

Shotgun metagenomics on indoor air for surveillance of respiratory, enteric, and skin viruses in a Belgian daycare setting, January to December 2022

Mustafa Karatas¹, Caspar Geenen², Els Keyaerts³, Lore Budts³, Joren Raymenants², Charlotte Eggers¹, Bastiaan Craessaerts³, Emmanuel André^{2,3,*}, Jelle Matthijnsens^{1,*}

1. Laboratory of Viral Metagenomics, Clinical and Epidemiological Virology, Rega Institute, Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium
2. Laboratory of Clinical Microbiology, KU Leuven, Leuven, Belgium
3. Department of Laboratory Medicine, National Reference Center of Respiratory Pathogens, University Hospitals Leuven, Leuven, Belgium

* These authors contributed equally.

Correspondence: Jelle Matthijnsens (jelle.matthijnsens@kuleuven.be)

Citation style for this article:

Karatas Mustafa, Geenen Caspar, Keyaerts Els, Budts Lore, Raymenants Joren, Eggers Charlotte, Craessaerts Bastiaan, André Emmanuel, Matthijnsens Jelle. Shotgun metagenomics on indoor air for surveillance of respiratory, enteric, and skin viruses in a Belgian daycare setting, January to December 2022. *Euro Surveill.* 2025;30(38):pii=2400711. <https://doi.org/10.2807/1560-7917.ES.2025.30.38.2400711>

Article received on 28 Oct 2024 / Accepted on 24 Feb 2025 / Published on 25 Sep 2025

Background: Hospital-based communicable disease surveillance may be costly during large outbreaks and often misses mild or asymptomatic infections. It can be enhanced by environmental surveillance, which monitors circulating pathogens, even from asymptomatic carriers. **Aim:** We investigated if tracking viruses in indoor air could be used for their surveillance in a community setting. We also tested the value of untargeted metagenomics to identify viruses in air samples. **Methods:** Weekly indoor air samples were collected with active air samplers from January until December 2022 from a daycare centre in Leuven, Belgium. Samples were analysed using respiratory and enteric quantitative (q)PCR panels, as well as with untargeted metagenomics, enabling both targeted and agnostic viral detections. **Results:** Human-associated viruses were detected in 40 of 42 samples across the study period, with MW polyomavirus being most prevalent (33 samples). Respiratory agents such as rhinoviruses and RSV-B and enteric viruses including rotavirus A, astrovirus, and adenovirus appeared at epidemiologically expected times. Skin-associated viruses were also observed, notably Merkel cell polyomavirus and STL polyomavirus. Metagenomics enabled reconstructing multiple complete genomes, distinguishing viral subtypes and detecting copresence of closely related variants. Additionally, several animal, insect, fungal, and plant viruses were found, reflecting both indoor and outdoor environmental exposure. **Conclusion:** Indoor air monitoring, combined with untargeted metagenomics, demonstrates a potential to support virus surveillance. This approach can allow monitoring circulation of viruses in community settings, including those causing asymptomatic or mild infections. By enabling to reconstruct complete viral genomes, it

allows detailed variant tracking, facilitating adapted public health responses.

Introduction

Conventionally, infectious disease surveillance has revolved around patient-centred approaches, focusing on specific diseases or syndromes (e.g. respiratory, enteric symptoms) using methods such as pathogen culturing and molecular techniques (PCR or quantitative (q)PCR panels) on individual patient samples [1]. While individual diagnostics can be valuable for immediate patient care or to assess the extent of an outbreak, these methods are often underused for broader surveillance due to cost, invasiveness and failure to detect less common or novel viruses [2]. Moreover, due to their focus on illness, they tend to miss mild or asymptomatic infections. Therefore, environmental surveillance has emerged as a relevant complement to individual testing for tracking pathogens in larger populations, allowing to monitor viral circulation in communities without relying on patient sampling.

Since the 2010s, environmental surveillance for pathogens in media such as wastewater, soil, rivers, and indoor air, has attracted growing interest, with wastewater being the most widely used method [3-5]. Wastewater analysis has effectively complemented other surveillance systems by tracking severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the variants of concern of this virus, and other respiratory and enteric pathogens using targeted PCR methods [6-8]. However, wastewater surveillance has limitations, including reduced precision in pinpointing specific locations or population subgroups, potential delays or signal dilution due to rainfall, and

KEY PUBLIC HEALTH MESSAGE

What did you want to address in this study and why?

Hospital diagnostics are vital for patient care but limited for timely surveillance: they miss mild or asymptomatic infections, may detect viruses only after community spread, and overlook pathogens not routinely tested. We investigated whether indoor-air monitoring can reveal viruses circulating in the community.

What have you learnt from this study?

In 2022, air samples from a daycare centre were tested for genetic material of either (i) specific viruses with qPCR techniques or (ii) of any virus with a metagenomics method. In 40 of 42 total samples collected during the year, human viruses were detected, despite children at the daycare appearing healthy. Tracking skin, respiratory, or enteric viruses through indoor air sampling was feasible. Known and new viruses could be simultaneously detected.

What are the implications of your findings for public health?

Indoor air monitoring can complement current virus surveillance systems by detecting known and novel viruses. Tracking viruses in the community before detections in hospitals occur, may allow health officials to more timely prevent their circulation among people, gain more information on their epidemiology and improve outbreak responses.

dependency on water usage [9]. As an alternative approach, indoor air sampling can be performed using active or passive samplers, with sampling conducted at specific locations for precise surveillance purposes. Active air sampling facilitates immediate sample collection within a short time frame and offers higher sensitivity, as demonstrated especially in studies on SARS-CoV-2 [10]. In 2023 and 2024, our teams and others have demonstrated the potential of active indoor air sampling for monitoring viruses [11–14]. This method has successfully identified a variety of viruses, including respiratory and enteric viruses (e.g. respiratory syncytial virus (RSV), influenza C, human parainfluenza virus) in community settings.

The methodologies employed to detect pathogens in environmental surveillance are usually similar to those employed in clinical diagnostics [1,15], frequently using molecular based single- or multi-targeted (qPCR or multiplex qPCR) tests. Targeted test methods tend to have high sensitivity and specificity. Previously, some studies have focused on the detection of viruses and bacteria with targeted methods from indoor air of various community settings [11,13,14], revealing temporal trends of virus circulation among specific populations and highlighting the early warning potential of indoor air sampling in real-world situations. However, for more precise genetic identification or subtyping, additional techniques such as Sanger sequencing or whole genome sequencing (WGS) are often necessary. Characterising novel or variant viruses can be challenging with both methods, because they require specific genetic sequence knowledge of viruses of interest.

