

RESEARCH

Open Access



Metagenomic analysis reveals differences in the co-occurrence and abundance of viral species in SARS-CoV-2 patients with different severity of disease

Pavel Iša^{1*}, Blanca Taboada¹, Rodrigo García-López¹, Celia Boukadida², José Ernesto Ramírez-González³, Joel Armando Vázquez-Pérez⁴, Alejandra Hernández-Terán⁴, José Ángel Romero-Espinoza⁴, José Esteban Muñoz-Medina⁵, Concepción Grajales-Muñiz⁵, Alma Rincón-Rubio², Margarita Matías-Florentino², Alejandro Sanchez-Flores⁶, Edgar Mendieta-Condado³, Gisela Barrera-Badillo³, Susana López¹, Lucía Hernández-Rivas³, Irma López-Martínez³, Santiago Ávila-Ríos² and Carlos F. Arias¹

Abstract

Background: SARS-CoV-2 infections have a wide spectrum of clinical manifestations whose causes are not completely understood. Some human conditions predispose to severe outcome, like old age or the presence of comorbidities, but many other facets, including coinfections with other viruses, remain poorly characterized.

Methods: In this study, the eukaryotic fraction of the respiratory virome of 120 COVID-19 patients was characterized through whole metagenomic sequencing.

Results: Genetic material from respiratory viruses was detected in 25% of all samples, whereas human viruses other than SARS-CoV-2 were found in 80% of them. Samples from hospitalized and deceased patients presented a higher prevalence of different viruses when compared to ambulatory individuals. Small circular DNA viruses from the *Anelloviridae* (Torque teno midi virus 8, TTV-like mini virus 19 and 26) and *Cycloviridae* families (Human associated cyclovirus 10), Human betaherpesvirus 6, were found to be significantly more abundant in samples from deceased and hospitalized patients compared to samples from ambulatory individuals. Similarly, *Rotavirus A*, *Measles morbillivirus* and *Alphapapillomavirus 10* were significantly more prevalent in deceased patients compared to hospitalized and ambulatory individuals.

Conclusions: Results show the suitability of using metagenomics to characterize a broader peripheral virological landscape of the eukaryotic virome in SARS-CoV-2 infected patients with distinct disease outcomes. Identified prevalent viruses in hospitalized and deceased patients may prove important for the targeted exploration of coinfections that may impact prognosis.

Keywords: SARS-CoV-2, Respiratory viruses, Metagenomic sequencing, COVID-19, Disease severity

*Correspondence: pavel.isa@ibt.unam.mx

¹ Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

Full list of author information is available at the end of the article

Introduction

The World Health Organization declared SARS-CoV-2 a world pandemic on March 11th, 2020. Since it was first reported in late 2019 in Wuhan, China, it has spread worldwide. SARS-CoV-2 is the etiologic agent of the



respiratory disease COVID-19, which can have different pathological presentations, from asymptomatic and mild respiratory illness to severe pneumonia with high fatality rates. Different factors, including age [1], presence of comorbidities [2–4], viral genetic factors [5, 6], host genetic constellations [7], and even coinfection with other pathogens and the composition of the microbiome [8], have been reported to affect the outcome and severity of the disease. In Mexico, there is a high prevalence of several comorbidities associated with the severity of COVID-19, specifically overweight and obesity, which are associated with diabetes type 2 and high blood pressure [2, 4, 9].

Since the first observation of the new coronavirus, there have been reports of the occasional presence of coinfections with other viruses [10, 11], and some of these coinfections have been reported in patients with a severe pathological course [12, 13]. However, despite the importance of the presence of other viruses, data on viral coinfections remain limited, focusing primarily on specific respiratory viruses through RT-PCR assays [10–20], and few studies have characterized the samples using high-throughput sequencing [21–25]. Some studies have reported a relatively high presence of respiratory viruses other than SARS-CoV-2, with 20–45% of samples being positive for other respiratory viruses [12, 19, 22]; however, generally, low levels of coinfections have been reported (less than 5% of samples containing other respiratory viruses) [10, 14, 17, 20, 26, 27].

We characterized respiratory samples from 120 subjects with different clinical symptoms in the present study using a classical metagenomic random-amplification (shotgun) method. The global analysis of viral populations showed many different human viral species and viruses from other origin such as plants and insects. We observed differences in the abundance of human viral species when samples from ambulatory, hospitalized, and deceased individuals were compared, being higher in hospitalized and deceased individuals compared to ambulatory patients.

Methods

The ethical statement, sample collection, and diagnostics

Samples and metadata collected for this work are considered part of the national response to COVID-19 and are directly related to prevention and disease control. Samples used were collected between the 13th of March and the 1st of May 2020 and processed under the Mexican Official NOM-017-SSA2-2012 (http://sersalud.cdmx.gob.mx/portalut/archivo/Art121FI/Normatividad_SSPDF/NOM-017-SSA2-2012.pdf) for epidemiological surveillance of Viral Respiratory Disease, emitted and approved by the CONAVE (National Counsel of Epidemiology

Surveillance) of the Ministry of Health of the Government of Mexico, and based on this norm ethical approval is not required. Clinical samples were collected at the “Instituto de Diagnóstico y Referencia Epidemiológicos” (InDRE), and “Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas” (INER), by guidance and regulations of declaration of Helsinki, as part of the early diagnostics scheme in public health laboratories and hospitals in Mexico City (Red Nacional de Laboratorios Estatales de Salud Pública, RNLSP; Instituto Nacional de Enfermedades Respiratorias, INER; and Instituto Mexicano del Seguro Social, IMSS). Based on the Mexican Official NOM-017-SSA2-2012 informed consent from patients was not required. All samples were anonymized before use.

Oro- and/or nasopharyngeal swabs, as well as tracheal aspirates were collected and placed in virus transport medium upon collection. The diagnosis was done using validated protocols for SARS-CoV-2 RNA detection, as approved by InDRE and by the World Health Organization (WHO). Depending on the indications of medical staff, in some cases, a panel of respiratory viruses xTAG RVPv1 (Luminex Molecular Diagnostics, Austin, TX) or only influenza virus detection: H1N1pdm09 (https://www.who.int/csr/resources/publications/swineflu/CDCRealtimeRTPCR_SwineH1Assay-2009_20090430.pdf), H3N2 and Influenza B (https://www.who.int/influenza/gisrs_laboratory/CDC_Laboratory_Support_for_Influenza_Surveillance_Info_Sheet_Aug2017.pdf) were used for additional screening and virus identification. We have included samples obtained from three different types of individuals: deceased, hospitalized (severe condition), or ambulatory patients. The demographic and clinical characteristics of the participants, such as age, sex, and comorbidities, are included in the Additional file 1: Table S1.

