

## Epigenome-wide differential DNA methylation between HIV-infected and uninfected individuals

Xinyu Zhang<sup>a,b</sup>, Amy C. Justice<sup>c</sup>, Ying Hu<sup>d</sup>, Zuoheng Wang<sup>e</sup>, Hongyu Zhao<sup>f</sup>, Guilin Wang<sup>g</sup>, Eric O. Johnson<sup>h</sup>, Brinda Emu<sup>e</sup>, Richard E. Sutton<sup>e</sup>, John H. Krystal<sup>a,b</sup>, and Ke Xu<sup>a,b</sup>

<sup>a</sup>Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA; <sup>b</sup>Connecticut Veteran Health System, West Haven, CT, USA; <sup>c</sup>Yale University School of Medicine, New Haven Veterans Affairs Connecticut Healthcare System, West Haven, CT, USA; <sup>d</sup>Center for Biomedical Informatics & Information Technology, National Cancer Institute, Bethesda, MD, USA; <sup>e</sup>Department of Internal Medicine, Division of Infectious Disease, Yale University School of Medicine, New Haven, CT, USA; <sup>f</sup>Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA; <sup>g</sup>Yale Center of Genomic Analysis, West Campus, Orange, CT, USA; <sup>h</sup>Fellow Program and Behavioral Health and Criminal Justice Division, RTI International, Research Triangle Park, NC, USA

### ABSTRACT

Epigenetic control of human immunodeficiency virus-1 (HIV-1) genes is critical for viral integration and latency. However, epigenetic changes in the HIV-1-infected host genome have not been well characterized. Here, we report the first large-scale epigenome-wide association study of DNA methylation for HIV-1 infection. We recruited HIV-infected ( $n = 261$ ) and uninfected ( $n = 117$ ) patients from the Veteran Aging Cohort Study (VACS) and all samples were profiled for 485,521 CpG sites in DNA extracted from the blood. After adjusting for cell type and clinical confounders, we identified 20 epigenome-wide significant CpGs for HIV-1 infection. Importantly, 2 CpGs in the promoter of the *NLR family, CARD domain containing gene 5* (*NLRCS*), a key regulator of major histocompatibility complex class I gene expression, showed significantly lower methylation in HIV-infected subjects than in uninfected subjects (cg07839457:  $t = -6.03$ ,  $P_{\text{nominal}} = 4.96 \times 10^{-9}$ ; cg16411857:  $t = -7.63$ ,  $P_{\text{nominal}} = 3.07 \times 10^{-13}$ ). Hypomethylation of these 2 CpGs was replicated in an independent sample (GSE67705: cg07839457:  $t = -4.44$ ,  $P_{\text{nominal}} = 1.61 \times 10^{-5}$ ; cg16411857:  $t = -5.90$ ;  $P = 1.99 \times 10^{-8}$ ). Methylation of these 2 CpGs in *NLRCS* was negatively correlated with viral load in the 2 HIV-infected samples (cg07839457:  $P = 1.8 \times 10^{-4}$ ; cg16411857:  $P = 0.03$  in the VACS; and cg07839457:  $P = 0.04$ ; cg16411857:  $P = 0.01$  in GSE53840). Our findings demonstrate that differential DNA methylation is associated with HIV infection and suggest the involvement of a novel host gene, *NLRCS*, in HIV pathogenesis.

### ARTICLE HISTORY

Received 22 April 2016  
Revised 30 June 2016  
Accepted 2 August 2016

### KEYWORDS

DNA methylation;  
Epigenome-wide association;  
HIV-1 infection; *NLRCS*; viral load

## Introduction

Epigenetic mechanisms are involved in human immunodeficiency virus-1 (HIV-1) integration into the host genome<sup>1–3</sup> and viral latency.<sup>4–6</sup> Epigenetic modifications regulate gene transcription and expression without changing the DNA sequence. Histone deacetylation has been shown to be important for the quiescence of HIV gene expression in HIV-1-infected T cells.<sup>7,8</sup> In contrast, DNA hypermethylation is associated with viral latency, likely by inhibiting transcriptional initiation of the integrated provirus and inhibiting viral post-integration reactivation from latency.<sup>7,9</sup> Therefore, epigenetic reprogramming regulates the virus life cycle at multiple stages and modulates the interaction between the viral and host genomes. However, epigenetic characterization of the host genome in HIV-infected vs. uninfected individuals has not been thoroughly conducted.

Epigenetic control specifically targets genes involved in cell cycle progression, senescence, survival, inflammation, and immunity. The hallmark of HIV infection is the depletion of peripheral CD4<sup>+</sup> T cells, which compromises host defense. HIV-1 infection

may initiate or influence epigenetic processes in host cells and manipulate host genome epigenetic reprogramming.<sup>10</sup> For example, HIV-1 increases DNA methyltransferase expression in HIV-1-infected CD4<sup>+</sup> T cells, which is responsible for *de novo* methylation.<sup>11,12</sup> DNA methylation of host genes in response to HIV-1 potentially impacts HIV-1 transcription and replication. Therefore, characterizing DNA methylation in the host genome may provide insight into the pathogenesis of HIV infection.

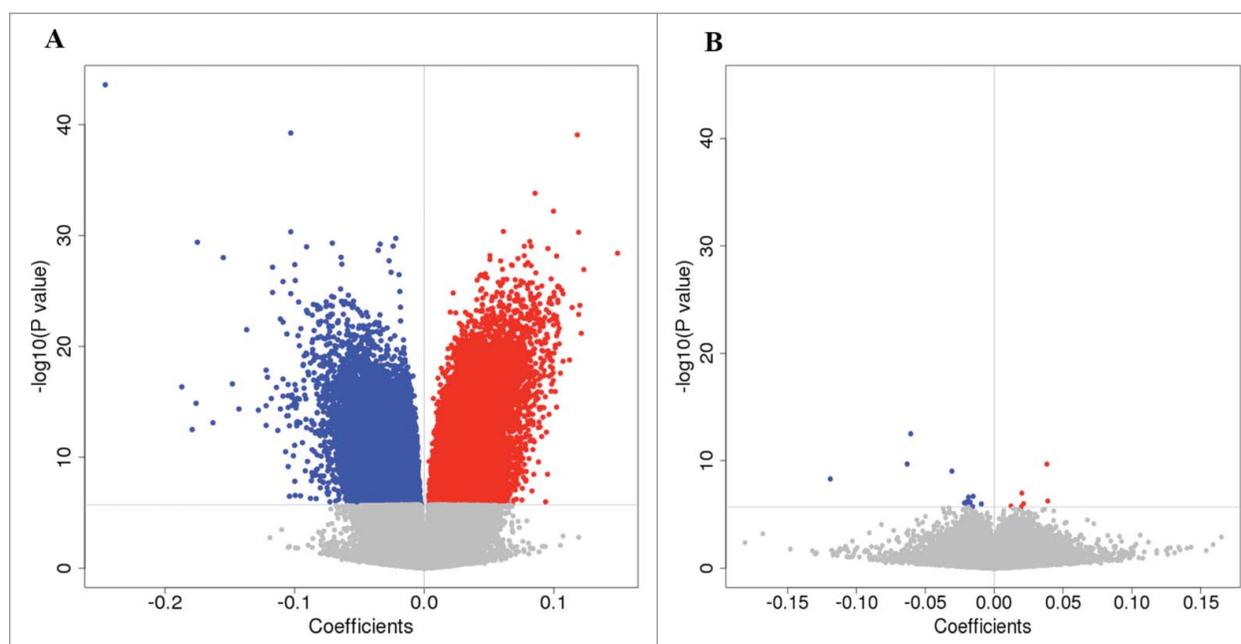
Candidate gene/region studies of DNA methylation associated with HIV infection have shown that CpG sites in the promoter of *FOXP3* were significantly less methylated in HIV-infected patients than in uninfected patients in colonic mucosa and blood samples.<sup>13</sup> A previous study reported that CpGs in the *IL2* promoter were completely methylated in naïve CD4<sup>+</sup> T cells and significantly demethylated in memory CD4<sup>+</sup> T cells during chronic HIV infection,<sup>14</sup> indicating that alterations in DNA methylation occur in different subpopulations of CD4<sup>+</sup> T cells during HIV infection. Interestingly, a recent study revealed hypomethylation in the region encoding the human leukocyte antigen locus in

**CONTACT** Ke Xu, MD, PhD,  [ke.xu@yale.edu](mailto:ke.xu@yale.edu)  Department of Psychiatry, Yale School of Medicine, 300 George St., Suite 901, New Haven, CT 06511, USA.

 Supplemental data for this article can be accessed on the [publisher's website](#).

Published with license by Taylor & Francis Group, LLC © Xinyu Zhang, Amy C. Justice, Ying Hu, Zuoheng Wang, Hongyu Zhao, Guilin Wang, Eric O. Johnson, Brinda Emu, Richard E. Sutton, John H. Krystal, and Ke Xu.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.



**Figure 1.** Effect of cell composition on DNA methylation. Volcano plots show the coefficient at a CpG site vs.  $-\log_{10}(P\text{-value})$ . Blue circles indicate hypomethylated CpG sites, while red circles indicate hypermethylated CpG sites. (A) Volcano plot showing association results without correction for cell types. (B) Volcano plot with correction for 6 cell types in the blood.

chronic HIV-infected patients.<sup>15</sup> These findings indicate that methylation should be examined at the epigenome level.