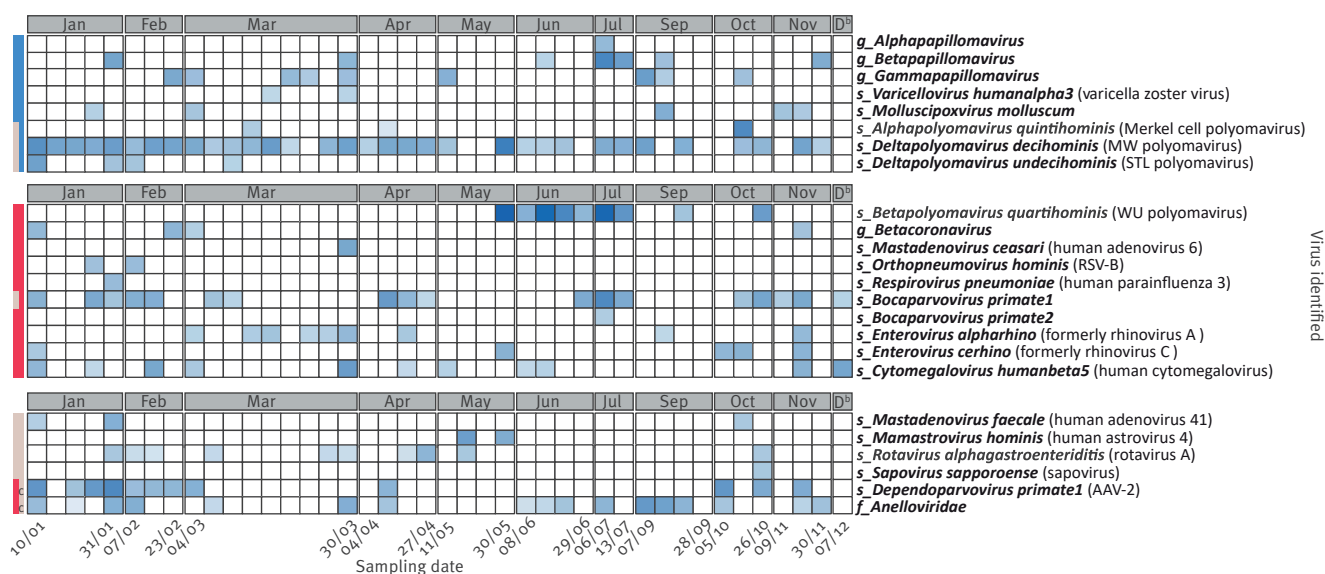
Next generation sequencing (NGS)-based ‘untargeted’ shotgun metagenomic methods can overcome this limitation. These methods often use a combination of viral enrichment, followed by deep sequencing using various NGS technologies and appropriate bioinformatic tools [16]. An established method, Novel Enrichment Techniques of Viromes (NetoVIR) [17], has proven effective for viral enrichment across various sample types, including faecal [18,19], insect [20,21], and plant samples [22], facilitating virome analysis with minimal bias and the discovery of novel viruses. This approach allows to identify pathogens and provide genomic insights as a ‘pathogen-agnostic’ tool [23]. For indoor air surveillance, NGS faces challenges such as low biomass and contaminating nucleic acid, despite its relatively cleaner nature compared with soil, wastewater or stool [24]. Nevertheless, Minor et al. used RNA virus-targeted NGS to analyse 20 samples from various indoor community settings. They detected not only respiratory viruses such as influenza A and C, but also rotaviruses and astrovirus, which are transmitted through the faecal–oral route [25].

In this study, air samples spanning a full year at a daycare centre were subjected to qPCR and untargeted metagenomics, to investigate the viruses they contained and better understand the indoor air virome and its dynamics. In addition, our objective was to explore the potential for viral surveillance of air sampling combined with metagenomics with reference-guided and *de novo* bioinformatics approaches.

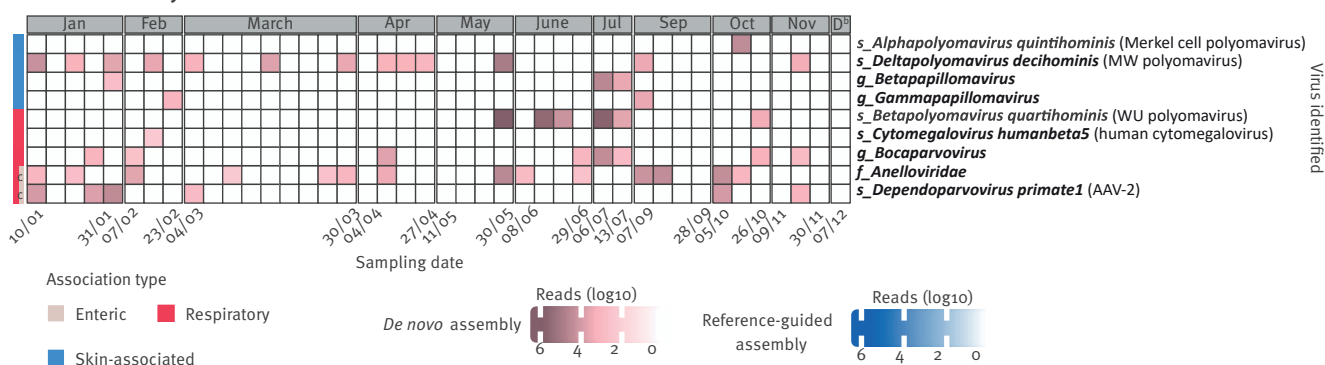
FIGURE 1

Human-associated viruses identified in indoor air samples of a daycare centre via shotgun metagenomics with (A) reference-guided and (B) *de novo* assembly, Leuven, Belgium, January–December 2022^a (n = 42 air samples shown)

A. Reference-guided assembly



B. De novo assembly



AAV-2: adeno-associated virus 2; f: family; g: genus; s: species.

^a The daycare centre was closed between mid-July and September, so no samples were collected during this period. Dates with no sampling including all the month of August 2022 are omitted from the graph.

^b D: December.

^c AAV-2 is known to infect a wide range of human cells and anelloviruses do not have a clear host cell, however both are reported to shed from the enteric and respiratory tracts.

On the X-axis, the first and last sample collection in each month is labelled.

Methods

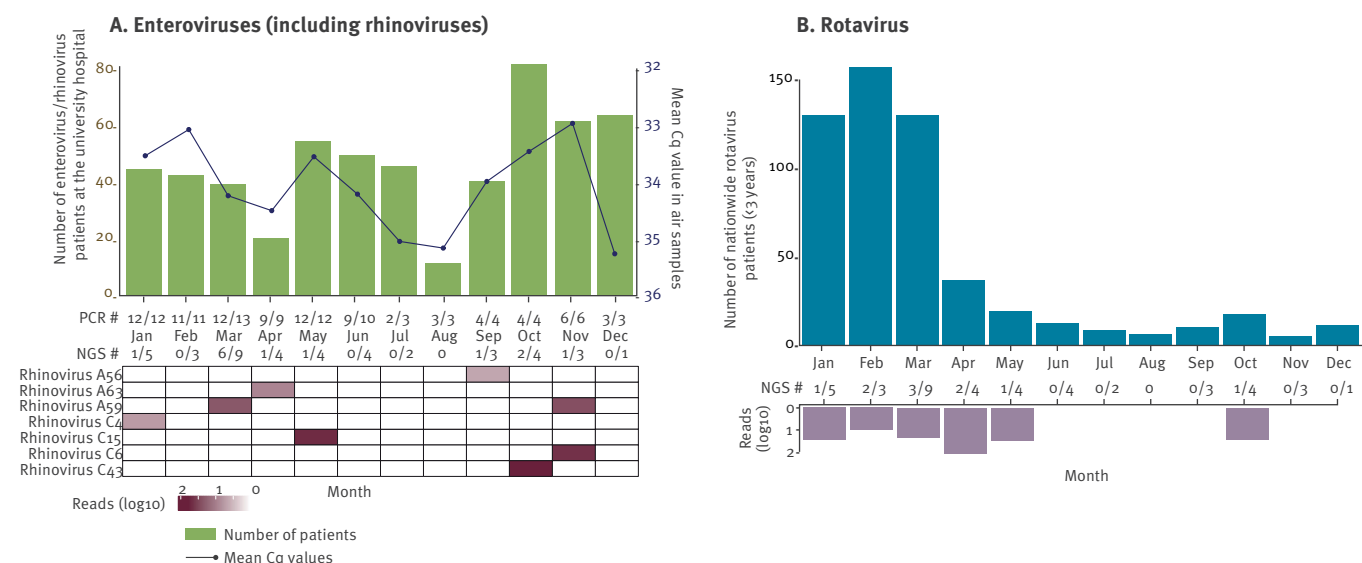
Setting and sample collection

Between January 2022 and December 2022, 90 indoor air samples were collected from a daycare centre in Leuven, hosting children aged 0 to 3 years old (with on an average day, 15 children and 3 adults aged ≥ 18 years old occupying the establishment).

