Metagenomic analysis of the virome of pooled plasma from healthy blood donors

Zhengang Shan (gzbc2016@126.com)
Guangzhou Blood Center https://orcid.org/0000-0001-5534-6947
Jieting Huang
Guangzhou Blood Center
Min Wang
Guangzhou Blood Center
Ru Xu
Guangzhou Blood Center
Qiao Liao
Guangzhou Blood Center
Huishan Zhong
Guangzhou Blood Center
Bochao Liu
Guangzhou Blood Center
Hualong Yang
Southern Medical University
Yongshui Fu
Guangzhou Blood Center
Xia Rong
Guangzhou Blood center https://orcid.org/0000-0002-2463-6340

Research Article

Keywords:

Posted Date: August 23rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3268789/v1

Abstract

Qualified blood from healthy donors may still contain infectious pathogens that could become important threats to human life. The objective of this study was to screen for blood-borne viruses and analyze the virome profiles in plasma via metagenomic sequencing. A total of 1,200 plasma samples were collected and mixed into 12 pools based on sampling time, then DNA and cDNA libraries were constructed for sequencing which was carried out by Illumina NovaSeq 6000 system. After data cleaning and filtering using bioinformatic tools, 2336 viral reads via DNA sequencing and 61985 viral reads via cDNA sequencing were obtained. Seven DNA viruses belonging to three families and one RNA virus were identified from these reads. Most of the viruses found in pooled plasma were considered non- pathogenic, especially anellovirus found in all 12 pools and human pegivirus detected in 7 of the 12 pools. Three pathogenic viruses were found by DNA sequencing, including human herpes virus 6A, human cytomegalovirus and Epstein-Barr virus. All of them belong to the family *Herpesviridae* and their reads accounted for 7.9% (184/2336). Currently, data on anellovirus and human pegivirus in China are scarce. The herpes virus, which has a high sero-prevalence, is not a mandatory marker for blood screening. This study preliminary elucidated the virome spectra in blood from qualified donors in Guangzhou and contributed to information regarding blood transfusion safety in China.

Introduction

To improve the safety of blood supply, donors are routinely screened for a range of blood-borne pathogens with highly sensitive screening tests[1]. In China, four pathogens were statutory tested: hepatitis b virus, hepatitis c virus, human immunodeficiency virus and treponema pallidum. However, these are not the only pathogens that can be transmitted through blood transfusion. As a matter of fact, qualified blood from healthy donors may still contain infectious pathogens that could become important threats to human life[2, 3]. Although these pathogens may affect the treatment of the recipients and even cause serious diseases, it is unpractical to systematically screen for all known transfusion-transmitted agents in donated blood. Blood safety measures such as nucleic acid testing can be only applied to a limited variety of pathogens[4]. Thus, it remains imperative to improve the blood safety, not only by routine methods to surveil well-characterized pathogens, but also by new methodologies to discover unknown or unexpected pathogens[1].

Take advantage of next-generation sequencing (NGS) technology, it is possible to sequence all nucleic acids present in a given sample and explore the microbial landscapes in a wide variety of samples, which is termed “metagenomics” [5]. The application and improvements of metagenomics related to continuous advances in NGS and bioinformatic approaches have shown that the viral abundance of clinical samples is much greater than previously suspected[6]. Hence, we conducted a metagenomic analysis of virome profiles in pooled plasma from qualified donors in Guangzhou to screen for blood-borne viruses that were otherwise missed using standard protocols and simultaneously surveil emerging/re-emerging infectious diseases.

Materials and methods

Sample collection and processing

A total of 1,200 plasma samples were collected from qualified blood donors in Guangzhou. These plasma were mixed into 12 pools based on sampling time (February 2021 to July 2022). Each pool contained 100 samples and 40ml (400μl per sample) of plasma, of which 10ml was filtered through PVDF membrane (Millex-GV, 0.22μm filter unit) to remove blood cells and cellular microbes. The filtered samples were stored in a sterile container and preserved in -80℃ refrigerator before further processing.

Library construction and sequencing

1.4ml filtered plasma from each pool was divided into two aliquots for DNA and RNA extraction using QIAsymphony Circulating NA Kit. TURBO DNase was added into RNA extracts of 5 pools (pool 8–12) before incubating for 15 minutes at 37℃ to eliminate residual cell-free DNA (Table S1). Next, the DNA/cDNA libraries (PE150) were prepared using TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme, TD503) and The Nucleic Acid Detection Kit for Pathogens (Kingcreat, KS619-DNAmN24). Sequencing library from pool 1 to 7 was constructed using transposase. In contrast, sequencing library from pool 8 to 12 was constructed using endonuclease. The concentration of library was measured by Qubit. Then the sequencing was carried out by Illumina NovaSeq 6000 system with 300 cycles Reagent Kit.