Epigenome-wide analysis of DNA methylation is a robust and unbiased approach for investigating the association of DNA methylation with complex disease phenotypes.<sup>16–18</sup> However, the epigenome-wide DNA methylation patterns in HIV-1 infection are largely unknown. Recently, a study reported 4,679 differentially methylated regions in the genomes of a single monozygotic twin pair in which one sibling was HIV-infected, while the other was HIV-uninfected.<sup>19</sup> This study found that 2 genes, *IGFBP6* and *SATB2*, were hypermethylated in the HIV-infected individual compared to the uninfected individual in 8 patients. This study used a small sample size, and thus studies in larger samples are necessary.

The goal of this study was to identify differentially methylated CpGs in the host genome between HIV-infected and uninfected subjects and to determine whether HIV-associated methylation sites are correlated with viral load (VL). Here, we report an epigenome-wide association study (EWAS) of HIV-infected and uninfected individuals selected from a well-established HIV cohort, the Veterans Aging Cohort Study (VACS). We replicated the biologically meaningful findings in an independent publicly available sample (GSE67705, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67705>). We also examined the relationship between HIV-associated CpG methylation and VL in HIV-infected patients from the VACS sample and an independent publicly available sample (GSE67751, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67751>).

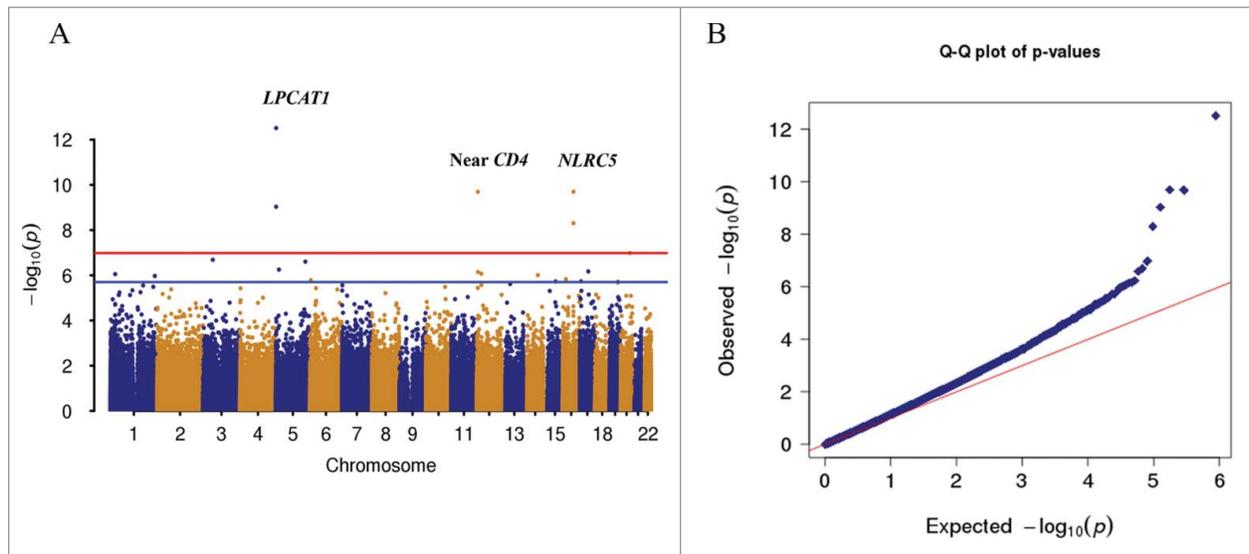
## Results

### ***EWAS revealed significant differential methylation CpG sites between HIV-1-infected and uninfected patients***

To correct for technical bias and biological confounders, we adapted a recently developed comprehensive pipeline for

EWAS to detect CpG sites related to HIV infection.<sup>20</sup> We first examined whether cell type composition was associated with methylation effects in HIV infection. Using a well-established method,<sup>21,22</sup> we estimated the cell type proportion in each sample. We compared the EWAS results from models with and without cell type correction. Fig. 1 shows the volcano plots with the  $P$ -value and coefficient at a CpG site between HIV-infected and uninfected patients without and with correction for cell type composition. A model without cell type correction revealed 405 differential CpG sites associated with HIV infection (Fig. 1A). The quantile–quantile plot showed significant inflation (Supplementary Fig. S1). An EWAS corrected for cell type proportion identified 20 significant CpG sites associated with HIV infection (Fig. 1B). After correcting for cell type composition, we found that none of the 20 HIV-associated CpG sites were correlated with any of the 6 cell types ( $P > 0.05$ ) (Supplementary Fig. S2). These results suggest that cell type strongly affects EWAS results for HIV infection and that it is essential to correct for cell types in EWAS analysis. We then examined whether demographic, clinical, and batch variables confounded the EWAS analysis. Using control probe intensities designed for the Illumina Infinium HumanMethylation450 Beadchip (450K array), we detected multiple confounders in addition to cell types, including batch, age, sex, smoking, and cell type (Supplementary Fig. S3A). After adjusting for these confounding factors, no correlation between residual methylation and confounders, except smoking status, was observed (Supplementary Fig. S3B), further suggesting that our analytic strategy sufficiently corrected for methylation signals from different cell types and other confounders. Batch showed no confounding effects in our analytic pipeline. Smoking status was adjusted in the model.

As shown in Fig. 2, 20 CpG sites reached epigenome-wide significance for HIV infection (false discovery rate, FDR  $q < 0.05$ ).



**Figure 2.** Epigenome-wide association analysis of 437,722 CpG sites in 378 samples. A. Manhattan plot showing chromosome location of 22 autosomes (X-axis) and  $-\log_{10}(P\text{-value})$  for each CpG (Y-axis). Blue line indicates the threshold for significance (FDR  $q < 0.05$ ). Red line indicates the significance level after Bonferroni correction ( $P_{\text{adj}} < 0.05$ ). Two CpG sites in the promoter of *NLR5* on chromosome 16 reached epigenome-wide significance (cg07839457:  $P = 4.96\text{E-}09$ ; cg16411857:  $P = 2.01\text{E-}10$ ). B. Quantile-quantile plot showing observed vs. expected  $-\log_{10}(P\text{-value})$  for association at all CpG sites ( $\lambda = 1.17$ ). A regression model was applied for EWAS, adjusted for batch, cell types, and other clinical confounding factors.

Significant CpGs, relationship with gene region, and effect sizes are presented in Table 1. Consistent with previous findings, global methylation in HIV-infected individuals was lower than in uninfected individuals. Among the 20 significant CpG sites, 14 CpGs were hypomethylated, while 6 CpGs were hypermethylated in the HIV-infected samples. Eighteen of 20 CpGs were methylated at an intermediate level ( $0.2 < \beta < 0.8$ ). Methylation at 14 significant CpG sites differed by more than 5% (5–25%) between HIV-infected and uninfected subjects. The average methylation at each significant CpG in these 14 sites in HIV-infected and uninfected groups is shown in Fig. 3A. The absolute fold-changes in these 14 CpGs in HIV-infected and uninfected groups were 1.1–2.0, reflecting moderate to large effects of methylation alteration in HIV-infected individuals relative to uninfected subjects (Fig. 3B). Notably, CpGs located in

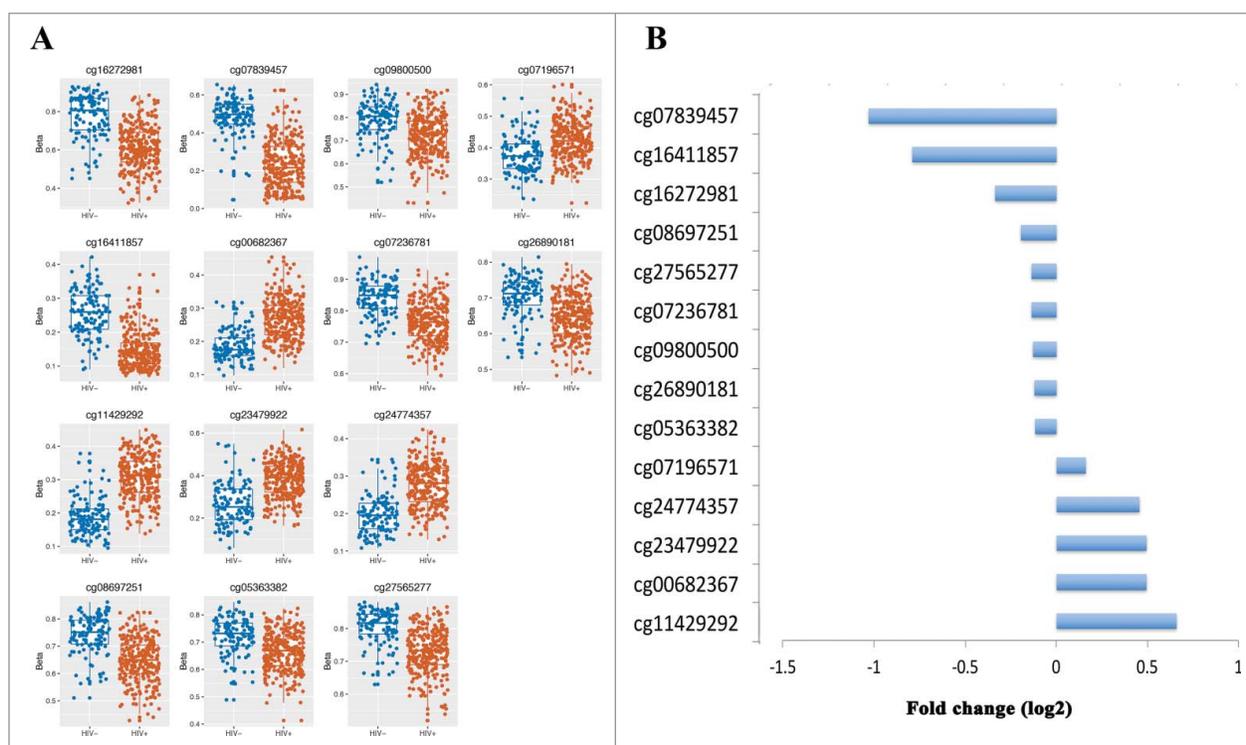
islands presented larger fold-changes between HIV-infected and uninfected groups compared to CpGs near islands.

