Traces of SARS-CoV-2 RNA in the Blood of COVID-19 Patients

Ahmed Moustafa[1,*], Ramy K. Aziz[2,3,*]

Affiliations

[1] Department of Biology, American University in Cairo, New Cairo, Egypt

[2] Department of Microbiology and Immunology, Faculty of Pharmacy, Cairo University, Cairo, Egypt

[3] Center for Genome and Microbiome Research, Cairo University, Cairo, Egypt

[*] Corresponding author

Emails

AM: amoustafa@aucegypt.edu

RKA: ramy.aziz@gmail.com

Short title

SARS-CoV-2 RNA in COVID-19 PBMC

Keywords

SARS-CoV-2, Coronavirus, COVID-19, Blood, PBMC, RNA-Seq

Abstract

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the third virus that caused coronavirus-related outbreaks over the past 20 years. The outbreak was first reported in December 2019 in Wuhan, China, but rapidly progressed into a pandemic of an unprecedented scale since the 1918 flu pandemic. Besides respiratory complications in COVID-19 patients, clinical characterizations of severe infection cases showed several other comorbidities, including multiple organ failure (liver, kidney, and heart) and septic shock. To better understand COVID-19 pathogenesis in different human organs, we interrogated the presence of the virus in the blood, or any of its components, which might provide a form of trafficking or hiding to the virus. By computationally analyzing high-throughput sequence data from patients with active COVID-19 infection, we found evidence of only traces of SARS-CoV-2 RNA in peripheral blood mononuclear cells (PBMC), while the virus RNA was abundant in bronchoalveolar lavage specimens from the same patients. We also devised a 'viral spike-to-actin' RNA normalization, as a metric to compare across various samples and minimize errors caused by intersample variability in human RNA. To the best of our knowledge, the presence of SARS-CoV-2 RNA in the PBMC of COVID-19 patients has not been reported before, and this observation could suggest immune presentation, but discounts the possibility of extensive viral infection of lymphocytes or monocytes.

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the third virus that caused coronavirus-related outbreaks over the past 20 years. The first outbreak occurred in Asia in 2002-2003 causing Severe Acute Respiratory Syndrome (SARS), hence, the name, SARS-CoV, which back then was not related to any of the known viruses (Marra et al. 2003; Rota et al. 2003). Between 2002 and 2003, 8,098 people became sick with SARS, and of those 774 died (i.e., a mortality rate of 9.5%). Since 2004, there have been no more reports of SARS cases (via NHS, WHO, and CDC).

The second coronavirus-related outbreak started in the Arabian Peninsula in 2012 (Zaki et al. 2012) causing a more fatal disease, Middle East respiratory syndrome (MERS), with a

significantly higher mortality rate of 40% of the cases infected by MERS-CoV virus (Zumla, Hui, and Perlman 2015).

More recently, in December 2019, the third coronavirus-related outbreak was first reported in Wuhan, China, by the emergence of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), initially dubbed "the 2019 novel coronavirus" (2019-nCoV). The spread of the virus led to a pandemic of an unprecedented scale since the 1918 flu pandemic.

As of May 30, nearly 6 million confirmed COVID-19 cases globally and > 350,000 deaths have been reported by WHO (WHO Dashboard, continuously updated). Besides the respiratory complications in COVID-19 patients, clinical characterizations of severe infection cases indicated further comorbidities, including multiple organ failure (liver, kidney, and heart) and septic shock (Poston, Patel, and Davis 2020; Cascella et al. 2020; Hui Li et al. 2020).

The genome sequence of SARS-COV-2 has been determined and made public (Lu et al. 2020), and since then thousands (over 35,000) of genomes have been sequenced from all around the world (Shu and McCauley 2017)https://www.gisaid.org(Shu and McCauley 2017). The availability of those genomic sequences allows the rapid screening of viral RNA in human tissues as well as environmental samples (e.g., sewage (Bibby and Peccia 2013)) using multi-omic wet lab technologies, and *in silico* screening tools, for publicly available metatranscriptomic samples.

To better understand COVID-19 pathogenesis in different human organs, we conducted this study to interrogate the presence of the virus in the blood, or any of its components, as it might provide a form of trafficking or hiding to the virus, notably that some precarious studies reported the ability of the virus to infect lymphocytes (X. Wang et al. 2020) whereas others suggested the virus exerts its pathogenesis through "attacking hemoglobin," although this hypothesis has been heavily criticized (Read 2020). The virus was reported to be found in the plasma of COVID-19 patients (Huang et al. 2020). Finally, peripheral blood mononuclear cells (PBMC) were shown to harbor other infectious viruses, such as HIV, HCV, and HBV (W.-K. Wang et al. 2002; Z. Li, Hou, and Cao 2015).

