrosequencing with methylation levels of each CpG nucleotide (A-E) and COBRA HERV-E LTR2C (F). (Continue)



E



F

Figure 13. The correlation between mean methylation levels of all CpG by pyrosequencing with methylation levels of each CpG nucleotide (A-E) and COBRA HERV-E LTR2C (F). (Continue)

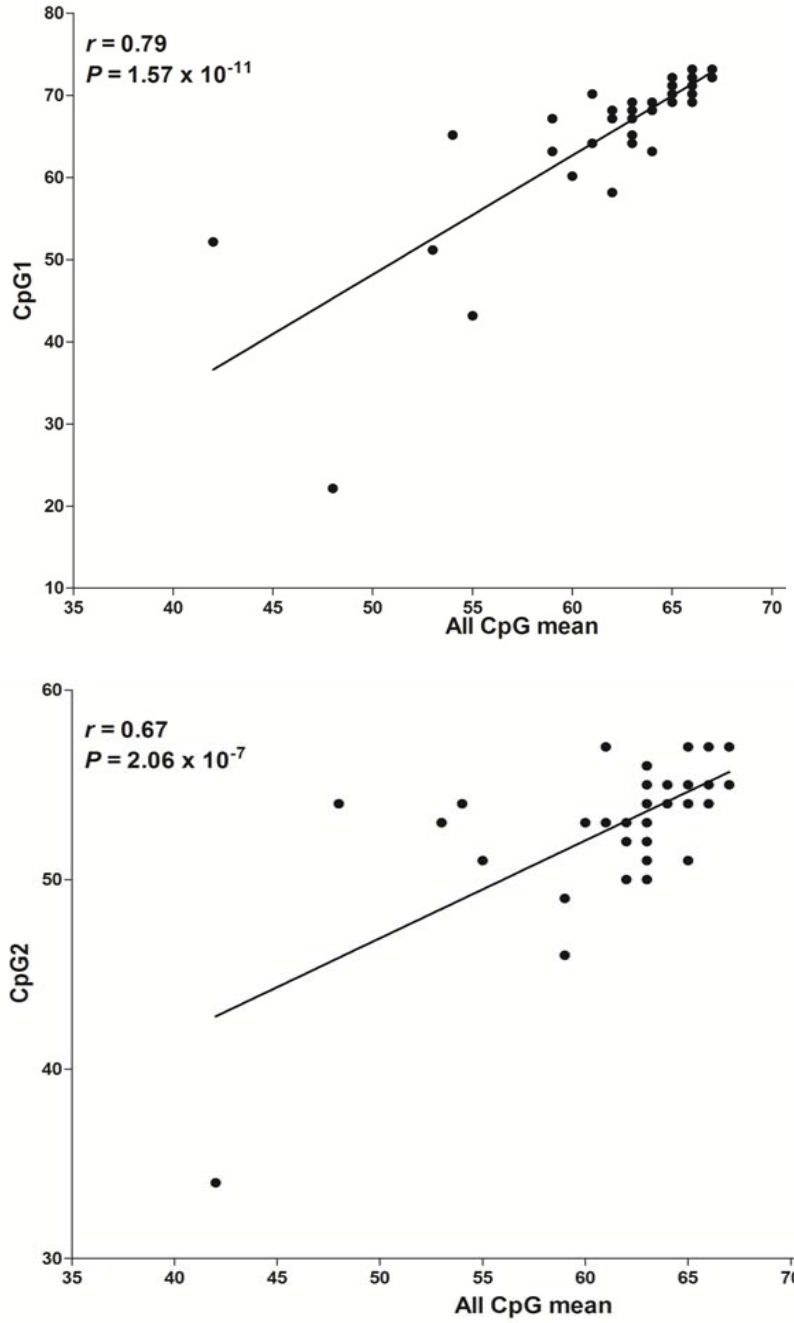
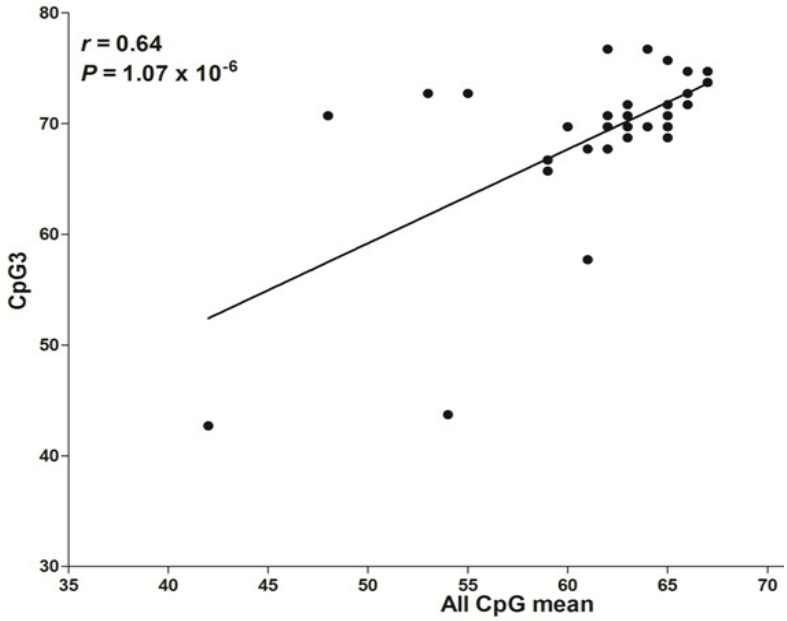
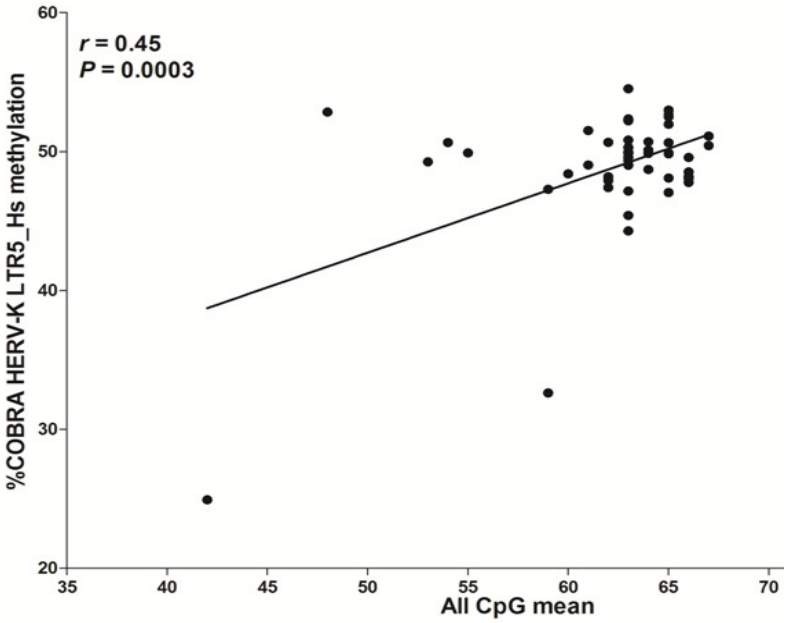


Figure 14. The correlation between mean methylation levels of all CpG by pyrosequencing with methylation levels of each CpG nucleotide (A-C) and COBRA HERV-K LTR5_Hs (D). B is *Tai*I COBRA HERV-K LTR5-Hs sites and means methylation levels. Each dot represents an individual sample. *P*-value calculated by Pearson's correlation coefficient test.



C



D

Figure 14. The correlation between mean methylation levels of all CpG by pyrosequencing with methylation levels of each CpG nucleotide (A-C) and COBRA HERV-K LTR5_Hs (D). (Continue)

Cell type-specific of DNA methylation at HERV-LTR

This study was designed to examine whether methylation levels of HERV-E LTR2C and HERV-K LTR5_Hs differ among cell types. Comparison of the methylation levels showed clearly that both of LTRs were difference in various cell types (Figure 15).

Interestingly, hypomethylation of HeLa cell (32.66±1.27% and 34.31±1.57% in HERV-E LTR2C and HERV-K LTR5_Hs, respectively) were obtained in HERV-E LTR2C and HERV-K LTR5_Hs when compared with other cervical cancer cell lines, which SiHa (45.45±2.15% and 46.49±0.15% in HERV-E LTR2C and HERV-K LTR5_Hs, respectively) and CaSki (43.17±0.93% and 48.46±1.29% in HERV-E LTR2C and HERV-K LTR5_Hs, respectively).

In lymphoma cell lines, Daudi, which is B cell line, have hypermethylation in HERV-E LTR2C and HERV-K LTR5_Hs (46.21±1.01% and 48.36±1.07%, respectively) when compared to two cell lines that have been derived from T cell, Jurkat (39.11±0.78% and 43.33±1.53%) and MOLT-4 (31.38±2.22% and 39.51±1.74%). Furthermore, HERV-E LTR2C and HERV-K LTR5_Hs Methylation in K-562 cell line (22.46±1.55% and 25.82±1.81%) were lower than other leukemic cells.

Lymphocyte subsets were isolated from 10 females healthy PBMCs (mean of age 24.4±2.1). Interestingly, we found that methylation levels of HERV-E LTR2C in B cells and CD3+CD8+ T cells were significantly decreased compared to CD3+CD4+ T cells (P = 0.0082 and P = 0.012, respectively) (Figure 15A). Whereas, the decrease of methylation levels of HERV-K LTR5_Hs in B cells was found when compared with CD3+CD4+ (P = 0.027) (Figure 15B). Besides the study in lymphocytes, oral epithelium (ORE) samples were collected from 11 females and 2 males (mean of age 28.15±4.49), 5 samples were collected from the same blood donors. Comparison of LTRs Methylation in ORE with lymphocytes, HERV-K LTR5_Hs methylation levels of normal ORE decreased when compared with CD3+CD4+ T cells (P = 0.0002) (Figure 15B), but there was no significant difference in HERV-E LTR2C methylation level between ORE and lymphocyte subsets (Figure 15A). The results indicated that methylation levels of HERV-LTRs are cell type specific manner.

Comparison of methylation levels between lymphoma cell lines and normal cells

Interestingly, Daudi showed specifically hypomethylation of HERV-LTRs. This cell was derived from Burkitt's lymphoma and characterized B lymphoblast cell. We detected significantly different of HERV-E LTR2C hypomethylation in Daudi cell compared to normal B cell (P = 0.0001, Figure 15A), but was not different in HERV-K LTR5_Hs methylation (P = 0.44) (Figure 15B). Acute T lymphoblastic leukemia cells (Jurkat and MOLT-4) were hypomethylation in both of LTRs when compared with normal CD3+CD4+ and CD4+CD8+ T cell subsets (P<0.0003) (Figure 15A and 15B). K-562 cell, that was derived from chronic myelogenous leukemia patient and characterized as undifferentiated granulocyte, showed least of methylation levels in both of LTRs when comparing with normal lymphocytes and other leukemic cells (P<0.0001) (Figure 15A and 15B). .

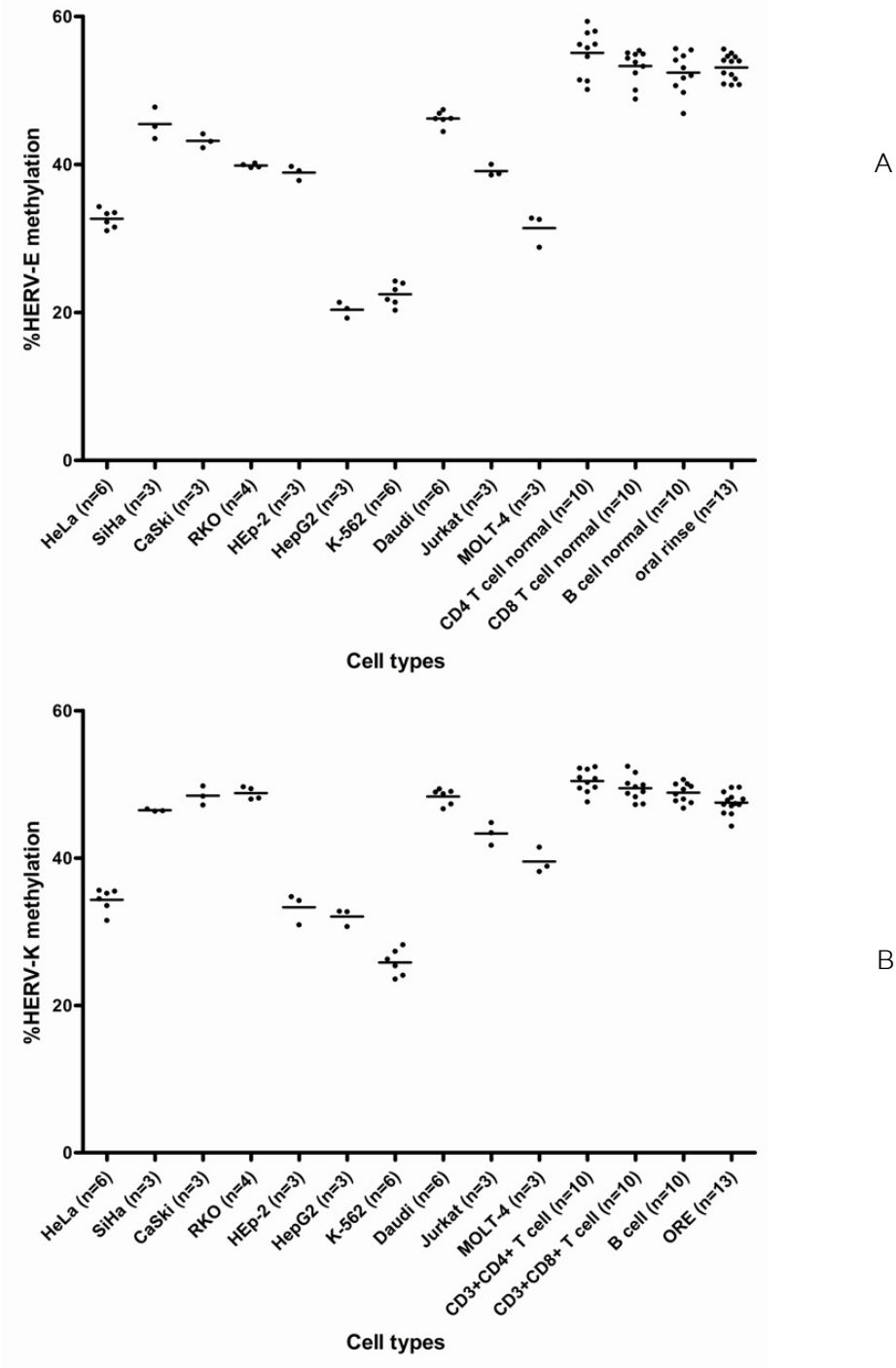
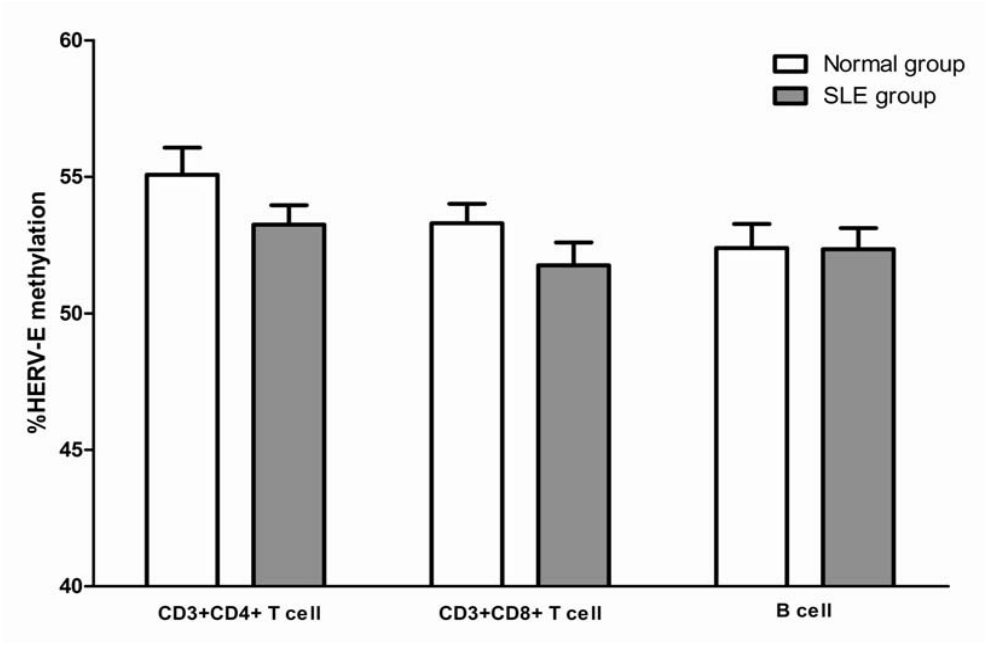


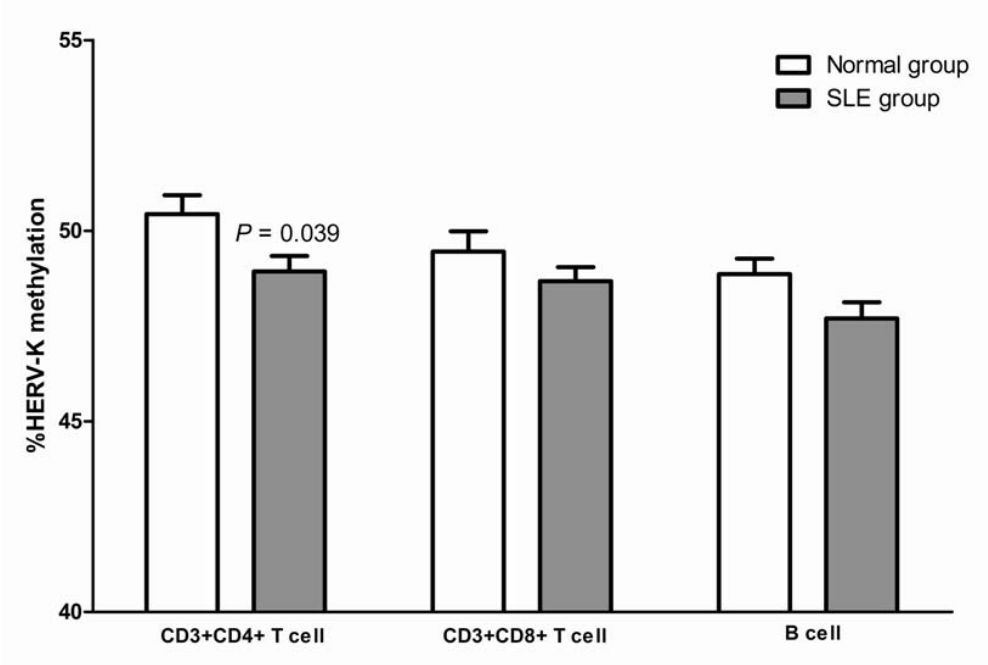
Figure 15. DNA Methylation levels of HERV-E LTR2C (A) and HERV-K LTR5_Hs (B) in different cell types. Numbers in the parentheses indicate the number of samples. Each symbol represents an individual sample and horizontal lines show mean values.

Differences in DNA methylation levels for HERV-E LTR2C, HERV-K LTR5_Hs, LINE-1 and Alu between CD4+, CD8+ T cells and B cells from SLE patients versus normal controls

DNA methylation levels of HERV-E LTR2C and HERV-K LTR5_Hs were compared among SLE and normal individuals. We found that CD3+CD4+ T cell of SLE patients have hypomethylation of HERV-K LTR5_Hs in comparing with normal group (P = 0.039) (Figure 16B). However, there was no difference of HERV-E LTR2C methylation in lymphocytes subsets between two groups (Figure 16A). DNA methylation of LINE1 and Alu, non-LTR Retrotransposon, were also examined. This study showed that SLE patients have hypomethylation of LINE-1 in CD4+, CD8+ T cells, and B cells when compared to normal controls (P = 0.005, 0.002, and 0.007, respectively) (Figure 16C). However, there were no differences in methylation profile for Alu between SLE patients and the normal controls for any lymphocyte subsets (Figure 16D).

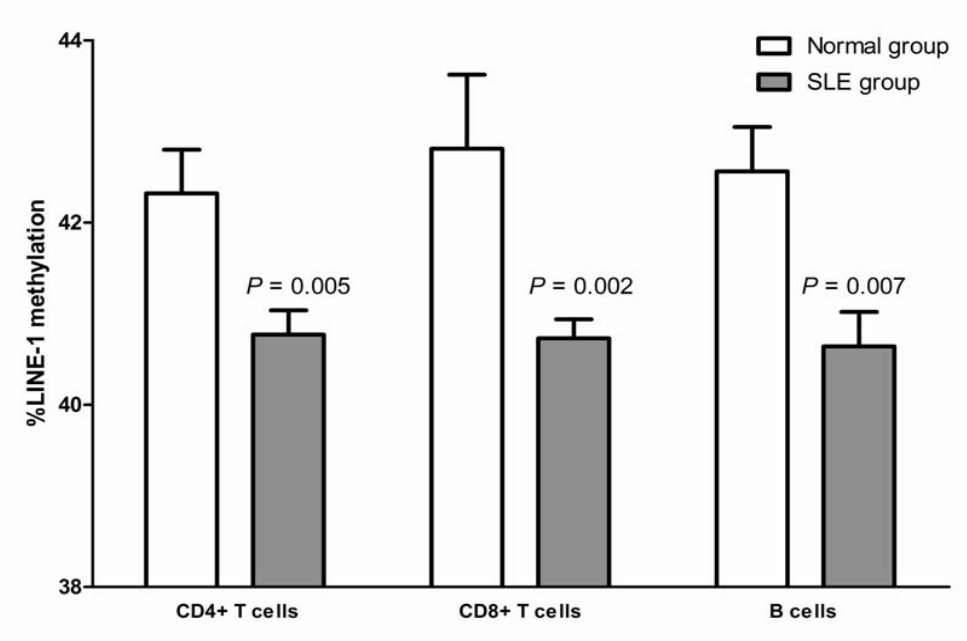


A

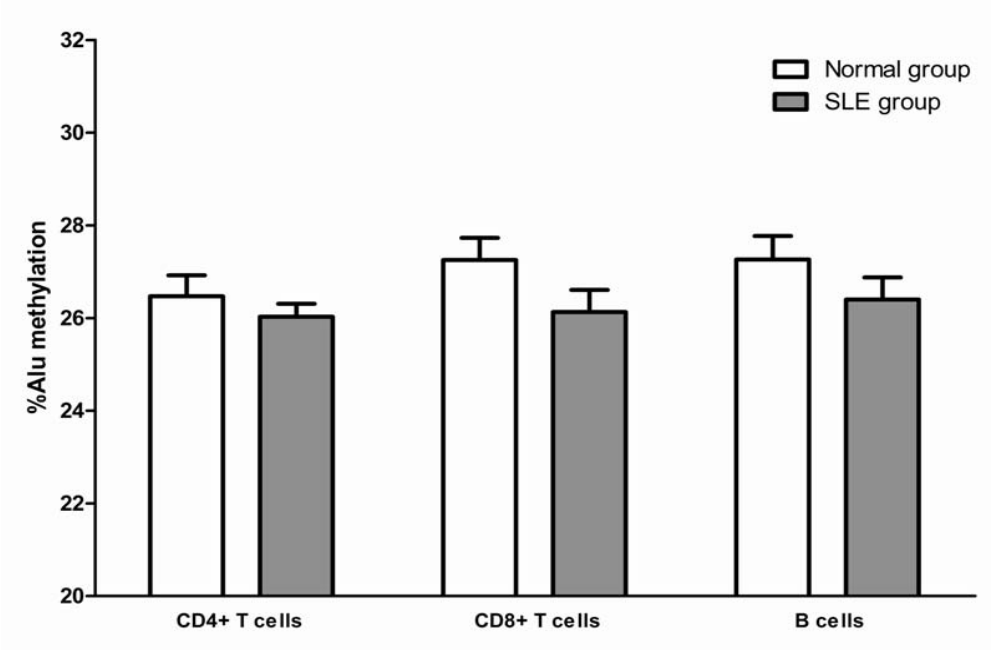


B

Figure 16. DNA Methylation levels in SLE CD4+ T cells, CD8+ T cells, and B cells compared to normal control of (A) HERV-E LTR2C (B) HERV-K LTR5_Hs (C) LINE-1 and (D) Alu were shown in Mean ± SEM. *P*-value calculated by unpaired t-test two-tailed analysis.



C



D

Figure 16. DNA Methylation levels in SLE CD4+ T cells, CD8+ T cells, and B cells compared to normal control of (A) HERV-E LTR2C (B) HERV-K LTR5_Hs (C) LINE-1 and (D) Alu were shown in Mean ± SEM. (Continue)

Changes in HERV-E LTR2C, HERV-K LTR5_Hs, LINE-1 and Alu methylation levels in SLE patients according to disease activity

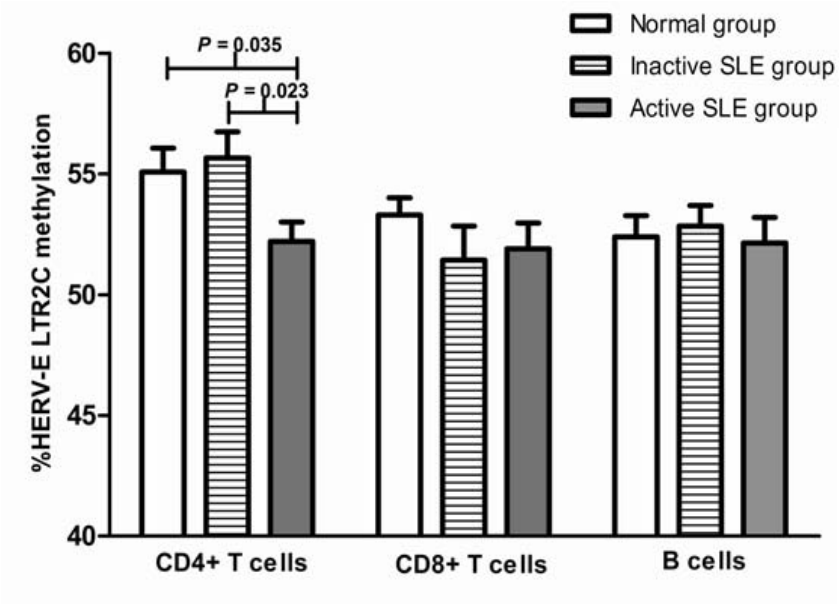
Methylation levels in lymphocyte subsets from patient with inactive and active SLE with control group were compared. SLE patients were divided into active and inactive groups based on their SLEDAI-2K score. Hypomethylation of LINE-1 was significantly observed in B cells from both groups (P = 0.028 and P = 0.0042, respectively) (Figure 17C). Interestingly, hypomethylation of LINE-1 was significantly more prominent in the active SLE group for both CD4+ and CD8+ T cells when compared to the inactive group. Even the hypomethylation in the active SLE group for CD4+ and CD8+ T cells were significantly different when compared to the normal controls (P = 0.0011 and 0.0025, respectively). Among the inactive SLE group, a trend of hypomethylation in CD4+ and CD8+ T cells was observed when compared to the normal controls (P = 0.2566 and 0.1555, respectively) but this was not statistically significant (Figure 17C). Surprisingly, Alu hypomethylation was detected in CD8+ lymphocytes from the inactive SLE group compared to the normal controls (P = 0.006) (Figure 17D).

The hypomethylation of HERV-E LTR2C was observed in CD4+ T cells of active SLE compared with inactive SLE and normals (P = 0.023 and P = 0.035, respectively) (Figure 17A). In contrast, inactive SLE showed demethylation of HERV-K LTR5_Hs in CD4+ T cells when compared with normal and active group (P = 0.002 and P = 0.027, respectively). Moreover, demethylation of HERV-K LTR5_Hs in B cells was also found when compared with normal group (P = 0.048) (Figure 17B).

