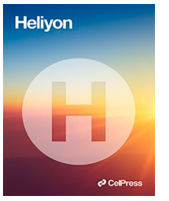
Heliyon 10 (2024) e31556

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Target enrichment metaviromics enables comprehensive surveillance of coronaviruses in environmental and animal samples

Sandra Martínez-Puchol [a,g], Maria Tarradas-Alemany [a,b,e], Cristina Mejías-Molina [a,d], Marta Itarte [a,d], Marta Rusiñol [a,d], Jordi Baliellas [h], Nerea Abasolo [i], Núria Canela [i], Abir Monastiri [c,f], Marc López-Roig [c,f], Jordi Serra-Cobo [c,f], Josep F. Abril [b,e], Sílvia Bofill-Mas [a,d,*]

[a] *Laboratory of Viruses Contaminants of Water and Food, Departament de Genètica. Microbiologia i Estadística, Universitat de Barcelona, Barcelona, Catalonia, Spain*
[b] *Computational Genomics Lab, Departament de Genètica. Microbiologia i Estadística, Universitat de Barcelona, Barcelona, Catalonia, Spain*
[c] *Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Universitat de Barcelona, Catalonia, Spain*
[d] *Institut de Recerca de l'Aigua (IdRA), Universitat de Barcelona, Barcelona, Catalonia, Spain*
[e] *Institut de Biomedicina (IBUB), Universitat de Barcelona, Barcelona, Catalonia, Spain*
[f] *Institut de Recerca de la Biodiversitat (IRBiO), Universitat de Barcelona, Catalonia, Spain*
[g] *Vicerectorat de Recerca. Universitat de Barcelona, Barcelona, Catalonia, Spain*
[h] *Grup Sanejament Porcí, Lleida, Catalonia, Spain*
[i] *Centre for Omic Sciences (COS), Joint Unit Universitat Rovira i Virgili-EURECAT, Unique Scientific and Technical Infrastructures (ICTS), Eurecat, Centre Tecnològic de Catalunya, Catalonia, Spain*

A B S T R A C T

The COVID-19 pandemic has underscored the importance of understanding the role of animals in the transmission of coronaviruses (CoVs) and their impact on human health. A One Health approach, integrating human, animal, and environmental health, is essential for effective CoVs control. Next-generation sequencing has played a pivotal role in identifying and monitoring the evolution of novel CoVs strains, like SARS-CoV-2. However, viral occurrence and diversity studies in environmental and animal samples are challenging because of the complexity of viral communities and low abundance of viruses in these samples.

Target enrichment sequencing (TES) has emerged as a valuable tool for investigating viral families in challenging samples. This approach involves the specific capture and enrichment of viral genomes using sequence-specific probes, thereby enhancing the efficiency of detection and characterization.

In this study, we aimed to develop and validate a TES panel to study CoVs in various complex environmental and animal derived samples. The results demonstrated the panel's effectiveness in capturing and sequencing a wide diversity of CoVs providing valuable insights into their abundance and host diversity in urban wastewater, farm animal corpses lixiviates and bat guano samples. In sewage samples, CoVs were detected solely when TES was employed while in guano samples, sequencing of CoVs species was achieved in 2 out of 4 samples showing an almost three-logarithmic increase in the number of reads obtained in comparison with the untargeted approach. For animal lixiviates, only the TES application enabled the acquisition of CoVs reads. The information obtained can significantly contribute to early detection, surveillance, and control measures for CoVs, including viral discovery and potential spillover events. Additionally, this sequencing panel shows potential for studying other significant viral families and monitoring viral diversity in different animal populations.

* Corresponding author. Laboratory of Viruses Contaminants of Water and Food, Departament de Genètica. Microbiologia i Estadística, Universitat de Barcelona, Barcelona, Catalonia, Spain.
E-mail address: sbofill@ub.edu (S. Bofill-Mas).

https://doi.org/10.1016/j.heliyon.2024.e31556
Received 30 October 2023; Received in revised form 8 May 2024; Accepted 17 May 2024
Available online 19 May 2024

1. Introduction

The COVID-19 pandemic highlighted the necessity of understanding the role of animals in the transmission of coronaviruses (CoVs) and their potential impact on human health [1]. CoVs can cause a range of illnesses in infected humans and animals. Their transmission occurs through respiratory droplets, but other bodily fluids could be involved and for this, studying their presence in human and animal excreta can provide valuable information about CoVs circulation [2]. CoVs are zoonotic pathogens capable of cross-species transmission, resulting in outbreaks in both humans and animals. Consequently, the implementation of effective strategies for controlling CoVs requires adopting a One Health approach that integrates human, animal, and environmental health [3,4]. Therefore, it becomes imperative to investigate the interactions between humans, animals, and their shared environments to identify potential reservoirs and intermediate CoVs hosts. This knowledge plays a crucial role in proactively preventing future spillover events and emerging infectious diseases, helping early detection of potential outbreaks and prompt response measures [5].

To identify potential sources of human infections or the emergence of new zoonotic species or strains, it is essential to study the genetic diversity and evolutionary dynamics of CoVs in both domesticated and wild animals, as well as in the environment [6,7]. Next-generation sequencing (NGS) has played a crucial role in identifying novel CoVs strains, including those responsible for recent outbreaks like SARS-CoV-2, and has enabled the monitoring of their evolution [8,9]. This technology facilitates the detection of genetic mutations, and the tracking of the emergence and spread of new variants within populations.

Virome studies on environmental and animal samples present significant challenges due to the complex and diverse nature of viral communities within these samples. One of the primary obstacles in these studies lies in the extensive viral diversity observed in environmental samples [10], which often remains unexplored or insufficiently characterized. Traditional virological methods often have limitations in capturing and identifying the full extent of viral genomes within a sample. Moreover, viruses may exhibit low abundance compared to other microorganisms, further complicating their detection and analysis [11].