For the aforementioned reasons, we computationally analyzed high-throughput sequence data from patients with active COVID-19 infection, and found evidence of traces of SARS-CoV-2 RNA in their PMBCs, while their bronchoalveolar lavage samples had large amounts of the viral RNA.

Methods

Dataset

Publicly accessible raw RNA-Seq FASTQ sequences published by Xiong et al. (Xiong et al. 2020) were obtained from the Genome Sequence Archive (Y. Wang et al. 2017) (GSA accession CRA002390). Raw sequences were processed for quality control using fastp (Chen et al. 2018).

Comparison of RNA Abundance and Gene Expression Profiles

Filtered FASTQ sequences were searched with blastx (Altschul et al. 1990) against the RefSeq protein database (O'Leary et al. 2016) (release 99) using DIAMOND (Buchfink, Xie, and Huson 2015) with an *e*-value cutoff < 1e-10. The counts of the matching RefSeq accessions were normalized to the total number of reads per library. Principal Component Analysis (PCA) was performed on the normalized abundance of the RefSeq proteins. For functional annotation, the matching RefSeq proteins were annotated by Pfam (Finn et al. 2014) using HMMER (Eddy 2011).

Detection of Viral RNA

Filtered FASTQ sequences were aligned to the SARS-CoV-2 reference genome (GenBank accession NC_045512) using the Burrows-Wheeler aligner (BWA, (Heng Li and Durbin 2009)). Generated binary alignment map (BAM) files were filtered for mapped sequences with a quality score > 40 and alignment score > 90 using Sambamba (Tarasov et al. 2015). Identified SARS-CoV-2 matching sequences were manually inspected and searched against NCBI "nt" by BLAST (blastn) for verification (Altschul et al. 1990).

Data normalization

To estimate viral RNA abundance in a given sample, we normalized any positive hits to SARS-CoV-2 to the total number of reads within that sample. Additionally, we used the spike gene, being a highly specific SARS-CoV-2 gene to estimate the extent of viral RNA load in a given sample, and we normalized the abundance of spike genes to human actin RNA, being a transcript of a housekeeping gene. This normalization generated a 'spike-to-actin' ratio that could be used as an accurate metric for viral RNA abundance relative to human RNA.

Results

We analyzed the RNA-Seq dataset published by Xiong et al. (Xiong et al. 2020). **Table S1** describes the processed samples, and the sequence reads statistics of each sample:

Based on the global gene expression in PBMC, PCA showed a strong, expected separation between BALF and PBMC samples, and further slight separation in the PBMC from healthy controls and COVID-19 patients (**Figure 1**). However, one healthy control was placed within the COVID-19 cluster.

We found viral sequences in all the BALF samples with a median of abundance 2.15% of the total reads (**Table S1**). We also identified 2 paired-ended viral reads in a PBMC sample of a COVID-19 patient, matching the SARS-CoV-2 polyprotein (pp1ab) (accession NP_828849) and SARS-CoV-2 surface glycoprotein (accession YP_009724390) (**Figures 2, 3**).

When we calculated a spike-to-actin RNA ratio for each sample, values for all four BALF samples, as well as the one PBMC sample with SARS-CoV-2 transcripts, followed the same pattern of viral RNA abundance values, and strongly correlated (Spearman correlation $p = 0.96$, p-value = 0.00047, **Figure 4C**)

Discussion

Coronavirus-related infections have been reported to be associated with hematological changes including lymphopenia, thrombocytopenia, and leukopenia through infecting blood cells, bone marrow stromal cells, or inducing autoantibodies (Yang et al. 2003).

In a former study characterizing the clinical features of COVID-19 patients, Huang and coworkers showed that using RT-PCR allowed them to detect coronavirus in plasma isolated samples from the patients (Huang et al. 2020). In their report, they preferred to use the term "RNAaemia," rather than "viraemia," which they defined as the presence of a virus in the blood because they did not perform tests to confirm the presence of an infectious SARS-CoV-2 virus in the blood of the patients.

There have been no reports of detecting viral RNA in blood cells, notably PBMC. In their preliminary analysis of RNA isolated from PBMC, Corley et al confirmed that they did not detect viral sequences (a preprint 10.1101/2020.04.13.039263v1).

Using high-throughput sequencing has been repeatedly demonstrated to be an effective approach for the identification and quantification of viruses in the blood (Moustafa et al. 2017) following similar methods as those used in viral metagenomics (Breitbart et al. 2003);(Aziz et al. 2015)) and uncultivated viral genomics ((Roux et al. 2019).

Therefore, we planned to exhaustively look for SARS-CoV-2 RNA sequences in publicly available RNA-Seq PBMC datasets. As of the writing of this report, there has been only one publicly available RNA-Seq PBMC dataset published by (Xiong et al. 2020), in which the group profiled global gene expression in BALF and PBMC specimens of COVID-19. Predictably, we detected viral RNA in all BALF samples (2 patients, 2 replicates each) with an average abundance of 2.15% of the total RNA in those samples, which included human RNA (**Table S1 and Figure 4)**.