Sample processing and whole metagenome sequencing

Clinical samples were extracted in BSL2 or BSL3 laboratories using required biosafety operational standards, including respirators, protective clothing, and head and eye protection. All samples were prepared for RNA extraction as described previously [28]. Briefly, centrifuged and 0.45 μ m filtered supernatants were treated with Turbo DNase and RNase. Nucleic acids were extracted using the PureLink™ Viral RNA/DNA Kit (ThermoFisher), or QIAamp viral RNA mini-kit (Qiagen). Total cDNA was synthesized using the SuperScript III Reverse Transcriptase System (ThermoFisher) and primer A1 (5'-GTTTCCCAGTAG GTCTCN9-3') or primer B1 (5'-GCCGGAGCTCTG CAGATATCN9-3'), both of which contained a degenerated 9-mer sequence at the 3' end. The second strand

was generated by two rounds of synthesis with Sequenase 2.0 (Affymetrix, USB, Ohio, USA), or Klenow fragment polymerase (New England Biolabs), followed by 15 cycles of amplification using Phusion DNA polymerase with primer A2 (5'-GTTTCCCAGTAGGTCTC-3'), or 25 cycles of amplification using Expand High Fidelity DNA polymerase (Roche) and primer B2 (5'-GCCGGAGCTCTGCAGATATC-3'). Primers A2 and B2 contain specific sequences of primers A1 and B1, which were used to prepare cDNA, and therefore amplify only particular products produced by the A1 and B1 primers. Purified dsDNA was used as input to generate whole-metagenome sequencing libraries using Nextera XT DNA library preparation kits (Illumina) (https://support.illumina.com/sequencing/sequencing_kits/nextera_xt_dna_kit/documentation.html). Finally, the samples were sequenced on an Illumina NextSeq 500 platform using 2 × 75-cycle or 2 × 150-cycle high-output kits to obtain paired-end reads. Sequencing yields are reported in the Additional file 1: Table S2.

Metagenomic data analysis

A viral metagenomics pipeline, including quality controls, and taxonomic classification was applied as previously described [29]. Briefly, adapters and low-quality bases from 5' and 3' ends were trimmed using fastp v.0.20.0 [30], and low complexity reads and those shorter than 40 bases were removed. Exact duplicate reads were excluded using CD-HIT-DUP v.4.8.1 [31]. Ribosomal RNA and human-derived reads were removed by aligning against ribosomal sequences from SILVA database (DB) [32] and human genomes sequences from GenBank, respectively, using Bowtie2 v.2.3.4.3 [33]. The remaining reads were used for downstream analyses. For the taxonomic classification, valid reads were mapped against a viral reference nt DB (minimally non-redundant nucleotide DB) from NCBI (NCBI Resource Coordinators 2020), using SMALT v.0.7.6 (<https://www.sanger.ac.uk/tool/smalt-0/>) at 70% of identity. Mapped reads were assembled using IDBA v.1.1.3 [34], and contigs and unassembled singleton reads, were compared against all nt DB using BLASTn [35] to remove false positives. Non-mapping reads were assembled using IDBA, and contigs longer than 200 nt were compared to all proteins in NCBI's nr DB (minimally non-redundant protein database) using BLASTx. The top 20 hits were considered from the resulting alignments. Then, MEGAN 6.21.2 [36] was used to taxonomically assign reads and contigs using its last common ancestor algorithm. Finally, taxa (viral species) with less than three assigned reads, seen in less than three samples each, were eliminated to reduce viruses that could be false positives.

Differential virus abundance analysis

To compare groups of patients, read counts were normalized to reads per million (RPM) at 10 million reads to reduce differences due to uneven sequencing depths. Then, viral species that were differentially abundant in the various groups were identified with EdgeR v3.13 [37], using trimmed mean of M-values with singleton pairing (TMMwsp) normalization, which is the preferred method for data with a high proportion of zeros (high data sparsity), followed by dispersion estimation using the CR method with an offset of 0.01. Pairwise comparisons were carried out using a negative binomial GLM with quasi-likelihood tests. Only taxa with differences in associated FDR-adjusted p-values ≤ 0.05 in at least one comparison were kept.

Multivariate group analyses

RPM tables, described above, were also used for multivariate analyses with the R package vegan v2.5-7 [38]. The Bray–Curtis semi-metric was used to calculate weighted distances between each pair of samples. The resulting dissimilarity matrix (DM) was used to assess group variation with a PERMANOVA test (adonis) and a Multivariate homogeneity of groups dispersions test (betadisper). Three-way (ambulatory-hospitalized-deceased) group comparison and pairwise post-hoc adonis and betadisper tests for all pairwise group permutations were performed. The DM was also subjected to multi-dimensional scaling using distance-based redundancy analysis (dbRDA) with vegan's capscale function. Since three groups were present, only two constrained linear combinations were created based on the type of patient.

Assembly of viral genomes

For each viral species with high genomic coverage, the corresponding reference genome sequence was downloaded from GenBank to map all reads using Bowtie2 v2.3.4.3 [33]. Mapped reads were used to perform de novo assembly using SPADes v3.13.0 [39]. In case complete genomes were not obtained, consensus sequences were generated with iVar (v1.3.1) [40], using Phred score $Q > 20$ and a minimum read coverage depth of 5X to call a base of N for lower values. A threshold of at least 55% of the majority base rule was used. Finally, each contig or complete genome was verified using Blast v2.9.1 [35] against its corresponding reference.

Results

Sample collection and generation of SARS-CoV-2 genome sequences

We characterized the virome of SARS-CoV-2 positive samples from 120 patients, categorized into three

different clinical outcome groups; ambulatory (36 samples), hospitalized (26 samples), and deceased (58 samples) (Table 1). Out of these, 41.7% of the participants had at least one comorbidity (Table 1 and Additional file 1: Table S1). As expected, a higher proportion of comorbidities were present in deceased patients (67.2%), with 40.4% having more than one comorbidity. Also, 38.5% of hospitalized patients presented comorbidities (11.5% having more than one), while only 2.8% of persons in the ambulatory group had comorbidity.

The age of patients varied across groups depending on the symptomatology (Additional file 2: Figure S1). Despite extant outliers, deceased patients comprised the most homogeneous set (Median = 36; IQR = 8) compared to hospitalized (Median = 41; IQR = 23.5) and ambulatory patients (Median = 36; IQR = 17). Regardless, no statistically significant difference was observed between any pairwise group permutations (Mann–Whitney; $\alpha = 0.05$).

The presence of other respiratory pathogens was tested by RT-PCR assay in 19 samples for diagnostic purposes before metagenomic sequencing. No sample was found positive for any respiratory virus other than SARS-CoV-2 (Additional file 1: Table S1). Randomly amplified libraries for high-throughput sequencing were successfully prepared from all samples. In total, we obtained 3,486,527,880 paired-end reads, ranging from 3.6×10^6 up to 77×10^6 paired-end reads per sample (Additional file 1: Table S2). Out of 120 samples used, we obtained 70 complete SARS-CoV-2 genome sequences (>98% genome coverage). In addition, 15 samples had a genome coverage of >50%, nine samples had a coverage of 10 to 50%, and we could obtain only isolated SARS-CoV-2 reads from the remaining 26 samples (less than 10% coverage) (Additional file 1: Table S2). Whole genome sequences of SARS-CoV-2 generated in this work are described in more detail

elsewhere and will not be discussed further in this work [41].

Respiratory and other human viruses

Out of 120 samples analysed, 13 human respiratory viruses other than SARS-CoV-2 were found in 30 samples (25%). Five of these samples were from ambulatory patients, seven from hospitalized patients, and 18 samples from deceased patients (Additional file 1: Table S3). Nine samples (7.5%) harbored two additional respiratory viruses other than SARS-CoV-2; two were from ambulatory cases, one from a severe disease patient, and the six remaining samples were from deceased persons (Additional file 1: Table S3). The most common non-SARS-CoV-2 respiratory virus found in the samples was *Human mastadenovirus C*, which was found in 20 samples (16.7%) (Table 2, Additional file 1: Table S3), albeit always in low abundance (mostly < 100 reads per sample); *Human coronavirus HKU1* and *Rhinovirus B* were found in four samples; *Influenza B virus* was identified in two samples, and the remaining viruses were single occurrences (Table 2). None of the samples tested for other respiratory viruses by RT-PCR contained sequence reads similar to the viral species tested.