To overcome these challenges, target enrichment sequencing (TES) has emerged as a valuable tool, proving to be highly beneficial in the investigation of various viral families, including CoVs, in clinical and environmental samples. However, previous attempts using a commercial probe-panel directed to vertebrate viruses only yielded a few reads from SARS-CoV-2 and other human and animal CoVs in sewage [12,13]. This technique enables the specific capture and enrichment of viral genomes using sequence-specific probes, thereby enhancing the sensitivity and efficiency of detection and characterization [14,15]. The use of targeted panels designed to capture specific viral families of interest within challenging samples can enhance NGS approaches, resulting in increased sequencing depth and providing a more comprehensive understanding of viral diversity. This becomes particularly crucial when studying reservoir animals such as bats [16] or potential intermediate hosts like farm animals [17].

Consequently, the aim of this study was to develop and validate the efficacy of a TES panel specifically designed to capture members of the *Coronaviridae* family to study the diversity of CoVs in sewage, bat guano, and farm-derived animal samples.

2. Materials and methods

2.1. Sample collection, viral concentration, and nucleic acid extraction

Urban 24-h composite sewage samples were collected on November 16th (Sewage02), and December 14th (Sewage03), 2021 and on January 3rd, 2022 (Sewage01) from a wastewater treatment plant (WWTP) located in the city of Barcelona. This plant serves up to 2.8 million population equivalents and receives both domestic and industrial waste from the sewer system. Upon collection, the samples were placed in a sterile container and kept at 4 °C until viral particles were concentrated from 100 ml of sewage. In brief, the samples underwent initial debris removal through centrifugation at 4750×*g* for 30 min and resulting supernatant (80 mL) was ultrafiltered using the automatic Concentration Pipette (CP-Select™) with 150 kDa tips (Innovaprep, Missouri, USA) [18]. After the process, viral concentrates were recovered into a volume of 300 μL. Nucleic acids were extracted using QIAmp RNA Viral Mini Kit (Qiagen, Hilden, Germany).

Pools of guano were collected from bat roosts in four different locations in Catalonia (33) and Balearic Islands (1). Refuge 1 (June 2019, guano 01) shelters various bat species, notably including significant colonies of *Miniopterus schreibersii* and *Myotis myotis* of hundred individuals. Refuge 2 (March 2016, Guano 02) hosts a hibernating colony of over 2000 *Miniopterus schreibersii*. Refuge 3 (August 2017, Guano 03) hosts various bat species, including significant breeding colonies of *Miniopterus schreibersii* and *Myotis myotis* of hundred individuals. Refuge 4 (December 2015, Guano 04) hosts a very important colony of 15,000–17,000 *Miniopterus schreibersii*. All guano samples were obtained either before or after the hibernating or breeding periods and were preserved using RNAlater (Invitrogen, Massachusetts, USA) and stored at −80 °C.

From each guano pool, viral particles were concentrated from 1 gr of feces using glycine (0.25 N, 9.5 pH) and ultracentrifugation protocol described by Pina et al. [19]. To eliminate free-DNA from sewage and guano viral concentrates, Turbo DNAse (Invitrogen, Massachusetts, USA) was used, followed by nucleic acid extraction using QIAamp Viral RNA mini kit (Qiagen, Hilden, Germany) with an initial concentrate volume of 280 μL and an elution volume of 80 μL.

Animal lixiviate samples from chicken, porcine and rabbit from animal corpse trucks were collected using sterile containers directly from the leachate generated during discharge at a rendering plant located in Catalonia and stored at 4 °C during transport to the laboratory where they were analyzed. Viruses were released from 2 gr of lixiviate by vortexing and then gentle shaking with 20 ml of glycine buffer (0.25 N, 9.5 pH) for 20 min at room temperature. The mixture was then centrifuged at 8000 ×*g* for 15 min. Total nucleic acids were purified from the supernatant using the Maxwell® Enviro Wastewater TNA Kit (Promega, Wisconsin, USA), the VacMan®

Vacuum Manifold and the automated Maxwell RSC instrument into a final volume of 60 μL.

From all the samples tested, SARS-CoV-2 N1 gene quantification was performed by (RT)-qPCR (probe, primers and cycling conditions described in the CDC-006-00019 CDC/DDID/NCIRD/Division of Viral Diseases protocol).

Division of Viral Diseases protocol).

2.2. Library preparation and probe-based target enrichment sequencing

Prior to library preparation, nucleic acids and a negative control were reverse transcribed into cDNA, which was then tagged and complemented to obtain double-stranded DNA (dsDNA). This randomly tagged dsDNA, representing the viral content, was subsequently amplified through 25 cycles of PCR amplification to generate enough DNA for library preparation, following the methodology described by Fernandez-Cassi et al. [20]. Libraries preparation was carried out in duplicate using the KAPA HyperPrep Kit, following the manufacturer's instructions (Roche-Kapa Biosystems, Basel, Switzerland). In brief, the dsDNA was enzymatically fragmented, indexed with Dual Indexes, and amplified. The resulting DNA fragments' quality and concentration were assessed using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Massachusetts, USA).

One of the libraries replicates was then subjected to hybridization with probes specifically designed to capture target species from the *Coronaviridae* family utilizing the KAPA HyperCap Enrichment Kit (Basel, Switzerland). Capture probes were designed based on the genomic sequences of 18 species from *Coronaviride* family with the purpose to maximize the hybridization of the probes with these sequences and related ones, and to minimize the capture of any other genomic material that might be present in the environment. Species used for the design were: Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/IND/CD211295/2021, SARS coronavirus Tor2, Middle East respiratory syndrome-related coronavirus isolate HCoV-EMC/2012, Human coronavirus OC43 strain ATCC VR-759, Human coronavirus 229E, Human Coronavirus NL63, Human coronavirus HKU1, Porcine coronavirus HKU15 strain HKU15-155, Bovine coronavirus isolate BCoV-ENT, Murine hepatitis virus strain A59, Transmissible gastroenteritis virus, Feline infectious peritonitis virus, Avian infectious bronchitis virus, Alphacoronavirus Bat-CoV/P.kuhlii/Italy/3398-19/2015, Bat coronavirus 1A, Rousettus bat coronavirus HKU10, Bat coronavirus HKU8 and Bat coronavirus HKU2. Reference genomic sequences were obtained from RefSeq and Genbank (see Supplementary Information 1).

