Detection and identification of plasma bacterial and viral elements in HIV/AIDS patients in comparison to healthy adults

S.-K. Li[1], R. K-K. Leung[2], H.-X. Guo[3], J.-F. Wei[4], J.-H. Wang[5], K.-T. Kwong[2], S.-S. Lee[6], C. Zhang[7] and S. K.-W. Tsui[1,2]

1) *Centre for Microbial Genomics and Proteomics, Faculty of Medicine,* 2) *The Hong Kong Bioinformatics Centre, The Chinese University of Hong Kong, Shatin, NT, Hong Kong,* 3) *Jiangsu Provincial Center for Disease Prevention and Control,* 4) *The Clinical Experiment Center, First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu,* 5) *Key Laboratory of Molecular Virology and Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences,* 6) *Department of Microbiology, The Chinese University of Hong Kong, Shatin, NT, Hong Kong and* 7) *Institute of Life Sciences, Jiangsu University, Zhenjiang, Jiangsu, China*

Abstract

A low level of CD4+ lymphocyte cells makes end-stage HIV/AIDS patients highly susceptible to microbial infections. We have adopted the next generation sequencing method to identify the spectrum of bacterial plasma and viral elements that might be present in these patients. The HIV/AIDS plasma microbiome was dominated by bacterial elements in the taxonomical order Pseudomonadales, while healthy people carried fewer bacterial DNA in the plasma. We have found that many of the bacterial elements in HIV/AIDS plasma are similar to those of the microbes found in the human gut, suggesting potential acquisition of microbial elements from the gut. The HIV/AIDS and normal plasma DNA virome shared some similarities in the presence of common ubiquitous eukaryotic viruses. The normal DNA virome was mainly composed of viruses from Anelloviridae. In contrast, the HIV/AIDS DNA virome contained a large proportion of bacteriophages, endogenous retroviruses and a non-human virus. In addition, several sequences, which might belong to novel bacteria or endogenous retroviruses, were identified. Taken together, the use of high-throughput sequencing technology in unveiling microbial metagenomics may facilitate future research in combating HIV/AIDS and its associated microbial complications.

Keywords: HIV, microbiome, sequencing, virome
Original Submission: 22 June 2011; **Revised Submission:** 8 September 2011; **Accepted:** 29 September 2011
Editor: D. Raoult
Article published online: 5 October 2011
Clin Microbiol Infect 2012; **18:** 1126–1133
10.1111/j.1469-0691.2011.03690.x

Corresponding author: S. K. W. Tsui PhD, School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China
E-mail: kwtsui@cuhk.edu.hk

Introduction

Depletion of CD4+ T lymphocytes is one of the distinctive features of HIV infection [1]. As a result of CD4+ lymphocytopenia, HIV/AIDS patients without antiretroviral therapy are highly susceptible to infections by different microbes in the natural environment [2]. Microbial infection is in fact a hallmark of AIDS, which causes significant morbidity and mortality [3,4]. The severe complications of such infections can be preceded by asymptomatic bacteraemia or viraemia, the characterization of which would be useful for supporting clinical diagnosis and for the introduction of prophylaxis. Elucidating the difference in the plasma microbiome and virome between HIV/AIDS patients and healthy adults is important because of the newly emerging paradigm that an illness may be defined by a disruption of the normal 'healthy' microbiome or virome. However, the identification of the metagenome of complex microbial elements may not be possible using traditional culture methods [5,6]. Bacteria cultivation usually takes a long time and is not feasible for uncultivable species. Although 16S rRNA gene sequencing is useful in clinical laboratories to identify the agent of infection [7,8], it will be too labour intensive to use this tech-

nique to study complex samples such as environmental samples. Moreover, amplification of the 16S rRNA gene may not always be feasible in studying clinical samples such as plasma [9]. Unlike bacteria, which can be studied using universal primers for the 16S rRNA gene that is present in every bacterial genome, viruses do not carry any conserved genes for a 'one-for-all' detection. The classical method of viral detection is by culturing the virus in susceptible cells and detecting the virus signatures by different means. Molecular detection of the viral DNA or RNA by polymerase chain reaction (PCR) has been widely used in virology laboratories for diagnosing viral infection [10]. Other typical methods used in the recognition of viral antigens are enzyme-linked immunosorbent assay (ELISA), Western blotting and immunofluorescence. When the metagenome of a complex sample is studied, a high throughput detection method should be used. One of the suggested high-throughput methods is detecting the viruses by means of microarray [11]. However, that will involve bioinformatic analysis of conserved regions in a large diversity of viral antigens and is limited to known viruses.