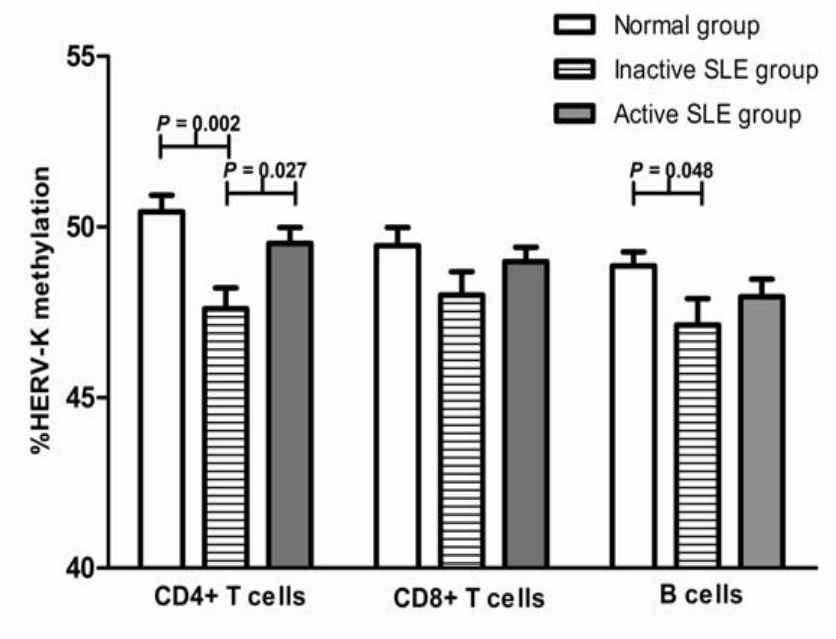
Each band of COBRA HERV-E LTR2C indicated methylation levels from difference sequence of PCR products (Figure 18). Therefore, each band of HERV-E LTR2C methylation was compared between normal and SLE patients. Interestingly, there was hypomethylation at 53 bp-methylated band (band 5) in CD3+CD4+ T cells and CD3+CD8+ T cells of active SLE when compared with normal (P = 0.01 and P = 0.038, respectively) (Figure 18D). Furthermore, although decreasing of methylation levels in B cells of active SLE patients was observed, there were no significant difference when compared to normal group (P = 0.19) (Figure 18D). The correlation between methylation levels of HERV-E

LTR2C and 53 bp-methylated band were evaluated. There were significant positive correlations of normal and active SLE in all cell subset ($P < 0.005$) (Figure 19). Therefore, hypomethylation at the 53 bp-methylated band in CD3+CD4+ T cells of active SLE may represents results of the COBRA HERV-E LTR2C.

Hypermethylation at 82 bp-methylated band (band 5) were observed in CD3+CD4+ T cells of inactive SLE in comparing with normal and active group ($P = 0.026$ and $P = 0.013$, respectively)(Figure 18B). Furthermore, there was hypomethylation at 73 bp-methylated band (band 4) in CD3+CD8+ T cells of inactive SLE compared with normal group ($P = 0.003$) (Figure 18C). Therefore, methylation status of HERV-E LTR2C seem to be different depending on their location.

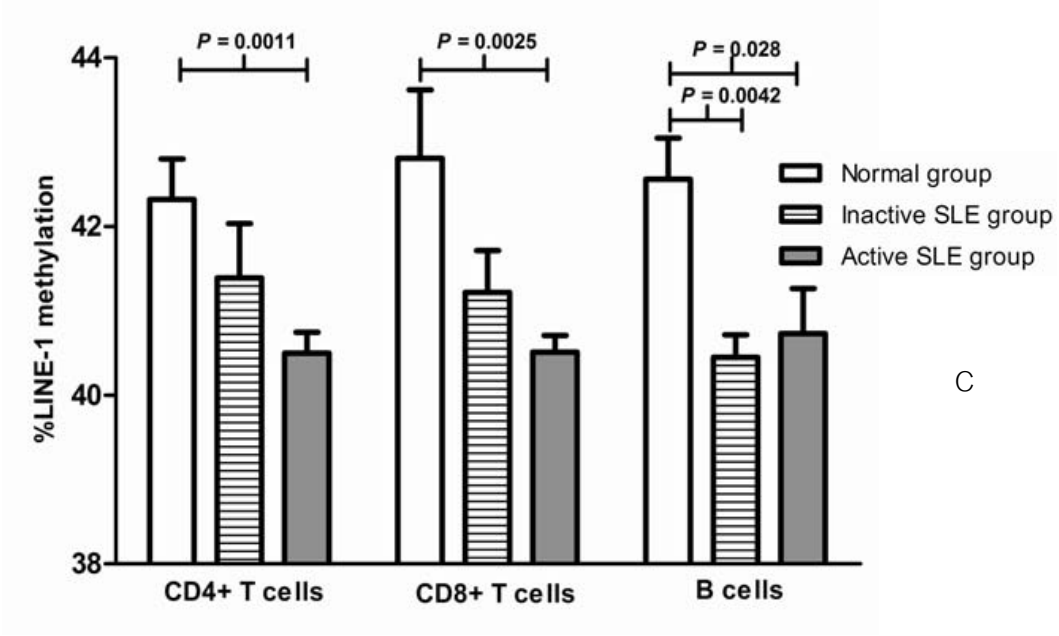


A

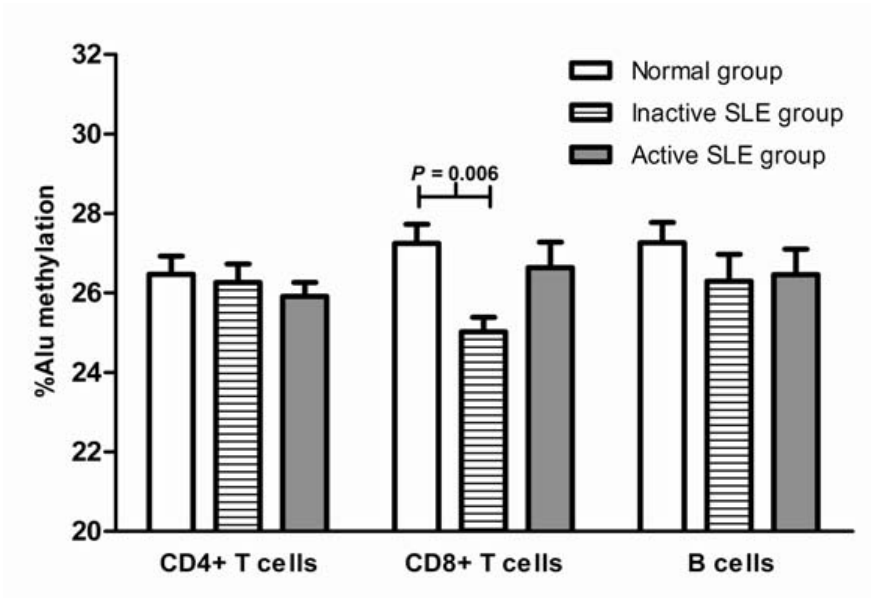


B

Figure 17. DNA Methylation levels in different SLE disease activity. Methylation levels of (A) HERV-E LTR2C (B) HERV-K LTR5_Hs (C) LINE-1 and (D) Alu were shown in Mean ± SEM. *P*-value calculated by unpaired t-test two-tailed analysis.

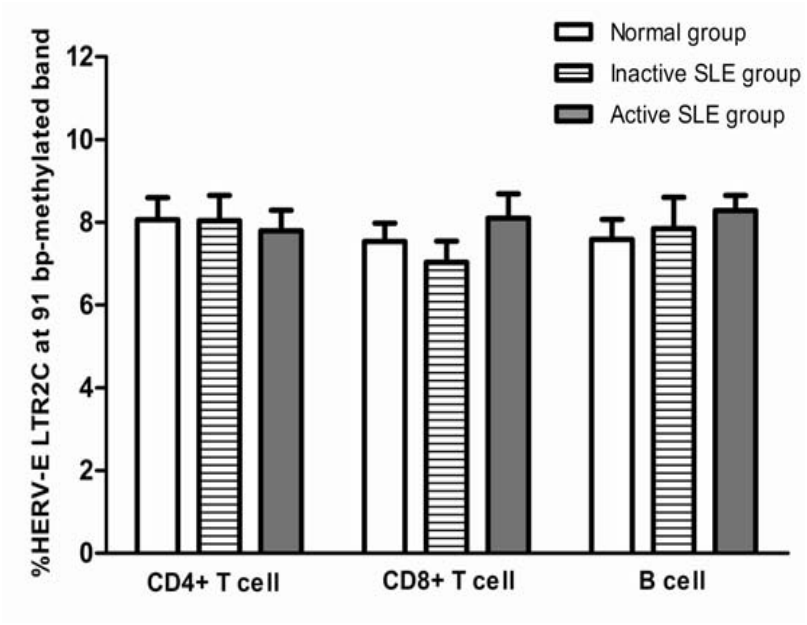


C

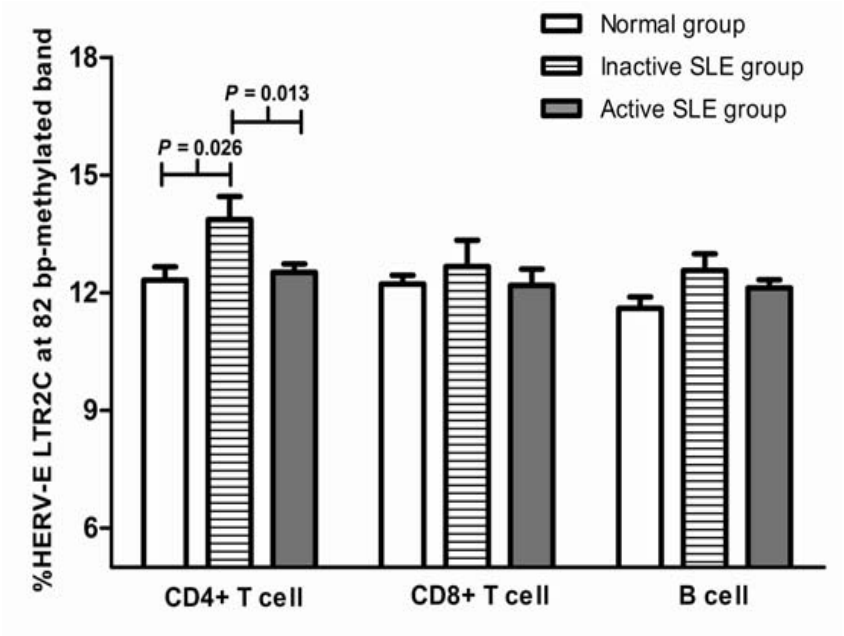


D

Figure 17. DNA Methylation levels in different SLE disease activity. Methylation levels of (A) HERV-E LTR2C (B) HERV-K LTR5_Hs (C) LINE-1 and (D) Alu were shown in Mean ± SEM. (Continue)

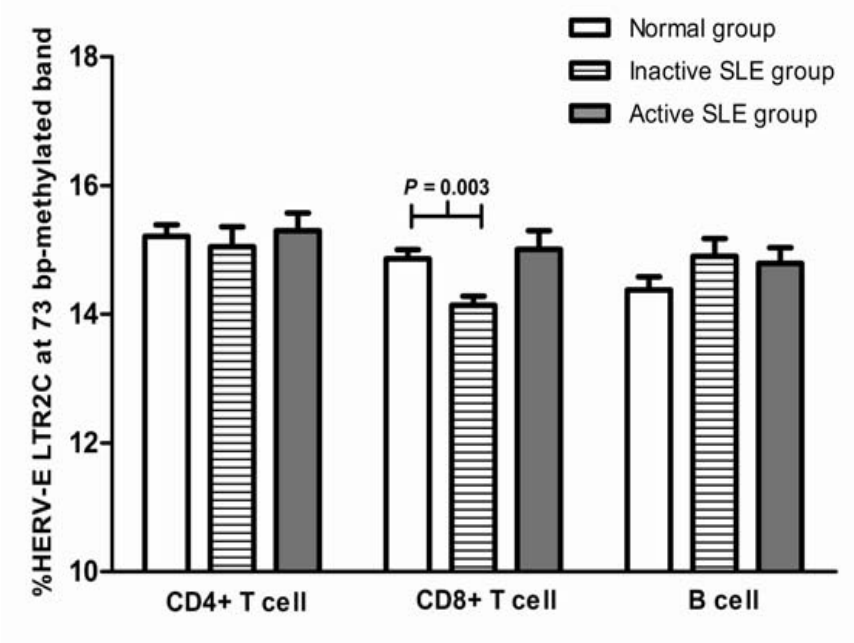


A

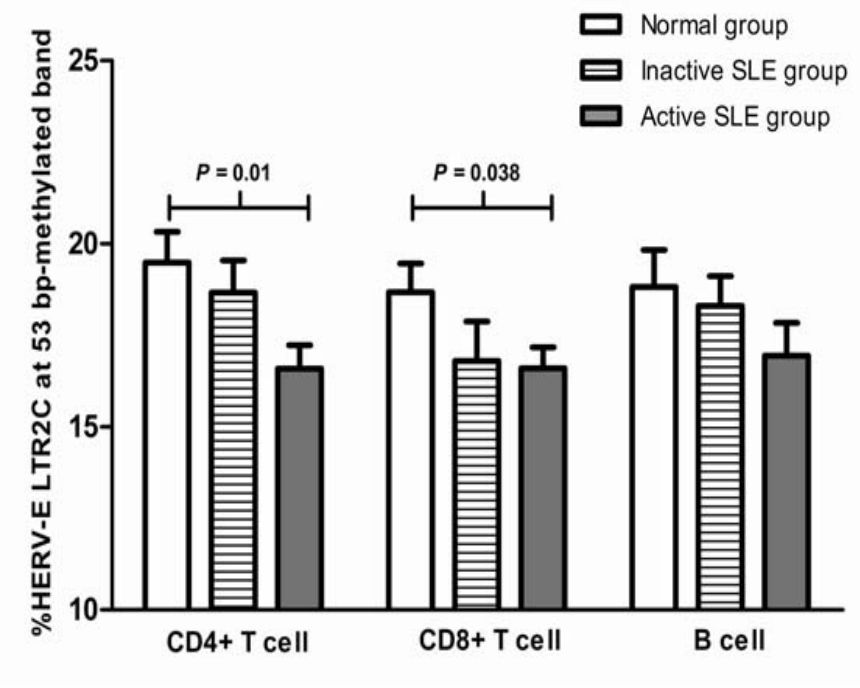


B

Figure 18. DNA Methylation levels of HERV-E LTR2C in each band. Methylation levels of (A) 91 bp, (B) 82 bp, (C) 73 bp and (D) 53 bp of HERV-E LTR2C in normal and SLE lymphocyte subsets were shown in Mean±SEM. *P*-value calculated by unpaired t-test two-tailed analysis.

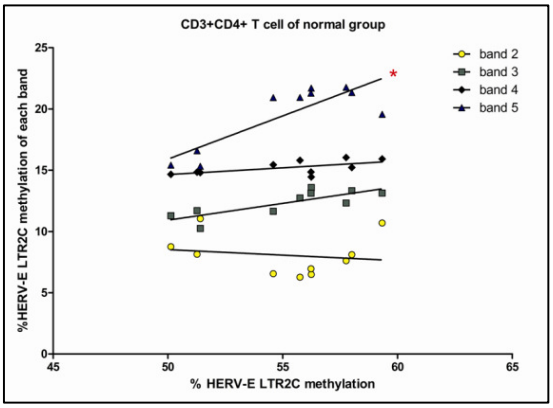


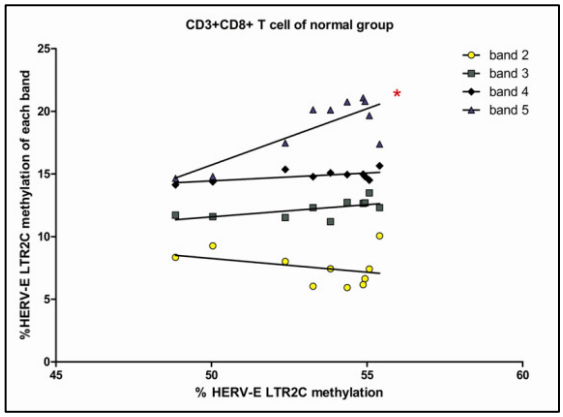
C



D

Figure 18. DNA Methylation levels of HERV-E LTR2C in each band. Methylation levels of (A) 91 bp, (B) 82 bp, (C) 73 bp and (D) 53 bp of HERV-E LTR2C in normal and SLE lymphocyte subsets were shown in Mean±SEM. (Continue)





A

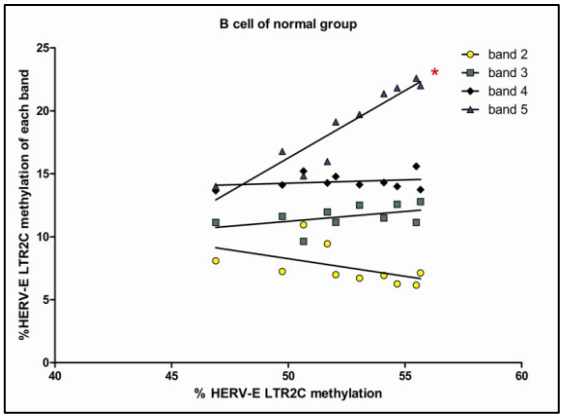
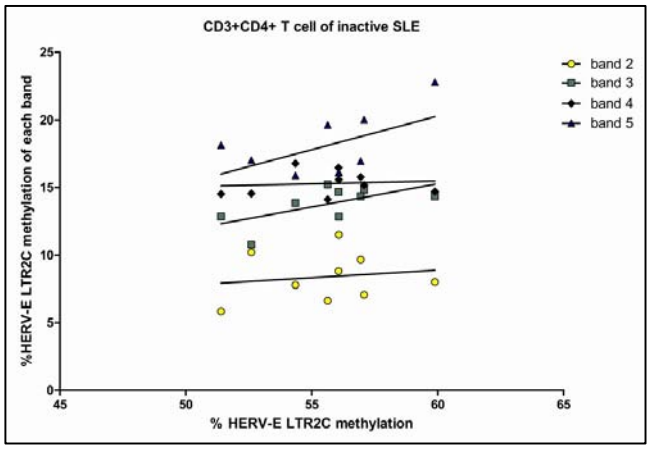
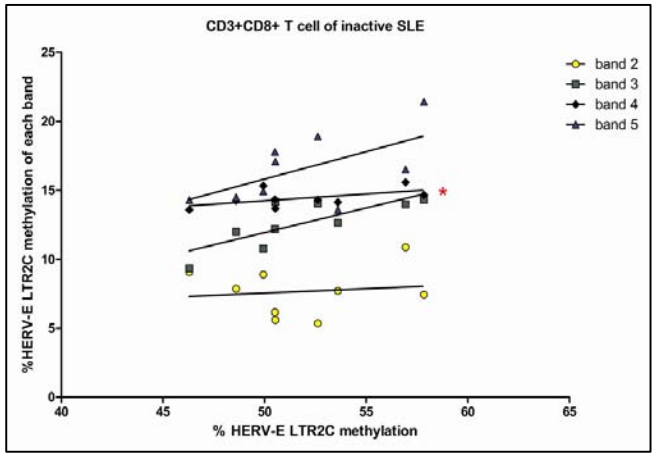


Figure 19. Correlation between DNA Methylation levels of HERV-E LTR2C in each band and global HERV-E LTR2C in normal (A) and SLE lymphocyte subsets (B-C). Each dot represents an individual sample. The asterisks indicate significant differences in hypomethylation levels between normal and SLE groups at $P < 0.05$ (*). *P*-value calculated by Pearson's correlation coefficient test.





B

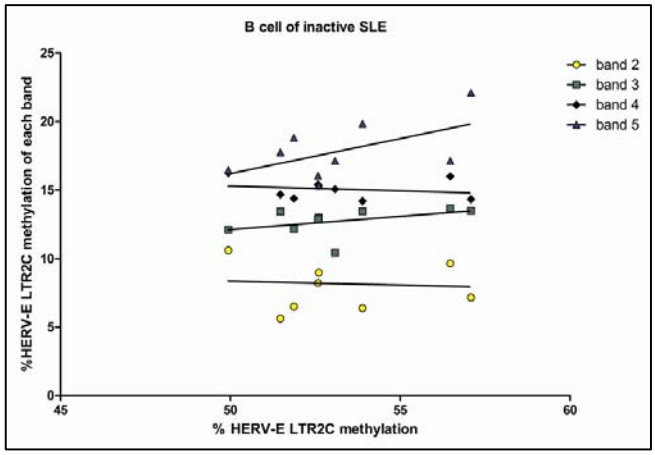
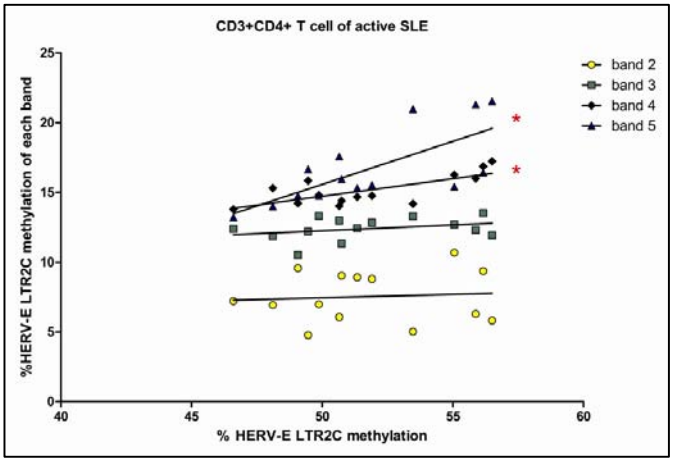
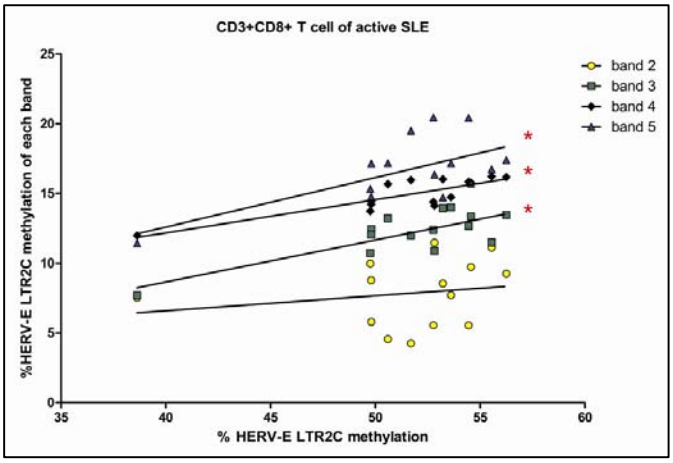


Figure 19. Correlation between DNA Methylation levels of HERV-E LTR2C in each band and global HERV-E LTR2C in SLE lymphocyte subsets (B-C). (Continue)





C

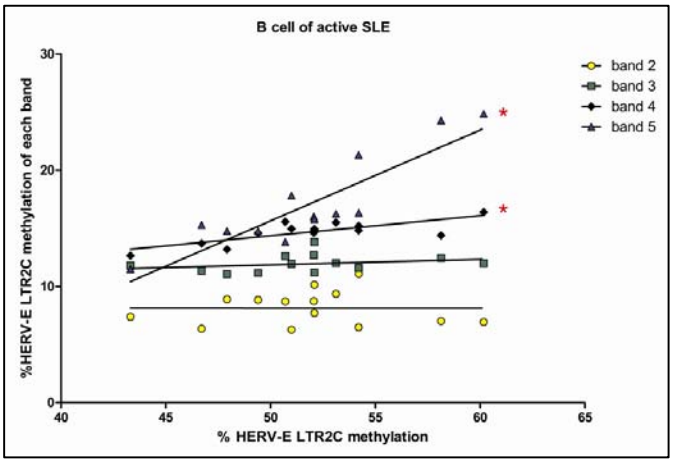


Figure 19. Correlation between DNA Methylation levels of HERV-E LTR2C in each band and global HERV-E LTR2C in SLE lymphocyte subsets (B-C). (Continue)

Association of HERV-E LTR2C, HERV-K LTR5_Hs, LINE-1 and Alu methylation and clinical parameters

We also correlated methylation levels of retroelements with SLEDAI score and clinical parameters including white blood counts, % lymphocytes, platelet counts, anti-dsDNA and antinuclear antibodies titer, complement activity (CH50) and quantitative complement levels (C3) (Table. 4). Methylation levels of HERV-E LTR2C and each band of COBRA HERV-E LTR2C in CD3+CD4+ T cells were correlated with leucopenia (P = 0.016, r = 0.51) and lymphopenia (P = 0.023, r = 0.48) (Figure 20A and Table 4). We found positive correlation of HERV-K LTR5_Hs, and negative correlation of HERV-E LTR2C band 3 in CD4+ T cells with SLEDAI score (P = 0.02, r = 0.48 and P = 0.008, r = -0.53, respectively) (Figure 20B and Table 4). The hypomethylation of HERV-K LTR5_Hs in CD3+CD4+ T cells was negative correlated with complement activity (P = 0.043, r = -0.43) (Figure 20B). Additional, correlation of methylation levels at HERV-E LTR2C band 2 in CD8+ T cells with Anti-dsDNA titer and C3 level were found (P = 0.01, r = 0.49) (Table 4). However, there were no significant correlations between methylation levels of LINE-1 and Alu to any of the clinical parameters studied (Table. 4). Therefore, hypomethylation of HERV-E and HERV-K have potential to be a prognostic marker in SLE disease activity. Further study to find out the functional role of these hypomethylations on clinical outcome might lead to the discovery of novel SLE pathogenesis pathway.

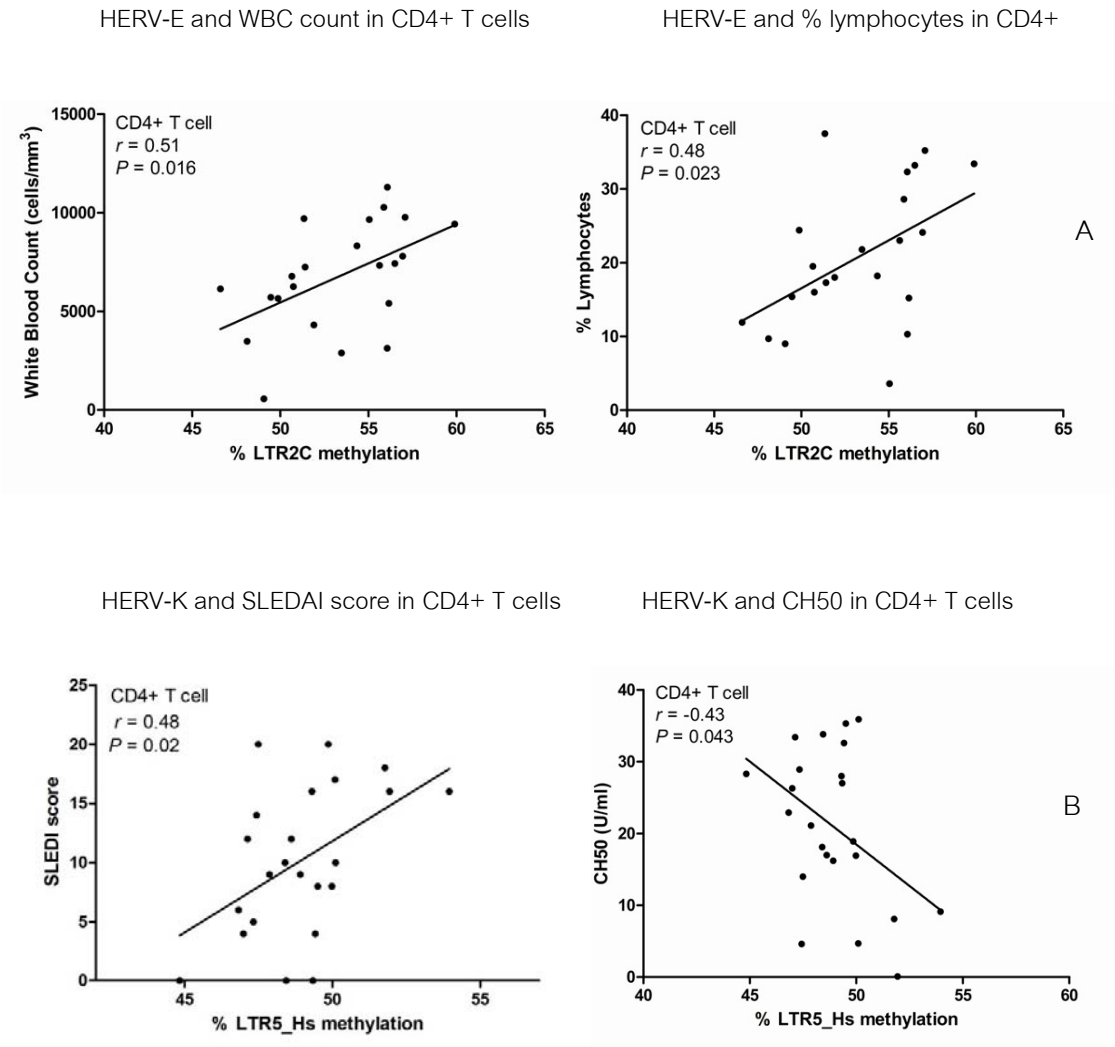


Figure 20. Correlation between DNA Methylation levels with disease activity and clinical characteristics. The relationship of HERV-E LTR2C (A) and HERV-K LTR5_Hs (B) methylation levels in lymphocyte subsets with clinical data and SLEDAI score of SLE patients were shown. Each dot represents an individual patient. *P*-value calculated by Pearson's correlation coefficient test.