We aimed to have from each month at least four samples (collected on a weekly basis), to explore different methods of virus detection and genetic analysis. However, after using parts of each sample for testing with the respiratory qPCR panel and initial optimisation of our assays, sufficient material was left to conduct metagenomic and enteric qPCR panel testing on 42 samples.

FIGURE 2

Monthly numbers of patients testing (A) enterovirus positive at the University Hospitals Leuven laboratory and (B) rotavirus positive at the NRC together with detection data for these respective viruses in a daycare centre air samples in Leuven, Belgium, 2022



Cq: quantification cycle; NGS: next generation sequencing; NRC: national reference centre; qPCR: quantitative PCR.

A. Chart with monthly enterovirus positive (including rhinoviruses) patients of all ages at the University Hospitals Leuven laboratory (green bars) and average qPCR Cq values per month (black line) measured from indoor air samples collected at the daycare centre between January–December 2022. The qPCR positivity rates are displayed on the X axis and below these the month of testing. Below the chart, a heatmap created using data from shotgun metagenomics of the daycare air is presented. Above the heatmap, the positivity rates of the NGS samples are presented for each month. B. Chart with monthly rotavirus cases under 3 years old in Belgium in 2022 coupled with an inverted bar plot of aligned reads of rotavirus A (log₁₀ scale) per month, obtained from our analyses. The monthly positivity rates obtained by NGS are displayed on the X axis of the inverted bar plot.

Sampling was done for 2 hours between 9.00 and 11.00 am in the changing room of the daycare centre which has the air extraction vents, and which is connected by two open doors to the main area where children are present. When the daycare centre was closed (e.g. holidays), no samples were collected.

For sampling, an AerosolSense active air sampler was used [11]. Ambient air was sampled with a rate of 200 L/min through a vertical collection pipe and impacted onto the AerosolSense Capture Media (Thermo Fisher Scientific, Waltham, Massachusetts (MA)). This method captures particles suspended in the air—which may include resuspended surface dust—while avoiding particles following a semi-ballistic trajectory.

Concentrations of carbon dioxide (CO₂) were measured using a remote climate sensor (Elsys ERC CO₂, Umeå, Sweden; placed next to the air sampler). Natural ventilation was also estimated by a Likert scale (no natural ventilation, one window open, door open, multiple windows open, door and window open).

Sample processing

After removing the standard cartridges from the sampler, samples were transported to the laboratory on the same day that they were collected. Samples were dissolved in 2 mL of universal transport medium (UTM),

to be used for respiratory and enteric multiplex qPCR panels as well as metagenomics (see below). Samples were stored at 4 °C until processing with the respiratory multiplex qPCR panel, which was conducted, whenever possible, on the same day as the sample collection. If storage was required over the weekend, samples were frozen at –80 °C. For the respiratory panel, 200 µL of each sample was used. The remaining volume was immediately stored at –20 °C until proceeding with metagenomics and the enteric qPCR panel, which occurred in February 2023, on samples which had undergone two or less freeze–thaw cycles. As negative controls, UTM (used to dissolve samples) was processed in parallel and sequenced as deep as samples.

Respiratory and enteric multiplex qPCR panels

Samples dissolved in UTM were subjected to multiplex qPCR testing. The detailed protocols for these procedures have been previously described [11,26]. The respiratory panel is designed to detect 22 viruses and seven bacterial pathogens and was applied on 90 samples when they were collected. In addition, the enteric panel detects 17 non-viral pathogens (4 parasites and 13 bacteria) and six viruses, specifically human adenovirus F40/F41, norovirus GI, norovirus GII, rotavirus A, astrovirus, and sapovirus (genogroups GI, GII, GIV, and GV). The enteric panel was applied to the same 42 samples used for shotgun metagenomics (see below).

Shotgun metagenomics

Samples were subjected to NetoVIR [27]. They were homogenised by vortexing for 10 seconds and homogeneous samples were centrifuged at 17,000g for 3 min and filtered with a 0.8 µm polyethersulfone (PES) centrifugal filter to eliminate eukaryotic cells and bacteria. Free floating nucleic acids were removed by benzonase and micrococcal nuclease while being incubated for 2 hours at 37°C. A modified Whole Transcriptome Amplification 2 (WTA2) kit was used to amplify both DNA and RNA. To obtain DNA amplification, the initial denaturation step was performed at 95°C instead of 70°C. Library preparation was done using a Nextera XT DNA Sample Preparation kit and sequencing was performed on a NovaSeq platform (Illumina).

Clinical data collection

At the daycare centre — where samples were collected — parents were advised to keep their children at home if they exhibited any symptoms of illness such as fever, runny nose, or diarrhoea. This led to our study population consisting of seemingly ‘healthy’ or asymptomatic individuals.

We collected the results of respiratory panel tests for enterovirus (including rhinovirus) performed at the clinical laboratory of University Hospitals Leuven during the same period as the air sample collection. The PCR targets also included non-rhinovirus enteroviruses. Clinical respiratory qPCR panels are performed according to specific clinical indications. In immunocompetent individuals, these panels are used for respiratory infections requiring intensive care admission or those unresponsive to initial therapy for a respiratory infection. For immunocompromised patients, respiratory panels are more commonly used when lower respiratory tract infections are suspected.

For rotavirus A infections in the age group of 0 to 3 years old, we used data from the national reference centre (NRC) for rotavirus in Leuven, Belgium in 2022. Samples of the NRC were collected through the nationwide laboratory system on a voluntary basis and under the regulations defined by the Royal Decree of 09/02/2011 [28].

Bioinformatics analyses

We employed reference-guided assembly and *de novo* assembly on the reads obtained from sequencing. Analyses were conducted on the Flemish Supercomputer Center ‘Vlaams Supercomputer Centrum’ (VSC) KU Leuven servers.

Reference-guided assembly

To identify human and animal-associated viruses, we used EsViritu [29] (v0.2.3) and nucleotide (nt) basic local alignment tool (BLASTN). EsViritu employs a database of all human and animal viruses in GenBank up to 16 November 2022 (EsViritu DB v2.0.2), curated to include one representative genome per 95% average nt identity cluster. Reads underwent quality control,

adapter trimming, and mapping to reference sequences with 90% identity and 90% coverage. Viruses were considered present if the consensus sequence exceeded 200bp or 10% of the reference genome and at least two reads mapped to it. Consensus sequences were obtained using Samtools consensus [30] and subsequently verified using BLASTN. Reference-guided assembly only showed human papillomavirus (HPV)-21 with 39 reads in one negative control, which resulted in a 236-nt long consensus sequence, found to be identical to that from a higher viral load sample with the same virus, and was therefore classified as cross-contamination from this sample.