In this study, we have elucidated the metagenome of both the bacterial and viral DNA elements in the plasma of treatment-naïve HIV/AIDS patients, as well as in healthy adults, by means of primer-free next-generation sequencing. Knowing the plasma microbiome information in HIV/AIDS patients will be beneficial for the identification of the common bacteria and viruses associated with these patients, thus improving future HIV disease management.

Patients and Methods

Ethics statement

This study was conducted according to the principles expressed in the Declaration of Helsinki. The study was approved by the Institutional Review Board of Jiangsu University, China, and the Institutional Review Board of the Chinese University of Hong Kong. All the studied subjects provided written informed consent for the collection of samples and subsequent analysis.

Collection of plasma from HIV/AIDS patients and control groups

Ten treatment-naïve HIV/AIDS patients with a CD4+ lymphocyte count of 4–125 cells per microlitre (μL) were recruited in Jiangsu Province, China. These patients (age, 36 ± 7 years; four males, six females) were not under antiretroviral therapy at the time that they were recruited. Medical follow-up was given afterwards. An independent control group of 10 healthy adults (24 ± 3 years; six males, four females) were recruited from Hong Kong. Blood samples were collected from the study subjects by applying standard aseptic techniques and put into tubes containing EDTA.

Extraction of plasma DNA

DNA was extracted from 200 μL plasma using the QIAamp DNA blood mini kit (QIAGEN). The extracted plasma DNA was dissolved in 200 μL H_2O, which had been treated with DNase. The DNase in the water was inactivated before use. The extraction process was performed in a hood, pre-illuminated with ultra-violet light for 30 min. Designated pipettes, filter tips and other necessary equipment were used in this project to avoid contamination of microbial DNA from the environment. Unless otherwise specified, similar precautionary measures were adopted in the other parts of the study, including polymerase chain reactions.

Whole genome amplification of plasma DNA and Illumina Solexa sequencing

The yield of plasma DNA extracted from 200 μL of plasma from HIV/AIDS patients was only about 10 ng/μL, the samples were therefore amplified in order to have 10 μg DNA for Illumina Solexa sequencing. To minimize individual variations of opportunistic infections by heterogeneous sources of pathogens, 5 ng of total plasma DNA was taken from each of the HIV/AIDS patients and mixed together. Similarly, 5 ng plasma DNA was taken from each of the ten healthy individuals and mixed together. The pooled plasma DNA was amplified by the multiple displacement amplification method using GenomiPhi V2 kit (GE Medical Systems). A reaction without DNA template was included as a negative control.

Sequencing data analysis

The amplified and purified plasma DNA samples were sequenced by paired-end Solexa short-read sequencing technology with an insert size of 250 base pairs (bp). The sequence signals were analyzed using Illumina Genome Analyzer. All sequenced reads were assembled using Velvet [12] with a hash length of K21. The sequence contigs assembled using Velvet with K21 were filtered to remove those of less than 100 nucleotides. Gene identity was determined by Basic Local Alignment Search Tool (BLAST) using nucleotide BLAST (BLASTn) and BLASTx searches against the microbial genomes available on the NCBI dataset (http://www.ncbi.nlm.nih.gov/) in January 2010 with an e-value of $<1 \times 10^{-10}$ and a positive hit length of >50% as the cut-off. Computer scripts were written by the Hong Kong Bioinformatics Centre of the Chinese University of Hong Kong, to

automatically align the numerous contigs to the nucleotide/protein databases. A separate set of scripts were written for the taxonomic analysis of the aligned microbial genomes, which was retrieved based on the NCBI accession number.

Results

Metagenomic sequencing of the HIV/AIDS plasma microbiome

Figure 1 shows the workflow of the experiments for analysing the HIV/AIDS plasma microbiome. A total of 658 535 877 bp were generated for the HIV/AIDS plasma microbiome in the Solexa sequencing including 1.43% of human sequences. After filtering contigs with a length shorter than 100 bp, about 16 308 contigs of 100 bp to 9.5 kb were left. The human and non-microbe contigs were removed, leaving 2100 contigs (equivalent to 2379 hits) and 3878 contigs (equivalent to 4277 hits) from BLASTn and BLASTx searches, respectively. The relative abundance of microbial genomes was calculated by taking into account the matched length of contigs and number of reads constituting the region of matched contigs (coverage). The analysed contigs were equivalent to a total of 48 811 322 bp, which contributes to 7.41% of the total sequenced nucleotides. The HIV-associated microbiome was dominated by bacteria from the orders Pseudomonadales (35 297 430 bp, 72.31%), Lactobacillus (2 400 333 bp, 4.92%), Burkholderiales (1 717 225 bp, 3.52%), Bacillales (1 264 882 bp, 2.59%) and Enterobacteriales (1 212 526 bp, 2.48%) (Fig. 2a, Table S1). We noticed that some of the microbes could be found in the normal human gut (Fig. S1).