Interestingly, reads corresponding to at least one human non-respiratory virus were found in 97 samples (80.8%) (Additional file 1: Table S4). They were present in 31 out of 36 ambulatory cases (86.1%), 20 of 26 hospitalized patients (76.9%), and 46 out of 56 deceased patients (82.1%). In total, we have identified 27 additional human non-respiratory viruses belonging to 10 viral families, with four viruses being unclassified at the family level. Among the most prevalent were members of the *Anneloviridae* family, identified in 37% of the samples, with seven viral species, followed by viruses in the family *Herpesviridae*, found in 41.7% of the samples, with five species, and viruses in the *Papillomaviridae* family, in 34.2% of the samples, with four species (Fig. 1; Additional file 1: Table S4). Other reads belonging to viruses in families *Picobirnaviridae* (25% of prevalence), *Reoviridae* (10.8%), *Circoviridae* (8.3%), *Paramyxoviridae* (6.7%), *Astroviridae* (3.3%), *Flaviviridae* (2.5%) and unclassified viral sequences (11.7%), were also identified in some samples. One sample (number 383 from a deceased patient without comorbidities) did not present any other viral reads than SARS-CoV-2 and was excluded from further analyses.

Apart from viruses known to infect humans, we identified many sequences that showed similarity to viruses that infect other organisms, including animal species, plants, invertebrates, yeast, and amoebas (Additional file 1: Tables S5 and S6). These results are not discussed

Table 1 Categories of SARS-2 positive samples used in this study

	All samples	Deceased	Severe	Ambulatory
Total	120 ^a	58	26	36
No comorbidities	70	19	16	35
Comorbidities ^b	50 (41.7%)	39 (67.2%)	10 (38.5%)	1 (2.8%)
1 comorbidity	24 (20%)	16 (27.6%)	7 (27%)	1 (2.8%)
≥ 2 comorbidities	26 (21.7%)	23 (40.4%)	3 (11.5%)	0

^a Number of samples in each category

^b samples from patients with comorbidities were divided into two groups; with 1 comorbidity, and with 2 or more comorbidities

Conditions considered comorbidity: Obesity BMI index > 30, smoker, high blood pressure, diabetes, immunosuppression, HIV, COPD, kidney insufficiency, asthma, chronic disease

Table 2 Presence of human viruses in SARS-2 positive samples

Virus species	Number of positive samples	severity of infection ^b									
Human respiratory viruses^a											
Coronavirus HKU1	4	■	■	■	■						
Coronavirus NL63	1	■									
Enterovirus B	1	■									
Enterovirus C	1	■									
Human bocavirus	1	■									
Influenza A	2	■	■								
Influenza B	2	■	■								
Mastadenovirus C	20	■	■	■	■	■	■	■	■	■	■
Mastadenovirus F	1	■									
Metapneumovirus	1	■									
Rhinovirus A	1	■									
Rhinovirus B	4	■	■	■	■						
Rhinovirus C	1	■									
Other human viruses											
Alphapapillomavirus	7 ^c	■	■	■	■	■	■	■	■	■	■
Anelloviridae sp.	34	■	■	■	■	■	■	■	■	■	■
Aroa virus	3	■	■	■							
Betapapillomavirus	13	■	■	■	■	■	■	■	■	■	■
Gammapapillomavirus	31	■	■	■	■	■	■	■	■	■	■
Human associated cyclovirus 10	7	■	■	■	■	■	■	■	■	■	■
Human betaherpesvirus 5	3	■	■	■							
Human betaherpesvirus 6	18	■	■	■	■	■	■	■	■	■	■
Human betaherpesvirus 7	19	■	■	■	■	■	■	■	■	■	■
Human gammaherpesvirus 4	24	■	■	■	■	■	■	■	■	■	■
Human lung-associated virovirus JB	8	■	■	■	■	■	■	■	■	■	■
Human oral-associated virovirus XM	10	■	■	■	■	■	■	■	■	■	■
Human respiratory circular DNA virus	3	■	■	■							
Lysoka partiti-like virus	30	■	■	■	■	■	■	■	■	■	■
Mamastrovirus 1	4	■	■	■	■						
Measles morbillivirus	8	■	■	■	■	■	■	■	■	■	■
Rotavirus A	14	■	■	■	■	■	■	■	■	■	■
SEN virus	9	■	■	■	■	■	■	■	■	■	■
Small anellovirus	4	■	■	■	■						
TTV-like mini virus	9	■	■	■	■	■	■	■	■	■	■
Torque teno midi virus 8	10	■	■	■	■	■	■	■	■	■	■
Torque teno mini virus SHA	8	■	■	■	■	■	■	■	■	■	■
Torque teno virus	26	■	■	■	■	■	■	■	■	■	■

^a Human viruses known to cause respiratory illness

^b Severity of infection: red squares, deceased patients; orange, hospitalized patients; green, ambulatory cases

^c Only viruses present in minimum of three samples are shown

further in this work and will be addressed in future studies.

Some virus species are associated with different severity of disease groups

We have observed that some viral species were present predominantly in some of the severity groups analysed in this study. Table 3 shows the viral species that were differentially distributed among paired study groups (pairwise comparisons of samples from ambulatory, hospitalized, and deceased patients). Eight viral species were significantly more prevalent in deceased patients compared to ambulatory ones: *Torque teno midi virus 8* (TTMDV 8), *TTV-like mini virus 19* (TTMV 19), *Torque teno mini virus SHA* (TTMV 26), *Human betaherpesvirus 6* (HHV 6), *Human associated acyclovir 10* (HuACyV 10), *Rotavirus A* (RV-A), *Measles morbillivirus* (MV) and *Alphapapillomavirus 10* (HPV-6). Likewise, five of these species were also more prevalent in hospitalized than in ambulatory patients (TTMDV 8, TTMV 19, TTMV 26, HHV 6, and HuACyV 10). Interestingly, there was also different

distribution of some of these human viral species when samples from deceased patients were compared to samples from hospitalized patients, with three being more prevalent (RV-A, MV, and HPV 6). All other group comparisons showed no statistically significant differences for other viruses.

To further evaluate the differences in the composition of viral communities in samples from different groups of patients, Bray–Curtis dissimilarities distances were calculated using the relative abundance of human viruses, and statistical comparisons were carried out by attribute and in pairwise post hoc analyses (Fig. 2 shows the resulting dBRDA from the human virus matrix). These results indicated that differences between ambulatory patients and the other groups were significant (PERMANOVA for ambulatory-deceased p-value=0.03, ambulatory-hospitalized p-value=0.04), but the difference between the hospitalized and deceased patients was not statistically significant (p-value=0.49). We also corroborated that the variances between groups were comparable (betadisper p-value=0.2827). Furthermore, other features, such