De novo assembly

For *de novo* assembly of viral genomes (or contigs), we used the Virome Paired-End Reads pipeline (ViPER) [31]. The pipeline processed quality-controlled and adapter-trimmed reads, performing *de novo* triple-assembly using metaSPAdes [32]. Contigs were then clustered at 95% identity and 85% coverage to reduce redundancy, first within individual samples and then among all samples. Reads were mapped back to contigs >1 kbp, considering those with over 50% horizontal coverage present in each sample. Contig identification was subsequently performed using DIAMOND [33] for eukaryotic viruses, geNomad [34] for bacteriophages, and BLASTN for identification of genetic material associated to other entities (i.e. fungi, plant, human). For *de novo* assembly, all contigs present in negative controls were treated as contaminants and removed before downstream analyses.

Statistical analyses

Statistical analyses for the difference of detections in the daycare’s open and closed window periods of non-human infecting and human-infecting viruses were conducted using the chi-square test and performed in RStudio. The analysis assessed the association between seasonality and the presence of these viruses. A p value of less than 0.05 was considered statistically significant.

Results

On average, 25 million reads were produced per sample. The findings based on their analyses by both EsViritu (thereafter referred to as ‘reference-guided assembly’), and ViPER, (referred to as ‘*de novo* assembly’) are further described together unless otherwise specified.

Respiratory, enteric and skin viruses revealed by untargeted shotgun metagenomics in indoor air

Using reference-guided assembly, we identified human-associated viruses in 40 of 42 samples; the list of viruses identified in each sample with read counts can be found in the Supplementary information 1. *De novo* assembly identified human-associated viruses in 29 of 42 samples (Figure 1). Several respiratory viruses were detected during the study period, including human bocavirus 1, 2 and 3 (species *Bocaparvovirus*

FIGURE 3

(A–N) Sequencing coverage plots of genetic sequences of viruses detected in indoor air samples of a daycare and phylogenetic analyses of sequences of (O) two human bocaviruses and (P) one RSV identified in some samples, Belgium, January–December 2022



HBov: human bocavirus; NSP4: non-structural protein 4; RSV: respiratory syncytial virus.

A–N. Coverage plots of viruses found in indoor air samples. The Y-axis for each panel is different and shows the depth of coverage, while the X-axis shows nt positions. O–P. Phylogenetic trees of two human bocaviruses and RSV found in indoor air samples. Trees are constructed by the maximum likelihood method and numbers above the branch points denote the confidence level of the relationship of the paired sequences determined by bootstraps ($n=100$). Branch lengths are measured in substitutions per site (e.g. the 0.02 scale bar indicates ~2% sequence divergence).

primate1–2, number of detections (n_{dt})=18), *Enterovirus alpharhino* and *Enterovirus cerhino* (formerly *Rhinovirus A* (n_{dt} =9) and C (n_{dt} =5)), human adenovirus 6 (n_{dt} =1), viruses belonging to the genus *Betacoronavirus* (human coronavirus HKU1, n_{dt} =3 and OC43, n_{dt} =1) and RSV-B (n_{dt} =2). WU polyomavirus was consistently detected from the end of May until the summer holiday closure in July (n_{dt} =7 during that period), and it accounted for the highest total number of human-infecting virus reads in our study. Among the identified respiratory viruses, we obtained complete genomes of WU polyomavirus and human bocavirus 1.

Enteric virus reads detected in the air included rotavirus A (n_{dt} =10), human astrovirus 4 (n_{dt} =2), sapovirus (n_{dt} =1), and human adenovirus 41 (n_{dt} =3). The highest horizontal genome coverage from these strictly enteric pathogens was reached for human astrovirus 4 with more than 70% coverage. While these pathogens are associated with gastroenteritis or respiratory disease in young children [35,36], other viruses identified in the current study, such as MW polyomavirus [37], STL polyomavirus [38,39], human bocavirus 1–3 [40,41], anelloviruses [42], and adeno-associated virus-2 (AAV-2) [43] are described in stool without being associated with disease. Anelloviruses were detected in 22 samples, and AAV-2 was identified in 12 samples. In five of the 22 samples testing positive for anelloviruses, these viruses were only identified through *de novo* assembly, highlighting their high genetic variability, which may have complicated detection by reference-based approaches.

Notably, the most frequently identified virus was MW polyomavirus, present in 33 of the 42 samples collected throughout the year. This virus showed 98.8% nt similarity to a strain reported to be shedding from the skin of healthy individuals in the United States (US) [44]. Several skin-associated polyomaviruses were detected, including STL polyomavirus and Merkel cell polyomavirus [45]. Additionally, varicella zoster virus (VZV) and variety of HPVs (18 different types, including HPV-151, HPV-120, HPV-129, HPV-9, and HPV-76, detected in 14 samples) were identified.

Seasonality of respiratory viruses was reflected in the air

Examination of the seasonality of respiratory viruses with shotgun metagenomics revealed distinct patterns. Rhinoviruses were detected throughout the year, showing no clear seasonal pattern. Bocaparvoviruses were identified frequently in blocks of several consecutive positive samples (mid-January, April, end-June, mid-October), presumably representing outbreaks. In contrast, viruses with well-established seasonal patterns, such as human coronaviruses and RSV, typically peak during the winter months [46,47] and were only detected in the months of January, February, and November in our study (Figure 1A).

To increase the resolution for seasonality analyses, we used qPCR respiratory panel results from 90 samples (the same 42 samples as above, plus 48 additional indoor air samples from the same year and location, see methods). RSV was detected only from January to April (n_{dt} =12, quantification cycle (Cq) range: 32.0–39.1) and November to December (n_{dt} =7, Cq range: 32.9–35.3). Similarly, human parainfluenza virus 3 was found from January to March (n_{dt} =12, Cq range: 33.5–38.7) and once in May (n_{dt} =1, Cq: 35.3). Human coronaviruses (HKU-1 and OC43) were detected 42 times, mostly between January and March (n_{dt} =31, Cq range: 27.9–37.8) and in November to December (n_{dt} =6, Cq range: 27.9–37.8). Furthermore, the respiratory qPCR panel and shotgun metagenomics did not detect any SARS-CoV-2, whereas a more sensitive TaqPath assay detected SARS-CoV-2 in 45 of 90 samples; qPCR results for respective viruses can be found in Supplementary information 2. S1.

Indoor air surveillance as a non-invasive way to follow temporal trends of rhinovirus and rotavirus A

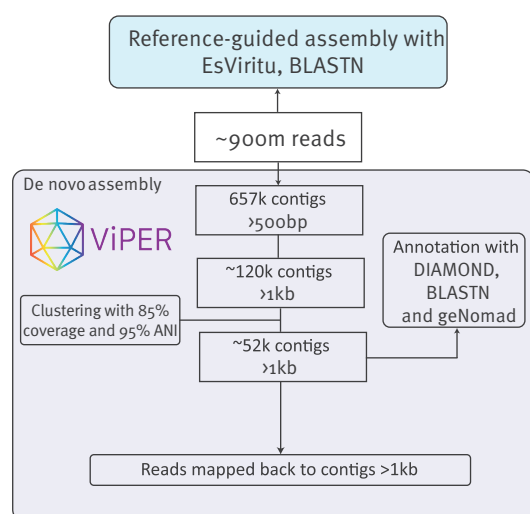
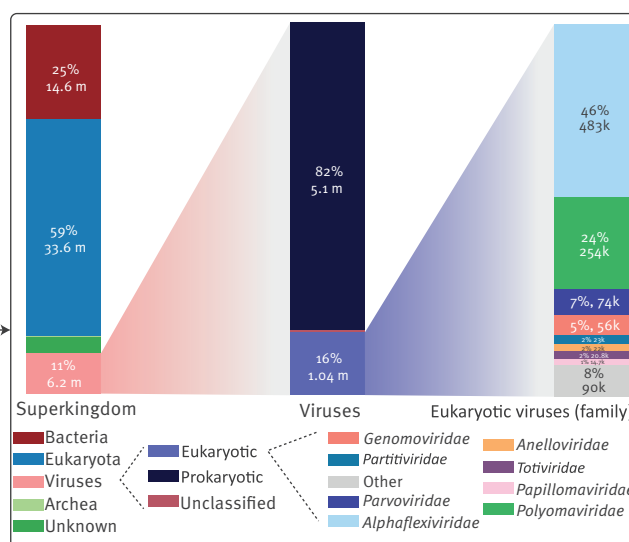
A recent study revealed that the number of infant hospitalisations due to human rhinoviruses represent up to two-thirds of those due to RSV [48], while rotavirus remains the most common gastroenteritis cause in children younger than 5 years old worldwide. In our data RSV was only identified in two indoor samples using shotgun metagenomics, possibly because infants with more severe disease were kept away from the daycare centre. On the other hand, rotavirus and rhinoviruses were consistently detected (Figure 1A), and therefore, they were selected for a more thorough comparison to other openly available surveillance data.