Of note, to estimate the viral RNA load in a given sample and to compare between different samples, we normalized viral read counts to the total number of reads in each data set. Since the total number of human-related reads may be strongly affected by potential microbial RNA (from pulmonary microbiota, for example) or by extensive viral counts, we also sought to

estimate the ratio of viral genomic RNA to human cellular RNA. For that purpose, we used actin RNA, being a transcript for a human housekeeping gene—not expected to be seen in bacterial cells or viral genomes, and we estimated the ratio of hits to RNA encoding the spike protein (S) copies to actin transcripts in the sample. Quite interestingly, the spike-to-actin ratio strongly correlated with viral RNA abundance values (computed as SARS-CoV-2 RNA reads normalized to the total number of reads), **Figure 4C**.

As for blood samples, we identified two paired-end reads in RNA isolated from PBMC from only one (accession CRR119891) out of three patients, which aligned to the SARS-CoV-2 genome. Expectedly, no SARS-CoV-2 RNA was detected in the PBMC from healthy controls (three donors).

These RNA traces are certainly quite rare; however, they confidently and specifically belong to SARS-COV-2 (**Figure 3**). One viral RNA read translates into polyprotein (pp1ab, accession NP_828849), which is the largest protein of coronaviruses and involved in the replication and transcription of the viral genome. The other viral RNA translates into surface (spike) glycoprotein (accession YP_009724390), which mediates the entry of the viral into the human cells expressing human angiotensin-converting enzyme 2 (hACE2) (Ou et al. 2020; Walls et al. 2020). Although we are not rejecting the possibility of cross-contamination or barcode bleeding (Mitra et al. 2015; Kircher, Sawyer, and Meyer 2012) for detecting the viral RNA in one PBMC RNA-Seq sample, such possibility is unlikely, given that control samples had zero hits to SARS-CoV-2 RNA. We are also considering the possibility of SARS-CoV-2 being sampled by antigen-presenting cells (most likely dendritic cells) or presented to T lymphocytes, which are in the PBMC population.

One more possibility, which we believe needs many more samples to consider, is that SARS-CoV-2 may be specifically or coincidentally internalized by one of the mononuclear cell types, which may suggest a mechanism for the chronicity of the SARS-CoV-2 infection. This hypothesis requires further testing. However, in the light of our data, it is hard to support the early reports that SARS-COV-2 targets T-lymphocytes *in vivo*, as suggested earlier, in a correspondence, based on cell culture experiment with pseudotyped viruses (X. Wang et al. 2020).

With more data becoming publicly available, it will be possible to revisit this hypothesis and others to improve our understanding of the progression and replication of SARS-CoV-2 in infected individuals.

References

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. “Basic Local Alignment Search Tool.” *Journal of Molecular Biology* 215 (3): 403–10.

Aziz, Ramy K., Bhakti Dwivedi, Sajia Akhter, Mya Breitbart, and Robert A. Edwards. 2015. “Multidimensional Metrics for Estimating Phage Abundance, Distribution, Gene Density, and Sequence Coverage in Metagenomes.” *Frontiers in Microbiology* 6 (May): 381.

Bibby, Kyle, and Jordan Peccia. 2013. “Identification of Viral Pathogen Diversity in Sewage Sludge by Metagenome Analysis.” *Environmental Science & Technology* 47 (4): 1945–51.

Breitbart, Mya, Ian Hewson, Ben Felts, Joseph M. Mahaffy, James Nulton, Peter Salamon, and Forest Rohwer. 2003. “Metagenomic Analyses of an Uncultured Viral Community from Human Feces.” *Journal of Bacteriology* 185 (20): 6220–23.

Buchfink, Benjamin, Chao Xie, and Daniel H. Huson. 2015. “Fast and Sensitive Protein Alignment Using DIAMOND.” *Nature Methods* 12 (1): 59–60.

Cascella, Marco, Michael Rajnik, Arturo Cuomo, Scott C. Dulebohn, and Raffaela Di Napoli. 2020. “Features, Evaluation and Treatment Coronavirus (COVID-19).” In *StatPearls*. Treasure Island (FL): StatPearls Publishing.

Chen, Shifu, Yanqing Zhou, Yaru Chen, and Jia Gu. 2018. “Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor.” *Bioinformatics* 34 (17): i884–90.

Eddy, Sean R. 2011. “Accelerated Profile HMM Searches.” *PLoS Computational Biology* 7 (10): e1002195.

Finn, Robert D., Alex Bateman, Jody Clements, Penelope Coggill, Ruth Y. Eberhardt, Sean R. Eddy, Andreas Heger, et al. 2014. “Pfam: The Protein Families Database.” *Nucleic Acids Research* 42 (Database issue): D222–30.