Table 4. Correlation between DNA Methylation levels with disease activity and clinical characteristics.

		SLEDAI-2K		white blood counts (6755 ± 586.5 cells/mm^3)		% Lymphocytes (20.8 ± 2.02)		Platelet counts (252587 ± 22447 cells/mm^3)	
		r	p-value	r	p-value	r	p-value	r	p-value
LINE-1	CD4+ T cells	-0.07	0.74	0.08	0.72	0.18	0.43	0.10	0.66
	CD8+ T cells	-0.31	0.15	0.36	0.10	0.05	0.80	0.25	0.27
	B cells	0.13	0.57	0.20	0.37	0.18	0.41	-0.04	0.87
Alu	CD4+ T cells	-0.01	0.99	0.31	0.21	-0.09	0.71	0.33	0.20
	CD8+ T cells	0.44	0.06	0.06	0.80	-0.18	0.48	-0.16	0.54
	B cells	0.16	0.52	-0.10	0.71	0.27	0.29	-0.02	0.95
HERV-K	CD4+ T cells	0.48	0.02*	-0.41	0.06	-0.29	0.18	-0.09	0.71
	CD8+ T cells	0.15	0.48	-0.06	0.79	-0.24	0.29	0.17	0.45
	B cells	0.26	0.23	-0.10	0.66	0.08	0.72	0.23	0.32
HERV-E	CD4+ T cells	-0.32	0.14	0.51	0.016*	0.48	0.023*	0.41	0.06
	CD8+ T cells	-0.05	0.80	0.03	0.90	-0.35	0.11	0.17	0.45
	B cells	-0.01	0.96	0.14	0.55	0.43	0.047*	0.15	0.53
HERV-E	CD4+ T cells	-0.05	0.81	0.29	0.18	-0.28	0.20	0.02	0.91
Band 2 (91 bp)	CD8+ T cells	0.09	0.65	-0.11	0.59	-0.58	0.004*	-0.29	0.17
	B cells	0.08	0.70	0.04	0.83	-0.31	0.15	-0.15	0.49
HERV-E	CD4+ T cells	-0.53	0.008*	0.03	0.87	0.45	0.03*	0.04	0.86
Band 3 (82 bp)	CD8+ T cells	-0.21	0.34	-0.07	0.75	-0.15	0.48	0.16	0.45
	B cells	-0.37	0.08	-0.06	0.78	0.28	0.19	-0.15	0.49
HERV-E	CD4+ T cells	-0.01	0.97	0.11	0.61	0.09	0.66	0.29	0.16
Band 4 (7 bp)	CD8+ T cells	0.07	0.73	-0.10	0.65	-0.38	0.08	0.17	0.44
	B cells	-0.09	0.66	0.08	0.71	0.27	0.22	0.02	0.92
HERV-E	CD4+ T cells	-0.12	0.58	0.27	0.20	0.58	0.004*	0.33	0.12
Band 5 (53 bp)	CD8+ T cells	0.09	0.65	-0.05	0.81	0.18	0.42	0.14	0.53
	B cells	-0.02	0.92	0.14	0.50	0.51	0.014*	0.22	0.31

* indicate significant differences by Pearson's correlation (r) coefficient test.

Numbers in the parentheses indicate mean ± SEM.

Table 4. Correlation between DNA Methylation levels with disease activity and clinical characteristics. (Continue)

		Anti-dsDNA titer (131.4 ± 60.3)		ANA titer (1657 ± 383.7)		CH50 (20.9 ± 2.3 U/ml)		C3 level (77.6 ± 5.9 mg/dl)	
		R	p-value	r	p-value	r	p-value	r	p-value
LINE-1	CD4+ T cells	0.12	0.62	0.06	0.77	-0.02	0.93	-0.19	0.38
	CD8+ T cells	-0.07	0.74	0.15	0.49	-0.06	0.78	-0.09	0.69
	B cells	-0.22	0.31	-0.00	0.99	-0.27	0.22	-0.16	0.46
Alu	CD4+ T cells	-0.14	0.57	0.02	0.92	0.30	0.21	0.27	0.27
	CD8+ T cells	0.03	0.91	-0.16	0.51	-0.20	0.41	-0.22	0.36
	B cells	-0.16	0.52	0.04	0.88	-0.01	0.97	0.03	0.89
HERV-K	CD4+ T cells	0.09	0.68	-0.14	0.54	-0.43	0.043*	-0.33	0.13
	CD8+ T cells	0.12	0.57	-0.00	0.99	-0.36	0.095	-0.30	0.16
	B cells	-0.12	0.58	-0.09	0.66	-0.00	0.98	0.09	0.68
HERV-E	CD4+ T cells	-0.27	0.21	-0.13	0.57	0.19	0.39	0.18	0.39
	CD8+ T cells	0.18	0.39	0.16	0.47	-0.06	0.79	-0.17	0.44
	B cells	-0.03	0.89	-0.17	0.43	0.07	0.72	0.07	0.73
HERV-E	CD4+ T cells	-0.01	0.94	0.11	0.62	-0.12	0.58	-0.23	0.29
Band 2 (91 bp)	CD8+ T cells	0.49	0.01*	0.28	0.19	-0.28	0.19	-0.44	0.03*
	B cells	0.26	0.22	0.18	0.39	-0.19	0.39	-0.35	0.10
HERV-E	CD4+ T cells	-0.42	0.04*	-0.12	0.58	0.28	0.18	0.41	0.054
Band 3 (82 bp)	CD8+ T cells	-0.24	0.25	-0.02	0.93	0.11	0.63	0.15	0.50
	B cells	-0.06	0.78	-0.17	0.42	0.16	0.46	0.31	0.15
HERV-E	CD4+ T cells	-0.06	0.78	0.02	0.92	0.17	0.42	0.04	0.85
Band 4 (7 bp)	CD8+ T cells	0.22	0.32	0.12	0.58	0.04	0.84	-0.01	0.95
	B cells	0.24	0.27	-0.03	0.88	0.19	0.39	-0.01	0.98
HERV-E	CD4+ T cells	-0.33	0.12	-0.19	0.37	0.13	0.56	0.19	0.38
Band 5 (53 bp)	CD8+ T cells	-0.01	0.96	-0.07	0.75	-0.01	0.95	-0.04	0.87
	B cells	-0.24	0.26	-0.19	0.38	0.14	0.53	0.21	0.34

* indicate significant differences by Pearson's correlation (r) coefficient test.

Numbers in the parentheses indicate mean ± SEM.

Association between HERV-E LTR2C, HERV-K LTR5_Hs, LINE-1 and Alu methylation

In order to investigate whether the hypomethylation of each interspersed repetitive sequences (IRS) occur via the same process or not, the correlations of methylation levels between IRS were evaluated. There was correlation of LINE-1 and Alu methylation in normal group when analyzed in all cell types (P = 0.023, r = 0.7) (Table. 5). Interestingly, this correlation was found only in active SLE (P = 0.04, r = 0.32) (Table. 5). However, there was no correlation between HERV-E LTR2C and HERV-K LTR5_Hs methylation levels in normal and SLE (Table. 5). For lymphocyte subset, a significant positive correlation was found between Alu and LINE-1 in CD4+ T cells of normal control (P = 0.023, r = 0.7) (Table. 6). The association between IRS proved that hypomethylation is not a generalized process. Therefore, there are specific mechanisms of demethylation for each LTR subclass.

Table 5. Correlation between HERV-E LTR2C, HERV-K LTR5_Hs, LINE-1 and Alu methylation

IRS and cell type of normal group	HERV-E all cell types		HERV-K all cell types		LINE-1 all cell types	
	R	p-value	r	p-value	r	p-value
HERV-K all cell types	0.10	0.58	1		-0.08	0.67
LINE-1 all cell types	-0.085	0.65	-0.08	0.67	1	
Alu all cell types	-0.45	0.19	0.29	0.40	0.70	0.023*

IRS and cell type of SLE group	HERV-E all cell types		HERV-K all cell types		LINE-1 all cell types	
	R	p-value	r	p-value	r	p-value
HERV-K all cell types	0.10	0.40	1		0.05	0.67
LINE-1 all cell types	0.22	0.06	0.05	0.67	1	
Alu all cell types	0.09	0.51	0.10	0.45	0.11	0.42

IRS and cell type of inactive SLE	HERV-E all cell types		HERV-K all cell types		LINE-1 all cell types	
	r	p-value	r	p-value	r	p-value
HERV-K all cell types	0.13	0.57	1		-0.17	0.45
LINE-1 all cell types	0.43	0.05	-0.17	0.45	1	
Alu all cell types	0.26	0.28	0.06	0.81	-0.25	0.29

IRS and cell type of active SLE	HERV-E all cell types		HERV-K all cell types		LINE-1 all cell types	
	r	p-value	r	p-value	r	p-value
HERV-K all cell types	0.16	0.26	1		0.19	0.17
LINE-1 all cell types	0.13	0.37	0.19	0.17	1	
Alu all cell types	0.06	0.70	0.04	0.80	0.32	0.04*

* indicate significant differences by Pearson's correlation (r) coefficient test.

Table 6. Correlation between HERV-E LTR2C, HERV-K LTR5_Hs, LINE-1 and Alu methylation in CD4+ T cell normal group

IRS and cell type in normal group	HERV-E CD4+ T	HERV-K CD4+ T	LINE-1 CD4+ T	Alu CD4+ T
HERV-K CD4+ T	r = -0.042 *P* = 0.91			
LINE-1 CD4+ T	r = -0.22 *P* = 0.52	r = 0.36 *P* = 0.29		
Alu CD4+ T	r = -0.45 *P* = 0.19	r = 0.29 *P* = 0.41	r = 0.70 *P* = 0.023*	
HERV-E CD8+ T	r = 0.83 *P* = 0.002*	r = -0.04 *P* = 0.91	r = -0.16 *P* = 0.64	r = -0.60 *P* = 0.06
HERV-K CD8+ T	r = -0.59 *P* = 0.071	r = 0.101 *P* = 0.78	r = -0.22 *P* = 0.52	r = -0.15 *P* = 0.66
LINE-1 CD8+ T	r = -0.13 *P* = 0.71	r = 0.46 *P* = 0.17	r = 0.41 *P* = 0.23	r = 0.40 *P* = 0.25
Alu CD8+ T	r = 0.061 *P* = 0.86	r = 0.18 *P* = 0.60	r = 0.13 *P* = 0.71	r = 0.58 *P* = 0.07
HERV-E B cells	r = 0.64 *P* = 0.042*	r = 0.40 *P* = 0.24	r = 0.15 *P* = 0.66	r = 0.11 *P* = 0.76
HERV-K B cells	r = -0.35 *P* = 0.31	r = 0.12 *P* = 0.73	r = 0.03 *P* = 0.92	r = 0.33 *P* = 0.33
LINE-1 B cells	r = -0.33 *P* = 0.34	r = 0.33 *P* = 0.33	r = 0.75 *P* = 0.01*	r = 0.71 *P* = 0.02*
Alu B cells	r = 0.51 *P* = 0.12	r = 0.02 *P* = 0.94	r = -0.05 *P* = 0.88	r = -0.17 *P* = 0.62

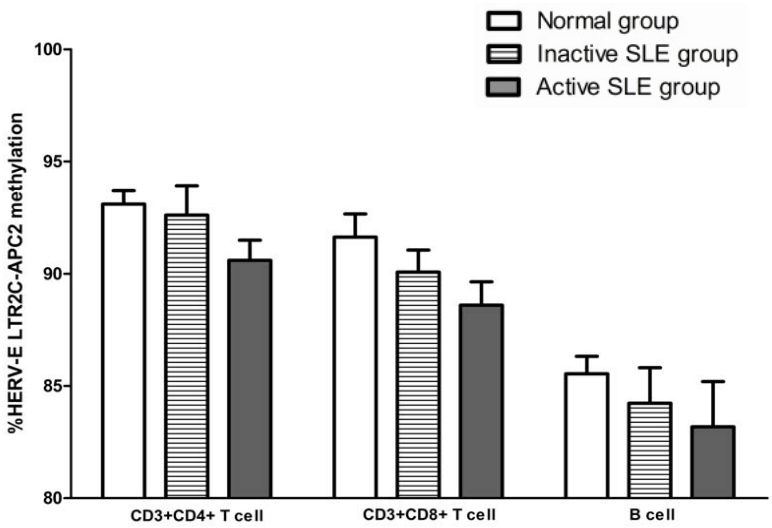
* indicate significant differences by Pearson's correlation (r) coefficient test.

DNA methylation of LTR2C loci specific

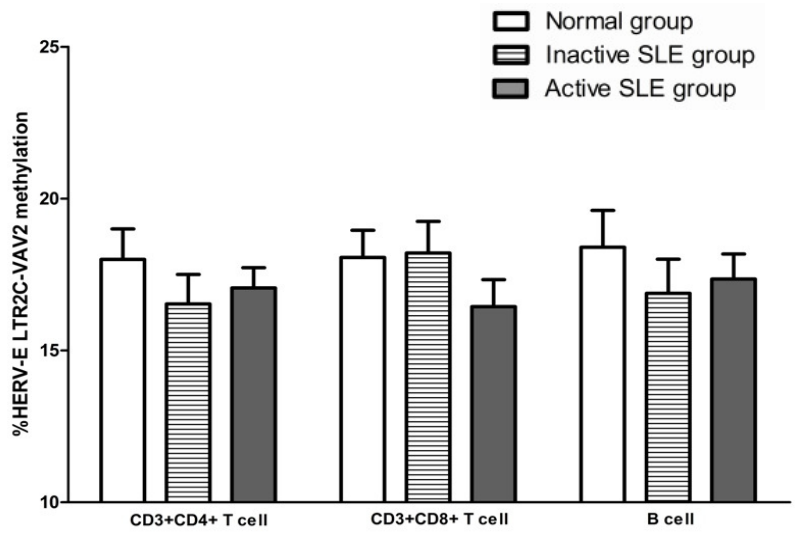
Since hypomethylstion of HERV-E LTR2C was investigated in CD4+ T cell of active SLE. Several studies have been reported the functional role of LTR-retroelements in human genome as alternative promoter, a new exon coding protein and a polyadenylation signal to neighboring genes (79, 89, 91, 100) To investigate the hypomethylated LTR may lead to identify genes that associated with SLE. A total of 295 (LTR2C) and 645 (LTR5_Hs) were mapped in human genome (GRCh37) by using Table Browser of UCSC Browser (125) Individual LTRs were chosen as follows i) potential of LTRs that were amplified by our primer based on an *in silico* PCR of BiSearch with bisulfite converted-human genome DNA (123, 124) ii) location and annotation of LTRs were performed by using UCSC Genome Browser (125), blast 2 sequences alignment (http://www.ncbi.nlm.nih.gov), and RepeatMasker (126) HERV-E LTR2C loci that located within approximately 15 kb of genes were shown in Table 7. The most type of HERV-E LTR2C is a single (solitary) LTR as previously reported (129) We selected a specific HERV-E LTR2C that created 53 bp-methylated band size from COBRA HERV-E LTR2C. We compared LTR2C methylation levels of 2 loci between normal and SLE CD4+ T lymphocytes, CD8+ T lymphocytes and B lymphocytes by using COBRA for unique HERV-E LTR2C sequence. HERV-E LTR2C-APC2 is located at chromosome 19 approximately 10 kb at 5' of adenomatous polyposis coli 2 (APC2) genes. Another loci, HERV-E LTR2C-VAV2 is located at chromosome 9 approximately 10 kb 3' of vav 2 guanine nucleotide exchange factor (VAV2). In normal lymphocytes, we found that methylation levels of HERV-E LTR2C-APC2 in B cells showed significantly decreased to CD3+CD8+ T cells and CD3+CD4+ T cells ($P = 0.0017$ and $P < 0.0001$, respectively) (Figure 21A), but not in HERV-E LTR2C-VAV2 methylation (Figure 21B). These results indicated that methylation levels of HERV-E LTR2C-APC2 are cell type specific manner. Furthermore, although decreasing of methylation levels in CD3+CD4+ T cells and CD3+CD8+ T cells of active SLE patients were observed, there were no significant difference when compared to normal group ($P = 0.053$ and $P = 0.062$) (Figure 21A and 23). However, the positive finding came from a limited sample size, so larger study should be performed to validate this observation. Meanwhile, there were no

significant differences of methylation levels at HERV-E LTR2C-VAV2 (Figure 21B). Because a significant positive correlation were found between global HERV-E LTR2C and both of LTR2C specific loci ($P < 0.0001$, r = 0.42 in HERE LTR2C-APC2, Figure 22A and $P = 0.023$, r = 0.23 in HERV-E LTR2C-VAV2, Figure 22B). Therefore, the methylation levels of the HERV-E LTR2C represent methylation levels of each HERV-E LTR2C loci in a positive correlated fashion.

However, it is remained to be determined what the effect of these hypomethylation on lymphocytes. Since LTR-retrotransposons can impact on human gene expression. We evaluate whether HERV-E LTR2C is associated in human transcripts by using human express sequence tag (EST) and mRNA databases to identify HERV-E LTR2C fusion transcripts. Our results showed that HERV-E LTR2C-associated transcripts were detected in various normal tissues and cancerous cell lines (Table 8). However, the functional consequences of HERV-E LTR2C-fusion transcripts are limited. To study the important role of HERV-E LTR2C sequence and DNA methylation in these alternative transcripts are necessary, and might lead to the discovery of novel SLE pathogenesis pathway.



A



B

Figure 21. HERV-E LTR2C APC2 and HERV-E LTR2C VAV2 methylation levels. Methylation levels of HERV-E LTR2C APC2 (A) and HERV-E LTR2C VAV2 (B) in normal and SLE lymphocyte subsets were shown in Mean±SEM. *P*-value calculated by unpaired t-test two-tailed analysis.

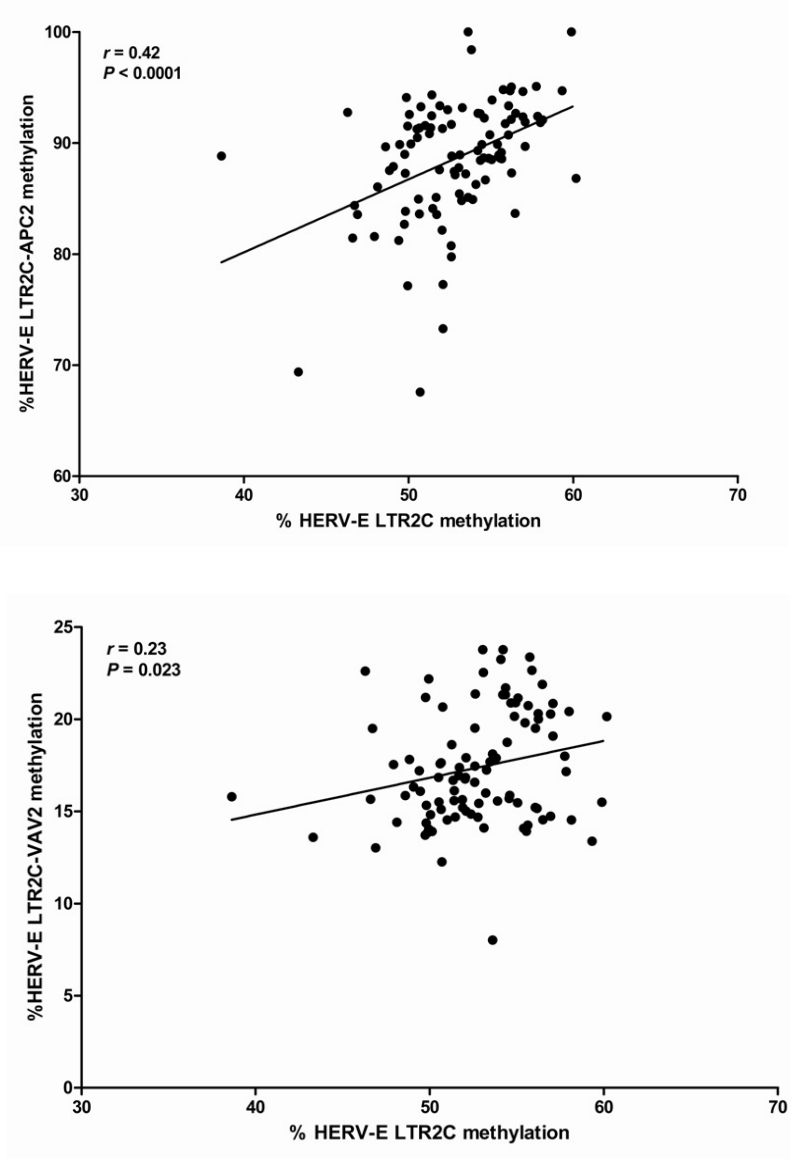


Figure 22. Correlation between DNA Methylation levels of HERV-E LTR2C APC2 (A) and HERV-E LTR2C VAV2 (B) with HERV-E LTR2C. Each dot represents an individual sample. *P*-value calculated by Pearson's correlation coefficient test.

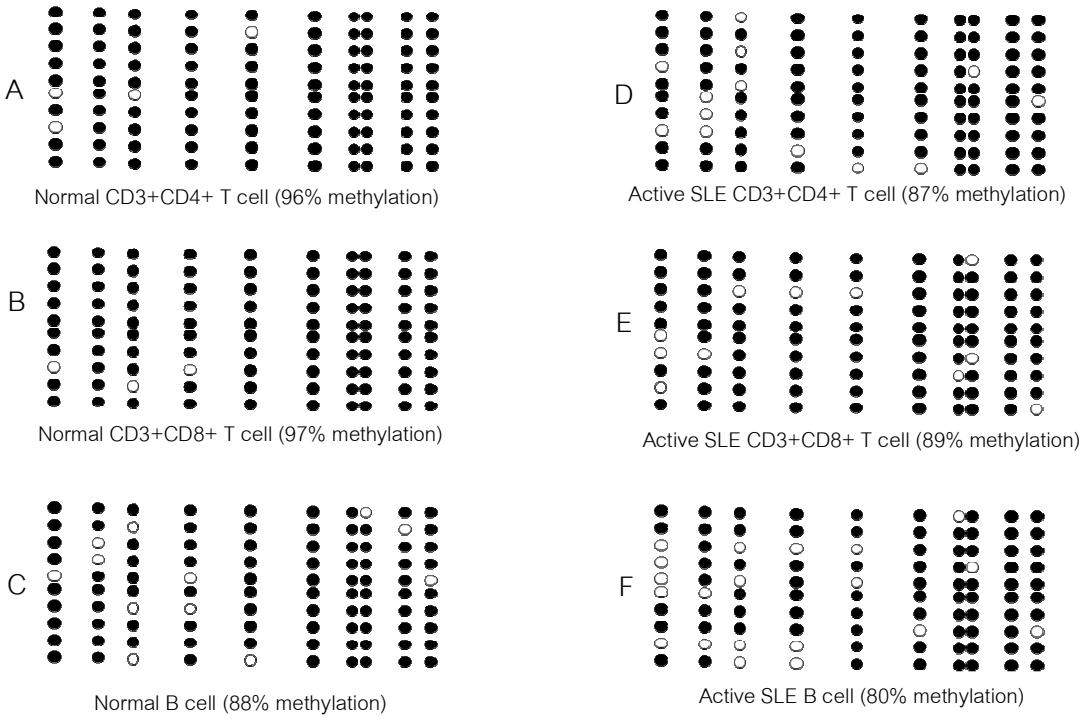


Figure 23. HERV-E LTR2C-APC2 bisulfite sequences. Bisulfite-treated DNA from normal and active SLE were amplified, cloned, and 10 fragments per donor were sequenced. Black circles correspond to methylated CpGs and white circles to unmethylated CpGs.

Table 7. Localization of HERV-E LTR2C in Human genome.

Adjacent gene (gene symbol)	Chromosome (position)	Type
Budding uninhibited by benzimidazoles 1 homolog (BUB1)	2q13 (111436874-111437374)	solitary LTR
Anoctamin 7 (ANO7) and high density lipoprotein binding protein (HDLBP)	2q37.3 (242158045-242158536)	solitary LTR
UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7 (B3GNT7)	2q37.1 (232265423- 232273554)	complete retrovirus
Family with sequence similarity 19 (chemokine (C-C motif)-like), member A1 (FAM19A1)	3p14.1 (68480892-68481390)	solitary LTR
Phosphoglucomutase 2 (PGM2)	4p14 (37827195-37827693)	solitary LTR
Williams-Beuren syndrome chromosome region 17 (WBSCR17)	7q11.22 (71182684- 71183206)	solitary LTR
Liver-related low express protein 1 (LRLE1)	8p23.1 (8037610-8038106)	solitary LTR
Zinc finger CCCH-type containing 3 (ZC3H3)	8q24.3 (144629745- 144630540)	solitary LTR
Gasdermin D (GSDMD)	8q24.3 (144630235-144630540)	solitary LTR
Vav 2 guanine nucleotide exchange factor (VAV2)	9q34.2 (136618800-136619288)	solitary LTR
Lipocalin 9 (LCN9) and spermatogenesis and oogenesis specific basic helix-loop-helix 1 (SOHLH1)	9q34.3 (138572173-138572655)	solitary LTR
Barrier to autointegration factor 1 (BANF1) and Eukaryotic translation initiation factor 1A domain containing (EIF1AD)	11q13.1 (65768120-65768619)	solitary LTR
Down syndrome cell adhesion molecule like 1 (DSCAML1)	11q23.3 (117429319-117429812)	solitary LTR
Hydroxysteroid (17-beta) dehydrogenase 12 (HSD17B12)	11p11.2 (43805542-43807048)	solitary LTR
Unc-93 homolog B1 (UNC93B1) and Aldehyde dehydrogenase 3B1 isoform a (ALDH3B1)	11q13.2 (67776867-67777348)	solitary LTR
Growth differentiation factor 3 precursor 9 (GDF3) and Developmental pluripotency associated 3 (DPPA3)	12p13.31 (7856233-7856732)	solitary LTR
Scavenger receptor class B, member 1 (SCARB1)	12q24.31 (125311650-125312153)	solitary LTR

Adjacent gene (gene symbol)	Chromosome (position)	Type
Myosin XVI (MYO16)	13q33.3 (109791863-109792342)	solitary LTR
G protein-coupled receptor 132 (GPR132)	14q32.33 (105504753-105505283)	solitary LTR
Eukaryotic translation initiation factor 3, subunit C-like (EIF3CL) and Nuclear pore complex-interacting protein-like 2 (NPIPL2)	16p11.2 (28381622-28382102)	solitary LTR
Quinolinate phosphoribosyltransferase (QPRT)	16p11.2 (29709572-29710062)	solitary LTR
Bleomycin hydrolase (BLMH)	17q11.2 (28567351-28568152)	solitary LTR
GTPase activating Rap/RanGAP domain-like 4 isoform 1 (RAP1GAP2)	17p13.3 (2688605-2689092)	solitary LTR
Free fatty acid receptor 3 (FFAR3)	19q13.12 (35873352-35873849)	solitary LTR
Zinc finger protein 816A (ZNF816A)	19q13.41 (53459980-53460457)	solitary LTR
Zinc finger protein 765 (ZNF765)	19q13.42 (53891569-53892069)	solitary LTR
Ribosomal protein S15 (RPS15) and Adenomatosis polyposis coli 2 (APC2)	19p13.3 (1442427-1442895)	solitary LTR
Zinc finger protein 66 (ZNF66)	19p12 (20929626-20938438)	complete retrovirus
Cystatin SA precursor (CST2)	20p11.21 (23803502-23803987)	solitary LTR
Integrin, beta 2 (ITGB2)	21q22.3 (46328139-46328643)	solitary LTR
Splicing factor 3a, subunit 1, 120kDa (SF3A1)	22q12.2 (30749326-30749820)	solitary LTR
Tubulin tyrosine ligase-like family, member 8 (TTLL8)	22q13.33 (50488954-50489439)	solitary LTR
Alpha-N-acetylgalactosaminidase (NAGA)	22q13.2 (42459334-42459823)	solitary LTR
Myosin XVIIIB (MYO18B)	22q12.1 (26218589-26219072)	solitary LTR
NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1 (NDUFA1) and A kinase (PRKA) anchor protein 14 isoform b (AKAP14)	Xq24 (119025972-119026456)	solitary LTR

Adjacent gene (gene symbol)	Chromosome (position)	Type
cancer/testis antigen CT45-6 (CT45A6) and sarcoma antigen 1 (SAGE1)	Xq26.3 (134971892-134972785	solitary LTR

Only genes mapping within approximately 15 kb were analyzed by using UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly (http://genome.ucsc.edu/cgi-bin/hgGateway).