For rhinovirus epidemiology, data from patients of all age with respiratory symptoms from the University Hospitals Leuven, who tested positive for enteroviruses (which includes rhinoviruses; see methods) using the respiratory qPCR panel, were used as a proxy (Figure 2A, brown bars). The same qPCR panel was applied to all indoor air samples (n =90), revealing a 96.7% positivity rate (87/90) (indicated with 'PCR #' in Figure 2A). Shotgun metagenomics identified rhinoviruses in 13 of the 42 samples and not every month, which is a lower positivity rate compared with qPCR panel tests. The inverse average Cq qPCR values per month showed similar trend to clinical cases between January–March, May–July and September–November. Despite low read counts (read counts can be found in Supplementary Information 1), our shotgun metagenomics data allowed identification and subtyping of rhinoviruses (Figure 2A, heatmap).

To compare indoor air monitoring data with actual rotavirus A circulation measured by surveillance, nationwide data from the NRC for rotavirus A (see methods) among children younger than 3 years old were used (Figure 2B). Unlike rhinoviruses, rotavirus A was not identified in daycare air using the qPCR enteric panel

FIGURE 4

(A) Bioinformatics pipelines for identifying viruses by untargeted metagenomics in indoor air samples of a daycare centre and (B) overview of obtained reads mapped back to contigs, Leuven, Belgium, January–December 2022

A. Bioinformatics pipelines**B. Overview of reads**

ANI: average nucleotide identity; BLASTN: nucleotide (nt) basic local alignment tool; VIPER: virome paired-end reads pipeline.

(0/42), but it was detected in 10 samples (January to May and October) using untargeted shotgun metagenomics. These detections corresponded well with nationwide data and the seasonality reported by the NRC for rotaviruses in Belgium (Figure 2B).

Untargeted shotgun metagenomics allows detailed phylogenetic analyses, and discrimination of distinct viral variants

In addition to the detection (absence or presence) of a particular pathogen, further molecular characterisation of viral genomes, such as subtyping or identifying mutations responsible for resistance against therapies can be important for public health purposes. Therefore, we aimed to further leverage the obtained NGS data for the molecular characteristics of viral genomes found in indoor air samples. The data allowed us to reconstruct several partial and (near) complete genomes of several skin-associated (Figure 3A–C) respiratory (Figure 3D–F) and enteric viruses (Figure 3G–I) as well as multiple near-complete genomes of AAV-2 (Figure 3J). Moreover, various animal (bird and cat) viruses (Figure 3K–M) and a complete genome of a novel densovirus were identified (Figure 3N). These obtained genomic data can further be used for phylogenetic analysis, as exemplified for the partial RSV genome data and the near-complete genomes of bocaviruses to reliably study their taxonomy and evolutionary relationships (Figure 3O–P). The partial genome of RSV clustered closely with RSV-B sequences, whereas bocavirus genomes showed over 99% similarity to human bocavirus 1 (OP255990) and human bocavirus 3 (NC_012564), reference sequences. Further examples of phylogenetic

analyses for AAV-2, WU and MW polyomaviruses, HPV-151, and the partial genomes of felis domesticus papillomavirus, a novel densovirus, a divergent canary polyomavirus, and human astrovirus 4, are provided in Supplementary information 2 S2.

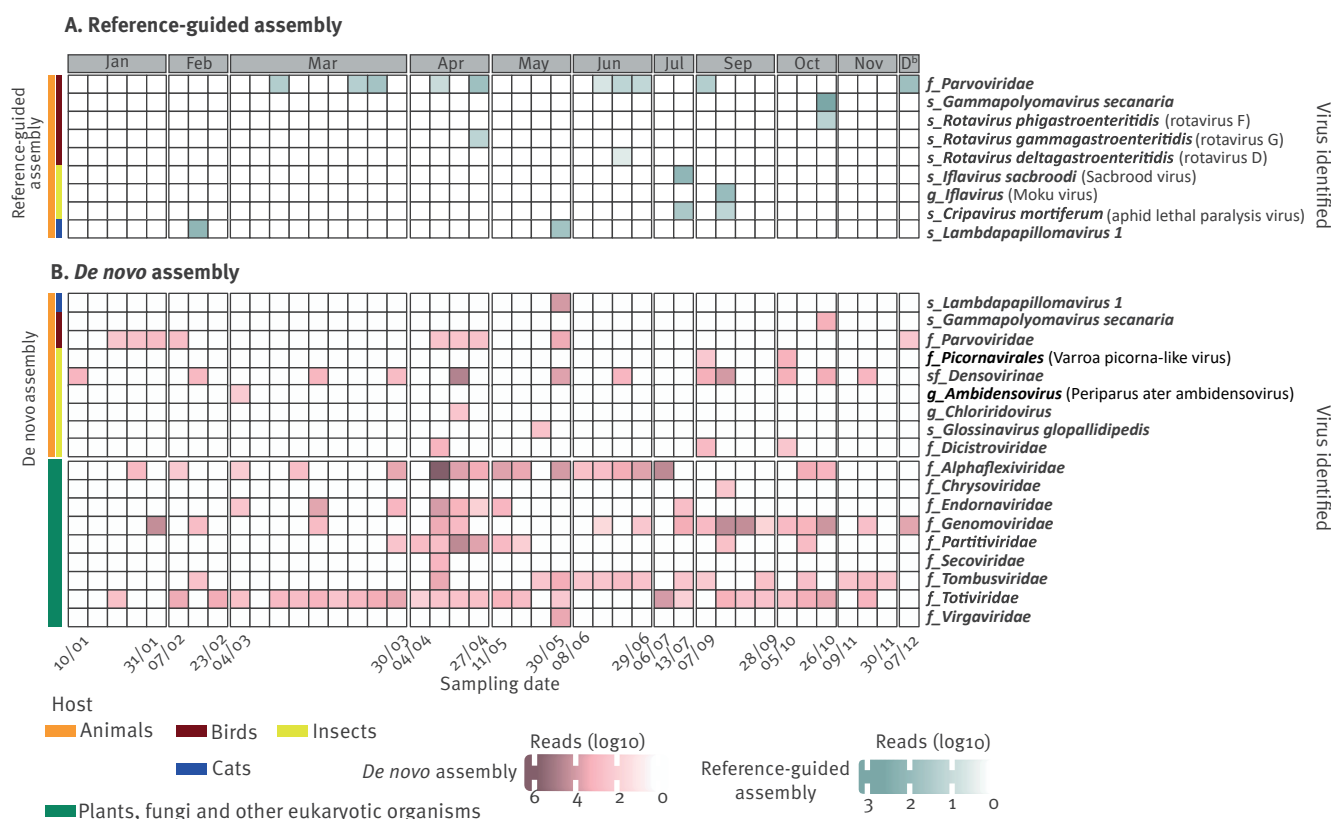
Interestingly, we detected WU polyomavirus in the daycare centre for the first time at the end of May, and it was consistently present for nearly 2 months with high abundance, until the start of the summer break (Figure 1A). This prompted us to investigate whether this prolonged detection was due to viral shedding of a single strain by one or more individuals or if there were multiple introductions of distinct viruses. Therefore, we reconstructed four near-complete genomes of WU polyomavirus from four samples collected between the end of May and July (30 May, 15 June, 22 June and 6 July). We show that the viral genomes collected at the first two time points are identical, whereas the third sample showed 12 mutations, suggesting the introduction of a second WU polyomavirus variant into the daycare centre (mutations in different variants and read level analysis can be found in Supplementary information 2. S3). Finally, our in-depth read-level analysis suggests that both of these variants were present in the fourth sample.