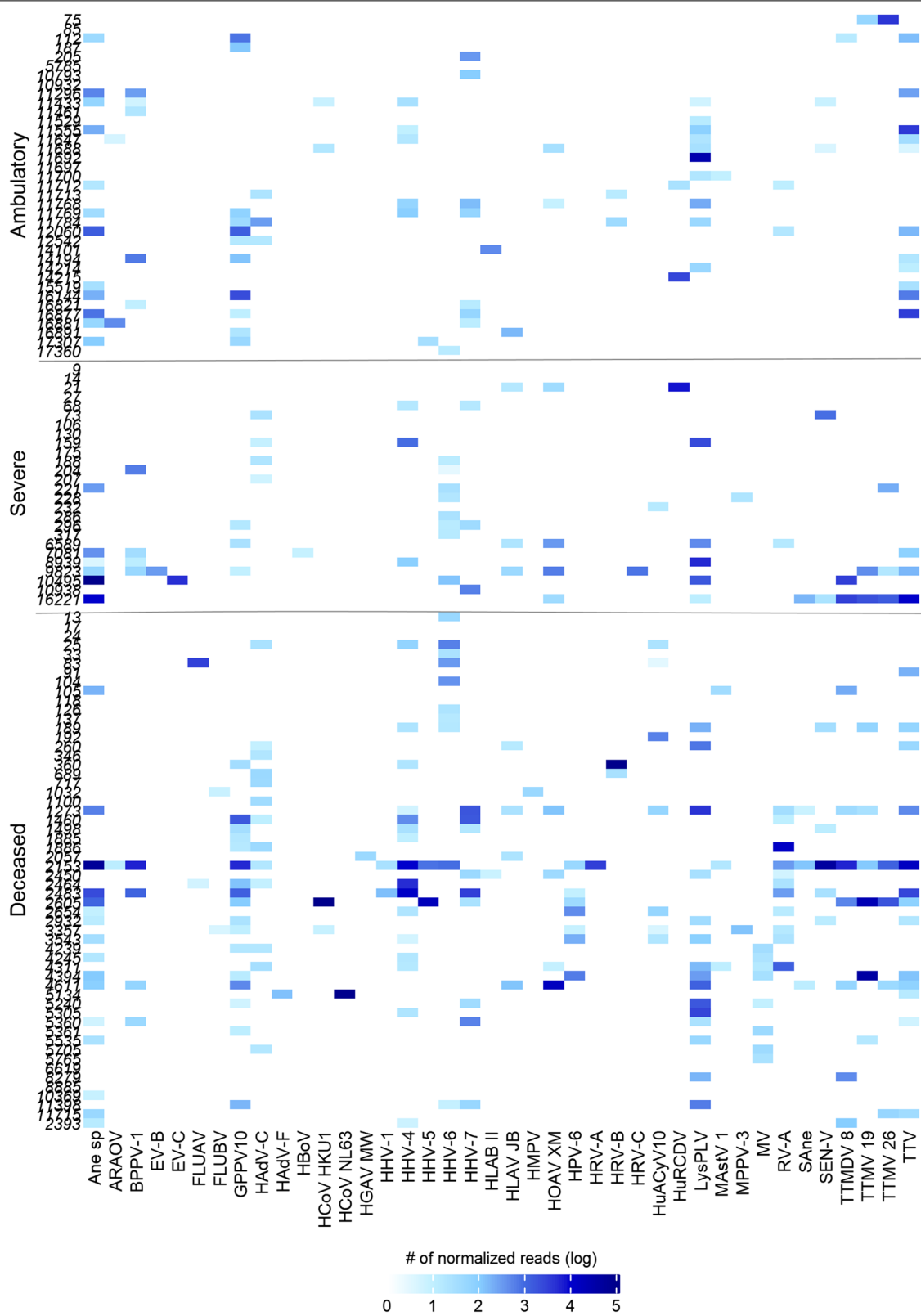


Fig. 1 Abundance of human viral species in SARS-CoV-2 positive samples. Sequence reads were normalized according to the total number of reads after quality filtering. The abundance is shown in logarithmic scale (log10). Samples from ambulatory, severe and deceased patients are separated vertically

Table 3 Virus species differentially abundant in patients with diverse severity of infection

Viral species	All samples	Number of positive samples			Comparing dec/amb		Comparing hosp/amb		Comparing dec/hosp	
		Deceased	Hospitalized	Ambulatory						
		57	26	37	lfch ^a	FDR ^b	lfch	FDR	lfch	FDR
TTMDV 8	10	6	3	1	4.201^c	0.025	4.279	0.021	0.078	0.97
TTMV 19	9	6	2	1	16.118	3.04E-07	9.947	0.0004	− 6.171	0.093
TTMV 26	8	4	3	1	4.815	0.003	4.298	0.021	− 0.517	0.944
HHV-6	19	12	7	0	9.986	1.98E-06	11.361	1.24E-06	1.374	0.595
HuACyV10	8	7	1	0	12.793	1.07E-06	13.305	1.81E-06	0.511	0.944
RVA	14	11	1	2	4.516	0.013	− 1.925	0.395	− 6.442	0.016
MV	8	8	0	0	14.037	4.90E-07	0	1	− 14.037	0.0005
HPV-6	7	7	0	0	14.851	4.90E-07	0	1	− 14.851	0.00003

^a log fold change

^b FDR represents the p-value adjusted by FDR using Benjamini–Hochberg correction

^c Significant results are in bold

as comorbidities, sex, and sampling location, were compared between groups, but no significant differences were observed (results not shown).

Generating complete genomes of other viruses

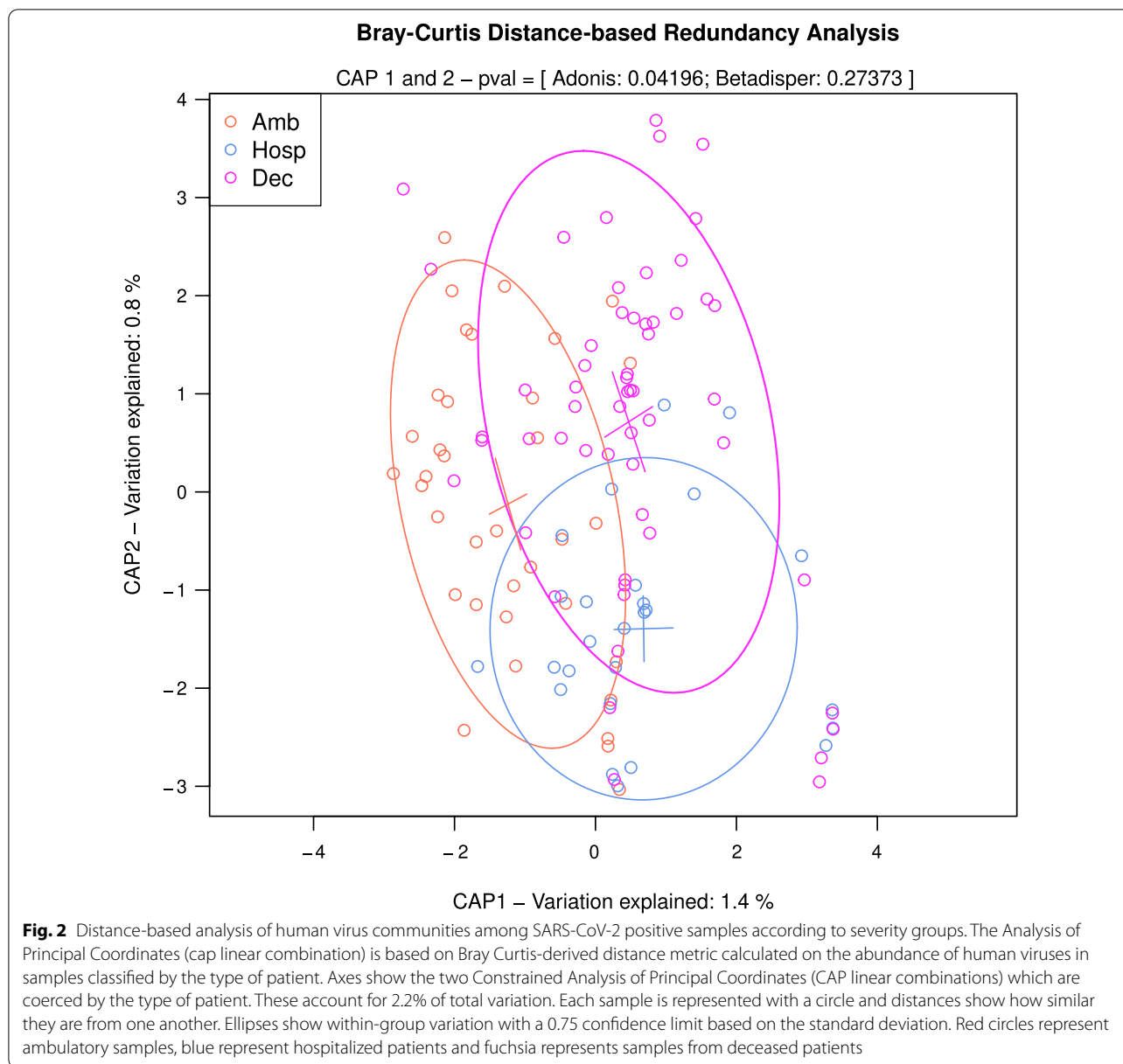
Our strategy allowed us to obtain enough viral reads in some of the samples to assemble complete or partial genomes of other viruses, in addition to SARS-CoV-2. We could assemble at least 25% of the genome for 70 viruses other than SARS-CoV-2 belonging to 14 different families, with three viruses being unclassified (listed in Additional file 1: Table S7). The most common family with complete genomes was *Anelloviridae*, found in 16 samples, and out of 28 anelloviruses, seven had a genome coverage >90%. Interestingly, we assembled the complete genome of two other coronavirus species (*Human coronavirus HKU1* and *Human coronavirus NL63* in one sample each) and one other respiratory virus (Human rhinovirus B3 in one sample). The remaining viruses with >25% genome coverage belong to different groups; human viruses (diverse papillomaviruses, human-associated cyclovirus, gut, and oral-associated videntovirus, circular DNA viruses, and enterovirus B), and diverse plant viruses.