Table 8. Expression of human endogenous retrovirus-E LTR2C-associated transcript in humans

mRNA or EST	Chromosome (position) of LTR2C	Organ / tissue type
BC041651, BM551788	1p36.12 (20486885-20487006)	Brain/astrocytoma grade IV (cell line)
DB042498 DA619311	1q21.1 (146697347-146697807)	Testis Leukemia cell line (myelogenous)
DB080195	1q24.2 (167879066-167879235)	Testis
AJ629456	1q32.1 (202137571-202137626)	MCF7 (breast cancer cell line)
AK125777	2p14 (65114546-65114656)	Testis
BC040607	2q13 (111436874-111437374)	Testis
AK126207 AK000770 CR626252	2q37.1 (232265423-232265917)	n/a Colon Neuroblastoma
BC041468, BI912296	4p16.1 (6689133-6689629)	Leukocyte
DA016574	4p16.1 (8441571-8442053)	Adrenal gland
CA773740	4q35.1 (184385713-184385852)	Pancreas/ Purified pancreatic islet
AK125006	6q15 (89380261-89380669)	Thalamus
BC101978	7p22.1 (6990040-6990535)	n/a
AK092714 AX747686 AK127989	7p14.3 (32802082-32802580)	Neuroblastoma n/a Testis
BG333736, BE277865, BE274464, BE728363, BE727301, BQ422965 BE389761, BE391513	8p23.1 (9106846-9107034)	Skin/melanotic melanoma Uterus/endometrium, adenocarcinoma cell line
BE938057	10p13 (13343333-13343586)	Testis
AF315099	11p14.3 (25489246-25489554)	MCF7 (breast cancer cell line)
DA162968	11q13.1 (65768120-65768619)	Amygdala

mRNA or EST	Chromosome (position) of LTR2C	Organ / tissue type
AL137634 BI908469	11q13.2 (67776867-67777348)	Testis Leukocyte
AK024373	19p13.3 (1442427-1442895)	Placenta
CD695042	19q13.42 (53895441-53895506)	Normal nasopharynx
AK094771	21q22.12 (37481547-37482017)	Brain
EL950881	21q22.3 (46710070-46710255)	Eyeball
DA182815	Xq26.3 (134971892-134972375)	Amygdala

LTR2C- associated transcripts were analyzed by using UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly (http://genome.ucsc.edu/cgi-bin/hgGateway).

The identification of methylation sensitive genes in CD4+ T cells of SLE patients

We used RNA sequencing technique to identify the methylation sensitive transcript. Whole transcriptome libraries were constructed from untreated and 5-azacytidine treated CD4+ T cells. Then, we performed RNA-sequencing by using a SOLiD next generation sequencer. More than 80% of 20 million reads were mapped to Human genome sequencing (Ref. GRCh37/hg19) (Table 9). Based on the values of RPKM, 8,085 differentially expressed genes were found between the two libraries, which 5,516 genes were up-regulated and 2,569 down-regulated in 5-azacytidine treated CD4+ T cells compared with untreated. We performed a chi-square to validate the results with expression microarray data available from GEO, GSE17922 (130) Down-or up-regulated genes in 5-azacytidine treated CD4+ T cells were also found in T cell after 4 days of culture with 5-azacytidine, $P = 4.51 \times 10^{-15}$; OR = 1.79: 95% CI = 1.54-2.07 and $P = 1.29 \times 10^{-100}$; OR = 3.61: 95% CI = 3.19-4.09, respectively (Table 11). In contrast, there were no connection in T cell after 2 days of culture with 5-azacytidine (Table 10). Therefore,

detected of CD4+ T cells- methylation sensitive genes by RNA-Seq can also confirmed in published 5-azacytidine treated T cells by specific time point. Then, to identify methylation sensitive genes that differently expressed in CD4+ T cells of SLE patients. Overlap between down-or up-regulated gene in demethylated CD4+ T cells and in SLE CD4+ T cells were determined by using a chi-square test. Two microarray data sets that compared between CD4+ Tcells of SLE patients and controls were available in GEO. Genes which were down-regulated in 5-azacytidine treated T cells were found to preferentially have lower mRNA levels in the CD4+ T cells of SLE patient ($P = 5.37 \times 10^{-8}$; OR = 2.93: 95% CI = 1.95-4.40 for GSE10325 and $P = 9.09 \times 10^{-27}$; OR = 2.85: 95% CI = 2.83-3.48 for GSE4588,Table 12 and 13). Moreover, genes which were up-regulated in demethylated CD4+ T cells were also increased mRNA levels in the CD4+ T cells of SLE patient ($P = 1.52 \times 10^{-20}$; OR = 3.71: 95% CI = 2.76-4.99 for GSE10325 and $P = 4.12 \times 10^{-95}$; OR = 4.62: 95% CI = 3.95-5.41 for GSE4588,Table 12 and 13). Therefore, this supports the hypothesis that hypomethylation contribute different genes expression in CD4+ T cells of SLE patients.

Table 9. The statistical mapping summary of the RNA Seq

	Control (Untreated)	5-azacytidine treated
Total reads	21,121,594 (100%)	22,738,518 (100%)
Reads mapped	17,131,276 (81.1%)	18,683,076 (82.2%)
Reads filtered	2,399,373 (11.4%)	1,697,512 (7.5%)
Reads mapped, not filtered	14,732,818 (100%)	16,987,145 (100%)
Reads with too many mappings (N >= 10)	2,419,090 (16.4%)	2,555,487 (15%)
Reads with number of mappings in proper range (N < 10)	12,313,728 (83.6%)	14,431,658 (85.0%)
Reads uniquely aligned (score clear zone = 3)	9,874,040 (67%)	11,873,064 (69.9%)
Reads uniquely aligned to junctions	1,242,793 (8.4%)	1,511,209 (8.9%)
Reads uniquely aligned to known junctions	1,233,517 (8.4%)	1,500,501 (8.8%)

Table 10. The intersections between differential expression of 5-AzaC-treated (2 days) T cell (GSE17922) and 5-AzaC treated CD4+ T cell. The down-or up-regulation of a gene (denoted by "Up" and "Down") was determined by paired t-test (P-value threshold is set at 0.01). Meanwhile, the down-or up-regulated genes from RNA-seq (also denoted by "Up" and "Down") were determined by fold change value between the two libraries. The p-values of 2x2 tables were obtained from Chi-square distribution.

		GSE17922 5-AzaC treated T cell (2 days)			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Down	4	2,300	4	2,300
	Not-Down	29	17,816	27	17,818
		P value = 0.90		P value = 0.79	
		Odds ratio = 1.07		Odds ratio = 1.15	
		Lower 95% CI = 0.38		Lower 95% CI = 0.4	
		Upper 95% CI = 3.04		Upper 95% CI = 3.28	

		GSE17922 5-AzaC treated T cell (2 days)			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Up	10	5,218	7	5,221
	Not-Up	23	14,898	24	14,897
		P value = 0.57		P value = 0.67	
		Odds ratio = 1.24		Odds ratio = 0.83	
		Lower 95% CI = 0.59		Lower 95% CI = 0.36	
		Upper 95% CI = 2.61		Upper 95% CI = 1.93	

Table 11. The intersections between differential expression of 5-AzaC treated (4 days) T cell (GSE17922) and 5-AzaC treated CD4+ T cell. The down-or up-regulation of a gene (denoted by "Up" and "Down") was determined by paired t-test (P-value threshold is set at 0.01). Meanwhile, the down-or up-regulated genes from RNA-seq (also denoted by "Up" and "Down") were determined by fold change value between the two libraries. The p-values of 2x2 tables were obtained from Chi-square distribution.

		GSE17922 5-AzaC treated T cell (4 days)			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Down	239	2,065	149	2,155
	Not-Down	1,084	16,761	914	16,931
		P value = 4.51E-15		P value = 6.57E-03	
		Odds ratio = 1.79		Odds ratio = 1.28	
		Lower 95% CI = 1.54		Lower 95% CI = 1.07	
		Upper 95% CI = 2.07		Upper 95% CI = 1.53	

		GSE17922 5-AzaC treated T cell (4 days)			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Up	563	4,665	572	4,656
	Not-Up	760	14,161	491	14,430
		P value = 4.04E-46		P value = 1.29E-100	
		Odds ratio = 2.25		Odds ratio = 3.61	
		Lower 95% CI = 2.01		Lower 95% CI = 3.19	
		Upper 95% CI = 2.52		Upper 95% CI = 4.09	

Table 12. The intersections between differential expression of SLE CD4+ T cell vs. normal control (GSE10325) and 5-AzaC treated CD4+ T cell. The down-or up-regulation of a gene (denoted by "Up" and "Down") was determined by paired t-test (P-value threshold is set at 0.01). Meanwhile, the down-or up-regulated genes from RNA-seq (also denoted by "Up" and "Down") were determined by fold change value between the two libraries. The p-values of 2x2 tables were obtained from Chi-square distribution.

		GSE10325 CD4+ T cell SLE			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Down	34	1,676	6	1,704
	Not-Down	78	11,270	185	11,163
		P value = 5.37E-08		P value = 3.99E-05	
		Odds ratio = 2.93		Odds ratio = 0.21	
		Lower 95% CI = 1.95		Lower 95% CI = 0.09	
		Upper 95% CI = 4.40		Upper 95% CI = 0.48	

		GSE10325 CD4+ T cell SLE			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Up	40	4,107	120	4,027
	Not-Up	72	8,839	71	8,840
		P value = 0.37		P value = 1.52E-20	
		Odds ratio = 1.2		Odds ratio = 3.71	
		Lower 95% CI = 0.81		Lower 95% CI = 2.76	
		Upper 95% CI = 1.76		Upper 95% CI = 4.99	

Table 13. The intersections between differential expression of SLE CD4+ T cell vs. normal control (GSE4588) and 5-AzaC treated CD4+ T cell. The down-or up-regulation of a gene (denoted by "Up" and "Down") was determined by paired t-test (P-value threshold is set at 0.01). Meanwhile, the down-or up-regulated genes from RNA-seq (also denoted by "Up" and "Down") were determined by fold change value between the two libraries. The p-values of 2x2 tables were obtained from Chi-square distribution.

		GSE4588 CD4+ T cell SLE			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Down	136	2,130	79	2,187
	Not-Down	397	17,711	589	17,519
		P value = 9.09E-27		P value = 0.56	
		Odds ratio = 2.85		Odds ratio = 1.07	
		Lower 95% CI = 2.33		Lower 95% CI = 0.85	
		Upper 95% CI = 3.48		Upper 95% CI = 1.36	

		GSE4588 CD4+ T cell SLE			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Up	165	4,997	398	4,764
	Not-Up	368	14,844	270	14,942
		P value = 0.0025		P value = 4.12E-95	
		Odds ratio = 1.33		Odds ratio = 4.62	
		Lower 95% CI = 1.11		Lower 95% CI = 3.95	
		Upper 95% CI = 1.6		Upper 95% CI = 5.41	

Demethylation preferentially activates the Interferon signaling pathway in CD4+ T cells

To identify such functional networks, genes were defined as methylation sensitive genes in CD4+ T cells of SLE were performed gene set enrichment analysis using the online available DAVID Bioinformatics Resources. The annotation cluster with a group enrichment score less than or equal to 0.05 (1.3 in minus log scale) were shown in Table 17 and 18. From 398 up-regulation genes (Table 1A in appendix B), the highest-enrichment score was antiviral defense cluster (Table 15). Interestingly, many IFN-induced genes involved in the IFN signaling pathway and activation of IRF by Cytosolic Pattern Recognition Receptors pathway were present in this cluster (Figure 24 and 25) such as IFIT3, IFITM, IFI35, IRF7, IRF9, MX1, OAS1 and STAT1. Meanwhile, highest-enrichment score from performing by 136 down-regulation genes (Table 2A in appendix B) was protein biosynthesis as shown in Table 14 (example of related genes: RPL35A, EEF1B2, RPL13, SECISBP2, RPL37, EEF2, QARS, RPL24, RPS2, RPS6, RPS8, EIF4B, RPS16, RPL13A, RPL22, EIF3F, EIF3K, RPL3, RPL10, EIF3L, EEF1G, RPL5, EEF1D, UBA52). However, there was significant in immune-response term include T cell receptor signaling as shown in Figure 26 (related genes: PIK3RI, FYN, MALT1) and PI3K signaling in B lymphocyte as shown in Figure 27 (related genes: FYN, IRS2, PKCB, PIK3).

Previous studies reported that demethylated CD4+ T cells cause an autoreactive. Subsequently, the autoreactivity was found to be partly due to an overexpression and demethylation at promoter sequence, as observed in SLE patients (9-11) Together, these observations suggested that CD4+ T cell hypomethylation may contribute to the development of SLE by demethylation at promoter of several genes which contributed to disease pathogenesis. Therefore, we hypothesized that up-regulation of many IFN-induced genes in CD4+ T cells of SLE and demethylated CD4+ T cell may also effects of promoter hypomethylation. To demonstrate this hypothesis the down or up-regulated genes in 5-azacytidine treated CD4+ T cells and genome-wide promoter methylation array comparing between CD4+ T cells from SLE patients and normal controls, as GSE27895 (131) were analyzed by using a chi-square test (Table 16). An analysis revealed that 122 up-regulated genes in demethylated CD4+ T cell significantly found in gene which have promoter

hypomethylation in CD4+ T cells of SLE (P = 7.82 x 10^{-8}; OR = 1.87: 95% CI = 1.48-2.36) (Table 3A in appendix B). Our result supports the proposed mechanism by which global hypomethylation influences an overexpression of methylation sensitive gene by demethylation at promoter making T cells autoreactive in SLE (132) Unexpectedly, the functional annotation clustering of all 122 genes is less like to be IFN-induced genes. The highest-enrichment score is cellular component function (Table 17). Therefore, increase expression of IFN-induced genes in CD4+ T cells of SLE and demethylated CD4+ T cell may not seem to be regulation by promoter methylation.

Table 14. The annotation clusters identified for down-regulation genes by DAVID Functional Classification Tool

Group no.	Associated biology	Enrichment score
1	Protein biosynthesis	7.50
2	Translation factor activity	3.86
3	RNA-binding	1.49
4	Nucleolus	1.48
5	T cell activation	1.14
6	Zinc-finger	1.06

Table 15. The annotation clusters identified for up-regulation genes by DAVID Functional Classification Tool

Group no.	Associated biology	Enrichment score
1	Antiviral defense	10.37
2	Nucleotide-binding	5.71
3	Chromosomal protein	5.61
4	Nucleus	5.28
5	Ligase	3.81
6	CS-Like domain	3.61
7	Cell cycle	2.98
8	ncRNA metabolic process	2.71
9	DNA metabolic process	2.28
10	Nucleotidyltransferase	2.12
11	NAD+ ADP-ribosyltransferase activity	2.00
12	Chromosome, centromeric region	1.99
13	DNA replication	1.96
14	Purine biosynthesis	1.83
15	Macromolecular complex assembly	1.82
16	Mitochondrion	1.79
17	Iron ion binding	1.77
18	Chelation	1.57
19	Helicase activity	1.45
20	Zinc finger, PHD-type	1.35
21	Viral reproduction	1.34

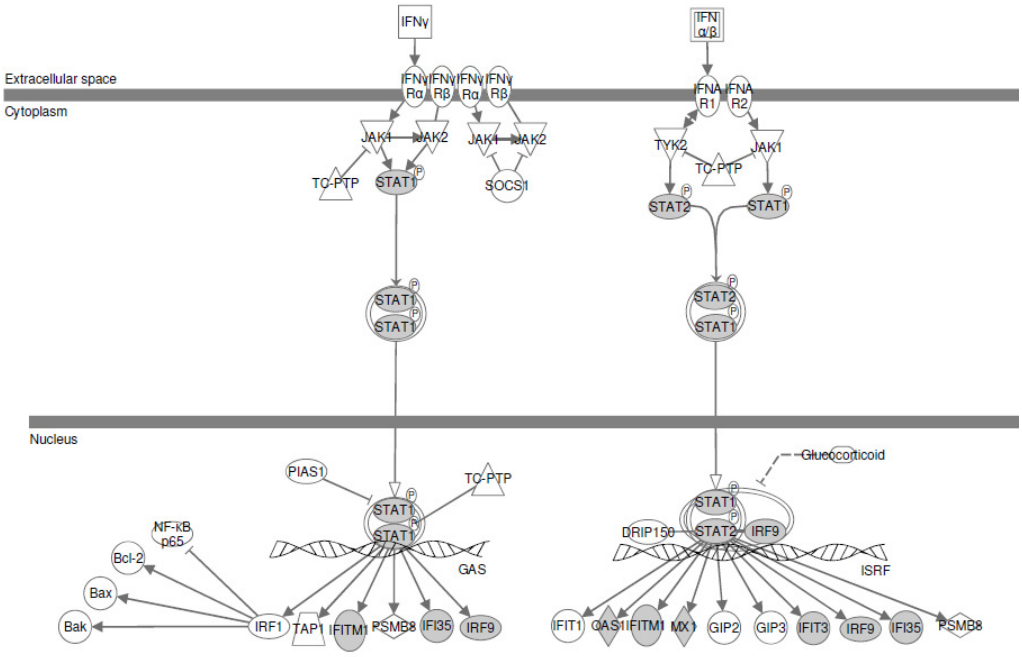


Figure 24. Interferon signaling pathway. Genes that are up-regulated are shown in dark fill.

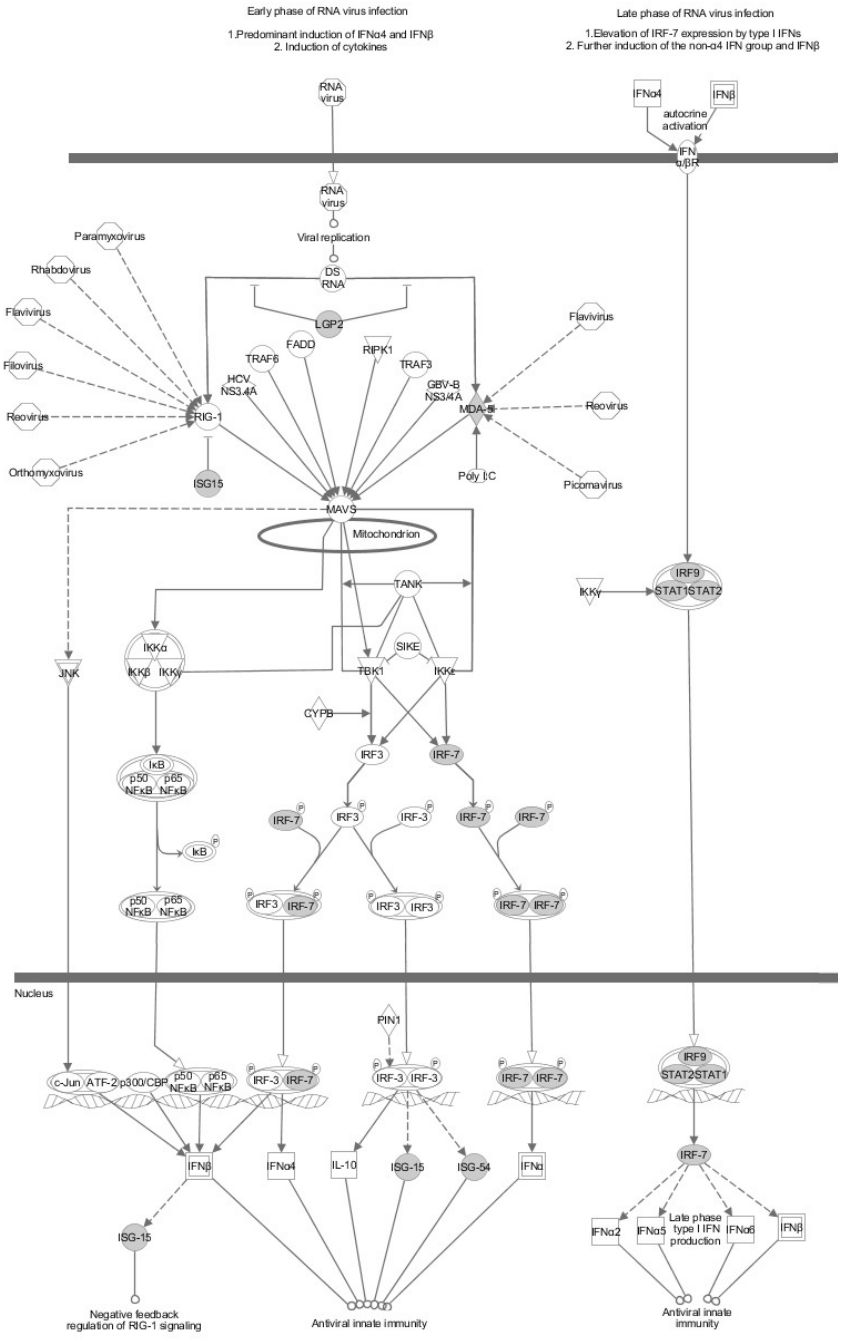


Figure 25. Activation of IRF by Cytosolic Pattern Recognition Receptors pathway. Genes that are up-regulated are shown in dark fill.

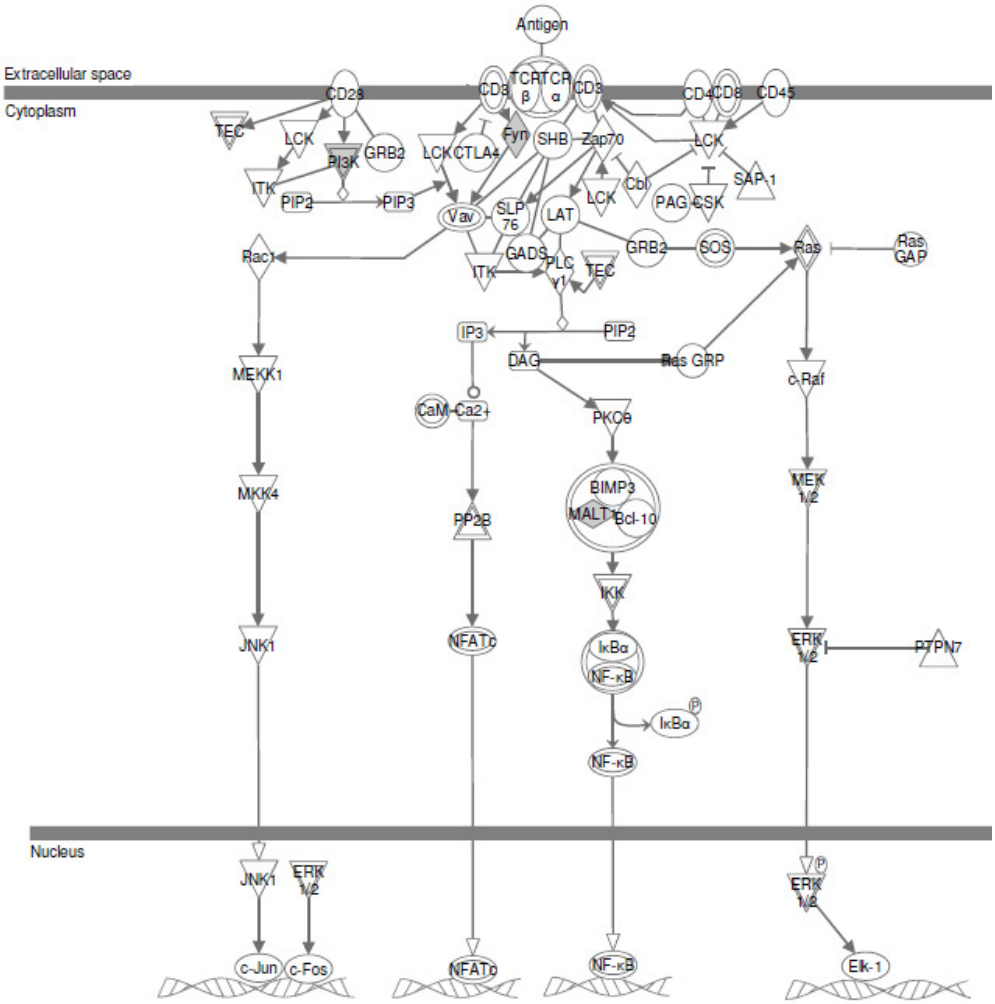


Figure 26. T cell receptor signaling. Genes that are down-regulated are shown in dark fill.

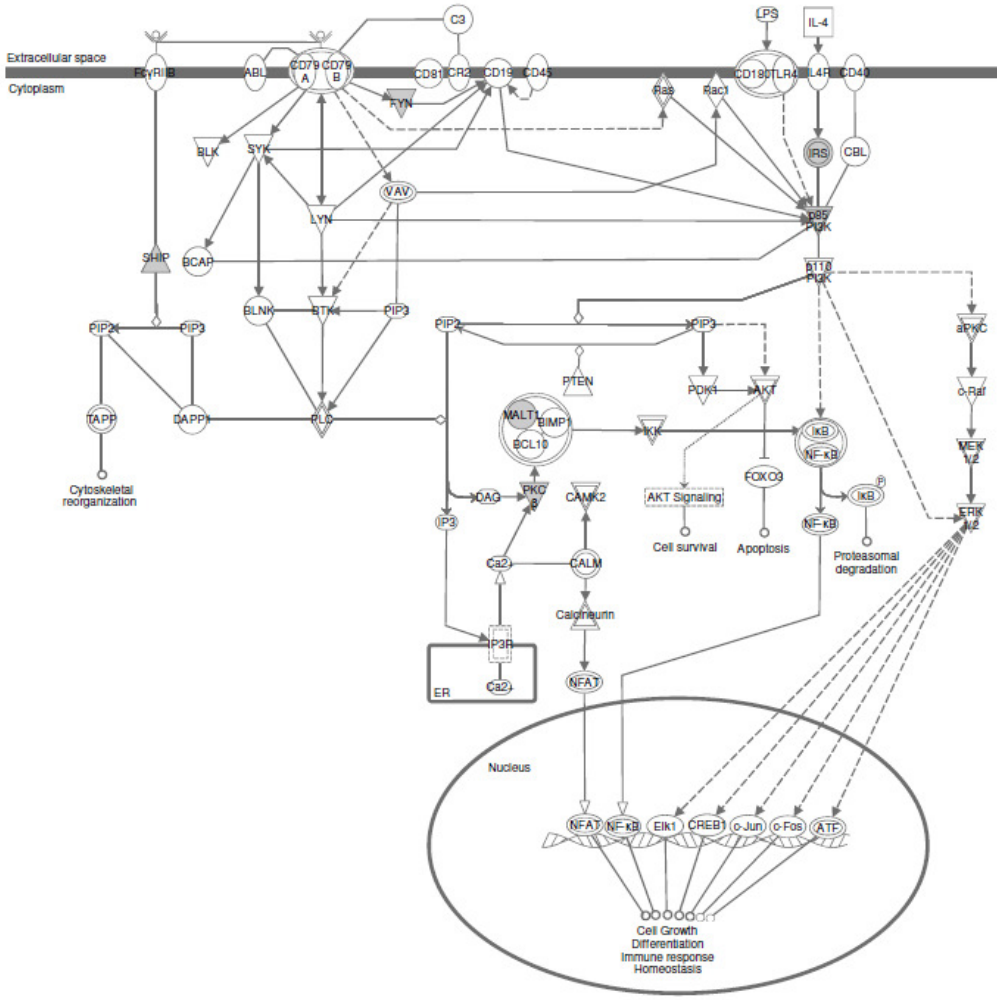


Figure 27. PI3K Signaling in B lymphocytes. Genes that are down-regulated are shown in dark fill.