Untargeted shotgun metagenomics on indoor air samples can reflect temporal dynamics of both the indoor and outdoor environment

Next, we focused on the air virome linked to other entities such as plants, animals (including insects), and fungi. Although contigs derived from bacteriophages

FIGURE 5

Detections of viruses with animal (except humans), plant, fungi or other eukaryotic hosts by (A) reference-based assembly or (B) *de novo* assembly in indoor air samples from a daycare centre, Belgium, January–December 2022^a (n = 42 samples)



f: family g: genus, s: species, sf: subfamily.

^a The daycare centre was closed between mid-July and September, so no samples were collected during this period. Dates with no sampling including all the month of August 2022 are omitted from the graph.

^b D: December.

On the X-axis, the first and last sample collection in each month is visible and the Y-axis represents viruses at taxonomical level.

(prokaryotic viruses) were also abundantly identified, due to their very short average contig length (ca 1kb), we did not further explore them (Figure 4).

Untargeted deep sequencing, combined with reference-based assembly and *de novo* assembly (Figure 4), allowed us to detect insect and animal-associated viruses in 27 of 42 samples (Figure 5). Avian species-infecting viruses, such as avian rotaviruses (detected in April, June, and October), and members of the *Gammaherpesvirus secanaria* (October) were detected. Additionally, honeybee-related viruses, such as Sacbrood virus and Moku virus (Figure 5), were identified in July and September, along with novel and divergent insect infecting densovirus throughout the year (in 12 of 42 samples). As expected, divergent (and novel) densovirus were only identified by *de novo* assembly, while viruses with low read counts but high similarity to known references, such as avian

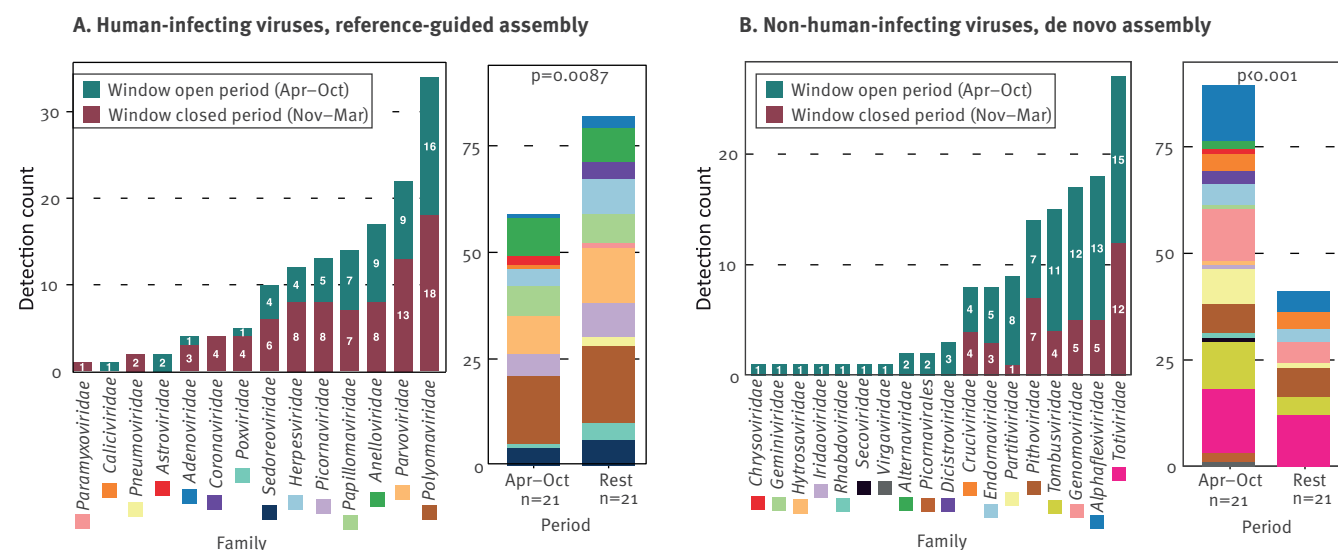
rotaviruses, were only identified by reference-based assembly.

The daycare centre typically kept its windows open from April to October (metadata can be found as Supplementary information 2 S4). Therefore, we categorised detections into two periods: April–October and the rest of the year. This categorisation revealed a significantly higher prevalence of human-infecting viruses during the closed-window periods, while non-human-infecting viruses were more prevalent during the open-window period ($p < 0.01$, Figure 6).

Additionally, we detected totiviruses, matching with *Malasseziaceae* (fungi) reads, suggesting that these viruses were *Malasseziaceae*-associated, as further explained in Supplementary information 2 S5. We conducted correlation coefficient analyses on environmental surveillance data to investigate whether any

FIGURE 6

Number detections by untargeted metagenomics of different viruses in the indoor air of a daycare, when windows were open or closed, stratified according to (A) human- and (B) non-human infecting viruses, Belgium, January–December 2022



df: degrees of freedom.

A. Number of detections of human-infecting viruses for each family (panel A: graph on the left) during period with open or closed windows. Aggregated detection counts of human-infecting viruses for the periods between April–October and rest of the year (panel A: graph on the right), $X^2=6.86$, $df=1$, $p=0.0087$. B. Number of detections of non-human-infecting viruses for each family (panel B: graph on the left). Aggregated detection counts of non-human-infecting viruses for the periods between April–October and rest of the year (panel B: graph on the right), $X^2=33.985$, $df=1$, $p<0.001$.

environmental parameters could serve as proxy for the species identified per sample, which did not reveal any significant results ($p>0.05$, analyses and further explanation can be found in Supplementary information 2 S6).

Discussion

In this study, we employed longitudinal indoor air sampling and untargeted viral metagenomics over a 1-year period at a daycare centre. We used the NetoVIR protocol to enrich in virus-like particle (VLP), combined with short-read deep sequencing, to enhance our ability to detect low-abundance viral species with greater sensitivity. Throughout the 12-month period, we successfully uncovered a diverse array of human-associated viruses, including respiratory, enteric, and skin viruses, alongside non-human-associated viruses, which were linked to animals, plants, and fungi. Additionally, we conducted an in-depth bioinformatics analysis, combining *de novo* assembly and reference-guided assembly techniques. These methods led us to uncover not only known viruses with very low abundance, but also novel and divergent viruses, making it potentially useful to identify ‘pathogen X’—an epidemic or pandemic viral pathogen that is not yet characterised [49].

