Data in Brief

Genome-wide analysis of DNA methylation associated with HIV infection based on a pair of monozygotic twins

Yinfeng Zhang a, Sai-Kam Li a, Stephen Kwok-Wing Tsui a,b,c,⁎

a School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong
b Hong Kong Bioinformatics Center, The Chinese University of Hong Kong, Hong Kong
c Centre for Microbial Genomics and Proteomics, The Chinese University of Hong Kong, Hong Kong

Abstract

Alteration of DNA methylation in mammalian cells could be elicited by many factors, including viral infections [1]. HIV has shown the ability to interact with host cellular factors to change the methylation status of some genes [2–4]. However, the change of the DNA methylation associated with HIV infection based on the whole genome has not been well illustrated. In this study, a unique pair of monozygotic twins was recruited: one of the twins was infected with HIV without further anti-retroviral therapy while the other one was healthy, which could be considered as a relatively ideal model for profiling the alterations of DNA methylation associated with HIV infection. Therefore, using methylated DNA immunoprecipitation–microarray method (MeDIP–microarray), we found the increased DNA methylation level in peripheral blood mononuclear cells from HIV infected twin compared to her normal sibling. Moreover, several distinguished differential methylation regions (DMRs) in HIV infected twin worth further study. The raw data has been deposited in Gene Expression Omnibus (GEO) datasets with reference number GSE68028.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Specifications

Organism/cell line/tissue: Human peripheral blood mononuclear cells
Sex: Female
Sequencer or array type: NimbleGen Human DNA Methylation 2.1M Deluxe Promoter Array
Data format: Raw and processed
Experimental factors: One of the monozygotic twins was infected with HIV while the other one was not. In addition, the HIV+ twin did not take any anti-retrovirus therapy.
Experimental features: DNA methylation from both twins was compared using the MeDIP–microarray to identify the alterations associated with HIV infection.
Consent: N/A
Sample source location: China

1. Deposited data

1.1. Direct link to deposited files

http://datalink.elsevier.com/midas/datalink/api/downloadfiles?items=16069-16070-16071.

1.2. Direct link to deposited genomic data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE68028.

2. Experimental design, materials and methods

2.1. Peripheral blood mononuclear cell isolation and genomic DNA extraction

Peripheral blood mononuclear cells (PBMCs) were separated from the EDTA-blood collected from the pair of monozygotic twins. Both twins were 15 years old at the time of blood collection. In addition, it was around 7 years since the HIV infection occurred in HIV infected twin without receiving any anti-retrovirus treatments. The whole blood was mixed with the OptiPrep™ by repeated inversion (1/8 volume of the blood). Then, 0.5–1 ml RPMI1640 medium was added and
centrifuged at 900 g for 30 min at 20 °C. Later, the middle layer comprised of the PBMC was collected and washed by phosphate buffered saline twice. Total genomic DNA was extracted from the PBMCs using QIAamp DNA blood mini Kit (QIAGEN) according to the manufacturer’s instructions. Genomic DNA quality and quantity were determined by a NanoDrop 2000c (Thermo Scientific).

2.2. MeDIP–microarray analysis

The sonication was employed for genomic DNA fragmentation ranging from 250 to 500 bp in length. Then the Methylated-DNA IP Kit (Zymo Research) was used for the immunoprecipitation of cytosine methylated DNA fragments from both twins. Afterwards, DNA labeling and hybridization were performed according to NimbleGen’s standard protocol. The immunoprecipitated CpG-methylated DNA from HIV positive subject (test) and from HIV negative subject (input control) was labeled with fluorescent dyes Cy5 and Cy3 using NimbleGen Dual-Color DNA Labeling Kit (Roche–NimbleGen). The combined test and input control DNA samples were suspended in hybridization buffer (Roche–NimbleGen) and co-hybridized onto NimbleGen Human DNA Methylation 2.1M Deluxe Promoter Array for 20 h at 42 °C, following washing with the Wash Kit (Roche–NimbleGen).

2.3. Microarray data analysis

The raw image files were obtained by MS 200 Microarray Scanner and MS 200 Data Collection Software. Then, the DEVA version 1.2.1 software (Data Extraction Visualization Analysis software, Roche.

Table 1
The differential methylated regions in HIV+ twin with peak value above 5.0.

| Data index | Chromosome | Start   | End     | Peak value | Feature track            | Feature strand | Name        | CpG island |
|------------|------------|---------|---------|------------|--------------------------|----------------|-------------|------------|
| 3115       | chr19      | 52,390,044 | 52,391,710 | 5.9178     | Transcription start site | ZNF577         | Yes         |
| 4258       | chr12      | 133,308,491 | 133,310,370 | 5.9045     | Primary Transcript       | ANKLE2         | Yes         |

Table 2
The differential methylated regions in HIV+ twin with peak value above 5.0.