Data analysis and interpretation

High-quality sequencing data were obtained after filtering out low-quality reads by the use of fastq (v0.22) software. Human nucleic acid sequences were then removed by mapping reads to human reference genome using bwa (v0.7.10) software. The remaining data were aligned to the microbial genome database using bwa (v0.7.10) software. The whole database contained 36,497 microorganisms and 9,808 of them had integral sequence of whole genome and detailed clinical analysis, which covered most of the known pathogenic viruses. The number of reads, relative abundance and other indicators such as RPM-r (RPM defined as Reads per million, RPM-r defined as RPM sample/RPM NTC) were used to classify and filter the viruses in clean data. Secondary filtration of the pathogenic reads were obtained by self-developed machine learning algorithm, with the most plausible reads retained and the less plausible ones removed.

Experimental procedures are represented in the schematic workflow in Figure S1.

Results

Next-generation sequencing results

The adaptor sequences, duplicate and low-quality reads were removed from the raw data and more than 6 billion clean reads were generated from all these sample pools, with the lowest Q30 value of 86.17% (Table S2 and Table S3). In total, more than 700GB of clean data were obtained, with each plasma pool producing an average of 60 GB, consisting of approximately half DNA sequencing data and half cDNA sequencing data. After further removal of sequences mapped to human, rRNA and carrier protein, 1,398,378,808 filtered reads were aligned to the microbial genome database. The results are shown in Table 1 and Fig. 1.

Table 1
Statistical results of metagenomic sequencing

	raw bases	clean bases	clean reads	filtered reads
DNA sequencing	471,360,956,542	380,417,044,928	3,473,711,826	147,899,895
cDNA sequencing	466,978,164,490	358,171,050,246	3,351,172,578	1,250,478,913
Total	938,339,121,032	738,588,095,174	6,824,884,404	1,398,378,808

Filtered reads: sequencing data after removing adaptor sequences, duplicate sequences, low-quality sequences and sequences mapped to human, rRNA and carrier protein.

Virus identification in pooled plasma

After data cleaning and filtering, a total of 64,321 viral reads were obtained (Table 2). Seven DNA viruses belonging to three families and one RNA virus were identified from these reads. Among the reads from DNA sequencing (Fig. 2A), nucleotide sequences of the family *Anelloviridae* (Torque teno virus, TTV; Torque teno midi virus, TTMDV; Torque teno mini virus, TTMV) accounted for 91.8% (2145/2336) and sequences of the family *Herpesviridae* (human herpes virus 6A, HHV-6A; human cytomegalovirus, HCMV; Epstein-Barr virus, EBV) accounted for 7.9% (184/2336). The remaining 0.3% (7/2336) reads were identified as TTV-like mini virus (TLMV). Among the reads from cDNA sequencing (Fig. 2B), GB virus C/human pegivirus (GBV-C/HPgV) accounted for 96.2% (59602/61985), and anellovirus accounted for 3.8% (2383/61985).

Table 2
The numbers of viral sequences in plasma pools

	virus reads via DNA sequencing	virus reads via cDNA sequencing
DNA virus reads	2336	2383
RNA virus reads	0	59602
Total	2336	61985

The viruses found by NGS and their reads in each plasma pool were detailed in Table 3. Most of the viruses found in this study were generally considered to be non-pathogenic, and no highly pathogenic virus was identified. Upon DNA sequencing, anelloviruses which were highly diverse but non-pathogenic were found in all 12 pools. TLMV were identified in pool 11 and 12, but as with anellovirus, the pathogenicity of this virus remained to be elucidated. Only three pathogenic viruses were detected: HHV-6A and CMV in pool 1, EBV in pool 4. The cDNA sequencing led to the discovery of HPgV, an RNA virus belonging to the family *Flaviviridae*, in 7 out of 12 pools, with unclear pathogenicity. Significantly, anelloviruses were detected by cDNA sequencing in 10 plasma pools, even though they were DNA viruses.

Table 3
Virus detected in each pool by metagenomic sequencing

Pool	DNA sequencing (reads)							cDNA sequencing (reads)			
	TTV	TTMDV	TTMV	HHV-6A	CMV	EBV	TLMV	TTV	TTMDV	TTMV	GBV-C
1	13			180	2			2	20		1092
2	13								59		1223
3		2						90	87	255	
4	8					2		18	476	76	
5	8	1						336	4		271
6	8							261	216	20	
7	4							89	287	28	691
8	245	223	30					5	20	4	
9	169	83	5					2	2		7493
10	115	196	4								22777
11	234	359	76				3				
12	155	151	43				4	24	2		26055

Virus spectra of 12 plasma pools

Although virus abundance in each plasma pool was not high, the virus profiles of these pools varied considerably, as showed in the relative abundance map (Fig. 3). This variation was mainly related to the abundance and diversity of anelloviruses, which were detected in all 12 plasma pools. Pools 8 to 12 had a higher anellovirus abundance than pools 1 to 7 in DNA sequencing. But in cDNA sequencing, relatively more anellovirus reads were identified from the latter (Fig. 4).