For the control group of healthy adults, a total of 410 000 000 bp were generated. Contigs assembly was performed on this dataset in the same way as analysing the HIV/AIDS plasma microbiome. A total of 337 857 contigs longer than 100 bp were analysed for their taxonomy and abundance. It included 140 653 contigs of human sequences, which were made up of 141 246 585 bp (34.45% of sequenced nucleotide). A total of 25 contigs with bacterial hits were found in BLASTn and BLASTx searches. These microbial contigs consisted of 34 408 bp, which was equivalent to 0.0084% of the total sequenced nucleotides. Unlike the HIV/AIDS plasma microbiome, the most dominant order of microbes in normal plasma was Clostridiales (14 758 bp, 42.89%), followed by Verrucomicrobiales (8141 bp, 23.66%) and Burkholderiales (5528 bp, 16.07%) (Fig. 2b, Table S2). Intriguingly, there was only one contig for Clostridiales, which contributed to more than 14 kb in weight.

For comparison, HIV/AIDS and normal plasma microbiomes were presented as relative weight in $\log_{10}$ scale against the taxonomy of microbes in the finest rank, given that the preciseness of classification was maintained (Fig. 2c). All the genetic materials of bacteria found in normal healthy adults

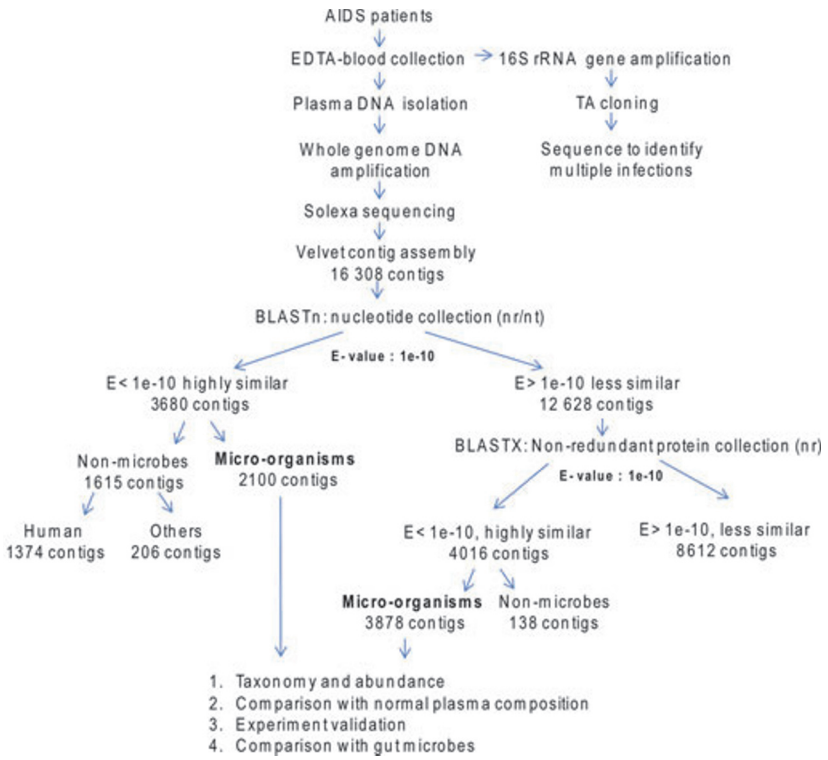


FIG. 1. Workflow of analysis of plasma bacterial elements in HIV/AIDS patients. The plasma DNA extracted from AIDS patients was amplified and sequenced using Illumina Solexa sequencing. The sequence reads were assembled in contigs and aligned against the reference nucleotide collection (nt). The contigs without matches in BLASTn were aligned with the protein sequences collection (nr) using BLASTx. The contigs with matches were further analysed for the identity and relative abundance.