Following the capture step, the quality and concentration of the captured libraries, non-captured libraries, and negative controls were reassessed. Subsequently, sequencing of these libraries was conducted on an Illumina NextSeq platform (Illumina, California USA), generating an output of 400 million reads.

2.3. Bioinformatic analysis

Raw sequencing reads from samples with and without capture—for each set of Sewage (Sewage01, Sewage02 and Sewage03) and Bat Guano (Guano01, Guano02, Guano03 and Guano04) samples, were cleaned, filtered, and assembled using CAPTVRED pipeline, an automated protocol designed to assess the virome present in complex samples, specially focused on those obtained by capture-based metagenomics approach. The analysis is sensible to degraded genomic sequences that may be present in the environment and provides comprehensive, reproducible and accessible results centered on viral outcomes (Tarradas-Alemany et al., submitted manuscript). It was run locally in a Debian server with 32 threads. Main steps of the pipeline and relevant parameters are briefly described hereafter. For the pipeline cleaning step, the paired FASTQ files were processed with BBDuk (BBMap version 38.96; [21]), and reads quality was assessed using FastQC (version 0.11.9; [22]) and MultiQC (version 1.9; [23]). A second filtering step was performed using Kaiju (version 1.9.0; [24]) to discard reads corresponding to non-viral species (such as eukaryotic, archaea, or bacterial sequences remnant on the samples). After that, paired and unpaired reads were assembled into contigs by Megahit (version 1.0; [25]); only those contig sequences and non-assembled reads (singletons) over 100 bp were selected. All programs were run using the pipeline default parameters as described in the pipeline documentation. Then, NCBI-BLASTn (version 2.11.0+; [26]) was run using C-RVDB [27] database as reference (identity >50 % and E-value <10E-10). Sequenced raw reads obtained from the chicken, porcine and rabbit lixiviates samples were fully processed using CAPTVRED complete pipeline which already implemented the BLAST step. The default parameters were used for this analysis.

3. Results and discussion

The use of new targeted methodologies could favor the discovery of potential sources of human infections or the emergence of new zoonotic species or strains. In this study we aimed to validate the efficacy of a TES panel specifically designed for studying members of the *Coronaviridae* family to study the diversity of coronaviruses in sewage and animal samples.

For the samples analyzed, SARS-CoV-2 concentrations ranged from 5,74E+01 GC/ml to 4,10E+03 GC/ml in sewage samples. Guano and animal lixiviate samples tested negative.

One of the objectives of the use of capture probes is to detect sequences with medium to high similarity to the genomes included in the panel, allowing the application of viral discovery analyses if appropriate. For this reason, the identity threshold is set to 50 % as a default parameter in the CAPTVRED protocol. Considering the environmental origin of the samples and that performance of the probes was unknown, this 50 % identity threshold is a safe and flexible threshold to explore the sensitivity and accuracy of the probes. In addition, same parameters must be used for the targeted and untargeted analyses to allow fair comparison between the two approaches. The results reported by the pipeline are always curated and in the case of obtaining a doubtful assignation this can be

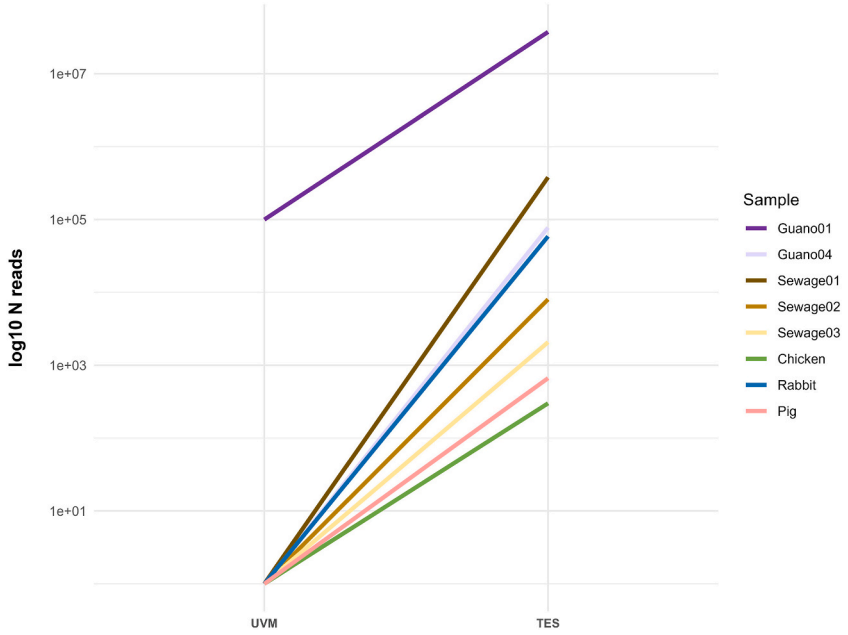


Fig. 1. Representation of the Log10 of number of CoV reads showing increase when applying the designed TES approach.