Discussion

Blood screening, especially nucleic acid testing, has greatly improved the safety of the blood products, however, these measures can only cover a limited range of pathogens. In recent years, metagenomics technology with the advantages including high efficiency and broad pathogen coverage has been used to analyze various biological samples, and has also proven remarkably useful in studies of infectious diseases[3]. For instance, Cache Valley virus, a mosquito-borne virus that is a rare cause of disease in humans, has recently been discovered to be transmitted through blood transfusions and cause meningoencephalitis in recipients by metagenomic next-generation sequencing after ruling out more common etiologies [7]. Since screening for all viruses in blood respectively is obviously impossible, metagenomic sequencing is regarded as a promising method to understand the microbiome in blood and thereby improve the safety of blood supply. In this study, we used Illumina NovaSeq 6000 sequencing system

for metagenomic analysis to reveal the virus composition of pooled plasma from 1200 qualified blood donors in Guangzhou.

This study has revealed the common presence of apparently non-pathogenic viruses in blood, particularly anellovirus and human pegivirus. Anellovirus, probably the most abundant viruses of the human blood virome, was observed among healthy people in many countries and regions[4, 6, 8–12]. The prevalence of anellovirus is highly variable, the infection rate in healthy populations in Japan is close to 100%, while in Britain and America are approximately 10%[11–13]. Currently, data on anellovirus in China are scarce[3], but we speculated that the prevalence may be high because all plasma pools in this study contained the nucleotide sequence of the virus. The high prevalence of anellovirus is a consequence of the multiple transmission routes, including parenteral, sexual, and vertical routes[8]. Because of the extensive polytropism, anellovirus have been observed in many other body compartments, other than in human blood[14]. In addition, the discovery of anellovirus by cDNA sequencing in this study indicated that these viruses might be constantly active and infective. Since anellovirus bloom in diversity or quantitative abundance was documented in conditions related to immune suppression[15–17], it has been suggested that pathogenesis may be conditional, acting as an aggravating factor or as an opportunistic agent[18]. However, whether a high viral load of anelloviruses in blood can cause clinical symptoms even disease in humans remains ambiguous[3, 19, 20] due to the available but contradictory data[14]. More detailed studies are needed to evaluate if anellovirus might exert pathogenic effect. At present, anelloviruses are still considered part of the natural human virome because of their high prevalence and largely asymptomatic persistence[21].

HPgV, another frequently described viral component of the healthy blood donor virome, is also considered a commensal virus without pathogenicity and clinical importance[6, 22]. Similar to the anellovirus, serological prevalence of HPgV has shown large variability in the general population[22]. Epidemiologic studies suggest that 1–4% of healthy donors in most developed countries are viraemic at the time of blood donation, while in developing countries, blood donor viraemia prevalence is higher, approaching 20% in some regions of the world[22]. In this study, more than half pools contained HPgV. The number of HPgV reads in two pools was particularly high, which might be caused by individual samples. HPgV infection may persist for decades, but most healthy individuals clear viremia within 2 years of infection[23]. By lack of clarification on associations between HPgV and disease, it is not a target agent for blood screening so far. Moreover, several studies have suggested that HPgV may impact HIV disease and improve survival in positive individuals[24–28], but others have not found such potential association[8, 29, 30]. There needs to be more evidence to elucidate the potential benefits of HPgV chronic infection.

Only three definitely pathogenic viruses had been found, CMV, EBV and HHV-6A. They all belong to the family *Herpesviridae*. The blood collection and supply system in China does not perform routine screening for herpes virus. However, whether herpes virus detection should be performed for certain blood recipients is worth consideration. For instance, CMV caused mostly mild febrile symptoms in immunocompetent individuals[31]. But populations undergoing frequent transfusions with low and defective immune function are particularly susceptible to acquiring CMV infection from blood transfusion and can suffer severe consequences[3, 32]. In some countries, such as Brazil, CMV serological test is requested when blood

products are used by specific vulnerable patient populations[33]. In the U.S., CMV antibody testing is performed on a portion of blood donations, including all platelet donations, to meet the demand of hospital for CMV seronegative blood products. According to an epidemiological study in China, the sero-prevalence of CMV, EBV and HHV-6 are 97.3%, 99.8% and 62.0%, respectively[34]. Not many herpes virus sequences were detected in this study, probably because the herpes virus does not replicate actively in healthy individuals. In this study, we did not perform a comprehensive research and analysis that can be accurately reflective of the prevalence of herpes virus among healthy/qualified blood donors in Guangzhou. But such a high sero-prevalence of herpes virus in China should be an area of focus for public health and transfusion medicine, and we would encourage blood centers to initiate epidemiological surveys targeting herpes viruses to understand the prevalence and potential health risks of these viruses and adjust screening strategies according to local epidemiology.