Five of 20 CpGs located in 3 genes remained significant after more stringent Bonferroni correction ( $P_{\text{nominal}} < 10^{-7}$ , Bonferroni  $P < 0.05$ ) (cg16272981, cg16411857, cg11429292, cg08697251, cg07839457). Among these 3 genes, a novel gene for HIV infection, *NLR family, CARD domain containing gene 5 (NLRC5)*, showed the largest effect size on methylation alteration and biological relevant to HIV infection. *NLRC5* was recently identified as a major histocompatibility complex (MHC) class I transcriptional activator through genome-wide gene expression profiling of human cell lines expressing either wild type or mutant *NLRC5*.<sup>23</sup> We found that 2 CpGs in the promoter of *NLRC5*, cg07839457 and cg16411857, showed lower methylation levels in HIV-infected subjects compared to

**Table 1.** Epigenome-wide association analysis identified 20 CpG sites for HIV infection in the Veteran Aging Cohort Study.

CpG	CHR	Position	Gene	Relation to Gene	Relation to CpG Island	Beta (HIV+)	Beta (HIV-)	Fold Change ( $\log_2$ )	t	P	FDR q
<b>cg16272981</b>	5	1489889	<i>LPCAT1</i>	Body	Island	0.62	0.78	0.79	-7.63	3.07E-13	1.34E-07
<b>cg16411857</b>	16	57023191	<i>NLRC5</i>	TSS1500	Island	0.15	0.25	0.58	-6.59	2.01E-10	2.97E-05
<b>cg11429292</b>	12	6888035	<i>Near CD4</i>			0.31	0.19	1.58	6.58	2.04E-10	2.97E-05
<b>cg08697251</b>	5	1489875	<i>LPCAT1</i>	Body	Island	0.65	0.74	0.87	-6.32	9.33E-10	0.0001
<b>cg07839457</b>	16	57023022	<i>NLRC5</i>	TSS1500	N_Shore	0.24	0.48	0.49	-6.03	4.96E-09	0.0004
cg00682367	20	48963280			S_Shelf	0.27	0.19	1.41	5.45	1.04E-07	0.0076
cg18727075	3	52847250	<i>ITIH4</i>	3UTR		0.85	0.86	0.98	-5.32	2.06E-07	0.0129
cg26790059	5	156772562	<i>CYFIP2</i>	5UTR		0.79	0.85	0.93	-5.28	2.49E-07	0.0136
cg23479922	5	16179633	<i>MARCH11</i>	1stExon	Island	0.38	0.27	1.41	5.12	5.60E-07	0.0272
cg05363382	17	45767653			N_Shelf	0.67	0.72	0.92	-5.08	6.77E-07	0.0284
cg10124440	12	6960295	<i>USP5</i>	TSS1500	N_Shore	0.47	0.47	0.99	-5.07	7.14E-07	0.0284
cg09800500	12	24992256	<i>BCAT1</i>	Body		0.73	0.79	0.92	-5.03	8.49E-07	0.0299
cg07236781	1	25291041	<i>RUNX3</i>	1stExon		0.77	0.84	0.91	-5.02	8.90E-07	0.0299
cg24774357	14	77504364				0.28	0.2	1.37	5	9.77E-07	0.0305
cg26270848	1	236176807	<i>NID1</i>	Body		0.93	0.94	1	-4.98	1.07E-06	0.0312
cg27565277	16	17228474	<i>XYLT1</i>	Body		0.73	0.81	0.91	-4.91	1.49E-06	0.0407
cg23374256	6	5192344	<i>LYRM4</i>	Body		0.9	0.88	1.02	4.89	1.64E-06	0.0422
cg13177052	17	7319319	<i>NLGN2</i>	Body	N_Shore	0.81	0.82	0.99	-4.86	1.87E-06	0.0433
cg07196571	15	64442578	<i>SNX22</i>	TSS1500	N_Shore	0.42	0.38	1.12	4.86	1.88E-06	0.0433
cg26890181	19	44285940	<i>KCNN4</i>	TSS1500		0.65	0.71	0.92	-4.85	2.00E-06	0.0437

CpG sites in bold are corrected for Bonferroni test ( $P < 0.05$ )



**Figure 3.** Epigenome-wide significant differences at 14 CpG sites. **A.** Box plots of methylation  $\beta$  values for 14 significant CpG sites with average methylation difference greater than 5% between HIV-infected and uninfected individuals. **B.** Log<sub>2</sub> value of fold-changes for 14 significant CpG sites in HIV-infected vs. HIV-uninfected patients.

uninfected subjects (cg07839457:  $t = -6.03$ ,  $P_{\text{nominal}} = 4.96 \times 10^{-9}$  and cg16411857:  $t = -6.59$ ,  $P_{\text{nominal}} = 2.01 \times 10^{-10}$ ) (Fig. 4A). The distribution of methylation levels in these 2 CpGs was significantly different between the HIV-infected and uninfected groups (Fig. 4B). The methylation fold-changes between HIV-infected and uninfected subjects for cg07839457 and cg16411857 were 0.49 and 0.58, respectively. These two CpGs were located approximately 350 bp apart in the TSS1500 within a CpG island (cg16411857) or surrounding a CpG island (cg07839457) in *NLRC5*. Histone modification H3K27ac was enriched in this region. Interestingly, multiple transcription factors, such as *TEF-1*, *AP-2rep*, and *MAX*, are located within this region.<sup>24–26</sup> These data suggest that methylation alterations at these 2 CpGs may be involved in dysregulation of *NLRC5* expression in HIV infection. However, the relationship between *NLRC5* expression and HIV infection has not been investigated.

Another significant CpG site, cg11429292, which is 10 kb upstream of the *CD4* gene, was hypermethylated in HIV-infected patients compared to uninfected patients ( $t = 6.58$ ,  $P_{\text{nominal}} = 2.04 \times 10^{-10}$ ). *CD4* encodes the primary receptor for HIV-1 and regulates T helper cell function. Two highly significant CpGs (cg16272981 and cg08697251) in *LPCAT1* were differentially methylated between HIV-infected and uninfected subjects, but their relationship with HIV infection is unknown.