Table 16. The intersections between promoter methylation array comparing between CD4+ T cells from SLE and normal controls (GSE27895) and 5-AzaC treated CD4+ T cell. Hyper or hypomethylation of promoter (denoted by "Up" and "Down") was also determined by paired t-test (P-value threshold is set at 0.01). Meanwhile, the down or up-regulated genes from RNA-seq (also denoted by "Up" and "Down") were determined by fold change value between the two libraries. The p-values of 2x2 tables were obtained from Chi-square distribution.

		GSE27895 CD4+ T cell SLE			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Down	37	1,531	23	1,545
	Not-Down	270	12,615	256	12,629
		P value = 0.493		P value = 0.158	
		Odds ratio = 1.13		Odds ratio = 0.73	
		Lower 95% CI = 0.8		Lower 95% CI = 0.48	
		Upper 95% CI = 1.6		Upper 95% CI = 1.13	

		GSE27895 CD4+ T cell SLE			
		Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
5-AzaC treated CD4+ T cell by RNA-Seq	Up	122	3,690	58	3,754
	Not-Up	185	10,456	221	10,420
		P value = 7.82E-08		P value = 0.0325	
		Odds ratio = 1.87		Odds ratio = 0.73	
		Lower 95% CI = 1.48		Lower 95% CI = 0.54	
		Upper 95% CI = 2.36		Upper 95% CI = 0.98	

Table 17. The annotation clusters identified for up-regulation genes in demethylated CD4+ T cell and hypomethylation at promoter in CD4+ T cells of SLE by DAVID Functional Classification Tool

Group no.	Associated biology	Enrichment score
1	Mitochondrion	2.22
2	Non-membrane-bounded organelle	2.05
3	Ubl conjugation	1.98
4	Cellular macromolecular complex subunit organization	1.91
5	Membrane-enclosed lumen	1.74
6	Regulation of cell proliferation and apoptosis	1.65
7	Cell cycle process	1.36

The role of LINE-1 in CD4+ T cell of SLE

LINE-1 is striking hypomethylated in CD4+ T cells of SLE. Recently, Aporntrwan et al. proved that the consequences of intragenic LINE-1 (within genes) hypomethylation in cancer cells is repression host genes via the nuclear RNA-induced silencing complex (RISC) (18) To investigate whether intragenic LINE-1 control host genes in SLE CD4+ T cell, gene expression in CD4+ T cells was compared with genes possessing intragenic LINE-1 and the rest. Genes processing LINE-1 were determined by L1base (133) and listed by Aporntrwan et al. (18) Table 18 demonstrated of chi-square tests of genes expression in SLE CD4+ T cell by using experiment GSE10325 and GSE4588. There is no statistic significant among 1,340 gene containing LINE-1 with up-or down-regulation genes from both experiments. Furthermore, a chi-square test of down-or up-regulation genes in both 5-azacytidine treated CD4+ T cells and SLE CD4+ T cell with intragenic LINE-1 were performed to explore whether intragenic LINE-1 can control the expression of sensitive genes in CD4+ T cells of SLE (Table 19). The result showed that the event of up- or down-regulation of genes in both azacytidine treated CD4+ T cells and SLE CD4+ T cell is more prevalent in genes without LINE-1 than in gene containing LINE-1 (Table 19). Therefore, hypomethylation intragenic LINE-1 represses transcription through a nuclear siRNA mediated *cis*-regulation element is unlikely to be a mechanism to control methylation sensitive genes in CD4+ T cells of SLE.

Table 18. 2x2 table of chi-square test to evaluate intragenic LINE-1 influence SLE CD4+ T cell expression by experiment GSE10325 and GSE4588. The down-or up-regulation of a gene (denoted by "Up" and "Down") was determined by paired t-test (P-value threshold is set at 0.01). Meanwhile, the genes containing or without intragenic LINE-1 were denoted by LINE-1 and No LINE-1, respectively. The p-values of 2x2 tables were obtained from Chi-square distribution.

	GSE10325 CD4+ T cell SLE			
	Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
LINE-1	9	912	19	902
No LINE-1	103	12,034	172	11,965
	P value = 0.683		P value = 0.116	
	Odds ratio = 1.15		Odds ratio = 1.47	
	Lower 95% CI = 0.58		Lower 95% CI = 0.91	
	Upper 95% CI = 2.29		Upper 95% CI = 2.36	

	GSE4588 CD4+ T cell SLE			
	Down (P<0.01)	Not-Down	Up (P<0.01)	Not-Up
LINE-1	41	1,301	36	1,306
No LINE-1	492	18,540	632	18,400
	P value = 0.297		P value = 0.205	
	Odds ratio = 1.19		Odds ratio = 0.80	
	Lower 95% CI = 0.86		Lower 95% CI = 0.57	
	Upper 95% CI = 1.64		Upper 95% CI = 1.13	

Table 19. 2x2 table of chi-square test to evaluate intragenic LINE-1 influence methylation sensitive genes in CD4+ T cells of SLE. The down-or up-regulation genes in both 5-azacytidine treated CD4+ T cells and SLE CD4+ T cell were denoted by Up and Down. Meanwhile, the genes containing or without intragenic LINE-1 were denoted by LINE-1 and No LINE-1, respectively. The p-values of 2x2 tables were obtained from Chi-square distribution.

	CD4+ T cells of SLE and 5-Aza-dC treatment			
	Down and Down	Not-Down and Not-Down	Up and Up	Not-Up and Not-Up
LINE-1	3	1,251	13	1,181
No LINE-1	133	16,460	385	13,761
	P value = 0.0414		P value = 0.00092	
	Odds ratio = 0.30		Odds ratio = 0.39	
	Lower 95% CI = 0.08		Lower 95% CI = 0.22	
	Upper 95% CI = 0.96		Upper 95% CI = 0.70	

	CD4+ T cells of SLE and 5-Aza-dC treatment			
	Up and Down	Not-Up and Not-Down	Down and Up	Not-Down and Not-Up
LINE-1	2	1,255	9	1,172
No LINE-1	77	16,264	156	13,672
	P value = 0.71		P value = 0.31	
	Odds ratio = 0.34		Odds ratio = 0.67	
	Lower 95% CI = 0.06		Lower 95% CI = 0.32	
	Upper 95% CI = 1.39		Upper 95% CI = 1.36	

CHAPTER V

DISCUSSION AND CONCLUSION

The global DNA hypomethylation of T lymphocytes in SLE patients have been previously demonstrated by using reverse-phase high performance liquid chromatography (HPLC) and an enzyme-linked immunosorbent assay-like reaction using 5-methylcytosine antibody (4-6, 66) These global aberrant in DNA methylation is not likely be explained by methylation level of certain genes. Although, a recent study in twins and sibling pairs discordant for SLE reported significantly decreased methylation of the 18S and 28S ribosomal genes, but have no significant difference of non-LTR retrotransposons including LINE-1 and Alu, satellite 2 repeats, NBL2 and D4Z4 non-satellite subtelomeric repeats methylation level in mixed WBCs (134) Many studies have suggested the potential role of IRSs to regulate cellular functions. Therefore, hypomethylation of IRSs may affect cellular gene expression and contribute to the pathogenesis and activity of SLE. Studies have shown that treating normal lymphocyte with 5-azacytidine can reactivate IRSs transcript (HERV) (31) and some IRS (LINE-1 and HERV-E transcripts) were detected in patients with SLE (105, 135) Our result confirms the up-regulation of HERV-E 4-1 in PBMCs of patients as previous reports. There was no significant difference in the HERV-K10 *gag* expression between two groups. PBMCs of rheumatoid arthritis patients showed a significantly higher level of HERV-K10 *gag* mRNA expression than normal controls (33) From our results indicated that the up-regulation of HERV-*gag* mRNA in PBMCs of SLE patients seems to be HERV class-specific. As a result of this, it was hypothesized that the up-regulation of LINE-1 or HERV was likely to have resulted from demethylation of these IRSs but there was no direct evidence to support this. First, this study aimed to investigate whether the methylation levels of LTR and also non-LTR retrotransposons in CD3+CD4+ T cells, CD3+CD8+ T cells and B cell differ between SLE and normal group. Second, we used the

comprehensive analysis of genome and expression to investigate the role of IRSs in CD4+ T cells of SLE patients.

COBRA-IRS

This study we developed COBRA method to examine methylation status of HERV-E LTR2C and HERV-K LTR5_Hs. We opted not to use pyrosequencing. The reason for this is because the COBRA-IRS can detect methylation of thousands of CpG loci by just using one set of conserved primers for each IRS. However, we were concerned that COBRA detects methylation of only one CpG dinucleotide. If there is a mutation at restriction site will be misinterpreted. So we have validated the COBRA method by comparing it to pyrosequencing technique. We found positive correlations between methylation levels of means and each CpG dinucleotides including *Tai*I CpG sites of COBRA-HERV-E LTR2C and HERV-K LTR5_Hs as previous reported in COBRA-LINE-1 and Alu (109) Moreover, positive correlations between methylation levels of means from pyrosequencing with COBRA HERV-E LTR2C and HERV-K LTR5_Hs were observed. Therefore, COBRA-IRS is very accurate and reliable technique. Moreover, a recent report showed that the methylation levels of LINE-1 loci can be divided into three groups based on the number of methylated and unmethylated CpG dinucleotides. The first group is hypermethylated. Second is partially methylated which cannot define by pyrosequecing, and third is hypomethylated (136) They found that the percentage of partial methylation in normal oral rinse was significantly higher than in head and neck squamous cell carcinoma cell line (136) However, only percent methylation of hypermethylation and hypomethylation in normal oral rinse and cancer patients were significantly different (136) Therefore, they suggested that the pattern of methylation is crucial for better understanding mechanisms and improving diagnostic applications of global hypomethylation in cancer. Our results showed that overall LINE-1 hypomethylation have also been found in lymphocyte of SLE (137) However, there is no report of increased cancer development risk in individuals with SLE. To investigate whether LINE-1 methylation pattern in CD3+CD4+ T cells, CD3+CD8+ T cells and B cell differ between SLE and normal group is necessary. It is important to

understand how the LINE-1 hypomethylation is associated with a variety of cellular phenotypes.

DNA methylation at HERV-LTRs is cell type-specific

Our experiment showed methylation patterns of HERV-E LTR2C and HERV-K LTR5_Hs different among normal cell-types. Interestingly, methylation of each LTR showed distinct pattern between B, T lymphocytes and normal ORE. To support results, recent study reported that there was lower global methylation of HERV-E LTR and HERV-K LTR in normal placenta than peripheral blood lymphocytes (21) In lymphocytes, we found that methylation patterns of HERV-E LTR2C and HERV-K LTR5_Hs differ between T and B cells. Global DNA methylations of lymphocytes have been previously examined by using several techniques. T cells showed higher of %deoxymethylcytosine content than T cell-depleted PBMC (non-T cells) by using reverse-phase high performance liquid chromatography (HPLC) (3.81±0.09 and 3.65±0.17) (66) There were differences of methylation indices between CD4+T cell and CD8+T cell or non-CD4+ T cell by an enzyme-linked immunosorbent assay-like reaction using 5-methylcytosine antibody (5, 6) Alteration of methylation in promoter regions and transposable elements can lead to change cellular gene expression. We suggest that dissimilarity of methylation levels influence to regulate cell-type specific genes expression. Differential gene expressions between CD8+, CD4+ T cells and B cell have been reported (138, 139) HERV-E LTR-derived promoters which function as alternative gene promoters were reported in pleiotrophin, endothelin B receptor and midline 1 genes (21) These LTRs are unmethylated in placenta but heavily methylated in peripheral blood lymphocytes (21) Moreover, HERV-E located at first exon of CD5 created alternative isoform, designed E1B, which is exclusively transcribed in B cell but did not make in T cell (97) Down regulation of DNA methyltransferase (DNMT) 1 in B cell leukemic can lead to un-regulated in proportion of E1A isoform /E1B isoform (140) This study demonstrated that patterns of methylation are difference between normal ORE and lymphocytes as previously reported (20) Interestingly, methylation patterns in normal cells depend on type of LTRs sequences. Thus, variation of HERV-LTRs methylation is cell-type

specific manners. This is may be an important regulation of cell development and differentiation.

Our results showed that there were different patterns of methylation and expression among cervical-derived cell lines. Variation in methylation levels of gene promoters between HeLa, SiHa and CaSki have been reported (141) Overall, leukemic cell lines were hypomethylated in HERV-LTRs. However, we found distinctiveness of hypomethylation in different LTRs sequences as obtained in Daudi cell. Therefore, hypomethylation in HERV sequence of leukemic cells is depend on cell type and LTR subclass. These results suggest that the distinct methylation patterns among cancer cells may be due to unique characteristics and process of cancer. To study the global HERV-LTRs and LTR-specific loci methylation in cancer are necessary in the future. These may be lead to understand the role of HERV in progression of cancer. In conclusion, this is the first report that i) methylation patterns of HERV-E LTR2C and HERV-K LTR5_Hs are difference among normal lymphocytes and ORE ii) cancer cells exhibit LTR-type specific hypomethylation.

Hypomethylation of SLE is IRSs and cell type specific

In LTR retroelements, our study clearly showed the HERV-E LTR2C hypomethylation in CD3+CD4+ T cells of active SLE. We were concerned that medications can be responsible for DNA hypomethylation. However, most of SLE patients in this study received steroid treatment. So, hypomethylation in SLE patients are not seemed to be effect of medications as previous described (5, 6) Moreover, when we analyzed the correlation between LTR methylation with clinical parameter, we found that the hypomethylation of HERV-E LTR2C was correlated with leucopenia and lymphopenia in active SLE. Although we cannot assess the association between HERV-E 4-1 with anti-U1 RNP and anti-Sm due to lack of these autoantibodies' result in our patients, the association of HERV-E LTR2C hypomethylation with leucopenia is very much interesting. One recent study also reported that leucopenia is linked to the methylation level (60) Wang GS et al found that the methylation levels in SLE patients with leucopenia were significantly lower than that in patients without leucopenia following UVB irradiation. It is possible that expression of

HERV-E antigen on the leukocyte surface might induce the autoantibodies against leukocytes or the hypomethylation of specific sequences can lead to an up-regulation of certain genes responsible for this phenomenon. Further studies are needed to confirm and explain this observation.

Surprisingly, our result showed that the inactive SLE group has hypomethylation of HERV-K LTR5_Hs in CD3+CD4+ T cells and B cells when compared with normal group. These LTR5 hypomethylation was also significantly correlated with SLEDAI score and complement activity. We hypothesize that the methylation level of certain gene areas might associated with specific clinical manifestation and even contribute as a disease protective effect. Besides, there is no correlation of DNA methylation levels among these two sequences. The association between HERV-E LTR2C and HERV-E LTR5_Hs proved that hypomethylation is not a generalized process. There may be specific mechanisms of demethylation for each LTR.

In summary, HERV-E LTR2C hypomethylation is found in CD3+CD4+ T cell of active SLE. Because the methylation levels of HERV-E LTR2C-APC2 and HERV-E LTR2C-VAV2 in lymphocyte subsets are positively correlated with genome wide level. So, these associations proved that mechanism of HERV-LTR2C hypomethylation seem to be a generalized process. However, only hypomethylation at HERV-E LTR2C-APC2 was observed in T cells of active SLE. Therefore, locus-specific methylation of HERV-E LTR2C can differentially be influenced by SLE pathogenesis depending on location of HERV-E LTR2C in the genome.

In non-LTR retroelements, our results showed that the hypomethylation in each lymphocyte subset of SLE is mainly found in LINE-1 rather than Alu. This finding also confirms the result previously reported (142) Although, in the PBMCs from the aging population, the hypomethylation of Alu and HERK-K were detected (109) It was hypothesized that the hypomethylation in aging might contribute to the pathogenesis of autoimmune diseases in elderly (143) However, there was no significant difference of Alu methylation in different lymphocyte subsets of SLE patients. This may be due to the

different pathogenesis seen in both groups. SLE, unlike in the aging population, predominantly develops in the younger age group. It is also possible that there may be other unknown factors aside from genetics and environment that may selectively influence certain IRS type. However, when we looked at the results obtained from the active, inactive SLE groups and normal controls, we noticed that Alu hypomethylation was significantly prominent in CD8+ lymphocytes of the inactive SLE group. Future studies are warranted to confirm the existence of Alu hypomethylation in CD8+ T lymphocytes of patients with inactive SLE as well as investigate the functional consequences of Alu methylation.

In contrast with the results observed from the SLE twin discordant study, (65) we detected significant hypomethylation changes in LINE-1 repeats in CD4+, CD8+ T and B lymphocytes subsets from patients with SLE when compared to normal controls. This discrepancy may be attributed to the WBCs that they had used to detect hypomethylation whereas in our study, we had separated our cells into CD4+, CD8+ T and B lymphocytes. Even though we were not able to detect Alu hypomethylation, but we were able to observe LINE-1 hypomethylation in the active group for both CD4+ and CD8+ T lymphocytes. This methylation was significantly distinct when compared to normal controls. Our result corroborates previous reports that T lymphocytes from active lupus have a reduced number in the total deoxymethylcytosine content (66) This observation is correlated with the decreased DNMT transcripts and decreased ERK phosphorylation in T lymphocytes upon stimulation (66, 74) Our results suggest that the methylation of LINE-1 may contribute to the global hypomethylation in T lymphocytes of patients with active SLE.

The role of IRSs methylation in SLE

Previous studies reported that demethylated CD4+ T cells cause an autoreactive. Subsequently, the autoreactivity was found to be partly due to an overexpression and demethylation at promoter sequence, as observed in SLE patients and lupus mice model (9-11) Together, these observations suggested that CD4+ T cell hypomethylation may contribute to the development of SLE by demethylation at promoter of several genes which contributed to disease pathogenesis. By using comprehensive analysis of expression and

promoter array database, we identified methylation sensitive genes which have promoter hypomethylation in CD4+ T cells of SLE. Our result supports the proposed mechanism by which global hypomethylation influences an overexpression of methylation sensitive gene by demethylation at promoter making T cells autoreactive in SLE (132) Many of IFN-induced genes showed up-regulation in methylation sensitive gene in CD4+ T cells of SLE. Interestingly, interferon-regulated genes were also up-regulated in both of the spleen of a dominant negative MEK transgenic mice model which decreased expression of Dnmt1 lead to lupus-like disease and PBMCs of SLE patients (76, 144) Nevertheless, there were no promoter hypomethylation of IFN-induced genes in CD4+ T cells of SLE. Therefore, the effect of promoter hypomethylation on certain genes overexpression may be the one of mechanisms that promotes SLE. However, the consequence of global hypomethylation in SLE development is largely unknown.

A LINE-1 Hypomethylation is common in most cancer. The consequences of intragenic LINE-1 hypomethylation in cancer cells is repression host genes via the nuclear RNA-induced silencing complex (RISC) (18) Whereas, it is unlikely to be a mechanism to control methylation sensitive genes in CD4+ T cells of SLE. Even though LINE-1 hypomethylation was also observed in SLE but the consequence may be different. Therefore, it is reasonable to hypothesize that there are other mechanisms can occur from the hypomethylation of LINE1. Interestingly, it was recently suggested that the increased expression of LINE-1 transcripts in lupus patients could be the inducer of IFN type I (135) Our results demonstrated event that SLE have LINE-1 hypomethylation and up-regulation of IFN-induced gene. Therefore, future studies are in the works to look at LINE-1 transcripts in correlation with LINE-1 hypomethylation and IFN-induced genes expression.

However, we did not rule out the role of IRSs methylation regulates gene expression in *cis*. Interestingly, recent study showed that SLE patients have demethylation in HERV-E alternative promoter of CD5 gene and can lead to up-regulation of CD5-E1B isoform in SLE B cells (36) Since alteration of methylation in IRSs might have an effect in *cis* e.g, controlling neighboring gene expression or alternative splicing, or in *trans* e.g, controlling distant genes or encoding protein product that can act as antigens, it is possible that some

specific retroelement might contribute to pathogenesis of SLE. Therefore, the demethylation may lead to up-regulation of repetitive sequence-related antigens directly. Evidence of molecular mimicry of HERV and autoantigen in SLE patient has been reported. For example, Hishikawa et al. demonstrated that approximate 50% of SLE patients have antibodies to recombinant Gag protein from HERV-E 4-1sequence, but not in normal control serum (145) Moreover, the correlation of HERV-E 4-1 *gag* mRNA with blood plasma concentrations of anti-U1 RNP and anti-Sm antibodies in SLE patients was found (32) Interestingly, this finding is consistent with previous finding that HERV-E 4-1 transcript was un-regulated in PBMCs of SLE patients (31, 32)

In conclusion, the hypomethylation in each lymphocyte subset of SLE was interspersed repetitive sequences (IRSs) specific, mainly found in LINE-1 rather than Alu. Meanwhile, the hypomethylation of HERV sequences in each lymphocyte subset of SLE is LTR-type specific. From comprehensive analysis of genome and gene expression in CD4+ T cells, we proposed consequence mechanism of LINE-1 hypomethylation in SLE that LINE-1 transcripts might have an effect in *trans* to up-regulate the IFN-induced gene. Further study to find out the functional role of IRSs hypomethylation might lead to the discovery of novel pathogenesis pathway in SLE.

REFERENCES

[1] Nguyen, C., Limaye, N., Wakeland, E.K. Susceptibility genes in the pathogenesis of murine lupus. Arthritis Res 4 Suppl 3 (2002): S255-263.

[2] Rhodes, B., Vyse, T.J. The genetics of SLE: an update in the light of genome-wide association studies. Rheumatology (Oxford) 47 (Nov 2008): 1603-1611.

[3] Rodenhiser, D., Mann, M. Epigenetics and human disease: translating basic biology into clinical applications. Cmaj 174 (Jan 31 2006): 341-348.

[4] Balada, E., Ordi-Ros, J., Serrano-Acedo, S., Martinez-Lostao, L., Vilardell-Tarres, M. Transcript overexpression of the MBD2 and MBD4 genes in CD4+ T cells from systemic lupus erythematosus patients. J Leukoc Biol 81 (Jun 2007): 1609-1616.

[5] Balada, E., et al. Transcript levels of DNA methyltransferases DNMT1, DNMT3A and DNMT3B in CD4+ T cells from patients with systemic lupus erythematosus. Immunology 124 (Jul 2008): 339-347.

[6] Luo, Y., et al. Abnormal DNA methylation in T cells from patients with subacute cutaneous lupus erythematosus. Br J Dermatol 159 (Sep 2008): 827-833.

[7] Quddus, J., et al. Treating activated CD4+ T cells with either of two distinct DNA methyltransferase inhibitors, 5-azacytidine or procainamide, is sufficient to cause a lupus-like disease in syngeneic mice. J Clin Invest 92 (Jul 1993): 38-53.

[8] Yung, R., Chang, S., Hemati, N., Johnson, K., Richardson, B. Mechanisms of drug-induced lupus. IV. Comparison of procainamide and hydralazine with analogs in vitro and in vivo. Arthritis Rheum 40 (Aug 1997): 1436-1443.

[9] Yung, R., et al. Mechanisms of drug-induced lupus. II. T cells overexpressing lymphocyte function-associated antigen 1 become autoreactive and cause a lupuslike disease in syngeneic mice. J Clin Invest 97 (Jun 15 1996): 2866-2871.

[10] Kaplan, M.J., Lu, Q., Wu, A., Attwood, J., Richardson, B. Demethylation of promoter regulatory elements contributes to perforin overexpression in CD4+ lupus T cells. J Immunol 172 (Mar 15 2004): 3652-3661.

[11] Lu, Q., Wu, A., Richardson, B.C. Demethylation of the same promoter sequence increases CD70 expression in lupus T cells and T cells treated with lupus-inducing drugs. J Immunol 174 (May 15 2005): 6212-6219.

[12] Lu, Q., et al. Demethylation of CD40LG on the inactive X in T cells from women with lupus. J Immunol 179 (Nov 1 2007): 6352-6358.

[13] Sugimura, T., Ushijima, T. Genetic and epigenetic alterations in carcinogenesis. Mutat Res 462 (Apr 2000): 235-246.

[14] Kaneda, A., et al. Frequent hypomethylation in multiple promoter CpG islands is associated with global hypomethylation, but not with frequent promoter hypermethylation. Cancer Sci 95 (Jan 2004): 58-64.

[15] Bannert, N., Kurth, R. Retroelements and the human genome: new perspectives on an old relation. Proc Natl Acad Sci U S A 101 Suppl 2 (Oct 5 2004): 14572-14579.

[16] Levy, S., et al. The diploid genome sequence of an individual human. PLoS Biol 5 (Sep 4 2007): e254.

[17] Jordan, I.K., Rogozin, I.B., Glazko, G.V., Koonin, E.V. Origin of a substantial fraction of human regulatory sequences from transposable elements. Trends Genet 19 (Feb 2003): 68-72.

[18] Aporntewan, C., et al. Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2. PLoS ONE 6 (2011): e17934.

[19] Cordaux, R., Batzer, M.A. The impact of retrotransposons on human genome evolution. Nat Rev Genet 10 (Oct 2009): 691-703.

[20] Phokaew, C., Kowudtitham, S., Subbalekha, K., Shuangshoti, S., Mutirangura, A. LINE-1 methylation patterns of different loci in normal and cancerous cells. Nucleic Acids Res 36 (Oct 2008): 5704-5712.

[21] Reiss, D., Zhang, Y., Mager, D.L. Widely variable endogenous retroviral methylation levels in human placenta. Nucleic Acids Res 35 (2007): 4743-4754.

[22] Chalitchagorn, K., et al. Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. Oncogene 23 (Nov 18 2004): 8841-8846.

[23] Migeon, B.R., Holland, M.M., Driscoll, D.J., Robinson, J.C. Programmed demethylation in CpG islands during human fetal development. Somat Cell Mol Genet 17 (Mar 1991): 159-168.

[24] Kremenskoy, M., et al. Genome-wide analysis of DNA methylation status of CpG islands in embryoid bodies, teratomas, and fetuses. Biochem Biophys Res Commun 311 (Nov 28 2003): 884-890.

[25] Gonzalo, S. Epigenetic alterations in aging. J Appl Physiol 109 (Aug 2010): 586-597.

[26] Wang, L., et al. Relation between hypomethylation of long interspersed nucleotide elements and risk of neural tube defects. Am J Clin Nutr 91 (May 2010): 1359-1367.

[27] Kitkumthorn, N., Mutirangura, A. Long interspersed nuclear element-1 hypomethylation in cancer: biology and clinical applications. Clinical Epigenetics 2 (Aug 2011): 315-330.