Another interesting finding in our study was the effect of different enzymes used to construct libraries on sequencing results. The number of anellovirus reads in DNA sequencing libraries constructed with transposase was lower than that of libraries constructed with endonuclease. One possible explanation is that the genome of anellovirus is single-stranded DNA, while transposase is more likely to recognize double-stranded DNA. DNase was not applied in pool 1–7 before construction of cDNA sequencing library. The single-stranded DNA remaining in the RNA extracts from these sample pools might be synthesized into cDNA by reverse transcriptase, leading to more anellovirus reads in the cDNA sequencing. These findings strongly suggested that metagenomic sequencing results must be interpreted with caution, owing to non-standardized experimental procedures.

For economy of cost, we adopted the strategy of mixing every 100 samples into one pool, which might lead to some low-load viruses being missed. In addition, determining whether the virus is infectious is also difficult, beacaus only genetic sequences can be obtained[35]. Due to these defects, metagenomic analysis for monitoring the additional risk of transfusion-borne pathogens is not yet in general use at present. Nonetheless, as a superior technology for the discovery of unknown or unexpected pathogens compared to the traditional methods, with further improvements, metagenomic analysis combined with next-generation sequencing would become a useful tool in blood transfusion safety and epidemiological surveillance.

Declarations

Consent to participate

All enrolled blood donors in this study had given written consent to participate in screening for transfusion-borne pathogens before donation.

Ethical approval

All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki and its later amendment. This study was approved by the Ethics Committee of Guangzhou Blood Center.

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Zhengang Shan. The first draft of the manuscript was written by Zhengang Shan and Jieting Huang. All authors commented on previous versions of the manuscript and approved the final manuscript.

Competing Interests

The authors declare they have no relevant financial or non-financial interests to disclose.

Funding

This study was funded by Science and Technology Projects in Guangzhou (202002030192, 2023A03J1000), the Key Medical Disciplines and Specialties Program of Guangzhou (2021-2023) and Medical Scientific Research Foundation of Guangdong Province (B2022150).

References

1. Sauvage V, Eloit M (2016) Viral metagenomics and blood safety. Transfus Clin Biol 23:28-38. https://doi.org/10.1016/j.tracli.2015.12.002
2. Stramer SL (2014) Current perspectives in transfusion-transmitted infectious diseases: emerging and re-emerging infections. ISBT Sci Ser 9:30-36. https://doi.org/10.1111/voxs.12070
3. Xu M, Gao J, Li S et al (2020) Metagenomic analysis and identification of emerging pathogens in blood from healthy donors. Sci Rep 10:15809. https://doi.org/10.1038/s41598-020-72808-8
4. Waldvogel-Abramowski S, Taleb S, Alessandrini M et al (2019) Viral Metagenomics of Blood Donors and Blood-Derived Products Using Next-Generation Sequencing. Transfusion medicine and hemotherapy : offizielles Organ der Deutschen Gesellschaft fur Transfusionsmedizin und Immunhamatologie 46:87-93. https://doi.org/10.1159/000499088
5. Sauvage V, Gomez J, Boizeau L et al (2017) The potential of viral metagenomics in blood transfusion safety. Transfus Clin Biol 24:218-222. https://doi.org/10.1016/j.tracli.2017.06.018
6. Slavov SN (2022) Viral Metagenomics for Identification of Emerging Viruses in Transfusion Medicine. Viruses 14https://doi.org/10.3390/v14112448
7. Al-Heeti O, Wu EL, Ison MG et al (2023) Transfusion-Transmitted Cache Valley Virus Infection in a Kidney Transplant Recipient With Meningoencephalitis. Clin Infect Dis 76:e1320-e1327. https://doi.org/10.1093/cid/ciac566
8. Cebria-Mendoza M, Bracho MA, Arbona C et al (2021) Exploring the Diversity of the Human Blood Virome. Viruses 13:2322. https://doi.org/10.3390/v13112322
9. Ye H, Wang K, Wang M et al (2019) Bioinspired nanoplatelets for chemo-photothermal therapy of breast cancer metastasis inhibition. Biomaterials 206:1-12. https://doi.org/10.1016/j.biomaterials.2019.03.024
10. Cebria-Mendoza M, Arbona C, Larrea L et al (2021) Deep viral blood metagenomics reveals extensive anellovirus diversity in healthy humans. Sci Rep 11:6921. https://doi.org/10.1038/s41598-021-86427-4