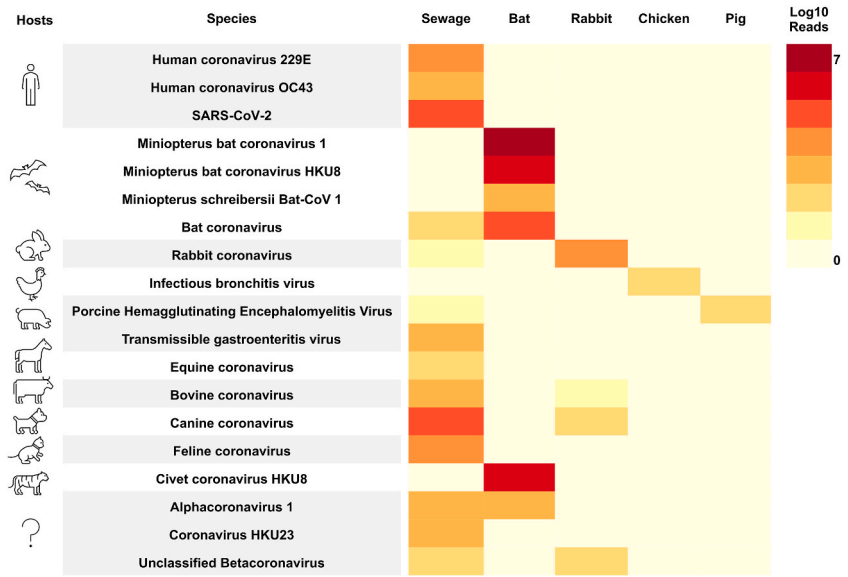


Fig. 2. Heatmap representing the CoV species abundance (Log10 number of reads) detected in the analyzed samples.

discarded.

3.1. Panel performance on Coronaviridae family

The abundance of CoV in the studied samples in comparison to the analysis of the same samples by UVM (Untargeted Viral Metagenomics) is graphically represented in Fig. 1. In sewage samples, CoV were detected solely when TES was employed. A total of 390,553 reads assigned to diverse CoV species were observed across the three analyzed sewage samples.

Concerning guano samples, sequencing of CoV species was achieved in 2 out of 4 samples. In one of the positive samples, Guano 01, CoV reads were obtained even without using the enrichment panel, resulting in 99,678 reads. However, the implementation of the panel significantly increased the number of reads obtained to 37,537,788, representing an almost three-logarithmic increase.

In the case of animal lixiviate samples, only the enrichment strategy enabled the acquisition of CoV reads. Specifically, 58,767 reads

were obtained for rabbit lixiviates, while pig and chicken lixiviate yielded 665 and 300 reads, respectively.

3.2. CoV diversity using TES

A wide range of species was detected throughout the samples analyzed when using the enrichment panel, as shown in Fig. 2, which presents their abundance and host diversity.

Interestingly, sewage samples exhibited the highest diversity of CoVs. This observation is not surprising, as these complex samples are known to harbor an extensive range of viral diversity [12,20]. Consequently, studying less prevalent species, such as CoVs, within these samples can be challenging and the use of TES methodologies represents the most effective approach to address this limitation.

In this study, the application of a CoV probe panel represented a considerable improvement in reference to other targeted efforts. Previous attempts using a commercial probe-panel directed to vertebrate viruses in urban sewage only yielded 8 contigs for SARS-CoV-2, 1 for canine CoV 1 for feline CoV 1 for human CoV OC43 and 5 contigs for Lucheng Rn rat CoV [13].

As shown in Fig. 2, SARS-CoV-2 and Canine CoV exhibited the highest number of reads followed by HCoV-229E and Feline CoV. The presence of these species can be attributed to their well-established epidemiological status. SARS-CoV-2 was associated with the ongoing pandemic during the study period (November 2021–January 2022), while HCoV-229E and HCoV OC43 are commonly linked to the occurrence of the common cold [28]. The presence of Canine CoV in the samples can be explained by canine waste disposal through toilets or through streets cleaning which would be applicable to another pet associated CoVs.

The remaining detected species display infectivity across a wide range of animal hosts, providing valuable insights into the contribution of different animal types (domesticated, farm, and wild animals) to the sewage in the studied area. The utilization of these targeted panels not only enhances the understanding of circulating strains but also serves as a potential tool for early detection and response. This highlights the usefulness of this approach in monitoring CoV and other significant viral families, including certain enteroviruses and poxviruses, as suggested in previous studies [11,12,29].

In the bat guano samples, despite the higher read count compared to other samples, the observed diversity is relatively limited, primarily consisting of four distinct species of CoVs known to infect bats. The more abundant of these species is *Miniopterus bat coronavirus 1*, also known as Bat-CoV MOP1, which belongs to the Alphacoronavirus genus and was the first CoV identified in bats [30]. Bat-CoV MOP1 is closely related to the other CoV found in the analyzed samples, such as HKU8 and Bat-CoV 1, all of which have a documented impact on *Miniopterus* spp., commonly known as bent-winged bats [31]. These bats are considered reservoirs of Alphacoronavirus species [32].

In addition to this species, a substantial number of reads associated with Civet CoV HKU8 were also detected in two different guano samples when the enrichment panel was utilized but also without enrichment. Notably, Civet CoV HKU8 was detected in civets, and it was speculated about a potential cross-species transmission event involving a bat-associated CoV [33].

The successful use of the enrichment panel for CoV characterization in bats has been documented in previous studies [34,35]. Therefore, the feasibility of conducting targeted, cost-effective, large-scale genome-level surveillance of bat CoVs has been firmly established. The information gained from such surveillance efforts could prove invaluable in the prevention and control of potential spillover events from bats to adjacent mammalian species. Migratory species may contribute to pathogen persistence in species encountered along the migratory path. Given the potential for long-range seasonal movements of *Miniopterus schreibersii* [36–38], this specie may represent a central vector for spatial dispersion of viruses in southern Europe, where this bat species is abundant. Determining which bat colonies exhibit a higher prevalence of viruses with zoonotic potential enables the implementation of preventive measures to mitigate the transmission to the human population, livestock, and pets.