Rosario et al. previously observed Merkel cell polyomavirus, MW polyomavirus and papillomaviruses in air filter samples collected in dormitory rooms [50].

Similarly, polyomaviruses were the most abundant human viruses revealed through our work, including WU polyomavirus, a respiratory virus reported in immunocompromised people and in children with minimal or no respiratory symptoms [51,52]. We also detected MW polyomavirus as well as Merkel cell polyomavirus, both also prior documented among asymptomatic individuals [53,54]. Interestingly, alongside the human viruses we found, the most abundant fungal viruses were related to commensal human skin fungi from the family *Malasseziaceae*, providing evidence for the high abundance of human skin-associated microorganisms in the environment. Therefore, our findings suggest a potential of indoor air sampling for surveillance of skin-associated viruses.

Several recent studies have investigated the potential of indoor air surveillance to monitor specific viral pathogen dynamics and epidemiology. These were conducted in settings such as hospitals, offices, and residential environments [10,11,13,25,55–58]. All of them relied on targeted tests requiring specific viral genomic knowledge, while we employed a non-targeted approach. The longitudinal design of our investigation, in combination with available epidemiological data from the University Hospitals Leuven and NRCs allowed us to use rhinovirus and rotavirus as examples to showcase the power of air sampling in epidemiological surveillance, especially for asymptomatic

circulation in a setting with no clinically ill individuals. In addition, our NGS-based untargeted methodology allowed us to subtype rhinoviruses, showing the co-circulation of various rhinoviruses throughout the year. In a non-longitudinal study, Minor et al. examined indoor air from multiple locations with metagenomics, detecting influenza C in a preschool and providing genomic data for this frequently untested virus [25]. We expanded their approach by seeking multiple complete or partial genomes of viruses, which may be associated with respiratory (e.g. WU polyomavirus) or skin diseases but are not currently included in the routine qPCR panels for respiratory diseases in hospitals. Additionally, we were able to identify and subtype several viruses (e.g. bocaviruses, RSV) and distinguish the co-circulation of closely related virus variants (WU polyomavirus) over time. This underscores the possible simultaneous use of shotgun metagenomics to determine the presence, frequency and changing genomic/epidemiological characteristics of multiple viruses.

Beyond surveillance of human viruses, indoor air and environmental surveillance with metagenomics could also be of interest for animal farms and greenhouses. Kwok et al. analysed dust samples and pooled chicken faeces from farms using shotgun metagenomics, identifying multiple viruses in dust that matched those found in the faeces from the same location [59]. In our study, we unexpectedly detected avian viruses (rotaviruses and a divergent polyomavirus), a cat papillomavirus, and several insect-infecting viruses, including a novel densovirus, in an animal-free environment. We suggest that the animal and plant viruses mentioned are highly resilient in the environment and may be transported indoors by insects, air currents, or even on human skin or garments that can subsequently shed them. Additionally, the presence of multiple plant viruses demonstrates that indoor air can serve as an effective medium for detecting pathogens that typically require labour-intensive sampling from various crops. Therefore, we propose that indoor air surveillance in strategic animal and plant production locations can be powerful for revealing epidemiological trends and potential outbreaks of viruses.

It was previously reported that indoor air shotgun metagenomics with other methods resulted in 1% to 3.2% of reads that could be assigned to viruses [60], underscoring the low viral biomass of indoor air samples. However, using the NetoVIR protocol, 10.9% of reads mapped to viral contigs, primarily phages, aligning with earlier findings [61]. Interestingly, shotgun metagenomics identified a broader diversity of enteric viruses, such as rotavirus and sapovirus, compared with the enteric qPCR panel, whereas the respiratory qPCR panel detected rhinovirus, RSV, adenovirus, and parainfluenza more consistently than metagenomics. These differences likely stem from the distinct sensitivities of the qPCR panels, which are tailored for detecting specific viral loads in defined sample types. The viral load in indoor air is likely at or near detection

thresholds, with commonly circulating human viruses comprising less than 0.005% of total reads, as reported by Prussin et al. [61]. In our study, we confirm this low viral load of human-associated viruses in indoor air apart from skin- and respiratory-associated polyomaviruses.

Given the longitudinal nature of our sample collection, it is possible to discuss seasonality of viruses. Prussin et al. employed passive indoor air sampling and shotgun metagenomics throughout the year in a daycare centre in Virginia, US. They found human-associated viruses to be much more diverse and dominant in the winter, while the summertime virome contained a higher proportion and diversity of plant-associated viruses [61]. Similarly, we also observed a higher prevalence of animal- or plant-associated viruses, particularly during the warmer months from April to October when windows were open and respiratory and enteric viruses mostly during the wintertime when windows were closed. Therefore, we suggest that active indoor air sampling can reflect dynamics of both the inside and outside environment. On the other hand, human parainfluenza virus 3 was detected in winter by indoor air sampling at the daycare centre, which is different than its defined peak in the literature, late spring and early summer.

Overall, untargeted viral metagenomics offers a strong and complementary tool for pathogen surveillance via indoor air, particularly in cases of suspected asymptomatic shedding. In practice, we would propose a surveillance system where samples are collected at regular intervals from a small number of strategically selected sentinel locations—ideally in ventilation shafts to capture air from multiple indoor spaces in each sample while further minimising the invasiveness of such sampling approach for the community [14]. This approach can leverage established specialised laboratory infrastructure and standardised workflows; with declining sequencing costs due to pooled/batched sample processing. This could result in early signals of beginning epidemic trends of seasonal, emerging and re-emerging pathogens, while at the same time providing genetic information for further characterisation. Timely information collected on viral (co-) circulation at the community level can be used to prepare the healthcare system for viral circulation. Moreover, metagenomics enables detailed genomic analyses and identification variants, which may eventually serve as a proxy for estimating the number of infected hosts or introductory events at a specific location. Future research should validate this approach at sentinel locations for skin, enteric, and respiratory viruses in targeted settings, ideally incorporating clinical data of individuals.

Our study has several limitations. Primarily, it was conducted at a single location, without access to information regarding the health status or medical history of individuals in the vicinity. Nonetheless, it was assumed

that individuals were unlikely to exhibit severe respiratory or gastroenteritis-related symptoms, as it was recommended that families keep children at home if they displayed such symptoms. Secondly, we could not acquire complete genomes for numerous viruses that were present, which may stem from some methodological constraints of sequencing and low viral loads in the air. Additionally, our results are not confirmed by culture-based techniques to determine if the detected viral genomes are from infectious viruses or not. This would be important information, with respect to the potential of airborne transmission of skin-associated and enteric viruses, since dose-risk relationships of many viruses are not well understood. On the other hand, we employed a longitudinal approach over a 12-month period, allowing to monitor seasonal dynamics and use advanced viral enrichment methods with comprehensive analyses, which enhanced our ability to detect even low-abundance viral species with greater sensitivity.

Conclusion

Our study highlights that indoor air sampling employed with viral metagenomics is a robust method for surveillance of both human-associated and non-human-associated viruses. We suggest that this method, implemented in sentinel locations and targeted populations, can be complementary to traditional patient-based surveillance methods.

Data and code availability

Raw data, excluding human reads, is available at the Sequence Read Archive (SRA) with Bioproject number PRJNA1158979. Accession numbers for respective sequences given in the manuscript originate from GenBank and can be found at <https://www.ncbi.nlm.nih.gov/genbank/>. Codes used to conduct the analyses and produce figures are available on https://github.com/Matthijnssenslab/Indoorair_daycare.