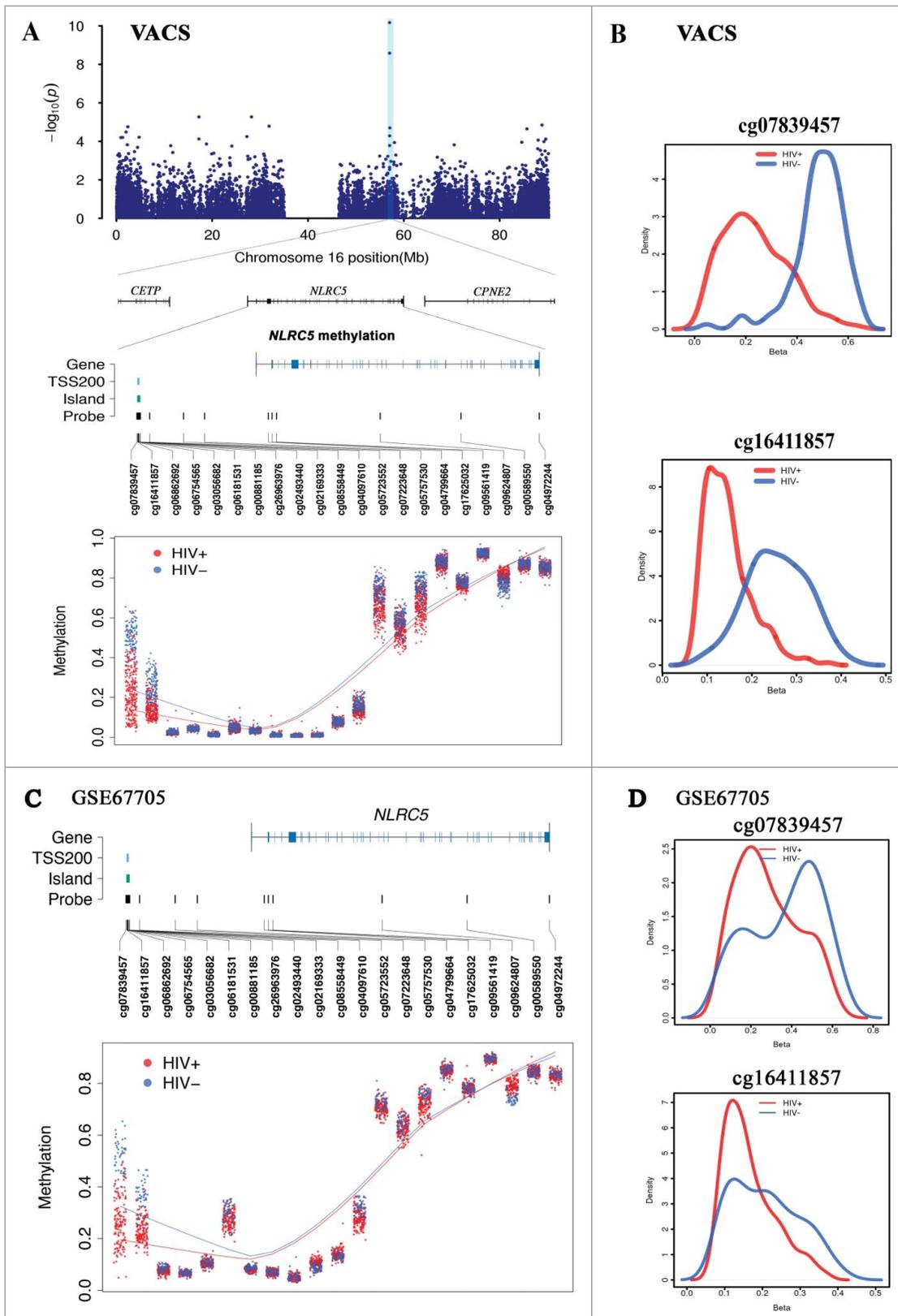
A number of other differential CpG sites reached an FDR  $q < 0.05$ , but were not retained after more stringent Bonferroni correction (Table 1). These CpGs may be biologically relevant, as they were located on genes involved in HIV-1 pathogenesis. For example, we observed reduced DNA methylation of cg07236781 in the 1st Exon of *RUNX3* ( $t = -5.02$ , FDR  $q = 0.03$ ). *RUNX3* is a member of the *RUNX* family and is a critical transcriptional regulator of  $CD4^+$  T cells. The *RUNX* protein complex may

bind to the HIV LTR and play a role in viral latency.<sup>27</sup> Another interesting differential methylated CpG site, cg07196571 in the promoter of *sorting nexin 22* *SNX22* (*SNX22*), may be involved in HIV-1 pathogenesis. Methylation of CpG in the promoter of *SNX22* showed 4% greater methylation in HIV-infected individuals than uninfected individuals in the VACS sample. A previous study showed that knockdown of *SNX22* significantly increased *CD4* expression in HIV-1-infected T cells, suggesting that *SNX22* is involved in HIV-mediated *CD4* downregulation.<sup>28</sup>

Previously reported CpG sites in candidate genes, *IL2*, *IGFBP6*, and *SATB2*, of HIV infection showed nominal significance, but were not retained after statistical correction. The CpG cg25065535, located in TSS1500 of *IL2*, was hypermethylated in HIV-infected individuals compared to uninfected individuals ( $t = 2.78$ ,  $P_{\text{nominal}} = 0.006$ ), which agrees with the results of a previous study.<sup>14</sup> We also detected hypermethylation of *IGFBP6* (cg04791129:  $t = 2.13$ ,  $P_{\text{nominal}} = 0.03$ ) and *SATB2* (cg05358815:  $t = 2.93$ ,  $P_{\text{nominal}} = 0.003$ ) in HIV-infected individuals compared to uninfected individuals, which is also consistent with the previous study.<sup>19</sup> None of these 2 CpG sites reached epigenome-wide significance.

#### Replication of methylation alterations in *NLRC5* in HIV infection (GSE67705)

Because the methylation alterations in the promoter of *NLRC5* showed large differences between HIV-infected and uninfected samples, we sought to replicate these findings in an independent sample. This dataset included 142 HIV-infected and 44 uninfected samples from 2 different cohorts. Using a similar analytic pipeline, we found significant differential methylation of the same 2 CpGs in the promoter of *NLRC5* in this sample



**Figure 4.** DNA methylation in *NLRC5* were associated with HIV infection in 2 independent data sets. A. Regional Manhattan plot of DNA methylation association analysis with HIV infection on chromosome 16, gene structure of *NLRC5*, and methylation level of 21 CpG sites in *NLRC5* between HIV-infected and uninfected groups in the Veteran Aging Cohort Study (VACS). B. Methylation distributions of cg07839457 and cg16411857 in HIV-infected patients and uninfected patients in the VACS sample. C. Methylation level of 21 CpG sites in *NLRC5* between HIV-infected and uninfected groups in the public dataset GSE67705. D. Methylation distributions of cg07839457 and cg16411857 in HIV-infected patients and uninfected patients in GSE67705.

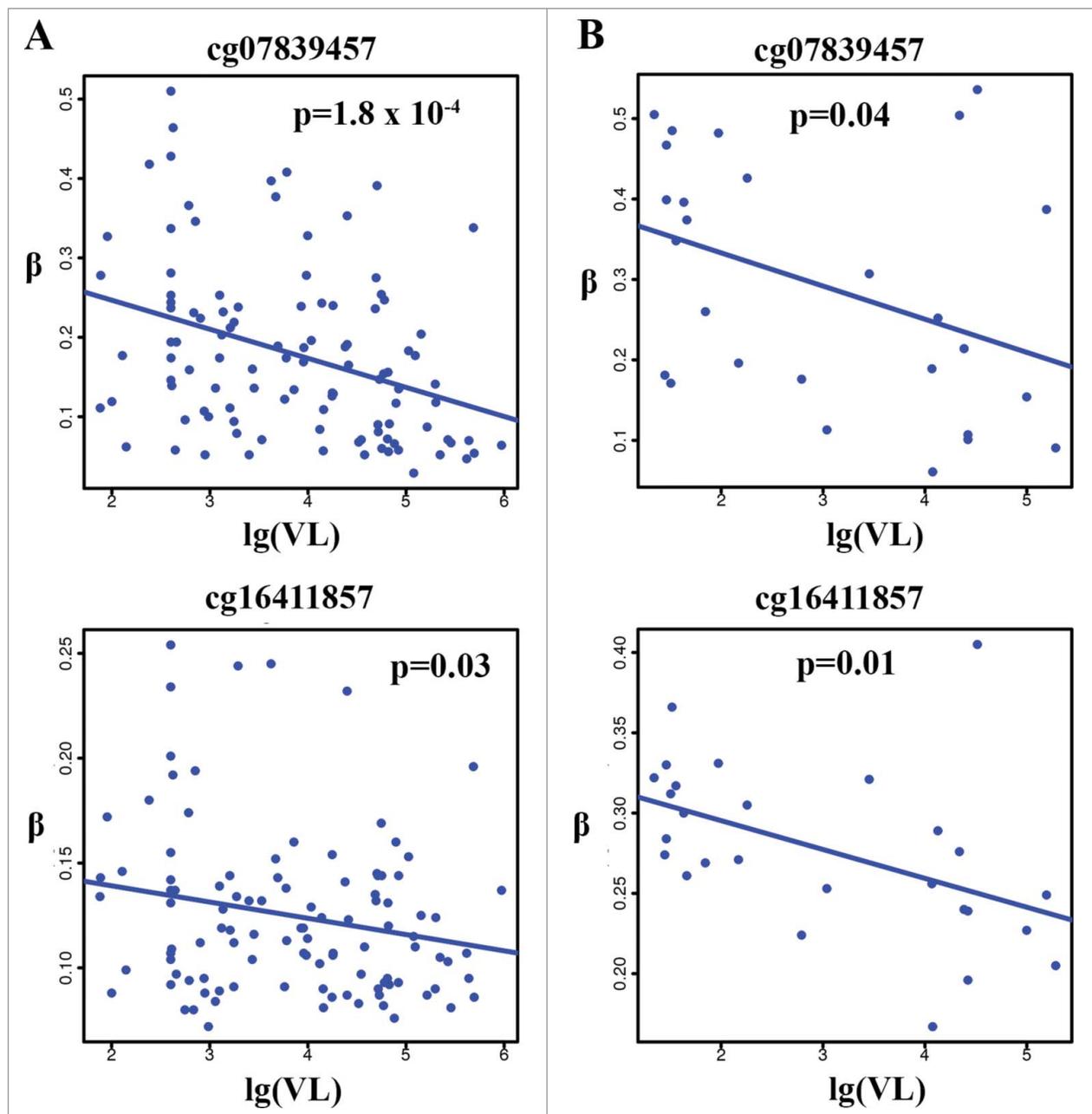
(cg07839457:  $t = -4.44$ ,  $P_{\text{nominal}} = 1.61 \times 10^{-5}$ ; cg16411857:  $t = -5.90$ ;  $P = 1.99 \times 10^{-8}$ ). Site cg07839457 showed 15% lower methylation and cg16411857 showed 13% lower methylation in the HIV-infected samples compared to the HIV-uninfected samples (Fig. 4C and D), which are consistent with the findings for *NLRC5* in the VACS sample. The results for 21 CpGs of *NLRC5* in the VACS sample and GSE67705 sample are shown in Supplemental Table 1. These findings further support the important role of *NLRC5* in HIV pathogenesis.