Huang, Chaolin, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, et al. 2020. “Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China.” *The Lancet* 395 (10223): 497–506.

Kircher, Martin, Susanna Sawyer, and Matthias Meyer. 2012. “Double Indexing Overcomes Inaccuracies in Multiplex Sequencing on the Illumina Platform.” *Nucleic Acids Research* 40 (1): e3.

Li, Heng, and Richard Durbin. 2009. “Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform.” *Bioinformatics* 25 (14): 1754–60.

Li, Hui, Liang Liu, Dingyu Zhang, Jiuyang Xu, Huaping Dai, Nan Tang, Xiao Su, and Bin Cao. 2020. “SARS-CoV-2 and Viral Sepsis: Observations and Hypotheses.” *The Lancet*, April. https://doi.org/10.1016/S0140-6736(20)30920-X.

Li, Zixiong, Xiaomei Hou, and Guangwen Cao. 2015. “Is Mother-to-Infant Transmission the Most Important Factor for Persistent HBV Infection?” *Emerging Microbes & Infections* 4 (5): e30.

Lu, Roujian, Xiang Zhao, Juan Li, Peihua Niu, Bo Yang, Honglong Wu, Wenling Wang, et al. 2020. “Genomic Characterisation and Epidemiology of 2019 Novel Coronavirus:

Implications for Virus Origins and Receptor Binding." *The Lancet* 395 (10224): 565–74.
Marra, Marco A., Steven J. M. Jones, Caroline R. Astell, Robert A. Holt, Angela Brooks-Wilson, Yaron S. N. Butterfield, Jaswinder Khattra, et al. 2003. "The Genome Sequence of the SARS-Associated Coronavirus." *Science* 300 (5624): 1399–1404.
Mitra, Abhishek, Magdalena Skrzypczak, Krzysztof Ginalski, and Maga Rowicka. 2015. "Strategies for Achieving High Sequencing Accuracy for Low Diversity Samples and Avoiding Sample Bleeding Using Illumina Platform." *PloS One* 10 (4): e0120520.
Moustafa, Ahmed, Chao Xie, Ewen Kirkness, William Biggs, Emily Wong, Yaron Turpaz, Kenneth Bloom, et al. 2017. "The Blood DNA Virome in 8,000 Humans." *PLoS Pathogens* 13 (3): e1006292.
O'Leary, Nuala A., Mathew W. Wright, J. Rodney Brister, Stacy Ciufo, Diana Haddad, Rich McVeigh, Bhanu Rajput, et al. 2016. "Reference Sequence (RefSeq) Database at NCBI: Current Status, Taxonomic Expansion, and Functional Annotation." *Nucleic Acids Research* 44 (D1): D733–45.
Ou, Xiuyuan, Yan Liu, Xiaobo Lei, Pei Li, Dan Mi, Lili Ren, Li Guo, et al. 2020. "Characterization of Spike Glycoprotein of SARS-CoV-2 on Virus Entry and Its Immune Cross-Reactivity with SARS-CoV." *Nature Communications* 11 (1): 1620.
Poston, Jason T., Bhakti K. Patel, and Andrew M. Davis. 2020. "Management of Critically Ill Adults With COVID-19." *JAMA: The Journal of the American Medical Association*, March. https://doi.org/10.1001/jama.2020.4914.
Read, Randy. 2020. "Flawed Methods in 'COVID-19: Attacks the 1-Beta Chain of Hemoglobin and Captures the Porphyrin to Inhibit Human Heme Metabolism.'"
Rota, Paul A., M. Steven Oberste, Stephan S. Monroe, W. Allan Nix, Ray Campagnoli, Joseph P. Icenogle, Silvia Peñaranda, et al. 2003. "Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome." *Science* 300 (5624): 1394–99.
Roux, Simon, Evelien M. Adriaenssens, Bas E. Dutilh, Eugene V. Koonin, Andrew M. Kropinski, Mart Krupovic, Jens H. Kuhn, et al. 2019. "Minimum Information about an Uncultivated Virus Genome (MIUViG)." *Nature Biotechnology* 37 (1): 29–37.
Shu, Yuelong, and John McCauley. 2017. "GISAID: Global Initiative on Sharing All Influenza Data - from Vision to Reality." *Euro Surveillance: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin* 22 (13). https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494.
Tarasov, Artem, Albert J. Vilella, Edwin Cuppen, Isaac J. Nijman, and Pjotr Prins. 2015. "Sambamba: Fast Processing of NGS Alignment Formats." *Bioinformatics* 31 (12): 2032–34.
Walls, Alexandra C., Young-Jun Park, M. Alejandra Tortorici, Abigail Wall, Andrew T. McGuire, and David Veesler. 2020. "Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein." *Cell* 181 (2): 281–92.e6.
Wang, Wei-Kung, Tzu-Ling Sung, Yu-Chen Tsai, Chuan-Liang Kao, Shu-Mei Chang, and Chwan-Chuen King. 2002. "Detection of Dengue Virus Replication in Peripheral Blood Mononuclear Cells from Dengue Virus Type 2-Infected Patients by a Reverse Transcription-Real-Time PCR Assay." *Journal of Clinical Microbiology* 40 (12): 4472–78.
Wang, Xinling, Wei Xu, Gaowei Hu, Shuai Xia, Zhiping Sun, Zezhong Liu, Youhua Xie, Rong Zhang, Shibo Jiang, and Lu Lu. 2020. "SARS-CoV-2 Infects T Lymphocytes through Its Spike Protein-Mediated Membrane Fusion." *Cellular & Molecular Immunology*, April. https://doi.org/10.1038/s41423-020-0424-9.
Wang, Yanqing, Fuhai Song, Junwei Zhu, Sisi Zhang, Yadong Yang, Tingting Chen, Bixia Tang, et al. 2017. "GSA: Genome Sequence Archive<sup/>." *Genomics, Proteomics &*