11. Biagini P, de Micco P, de Lamballerie X (2006) Identification of a third member of the Anellovirus genus ("small anellovirus") in French blood donors. Archives of virology 151:405-408. https://doi.org/10.1007/s00705-005-0660-4
12. Ninomiya M, Takahashi M, Nishizawa T et al (2008) Development of PCR assays with nested primers specific for differential detection of three human anelloviruses and early acquisition of dual or triple infection during infancy. J Clin Microbiol 46:507-514. https://doi.org/10.1128/JCM.01703-07
13. Spandole S, Cimponeriu D, Berca LM et al (2015) Human anelloviruses: an update of molecular, epidemiological and clinical aspects. Archives of virology 160:893-908. https://doi.org/10.1007/s00705-015-2363-9
14. Reshetnyak VI, Maev IV, Burmistrov AI et al (2020) Torque teno virus in liver diseases: On the way towards unity of view. World J Gastroenterol 26:1691-1707. https://doi.org/10.3748/wjg.v26.i15.1691
15. Thijssen M, Tacke F, Beller L et al (2020) Clinical relevance of plasma virome dynamics in liver transplant recipients. EBioMedicine 60:103009. https://doi.org/10.1016/j.ebiom.2020.103009
16. van Rijn AL, Wunderink HF, Sidorov IA et al (2021) Torque teno virus loads after kidney transplantation predict allograft rejection but not viral infection. J Clin Virol 140:104871. https://doi.org/10.1016/j.jcv.2021.104871
17. Gimenez E, Monzo C, Albert E et al (2021) Diversity and dynamic changes of anelloviruses in plasma following allogeneic hematopoietic stem cell transplantation. J Med Virol 93:5167-5172. https://doi.org/10.1002/jmv.26661
18. Spandole-Dinu S, Cimponeriu DG, Craciun AM et al (2018) Prevalence of human anelloviruses in Romanian healthy subjects and patients with common pathologies. BMC infectious diseases 18:334. https://doi.org/10.1186/s12879-018-3248-9
19. Piaggio F, Dodi F, Bottino G et al (2009) Torque Teno Virus–cause of viral liver disease following liver transplantation: a case report. Transplant Proc 41:1378-1379. https://doi.org/10.1016/j.transproceed.2009.03.047
20. Maggi F, Pifferi M, Michelucci A et al (2011) Torque teno virus viremia load size in patients with selected congenital defects of innate immunity. Clin Vaccine Immunol 18:692-694. https://doi.org/10.1128/CVI.00466-10
21. Cebria-Mendoza M, Bracho MA, Arbona C et al (2021) Exploring the Diversity of the Human Blood Virome. Viruses 13https://doi.org/10.3390/v13112322
22. Stapleton JT, Foung S, Muerhoff AS et al (2011) The GB viruses: a review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. The Journal of general virology 92:233-246. https://doi.org/10.1099/vir.0.027490-0
23. Chivero ET, Stapleton JT (2015) Tropism of human pegivirus (formerly known as GB virus C/hepatitis G virus) and host immunomodulation: insights into a highly successful viral infection. The Journal of general virology 96:1521-1532. https://doi.org/10.1099/vir.0.000086
24. Muerhoff AS, Tillmann HL, Manns MP et al (2003) GB virus C genotype determination in GB virus-C/HIV co-infected individuals. J Med Virol 70:141-149. https://doi.org/10.1002/jmv.10375