The other 3 significant CpG sites identified in the VACS samples showed no significant association with HIV infection in the replication sample GSE67705. The top CpG site in *LPCAT1*,

cg16272981, showed nominal significance ( $t = -2.4$ ,  $P = 0.02$ ). CpG sites associated with HIV infection in the VACS sample, cg11429292 and cg00682367, did not reach nominal significance level ( $P > 0.05$ ).

#### Methylation of *NLRC5* was correlated with HIV-1 viral load

Among the 20 HIV-associated CpGs, 3 CpGs showed a significant association with VL in the VACS samples, including 2 CpG sites in *NLRC5* (cg07839457:  $P = 1.80 \times 10^{-4}$ ; cg16411857:  $P = 0.03 \times 10^{-6}$ ). Fig. 5A shows the correlation of these 2 CpGs with unsuppressed VL ( $>75$ ) in the VACS samples. Methylation of significant CpG sites in *LPCAT1* was not correlated



**Figure 5.** Correlation between methylation levels of cg07839457 and cg16411857 in *NLRC5* and HIV-1 viral load in the 2 independent samples. Negative correlation with HIV-1 viral load in HIV-infected individuals (VL  $> 75$ ) ( $r^2 = 0.13$ ,  $P = 1.8 \times 10^{-4}$  for cg07839457 and  $r^2 = 0.04$ ,  $P = 0.03$  for cg16411857) in the VACS (A) and in an independent publicly available sample (GSE53840) (VL  $> 75$ ) ( $r^2 = 0.18$ ,  $P = 0.04$  for cg07839457 and  $r^2 = 0.15$ ,  $P = 0.01$  for cg16411857) (B).

with VL. These results suggest that alterations in DNA methylation, particularly in the promoters of *NLRC5*, are also involved in the control of HIV-1 replication in the blood. In an independent public data set from the Multicenter AIDS Cohort Study (MACS; GSE53840), methylation of cg07839457 and cg16411857 on *NLRC5* was negatively correlated with viral load in HIV-infected samples ( $P = 0.04$  for cg07839457 and  $P = 0.01$  for cg16411857), supporting the findings for the VACS samples. Fig. 5B shows the correlations of these 2 CpGs with VL in the unsuppressed samples. Despite the limited sample size, these results indicate that methylation changes in *NLRC5* play a role in HIV-1 replication.

## Discussion

HIV-1 infection is known to increase DNA methyltransferase activity in CD4<sup>+</sup> T cells *in vitro*.<sup>11,12</sup> However, this hypothesis has not been evaluated on an epigenome-wide scale in humans. Here, we report the epigenome-wide characterization of DNA methylation for HIV infection. Using an EWAS approach, we identified 20 CpGs that were differentially methylated between HIV-infected and uninfected samples. Specifically, hypomethylation of 2 CpGs (cg07839457 and cg16411857) in the promoter of *NLRC5* was associated with HIV infection. The associations of these 2 CpG sites in *NLRC5* with HIV infection were replicated in an independent sample (GSE67705). Furthermore, methylation of these 2 CpGs in *NLRC5* was negatively correlated with VL, which was also replicated in an independent cohort (MACS). Although differential DNA methylation of the candidate genes, *IL2*, *IGFBP6*, and *SATB2*, did not reach epigenome-wide significance, we found similar trends in methylation alteration as previous reports.<sup>14,19</sup> Our findings suggest that alterations in DNA methylation are associated with the host genomic response to HIV-1 infection and support a role for *NLRC5* in HIV-1 pathogenesis.

We observed significant alterations in DNA methylation in the HIV-infected host epigenome. Most of the significant CpGs showed lower methylation levels in HIV-infected subjects compared to uninfected subjects. Most significant CpGs showed intermediate levels of methylation, while a small number of CpGs showed extremely low or high methylation. The CpGs that were significantly associated with HIV infection were near or within genes known to be involved in immune activation, which is central to the pathophysiology associated with chronic viral infection. These findings support that mild to moderate methylation of these genes may have large impacts on the regulation of immune and inflammatory functions.<sup>29</sup> Although the mechanism of how HIV-1 impacts the host methylome is unknown, methylation alteration may be an epigenetic consequence of the integration of HIV-1 DNA into the infected host genome. This interpretation is based on the results of a study showing that foreign DNA fragment insertion can have important functional consequences that are beyond the site of foreign DNA insertion.<sup>30</sup> The effects of foreign DNA on the host genome were previously observed in studies in which tumor viral DNA such as adenovirus type 12 was inserted and remarkably altered methylation; such epigenetic genome-wide consequences play an important role in oncogenesis.<sup>31</sup> Recently,

insertion of a 5.6-kb plasmid in the human genome altered methylation in 3,791 of 480,000 CpG sites examined compared to non-transgenomic clones, suggesting epigenome destabilization of foreign DNA in human cells.<sup>32</sup> Therefore, it is reasonable to speculate that HIV-1 DNA insertion into the host genome can lead *de novo* methylation and decrease host epigenome stability. Future studies examining the epigenetic effects of HIV DNA insertion in HIV-infected cells will provide insight into methylation alteration in the HIV-infected genome observed in this study.

Among the significant HIV-associated CpG sites, the most interesting CpG sites were within the promoter of *NLRC5*. Both CpG sites were moderately under-methylated in HIV-infected subjects compared to uninfected subjects in 2 independent samples. NLR proteins constitute a protein family associated with the adaptive cellular immune response and play an important role in host recognition and the defense against pathogens.<sup>33-35</sup> *NLRC5* is located on chromosome 16 and is expressed in a wide variety of cell types and tissues.<sup>36</sup> *NLRC5* was recently identified as a specific transactivator of MHC class I genes and was found to interact with major class II transactivators with a similar MHC promoter-bound factor.<sup>36-39</sup> Precisely how *NLRC5* is involved in HIV-1 infection remains unknown. However, a recent study demonstrated that hypermethylation of the promoter of *NLRC5* downregulated *NLRC5* expression in cancer cells.<sup>40</sup> The expression of *NLRC5* and MHC class I genes was highly correlated with human cancers. Interestingly, methylation of *NLRC5* was also inversely associated with MHC class I gene expression.<sup>40</sup> These results suggest that *NLRC5* is a target for cancer immune evasion. Our findings of the hypomethylation of 2 CpGs in HIV-infected subjects suggests that upregulation of *NLRC5* expression enhances the transcription of MHC class I genes, which involves the mechanisms of host defense in HIV infected genome.

Additionally, methylation of *NLRC5* was inversely correlated with HIV-1 viral load in the VACS cohort, which was validated in an independent cohort, suggesting that *NLRC5* is involved in viral replication. This gene also plays a role in the cytokine response and antiviral immunity by inhibiting nuclear factor (NF)- $\kappa$ B activation and negatively regulating type I interferon (IFN) signaling pathways.<sup>41,42</sup> *NLRC5* negatively regulates the NF- $\kappa$ B and type I IFN signaling pathways and NF- $\kappa$ B and IFN upregulate HIV-1 replication.<sup>41</sup> The methylation alterations in *NLRC5* may be associated HIV replication via the upregulation of NF- $\kappa$ B and interferon pathways. However, the cause-consequence relationship between methylation of *NLRC5* and HIV infection is unknown. Hypomethylation of *NLRC5* may result from chronic HIV infection, but further studies are needed to clarify this.

The top significant signals in the EWAS were from *LPCAT1*, a newly identified protein coding gene on chromosome 5p15.33.<sup>43</sup> Two CpG sites (cg16272981 and cg08697251) in the gene body of *LPCAT1* were significantly hypomethylated in HIV-infected subjects compared to uninfected controls. However, these 2 CpG sites were not associated with HIV infection in the replication sample and neither of the 2 CpG sites was associated with VL. A previous study suggested that *LPCAT1* is involved in the regulation of inflammatory lipids.<sup>44</sup> The role of

*LPACT1* in HIV infection is unknown. Our findings suggest that the role of *LPACT1* in HIV disease should be further investigated.

We found nominally significant differential methylation in *IL2*, *IGFBP6*, and *SATB2* between HIV-infected and uninfected groups. A previous *in vitro* study reported the hypermethylation of *IL2* in CD4<sup>+</sup> T cells from HIV-infected non-controllers than in CD4<sup>+</sup> T cells from uninfected individuals.<sup>14</sup> We observed the same trend in hypermethylation for HIV infection in our patient-based sample. Similarly, we observed hypermethylation of *IGFBP6* and *SATB2* for HIV infection, which is consistent with the results of a previous twin study that included one HIV-infected twin and one uninfected twin.<sup>19</sup> These genes did not reach epigenome-wide significance, which may be because of the greater heterogeneity in our patient-based sample compared to laboratory-controlled studies and the twin sample.

One limitation of EWAS is cell type heterogeneity in the blood samples, as epigenetic changes are correlated with cellular differentiation into specific lineages.<sup>45</sup> Cell compositions significantly differ between HIV-infected and uninfected subjects and can be particularly problematic in an EWAS. However, isolation of a specific cell type is often not feasible when evaluating large clinical samples because the total cell numbers are low, particularly in cryopreserved specimens. Our results showed that the statistical approach for estimating and correcting the cell composition can substantially reduce the confounding effects of cell type differences and improve the signals associated with the underlying pathophysiology. This approach has been applied in EWAS of other complex diseases.<sup>17,46,47</sup> Notably, despite the more conservative and comprehensive analytic approach used for EWAS, the results still showed slight inflation, suggesting that other unknown confounding factors were not completely corrected. However, the findings for *NRLC5* in 2 independent samples indicate that this gene is associated with HIV infection rather than a confounding effect.