Bioinformatics 15 (1): 14–18.

Xiong, Yong, Yuan Liu, Liu Cao, Dehe Wang, Ming Guo, Ao Jiang, Dong Guo, et al. 2020. "Transcriptomic Characteristics of Bronchoalveolar Lavage Fluid and Peripheral Blood Mononuclear Cells in COVID-19 Patients." *Emerging Microbes & Infections* 9 (1): 761–70.

Yang, Mo, Kam-Lun E. Hon, Karen Li, Tai-Fai Fok, and Chi-Kong Li. 2003. "The Effect of SARS Coronavirus on Blood System: Its Clinical Findings and the Pathophysiologic Hypothesis." *Zhongguo Shi Yan Xue Ye Xue Za Zhi / Zhongguo Bing Li Sheng Li Xue Hui = Journal of Experimental Hematology / Chinese Association of Pathophysiology* 11 (3): 217–21.

Zaki, Ali M., Sander van Boheemen, Theo M. Bestebroer, Albert D. M. E. Osterhaus, and Ron A. M. Fouchier. 2012. "Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia." *The New England Journal of Medicine* 367 (19): 1814–20.

Zumla, Alimuddin, David S. Hui, and Stanley Perlman. 2015. "Middle East Respiratory Syndrome." *The Lancet* 386 (9997): 995–1007.