In the analysis of animal lixiviate samples, a lower viral diversity was detected in comparison to the results obtained from wastewater and guano samples. This observation suggests that these highly complex samples may have a lower contribution of viral excretion, a higher degree of viral nucleic acid degradation, or may require an optimized methodology for viral isolation.

Infectious bronchitis virus (IBV) reads were detected in chicken lixiviates. IBV is a particularly relevant virus that infects various avian body compartments, including the respiratory tract, kidney, gut, and reproductive systems, and is a significant contributor to economic losses within the poultry industry [39]. Since live attenuated vaccine is available against IBV, the possibility that some of these sequences belong to excreted vaccine strains could not be ruled out.

The continuous emergence of novel viral strains through mutations and recombination events in the viral genome exacerbates the challenges associated with the identification and control of IBV [40]. Given its status as a major respiratory pathogen affecting the European poultry industry, advancements in IBV detection, particularly in conjunction with other relevant CoVs, utilizing enrichment panels, could hold the key to overcoming the apparent difficulties in its control.

Porcine hemagglutinating encephalomyelitis virus (PHEV) sequences were found in pig lixiviates samples. PHEV is responsible for causing vomiting and/or encephalomyelitis in pigs, being the sole identified neurotropic CoV known to affect pigs. PHEV exhibits a high prevalence and commonly circulates subclinically within most swine herds globally, not having significant clinical implications in many swine-producing countries [41]. While PHEV is not currently recognized as a zoonotic virus, it is crucial to acknowledge the emerging understanding of cross-species transmission and zoonotic potential demonstrated by other porcine CoVs such as Porcine delta CoV and Swine acute diarrhea syndrome CoV [42]. This knowledge underscores the importance of implementing effective control measures for CoVs present in pigs.

In the sequencing analysis of rabbit lixiviates, the presence of Rabbit CoV, Bovine CoV, and Canine CoV was observed. Rabbit CoV is recognized as the causative agent of common enteric infections in rabbit colonies [43]. However, the extent of its involvement in cross-species transmission events, as well as the implications of Bovine CoV, remain largely unknown. Canine CoV, which was also detected in sewage samples, has been associated with febrile or lower respiratory symptoms in humans [44,45], suggesting

Table 1
CoV assignations obtained by TES and UVM from analyzed samples including genome coverage, and nucleotide identity represented in percentages.

Sample	*Coronaviridae* species	TES			UVM		
		Reads (log10 N)	Genome coverage (%)	Identity (%)	Reads (log10 N)	Genome coverage (%)	Identity (%)
Sewage	Alphacoronavirus 1	3,45	96,50	99,87	ND	–	–
	Bat coronavirus	2,97	99,97	97,89	ND	–	–
	Bovine coronavirus	3,94	4,64	97,29	ND	–	–
	Canine coronavirus	2,96	95,65	92,60	ND	–	–
	Coronavirus HKU23	3,95	73,08	98,07	ND	–	–
	Equine coronavirus	2,43	71,04	93,69	ND	–	–
	Feline coronavirus	3,97	92,30	83,87	ND	–	–
	Human coronavirus 229E	4,34	85,59	99,17	ND	–	–
	Human coronavirus OC43	3,93	74,36	95,33	ND	–	–
	Porcine hemagglutinating encephalomyelitis virus	1,76	73,43	95,22	ND	–	–
	Rabbit coronavirus	1,32	71,60	93,44	ND	–	–
	SARS-CoV-2	5,80	99,96	99,89	ND	–	–
	Transmissible gastroenteritis virus	5,80	97,18	92,24	ND	–	–
	Unclassified Betacoronavirus	1,89	99,51	90,98	ND	–	–
Batguano	Alphacoronavirus 1	3,94	73,01	74,45	ND	–	–
	Bat coronavirus	5,57	99,96	79,81	2,77	91,02	79,60
	Civet coronavirus HKU8	6,14	85,86	96,63	2,31	82,22	96,37
	Miniopterus bat coronavirus 1	7,54	98,45	97,25	4,99	98,27	96,06
	Miniopterus bat coronavirus HKU8	6,09	98,33	86,11	2,40	88,56	84,69
	Miniopterus schreibersii bat coronavirus 1	3,79	93,67	74,68	ND	–	–
Chicken	Infectious bronchitis virus	2,48	1,71	100,00	ND	–	–
Rabbit	Rabbit coronavirus	4,76	2,85	97,74	ND	–	–
	Canine coronavirus	2,94	2,43	96,43	ND	–	–
	Bovine coronavirus	1,86	0,75	98,21	ND	–	–
	Unclassified Betacoronavirus	2,49	0,73	99,04	ND	–	–
Pig	Porcine hemagglutinating encephalomyelitis virus	2,82	1,16	98,42	ND	–	–

possibilities for cross-species transmission. Therefore, controlling its spread is significant to prevent potential recombinant transmissions among various mammalian species.

In addition to identifying the obtained species from each sample, further characterization of these viral assignments was conducted based on genomic coverage and identity, as outlined in Table 1.

The use of TES led to notable differences in read acquisition, especially in sewage and bat samples, where substantial genome coverage was achieved for viral assignments when employing the enrichment panel. Conversely, in the unenriched sample (UVM) where reads were obtained without the enrichment, a lower level of genomic coverage was observed, consistent with previous comparisons of these NGS techniques [13].

The coverage observed in the animal lixiviate samples was relatively low, potentially attributed to the lower quantity of reads obtained for certain species that could have been caused, as explained before, for the lower contribution of viral excretion in these samples or for the viral nucleic acid degradation.

4. Conclusions

The probe-based enrichment panel designed for sequencing members of the *Coronaviridae* family demonstrated success in capturing and sequencing not only the targeted species but also other viral species within the family. A higher number of different CoV species were detected in comparison with untargeted metagenomics or with vertebrate viruses directed panels applied in other studies by this research group. This enrichment panel provided increased sequencing depth for the intended viral species. It also showed potential for viral discovery which could enable in the future comprehensive studies on both known and potential animal species susceptible to spillover events. This approach holds promise for enhancing our understanding of viral dynamics and their potential implications for zoonotic transmission.