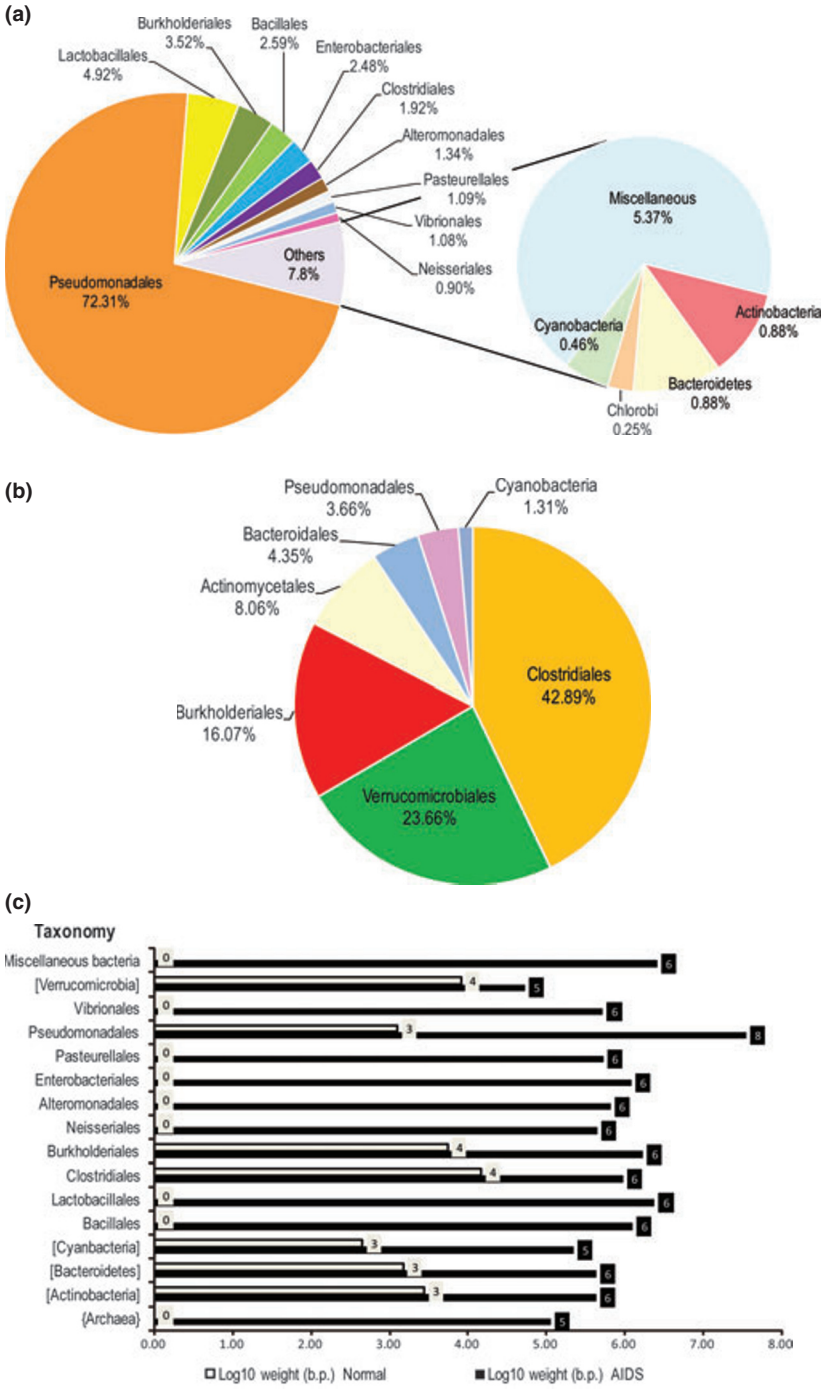


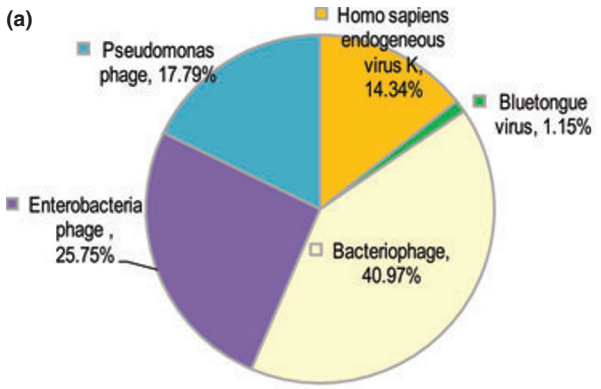
FIG. 2. Plasma microbiomes in HIV/AIDS patients and healthy adults. (a) The relative abundance of bacterial elements found in the HIV/AIDS patients. The left panel shows the most abundant order of bacteria, dominated by Pseudomonadales (72.31%), Lactobacillales (4.92%) and Burkholderiales (3.52%), etc. The right panel shows the less abundant microorganisms per phylum. (b) The relative abundance of microbial genomes in healthy adults. The relative abundance of microbes in a normal control group was denoted in order and was dominated by Clostridiales (42.89%), followed by Verrucomicrobiales (23.66%) and Burkholderiales (16.07%). (c) Relative abundance of microbes in HIV/AIDS patients compared with the normal control group. The relative abundance of different groups of microbes was presented in Log_{10} scale with respect to the finest rank of taxonomy, basically in order rank. The broader level of taxonomy includes phylum, indicated in square bracket [ ] or super-kingdom in curly brace { }.

were also found in HIV/AIDS patients. These include Verrucomicrobia, Pseudomonadales, Burkholderiales, Clostridiales, Cyanobacteria, Bacteroidetes and Actinobacteria. By comparison, the HIV/AIDS plasma microbiome carried some specific bacteria from Vibrionales, Pasteurellales, Enterobacteriales, Alteromonadales, Neisseriales, Lactobacillales and Bacillales. It also carried some Archaea materials and a number of sequences for miscellaneous bacteria. The presence of some of these bacterial elements in individual samples was validated by conventional PCR (Fig. S2).