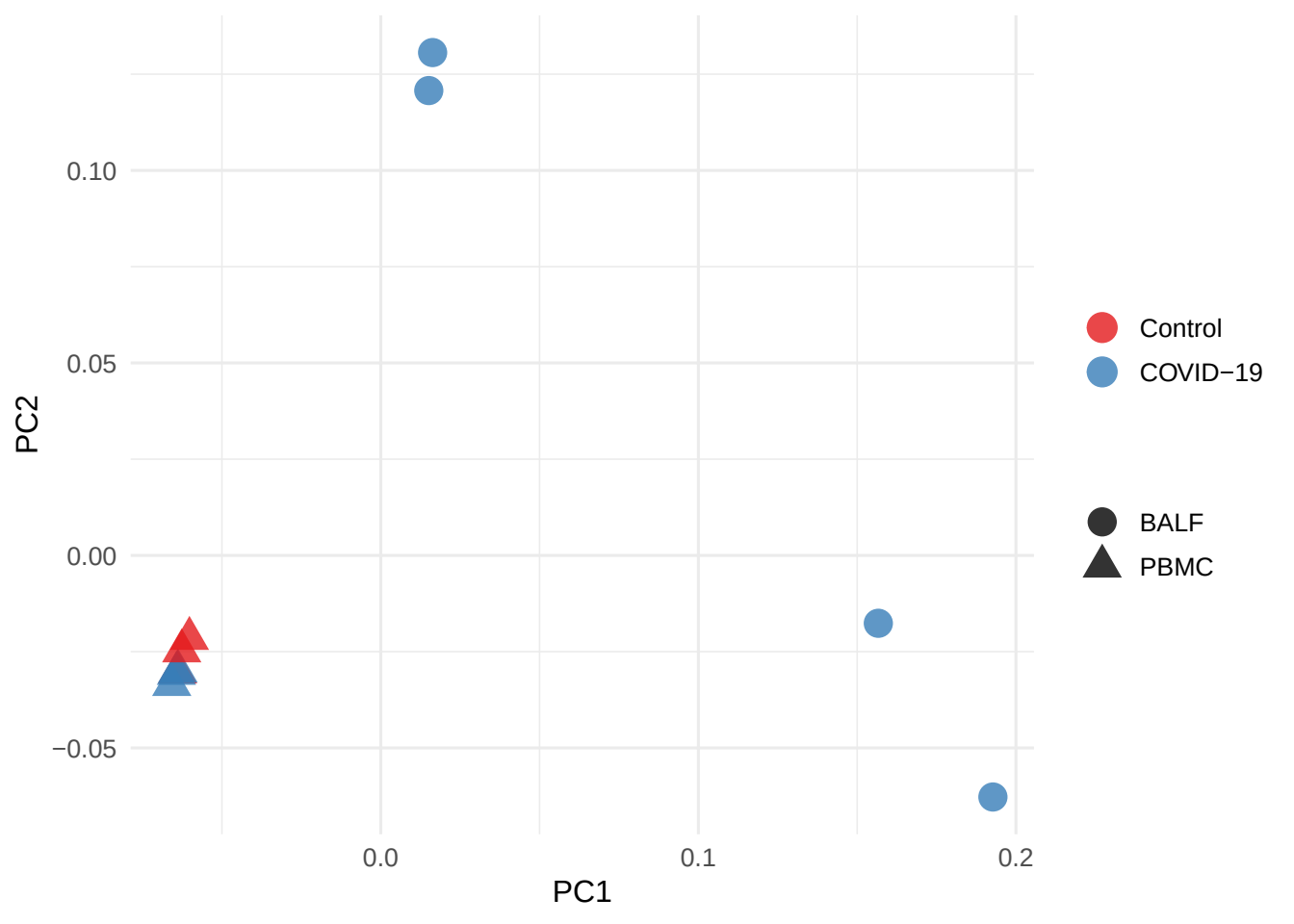
Figures

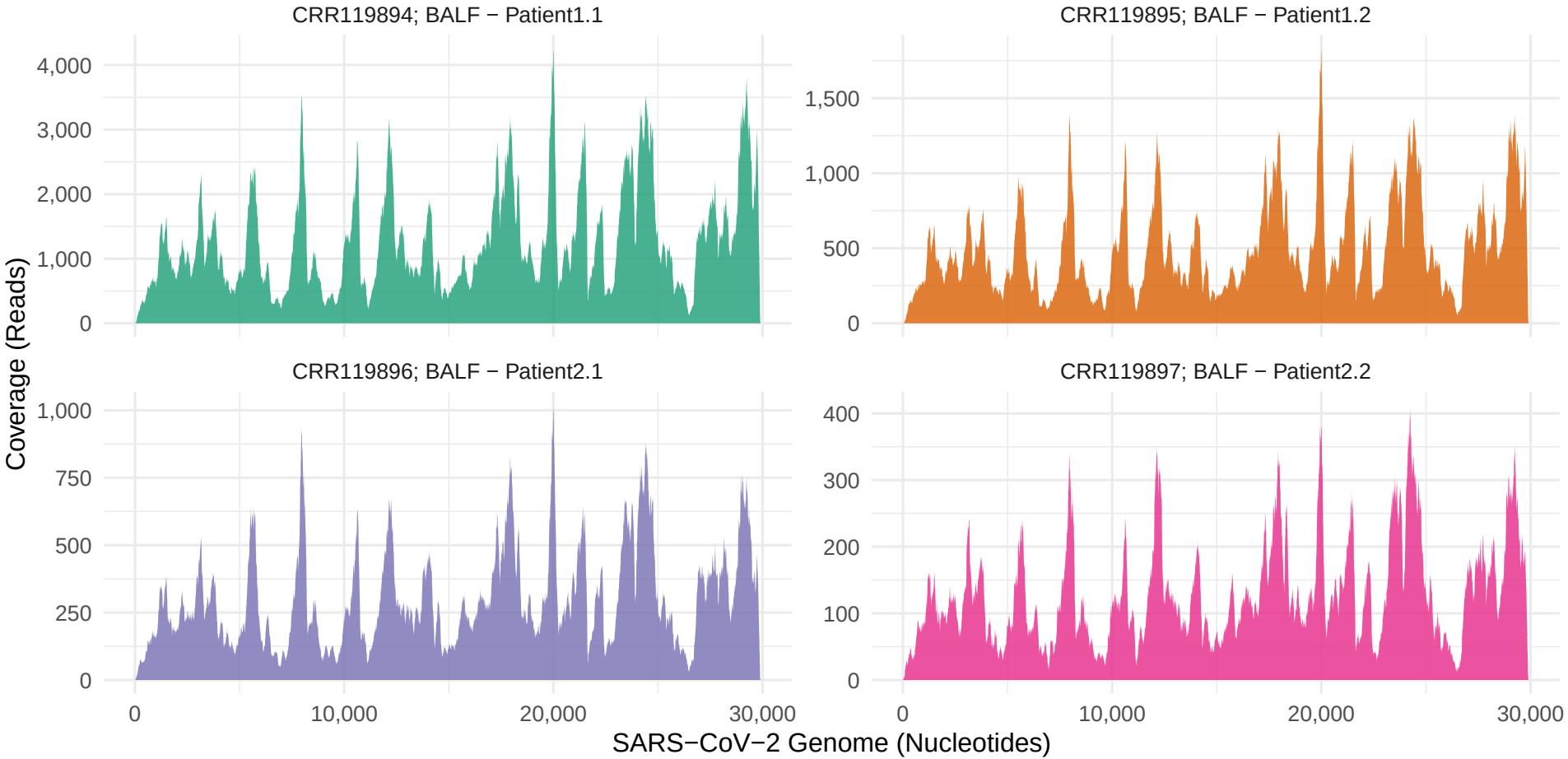
Figure 1. Principal Component Analysis based on RNA abundance data.
Samples are color-coded; red for healthy donors, and green for COVID-19 patients. PC1 explains 63% of the variance and PC2 explains 30% of the variance.