Availability of data and material

Raw sequencing data is available from:
https://compgen.bio.ub.edu/datasets/Target_Enrichment_MetaViromics_CoV/rawdata/

During the preparation of this work the authors used https://chat.openai.com/to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Sandra Martínez-Puchol: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Maria Tarradas-Alemany:** Writing – review & editing, Visualization, Software, Formal analysis. **Cristina Mejías-Molina:** Methodology. **Marta Itarte:** Writing – review & editing, Methodology. **Marta Rusiñol:** Supervision, Resources, Project administration, Investigation, Funding acquisition. **Jordi Baliellas:** Resources, Conceptualization. **Nerea Abasolo:** Methodology. **Núria Canela:** Validation, Project administration, Methodology, Funding acquisition. **Abir Monastiri:** Methodology. **Marc López-Roig:** Writing – review & editing, Methodology, Investigation. **Jordi Serra-Cobo:** Writing – review & editing, Investigation. **Josep F. Abril:** Writing – review & editing, Software, Methodology, Investigation. **Sílvia Bofill-Mas:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests Silvia Bofill Mas reports financial support was provided by "La Marató de TV3". Silvia Bofill-Mas reports financial support was provided by Spain Ministry of Science and Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors express their gratitude to the staff of the rendering and wastewater treatment plants involved in this study for their assistance and technical support during sampling.

The authors wish to express their gratitude to Ferran Briansó from Roche Diagnostics for the collaborative efforts in developing the capture probe panel.

This research was conducted as part of the EPISARS project, which received funding from donations obtained during “La Marató de TV3” dedicated to COVID-19 and was also funded by MCIN/AEI/10.13039/501100011033 and by “ERDF A way of making Europe”.

S. M-P holds a *Margarita Salas* grant funded by Ministerio de Universidades - NextGeneration EU. S. B-M is a Serra-Hunter fellow at the University of Barcelona. M. T- A. holds a predoctoral grant (FPI) funded by Ministerio de Universidades. C. M-M. holds a predoctoral grant (FI-SDUR) funded by the Catalan Government ’AGAUR’ at the University of Barcelona.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e31556.

References

[1] J.E. Pekar, A. Magee, E. Parker, N. Moshiri, K. Izhikevich, J.L. Havens, K. Gangavarapu, L.M.M. Serrano, A. Crits-Christoph, N.L. Matteson, M. Zeller, J.I. Levy, J. C. Wang, S. Hughes, J. Lee, H. Park, M.S. Park, K.C.Z. Yan, R.T.P. Lin, J.O. Wertheim, The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2, Science 377 (6609) (2022), https://doi.org/10.1126/science.abp8337.

[2] S. Maryam, I. Ul Haq, G. Yahya, M. Ul Haq, A.M. Algammal, S. Saber, S. Cavalu, COVID-19 surveillance in wastewater: an epidemiological tool for the monitoring of SARS-CoV-2, Front. Cell. Infect. Microbiol. 12 (2023) 978643, https://doi.org/10.3389/fcimb.2022.978643.

[3] M. Ruiz-Aravena, C. McKee, A. Gamble, T. Lunn, A. Morris, C.E. Snedden, C.K. Yinda, J.R. Port, D.W. Buchholz, Y.Y. Yeo, C. Faust, E. Jax, L. Dee, D.N. Jones, M. K. Kessler, C. Falvo, D. Crowley, N. Bharti, C.E. Brook, R.K. Plowright, Ecology, evolution and spillover of coronaviruses from bats, Nat. Rev. Microbiol. 20 (Issue 5) (2022), https://doi.org/10.1038/s41579-021-00652-2.

[4] K. Sharun, K. Dhama, A.M. Pawde, C. Gortázar, R. Tiwari, D.K. Bonilla-Aldana, A.J. Rodriguez-Morales, J. de la Fuente, I. Michalak, Y.A. Attia, SARS-CoV-2 in animals: potential for unknown reservoir hosts and public health implications, Vet. Q. 41 (1) (2021), https://doi.org/10.1080/01652176.2021.1921311.

[5] A. Islam, J. Ferdous, S. Islam, M.A. Sayeed, M.K. Rahman, O. Saha, M.M. Hassan, T. Shirin, Transmission dynamics and susceptibility patterns of SARS-CoV-2 in domestic, farmed and wild animals: sustainable One Health surveillance for conservation and public health to prevent future epidemics and pandemics, Transboundary and Emerging Diseases 69 (Issue 5) (2022), https://doi.org/10.1111/tbed.14356.

[6] C.J. Carlson, R.J. Gibb, G.F. Albery, L. Brierley, R.P. Connor, T.A. Dallas, E.A. Eskew, A.C. Fagre, M.J. Farrell, H.K. Frank, R.L. Muylaert, T. Poisot, A. L. Rasmussen, S.J. Ryan, S.N. Seifert, The global virome in one network (VIRION): an atlas of vertebrate-virus associations, mBio 13 (2) (2022), https://doi.org/10.1128/mbio.02985-21.

[7] W.K. Jo, C. Drosten, J.F. Drexler, The evolutionary dynamics of endemic human coronaviruses, Virus Evolution 7 (1) (2021), https://doi.org/10.1093/ve/veab020.

[8] C. Charre, C. Ginevra, M. Sabatier, H. Regue, G. Destras, S. Brun, G. Burfin, C. Scholtes, F. Morfin, M. Valette, B. Lina, A. Bal, L. Josset, Evaluation of NGS-based approaches for SARS-CoV-2 whole genome characterisation, Virus Evolution (2020), https://doi.org/10.1093/ve/veaa075.