25. Schwarze-Zander C, Blackard JT, Zheng H et al (2006) GB virus C (GBV-C) infection in hepatitis C virus (HCV)/HIV-coinfected patients receiving HCV treatment: importance of the GBV-C genotype. J Infect Dis 194:410-419. https://doi.org/10.1086/505713
26. Alcalde R, Nishiya A, Casseb J et al (2010) Prevalence and distribution of the GBV-C/HGV among HIV-1-infected patients under anti-retroviral therapy. Virus Res 151:148-152. https://doi.org/10.1016/j.virusres.2010.04.008
27. Blackard JT, Ma G, Welge JA et al (2017) Cytokine/chemokine expression associated with Human Pegivirus (HPgV) infection in women with HIV. J Med Virol 89:1904-1911. https://doi.org/10.1002/jmv.24836
28. Bhattarai N, Stapleton JT (2012) GB virus C: the good boy virus? Trends Microbiol 20:124-130. https://doi.org/10.1016/j.tim.2012.01.004
29. Berzsenyi MD, Bowden DS, Roberts SK et al (2009) GB virus C genotype 2 predominance in a hepatitis C virus/HIV infected population associated with reduced liver disease. J Gastroenterol Hepatol 24:1407-1410. https://doi.org/10.1111/j.1440-1746.2009.05920.x
30. Miao Z, Gao L, Song Y et al (2017) Prevalence and Clinical Impact of Human Pegivirus-1 Infection in HIV-1-Infected Individuals in Yunnan, China. Viruses 9:28. https://doi.org/10.3390/v9020028
31. Ziemann M, Thiele T (2017) Transfusion-transmitted CMV infection - current knowledge and future perspectives. Transfus Med 27:238-248. https://doi.org/10.1111/tme.12437
32. Marshall BC, Adler SP (2009) The frequency of pregnancy and exposure to cytomegalovirus infections among women with a young child in day care. Am J Obstet Gynecol 200:163 e161-165. https://doi.org/10.1016/j.ajog.2008.08.037
33. Souza MA, Passos AM, Treitinger A et al (2010) Seroprevalence of cytomegalovirus antibodies in blood donors in southern, Brazil. Rev Soc Bras Med Trop 43:359-361. https://doi.org/10.1590/s0037-86822010000400004
34. Yao P, Millwood I, Kartsonaki C et al (2022) Sero-prevalence of 19 infectious pathogens and associated factors among middle-aged and elderly Chinese adults: a cross-sectional study. BMJ Open 12:e058353. https://doi.org/10.1136/bmjopen-2021-058353
35. Xu M, Yang Y, Zhou Y et al (2018) Metagenomics in pooled plasma, with identification of potential emerging infectious pathogens. Transfusion 58:633-637. https://doi.org/10.1111/trf.14455

Figures

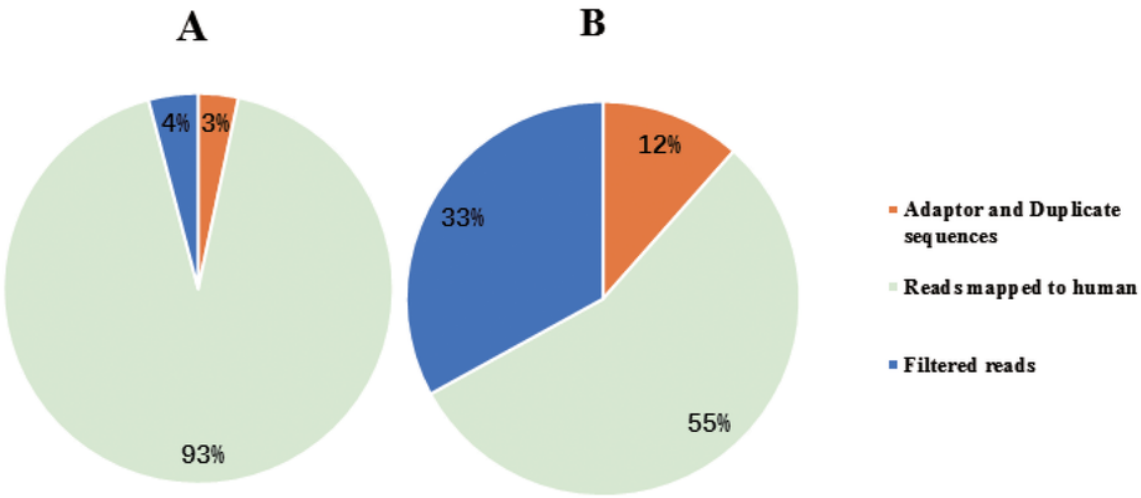

Figure 1

Pie charts of sequencing reads in plasma pools

The percentage of human sequences, filtered reads (non-human sequences), and low-complexity reads (adaptor and duplicate sequences) are represented by pie charts. (A) shows the result of DNA sequencing. (B) shows the result of cDNA sequencing.