Figure 2. Mapping of SARS-CoV-2 RNA against the SARS-CoV-2 genome.
Viral RNA sequences from the COVID-19 patients were mapped against the SARS-CoV-2 reference genome (GenBank accession NC_045512). The x-axis is the nucleotide position on the virus genome. The y-axis is the coverage (in reads; not normalized) of the genomic position by RNA-Seq viral sequences. Panel "A" represents the BALF samples. Panel "B" represents the PBMC samples.

Figure 3. Alignment of the detected viral RNA sequences to the SARS-CoV-2 genome
The two paired-end viral reads from a COVID-19 PBMC sample aligned to the SARS-CoV-2 reference genome (GenBank accession NC_045512). Panel A is the alignment first paired-end read. Panel B is the alignment second paired-end read. Perfect identities between the reference and the two overlapping mates are indicated by asterisks ("*").

Figure 4. SARS-CoV-2 viral abundance relative to the host gene expression
A) Normalized abundance of SARS-CoV-2 viral reads, calculated number of SARS-CoV-2 specific hits per million sequence reads. B) SARS-CoV-2 spike-to-human actin ratio for each sample. C) A scatter plot showing the correlation between SARS-CoV-2 spike-to-human actin ratio and SARS-CoV-2 viral reads abundance (in ppm).


0.10
0.05
0.00
−0.05
PC2
0.0
0.1
0.2
PC1
Control
COVID−19
BALF
PBMC


CRR119894; BALF − Patient1.1
CRR119895; BALF − Patient1.2
CRR119896; BALF − Patient2.1
CRR119897; BALF − Patient2.2
Coverage (Reads)
4,000
3,000
2,000
1,000
0
1,500
1,000
500
0
1,000
750
500
250
0
400
300
200
100
0
0
10,000
20,000
30,000
SARS−CoV−2 Genome (Nucleotides)

A

```
NC_045512        ttctttgctcaggatggtaatgctgctatcagcgattatgactactatcgttataatcta
_R_V300046811L2  ------------------------------------------------------------
V300046811L2C00  ----------------------------------------ctactatcgttataatcta

NC_045512        ccaacaatgtgtgatatcagacaactactatttgtagttgaagttgttgataagtacttt
_R_V300046811L2  -------------------------------------------------------------t
V300046811L2C00  ccaacaatgtgtgatatcagacaactactatttgtagttgaagttgttgataagtacttt
                                                                             *

NC_045512        gattgttacgatggtggctgtattaatgctaaccaagtcatcgtcaacaacctagacaaa
_R_V300046811L2  gattgttacgatggtggctgtattaatgctaaccaagtcatcgtcaacaacctagacaaa
V300046811L2C00  gattgttacgatggtggctgt--------------------------------------
                 **********************
```