Discussion

Data regarding SARS-CoV-2 coinfections with other viruses and the role these viruses could have in the severity of the disease are still limited. Most of the previous studies have used RT-qPCR assays focused on detecting a few specific respiratory viruses [11, 12, 14–18, 27], and only some studies have used metagenomic approaches [21–24, 42]. In this study, we used sequence-independent metagenomic sequencing to determine the whole virome

in three different severity groups of patients (deceased, hospitalized, and ambulatory) infected with SARS-CoV-2 and to explore whether there could be a correlation between the presence of viruses other than SARS-CoV-2, and severity of the disease. Despite processing samples to decrease the amount of contaminant genetic material, all samples still contained human reads. This is standard in viromic studies, suggesting that filtration and nuclease treatment were not absolute.

We found that 25% of the 120 samples analysed were positive for another respiratory virus than SARS-CoV-2, with 7.5% having two additional respiratory viruses. This high level of co-occurrence has been previously observed in other studies: 20.7% in North California [16], 28.4% in patients from Jiangsu Province, China [19], and 45% in patients in Iran [12]. Previously, rhinoviruses/enteroviruses have been found as the predominant coinfecting agents [16, 20, 27], with influenza being frequent in one study [12]. In this study, we identified Human mastadenovirus C as the most prevalent, being present in 20 samples (16.7%), although with low abundance (<100 reads per sample). This high level of co-occurrence with Human mastadenovirus C was not expected, as it is not recognized as a frequent pathogen in respiratory diseases, generally being found in less than 10% of samples of children under five years of age, without a clear seasonality, and instead all year around [43, 44]. Human mastadenovirus C had been reported in lower prevalence (4.8%) in SARS-CoV-2 positive samples in Brazil [45]. Regarding other respiratory viruses, similar to most other published works, [11, 16, 18, 27, 45], we observed Influenza A or B viruses in only a few samples. These low co-occurrence values in our samples could be explained by the time of sample collection, as the influenza season



in Mexico generally ends in March–April, and due to the anti-SARS-CoV-2 preventive measures implemented (wearing masks, social distancing) and the anti-influenza vaccination campaign in Mexico. We found *Human coronavirus HKU1* and *Rhinovirus B* in four samples each (3.3%), and other respiratory viruses were found in only one or two samples.

The presence of other human non-respiratory viruses in SARS-CoV-2 positive samples has been even less studied during SARS-CoV-2 infections. A study by Kim and collaborators [22] reported a high prevalence of other non-respiratory viruses in southern hemisphere samples (74%, 68 out of 92 samples), while another metagenomic

next-generation sequencing (mNGS) work did not present a detailed virome composition [21, 23, 24]. Our study identified other (non-respiratory) human viruses in 97 out of 120 samples (80.8%). We found sequences with homology to 27 different viruses (Additional file 1, Table S4). The viruses most frequently detected in our study differ from those identified by Kim et al. [22]. They reported mammarenaviruses, rodeoloviruses, and alphapolyomaviruses as the most frequent, followed by papillomaviruses and lymphocryptoviruses. Recent virome analysis of samples collected in Italy has identified six viral families in SARS-CoV-2 samples (*Retroviridae*, *Herpesviridae*, *Poxviridae*, *Pneumoviridae*,

Anelloviridae, and *Pandoraviridae*), with the first three families being most prevalent [46]. In our study, the most common families found were *Anelloviridae*, with seven members, *Herpesviridae* with five species, and *Papillomaviridae*, with four viral species. Differences observed may be due to geographic locations and differences in the populations studied. Regarding the evolution of the patients' disease, significant differences were found when the abundance of human viruses was compared between types of patients globally, as well as between ambulatory and hospitalized patients, and between ambulatory and deceased patients; no statistical significance was found in differences between hospitalized and deceased patients. When observing the virus species that could contribute to these differences, some were preferentially abundant in samples from deceased or/and hospitalized patients (Table 3); however, given the type of viruses identified and fact that they were only present in a small proportion of all patients, their association with the severity of the infection cannot be established, but rather, those viruses that were present, showed higher sample frequency and overall abundance in hospitalized or deceased patients.

For example, in this study, some Anelloviruses were identified in high frequency, principally in hospitalized and deceased patients. These are ubiquitous viruses that persistently and commonly infect humans [47], which could explain their frequent presence in samples. Alternatively, their presence could reflect an increased replication due to alterations of the immune responses during SARS-CoV-2 infection, as described for other infections [48, 49], and increased coinfection in severe cases of disease would be result of SARS-CoV-2 infection. In the case of measles virus, which in this work was identified abundantly and frequently in deceased patients, it has been described that its infection could cause immune suppression [50], which could lead to more severe disease. However, the impact of other coinfecting viruses on the COVID-19 severity remains to be explored. Recently, Paparoupa and collaborators [13] described an increase in coinfection of SARS-CoV-2 positive patients on invasive ventilation with *Herpes simplex virus* and *Cytomegalovirus*, possibly as a result of long treatment, suggesting that time spent in the hospital could have an effect on the viral presence, which is consistent with HHV6 and HuACyV 10 viruses being identified in higher frequency and abundance in deceased patients which had been hospitalized.