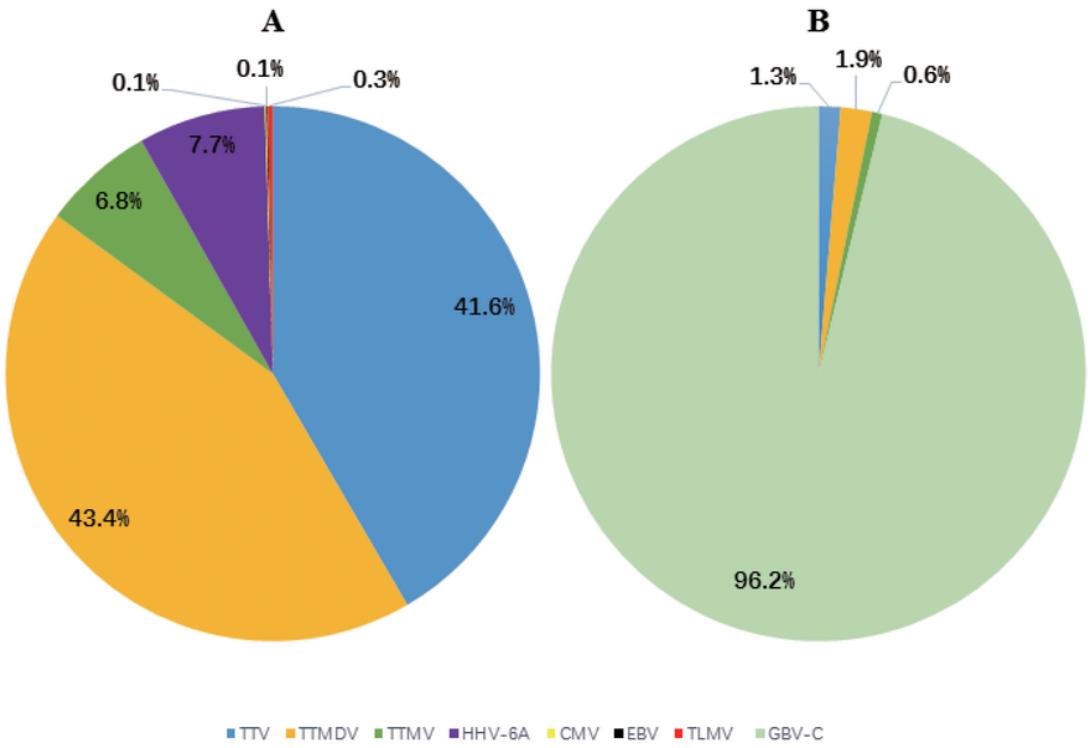

Figure 2

Summary of the metagenomics analysis of virus profiles in pooled plasma

The percentage for the areas corresponds to the percentage of number of reads detected in all the sequenced pools for each class of pathogens. (A) shows the result of DNA sequencing. The proportion of CMV, EBV and TLMV was 0.1%, 0.1% and 0.3%, respectively. (B) shows the result of cDNA sequencing.