| Data index | Chromosome | Start   | End     | Peak value | Feature track            | Feature strand | Name        | CpG island |
|------------|------------|---------|---------|------------|--------------------------|----------------|-------------|------------|
| 408        | chr4       | 3,310,903  | 3,312,010 | 5.4753     | Transcription start site | RCS12          | Yes         |
| 2640       | chrX       | 85,510,778 | 85,511,909 | 5.2014     | Primary Transcript       | DACH2          | No          |
| 1047       | chr11      | 1.05E+08  | 1.05E+08 | 5.1255     | Transcription start site | CASP5          | No          |
| 852        | chr8       | 1.44E+08  | 1.44E+08 | 5.0597     | Transcription start site | GPR182         | No          |
| 2119       | chr2       | 1.13E+08  | 1.13E+08 | 5.0144     | Transcription start site | c3orf79        | No          |

Y. Zhang et al. / Genomics Data 6 (2015) 12-15
NimbleGen) was used for further analysis. Then, the result of microarray was analyzed by DEVA version 1.2.1 software (Roche, NimbleGen) using default parameters. In detail, 100929_HG19_Deluxe_Prom_Meth_HX1.ncd, 100929_HG19_Deluxe_Prom_Meth_HX1.ndf, 100929_HG19_Deluxe_Prom_Meth_HX1.pos and 100929_HG19_Deluxe_Prom_Meth_HX1.gff files were selected for organisms. The proper annotation files were selected for the identification of the features of the probes. After analyzing, the results of log2 ratios, P-score derived from Kolmogorov–Smirnov (KS) test were processed and obtained. The final results were presented as peak value based on P-score. The larger peak value from the designated region indicated that its differential methylation level was higher. The threshold for defining the differential methylation regions (DMRs) was set to 3.0. In Tables 1 and 2, the resulting list of the most differentially methylated regions (peak value > 5.0) from the HIV+ twin and HIV− twin was shown. In addition, the hyper-methylated genes identified in HIV+ twin were combined with the reported host genes required for HIV infection [5] to predict the potential protein–protein interactions by STRING [6]. The potential or known relationships among the genes were shown in Fig. 1.

3. Discussion

Nowadays, viruses are known to be able to change the DNA methylation pattern in host [1], such as HIV [2–4]. In addition, high throughput microarray method has been widely used to investigate the DNA methylation status across the whole genome. We showed the differential DNA methylation pattern in a rare pair of identical twins, while the difference was potentially associated with HIV infection. The further analysis and interpretation of the results were included in Zhang et al. (2015) [7]. Although the dataset was valuable in identifying the host genes which may play important roles in the course of AIDS, it would be noted that the meticulous experimental validation should be performed before any conclusion could be drawn.

Acknowledgments

This project was partly funded by the Scheme B funding on the project “Establishment of the Centre for Microbial Genomics and Proteomics”, the Scheme D funding on the project “Enhancing the Capability of the Hong Kong Bioinformatics Centre” of the Focused Investment Scheme and two grants from Human Medical Research Fund of the Food and Health Bureau, Hong Kong Special Administrative Region (reference nos.: 12110582 and 14130302).

References

[1] R. Ferrari, A. Berk, S. Kurdistani, Viral manipulation of the host epigenome for oncogenic transformation. Nat. Rev. Genet. 10 (2009) 290–294, http://dx.doi.org/10.1038/nrg2539.
[2] B. Youngblood, N.O. Reich, The early expressed HIV-1 genes regulate DNMT1 expression. Epigenetics 3 (2008) 149–156, http://dx.doi.org/10.4161/epi.3.3.6372.
[3] M. Pion, D. Jaramillo-Ruiz, A. Martinez, M.A. Munoz-Fernandez, R. Correa-Rocha, HIV infection of human regulatory T cells downregulates FoXP3 expression by increasing DNMT3b levels and DNA methylation in the FOXP3 gene. AIDS 27 (2013) 2019–2029, http://dx.doi.org/10.1097/QAD.0b013e32836253fd.
[4] J.Y. Fang, J.A. Mikovits, R. Bagni, C.L. Petrow-Sadowski, F.W. Ruscetti, Infection of lymphoid cells by integration-defective human immunodeficiency virus type 1 increases de novo methylation. J. Virol. 75 (2001) 9753–9761, http://dx.doi.org/10.1128/JVI.75.20.9753-9761.2001.

[5] A. Brass, D. Dykxhoorn, Y. Benita, N. Yan, A. Engelman, R. Xavier, J. Lieberman, S. Elledge, Identification of host proteins required for HIV infection through a functional genomic screen. Science 319 (2008) 921–926, http://dx.doi.org/10.1126/science.1152725.

[6] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. Tsafou, M. Kuhn, P. Bork, L. Jensen, C. Mering, STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 43 (2015) D447–D452, http://dx.doi.org/10.1093/nar/gku1003.

[7] Y. Zhang, S.K. Li, K.Y. Yang, M. Liu, N. Lee, X. Tang, H. Wang, L. Liu, Z. Chen, C. Zhang, J. Wang, S.K. Tsui, Whole genome methylation array reveals the down-regulation of IGFBP6 and SATB2 by HIV-1. Sci. Rep. 5 (2015) 10806, http://dx.doi.org/10.1038/srep10806.