[28] Saito, K., et al. Long interspersed nuclear element 1 hypomethylation is a marker of poor prognosis in stage IA non-small cell lung cancer. Clin Cancer Res 16 (Apr 15 2010): 2418-2426.

[29] Pattamadilok, J., et al. LINE-1 hypomethylation level as a potential prognostic factor for epithelial ovarian cancer. Int J Gynecol Cancer 18 (Aug 2008): 711-717.

[30] Jintaridth, P., Mutirangura, A. Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. Physiol Genomics 41 (Feb 9 2010): 194-200.

[31] Okada, M., et al. Role of DNA methylation in transcription of human endogenous retrovirus in the pathogenesis of systemic lupus erythematosus. J Rheumatol 29 (Aug 2002): 1678-1682.

[32] Piotrowski, P.C., Duriagin, S., Jagodzinski, P.P. Expression of human endogenous retrovirus clone 4-1 may correlate with blood plasma concentration of anti-U1 RNP and anti-Sm nuclear antibodies. Clin Rheumatol 24 (Nov 2005): 620-624.

[33] Ejtehadi, H.D., et al. The potential role of human endogenous retrovirus K10 in the pathogenesis of rheumatoid arthritis: a preliminary study. Ann Rheum Dis 65 (May 2006): 612-616.

[34] Lavie, L., Kitova, M., Maldener, E., Meese, E., Mayer, J. CpG methylation directly regulates transcriptional activity of the human endogenous retrovirus family HERV-K(HML-2). J Virol 79 (Jan 2005): 876-883.

[35] Gotzinger, N., Sauter, M., Roemer, K., Mueller-Lantzsch, N. Regulation of human endogenous retrovirus-K Gag expression in teratocarcinoma cell lines and human tumours. J Gen Virol 77 (Dec 1996): 2983-2990.

[36] Garaud, S., et al. IL-6 modulates CD5 expression in B cells from patients with lupus by regulating DNA methylation. J Immunol 182 (May 1 2009): 5623-5632.

[37] Tan, E.M., et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 25 (Nov 1982): 1271-1277.

[38] Hochberg, M.C. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 40 (Sep 1997): 1725.

[39] Pons-Estel, G.J., Alarcon, G.S., Scofield, L., Reinlib, L., Cooper, G.S. Understanding the epidemiology and progression of systemic lupus erythematosus. Semin Arthritis Rheum 39 (Feb 2010): 257-268.

[40] Lau, C.S., Yin, G., Mok, M.Y. Ethnic and geographical differences in systemic lupus erythematosus: an overview. Lupus 15 (2006): 715-719.

[41] Alarcon-Segovia, D., et al. Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. Arthritis Rheum 52 (Apr 2005): 1138-1147.

[42] Deapen, D., et al. A revised estimate of twin concordance in systemic lupus erythematosus. Arthritis Rheum 35 (Mar 1992): 311-318.

[43] Block, S.R., Winfield, J.B., Lockshin, M.D., D'Angelo, W.A., Christian, C.L. Studies of twins with systemic lupus erythematosus. A review of the literature and presentation of 12 additional sets. Am J Med 59 (Oct 1975): 533-552.

[44] Harley, J.B., et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 40 (Feb 2008): 204-210.

[45] Hom, G., et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med 358 (Feb 28 2008): 900-909.

[46] Kozyrev, S.V., et al. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet 40 (Feb 2008): 211-216.

[47] Musone, S.L., et al. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat Genet 40 (Aug 1 2008): 1062-1064.

[48] Oishi, T., et al. A functional SNP in the NKX2.5-binding site of ITPR3 promoter is associated with susceptibility to systemic lupus erythematosus in Japanese population. J Hum Genet 53 (2008): 151-162.

[49] Kamatani, Y., et al. Identification of a significant association of a single nucleotide polymorphism in TNXB with systemic lupus erythematosus in a Japanese population. J Hum Genet 53 (2008): 64-73.

[50] Kaufman, K.M., Kirby, M.Y., Harley, J.B., James, J.A. Peptide mimics of a major lupus epitope of SmB/B'. Ann N Y Acad Sci 987 (Apr 2003): 215-229.

[51] McClain, M.T., et al. Early events in lupus humoral autoimmunity suggest initiation through molecular mimicry. Nat Med 11 (Jan 2005): 85-89.

[52] Pender, M.P. Infection of autoreactive B lymphocytes with EBV, causing chronic autoimmune diseases. Trends Immunol 24 (Nov 2003): 584-588.

[53] Harley, J.B., Harley, I.T., Guthridge, J.M., James, J.A. The curiously suspicious: a role for Epstein-Barr virus in lupus. Lupus 15 (2006): 768-777.

[54] Sarzi-Puttini, P., Atzeni, F., Iaccarino, L., Doria, A. Environment and systemic lupus erythematosus: an overview. Autoimmunity 38 (Nov 2005): 465-472.

[55] Wysenbeek, A.J., Block, D.A., Fries, J.F. Prevalence and expression of photosensitivity in systemic lupus erythematosus. Ann Rheum Dis 48 (Jun 1989): 461-463.

[56] Marzano, A.V., Vezzoli, P., Crosti, C. Drug-induced lupus: an update on its dermatologic aspects. Lupus 18 (Oct 2009): 935-940.

[57] Parks, C.G., Cooper, G.S. Occupational exposures and risk of systemic lupus erythematosus: a review of the evidence and exposure assessment methods in population- and clinic-based studies. Lupus 15 (2006): 728-736.

[58] Prete, P.E. The mechanism of action of L-canavanine in inducing autoimmune phenomena. Arthritis Rheum 28 (Oct 1985): 1198-1200.

[59] Costenbader, K.H., Karlson, E.W. Cigarette smoking and systemic lupus erythematosus: a smoking gun? Autoimmunity 38 (Nov 2005): 541-547.

[60] Wang, G.S., et al. Ultraviolet B exposure of peripheral blood mononuclear cells of patients with systemic lupus erythematosus inhibits DNA methylation. Lupus 18 (Oct 2009): 1037-1044.

[61] Takai, D., Jones, P.A. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A 99 (Mar 19 2002): 3740-3745.

[62] Issa, J.P. CpG island methylator phenotype in cancer. Nat Rev Cancer 4 (Dec 2004): 988-993.

[63] Chedin, F., Lieber, M.R., Hsieh, C.L. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proc Natl Acad Sci U S A 99 (Dec 24 2002): 16916-16921.

[64] Robertson, K.D., Wolffe, A.P. DNA methylation in health and disease. Nat Rev Genet 1 (Oct 2000): 11-19.

[65] Javierre, B.M., et al. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Res 20 (Feb 2010): 170-179.

[66] Richardson, B., et al. Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis. Arthritis Rheum 33 (Nov 1990): 1665-1673.

[67] Richardson, B. Effect of an inhibitor of DNA methylation on T cells. II. 5-Azacytidine induces self-reactivity in antigen-specific T4+ cells. Hum Immunol 17 (Dec 1986): 456-470.

[68] Cornacchia, E., et al. Hydralazine and procainamide inhibit T cell DNA methylation and induce autoreactivity. J Immunol 140 (Apr 1 1988): 2197-2200.

[69] Richardson, B., et al. N-acetylprocainamide is a less potent inducer of T cell autoreactivity than procainamide. Arthritis Rheum 31 (Aug 1988): 995-999.

[70] Richardson, B., Powers, D., Hooper, F., Yung, R.L., O'Rourke, K. Lymphocyte function-associated antigen 1 overexpression and T cell autoreactivity. Arthritis Rheum 37 (Sep 1994): 1363-1372.

[71] Lu, Q., et al. Demethylation of ITGAL (CD11a) regulatory sequences in systemic lupus erythematosus. Arthritis Rheum 46 (May 2002): 1282-1291.

[72] Deng, C., Yang, J., Scott, J., Hanash, S., Richardson, B.C. Role of the ras-MAPK signaling pathway in the DNA methyltransferase response to DNA hypomethylation. Biol Chem 379 (Sep 1998): 1113-1120.

[73] Deng, C., et al. Hydralazine may induce autoimmunity by inhibiting extracellular signal-regulated kinase pathway signaling. Arthritis Rheum 48 (Mar 2003): 746-756.

[74] Deng, C., et al. Decreased Ras-mitogen-activated protein kinase signaling may cause DNA hypomethylation in T lymphocytes from lupus patients. Arthritis Rheum 44 (Feb 2001): 397-407.

[75] Gorelik, G., Fang, J.Y., Wu, A., Sawalha, A.H., Richardson, B. Impaired T cell protein kinase C delta activation decreases ERK pathway signaling in idiopathic and hydralazine-induced lupus. J Immunol 179 (Oct 15 2007): 5553-5563.

[76] Sawalha, A.H., et al. Defective T-cell ERK signaling induces interferon-regulated gene expression and overexpression of methylation-sensitive genes similar to lupus patients. Genes Immun 9 (Jun 2008): 368-378.

[77] Kazazian, H.H., Jr. Mobile elements: drivers of genome evolution. Science 303 (Mar 12 2004): 1626-1632.

[78] Khodosevich, K., Lebedev, Y., Sverdlov, E. Endogenous retroviruses and human evolution. Comp Funct Genomics 3 (2002): 494-498.

[79] Conley, A.B., Piriyapongsa, J., Jordan, I.K. Retroviral promoters in the human genome. Bioinformatics 24 (Jul 15 2008): 1563-1567.

[80] Nelson, P.N., et al. Demystified. Human endogenous retroviruses. Mol Pathol 56 (Feb 2003): 11-18.

[81] Moyes, D., Griffiths, D.J., Venables, P.J. Insertional polymorphisms: a new lease of life for endogenous retroviruses in human disease. Trends Genet 23 (Jul 2007): 326-333.

[82] Gifford, R., Tristem, M. The evolution, distribution and diversity of endogenous retroviruses. Virus Genes 26 (May 2003): 291-315.

[83] Bock, M., Stoye, J.P. Endogenous retroviruses and the human germline. Curr Opin Genet Dev 10 (Dec 2000): 651-655.

[84] Lander, E.S., et al. Initial sequencing and analysis of the human genome. Nature 409 (Feb 15 2001): 860-921.

[85] Stoye, J.P. Endogenous retroviruses: still active after all these years? Curr Biol 11 (Nov 13 2001): R914-916.

[86] Buzdin, A. Human-specific endogenous retroviruses. ScientificWorldJournal 7 (2007): 1848-1868.

[87] Blomberg, J., Benachenhou, F., Blikstad, V., Sperber, G., Mayer, J. Classification and nomenclature of endogenous retroviral sequences (ERVs): problems and recommendations. Gene 448 (Dec 15 2009): 115-123.

[88] Ting, C.N., Rosenberg, M.P., Snow, C.M., Samuelson, L.C., Meisler, M.H. Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. Genes Dev 6 (Aug 1992): 1457-1465.

[89] Piriyapongsa, J., Polavarapu, N., Borodovsky, M., McDonald, J. Exonization of the LTR transposable elements in human genome. BMC Genomics 8 (2007): 291.

[90] Baust, C., Seifarth, W., Germaier, H., Hehlmann, R., Leib-Mosch, C. HERV-K-T47D-Related long terminal repeats mediate polyadenylation of cellular transcripts. Genomics 66 (May 15 2000): 98-103.

[91] Mager, D.L., Hunter, D.G., Schertzer, M., Freeman, J.D. Endogenous retroviruses provide the primary polyadenylation signal for two new human genes (HHLA2 and HHLA3). Genomics 59 (Aug 1 1999): 255-263.

[92] Cohen, C.J., Lock, W.M., Mager, D.L. Endogenous retroviral LTRs as promoters for human genes: a critical assessment. Gene 448 (Dec 15 2009): 105-114.

[93] Samuelson, L.C., Phillips, R.S., Swanberg, L.J. Amylase gene structures in primates: retroposon insertions and promoter evolution. Mol Biol Evol 13 (Jul 1996): 767-779.

[94] Medstrand, P., Landry, J.R., Mager, D.L. Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. J Biol Chem 276 (Jan 19 2001): 1896-1903.

[95] Landry, J.R., Rouhi, A., Medstrand, P., Mager, D.L. The Opitz syndrome gene Mid1 is transcribed from a human endogenous retroviral promoter. Mol Biol Evol 19 (Nov 2002): 1934-1942.

[96] Schulte, A.M., et al. Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus. Proc Natl Acad Sci U S A 93 (Dec 10 1996): 14759-14764.

[97] Renaudineau, Y., Hillion, S., Saraux, A., Mageed, R.A., Youinou, P. An alternative exon 1 of the CD5 gene regulates CD5 expression in human B lymphocytes. Blood 106 (Oct 15 2005): 2781-2789.

[98] Renaudineau, Y., et al. Characterization of the human CD5 endogenous retrovirus-E in B lymphocytes. Genes Immun 6 (Dec 2005): 663-671.

[99] Dunn, C.A., Romanish, M.T., Gutierrez, L.E., van de Lagemaat, L.N., Mager, D.L. Transcription of two human genes from a bidirectional endogenous retrovirus promoter. Gene 366 (Feb 1 2006): 335-342.

[100] Faulkner, G.J., et al. The regulated retrotransposon transcriptome of mammalian cells. Nat Genet 41 (May 2009): 563-571.

[101] Huh, J.W., et al. Transcriptional regulation of GSDML gene by antisense-oriented HERV-H LTR element. Arch Virol 153 (2008): 1201-1205.

[102] Gogvadze, E., Stukacheva, E., Buzdin, A., Sverdlov, E. Human-specific modulation of transcriptional activity provided by endogenous retroviral insertions. J Virol 83 (Jun 2009): 6098-6105.

[103] van de Lagemaat, L.N., Landry, J.R., Mager, D.L., Medstrand, P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. Trends Genet 19 (Oct 2003): 530-536.

[104] Guntaka, R.V. Transcription termination and polyadenylation in retroviruses. Microbiol Rev 57 (Sep 1993): 511-521.

[105] Ogasawara, H., et al. Quantitative comparison of human endogenous retrovirus mRNA between SLE and rheumatoid arthritis. Lupus 10 (2001): 517-518.

[106] Florl, A.R., Lower, R., Schmitz-Drager, B.J., Schulz, W.A. DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. Br J Cancer 80 (Jul 1999): 1312-1321.

[107] Menendez, L., Benigno, B.B., McDonald, J.F. L1 and HERV-W retrotransposons are hypomethylated in human ovarian carcinomas. Mol Cancer 3 (Apr 26 2004): 12.

[108] Gimenez, J., et al. Custom human endogenous retroviruses dedicated microarray identifies self-induced HERV-W family elements reactivated in testicular cancer upon methylation control. Nucleic Acids Res 38 (Apr 1 2010): 2229-2246.

[109] Jintaridth, P., Mutirangura, A. Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. Physiol Genomics 41 (Feb 9 2010): 194-200

[110] Gladman, D.D., Ibanez, D., Urowitz, M.B. Systemic lupus erythematosus disease activity index 2000. J Rheumatol 29 (Feb 2002): 288-291.

[111] Uribe, A.G., et al. The Systemic Lupus Activity Measure-revised, the Mexican Systemic Lupus Erythematosus Disease Activity Index (SLEDAI), and a modified SLEDAI-2K are adequate instruments to measure disease activity in systemic lupus erythematosus. J Rheumatol 31 (Oct 2004): 1934-1940.

[112] Miller, S.A., Dykes, D.D., Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16 (Feb 11 1988): 1215.

[113] Avihingsanon, Y., et al. Measurement of urinary chemokine and growth factor messenger RNAs: a noninvasive monitoring in lupus nephritis. Kidney Int 69 (Feb 2006): 747-753.

[114] Depil, S., Roche, C., Dussart, P., Prin, L. Expression of a human endogenous retrovirus, HERV-K, in the blood cells of leukemia patients. Leukemia 16 (Feb 2002): 254-259.

[115] Lu, Q., et al. DNA methylation and chromatin structure regulate T cell perforin gene expression. J Immunol 170 (May 15 2003): 5124-5132.

[116] Oelke, K., et al. Overexpression of CD70 and overstimulation of IgG synthesis by lupus T cells and T cells treated with DNA methylation inhibitors. Arthritis Rheum 50 (Jun 2004): 1850-1860.

[117] Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5 (Jul 2008): 621-628.

[118] Aporntewan, C., Mutirangura, A. Connection up- and down-regulation expression analysis of microarrays (CU-DREAM): a physiogenomic discovery tool. Asian Biomedicine 5 (Apr 2011): 257-262.

[119] Edgar, R., Domrachev, M., Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30 (Jan 1 2002): 207-210.

[120] Barrett, T., et al. NCBI GEO: archive for high-throughput functional genomic data. Nucleic Acids Res 37 (Jan 2009): D885-890.

[121] Huang da, W., et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 8 (2007): R183.

[122] Huang da, W., Sherman, B.T., Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4 (2009): 44-57.

[123] Tusnady, G.E., Simon, I., Varadi, A., Aranyi, T. BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. Nucleic Acids Res 33 (2005): e9.

[124] Aranyi, T., Varadi, A., Simon, I., Tusnady, G.E. The BiSearch web server. BMC Bioinformatics 7 (2006): 431.

[125] Kent, W.J., et al. The human genome browser at UCSC. Genome Res 12 (Jun 2002): 996-1006.

[126] Jurka, J. Repbase update: a database and an electronic journal of repetitive elements. Trends Genet 16 (Sep 2000): 418-420.

[127] Taruscio, D., Floridia, G., Zoraqi, G.K., Mantovani, A., Falbo, V. Organization and integration sites in the human genome of endogenous retroviral sequences belonging to HERV-E family. Mamm Genome 13 (Apr 2002): 216-222.

[128] Belshaw, R., et al. Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. J Virol 79 (Oct 2005): 12507-12514.

[129] van de Lagemaat, L.N., Medstrand, P., Mager, D.L. Multiple effects govern endogenous retrovirus survival patterns in human gene introns. Genome Biol 7 (2006): R86.

[130] Sanchez-Abarca, L.I., et al. Immunomodulatory effect of 5-azacytidine (5-azaC): potential role in the transplantation setting. Blood 115 (Jan 7 2010): 107-121.

[131] Jeffries, M.A., et al. Genome-wide DNA methylation patterns in CD4+ T cells from patients with systemic lupus erythematosus. Epigenetics 6 (May 2011): 593-601.

[132] Gorelik, G., Richardson, B. Aberrant T cell ERK pathway signaling and chromatin structure in lupus. Autoimmun Rev 8 (Jan 2009): 196-198.

[133] Penzkofer, T., Dandekar, T., Zemojtel, T. L1Base: from functional annotation to prediction of active LINE-1 elements. Nucleic Acids Res 33 (Jan 1 2005): D498-500.

[134] Javierre, B.M., et al. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Res (Dec 22 2009):

[135] Crow, M.K. Long interspersed nuclear elements (LINE-1): potential triggers of systemic autoimmune disease. Autoimmunity 43 (Feb 7-16.

[136] Pobsook, T., Subbalekha, K., Sannikorn, P., Mutirangura, A. Improved measurement of LINE-1 sequence methylation for cancer detection. Clin Chim Acta 412 (Jan 30 2011): 314-321.

[137] Nakkuntod, J., Avihingsanon, Y., Mutirangura, A., Hirankarn, N. Hypomethylation of LINE-1 but not Alu in lymphocyte subsets of systemic lupus erythematosus patients. Clin Chim Acta 412 (Jul 15 2011): 1457-1461.

[138] Palmer, C., Diehn, M., Alizadeh, A.A., Brown, P.O. Cell-type specific gene expression profiles of leukocytes in human peripheral blood. BMC Genomics 7 (2006): 115.

[139] Wang, M., Windgassen, D., Papoutsakis, E.T. Comparative analysis of transcriptional profiling of CD3+, CD4+ and CD8+ T cells identifies novel immune response players in T-cell activation. BMC Genomics 9 (2008): 225.

[140] Garaud, S., et al. Selection of the alternative exon 1 from the cd5 gene down-regulates membrane level of the protein in B lymphocytes. J Immunol 181 (Aug 1 2008): 2010-2018.

[141] Sova, P., et al. Discovery of novel methylation biomarkers in cervical carcinoma by global demethylation and microarray analysis. Cancer Epidemiol Biomarkers Prev 15 (Jan 2006): 114-123.

[142] Huck, S., Deveaud, E., Namane, A., Zouali, M. Abnormal DNA methylation and deoxycytosine-deoxyguanine content in nucleosomes from lymphocytes undergoing apoptosis. Faseb J 13 (Aug 1999): 1415-1422.

[143] Yung, R.L., Julius, A. Epigenetics, aging, and autoimmunity. Autoimmunity 41 (May 2008): 329-335.

[144] Arasappan, D., Tong, W., Mummaneni, P., Fang, H., Amur, S. Meta-analysis of microarray data using a pathway-based approach identifies a 37-gene expression signature for systemic lupus erythematosus in human peripheral blood mononuclear cells. BMC Med 9 (2011): 65.

[145] Hishikawa, T., et al. Detection of antibodies to a recombinant gag protein derived from human endogenous retrovirus clone 4-1 in autoimmune diseases. Viral Immunol 10 (1997): 137-147.

APPENDICES

APPENDIX A

BUFFER AND REAGENT

1. Nuclei Lysis Buffer (NLB)

1 M Tris (pH 8.0)	10	ml
5 M NaCl	0.5	ml
0.5 M EDTA (pH 8.0)	0.4	ml

Adjust volume to 100 ml with distilled water. Adjust pH to 7.2. Keep refrigerated. Shelf life is approximately 6 months.

2. 1 M Tris

Tris base	12.11	g
Distilled water	100	g

Adjust volume to 100 ml with distilled water. Adjust pH to 7.2. The solution was mixed and sterilizes by autoclaving at 121 ºC for 15 minutes.

3. 5 M NaCl

NaCl	29.22	g
Distilled water	100	ml

Adjust volume to 100 ml with distilled water. The solution was mixed and sterilized by autoclaving at 121 ºC for 15 minutes.

4. EDTA

EDTA	37.22	g
Distilled water	200	ml

Adjust volume to 200 ml with distilled water. The solution was mixed and sterilizes by autoclaving at 121 ºC for 15 minutes. Keep refrigerated.

5. 5.3 M NaCl

NaCl	15.5	g
Distilled water	50	ml

Adjust volume to 50 ml with distilled water. The solution was mixed and sterilizes by autoclaving at 121 ºC for 15 minutes. Keep refrigerated.

6. Proteinase K 10 mg/ml

Proteinase K	100	mg
Distilled water	10	ml

Mix the solution and store at -20 ºC

7. 10% SDS

SDS	10	g
Distilled water	100	ml

Adjust volume to 100 ml with distilled water. The solution was mixed and sterilizes by autoclaving at 121 ºC for 15 minutes.

8. 10X Tris-boric acid buffer (TBE) 1000 ml

Tris base	108.0	g
Boric acid	55	g
0.5 m EDTA Ph 8.0	40	ml

Adjust volume to 1 liter with distilled water. The solution was mixed and sterilized by autoclaving at 121 ºC for 15 min.

APPENDIX B

Table 1A. Up-regulated genes in demethylated CD4+ T cells were also increased mRNA levels in the CD4+ T cells of SLE patient

Gene symbol	Full Name
AARS	alanyl-tRNA synthetase
ACTR3	ARP3 actin-related protein 3 homolog (yeast)
ADA	adenosine deaminase
ADAR	adenosine deaminase, RNA-specific
AGRN	agrin
AHSA1	AHA1, activator of heat shock 90kDa protein ATPase homolog 1 (yeast)
AIM2	absent in melanoma 2
ALAS1	aminolevulinate, delta-, synthase 1
APBA3	amyloid beta (A4) precursor protein-binding, family A, member 3
APTX	aprataxin
ARPC5L	actin related protein 2/3 complex, subunit 5-like
ARV1	ARV1 homolog (S. cerevisiae)
ASF1B	ASF1 anti-silencing function 1 homolog B (S. cerevisiae)
ASH2L	ash2 (absent, small, or homeotic)-like (Drosophila)
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase
ATP5C1	ATP synthase, H+ transporting, mitochondrial F1 complex, gamma polypeptide 1
ATP6V0E1	ATPase, H+ transporting, lysosomal 9kDa, V0 subunit e1
ATP6V1C1	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C1
AURKA	aurora kinase A; aurora kinase A pseudogene 1
AURKB	aurora kinase B
AZI2	5-azacytidine induced 2
B3GNT2	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 1;

	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2
BANF1	similar to barrier-to-autointegration factor; barrier to autointegration factor 1
BATF	basic leucine zipper transcription factor, ATF-like
BCL2L14	BCL2-like 14 (apoptosis facilitator)
BIRC5	baculoviral IAP repeat-containing 5
BLM	Bloom syndrome, RecQ helicase-like
BST2	NPC-A-7; bone marrow stromal cell antigen 2
BTN2A2	butyrophilin, subfamily 2, member A2
C11orf31	chromosome 11 open reading frame 31
C11orf48	chromosome 11 open reading frame 48
C11orf58	chromosome 11 open reading frame 58 pseudogene; chromosome 11 open reading frame 58
C11orf83	chromosome 11 open reading frame 83
C12orf44	chromosome 12 open reading frame 44
C14orf156	chromosome 14 open reading frame 156
C16orf75	chromosome 16 open reading frame 75
C16orf88	chromosome 16 open reading frame 88
C19orf12	chromosome 19 open reading frame 12
C19orf66	chromosome 19 open reading frame 66
C1GALT1C1	C1GALT1-specific chaperone 1
C1orf144	chromosome 1 open reading frame 144
C1orf25	chromosome 1 open reading frame 25
C1orf31	chromosome 1 open reading frame 31
C20orf177	chromosome 20 open reading frame 177
C3orf34	chromosome 3 open reading frame 34
C6orf125	chromosome 6 open reading frame 125
C8orf55	chromosome 8 open reading frame 55

CALCOCO2	calcium binding and coiled-coil domain 2
CARS	cysteinyl-tRNA synthetase
CASP7	caspase 7, apoptosis-related cysteine peptidase
CBFB	core-binding factor, beta subunit
CCDC28B	coiled-coil domain containing 28B
CCDC50	coiled-coil domain containing 50
CCDC51	coiled-coil domain containing 51
CCDC90B	coiled-coil domain containing 90B
CCNB2	cyclin B2
CCNK	cyclin K
CCT5	chaperonin containing TCP1, subunit 5 (epsilon)
CD59	CD59 molecule, complement regulatory protein
CDC20	cell division cycle 20 homolog (S. cerevisiae)
CDC7	cell division cycle 7 homolog (S. cerevisiae)
CDCA7	cell division cycle associated 7
CDCA8	cell division cycle associated 8
CDKN3	cyclin-dependent kinase inhibitor 3
CENPA	centromere protein A
CENPN	centromere protein N
CEP55	centrosomal protein 55kDa
CFLAR	CASP8 and FADD-like apoptosis regulator
CHCHD1	coiled-coil-helix-coiled-coil-helix domain containing 1
CHCHD8	coiled-coil-helix-coiled-coil-helix domain containing 8
CHD4	chromodomain helicase DNA binding protein 4
CHI3L2	chitinase 3-like 2
CHMP5	chromatin modifying protein 5
CISD1	CDGSH iron sulfur domain 1
CISD2	CDGSH iron sulfur domain 2