B

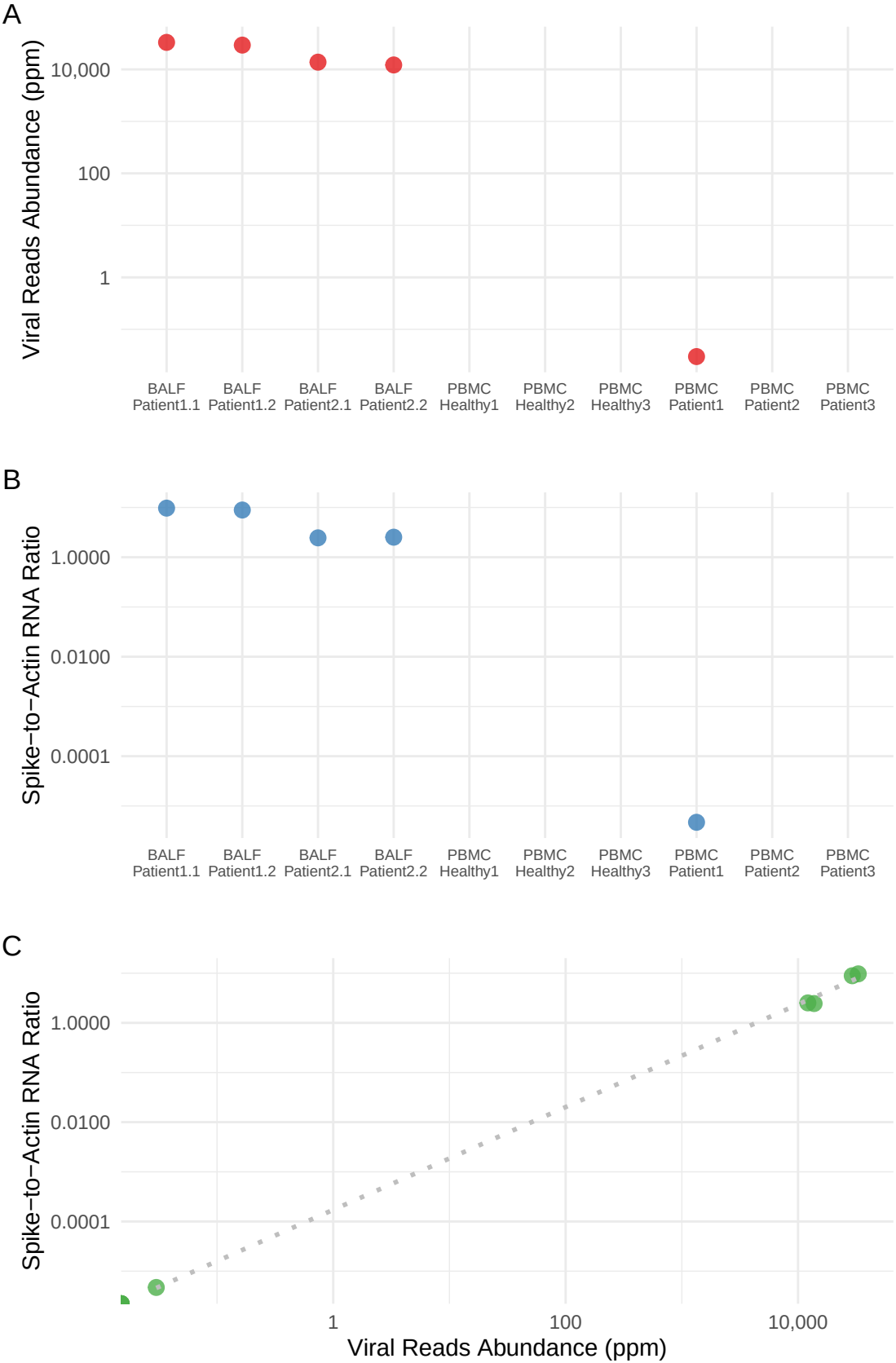
```
NC_045512        gttgcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttga
_R_New|V3000468  ------------------------------------------------------------
New|V300046811L  ------------------------------aaaccgtgctttaactggaatagctgttga

NC_045512        acaagacaaaaacacccaagaagtttttgcacaagtcaaacaaatttacaaaacaccacc
_R_New|V3000468  ------------------------------------------------------------
New|V300046811L  acaagacaaaaacacccaagaagtttttgcacaagtcaaacaaatttacaaaacaccacc

NC_045512        aattaaagattttggtggttttaatttttcacaaatattaccagatccatcaaaaccaag
_R_New|V3000468  ------------------------------------------------------------
New|V300046811L  aattaaagat--------------------------------------------------

NC_045512        caagaggtcatttattgaagatctacttttcaacaaagtgacacttgcagatgctggctt
_R_New|V3000468  ---gaggtcatttattgaagatctacttttcaacaaagtgacacttgcagatgctggctt
New|V300046811L  ------------------------------------------------------------

NC_045512        catcaaacaatatggtgattgccttggtgatattgctgctagagacctcatttgtgcaca
_R_New|V3000468  catcaaacaatatggtgattgccttggtgatattgctgctaga-----------------
New|V300046811L  ------------------------------------------------------------
```


A
Viral Reads Abundance (ppm)
10,000
100
1
BALF Patient1.1
BALF Patient1.2
BALF Patient2.1
BALF Patient2.2
PBMC Healthy1
PBMC Healthy2
PBMC Healthy3
PBMC Patient1
PBMC Patient2
PBMC Patient3
B
Spike−to−Actin RNA Ratio
1.0000
0.0100
0.0001
BALF Patient1.1
BALF Patient1.2
BALF Patient2.1
BALF Patient2.2
PBMC Healthy1
PBMC Healthy2
PBMC Healthy3
PBMC Patient1
PBMC Patient2
PBMC Patient3
C
Spike−to−Actin RNA Ratio
1.0000
0.0100
0.0001
1
100
10,000
Viral Reads Abundance (ppm)