Comparison of viral materials in plasma of HIV/AIDS patients and normal adults

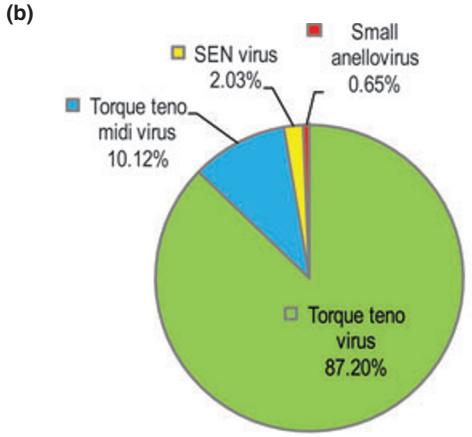
The HIV/AIDS plasma virome was dominated by bacterial phages, which contributed to 84.51% of total plasma viral groups (nine contigs of high coverage) (Fig. 3a). Human

(a)



Virus/phage	Weight (b.p.)	%	No. of contigs
Homo sapiens endogeneous virus K	27 764	14.34%	5
Bluetongue virus	2234	1.15%	1
Bacteriophage	79 318	40.97%	4
Enterobacteria phage	49 860	25.75%	4
Pseudomonas phage	34 441	17.79%	1
Total	**193 617**	**100.00%**	**6**

(b)



Virus/phage	Weight	%	No. of contigs
Torque teno virus	439 394	87.20%	17
Torque teno midi virus	51 003	10.12%	15
SEN virus	10 224	2.03%	1
Small anellovirus	3270	0.65%	2
Phage	0	0.00%	0
Total	**503 892**	**100.00%**	**35**

FIG. 3. Plasma viromes in HIV/AIDS patients and healthy adults. (a) The major population of the HIV/AIDS plasma virome was bacterial phages (84.51%), dominated by bacteriophage (40.97%) and enterobacteria phage (25.75%). Human endogenous retrovirus K contributed to 14.34% of the plasma virome. A non-human Bluetongue virus was also found (1.15%). (b) Unlike the HIV/AIDS plasma virome, the normal virome did not contain bacterial phages. Torque teno virus (87.20%) and Torque teno midi virus (10.12%) constituted the largest viral communities in this virome. SEN virus (2.03%) and small anellovirus (0.65%) were also found in the plasma virome.

endogenous retroviruses K contributed to 14.34% of the virome. The most abundant phage group was Bacteriophage (40.97%), followed by Enterobacteria phage (25.75%) and Pseudomonas phage (17.79%). Surprisingly, the non-human virus bluetongue virus was also found in HIV/AIDS patients. This virus, which infects ruminants, contributed to 1.15% of the virome [13,14]. In contrast, the normal plasma virome did not contain any bacterial phages. It consisted of a few species of Anelloviruses. The dominant group of viruses was Torque teno virus (TTV) (87.20%), followed by Torque teno midi virus (TTMDV) (10.12%). SEN virus and small anellovirus constituted 2.03% and 0.65% of the normal virome, respectively. No human endogenous retrovirus K DNA was sequenced (Fig. 3b).

Identification of novel bacteria and viruses in the sequencing data sets

We have analysed the longest contig in the HIV/AIDS microbiome by mapping it to the quickly expanding bacteria reference genomes. Some previously unmatched contigs were later found to belong to *Aerococcus viridans* (*A. viridans*) when its genome was first available in November 2010. We identified 24 genes of *A. viridans* in 13 contigs found in the HIV/AIDS plasma microbiome (Fig. 4). The 13 selected contigs in the HIV/AIDS plasma microbiome contributed to 58 343 bp in total, equivalent to 2.91% of the whole genome of *A. viridans*.

We noticed that many of the contigs representing the human endogenous retrovirus (HERV) K actually had less than 60% similarity to the current HERV genome on the NCBI nt/nr protein reference database. As shown in Table 1, the nucleotide BLAST results of all the five contigs of HERV showed a nearly 100% identity to human chromosome. However, when the contigs were searched by BLAST analysis against the reference nucleotide translated database using BLASTx, only human endogenous retrovirus K was found. Phylogenetic analysis shows that these HERV sequences were distinctively different from all the existing HERV-K genomes Fig. S3.