A limitation of this study was the inability to differentiate between methylation signals involved in HIV-1-induced pathophysiological processes and cell type-specific methylation. Here, we treated cell type as a confounding factor to limit the generation of spurious signals. However, a reduced CD4<sup>+</sup> T cell count is a clinical marker of HIV progression. The adjusted cell composition may suppress HIV-related methylation signals, which may be also methylation signatures of CD4<sup>+</sup> T cells.<sup>48</sup> This may explain why known genes involved in HIV pathogenesis, such as *CCR5*, did not reach epigenome-wide significance. Future independent studies including longitudinal studies of patients starting antiretroviral therapy are necessary to replicate these findings. We were also unable to examine the functional roles of the identified methylation sites in HIV infection. Future studies investigating the regulatory mechanism of methylation alteration in the identified genes will provide additional insight into the epigenetic mechanisms.

The study had several strengths: 1) This is the first EWAS with sufficient power to characterize altered DNA methylation patterns of the host genome associated with HIV disease; 2) Biologically meaningful methylation of *NLRC5* with large effect sizes was replicated in an independent sample; 3) Methylation in *NLRC5* was associated with HIV VL in 2 independent samples. The differential

methylation signals observed highlight the potential of EWAS to explore the pathophysiology of HIV disease in humans and is an important step toward improving the understanding of the mechanisms involved in host epigenetic regulation of HIV infection.

## Materials and methods

### Sample description

The study was approved by the Human Research Protection Program of Yale University and Institutional Research Board committee at the Connecticut Veteran Healthcare System, West Haven Campus. All subjects in the VACS sample were informed of the study and each subject provided written consent for participation.

### VACS dataset for EWAS

All subjects were participants in the VACS, an observational, prospective longitudinal study of HIV-infected and age, race/ethnicity, sex, and site-matched uninfected veterans in care at 8 US Department of Veterans Affairs medical centers across the US.<sup>49</sup> Total white blood cell counts and CD4<sup>+</sup> and CD8<sup>+</sup> T cell subsets were enumerated at time of peripheral blood sample collection. HIV VL was measured per standard of care by PCR as copies per milliliter. We selected 261 HIV-infected and 117 uninfected patients for this study. Genomic DNA was extracted from whole blood using a standard method. Demographic and clinical characteristics are presented in Table 2.

### EWAS replication data set (GSE67705)

A dataset from the study “Methylome-wide analysis of chronic HIV infected patients and healthy controls” (GSE67705) (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67705>) was used as a replication sample. A total of 186 samples (HIV-infected: 142; HIV-uninfected: 44) were available for analysis. HIV-infected samples were obtained from the CNS HIV Antiretroviral Therapy Effects Research (CHARTER) study (<http://www.charterresource.ucsd.edu/>). An additional 35 HIV-infected samples were included in the GSE67705. All subjects were male, of European descent, and 25 to 65 y old. HIV-infected subjects were on cART and VL was controlled at less

**Table 2.** Demographic and clinical characteristics of HIV-infected and HIV-uninfected patients from the Veteran Aging Cohort Study.

	HIV-infected (n = 261)	HIV-uninfected (n = 117)	P
Age (years)	48.80 ± 8.0	50.19 ± 8.71	0.155
Male (%)	96.9	88.8	0.004
Race (%)			0.744
EA	22.2	25.6	
AA	67	62.4	
Other	10.7	12.0	
Smoking (%)	56	48.3	0.203
AUDIT-C	4.12 ± 3.14	4.45 ± 3.43	0.401
CD4 <sup>+</sup> T cell	463 ± 311	NA	NA
Viral load (log <sub>10</sub> )	2.68 ± 1.22	NA	NA
cART (%)	79.1	NA	NA

EA: European American; AA: African American; AUDIT-C: Alcohol Use Identification Test; cART: combination antiretroviral therapy.

than 20 copies/mL. Because of the limited sample size, this data set was used to replicate the significant CpG sites related to HIV infection identified in the VACS sample.

### Multicenter AIDS cohort study (MACS, GSE53840)

This dataset was obtained from a public database (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53840>). This cohort included 109 HIV-infected men with a mean age of 52 years (31–68 years). Average CD4<sup>+</sup> and CD8<sup>+</sup> T cell percentages were 30% (2–54%) and 46% (19–79%), respectively. Non-detectable VL was defined as a VL lower than 10 copies/mL as reported in the database. Genomic DNA was extracted from peripheral blood mononuclear cells. HIV-uninfected subjects were not included in this cohort.<sup>50</sup> Thus, this data set was used as a validation sample to examine the correlations between HIV-associated significant CpG sites and VL identified from the VACS.

### DNA methylation profiling, cell type estimation, and batch effects in the VACS sample

DNA was extracted from whole-blood samples using PaxGene collection tubes (QIAGEN, Hilden, Germany) and FlexiGene DAN extraction kits (QIAGEN). Genomic DNA methylation profiling was conducted at the Yale Center for Genomic Analysis (<https://medicine.yale.edu/keck/ycga/>) using the Illumina Infinium HumanMethylation450 Beadchip (450K; Illumina, San Diego, CA, USA). To reduce batch effects, all samples were randomly placed on each array. The same scientist (G Wang at the Yale Center for Genomic Analysis) who was blinded to the HIV status conducted the microarray experiment. Quality control and probe normalization were performed as previously described by Lehne et al.<sup>20</sup> Detailed information regarding this procedure is presented in the supplementary material.

Using an algorithm developed by Houseman et al.,<sup>48,54</sup> we applied the 600 probes from the 450K array, which included the signature methylation sites of 6 cell types (CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, NK T cells, B cells, monocytes, and granulocytes) in the blood,<sup>48</sup> and estimated the cell composition in each sample using the minfi R package. To validate the accuracy of the estimated cell composition, we compared the correlation between the methylation-estimated and clinical laboratory reports of CD4<sup>+</sup> and CD8<sup>+</sup> T cell percentage in HIV-infected samples. The methylation-estimated CD4<sup>+</sup> and CD8<sup>+</sup> percentages were highly correlated with the laboratory measurements ( $r = 0.76$ ,  $P < 0.0001$  for CD4<sup>+</sup>,  $r = 0.77$ ,  $P < 0.0001$  for CD8<sup>+</sup>) (Supplementary Fig. S4A and B). As expected, the percentage of CD4<sup>+</sup> T cells was significantly lower ( $t = -12.1$ ,  $P = 1.03 \times 10^{-25}$ ) and the percentage of CD8<sup>+</sup> T cells was significantly higher ( $t = 16.6$ ,  $P = 1.38 \times 10^{-45}$ ) in HIV-infected patients than in HIV-uninfected patients (Supplementary Fig. S4C). Therefore, the 6 methylation-estimated cell type proportions were adjusted as covariates in the subsequent EWAS analytic model.

To further detect whether batch and other clinical variables confounded HIV infection in our sample, we estimated the correlation between each potential confounder and the first 30 principal components of control probes designed for the 450K array (Fig. S3A). We then performed a linear regression to

adjust significant confounders. Fig. S3B showed no significant correlation between principal components on residual methylation and confounders except smoking status, which was covariate in the final regression model. Only male samples were analyzed to avoid sex bias.

### Statistical analysis

To adjust for significant global confounding factors in the VACS sample as described by Lehne et al.,<sup>20</sup> we conducted 2 regression analyses to determine the associations between methylome-wide CpGs and HIV infection. The following steps were performed to correct for global covariations that may confound specific methylation in HIV infection.

We first estimated residual  $\beta$  using regression model (1):

$$\begin{aligned} \beta \text{ value (quantile normalized, QN)} &\sim \text{age} + \text{race} \\ &+ \text{smoking status} + \text{White Blood Cell (WBC)} \\ &+ 6 \text{ cell type proportions} + \text{PCs 1–30 on intensities of control probes} \end{aligned} \quad (1)$$

We then performed a second PCA on the resulting regression residual  $\beta$  values and regressed out the first 5 PCs to further control for confounders in the regression model (2).

$$\begin{aligned} \beta \text{ value (QN)} &\sim \text{HIV} + \text{age} + \text{race} + \text{smoking status} \\ &+ \text{WBC} + 6 \text{ cell types} + \text{PC 1–30 on control probes} \\ &+ \text{PCs 1–5 on residuals from model (1)}. \end{aligned} \quad (2)$$

Significance was set at a false discovery error (FDR)  $< 0.05$ . To further confirm the correction of global confounders, we performed Pearson correlation analysis between the first 30 PCs on residual methylation from model (1) and batch, demographic, clinical, and 6 cell type confounders.

To validate our analytic pipeline, we performed an EWAS for age because age-associated DNA methylation has been well established.<sup>51,52</sup> EWAS revealed 5,332 CpG sites for age in the VACS sample. The top significant aging-related CpG site was cg16867657 on *ELOVL2* (Supplementary Fig. S5); this gene is well known to be related to aging.<sup>53,54</sup> Other aging-related genes involved in differential methylation, such as *FHL2* and *KLF14*, were replicated among the most significant CpG sites in our sample.<sup>51</sup> Supplemental Table 2 lists the top 50 significant CpG sites associated with age in our sample. These results validated our data collection, processing, normalization, and analysis.