[9] H. Liu, J. Li, Y. Lin, X. Bo, H. Song, K. Li, P. Li, M. Ni, Assessment of two-pool multiplex long-amplicon nanopore sequencing of SARS-CoV-2, J. Med. Virol. 94 (1) (2022) 327–334, https://doi.org/10.1002/jmv.27336.

[10] P.G. Cantalupo, B. Calgua, G. Zhao, A. Hundesa, A.D. Wier, J.P. Katz, M. Grabe, R.W. Hendrix, R. Girones, D. Wang, J.M. Pipas, Raw sewage harbors diverse viral populations, mBio 2 (5) (2011), https://doi.org/10.1128/mBio.00180-11.

[11] X. Fernandez-Cassi, M. Rusiñol, S. Martínez-Puchol, Viral concentration and amplification from human serum samples prior to application of next-generation sequencing analysis, Methods Mol. Biol. 1838 (2018) 173–188, https://doi.org/10.1007/978-1-4939-8682-8_13.

[12] S. Martínez-Puchol, M. Rusiñol, X. Fernández-Cassi, N. Timoneda, M. Itarte, C. Andrés, A. Antón, J.F. Abril, R. Girones, S. Bofill-Mas, Characterisation of the sewage virome: comparison of NGS tools and occurrence of significant pathogens, Sci. Total Environ. 713 (2020), https://doi.org/10.1016/j.scitotenv.2020.136604.

[13] S. Martínez-Puchol, M. Itarte, M. Rusiñol, E. Forés, C. Mejías-Molina, C. Andrés, A. Antón, J. Quer, J.F. Abril, R. Girones, S. Bofill-Mas, Exploring the diversity of coronavirus in sewage during COVID-19 pandemic: don't miss the forest for the trees, Sci. Total Environ. 800 (2021), https://doi.org/10.1016/j.scitotenv.2021.149562.

[14] T. Briese, A. Kapoor, N. Mishra, K. Jain, A. Kumar, O.J. Jabado, W. Ian Lipkina, Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis, mBio 6 (5) (2015) 1–12, https://doi.org/10.1128/mBio.01491-15.

[15] A.C. Paskey, K.G. Frey, G. Schroth, S. Gross, T. Hamilton, K.A. Bishop-Lilly, Enrichment post-library preparation enhances the sensitivity of high-throughput sequencing-based detection and characterization of viruses from complex samples, BMC Genom. 20 (1) (2019) 155, https://doi.org/10.1186/s12864-019-5543-2.

[16] H. Zhou, X. Chen, T. Hu, J. Li, H. Song, Y. Liu, P. Wang, D. Liu, J. Yang, E.C. Holmes, A.C. Hughes, Y. Bi, W. Shi, A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein, Curr. Biol. 30 (11) (2020) 2196–2203.e3, https://doi.org/10.1016/j.cub.2020.05.023.

[17] M. Khamassi Khbou, M. Daaloul Jedidi, F. Bouaicha Zaafouri, M. Benzarti, Coronaviruses in farm animals: epidemiology and public health implications, Veterinary Medicine and Science 7 (Issue 2) (2021), https://doi.org/10.1002/vms3.359.

[18] E. Forés, S. Bofill-Mas, M. Itarte, S. Martínez-Puchol, A. Hundesa, M. Calvo, C.M. Borrego, L.L. Corominas, R. Girones, M. Rusiñol, Evaluation of two rapid ultrafiltration-based methods for SARS-CoV-2 concentration from wastewater, Sci. Total Environ. 768 (2021), https://doi.org/10.1016/j.scitotenv.2020.144786.

[19] S. Pina, M. Puig, F. Lucena, J. Jofre, R. Girones, Viral pollution in the environment and in shellfish: human adenovirus detection by PCR as an index of human viruses, Appl. Environ. Microbiol. 64 (9) (1998), https://doi.org/10.1128/aem.64.9.3376-3382.1998.

[20] X. Fernandez-Cassi, N. Timoneda, S. Martínez-Puchol, M. Rusiñol, J. Rodriguez-Manzano, N. Figuerola, S. Bofill-Mas, J.F. Abril, R. Girones, Metagenomics for the study of viruses in urban sewage as a tool for public health surveillance, Sci. Total Environ. 618 (August) (2018) 870–880, https://doi.org/10.1016/j.scitotenv.2017.08.249.

[21] J.G. Institute, BBDuk, BBMap tools. https://shorturl.at/qJOS4, 2022.

[22] S. Andrews, FastQC: A Quality Control Tool for High Throughput Sequence Data, Babraham Institute, 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

[23] P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: summarize analysis results for multiple tools and samples in a single report, Bioinformatics 32 (19) (2016) 3047–3048, https://doi.org/10.1093/bioinformatics/btw354.

[24] P. Menzel, K.L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju, Nat. Commun. 7 (2016) 11257, https://doi.org/10.1038/ncomms11257.

[25] D. Li, R. Luo, C.M. Liu, C.M. Leung, H.F. Ting, K. Sadakane, H. Yamashita, T.W. Lam, MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices, Methods 102 (2016) 3–11, https://doi.org/10.1016/j.ymeth.2016.02.020.

[26] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (3) (1990) 403–410, https://doi.org/10.1016/S0022-2836(05)80360-2.

[27] N. Goodacre, A. Aljanahi, S. Nandakumar, M. Mikailov, A.S. Khan, A reference viral database (RVDB) to enhance bioinformatics analysis of high-throughput sequencing for novel virus detection, mSphere 3 (2) (2018), https://doi.org/10.1128/mSphereDirect.00069-18.

[28] L. Van der Hoek, Human coronaviruses: what do they cause? Antivir. Ther. 12 (March) (2015).