Discussion

We have used the unbiased high-throughput sequencing technology in profiling the plasma microbiomes and viromes in two groups of Chinese subjects. The reliability of the contig assembly was validated by the BLAST searching and PCR results (Supporting information). The results of validation experiments suggested, for the first time, the association of opportunistic bacteria *Moraxella osloensis* and

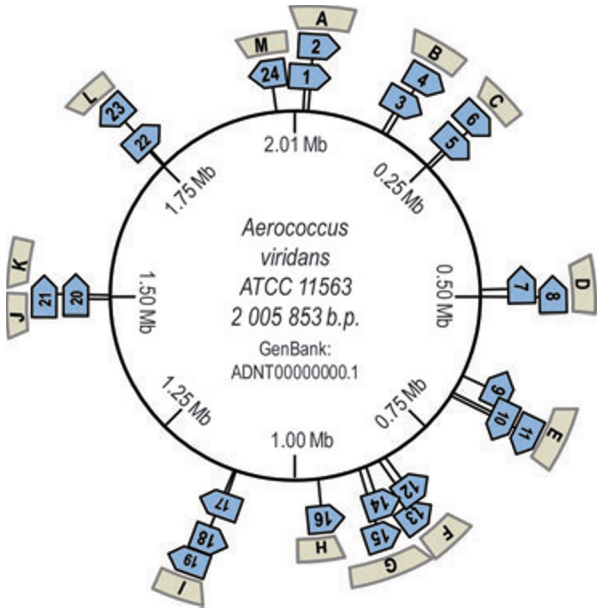


FIG. 4. Identification of potential novel microbial sequences in the HIV/AIDS microbiome. This figure shows the genetic map of *Aerococcus viridans ATCC 11563* and the mapped contigs. The localization of the contigs found in the HIV/AIDS plasma microbiome was shown in grey boxes. The blue arrows indicate the *A. viridans* genes aligned with the contigs with the orientation of transcription.

Psychrobacters with HIV infection. We have also detected a few microbial genetic elements in the plasma of healthy adults, suggesting the presence of bacterial genomes in apparently healthy people. The bacterial elements detected may belong to the bacteria that contributed to the major opportunistic infections in HIV/AIDS patients. Thus, the differences in the plasma microbiomes of normal adults and HIV/AIDS will be important in understanding the risk of opportunistic infections in immuno-compromised people. Alternatively, these bacterial sequences may be genetic materials of bacteria that await removal by the immune system, due to the fact that the presence of exogenous nucleic acids in our circulation, such as human DNA [15], viral DNA [6] or fetal DNA in pregnant women [16], does not provoke immune attack.

Here we also provide evidence that genetic materials of gut microbes are present in patient plasma (Supporting information) [17]. Our gastrointestinal (GI) tract is the most complex part of our immune system as the GI tract needs to differentiate between food proteins, drugs and oral DNA vaccines from the infectious agents [18]. In fact, the gut mucosal layer is more abundant in activated CD4+ cells than in the circulation system. Therefore, the majority of the depleted CD4+ T-cells in HIV infection reside in the gut [19,20]. It has been recently suggested that the microbial ele-

TABLE 1. Details of BLAST results for all the contigs representing the human endogenous retrovirus K in the HIV/AIDS plasma virome

				Nucleotide BLAST				BLASTX			
Contig	Reference no.	Length	Coverage	Matched target	Identity	*E*-value	Target details	Matched target	Identity	*E*-value	Target details
1	NODE_432919	392	2.90×	gi|21728162|dbj|AP005209.3|	99.74	0	Homo sapiens genomic DNA, chromosome 18 clone: RP11-138E9, complete sequence	sp|P63136.1|POK12_HUMAN	56.67	4.00E-16	Human endogenous retrovirus K 11q22 provirus ancestral pol protein
2	NODE_491839	843	4.03×	gi|23396223|gb|AC130456.2|	99.27	0	Homo sapiens chromosome 16 clone CTA-363E6, complete sequence	sp|P63136.1|POK17_HUMAN	65.83	4.00E-71	Human endogenous retrovirus K 11q22.1 provirus pol protein
3	NODE_253046	261	7.57×	gi|23396223|gb|AC130456.2|	99.20	5.00E-121	Homo sapiens chromosome 16 clone CTA-363E6, complete sequence	sp|Q9WJR5.2|POK_HUMAN	66.67	6.00E-24	Human endogenous retrovirus K 19q12 provirus ancestral pol protein
4	NODE_501411	212	4.65×	gi|23396223|gb|AC130456.2|	100.00	9.00E-99	Homo sapiens chromosome 16 clone CTA-363E6, complete sequence	gb|AAD51797.1|AF164614_1	66.67	1.00E-15	Human endogenous retrovirus K 6 gal-pro-pol protein
5	NODE_478229	149	4.03×	gi|29498303|emb|AL391099.12|	100.00	2.00E-62	Human DNA sequence from clone RP11-178A10 on chromosome 10 Contains a novel gene (FLJ34785 FLJ33470)	gb|ABD29044.1	70.45	2.00E-12	Human endogenous retrovirus K envelope glycoprotein