To examine whether HIV-associated CpG methylation identified by EWAS is correlated with HIV-1 viral replication, we performed linear regression for significant CpG sites derived from model (2) with VL, adjusted for age, sex, and race. We excluded HIV-1 suppressors with VL  $< 75$  in the VACS sample. The relationship between methylation at significant CpG sites and HIV-1 load identified in the VACS cohort was confirmed in an independent cohort from GSE53840 as described above.

## Data availability

Demographic, clinical variables, and methylation for the VACS sample were submitted to the GEO dataset (GSE77696).

## Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

## Acknowledgments

The authors appreciate the support of the Veteran Aging Study Cohort and Yale Center of Genomic Analysis.

## Funding

The work was funded by the National Institute on Drug Abuse [R03 DA039745, K12 DA000167, R01 DA038632, R01 DA042691]; National Institute on Alcoholism and Alcohol Abuse [R21AA022870]; and National Center for Post-traumatic Stress Disorder.

## Author contributions

Xinyu Zhang was responsible for bioinformatics data processing and statistical analysis. Amy Justice provided DNA samples and clinical data and contributed to the interpretation of results and manuscript preparation. Zuohe Wang was involved in statistical modeling and data analysis. Hongyu Zhao contributed to the analytic strategy and manuscript preparation. Ying Hu was involved in establishing the bioinformatics pipeline. Guilin Wang conducted the microarray-based DNA methylation assay. Eric Johnson contributed to manuscript preparation. Richard Sutton interpreted the results and assisted with manuscript preparation. Brinda Emu contributed to data interpretation and manuscript preparation. John Krystal contributed to data interpretation and manuscript preparation. Ke Xu was responsible for the study design, study protocol, sample preparation, data analysis, interpretation, and manuscript preparation.

## References

1. Ay E, Banati F, Mezei M, Bakos A, Niller HH, Buzas K, Minarovits J. Epigenetics of HIV infection: promising research areas and implications for therapy. *AIDS Rev* 2013; 15(3):181-8; PMID:24002202
2. Paschos K, Allday MJ. Epigenetic reprogramming of host genes in viral and microbial pathogenesis. *Trends Microbiol* 2010; 18(10):439-47; PMID:20724161; <http://dx.doi.org/10.1016/j.tim.2010.07.003>
3. Harbers K, Schnieke A, Stuhlmann H, Jahner D, Jaenisch R. DNA methylation and gene expression: endogenous retroviral genome becomes infectious after molecular cloning. *Proc Natl Acad Sci U S A* 1981; 78(12):7609-13; PMID:6950402; <http://dx.doi.org/10.1073/pnas.78.12.7609>
4. Kumar A, Darcis G, Van Lint C, Herbein G. Epigenetic control of HIV-1 post integration latency: implications for therapy. *Clin Epigenetics* 2015; 7(1):103; PMID:26405463; <http://dx.doi.org/10.1186/s13148-015-0137-6>
5. Park J, Lim CH, Ham S, Kim SS, Choi BS, Roh TY. Genome-wide analysis of histone modifications in latently HIV-1 infected T cells. *AIDS* 2014; 28(12):1719-28; PMID:24762674; <http://dx.doi.org/10.1097/QAD.0000000000000309>
6. Bednarik DP, Cook JA, Pitha PM. Inactivation of the HIV LTR by DNA CpG methylation: evidence for a role in latency. *EMBO J* 1990; 9(4):1157-64; PMID:2323336
7. Kauder SE, Bosque A, Lindqvist A, Planelles V, Verdin E. Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog* 2009; 5(6):e1000495; PMID:19557157; <http://dx.doi.org/10.1371/journal.ppat.1000495>
8. Tripathy MK, Abbas W, Herbein G. Epigenetic regulation of HIV-1 transcription. *Epigenomics* 2011; 3(4):487-502; PMID:22126207; <http://dx.doi.org/10.2217/epi.11.61>
9. Blazkova J, Trejbalova K, Gondois-Rey F, Halfon P, Philibert P, Guiguen A, Verdin E, Olive D, Van Lint C, Hejnar J, et al. CpG methylation controls reactivation of HIV from latency. *PLoS Pathog* 2009; 5(8):e1000554; PMID:19696893; <http://dx.doi.org/10.1371/journal.ppat.1000554>
10. Deeks SG, Walker BD. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity* 2007; 27(3):406-16; PMID:17892849; <http://dx.doi.org/10.1016/j.immuni.2007.08.010>
11. Mikovits JA, Young HA, Vertino P, Issa JP, Pitha PM, Turcoski-Corales S, Taub DD, Petrow CL, Baylin SB, Ruscetti FW. Infection with human immunodeficiency virus type 1 upregulates DNA methyltransferase, resulting in de novo methylation of the gamma interferon (IFN-gamma) promoter and subsequent downregulation of IFN-gamma production. *Mol Cell Biol* 1998; 18(9):5166-77; PMID:9710601; <http://dx.doi.org/10.1128/MCB.18.9.5166>
12. Fang JY, Mikovits JA, Bagni R, Petrow-Sadowski CL, Ruscetti FW. Infection of lymphoid cells by integration-defective human immunodeficiency virus type 1 increases de novo methylation. *J Virol* 2001; 75(20):9753-61; PMID:11559808; <http://dx.doi.org/10.1128/JVI.75.20.9753-9761.2001>
13. Pion M, Jaramillo-Ruiz D, Martinez A, Munoz-Fernandez MA, Correa-Rocha R. HIV infection of human regulatory T cells downregulates Foxp3 expression by increasing DNMT3b levels and DNA methylation in the FOXP3 gene. *AIDS* 2013; 27(13):2019-29; PMID:24201117; <http://dx.doi.org/10.1097/QAD.0b013e32836253fd>
14. Nakayama-Hosoya K, Ishida T, Youngblood B, Nakamura H, Hosoya N, Koga M, Koibuchi T, Iwamoto A, Kawana-Tachikawa A. Epigenetic repression of interleukin 2 expression in senescent CD4+ T cells during chronic HIV type 1 infection. *J Infect Dis* 2015; 211(1):28-39; PMID:25001463; <http://dx.doi.org/10.1093/infdis/jiu376>
15. Gross AM, Jaeger PA, Kreisberg JF, Licon K, Jepsen KL, Khosroheidari M, Morse BM, Swindells S, Shen H, Ng CT, et al. Methyloome-wide analysis of chronic HIV infection reveals five-year increase in biological age and epigenetic targeting of HLA. *Mol Cell* 2016; 62(2):157-68; PMID:27105112; <http://dx.doi.org/10.1016/j.molcel.2016.03.019>
16. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, Eaton ML, Keenan BT, Ernst J, McCabe C, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci* 2014; 17(9):1156-63; PMID:25129075; <http://dx.doi.org/10.1038/nn.3786>
17. Liang L, Willis-Owen SA, Laprise C, Wong KC, Davies GA, Hudson TJ, Binia A, Hopkin JM, Yang IV, Grundberg E, et al. An epigenome-wide association study of total serum immunoglobulin E concentration. *Nature* 2015; 520(7549):670-4; PMID:25707804; <http://dx.doi.org/10.1038/nature14125>
18. Paul DS, Beck S. Advances in epigenome-wide association studies for common diseases. *Trends Mol Med* 2014; 20(10):541-3; PMID:25092140; <http://dx.doi.org/10.1016/j.molmed.2014.07.002>
19. Zhang Y, Li SK, Yi Yang K, Liu M, Lee N, Tang X, Wang H, Liu L, Chen Z, Zhang C, et al. Whole genome methylation array reveals the down-regulation of IGFBP6 and SATB2 by HIV-1. *Sci Rep* 2015; 5:10806; PMID:26039376; <http://dx.doi.org/10.1038/srep10806>
20. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, Afzal U, Scott J, Jarvelin MR, Elliott P, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol* 2015; 16:37; PMID:25853392; <http://dx.doi.org/10.1186/s13059-015-0600-x>
21. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012; 13:86; PMID:22568884; <http://dx.doi.org/10.1186/1471-2105-13-86>
22. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol* 2014; 15(2):R31; PMID:24495553; <http://dx.doi.org/10.1186/gb-2014-15-2-r31>