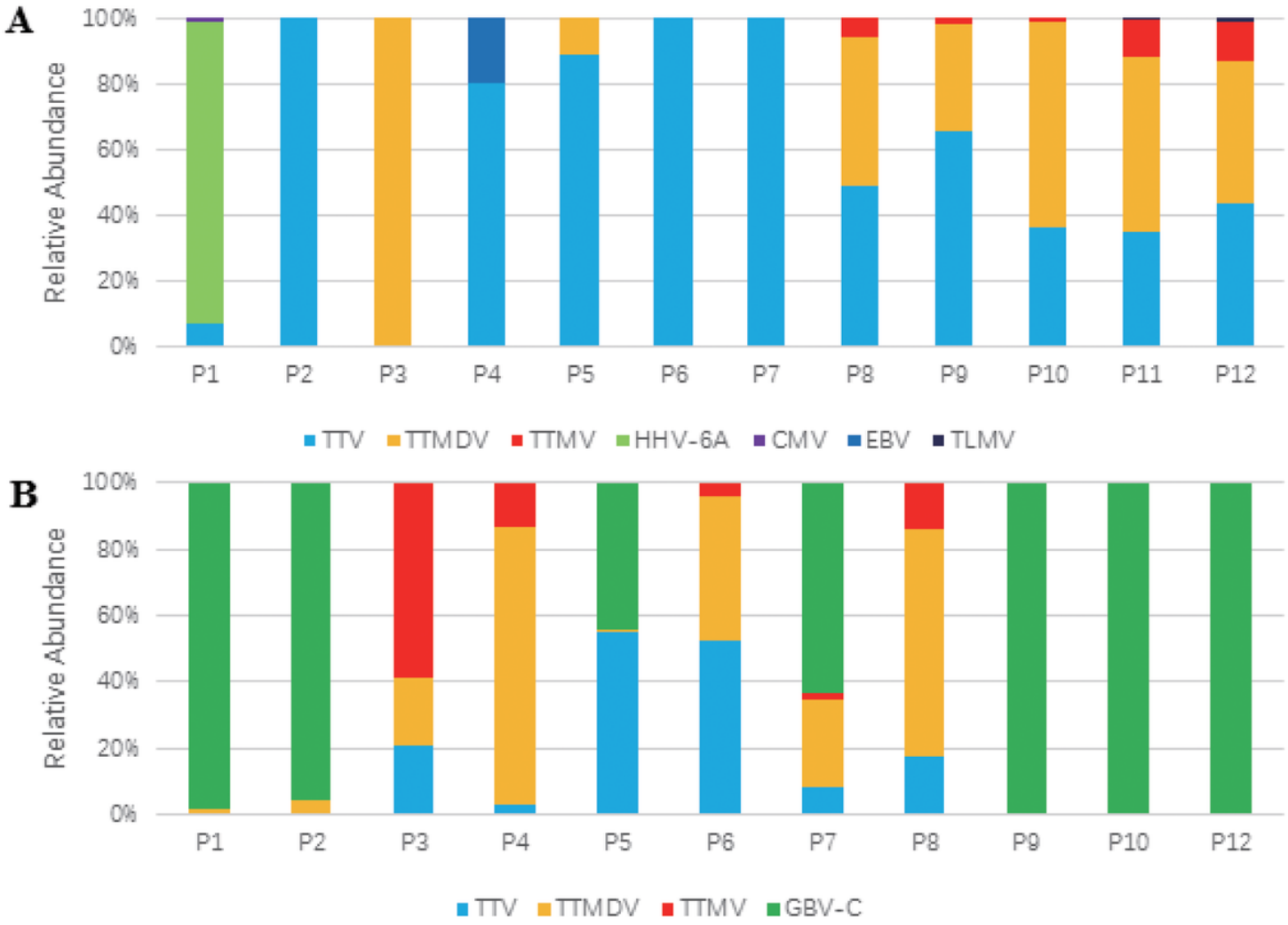

Figure 3

Relative abundance of nucleic acid sequences in each pool with virome analyzed using DNA sequencing (A) and cDNA sequencing (B)

The percentage ratio of read numbers of sequences matched with TTV (blue), GBV-C (green) and other viruses are indicated.

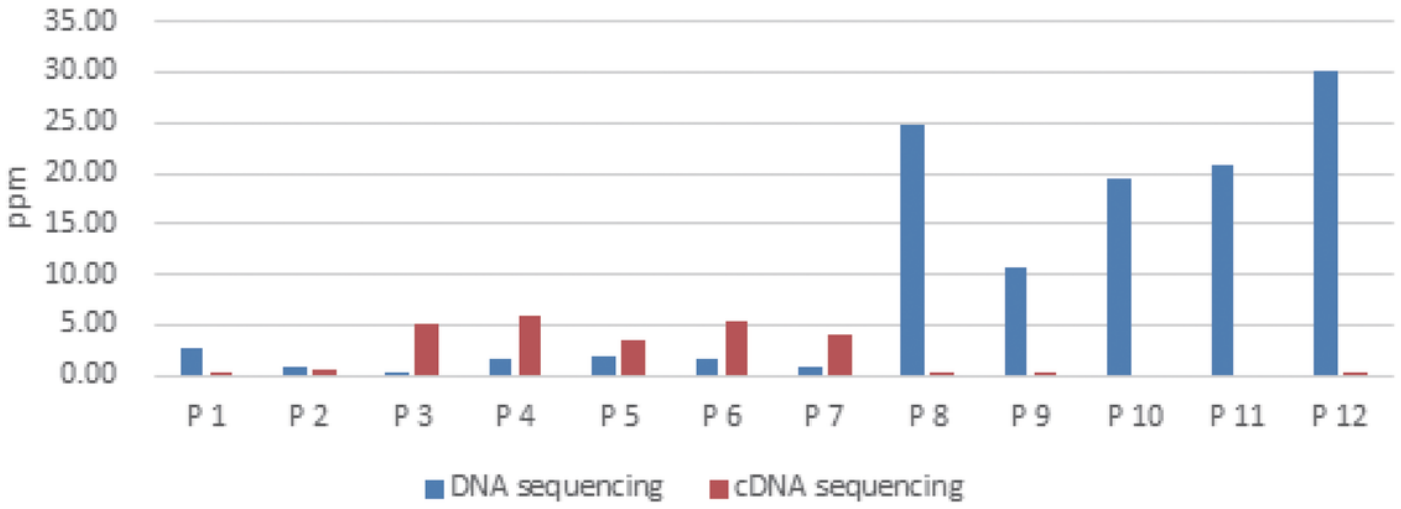

Figure 4

The abundance of anellovirus in each pool

Each bar represents the percentage ratio of read numbers in each plasma pool (anellovirus reads / filtered reads). ppm: parts per million

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Fig.S1.eps
- TableS1.docx
- TableS2.docx
- TableS3.docx