[29] A. Hata, M. Kitajima, E. Haramoto, S. Lee, M. Ihara, C.P. Gerba, H. Tanaka, Next-generation amplicon sequencing identifies genetically diverse human astroviruses, including recombinant strains, in environmental waters, Sci. Rep. 8 (1) (2018) 11837, https://doi.org/10.1038/s41598-018-30217-y.

[30] L.L.M. Poon, D.K.W. Chu, K.H. Chan, O.K. Wong, T.M. Ellis, Y.H.C. Leung, S.K.P. Lau, P.C.Y. Woo, K.Y. Suen, K.Y. Yuen, Y. Guan, J.S.M. Peiris, Identification of a novel coronavirus in bats, J. Virol. 79 (4) (2005) 2001–2009, https://doi.org/10.1128/jvi.79.4.2001-2009.2005.

[31] D.K.W. Chu, J.S.M. Peiris, H. Chen, Y. Guan, L.L.M. Poon, Genomic characterizations of bat coronaviruses (1A, 1B and HKU8) and evidence for co-infections in Miniopterus bats, J. Gen. Virol. 89 (5) (2008), https://doi.org/10.1099/vir.0.83605-0.

[32] T. Kimprasit, M. Nunome, K. Iida, Y. Murakami, M.L. Wong, C.H. Wu, R. Kobayashi, Y. Hengjan, H. Takemae, K. Yonemitsu, R. Kuwata, H. Shimoda, L. Si, J. H. Sohn, S. Asakawa, K. Ichiyanagi, K. Maeda, H.S. Oh, T. Mizutani, E. Hondo, Dispersal history of Miniopterus fuliginosus bats and their associated viruses in east Asia, PLoS One 16 (1 January) (2021), https://doi.org/10.1371/journal.pone.0244006.

[33] W.T. He, X. Hou, J. Zhao, J. Sun, H. He, W. Si, J. Wang, Z. Jiang, Z. Yan, G. Xing, M. Lu, M.A. Suchard, X. Ji, W. Gong, B. He, J. Li, P. Lemey, D. Guo, C. Tu, S. Su, Virome characterization of game animals in China reveals a spectrum of emerging pathogens, Cell 185 (7) (2022), https://doi.org/10.1016/j.cell.2022.02.014.

[34] B. Li, H.-R. Si, Y. Zhu, X.-L. Yang, D.E. Anderson, Z.-L. Shi, L.-F. Wang, P. Zhou, Discovery of bat coronaviruses through surveillance and probe capture-based next-generation sequencing, mSphere 5 (1) (2020) 1–10, https://doi.org/10.1128/msphere.00807-19.

[35] X.F. Lim, C.B. Lee, S.M. Pascoe, C.B. How, S. Chan, J.H. Tan, X. Yang, P. Zhou, Z. Shi, O.M. Sessions, L.F. Wang, L.C. Ng, D.E. Anderson, G. Yap, Detection and characterization of a novel bat-borne coronavirus in Singapore using multiple molecular approaches, J. Gen. Virol. 100 (10) (2019) 1363–1374, https://doi.org/10.1099/jgv.0.001307.

[36] J. Serra-Cobo, M. López-Roig, X. Bayer, B. Amengual, C. Guasch, Bats: Science and Myth, Barcelona, Publications and Editions, University of Barcelona, 2009, p. 268.

[37] J. Serra-Cobo, V. Sanz-Trullén, J.P. Martinez-Rica, Migratory movements of Miniopterus schreibersii in the north-east of Spain, Acta Theriol. 43 (2017) 271–283.

[38] D. Colombi, J. Serra-Cobo, R. Métras, A. Apolloni, Ch Poletto, M. López-Roig, H. Bourhy, V. Colizza, Mechanisms for lyssavirus persistence in non-synanthropic bats in Europe: insights from a modeling study, Sci. Rep. 9 (2019) 537, https://doi.org/10.1038/s41598-018-36485-y.

[39] D. Cavanagh, Coronavirus avian infectious bronchitis virus, Vet. Res. 38 (Issue 2) (2007), https://doi.org/10.1051/vetres:2006055.

[40] M.W. Jackwood, Review of infectious bronchitis virus around the world, Avian Dis. 56 (Issue 4) (2012), https://doi.org/10.1637/10227-043012-Review.1.

[41] J.C. Mora-Díaz, P.E. Piñeyro, E. Houston, J. Zimmerman, L.G. Giménez-Lirola, Porcine hemagglutinating encephalomyelitis virus: a review, Front. Vet. Sci. 6 (FEB) (2019), https://doi.org/10.3389/fvets.2019.00053.

[42] Z. Guo, Q. Jin, P. Li, G. Xing, Q. Lu, G. Zhang, Potential cross-species transmission risks of emerging swine enteric coronavirus to human beings, J. Med. Virol. 95 (7) (2023), https://doi.org/10.1002/jmv.28919.

[43] J.P. Descôteaux, G. Lussier, L. Berthiaume, R. Alain, C. Seguin, M. Trudel, An enteric coronavirus of the rabbit: detection by immunoelectron microscopy and identification of structural polypeptides, Arch. Virol. 84 (3–4) (1985), https://doi.org/10.1007/BF01378976.

[44] J.A. Lednicky, M.S. Tagliamonte, S.K. White, G.M. Blohm, M.M. Alam, N.M. Iovine, M. Salemi, C. Mavian, J.G. Morris, Isolation of a novel recombinant canine coronavirus from a visitor to Haiti: further evidence of transmission of coronaviruses of zoonotic origin to humans, Clin. Infect. Dis. 75 (1) (2022), https://doi.org/10.1093/cid/ciab924.

[45] A.N. Vlasova, A. Diaz, D. Damtie, L. Xiu, T.H. Toh, J.S.Y. Lee, L.J. Saif, G.C. Gray, Novel canine coronavirus isolated from a hospitalized patient with pneumonia in east Malaysia, Clin. Infect. Dis. 74 (3) (2022), https://doi.org/10.1093/cid/ciab456.