CISD3	CDGSH iron sulfur domain 3
CLDND1	claudin domain containing 1
CLPP	ClpP caseinolytic peptidase, ATP-dependent, proteolytic subunit homolog (E. coli)
CLPX	ClpX caseinolytic peptidase X homolog (E. coli)
CMPK2	cytidine monophosphate (UMP-CMP) kinase 2, mitochondrial
CNDP2	CNDP dipeptidase 2 (metallopeptidase M20 family)
CNP	2',3'-cyclic nucleotide 3' phosphodiesterase
COPB2	coatomer protein complex, subunit beta 2 (beta prime)
COQ4	coenzyme Q4 homolog (S. cerevisiae)
COQ9	coenzyme Q9 homolog (S. cerevisiae)
COX4NB	COX4 neighbor
COX5A	cytochrome c oxidase subunit Va
CPSF4	cleavage and polyadenylation specific factor 4, 30kDa
CRIPT	cysteine-rich PDZ-binding protein
DARS	aspartyl-tRNA synthetase
DCTPP1	dCTP pyrophosphatase 1
DDX39	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39
DDX51	DEAD (Asp-Glu-Ala-Asp) box polypeptide 51
DFFA	DNA fragmentation factor, 45kDa, alpha polypeptide
DGUOK	deoxyguanosine kinase
DHX58	DEXH (Asp-Glu-X-His) box polypeptide 58
DPP3	dipeptidyl-peptidase 3
DPY30	dpy-30 homolog (C. elegans)
DR1	down-regulator of transcription 1, TBP-binding (negative cofactor 2)
DRAP1	DR1-associated protein 1 (negative cofactor 2 alpha)
DRG1	developmentally regulated GTP binding protein 1
DTL	denticleless homolog (Drosophila)

DTX3L	deltex 3-like (Drosophila)
EBAG9	estrogen receptor binding site associated, antigen, 9
EBP	emopamil binding protein (sterol isomerase)
ECHDC1	enoyl Coenzyme A hydratase domain containing 1
EIF2AK2	eukaryotic translation initiation factor 2-alpha kinase 2
EIF2S1	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa
EIF3J	eukaryotic translation initiation factor 3, subunit J
EIF6	eukaryotic translation initiation factor 6
ELAC2	elaC homolog 2 (E. coli)
ELP3	elongation protein 3 homolog (S. cerevisiae)
ENO1	enolase 1, (alpha)
EPSTI1	epithelial stromal interaction 1 (breast)
ERGIC2	ERGIC and golgi 2
ESPL1	extra spindle pole bodies homolog 1 (S. cerevisiae)
ETFA	electron-transfer-flavoprotein, alpha polypeptide
ETNK1	ethanolamine kinase 1
EXOSC4	exosome component 4
EXOSC5	exosome component 5
EZH2	enhancer of zeste homolog 2 (Drosophila)
FAM100B	family with sequence similarity 100, member B
FAM20B	family with sequence similarity 20, member B
FAM48A	family with sequence similarity 48, member A
FAM96B	family with sequence similarity 96, member B
FAM98A	family with sequence similarity 98, member A
FANCA	Fanconi anemia, complementation group A
FANCG	Fanconi anemia, complementation group G
FARSA	phenylalanyl-tRNA synthetase, alpha subunit
FBXO22	FBXO22 opposite strand (non-protein coding); F-box protein 22

FBXO6	F-box protein 6
FCF1	FCF1 small subunit (SSU) processome component homolog (S. cerevisiae)
FH	fumarate hydratase
FTSJ2	FtsJ homolog 2 (E. coli)
FTSJD2	FtsJ methyltransferase domain containing 2
GARS	glycyl-tRNA synthetase
GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase
GBP1	guanylate binding protein 1, interferon-inducible, 67kDa
GBP4	guanylate binding protein 4
GINS2	GINS complex subunit 2 (Psf2 homolog)
GNG8	guanine nucleotide binding protein (G protein), gamma 8
GNPDA1	glucosamine-6-phosphate deaminase 1
GPATCH4	G patch domain containing 4
GPR171	G protein-coupled receptor 171
GRSF1	G-rich RNA sequence binding factor 1
GSDMD	gasdermin D
H1F0	H1 histone family, member 0
HERC6	hect domain and RLD 6
HIST1H2AC	histone cluster 1, H2ac
HIST1H2AG	histone cluster 1, H2ag; histone cluster 1, H2ah; histone cluster 1, H2ai; histone cluster 1, H2ak; histone cluster 1, H2al; histone cluster 1, H2am
HIST1H2AH	histone cluster 1, H2ag; histone cluster 1, H2ah; histone cluster 1, H2ai; histone cluster 1, H2ak; histone cluster 1, H2al; histone cluster 1, H2am

HIST1H2AJ	histone cluster 1, H2aj
HIST1H2BD	histone cluster 1, H2bd
HIST1H2BE	histone cluster 1, H2bi; histone cluster 1, H2bg; histone cluster 1, H2be; histone cluster 1, H2bf; histone cluster 1, H2bc
HIST1H2BF	histone cluster 1, H2bi; histone cluster 1, H2bg; histone cluster 1, H2be; histone cluster 1, H2bf; histone cluster 1, H2bc
HIST1H2BK	histone cluster 1, H2bk
HK1	hexokinase 1
HMGN4	high mobility group nucleosomal binding domain 4
HMMR	hyaluronan-mediated motility receptor (RHAMM)
HN1	hematological and neurological expressed 1
HNRNPC	heterogeneous nuclear ribonucleoprotein C (C1/C2)
HRSP12	heat-responsive protein 12
HSH2D	hematopoietic SH2 domain containing
IARS	isoleucyl-tRNA synthetase
IDH2	isocitrate dehydrogenase 2 (NADP+), mitochondrial
IFI16	interferon, gamma-inducible protein 16
IFI27	interferon, alpha-inducible protein 27
IFI35	interferon-induced protein 35
IFI44	interferon-induced protein 44
IFI44L	interferon-induced protein 44-like
IFIH1	interferon induced with helicase C domain 1
IFIT2	interferon-induced protein with tetratricopeptide repeats 2
IFIT3	interferon-induced protein with tetratricopeptide repeats 3
IFIT5	interferon-induced protein with tetratricopeptide repeats 5
IFITM1	interferon induced transmembrane protein 1 (9-27)
IFITM3	interferon induced transmembrane protein 3 (1-8U)
IKZF4	IKAROS family zinc finger 4 (Eos)

IL12RB2	interleukin 12 receptor, beta 2
IMMT	inner membrane protein, mitochondrial (mitofilin)
IPO4	importin 4
IRF7	interferon regulatory factor 7
IRF9	interferon regulatory factor 9
ISG15	ISG15 ubiquitin-like modifier
ISG20	interferon stimulated exonuclease gene 20kDa
KIAA0040	KIAA0040
KIAA0101	KIAA0101
KIF11	kinesin family member 11
KLHDC7B	kelch domain containing 7B
KPNB1	karyopherin (importin) beta 1
LAG3	lymphocyte-activation gene 3
LAGE3	L antigen family, member 3
LAMP3	lysosomal-associated membrane protein 3
LDLR	low density lipoprotein receptor
LGALS3BP	lectin, galactoside-binding, soluble, 3 binding protein
LGALS9	lectin, galactoside-binding, soluble, 9
LMAN2	lectin, mannose-binding 2
LOC100144603	hypothetical transcript
LXN	latexin
LY6E	lymphocyte antigen 6 complex, locus E
LYAR	Ly1 antibody reactive homolog (mouse)
LYPLAL1	lysophospholipase-like 1
LZTS2	leucine zipper, putative tumor suppressor 2
MARS	methionyl-tRNA synthetase
MCCC2	methylcrotonoyl-Coenzyme A carboxylase 2 (beta)
MCM3	minichromosome maintenance complex component 3

MCM7	minichromosome maintenance complex component 7
MCTS1	malignant T cell amplified sequence 1
MDH1	malate dehydrogenase 1, NAD (soluble)
MELK	maternal embryonic leucine zipper kinase
MITD1	MIT, microtubule interacting and transport, domain containing 1
MKI67	antigen identified by monoclonal antibody Ki-67
MOV10	Mov10, Moloney leukemia virus 10, homolog (mouse)
MRM1	mitochondrial rRNA methyltransferase 1 homolog (S. cerevisiae)
MRPL4	mitochondrial ribosomal protein L4
MRPL42	mitochondrial ribosomal protein L42
MRPL51	mitochondrial ribosomal protein L51
MRPS12	mitochondrial ribosomal protein S12
MRPS16	mitochondrial ribosomal protein S16
MRPS31	mitochondrial ribosomal protein S31
MT1F	metallothionein 1F
MT1X	metallothionein 1X
MT2A	metallothionein 2A
MTHFD1	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase
MX1	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)
MX2	myxovirus (influenza virus) resistance 2 (mouse)
N4BP1	NEDD4 binding protein 1
NAPA	N-ethylmaleimide-sensitive factor attachment protein, alpha
NCAPD2	non-SMC condensin I complex, subunit D2
NDUFA8	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 8, 19kDa
NDUFA9	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 9, 39kDa

NDUFAF1	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, assembly factor 1
NMI	N-myc (and STAT) interactor
NMT1	N-myristoyltransferase 1
NOLC1	nucleolar and coiled-body phosphoprotein 1
NOP10	NOP10 ribonucleoprotein homolog (yeast)
NRBP1	nuclear receptor binding protein 1
NUB1	negative regulator of ubiquitin-like proteins 1
NUDC	nuclear distribution gene C homolog (A. nidulans)
NUDCD1	NudC domain containing 1
NUDCD2	NudC domain containing 2
NUDCD3	NudC domain containing 3
NUP43	nucleoporin 43kDa
NUSAP1	nucleolar and spindle associated protein 1
OAS1	2',5'-oligoadenylate synthetase 1, 40/46kDa
OAS2	2'-5'-oligoadenylate synthetase 2, 69/71kDa
OAS3	2'-5'-oligoadenylate synthetase 3, 100kDa
OASL	2'-5'-oligoadenylate synthetase-like
OGFOD1	2-oxoglutarate and iron-dependent oxygenase domain containing 1
OIP5	Opa interacting protein 5
ORMDL2	ORM1-like 2 (S. cerevisiae)
PA2G4	proliferation-associated 2G4, 38kDa; proliferation-associated 2G4 pseudogene 4
PAICS	phosphoribosylaminoimidazole carboxylase, phosphoribosylaminoimidazole succinocarboxamide synthetase
PARP10	poly (ADP-ribose) polymerase family, member 10
PARP11	poly (ADP-ribose) polymerase family, member 11
PARP12	poly (ADP-ribose) polymerase family, member 12

PARP14	poly (ADP-ribose) polymerase family, member 14
PARP9	poly (ADP-ribose) polymerase family, member 9
PCK2	phosphoenolpyruvate carboxykinase 2 (mitochondrial)
PDCD2L	programmed cell death 2-like
PDHA1	pyruvate dehydrogenase (lipoamide) alpha 1
PDIA4	protein disulfide isomerase family A, member 4
PDLIM7	PDZ and LIM domain 7 (enigma)
PEF1	penta-EF-hand domain containing 1
PEX3	peroxisomal biogenesis factor 3
PHF11	PHD finger protein 11
PHF19	PHD finger protein 19
PHF20L1	PHD finger protein 20-like 1
PHF23	PHD finger protein 23
PIH1D1	PIH1 domain containing 1
PIM1	pim-1 oncogene
PLAC8	placenta-specific 8
PLSCR1	phospholipid scramblase 1
PML	promyelocytic leukemia; similar to promyelocytic leukemia protein isoform 1
PNPO	pyridoxamine 5'-phosphate oxidase
PNPT1	polyribonucleotide nucleotidyltransferase 1
POLDIP3	polymerase (DNA-directed), delta interacting protein 3
POLR3K	polymerase (RNA) III (DNA directed) polypeptide K, 12.3 kDa
PPIL1	peptidylprolyl isomerase (cyclophilin)-like 1
PPIL5	peptidylprolyl isomerase (cyclophilin)-like 5
PPP2R5A	protein phosphatase 2, regulatory subunit B', alpha isoform
PRDX4	peroxiredoxin 4
PREB	prolactin regulatory element binding

PRIC285	peroxisomal proliferator-activated receptor A interacting complex 285
PRIM1	primase, DNA, polypeptide 1 (49kDa)
PRMT5	protein arginine methyltransferase 5
PRPS2	phosphoribosyl pyrophosphate synthetase 2
PSAT1	chromosome 8 open reading frame 62; phosphoserine aminotransferase 1
PSMA4	proteasome (prosome, macropain) subunit, alpha type, 4
PSMA5	proteasome (prosome, macropain) subunit, alpha type, 5
PSMA7	proteasome (prosome, macropain) subunit, alpha type, 7
PSMB10	proteasome (prosome, macropain) subunit, beta type, 10
PSMB2	proteasome (prosome, macropain) subunit, beta type, 2
PSMD13	proteasome (prosome, macropain) 26S subunit, non-ATPase, 13
PSMD14	proteasome (prosome, macropain) 26S subunit, non-ATPase, 14
PSME2	proteasome (prosome, macropain) activator subunit 2 (PA28 beta)
PTGIR	prostaglandin I2 (prostacyclin) receptor (IP)
PTPLAD1	protein tyrosine phosphatase-like A domain containing 1
RAB11A	RAB11A, member RAS oncogene family
RAB7L1	RAB7, member RAS oncogene family-like 1
RACGAP1	Rac GTPase activating protein 1 pseudogene; Rac GTPase activating protein 1
RASGRP3	RAS guanyl releasing protein 3 (calcium and DAG-regulated)
RBX1	ring-box 1
RDH11	retinol dehydrogenase 11 (all-trans/9-cis/11-cis)
REC8	REC8 homolog (yeast)
REXO2	REX2, RNA exonuclease 2 homolog (S. cerevisiae)
RFC3	replication factor C (activator 1) 3, 38kDa
RNASEH2A	ribonuclease H2, subunit A
RNF170	ring finger protein 170

RNF213	ring finger protein 213
RPE	rcRPE; ribulose-5-phosphate-3-epimerase
RPS6KB2	ribosomal protein S6 kinase, 70kDa, polypeptide 2
RRAGA	Ras-related GTP binding A
RRM2	ribonucleotide reductase M2 polypeptide
RSAD2	radical S-adenosyl methionine domain containing 2
RSC1A1	regulatory solute carrier protein, family 1, member 1
RTP4	receptor (chemosensory) transporter protein 4
RUVBL2	RuvB-like 2 (E. coli)
SAMD9	sterile alpha motif domain containing 9
SAMD9L	sterile alpha motif domain containing 9-like
SBNO2	strawberry notch homolog 2 (Drosophila)
SCCPDH	saccharopine dehydrogenase (putative)
SEC22B	SEC22 vesicle trafficking protein homolog B (S. cerevisiae)
SERBP1	SERPINE1 mRNA binding protein 1
SF3B3	splicing factor 3b, subunit 3, 130kDa
SFXN4	sideroflexin 4
SLAMF7	SLAM family member 7
SLC38A5	solute carrier family 38, member 5
SLC39A14	solute carrier family 39 (zinc transporter), member 14
SMC1A	structural maintenance of chromosomes 1A
SNX3	sorting nexin 3
SNX6	sorting nexin 6
SOCS2	suppressor of cytokine signaling 2
SP100	SP100 nuclear antigen
SP110	SP110 nuclear body protein
SPG11	spastic paraplegia 11 (autosomal recessive)
SQRDL	sulfide quinone reductase-like (yeast)

SRM	spermidine synthase
SRR	serine racemase
SSB	Sjogren syndrome antigen B (autoantigen La)
STAG3	stromal antigen 3
STAMBP	STAM binding protein
STAMBPL1	STAM binding protein-like 1
STARD3NL	STARD3 N-terminal like
STARD7	StAR-related lipid transfer (START) domain containing 7
STAT1	signal transducer and activator of transcription 1, 91kDa
STAT2	signal transducer and activator of transcription 2, 113kDa
STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)
STIP1	stress-induced-phosphoprotein 1
STMN1	stathmin 1
STX4	syntaxin 4
SUB1	SUB1 homolog (S. cerevisiae)
SUCLG1	succinate-CoA ligase, alpha subunit
SUV39H2	suppressor of variegation 3-9 homolog 2 (Drosophila)
SYNCRIP	synaptotagmin binding, cytoplasmic RNA interacting protein
TAPBPL	TAP binding protein-like
TARS	threonyl-tRNA synthetase
TCF3	transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)
TDRD7	tudor domain containing 7
THOC5	THO complex 5
TIMM13	translocase of inner mitochondrial membrane 13 homolog (yeast)
TIMM17B	translocase of inner mitochondrial membrane 17 homolog B (yeast)
TINF2	TERF1 (TRF1)-interacting nuclear factor 2

TK1	thymidine kinase 1, soluble
TMEM126B	transmembrane protein 126B
TMEM173	transmembrane protein 173
TMOD3	tropomodulin 3 (ubiquitous)
TNFSF10	tumor necrosis factor (ligand) superfamily, member 10
TOR3A	torsin family 3, member A
TRAFD1	TRAF-type zinc finger domain containing 1
TRIM14	tripartite motif-containing 14
TRIM21	tripartite motif-containing 21
TRIM22	tripartite motif-containing 22
TRIM25	tripartite motif-containing 25
TRIM5	tripartite motif-containing 5
TSPAN31	tetraspanin 31
TUBG1	tubulin, gamma 1; similar to Tubulin, gamma 1
TXNDC5	thioredoxin domain containing 5 (endoplasmic reticulum); muted homolog (mouse)
TXNDC9	thioredoxin domain containing 9
TYMS	thymidylate synthetase
UBA1	ubiquitin-like modifier activating enzyme 1
UBA3	ubiquitin-like modifier activating enzyme 3
UBE2E1	ubiquitin-conjugating enzyme E2E 1 (UBC4/5 homolog, yeast)
UBE2G2	ubiquitin-conjugating enzyme E2G 2 (UBC7 homolog, yeast)
UBE2L3	ubiquitin-conjugating enzyme E2L 3
UBE2L6	ubiquitin-conjugating enzyme E2L 6
UBE2T	ubiquitin-conjugating enzyme E2T (putative)
UCHL3	ubiquitin carboxyl-terminal esterase L3 (ubiquitin thiolesterase)
UGP2	UDP-glucose pyrophosphorylase 2
UNG	uracil-DNA glycosylase

USP48	ubiquitin specific peptidase 48
XAF1	XIAP associated factor 1
XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining)
YARS	tyrosyl-tRNA synthetase
YEATS2	YEATS domain containing 2
YIF1A	Yip1 interacting factor homolog A (S. cerevisiae)
ZCRB1	zinc finger CCHC-type and RNA binding motif 1
ZDHHC4	zinc finger, DHHC-type containing 4
ZDHHC5	zinc finger, DHHC-type containing 5

Table 2A. Down-regulated genes in demethylated CD4+ T cells were also decreased mRNA levels in the CD4+ T cells of SLE patient

Gene symbol	Full Name
AHNAK	AHNAK nucleoprotein
AHSA2	AHA1, activator of heat shock 90kDa protein ATPase homolog 2 (yeast)
ANKH	ankylosis, progressive homolog (mouse)
ANKRD12	ankyrin repeat domain 12
ARIH1	ariadne homolog, ubiquitin-conjugating enzyme E2 binding protein, 1 (Drosophila)
ARL4C	ADP-ribosylation factor-like 4C
ASF1A	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)
ATP2B1	ATPase, Ca++ transporting, plasma membrane 1
BCL11B	B-cell CLL/lymphoma 11B (zinc finger protein)
BRF1	BRF1 homolog, subunit of RNA polymerase III transcription initiation factor IIIB (S. cerevisiae)
BTG1	B-cell translocation gene 1, anti-proliferative
C11orf2	chromosome 11 open reading frame2
C12orf57	chromosome 12 open reading frame 57
C14orf28	chromosome 14 open reading frame 28
C5orf41	chromosome 5 open reading frame 41
CD44	CD44 molecule (Indian blood group)
CD48	CD48 molecule
CDKN1B	cyclin-dependent kinase inhibitor 1B (p27, Kip1)
CERK	ceramide kinase
CIRBP	cold inducible RNA binding protein
CITED2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2

CLK4	CDC-like kinase 4
COX4I1	cytochrome c oxidase subunit IV isoform 1
CRBN	cereblon
CREBZF	CREB/ATF bZIP transcription factor
CRTAP	cartilage associated protein
CYB561	cytochrome b-561
DOCK8	dedicator of cytokinesis 8
DPEP2	dipeptidase 2
DPH5	DPH5 homolog (S. cerevisiae)
EEF1B2	eukaryotic translation elongation factor 1 beta 2; eukaryotic translation elongation factor 1 beta 2-like
EEF1D	eukaryotic translation elongation factor 1 delta (guanine nucleotide exchange protein)
EEF1G	eukaryotic translation elongation factor 1 gamma
EEF2	eukaryotic translation elongation factor 2
EIF3F	eukaryotic translation initiation factor 3, subunit F; similar to hCG2040283
EIF3K	eukaryotic translation initiation factor 3, subunit K
EIF3L	eukaryotic translation initiation factor 3, subunit L
EIF4B	similar to eukaryotic translation initiation factor 4H; eukaryotic translation initiation factor 4B
EIF4EBP3	ankyrin repeat and KH domain containing 1; ANKHD1-EIF4EBP3 readthrough transcript; eukaryotic translation initiation factor 4E binding protein 3
ELK3	ELK3, ETS-domain protein (SRF accessory protein 2)
EPC1	enhancer of polycomb homolog 1 (Drosophila)
EPM2AIP1	EPM2A (laforin) interacting protein 1
EVI2B	ecotropic viral integration site 2B

FAM134A	family with sequence similarity 134, member A
FBXL3	F-box and leucine-rich repeat protein 3
FBXO11	F-box protein 11
FEM1B	fem-1 homolog b (C. elegans)
FYN	FYN oncogene related to SRC, FGR, YES
GAS5	growth arrest-specific 5 (non-protein coding)
GCNT4	glucosaminyl (N-acetyl) transferase 4, core 2 (beta-1,6-N-acetylglucosaminyltransferase)
GLIPR1	GLI pathogenesis-related 1
GLTSCR2	glioma tumor suppressor candidate region gene 2; glioma tumor suppressor candidate region gene 2 pseudogene
GNB2L1	guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1
HADHA	hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase (trifunctional protein), alpha subunit
HECA	headcase homolog (Drosophila)
ICAM3	intercellular adhesion molecule 3
IDS	iduronate 2-sulfatase
IGFBP3	insulin-like growth factor binding protein 3
INPP5D	inositol polyphosphate-5-phosphatase, 145kDa
INPP5F	inositol polyphosphate-5-phosphatase F
IRS2	insulin receptor substrate 2
ITM2B	integral membrane protein 2B
ITPKB	inositol 1,4,5-trisphosphate 3-kinase B
JMJD1C	jumonji domain containing 1C
KLHL28	kelch-like 28 (Drosophila)
KLRB1	killer cell lectin-like receptor subfamily B, member 1

LEMD3	LEM domain containing 3
LRIG1	leucine-rich repeats and immunoglobulin-like domains 1
MALT1	mucosa associated lymphoid tissue lymphoma translocation gene 1
MAN2A2	mannosidase, alpha, class 2A, member 2
METTL9	methyltransferase like 9
MPZL3	myelin protein zero-like 3
MYST1	MYST histone acetyltransferase 1
NAP1L2	nucleosome assembly protein 1-like 2
NKTR	natural killer-tumor recognition sequence
NR2C2	nuclear receptor subfamily 2, group C, member 2
P2RY10	purinergic receptor P2Y, G-protein coupled, 10
PCBP2	poly(rC) binding protein 2
PCNX	pecanex homolog (Drosophila)
PDIK1L	PDLIM1 interacting kinase 1 like
PFDN5	prefoldin subunit 5
PHF3	PHD finger protein 3
PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)
PRKCB	protein kinase C, beta
QARS	glutaminyl-tRNA synthetase
RAB4A	RAB4A, member RAS oncogene family
RANBP6	RAN binding protein 6
RBM16	RNA binding motif protein 16
RBM23	RNA binding motif protein 23
RC3H2	ring finger and CCCH-type zinc finger domains 2
RCAN3	RCAN family member 3
RNF103	vacuolar protein sorting 24 homolog (S. cerevisiae); ring finger protein 103
RNF38	ring finger protein 38

RORA	RAR-related orphan receptor A
RPL10	ribosomal protein L10; ribosomal protein L10 pseudogene 15; ribosomal protein L10 pseudogene 6; ribosomal protein L10 pseudogene 16; ribosomal protein L10 pseudogene 9
RPL10A	ribosomal protein L10a pseudogene 6; ribosomal protein L10a; ribosomal protein L10a pseudogene 9
RPL10L	ribosomal protein L10-like
RPL13	ribosomal protein L13 pseudogene 12; ribosomal protein L13
RPL13A	ribosomal protein L13a pseudogene 7; ribosomal protein L13a pseudogene 5; ribosomal protein L13a pseudogene 16; ribosomal protein L13a; ribosomal protein L13a pseudogene 18
RPL22	ribosomal protein L22 pseudogene 11; ribosomal protein L22
RPL24	ribosomal protein L24; ribosomal protein L24 pseudogene 6
RPL3	ribosomal protein L3; similar to 60S ribosomal protein L3 (L4)
RPL35A	ribosomal protein L35a
RPL37	ribosomal protein L37
RPL5	ribosomal protein L5 pseudogene 34; ribosomal protein L5 pseudogene 1; ribosomal protein L5
RPS16	ribosomal protein S16 pseudogene 1; ribosomal protein S16 pseudogene 10; ribosomal protein S16
RPS2	ribosomal protein S2 pseudogene 8; ribosomal protein S2 pseudogene 11; ribosomal protein S2 pseudogene 5; ribosomal protein S2 pseudogene 12; ribosomal protein S2 pseudogene 51; ribosomal protein S2 pseudogene 17; ribosomal protein S2 pseudogene 55; ribosomal protein S2 pseudogene 20; ribosomal protein S2
RPS6	ribosomal protein S6 pseudogene 25; ribosomal protein S6; ribosomal protein S6 pseudogene 1

RPS8	ribosomal protein S8; ribosomal protein S8 pseudogene 8; ribosomal protein S8 pseudogene 10
SECISBP2	SECIS binding protein 2
SF3B1	splicing factor 3b, subunit 1, 155kDa
SFT2D3	SFT2 domain containing 3
SIRT1	sirtuin (silent mating type information regulation 2 homolog) 1 (S. cerevisiae)
SLC30A1	solute carrier family 30 (zinc transporter), member 1
SNRPN	small nuclear ribonucleoprotein polypeptide N; SNRPN upstream reading frame
SORL1	sortilin-related receptor, L(DLR class) A repeats-containing
SSR2	signal sequence receptor, beta (translocon-associated protein beta)
SYPL1	synaptophysin-like 1
TGIF1	TGFB-induced factor homeobox 1
TMEM9	transmembrane protein 9
TOMM20	similar to translocase of outer mitochondrial membrane 20 homolog; similar to mitochondrial outer membrane protein 19; translocase of outer mitochondrial membrane 20 homolog (yeast)
TSC22D3	TSC22 domain family, member 3; GRAM domain containing 4
TTC17	tetratricopeptide repeat domain 17
UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1
UBR4	ubiquitin protein ligase E3 component n-recognin 4
UBXN11	UBX domain protein 11
VCPIP1	valosin containing protein (p97)/p47 complex interacting protein 1
VPS25	vacuolar protein sorting 25 homolog (S. cerevisiae)
ZBTB44	zinc finger and BTB domain containing 44
ZC3H6	zinc finger CCCH-type containing 6
ZFAND1	zinc finger, AN1-type domain 1

ZFP36L2	zinc finger protein 36, C3H type-like 2
ZMYM6	hypothetical LOC100130633; zinc finger, MYM-type 6
ZNF18	zinc finger protein 18
ZNF671	zinc finger protein 671
ZNF75A	zinc finger protein 75a

Table 3A. Up-regulation genes in demethylated CD4+ T cell and hypomethylation at promoter in CD4+ T cells of SLE