ments, leaked from the damaged gut, are the drivers for progressive CD4+ T-cell depletions [21] and for most of the severe sepsis (with unknown causes) in patients with critical illness [22]. These studies also suggested that the translocation of bacterial DNA could be attributed to the depletion of CD4+ T cells in the mucosal layer of the gastrointestinal tract in chronic HIV infection [23,24].

We have also examined the plasma viromes in the two groups of study subjects. The presence of bacteriophages in the HIV/AIDS virome may, in another way, reflect the presence of bacteria in patients' plasma that are circulating with the systemic blood flow. Nevertheless, it is also possible that the phages enter the blood stream by passing through the digestive system, as exemplified by the plant virus Pepper Mild Mottle Virus (PMMV) in food, which was proven to enter the circulation via the gastrointestinal system [25]. We have also found the presence of ruminant virus bluetongue virus in the HIV/AIDS DNA virome, suggesting that patients with HIV/AIDS may be susceptible to infection by non-human viruses. On the contrary, the normal virome mainly contains Anelloviruses but no bacteriophage was found, suggesting that the bacterial elements found in the normal microbiome may be the remains of non-living bacteria. The non-predominance of Anellovirus in the HIV/AIDS virome can be explained by the presence of a large microbiome, acting as preferential MDA targets. Another limitation of our virome study is that the adopted protocol was restricted to the investigation of DNA viral sequences. Thus, viruses of RNA origin such as HIV cannot be revealed.

In aiming to reveal potential novel DNA in this study, the free human nucleic acids in the plasma were not pretreated by nucleases. Therefore, after removing the sequence of human DNA by computation analysis, we were able to find five contigs that might be the coding genes of human endogenous retroviruses (HERV). Actually, HERV sequences comprise about 8% of the human genome [26,27]. These viruses are thought to have infected humans and integrated into the human genome many millions of years ago. With the success in reconstituting the HERV-K genome, it was implicated that the 'dormant' retroviral gene fragments may still possess the ability to encode proteins such as the retrovirus polymerase and envelope protein [28,29]. More strikingly, it has been shown that the HERV-K fragment in the human genome can be 'reactivated' by other exogenous retroviruses such as HIV [27].

In this study, the next-generation sequencing technology was used in determining the bacterial and viral metagenomes. All subjects were of Chinese origin living in southern China with similar weather conditions and culture. While there is no information on the influence of environmental factors on the microbiome, this type of study will always benefit from recruiting subjects from the same living environment. Moreover, additional cohorts from different living environments should be investigated in the future in order to make more robust conclusions. Besides, there are always potentially contaminated sequences in the datasets. Unfortunately, we cannot 'kill' DNA and there is no universal standard of adequate precaution. In this manner, metagenomics are always presented in relative proportion of different species and only the dominating population is analysed critically. In summary, this study on the plasma microbiome and virome will help us to better understand the spectrum of circulating microbial elements in our blood, though their mode of 'existence' awaits further studies.

Acknowledgements

We thank Dr Ting-Fung Chan and Mr Chi-Shing Yu for technical help with the taxonomy analysis. We also thank Dr Terrence Chi-Kong Lau for technical advice on the PCR amplification of microbes.

Transparency Declaration

This project is supported by the Scheme B funding of the project 'Establishment of the Centre for Microbial Genomics and Proteomics' of the Focused Investment Scheme of The Chinese University of Hong Kong. All the authors declare that they have no conflicts of interest.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Abundance of gut microbes in HIV/AIDS patient plasma.

Figure S2. Amplification of bacterial genes from HIV/AIDS patients.

Figure S3. Phylogenic relationship of HERV contig 2 (NODE_491839) and its related HERV polymerase reference sequences.

Table S1. Summary of taxonomy information for the microbes found in the HIV/AIDS patients group (excluding viruses).

Table S2. Summary of taxonomy information for bacterial materials found in the healthy adults (excluding viruses).