When analysing the presence of other respiratory viruses in deceased patients from North Khorasan (Iran), a high level of coinfection between SARS-CoV-2 and Influenza A virus (22.3% of the samples) was observed [12]. Furthermore, Respiratory syncytial virus (RSV) and bocavirus were each present in 9.7% of the

samples. Although that work focused only on samples from deceased patients, without comparing the results with other populations, for example, ambulatory patients, it is tempting to speculate that coinfection with *Influenza A virus* and RSV (two respiratory pathogens that are the cause of severe disease) could play a role in the high mortality reported. Coinfection with Influenza and adenoviruses has also been reported to be significantly associated with increased mechanical ventilation (influenza viruses) or death (both viruses) [51]. Moreover, two recent studies in mouse and ferret models demonstrated that SARS-CoV-2 and *Influenza A virus* coinfection causes more severe pathology [52, 53]. In our study, these coinfections were rare for influenza viruses and null for RSV. Concerning adenoviruses homologs to *Mastadenovirus C* and *F* were frequent but could not be associated with pathogenicity. It is of interest that, in our case whole genome of three different respiratory viruses (human rhinovirus B and human coronaviruses HKU1 and NL63) were obtained from samples corresponding to deceased patients without comorbidities. Viral metagenomics approaches are considered semi-quantitative, and the ability to assemble whole viral genomes suggests a high load of a given virus in the sample, most probably due to an active infection rather than the presence of a virus that is just passing through. However, it is not known whether these coinfections had an impact on disease severity.

Our study has several limitations. Firstly, our samples were from patients with and without different comorbidities (Additional file 1: Table S1), making it difficult to draw conclusions concerning the role of other viruses on the disease outcome. Furthermore, samples were derived from nasopharyngeal or oropharyngeal swabs or from tracheal aspirates, and the sample type may affect virus detection [54, 55]. Finally, all the samples analysed were from SARS-CoV-2 infected subjects, and a group of SARS-CoV-2 negative patients was not included. These inconsistencies among the samples (presence of comorbidities and different types of samples), together with the expected diversity of the inter-individual respiratory virome [56], make it difficult to observe differences among viral communities in samples from different study groups (deceased, severe-hospitalized, and ambulatory cases). Despite these limitations, we were able to observe differences in the abundance of human viral species when samples from ambulatory, hospitalized, and deceased individuals were compared, being higher in severe and deceased persons as compared to ambulatory patients.

Abbreviations

SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; COVID-19: Coronavirus disease 2019; RT-PCR: Reverse transcription polymerase chain reaction; InDRE: Instituto de Diagnóstico y Referencia Epidemiológicos; INER:

Instituto Nacional de Enfermedades Respiratorias; RNLSP: Red Nacional de Laboratorios Estatales de Salud Pública; IMSS: Instituto Mexicano del Seguro Social; WHO: World Health Organization; RNA: Ribonucleic acid; cDNA: Complimentary deoxyribonucleic acid; DB: Data bank; NCBI: National Center for Biotechnology Information; RPM: Reads per million; DM: Dissimilarity matrix; dbRDA: Distance-based redundancy analysis; PCoA/NMDS: Principal coordinates analysis and multidimensional scaling; TTMDV 8: Torque teno midi virus 8; TTMV 19: TTV-like mini virus 19; TTMV 26: Torque teno mini virus SHA; HHV 6: Human betaherpesvirus 6; HuAcV10: Human associated cyclovirus 10; RV-A: Rotavirus A; MV: Measles morbillivirus; HPV-6: Alphapapillomavirus 10; RT-qPCR: Reverse transcription quantitative polymerase chain reaction; mNGS: Metagenomic next-generation sequencing; RSV: Respiratory syncytial virus.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12879-022-07783-8>.

Additional file 1: Table S1. Patients included in the study. Information of samples included in study, sample ID, age, and sex of patients, place of collection, date of sample and start of the symptoms, clinical status, comorbidities, sample type, and diagnostic test for other respiratory viruses if performed. **Table S2.** General sequencing data and SARS-CoV-2 sequenced genome characteristics. Number of reads (original total reads, filtered reads, SARS-CoV-2-specific reads, reads used for metagenomic analysis) and results of SARS-CoV-2 genome sequencing. **Table S3.** Respiratory viruses found in SARS-CoV-2 positive samples. List of samples with homologs to other (non-SARS-CoV-2) human respiratory viruses. **Table S4.** Other human (non-respiratory) viruses, found in SARS-CoV-2 positive samples. List of samples with homologs to other non-respiratory human viruses. **Table S5.** Viruses of other vertebrates found in SARS-CoV-2 infected patients. Contains list of samples with homologs to viruses of other vertebrate species found in SARS-CoV-2 infected patients. **Table S6.** Non-vertebrate viruses found in SARS-CoV-2 infected persons. Contains list of samples with homologs to non-vertebrate viruses. **Table S7.** Non-SARS-CoV-2 viruses with > 25% genome coverage. Contains list of viruses, where more than 25% genome was sequenced, from those found in SARS-CoV-2 positive samples.

Additional file 2: Figure S1. Age distribution in different outcome groups. Boxplot shows age quartiles for each group of patients (deceased, hospitalized and ambulatory). Boxes are delimited by Q_1 and Q_3 and whiskers show IQR*1.5 from either Q_1 or Q_3 . Medians (Q_2) are shown as black horizontal bars inside boxes.

Acknowledgements

We thank Jerome Verleyen and Juan Manuel Hurtado for their computer support. We also thank the Dirección General de Cómputo y de Tecnologías de la Información (DG TIC-UNAM) for providing Supercomputing resources on MIZTLI through the projects LANCAD-UNAM-DGTIC-350 and LANCAD-UNAM-DGTIC-396.

Author contributions

PI conceptualization, methodology, formal analyses, writing original-draft, manuscript revision, BT conceptualization, data curation, data analyses, methodology, manuscript revision, RGL formal analyses, data curation, data analyses, manuscript revision, CB conceptualization, methodology, formal analyses, data curation, manuscript revision, JERG conceptualization, coordination sample acquisition, manuscript revision, JAVP formal analyses, sample acquisition, data curation, funding, manuscript revision, AHT methodology, data curation, manuscript revision, JARE methodology, manuscript revision, JEMM methodology, resources, manuscript revision, CGM—methodology, resources, manuscript revision, ARR methodology, data curation, manuscript revision, MMF methodology, data curation, manuscript revision, ASF—methodology, supervision, manuscript review, EMC—resources, methodology, manuscript review, GBB resources, sample coordination, SL conceptualization, manuscript review, LHR resources, ILM—resources, supervision, manuscript review, SAR methodology, formal analyses, data curation, manuscript revision, CFA conceptualization, funding acquisition, project administration, manuscript review. All authors read and approved the final manuscript.

Funding

This work was partially supported by grants 'Epidemiología Genómica de los Virus SARS-CoV-2 Circulantes en México' (to CFA) from CONACYT, grant 057 from the 'Ministry of Education, Science, Technology and Innovation (SECTEI) of Mexico City' (to CFA), and CONACYT FORDECYT 2020, "Caracterización de la diversidad viral y bacteriana presenta en la enfermedad respiratoria aguda" (to JAVP). RGL is the recipient of a 2021 postdoctoral fellowship from CONACYT, linked to ProNacEs #11000/023/2021; C-08/2021.

Availability of data and materials

Sequences are readily available at NCBI's SRA under project PRJNA819439 with accession numbers: SRR19400449-SRR19400568. Available at: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA819439/>.