Gene symbol	Full Name
ACLY	ATP citrate lyase
AP1B1	adaptor-related protein complex 1, beta 1 subunit
AP3D1	adaptor-related protein complex 3, delta 1 subunit
APOBEC3B	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B
ARAF	v-raf murine sarcoma 3611 viral oncogene homolog
ARID5A	AT rich interactive domain 5A (MRF1-like)
ARMCX2	armadillo repeat containing, X-linked 2
ARRDC1	arrestin domain containing 1
ATP1B3	ATPase, Na+/K+ transporting, beta 3 polypeptide
ATP5J2	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit F2
BCL2	B-cell CLL/lymphoma 2
BCS1L	BCS1-like (yeast)
BICD2	bicaudal D homolog 2 (Drosophila)
BIVM	basic, immunoglobulin-like variable motif containing
BLMH	bleomycin hydrolase
BST2	NPC-A-7; bone marrow stromal cell antigen 2
C10orf84	chromosome 10 open reading frame 84
C10orf88	chromosome 10 open reading frame 88
C13orf23	chromosome 13 open reading frame 23
C6orf150	chromosome 6 open reading frame 150
CALR	calreticulin
CBS	cystathionine-beta-synthase
CD2AP	CD2-associated protein
CD9	CD9 molecule
CDCA7L	cell division cycle associated 7-like

COG1	component of oligomeric golgi complex 1
COPS7A	COP9 constitutive photomorphogenic homolog subunit 7A (Arabidopsis)
CRAT	carnitine acetyltransferase
CUL5	cullin 5
DAP	death-associated protein
DCTN3	dynactin 3 (p22)
EEF1E1	eukaryotic translation elongation factor 1 epsilon 1
ESRRA	estrogen-related receptor alpha
ETFA	electron-transfer-flavoprotein, alpha polypeptide
FADS1	fatty acid desaturase 1
FAM96A	family with sequence similarity 96, member A
FBXO22	FBXO22 opposite strand (non-protein coding); F-box protein 22
FEM1A	fem-1 homolog a (C. elegans); similar to fem-1 homolog a (C.elegans); similar to fem-1 homolog a
FEN1	flap structure-specific endonuclease 1
FSCN1	fascin homolog 1, actin-bundling protein (Strongylocentrotus purpuratus)
GABARAPL1	GABA(A) receptors associated protein like 3 (pseudogene); GABA(A) receptor-associated protein like 1
GCN1L1	GCN1 general control of amino-acid synthesis 1-like 1 (yeast)
GJB6	gap junction protein, beta 6, 30kDa
GNG5	guanine nucleotide binding protein (G protein), gamma 5
GPS1	G protein pathway suppressor 1
HEBP2	heme binding protein 2
IFI44L	interferon-induced protein 44-like
ILF3	interleukin enhancer binding factor 3, 90kDa
ING1	inhibitor of growth family, member 1

IPO13	importin 13
JAGN1	jagunal homolog 1 (Drosophila)
KIAA0196	KIAA0196
KLF10	Kruppel-like factor 10
LDHA	lactate dehydrogenase A
LZTFL1	leucine zipper transcription factor-like 1
MAFK	v-maf musculoaponeurotic fibrosarcoma oncogene homolog K (avian)
MARCKS	myristoylated alanine-rich protein kinase C substrate
MCCC2	methylcrotonoyl-Coenzyme A carboxylase 2 (beta)
MCM6	minichromosome maintenance complex component 6
MDM2	Mdm2 p53 binding protein homolog (mouse)
MLX	MAX-like protein X
MRPL15	mitochondrial ribosomal protein L15
MRPL20	similar to mitochondrial ribosomal protein L20; mitochondrial ribosomal protein L20
MRPL41	mitochondrial ribosomal protein L41
MRPL42	mitochondrial ribosomal protein L42
MRPS14	mitochondrial ribosomal protein S14
MTMR11	myotubularin related protein 11
NAP1L1	nucleosome assembly protein 1-like 1
NDUFA1	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1, 7.5kDa
NENF	neuron derived neurotrophic factor
NIP7	nuclear import 7 homolog (S. cerevisiae)
NUP43	nucleoporin 43kDa
OSM	oncostatin M
PAPOLA	poly(A) polymerase alpha
PCBP4	poly(rC) binding protein 4
PEA15	phosphoprotein enriched in astrocytes 15

PIP5K1A	phosphatidylinositol-4-phosphate 5-kinase, type I, alpha
PKNOX1	PBX/knotted 1 homeobox 1
POLR3D	polymerase (RNA) III (DNA directed) polypeptide D, 44kDa
PPAN	peter pan homolog (Drosophila)
PPOX	protoporphyrinogen oxidase
PPP2R5B	protein phosphatase 2, regulatory subunit B', beta isoform
PPP3CB	protein phosphatase 3 (formerly 2B), catalytic subunit, beta isoform
PRKACA	protein kinase, cAMP-dependent, catalytic, alpha
PSENEN	presenilin enhancer 2 homolog (C. elegans)
PSME2	proteasome (prosome, macropain) activator subunit 2 (PA28 beta)
PTPN1	protein tyrosine phosphatase, non-receptor type 1
PYCARD	PYD and CARD domain containing
RAB21	RAB21, member RAS oncogene family
RANBP2	RAN binding protein 2
RARG	retinoic acid receptor, gamma
RECQL	RecQ protein-like (DNA helicase Q1-like)
RPL28	ribosomal protein L28
SACM1L	SAC1 suppressor of actin mutations 1-like (yeast)
SCFD1	sec1 family domain containing 1
SLC30A5	solute carrier family 30 (zinc transporter), member 5
SLC37A3	solute carrier family 37 (glycerol-3-phosphate transporter), member 3
SMARCC1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1
SMU1	smu-1 suppressor of mec-8 and unc-52 homolog (C. elegans)
SNAPC3	small nuclear RNA activating complex, polypeptide 3, 50kDa
SUPT16H	suppressor of Ty 16 homolog (S. cerevisiae); suppressor of Ty 16 homolog (S. cerevisiae) pseudogene
TFAM	transcription factor A, mitochondrial

TM9SF1	transmembrane 9 superfamily member 1
TPM2	tropomyosin 2 (beta)
TRIM22	tripartite motif-containing 22
TRIM35	tripartite motif-containing 35
TRIM59	tripartite motif-containing 59
TRIP11	hypothetical LOC341378; thyroid hormone receptor interactor 11
TSR1	TSR1, 20S rRNA accumulation, homolog (S. cerevisiae)
TTC1	tetratricopeptide repeat domain 1
TTLL4	tubulin tyrosine ligase-like family, member 4
TUBB2A	tubulin, beta 2A
UHRF1	ubiquitin-like with PHD and ring finger domains 1
VAPA	VAMP (vesicle-associated membrane protein)-associated protein A, 33kDa
WDR37	WD repeat domain 37
YWHAG	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide
ZBED4	zinc finger, BED-type containing 4
ZDHHC5	zinc finger, DHHC-type containing 5
ZNF282	zinc finger protein 282
ZNF410	zinc finger protein 410
ZNF622	zinc finger protein 622
ZW10	ZW10, kinetochore associated, homolog (Drosophila)

Chromosome 1: 113213155-113213281

GTTCTGTTAGCTGATGTATGTAGCCCCTAGTCACGTTCCTCACACTTACTTGATCTATTACG
ACCTTTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTTG
GGG

Chromosome 2: 57103032-57103158

GTTCTGTTAGGTGATGCATGTAGCCCCCAGTCACGTTCCTCATGCTTGCTTGATCTATTATG
ACCCTTTCACGTGGATCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTC
GGAG

Chromosome 2: 111153611-111153741 (BUB1)

GTTCTGTTAGCCGATGCATGTAGCCCCCAGTCACGTTCCCCACGCTTGCTCCATCTATCA
CAACCGTTTCACATGGACCCCCCCCTTCAAGTTGTAAGCCCTTAAAAGGGCTAGGAATTT
CTTTTTTGGGG

Chromosome 2: 128157730-128157856

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTTCTCACGCTTACTTGATCTATTATG
ACTTTTTCACGTAGACCCCTTAGAGTTATAAGCCCTTAAAAGGGCTAGGAATTTCTTTCTCG
GGG

Chromosome 2: 149600363-149600489

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTTACGTTCCTCACGCTTACTTGATCTATTATG
ACTTTTTCATGTAGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCCAGGAATTTCTTTTTCG
GGG

Chromosome 4: 37503850-37503980 (phosphoglucomutase 2)

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATCATG
ACCCTTTCACGTAGACAGACCCCTTAGCGTTGTAAGCCCTTAAAATGGCTAGGAATTTCTT
TTTCAGGG

Chromosome 7: 57440239-57440365

GTTCTGTTAGCTGATGTATGTAGCCCCCGGTCACGTTCCTCACGCTTACTTGATCTATCAT
GACCCTTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAAGGCTAGGAATTTCTTTTT
CAAGG

Chromosome 7: 63692884-63693007

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTGCTCACCTTACTTGATCTATTATGA
CTTTCACGTAGACCCCTTAGAGTTGTAACCCCTTAAAAGGGCTAGGAATCTCTTTTTCGGG
G

Chromosome 8: 8075280-8075406 (similar to liver-related low express protein 1)

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCATGTTCCTCAAGCTTCCTTGATCTATTATG
ACTTTTTCACGTAGAACCCTTAGAGTTGTCAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCA
GGG

Chromosome 8: 144701449-144701575 (zinc finger CCCH-type containing 3 and gasdermin domain containing 1)

GTTCTGTGAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGAGCTATCAT
GACCCTTTCACGGGGACCCCTTAGAGTTGTAAGCTCTTAAAAGGGCTAGGAATTTCTTTTT
CAGGG

Chromosome 11: 65524959-65525085 (barrier to autointegration factor 1 and eukaryotic translation initiation factor 1A domain)

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTTCTCACGCTTACTTGATCTATTATG
ACTTTTTCACGTAGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCG
GGG

Chromosome 14: 101290266-101290392

GTTCTGTTAGCTGATGCATGTAGCCCCCAGTCACATTGCCCACGCTTGCTTGACCTATCAC
GACCCTTTCACATGTACCCCTTAGAGTTGTAAGCCCTTAAAAAGGCCAGGAATTTCTTTTTC
CAGG

Chromosome 14: 104576094-104576220 (G protein-coupled receptor 132)

GTTCTGTGAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTACCAT
GACCTTTTCACGTGGACCTCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTC
CGGG

Chromosome 18: 831763-831889

GTTCTGTTAGCTGATGCATGTAACCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTATG
ACCCTTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTC
GGGG

Chromosome 19: 40565455-40565581 (G protein-coupled receptor 42 and free fatty acid receptor 3)

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCATGCTTACTTGATCTATTATG
ACCCTTTCACATGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTTG
GGG

Chromosome 19: 58152036-58152162 (zinc finger protein 816A)

GTTCTGTTAGCTGATGCACGTAGTGTCAAGTCACATTCCCCACACTTGCTTGATCTATCAT
GACCCTTTCACGTGGATGCCTTAGAGCTGTAAGCCCTTAAAAGGGCCAGGAATTTCTTTTT
CAGGG

Chromosome 20: 24848122-24848248

GTTCTGTTAGCTGATGTATGTAGCCCCTAGTCACGTTCCTCACACTTACTTGATCTATCATG
ACCCTTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTC
GGGG

Chromosome 21: 45152839-45152965 (integrin, beta 2)

GTTCTGTCAGCTGATGCATGTAACCCCCAGTCAAGTTCCCCACACTTGCTCTATCTATCAC
AACCCTTTCATGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTC
GGGG

Chromosome 22: 29079586-29079711 (splicing factor 3a, subunit 1, 120kDa)

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTATG
ACTCTTTCACATAGACCTTTAGAGTTGTAAGCCCTTTAAAGGGCTAGGAATTTCTTTTTCGG
GG

Chromosome X: 14042-14168

GTTCTCTTAGCTGATGTATGTAGCCCCCACTCACGTTCCTCACACTTACTTGATCTATTATG
ACTCTTTCACATAGACCCCTTAGAGTTCTTAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCG
GGG

Chromosome X: 92777-92903

GTTCTCTTAGCTGATGTATGTAGCCCCCACTCACGTTCCTCACACTTACTTGATCTATTATG
ACTCTTTCACATAGACCCCTTAGAGTTCTTAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCG
GGG

Chromosome X: 44106884-44107011

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCATGTTCCTCACGCTTACTCGATCTATTATG
ACTCTTTCACGTAGACCCCTTAAGAGTTGTAAGCCCTTAAACGGGCTAGGAATTTCTTTTTC
AGGG

Chromosome X: 92571583-92571709

GTTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCATGCTTGCTTGATCCATTATG
ACTCTTTCACGTAGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCCAGGAATTTCTTTTTTG
GGG

Chromosome X: 92579910-92580036

ATTCTGTTAGCTAATGTATGTAGCCCCCAGTCATGTTCCTCACGCTTACTTAATCTATTATAA
CTCTTTCACGTAAACCCCTTAGAGTTATAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCGG
GG

Chromosome X: 118910249-118910375 (NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1 and A kinase (PRKA) anchor protein 14 isoform b)

GTTCTGTTACCTGATGTATGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTATG
ACTCTTTCACGTAGACTCCTTAGAGTTGTAAGGCCTTAAAAGGGCTAGGAATTTCTTTTTCG
GGG

Chromosome Y: 8791038-8791164

GTTCTGTTAGCGGATGTATGTAGCCCCCAGTCACATTCCTCGCACTTACTTGATCTTTTATG
AACCTTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAATGGCTAGGAATTTCTCTTTTG
GGG

Chromosome 3: 68563687-68563813 (family with sequence similarity 19 (chemokine (C-C motif)-like), member A1)

TGTTAGCTGATGTATGTAGCCCCCAGTCATGTTCCTCACGCTTACTTGATCTATTATGACCC
TTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCGGGG
AGCT

Chromosome 3: 131342551-131342677

TGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTATGACTT
TTTCACGTAGACCGCTTAGAGTTGTCAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCCGGG
AGCT

Chromosome 7: 70820728-70820854 (UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase)

TTCTGTTAGCTGATGTATGTAGCCCCCAGGCACGTTCCTCACGCTTACTTGATCTATTATGA
CCCTTTCATGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCA
GGGA

Chromosome 8: 11806653-11806779

TTCTGTTAGCTGATGTATGTAGCCCCCAGTCATGTTCCTCATACTTACTTGATCTATTATGAC
GTTTTCACATAGAACCCTGAGAGTGGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCAG
GGA

Chromosome 8: 12387068-12387194

TTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCAAGCTTCCTTGATCTATTATGA
CTTTTTCACGTAGAACCCTTAGAGTTGTCAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCAG
GGA

Chromosome 8: 146265664-146265790

CTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTATGA
CCCTTTCACGTGGACCCCTTAGAGTTGTAAACCCTTACAAGGGCTAGGAATTTATTTTTTG
GAGA

Chromosome 9: 25359744-25359870

TTCCGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTATGA
CTCTTTCACGCAGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCG
GGGA

Chromosome 11: 106630973-106631099

TTCTGTTAGCTGATGCACGTAGCCCCCAGTCACGTTCCTCACACTTACTTGATTTATCACG
ACCCTTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTGTTTTTC
AGGGA

Chromosome 11: 116934636-116934762 (Down syndrome cell adhesion molecule like 1)

TTCTGTTAGCTGATGTATGTAGCCCCCAGTCACCTTCCTCACGCTTACTTGATCTATTATGA
CCCTTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCG
GGGA

Chromosome 12: 7747609-7747735 (growth differentiation factor 3 precursor and stella)

TTCTGTAAGCTGATGTATGTAGCCCCCAGTCACGTTTCTCATGCTTACTTGATCTATTATAA
CTTTTTCACGTAGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCGG
GGA

Chromosome 12: 8326271-8326397

TTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCATGCTTACTTGATCTATTGTGA
CTTTTTCACGTAGAACCCTTAGAGTTGTCAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCAG
GGA

Chromosome 12: 123877710-123877836 (scavenger receptor class B, member 1)

TTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTCCTCACACTTACTCGATCTATTATGA
CCCTTTCACATGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCG
GGGA

Chromosome 14: 81176905-81177028

TTCTGTTAGCTGATGTATGTAGCCCCCAGTCACATTCCTCACGCTTTGATCTATTATGACCC
TTTCACATGGACCCCTTAGAGTTGCAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCGGAG
A

Chromosome 17: 25591583-25591709 (bleomycin hydrolase)

TTCTGTTAGCTGATGCATGTAGCCCCCAGTCATGTTCCCCACGCTTGCTCGATCTATCACA
ACCCTTTCACGTGTACCCCTTAGAGTGGTAAGCCCTTAAAAGGGCCAGGAATTTCTTTTTC
AGGGA

Chromosome 18: 24533102-24533228

TTCTGTTAGCTGATGTATGTAGCCCCCAGTCACGTTTCTCATGCTTACTTGATCTATTATGA
CTTTTTCAGGTAGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTCAG
GGA

Chromosome 19: 58583489-58583611 (zinc finger protein 765)

TTCTGTTAGCTGATGTATGTAGCCTCCAGTCACATTCCTCACACTTACTTGATTATGACCTTT
TCACGTAGACCCCTTAGAGCTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTTGGGGA

Chromosome 7: 67456983-67457109

GTTCTGTTAGCTGATGCATGTAGCCCCCAGTTGTATTTCTCATGCTTGCTTGATGTATCACG
ACCCTTTCACGTGGACCCCTTAAAGTTGAAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTTG
GAG

Chromosome 7: 63887260-63887386

GTTCTGTTAGCTGATGCATGTAGCCCCCAGTCACGTTTCCCATGCTTGCTTGATGTATCAC
GACCCTTTCACGTGGACCCCTTAAAGTTGCAAGCCTTTAAAAAGGCCAAGAATTTCTCTTT
CGGGG

Chromosome 9: 135608873-135609003 (vav 2 guanine nucleotide exchange factor)

GTTCTGTTCGCTGATGCATGCAGGCCCCAGTCACGTTTCCCATGCTTGTTCGATTTATCAC
GACCTTTTCACGTGGACCCCTTAGAGCTAGCTATAAGCCTTTAAAAAGGCCAAGAATTTCT
TTTTCAGGG

Chromosome 16: 28289370-28289496 (eukaryotic translation initiation factor 3, subunit C-like and Uncharacterized protein KIAA0220)

GTTCTGTTAGCTGATGCATGTAGCCCCCACTCACCTTCCCCATGCTTGCTCGATTTATCAC
GACACTTTCACGTGGACCCTTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTC
GGGG

Chromosome 17: 2635608:2635734 (GTPase activating Rap/RanGAP domain-like 4 isoform 1)

GTTCTGTTAGCTGATGCATGTAGCTCCCAGTCACATTTCCCACGCTTGCTTGATGTATCAC
GACCTTTTCACGTGGACCCCTTTAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTC
GGGA

Chromosome 19: 1393674-1393800 (ribosomal protein S15 and adenomatosis polyposis coli 2)

GTTGTGTTAGCTGATGCATGCAGCCCCCACTCACGTTTTCCACACTTGCTTGATGTATCAC
GACCCTTTCACGTGGACCCCTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATCTCTTTTT
CGGGG

Chromosome 20: 23751754-23751879 (cystatin SA precursor)

GTTCTGTTAGCTGATGCATGTAGCCCCAGTCAAGTTTCCCACGCTTGCTTGATGTATCATG
ACCCTTTCATGTGGACCCCTTAAAGCTGTAAGCCTATAAAAAGGCCAAGAATTTCTTTTTCC
GGG

Chromosome 22: 48831332-48831458 (tubulin tyrosine ligase-like family, member 8)

GTTCTGTTAGCTGATGCATGCAGCCCCCAGTCACGTTTCCCATGCTTACTTGATGTATCAC
GACCCTTTCACGTGGACCCCTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCCTTTT
CTGGG

Chromosome 22: 40789535-40789661 (alpha-N-acetylgalactosaminidase)

GTTCTGTTAGCTGATGCATGTAGCCTCCAGTCACGTTTCCCACGCTTGCGTGATGTATCAC
GACCCTTTCACATGGACCCCTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTC
AGGG

Chromosome X: 53067905-53068031

GTTCTGTTAGCTGATGCATGTAGCCCCCAGTCACGTTTCCCAAGCTTGCTTGATGTATCAC
GACCCTTTCACGTGGACCCCTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTC
GGGG

Chromosome X: 134799809-134799935 (cancer/testis antigen CT45-6 and sarcoma antigen 1)

GTTCTGTTAGCTGCTGCATGTAGCCCCCAGTCACATTTCCCACACTTGCTCAATTTATCAC
GACTCTTTCACGTGGACCCCTTAAAGTTGTAAGCCTCTAAAAAGGCCAAGAATTTCTTTTTC
GGGG

Chromosome 1: 159640697-159640823

TTCTGTTAGCTGATGCATGTAGCCCCCAGTCACGTTTCCCACATTTGCTCGATTTATCACG
ACCCTTTCATGTGGACCCCTTAAAGTTTTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTCCG
GGGA

Chromosome 2: 122160712-122160838 (hCG1812832)

TTCTGTTAGCTGATGCATGTAGCCTCCAGTCACATTTCCCACGCTTGCTCGATTTATCACG
ACCCTTTCACATGGACCCCTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTCG
GAGA

Chromosome 2: 241806827-241806952 (transmembrane protein 16G isoform NGEP long (anoctamin 7) and high density lipoprotein binding protein)

TACCGTTAGCTGATGTATGTAGCCCACAGTCACGTTCCTCACGCTTACTTGATCTATTATGA
CTCTTTCACGTAGACCCTTAGAGTTGTAAGCCCTTCAAAGGGCTAAGAATTTCTTTTTCGGC
GA

Chromosome 11: 25445928-25446054

TTCTGTTAGCTGATGCATGTAGCCCCCAGTCATGTTCCCTGCGCTGACTCGATCTATCACG
ACCCTTTCACGTGGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCCAAGAATTTCTTTTTC
AGGGA

Chromosome 11: 43762224:43762351 (hydroxysteroid (17-beta) dehydrogenase 12)

TTCTGTTAGCTTATGCATGTAGCCCCCAGTAACGTTTCCCATGCTTACTCGATCTATCATGA
CTCTTTCATGTGGACCCCTTAGGAGTTGTAAGCCCTTAAAAGGGCCAAGAATTTCTTTTTGA
GGGA

Chromosome 11: 67533546-67533672 (unc-93 homolog B1 and aldehyde dehydrogenase 3B1 isoform a)

TTCTGTTAGCTGACGCATGCAGCCCCCGGTCACGTTCCCCGAGCTTGCTCGATTTATCAC
GACCCTTTCACGTGGACCCCTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATGTCTTTTT
CAGGGA

Chromosome 13: 108589970-108590095 (myosin heavy chain Myr 8)

TTCTGTTAGCTGATGCATGTAACCCCAGTCACGTTCCCCACCCTTGCTCTATCTATTACGA
CCCTTTCACGTAGACCCCTTAAACTTGTAAGCCCTTAAAAGGGCCAAGAATTTCTTTTTCG
GAGA

Chromosome 16: 28663433-28663559

TTCTGTTAGCTGATGCATGTAGCCCGCACTCACCTTCCCCATGCTTGCTCGATTTATCACG
ACACTTTCACGTGGACCCTTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTCG
GGGA

Chromosome 19: 20721574-20721700 (zinc finger protein 66)

TTCTGTTAGCTGATGTATGTAGCCCCCAGTCATGTTTCTCACGCTTACTTGATCTATTATGA
CTTTTTCATGTAGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAAGAATTTCTTTTTCGG
GGA

Chromosome 22: 24548694-24548821 (myosin XVIIIB)

TTCTGTTAGCTGATGCATGTAGTCCCCCCAGTCACGTTTCCCAGCTTGCTCGATTTATCAC
GACCTCTTCATGCGGACCCCTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTC
AGGGA

Chromosome X: 135840772-135840900

TTCTGTTAGCTGATGCATGTAGCCCCCAGTCACGTTCCCCATGCTTGCCCAATTTATCATG
ACCTTTTCACGTGGACCCCTTAGAGTTGTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTT
CAGGGA

Chromosome X: 135846099-135846226

TTCTGTTAGCTGATGCATGTATCCCCCAGTCACGTTCCCCACACTTGCTCAATTTATCATGA
CCTTTTCATGTGGACCTCTTAAGAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTCA
GGGA

Chromosome 17: 23581780-23581906

TAGTGTTAGCTGATGTAGGTAGCCCACAGTCACATTCCTCACACTTACTTGATCTATTATGA
CTCTTTCACGTAGACCCCTTAGAGTTGTAAGCCCTTCAAAGGGCTAGGAATTTCTTTTTCG
GGGA

Chromosome 17: 23590105-23590231

TACTGTTAGCTGATGTAGGTAGCCCCCAGTCACGTTCCTCACACTTACTTGATCTGTGATG
ACTCTTTCACGTAGACCCCTTAGAGTTGTAAGCCCTTCAAAGGGCTAGGAATTTCTTTTTCG
GGGA

Chromosome 19: 20729892-20730018 (zinc finger protein 66)

TTCTGTTAGCTGATGTAGGTAGCCCCCAGTCATGTTTCTCACGCTTACTTGACCTATTATGA
CTTTTTCATGTAGACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTTGG
GGA

Chromosome 2: 231973927-231974053 (UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7)

GTCCTATTAGCTGATGTGTGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTATG
ACCCTTTCACGTGGACCCCTTAGAGTTGTAAGCTCTTAAAAGGGCTAGGAATTTCTTTTTC
GGGG

Chromosome 2: 231981564-231981690 (UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7)

GTCCTGTTAGCTGATGTGTGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTAT
GACCCTTTCACGTGGACCCCTTAGAGTTGTAAGCTCTTAAAAGGGCTAGGAGTTTCTTTCT
CGGGG

Chromosome 5: 180227496-180227622

GTTCTGTTAGCTGATGTGTGTAGCCCCCAGTCACGTTCCTCACGCTTACTTGATCTATTATG
ACCCTTTCACGTGAACCCCTTAGAGTTGTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTC
GGGG

Chromosome X: 2545:2670

GTTCTGTTATCTGATGCGTGTAGCCCCCAGTCACGTTCCGATGCTTGCTCGATCTATCACG
ACCCTTTCAAGTGAACCCCTTAGAGTCGTAAACCCTTAAAAGGGCCAGGAATTTCGTTTTC
GGGG

Chromosome X: 5935-6061

GTTCTGTTATCTGATGCGTGTAGCCCCCAGTCACGTTCCCCATGCTTCCTCGATCTATCAC
AACCCTTTCACGTAAACCCCTTAGAGTTGTAAACCCTTAAAAGGGCCAGGAATTTCGTTTT
CGGGG

Chromosome X: 85314-85440

GTTCTGTTATCTGATGCGTGTAGCCCCCAGTCACGTTCCCCACGCTTCCTCGATCTATCAC
AACCCTTTCACGTAAACCCCTTAGAGTTGTAAACCCTTAAAAGGGCTAGGAATTTCGTTTTC
GGGG

Chromosome 12: 58033524-58033651

TTCTGTTAGCTGATGTGTGTAGCCCCCAGTCACGTTCCTCACACTTACTTGATCTATTATGA
CCCTTTCACGTGGACCCCTTAGAGTTGTTAAGCCCTTAAAAGGGCTAGGAATTTCTTTTTC
GGGGA

Chromosome 16: 29617333-29617457 (quinolinate phosphoribosyltransferase)

GCTCTGTTAGCTGATGCGTGTAGCCCCCAGTCACGTTCCCCGCTTGCTTGAGATATCATG
ACCCTTTCACGTGGACCCCTTAGAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTCA
GGG

Chromosome 9: 137712101-137712227 (lipocalin 9 and spermatogenesis and oogenesis specific basic helix-loop)

TTCTGTTAGCGGATGCGTGTAGCCCCCGGTCACGTTTCCCACGCTTGCTTGATGTATCAC
GACCTTTTCACGTGGACCCCTTAAAGTTGTAAGCCTTTAAAAAGGCCAAGAATTTCTTTTTC
GGGGA

BIOGRAPHY

Mr. Jeerawat Nakkuntod was born on May 25th, 1979 in Nakhonratchasima, Thailand. He graduated with the Bachelor degree of Science in Microbiology from Ubonratchatani University in 2000 and Master degree of Science in Medical Microbiology from Chulalongkorn University in 2004. He got a Royal Golden Jubilee (RGJ) Ph.D. Scholarship from the Thailand Research Fund (TRF) and participated in Medical Microbiology program, Graduate School, Chulalongkorn University for philosophy degree in 2006.