References

1. Rudnicka D, Schwartz O. Intrusive HIV-1-infected cells. *Nat Immunol* 2009; 10: 933–934.
2. Mootsikapun P. Bacteremia in adult patients with acquired immunodeficiency syndrome in the northeast of Thailand. *Int J Infect Dis* 2007; 11: 226–231.
3. Antiretroviral Therapy Cohort Collaboration A. Causes of death in HIV-1-infected patients treated with antiretroviral therapy, 1996–2006: collaborative analysis of 13 HIV cohort studies. *Clin Infect Dis* 2010; 50: 1387–1396.
4. Palella FJ Jr, Baker RK, Moorman AC *et al.* Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study. *J Acquir Immune Defic Syndr* 2006; 43: 27–34.
5. Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 2004; 68: 669–685.
6. Leland DS, Ginocchio CC. Role of cell culture for virus detection in the age of technology. *Clin Microbiol Rev* 2007; 20: 49–78.
7. Petti CA, Polage CR, Schreckenberger P. The role of 16s rRNA gene sequencing in identification of microorganisms misidentified by conventional methods. *J Clin Microbiol* 2005; 43: 6123–6125.
8. Woo PC, Lau SK, Teng JL, Tse H, Yuen KY. Then and now: use of 16s rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clin Microbiol Infect* 2008; 14: 908–934.
9. Ferri E, Novati S, Casiraghi M *et al.* Plasma levels of bacterial DNA in HIV infection: the limits of quantitative polymerase chain reaction. *J Infect Dis* 2010; 202: 176–177; author reply 178.
10. Elnifro EM, Ashshi AM, Cooper RJ, Klapper PE. Multiplex PCR: optimization and application in diagnostic virology. *Clin Microbiol Rev* 2000; 13: 559–570.
11. Chou CC, Lee TT, Chen CH *et al.* Design of microarray probes for virus identification and detection of emerging viruses at the genus level. *BMC Bioinformatics* 2006; 7: 232.
12. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de bruijn graphs. *Genome Res* 2008; 18: 821–829.
13. Zhang X, Boyce M, Bhattacharya B, Schein S, Roy P, Zhou ZH. Bluetongue virus coat protein vp2 contains sialic acid-binding domains, and vp5 resembles enveloped virus fusion proteins. *Proc Natl Acad Sci U S A* 2010; 107: 6292–6297.
14. Maclachlan NJ, Guthrie AJ. Re-emergence of bluetongue, African horse sickness, and other orbivirus diseases. *Vet Res* 2010; 41: 35.
15. Stroun M, Anker P, Lyautey J, Lederrey C, Maurice PA. Isolation and characterization of DNA from the plasma of cancer patients. *Eur J Cancer Clin Oncol* 1987; 23: 707–712.
16. Chiu RW, Cantor CR, Lo YM. Non-invasive prenatal diagnosis by single molecule counting technologies. *Trends Genet* 2009; 25: 324–331.
17. Qin J, Li R, Raes J *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 2010; 464: 59–65.
18. Mowat AM. Anatomical basis of tolerance and immunity to intestinal antigens. *Nat Rev Immunol* 2003; 3: 331–341.
19. Brenchley JM, Schacker TW, Ruff LE *et al.* Cd4+ t cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J Exp Med* 2004; 200: 749–759.
20. Veazey RS, DeMaria M, Chalifoux LV *et al.* Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. *Science* 1998; 280: 427–431.
21. Lederman M. Why suppression of HIV replication does not always make everything better. *J Int AIDS Soc* 2010; 13 (suppl 4): K1.
22. Alverdy JC, Chang EB. The re-emerging role of the intestinal microflora in critical illness and inflammation: why the gut hypothesis of sepsis syndrome will not go away. *J Leukoc Biol* 2008; 83: 461–466.
23. Brenchley JM, Price DA, Douek DC. HIV disease: fallout from a mucosal catastrophe? *Nat Immunol* 2006; 7: 235–239.
24. Jiang W, Lederman MM, Hunt P *et al.* Plasma levels of bacterial DNA correlate with immune activation and the magnitude of immune restoration in persons with antiretroviral-treated HIV infection. *J Infect Dis* 2009; 199: 1177–1185.
25. Zhang T, Breitbart M, Lee WH *et al.* RNA viral community in human faeces: prevalence of plant pathogenic viruses. *PLoS Biol* 2006; 4: e3.
26. Belshaw R, Pereira V, Katzourakis A *et al.* Long-term reinfection of the human genome by endogenous retroviruses. *Proc Natl Acad Sci U S A* 2004; 101: 4894–4899.
27. Garrison KE, Jones RB, Meiklejohn DA *et al.* T cell responses to human endogenous retroviruses in HIV-1 infection. *PLoS Pathog* 2007; 3: e165.
28. Dewannieux M, Harper F, Richaud A *et al.* Identification of an infectious progenitor for the multiple-copy herv-k human endogenous retroelements. *Genome Res* 2006; 16: 1548–1556.
29. Lee YN, Bieniasz PD. Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog* 2007; 3: e10.