Declarations

Ethics approval and consent to participate

Samples were collected and processed as part of the national response to COVID-19 under the Mexican Official NOM-017-SSA2-2012 for epidemiological surveillance of Viral Respiratory Disease emitted and approved by the CONAVE (National Counsel of Epidemiology Surveillance) of the Ministry of Health of the Government of Mexico. Samples were collected at the "Instituto de Diagnóstico y Referencia Epidemiológicos" (InDRE) and "Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas" (INER), by guidance and regulations of declaration of Helsinki, as part of the early diagnostics scheme in public health laboratories and hospitals in Mexico City (Red Nacional de Laboratorios Estatales de Salud Pública, RNLSP; Instituto Nacional de Enfermedades Respiratorias, INER; and Instituto Mexicano del Seguro Social, IMSS). Based on Mexican Official NOM-017-SSA2-2012 Institutional Review Board (IRB), approval is not required. Nor informed consent from patients was required. All samples were anonymized before use.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico. ²Centro de Investigación en Enfermedades Infecciosas, Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Mexico City, Mexico. ³Instituto de Diagnóstico y Referencia Epidemiológicos, Dirección General de Epidemiología, Ciudad de México, Mexico. ⁴Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Mexico City, Mexico. ⁵Coordinación de Calidad de Insumos y Laboratorios Especializados, Instituto Mexicano del Seguro Social, Mexico City, Mexico. ⁶Unidad Universitaria de Secuenciación Masiva y Bioinformática, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico.

Received: 6 May 2022 Accepted: 6 October 2022

Published online: 19 October 2022

References

- Li X, Xu S, Yu M, Wang K, Tao Y, Zhou Y, Shi J, Zhou M, Wu B, Yang Z, et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J Allergy Clin Immunol*. 2020;146(1):110–8.
- Fernald LC, Gutierrez JP, Neufeld LM, Olaiz G, Bertozzi SM, Miettus-Snyder M, Gertler PJ. High prevalence of obesity among the poor in Mexico. *JAMA*. 2004;291(21):2544–5.
- Kwok S, Adam S, Ho JH, Iqbal Z, Turkington P, Razvi S, Le Roux CW, Soran H, Syed AA. Obesity: a critical risk factor in the COVID-19 pandemic. *Clin Obes*. 2020;10(6): e12403.
- Simonnet A, Chetboun M, Poissy J, Raverdy V, Noulette J, Duhamel A, Labreuche J, Mathieu D, Pattou F, Jourdain M, et al. High prevalence of obesity in severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2)

- requiring invasive mechanical ventilation. *Obesity* (Silver Spring). 2020;28(7):1195–9.
5. Greaney AJ, Loes AN, Crawford KHD, Starr TN, Malone KD, Chu HY, Bloom JD. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe*. 2021;29(3):463–76.
 6. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Hengartner N, Giorgi EE, Bhattacharya T, Foley B, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*. 2020;182(4):812–27.
 7. Anastassopoulou C, Gkizarioti Z, Patrinos GP, Tsakris A. Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity. *Hum Genomics*. 2020;14(1):40.
 8. Hernandez-Teran A, Mejia-Nepomuceno F, Herrera MT, Barreto O, Garcia E, Castillejos M, Boukadida C, Matias-Florentino M, Rincon-Rubio A, Avila-Rios S, et al. Dysbiosis and structural disruption of the respiratory microbiota in COVID-19 patients with severe and fatal outcomes. *Sci Rep*. 2021;11(1):21297.
 9. Bello-Chavolla OY, Rojas-Martinez R, Aguilar-Salinas CA, Hernandez-Avila M. Epidemiology of diabetes mellitus in Mexico. *Nutr Rev*. 2017;75(suppl 1):4–12.
 10. Lin D, Liu L, Zhang M, Hu Y, Yang Q, Guo J, Guo Y, Dai Y, Xu Y, Cai Y, et al. Co-infections of SARS-CoV-2 with multiple common respiratory pathogens in infected patients. *Sci China Life Sci*. 2020;63(4):606–9.
 11. Ma L, Wang W, Le Grange JM, Wang X, Du S, Li C, Wei J, Zhang JN. Coinfection of SARS-CoV-2 and other respiratory pathogens. *Infect Drug Resist*. 2020;13:3045–53.
 12. Hashemi SA, Safamanesh S, Ghasemzadeh-Moghaddam H, Ghafouri M, Azimian A. High prevalence of SARS-CoV-2 and influenza A virus (H1N1) coinfection in dead patients in Northeastern Iran. *J Med Virol*. 2021;93(2):1008–12.
 13. Paparoupa M, Aldemyati R, Roggenkamp H, Berinson B, Norz D, Olearo F, Kluge S, Roedl K, de Heer G, Wichmann D. The prevalence of early- and late-onset bacterial, viral, and fungal respiratory superinfections in invasively ventilated COVID-19 patients. *J Med Virol*. 2022;94(5):1920–5.
 14. Calderaro A, De Conto F, Buttrini M, Piccolo G, Montecchini S, Maccari C, Martinelli M, Di Maio A, Ferraglia F, Pinardi F, et al. Human respiratory viruses, including SARS-CoV-2, circulating in the winter season 2019–2020 in Parma, Northern Italy. *Int J Infect Dis*. 2021;102:79–84.
 15. Hazra A, Collison M, Pisano J, Kumar M, Oehler C, Ridgway JP. Coinfections with SARS-CoV-2 and other respiratory pathogens. *Infect Control Hosp Epidemiol*. 2020;41(10):1228–9.
 16. Kim D, Quinn J, Pinsky B, Shah NH, Brown I. Rates of co-infection between SARS-CoV-2 and other respiratory pathogens. *JAMA*. 2020;323(20):2085–6.
 17. Nowak MD, Sordillo EM, Gitman MR, Paniz Mondolfi AE. Coinfection in SARS-CoV-2 infected patients: where are influenza virus and rhinovirus/enterovirus? *J Med Virol*. 2020;92(10):1699–700.
 18. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, Barnaby DP, Becker LB, Chelico JD, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA*. 2020;323(20):2052–9.
 19. Zhu X, Ge Y, Wu T, Zhao K, Chen Y, Wu B, Zhu F, Zhu B, Cui L. Co-infection with respiratory pathogens among COVID-2019 cases. *Virus Res*. 2020;285: 198005.
 20. Cooksey GLS, Morales C, Linde L, Schildhauer S, Guevara H, Chan E, Gibb K, Wong J, Lin W, Bonin BJ, et al. Severe acute respiratory syndrome coronavirus 2 and respiratory virus sentinel surveillance, California, USA, May 10, 2020–June 12, 2021. *Emerg Infect Dis*. 2022;28(1):9–19.
 21. Babiker A, Bradley HL, Stittlerburg VD, Ingersoll JM, Key A, Kraft CS, Waggoner JJ, Piantadosi A. Metagenomic sequencing to detect respiratory viruses in persons under investigation for COVID-19. *J Clin Microbiol*. 2020; 59(1).
 22. Kim KW, Deveson IW, Pang CNI, Yeang M, Naing Z, Adikari T, Hammond JM, Stevanovski I, Beukers AG, Verich A, et al. Respiratory viral co-infections among SARS-CoV-2 cases confirmed by virome capture sequencing. *Sci Rep*. 2021;11(1):3934.
 23. Mostafa HH, Fissel JA, Fanelli B, Bergman Y, Gniazdowski V, Dadlani M, Carroll KC, Colwell RR, Simmer PJ. Metagenomic next-generation sequencing of nasopharyngeal specimens collected from confirmed and suspect COVID-19 patients. *mBio* 2020; 11(6).
 24. Shah SJ, Barish PN, Prasad PA, Kistler A, Neff N, Kamm J, Li LM, Chiu CY, Babik JM, Fang MC, et al. Clinical features, diagnostics, and outcomes of patients presenting with acute respiratory illness: a retrospective cohort study of patients with and without COVID-19. *EClinicalMedicine*. 2020;27: 100518.
 25. Zhou H, Li C, Hu T, Liu T, Ni N, Chen W, Zhao H, Ruan S, Li J, Wu H, et al. Total infectoromes of 162 SARS-CoV-2 cases using meta-transcriptomic sequencing. *J Infect*. 2021;82(1):e44–8.
 26. Dao TL, Hoang VT, Colson P, Million M, Gautret P. Co-infection of SARS-CoV-2 and influenza viruses: a systemic review and meta-analysis. *J Clin Virol Plus*. 2021;1:100036.
 27. Scott SJ, Pfothenauer B, Weiner JJ, Hilleshiem J, Khubbar M, Bhattacharyya S. Respiratory pathogen coinfections in SARS-CoV-2-positive patients in Southeastern Wisconsin: a retrospective analysis. *Microbiol Spectr*. 2021;9(2): e0083121.
 28. Taboada B, Vazquez-Perez JA, Munoz-Medina JE, Ramos-Cervantes P, Escalera-Zamudio M, Boukadida C, Sanchez-Flores A, Isa P, Mendieta-Condado E, Martinez-Orozco JA et al. Genomic analysis of early SARS-CoV-2 variants introduced in Mexico. *J Virol* 2020; 94(18).
 29. Taboada B, Isa P, Gutierrez-Escolano AL, Del Angel RM, Ludert JE, Vazquez N, Tapia-Palacios MA, Chavez P, Garrido E, Espinosa AC et al. The Geographic Structure of Viruses in the Cuatro Ciénegas Basin, a Unique Oasis in Northern Mexico, Reveals a Highly Diverse Population on a Small Geographic Scale. *Appl Environ Microbiol* 2018; 84(11).
 30. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884–90.
 31. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*. 2012;28(23):3150–2.
 32. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41:D590–596.
 33. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
 34. Peng Y, Leung HC, Yiu SM, Chin FY. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*. 2012;28(11):1420–8.
 35. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:421.
 36. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. *Genome Res*. 2007;17(3):377–86.
 37. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multi-factor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res*. 2012;40(10):4288–97.
 38. Oksanen J, Guillaume Blanchet F, Friendly M, Kind R, Legendre G, McGlenn D, Minchin PR, Brien JA, B. OHR, Simpson GL et al. Cpmminity Ecology Package. In., 2.5–7 edn; 2020: Ordination methods, diversity analysis and other functions for community and vegetation ecologists.
 39. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455–77.
 40. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, Tan AL, Paul LM, Brackney DE, Grewal S, et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol*. 2019;20(1):8.
 41. Boukadida C, Taboada B, Escalera-Zamudio M, Isa P, Ramirez-Gonzalez JE, Vazquez-Perez JA, Munoz-Medina JE, Grajalles-Muniz C, Gonzalez-Torres C, Gaytan-Cervantes FJ, et al. Genomic characterization of SARS-CoV-2 isolated from patients with distinct disease outcomes in Mexico. *Microbiol Spectr*. 2022;10(1): e0124921.
 42. Peddu V, Shean RC, Xie H, Shrestha L, Perchetti GA, Minot SS, Roychoudhury P, Huang ML, Nalla A, Reddy SB, et al. Metagenomic analysis reveals clinical SARS-CoV-2 infection and bacterial or viral superinfection and colonization. *Clin Chem*. 2020;66(7):966–72.
 43. Wong-Chew RM, Espinoza MA, Taboada B, Aponte FE, Arias-Ortiz MA, Monge-Martinez J, Rodriguez-Vazquez R, Diaz-Hernandez F, Zarate-Vidal F, Santos-Preciado JI, et al. Prevalence of respiratory virus in symptomatic children in private physician office settings in five communities of the state of Veracruz, Mexico. *BMC Res Notes*. 2015;8:261.