23. Meissner TB, Li A, Biswas A, Lee KH, Liu YJ, Bayir E, Iliopoulos D, van den Elsen PJ, Kobayashi KS. NLR family member NLRC5 is a transcriptional regulator of MHC class I genes. *Proc Natl Acad Sci U S A* 2010; 107(31):13794-9; PMID:20639463; <http://dx.doi.org/10.1073/pnas.1008684107>
24. Anbanandam A, Albarado DC, Nguyen CT, Halder G, Gao X, Veeraraghavan S. Insights into transcription enhancer factor 1 (TEF-1) activity from the solution structure of the TEA domain. *Proc Natl Acad Sci U S A* 2006; 103(46):17225-30; PMID:17085591; <http://dx.doi.org/10.1073/pnas.0607171103>
25. Roth C, Schuierer M, Gunther K, Buettner R. Genomic structure and DNA binding properties of the human zinc finger transcriptional repressor AP-2rep (KLF12). *Genomics* 2000; 63(3):384-90; PMID:10704285; <http://dx.doi.org/10.1006/geno.1999.6084>
26. Hurlin PJ, Ayer DE, Grandori C, Eisenman RN. The Max transcription factor network: involvement of Mad in differentiation and an approach to identification of target genes. *Cold Spring Harb Symp Quant Biol* 1994; 59:109-16; PMID:7587059; <http://dx.doi.org/10.1101/SQB.1994.059.01.014>
27. Klase Z, Yedavalli VS, Houzet L, Perkins M, Maldarelli F, Brenchley J, Strebel K, Liu P, Jeang KT. Activation of HIV-1 from latent infection via synergy of RUNX1 inhibitor Ro5-3335 and SAHA. *PLoS Pathog* 2014; 10(3):e1003997; PMID:24651404; <http://dx.doi.org/10.1371/journal.ppat.1003997>
28. Landi A, Vermeire J, Iannucci V, Vanderstraeten H, Naessens E, Bentahir M, Verhasselt B. Genome-wide shRNA screening identifies host factors involved in early endocytic events for HIV-1-induced CD4 down-regulation. *Retrovirology* 2014; 11:118; PMID:25496667; <http://dx.doi.org/10.1186/s12977-014-0118-4>
29. Mehta S, Jeffrey KL. Beyond receptors and signaling: epigenetic factors in the regulation of innate immunity. *Immunol Cell Biol* 2015; 93(3):233-44; PMID:25559622; <http://dx.doi.org/10.1038/icb.2014.101>
30. Doerfler W. Beware of manipulations on the genome: epigenetic destabilization through (foreign) DNA insertions. *Epigenomics* 2016; 8(5):587-91; PMID:26997469; <http://dx.doi.org/10.2217/epi-2016-0019>
31. Doerfler W. Impact of foreign DNA integration on tumor biology and on evolution via epigenetic alterations. *Epigenomics* 2012; 4(1):41-9; PMID:22332657; <http://dx.doi.org/10.2217/epi.11.111>
32. Weber S, Hofmann A, Herms S, Hoffmann P, Doerfler W. Destabilization of the human epigenome: consequences of foreign DNA insertions. *Epigenomics* 2015; 7(5):745-55; PMID:26088384; <http://dx.doi.org/10.2217/epi.15.40>
33. Kobayashi KS, van den Elsen PJ. NLRC5: a key regulator of MHC class I-dependent immune responses. *Nat Rev Immunol* 2012; 12(12):813-820; PMID:23175229; <http://dx.doi.org/10.1038/nri3339>
34. Lamkanfi M, Kanneganti TD. Regulation of immune pathways by the NOD-like receptor NLRC5. *Immunobiology* 2012; 217(1):13-6; PMID:22024701; <http://dx.doi.org/10.1016/j.imbio.2011.08.011>
35. Yao Y, Wang Y, Chen F, Huang Y, Zhu S, Leng Q, Wang H, Shi Y, Qian Y. NLRC5 regulates MHC class I antigen presentation in host defense against intracellular pathogens. *Cell Res* 2012; 22(5):836-47; PMID:22491475; <http://dx.doi.org/10.1038/cr.2012.56>
36. Staehli F, Ludigs K, Heinz LX, Seguin-Estevez Q, Ferrero I, Braun M, Schroder K, Rebsamen M, Tardivel A, Mattmann C, et al. NLRC5 deficiency selectively impairs MHC class I-dependent lymphocyte killing by cytotoxic T cells. *J Immunol* 2012; 188(8):3820-8; PMID:22412192; <http://dx.doi.org/10.4049/jimmunol.1102671>
37. Neerincx A, Rodriguez GM, Steimle V, Kufer TA. NLRC5 controls basal MHC class I gene expression in an MHC enhanceosome-dependent manner. *J Immunol* 2012; 188(10):4940-50; PMID:22490867; <http://dx.doi.org/10.4049/jimmunol.1103136>
38. Biswas A, Meissner TB, Kawai T, Kobayashi KS. Cutting edge: impaired MHC class I expression in mice deficient for Nlr5/class I transactivator. *J Immunol* 2012; 189(2):516-20; PMID:22711889; <http://dx.doi.org/10.4049/jimmunol.1200064>
39. Meissner TB, Liu YJ, Lee KH, Li A, Biswas A, van Eggermond MC, van den Elsen PJ, Kobayashi KS. NLRC5 cooperates with the RFX transcription factor complex to induce MHC class I gene expression. *J Immunol* 2012; 188(10):4951-8; PMID:22490869; <http://dx.doi.org/10.4049/jimmunol.1103160>
40. Yoshihama S, Roszik J, Downs I, Meissner TB, Vijayan S, Chapuy B, Sidiq T, Shipp MA, Lizee GA, Kobayashi KS. NLRC5/MHC class I transactivator is a target for immune evasion in cancer. *Proc Natl Acad Sci U S A* 2016; 113(21):5999-6004; PMID:27162338; <http://dx.doi.org/10.1073/pnas.1602069113>
41. Cui J, Zhu L, Xia X, Wang HY, Legras X, Hong J, Ji J, Shen P, Zheng S, Chen ZJ, et al. NLRC5 negatively regulates the NF-kappaB and type I interferon signaling pathways. *Cell* 2010; 141(3):483-96; PMID:20434986; <http://dx.doi.org/10.1016/j.cell.2010.03.040>
42. Tong Y, Cui J, Li Q, Zou J, Wang HY, Wang RF. Enhanced TLR-induced NF-kappaB signaling and type I interferon responses in NLRC5 deficient mice. *Cell Res* 2012; 22(5):822-35; PMID:22473004; <http://dx.doi.org/10.1038/cr.2012.53>
43. Soupene E, Kuypers FA. Phosphatidylcholine formation by LPCAT1 is regulated by Ca(2+) and the redox status of the cell. *BMC Biochem* 2012; 13:8; PMID:22676268; <http://dx.doi.org/10.1186/1471-2091-13-8>
44. Harayama T, Shindou H, Ogasawara R, Suwabe A, Shimizu T. Identification of a novel noninflammatory biosynthetic pathway of platelet-activating factor. *J Biol Chem* 2008; 283(17):11097-106; PMID:18285344; <http://dx.doi.org/10.1074/jbc.M708909200>
45. Verma M. Epigenetic regulation of HIV, AIDS, and AIDS-related malignancies. *Methods Mol Biol* 2015; 1238:381-403; PMID:25421672; [http://dx.doi.org/10.1007/978-1-4939-1804-1\\_21](http://dx.doi.org/10.1007/978-1-4939-1804-1_21)
46. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, Reinus L, Acevedo N, Taub M, Ronninger M, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat Biotechnol* 2013; 31(2):142-7; PMID:23334450; <http://dx.doi.org/10.1038/nbt.2487>
47. Langevin SM, Houseman EA, Accomando WP, Koestler DC, Christensen BC, Nelson HH, Karagas MR, Marsit CJ, Wiencke JK, Kelsey KT. Leukocyte-adjusted epigenome-wide association studies of blood from solid tumor patients. *Epigenetics* 2014; 9(6):884-95; PMID:24671036; <http://dx.doi.org/10.4161/epi.28575>
48. Nelson G, Jia L, Elloumi F, Binns E, Goedert JJ, Limou S, Winkler C. Adjusting Infinium methylation profiles to suppress signals from varying cell proportion. *Am Soc Hum Genet* 2015; PMID:26404086
49. Fultz SL, Skanderson M, Mole LA, Gandhi N, Bryant K, Crystal S, Justice AC. Development and verification of a "virtual" cohort using the National VA Health Information System. *Med Care* 2006; 44(8 Suppl 2):S25-30; PMID:16849965; <http://dx.doi.org/10.1097/01.mlr.0000223670.00890.74>
50. Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. *J Infect Dis* 2015; 212(10):1563-73; PMID:25969563; <http://dx.doi.org/10.1093/infdis/jiv277>
51. Reynolds LM, Taylor JR, Ding J, Lohman K, Johnson C, Siscovick D, Burke G, Post W, Shea S, Jacobs DR, Jr, et al. Age-related variations in the methylome associated with gene expression in human monocytes and T cells. *Nat Commun* 2014; 5:5366; PMID:25404168; <http://dx.doi.org/10.1038/ncomms6366>
52. Bacalini MG, Boattini A, Gentilini D, Giampieri E, Pirazzini C, Giuliani C, Fontanesi E, Remondini D, Capri M, Del Rio A, et al. A meta-analysis on age-associated changes in blood DNA methylation: results from an original analysis pipeline for Infinium 450k data. *Aging (Albany NY)* 2015; 7(2):97-109; PMID:25701668; <http://dx.doi.org/10.18632/aging.100718>
53. Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, Di Blasio AM, Gentilini D, Vitale G, Collino S, et al. Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell* 2012; 11(6):1132-4; PMID:23061750; <http://dx.doi.org/10.1111/accel.12005>
54. Bekaert B, Kamalandua A, Zapico SC, Van de Voorde W, Decorte R. Improved age determination of blood and teeth samples using a selected set of DNA methylation markers. *Epigenetics* 2015; 10(10):922-30; PMID:26280308; <http://dx.doi.org/10.1080/15592294.2015.1080413>