44. Wong-Chew RM, Garcia-Leon ML, Noyola DE, Perez Gonzalez LF, Gaitan Meza J, Vilasenor-Sierra A, Martinez-Aguilar G, Rivera-Nunez VH, Newton-Sanchez OA, Firo-Reyes V, et al. Respiratory viruses detected in Mexican children younger than 5 years old with community-acquired pneumonia: a national multicenter study. *Int J Infect Dis*. 2017;62:32–8.
45. Eisen AKA, Gualarte JS, Demoliner M, de AbreuGoes Pereira VM, Heldt FH, Filippi M, de Almeida PR, Hansen AW, Fleck JD, Spilki FR. Low circulation of influenza A and coinfection with SARS-CoV-2 among other respiratory viruses during the COVID-19 pandemic in a region of southern Brazil. *J Med Virol*. 2021;93(7):4392–8.
46. Ferravante C, Sanna G, Melone V, Fromentier A, Rocco T, D'Agostino Y, Lamberti J, Alexandrova E, Pecoraro G, Pagliano P, et al. Nasopharyngeal virome analysis of COVID-19 patients during three different waves in Campania region of Italy. *J Med Virol*. 2022;94(5):2275–83.
47. Focosi D, Antonelli G, Pistello M, Maggi F. Torquetenovirus: the human virome from bench to bedside. *Clin Microbiol Infect*. 2016;22(7):589–93.
48. Shibayama T, Masuda G, Ajisawa A, Takahashi M, Nishizawa T, Tsuda F, Okamoto H. Inverse relationship between the titre of TT virus DNA and the CD4 cell count in patients infected with HIV. *AIDS*. 2001;15(5):563–70.
49. Thom K, Petrik J. Progression towards AIDS leads to increased Torque teno virus and Torque teno minivirus titers in tissues of HIV infected individuals. *J Med Virol*. 2007;79(1):1–7.
50. Laksono BM, de Vries RD, Duprex WP, de Swart RL. Measles pathogenesis, immune suppression and animal models. *Curr Opin Virol*. 2020;41:31–7.
51. Swets MC, Russell CD, Harrison EM, Docherty AB, Lone N, Girvan M, Hardwick HE, Investigators IC, Visser LG, Openshaw PJM, et al. SARS-CoV-2 co-infection with influenza viruses, respiratory syncytial virus, or adenoviruses. *Lancet*. 2022;399(10334):1463–4.
52. Huang Y, Skarlupka AL, Jang H, Blas-Machado U, Holladay N, Hogan RJ, Ross TM. SARS-CoV-2 and influenza A virus coinfections in ferrets. *J Virol*. 2022;96(5): e0179121.
53. Kim EH, Nguyen TQ, Casel MAB, Rollon R, Kim SM, Kim YI, Yu KM, Jang SG, Yang J, Poo H, et al. Coinfection with SARS-CoV-2 and influenza A virus increases disease severity and impairs neutralizing antibody and CD4(+) T cell responses. *J Virol*. 2022;96(6): e0187321.
54. Ahmed B, Cox MJ, Cuthbertson L, James PL, Cookson WOC, Davies JC, Moffatt MF, Bush A. Comparison of the upper and lower airway microbiota in children with chronic lung diseases. *PLoS ONE*. 2018;13(8): e0201156.
55. Wen Z, Xie G, Zhou Q, Qiu C, Li J, Hu Q, Dai W, Li D, Zheng Y, Wen F. Distinct nasopharyngeal and oropharyngeal microbiota of children with influenza A virus compared with healthy children. *Biomed Res Int*. 2018;2018:6362716.
56. Xiao W, Ma ZS. Inter-individual diversity scaling analysis of the human virome with classic diversity-area relationship (DAR) modeling. *Front Genet*. 2021;12: 627128.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