Authors' contributions

Conceptualisation was carried out by JM, EA, EK, CG and MK. The project was managed by EK, LB, JR and EA. CE, BC, MK, EK, CG performed sample collection and processing. Data analysis, visualisation, and drafting of the initial manuscript were conducted by MK. Supervision was provided by JM, EA, and EK. All authors reviewed and approved the final manuscript.

Conflict of interests

The authors declare no competing interests.

Funding statement

Mustafa Karatas is supported by a Research Foundation Flanders (FWO) fundamental research scholarship (number: 11P7124N). Joren Raymenants acknowledges support of the Research Foundation Flanders (FWO, grant number: 1S88721N). C.G. and E.A. acknowledge support from the DURABLE project. The DURABLE project has been co-funded

by the European Union, under the EU4Health Programme (EU4H), project no. 101102733. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. The computing power in this work was provided by the VSC (Flemish Supercomputer Centre), financed by the FWO and the Flemish government–department EWI.

Ethical statementGo to section

The study was approved by the Ethics Committee Research UZ/KU Leuven (S66518, B3222022000873). No informed consent was required from occupants of the sampled environments.

Use of artificial intelligence tools

ChatGPT-4o has been used for proofreading purposes.

Preprint

Manuscript is deposited in the preprint server medRxiv (<https://www.medrxiv.org/content/10.1101/2024.10.11.24315306v1>).

Acknowledgements

We would like to thank Pieter Wets and Steven Traets for their help in collection of samples. Additionally, we thank the Rotavirus National Reference Center and National Reference Center for Respiratory Pathogens for their support throughout this research as well as Sciensano (Scientific Institute of Public Health, Belgium) for its support in reference centers.

References

1. Abat C, Chaudet H, Rolain JM, Colson P, Raoult D. Traditional and syndromic surveillance of infectious diseases and pathogens. *Int J Infect Dis*. 2016;48:22-8. <https://doi.org/10.1016/j.ijid.2016.04.021> PMID: 27143522
2. Leifels M, Khalilur Rahman O, Sam IC, Cheng D, Chua FJD, Nainani D, et al. The one health perspective to improve environmental surveillance of zoonotic viruses: lessons from COVID-19 and outlook beyond. *ISME Commun*. 2022;2(1):107. <https://doi.org/10.1038/s43705-022-00191-8> PMID: 36338866
3. Vieira CB, de Abreu Corrêa A, de Jesus MS, Luz SLB, Wyn-Jones P, Kay D, et al. Viruses Surveillance Under Different Season Scenarios of the Negro River Basin, Amazonia, Brazil. *Food Environ Virol*. 2016;8(1):57-69. <https://doi.org/10.1007/s12560-016-9226-8> PMID: 26783031
4. Iaconelli M, Muscillo M, Della Libera S, Fratini M, Meucci L, De Ceglia M, et al. One-year Surveillance of Human Enteric Viruses in Raw and Treated Wastewaters, Downstream River Waters, and Drinking Waters. *Food Environ Virol*. 2017;9(1):79-88. <https://doi.org/10.1007/s12560-016-9263-3> PMID: 27682315
5. Moresco V, Oliver DM, Weidmann M, Matallana-Surget S, Quilliam RS. Survival of human enteric and respiratory viruses on plastics in soil, freshwater, and marine environments. *Environ Res*. 2021;199:111367. <https://doi.org/10.1016/j.envres.2021.111367> PMID: 34029551
6. Kilaru P, Hill D, Anderson K, Collins MB, Green H, Kmush BL, et al. Wastewater Surveillance for Infectious Disease: A Systematic Review. *Am J Epidemiol*. 2023;192(2):305-22. <https://doi.org/10.1093/aje/kwac175> PMID: 36227259
7. Amman F, Markt R, Endler L, Hupfaut S, Agerer B, Schedl A, et al. Viral variant-resolved wastewater surveillance of SARS-CoV-2 at national scale. *Nat Biotechnol*. 2022;40(12):1814-22. <https://doi.org/10.1038/s41587-022-01387-y> PMID: 35851376
8. Jahn K, Dreifuss D, Topolsky I, Kull A, Ganesanandamoorthy P, Fernandez-Cassi X, et al. Early detection and surveillance of SARS-CoV-2 genomic variants in wastewater using COJAC. *Nat*

- Microbiol. 2022;7(8):1151-60. <https://doi.org/10.1038/s41564-022-01185-x> PMID: 35851854
9. Wastewater monitoring comes of age. *Nat Microbiol.* 2022;7(8):1101-2. <https://doi.org/10.1038/s41564-022-01201-0> PMID: 35918421
10. Dietz L, Constant DA, Fretz M, Horve PF, Olsen-Martinez A, Stenson J, et al. Exploring Integrated Environmental Viral Surveillance of Indoor Environments: A comparison of surface and bioaerosol environmental sampling in hospital rooms with COVID-19 patients. *medRxiv.* 2021. Preprint. <https://doi.org/10.1101/2021.03.26.21254416>
11. Raymenants J, Geenen C, Budts L, Thibaut J, Thijssen M, De Mulder H, et al. Indoor air surveillance and factors associated with respiratory pathogen detection in community settings in Belgium. *Nat Commun.* 2023;14(1):1332. <https://doi.org/10.1038/s41467-023-36986-z> PMID: 36898982
12. Ramuta MD, Newman CM, Brakefield SF, Stauss MR, Wiseman RW, Kita-Yarbro A, et al. SARS-CoV-2 and other respiratory pathogens are detected in continuous air samples from congregate settings. *Nat Commun.* 2022;13(1):4717. <https://doi.org/10.1038/s41467-022-32406-w> PMID: 35953484
13. Raymenants J, Van Gestel L, Coppens J, De Block T, Bangwen E, Rutgers J, et al. Detection of mpox virus in ambient air in a sexual health clinic. *Arch Virol.* 2023;168(8):210. <https://doi.org/10.1007/s00705-023-05837-z> PMID: 37486383
14. Happaerts M, Geenen C, Michiels J, Gorissen S, Swinnen J, Beuselinck K, et al. Centralised Air Sampling From a Ventilation System for the Surveillance of Respiratory Pathogens. *Indoor Air.* 2024;1(1):5176619. <https://doi.org/10.1155/2024/5176619>
15. Sinclair RG, Choi CY, Riley MR, Gerba CP. Pathogen surveillance through monitoring of sewer systems. *Adv Appl Microbiol.* 2008;65:249-69. [https://doi.org/10.1016/S0065-2164\(08\)00609-6](https://doi.org/10.1016/S0065-2164(08)00609-6) PMID: 19026868
16. Wang G, Li S, Yan Q, Guo R, Zhang Y, Chen F, et al. Optimization and evaluation of viral metagenomic amplification and sequencing procedures toward a genome-level resolution of the human fecal DNA virome. *J Adv Res.* 2023;48:75-86. <https://doi.org/10.1016/j.jare.2022.08.011> PMID: 35995413
17. Conceição-Neto N, Yinda KC, Van Ranst M, Matthijssens J. NetoVIR: Modular Approach to Customize Sample Preparation Procedures for Viral Metagenomics. *Methods Mol Biol.* 2018;1838:85-95. https://doi.org/10.1007/978-1-4939-8682-8_7 PMID: 30128991
18. Van Espen L, Bak EG, Beller L, Close L, Deboutte W, Juel HB, et al. A Previously Undescribed Highly Prevalent Phage Identified in a Danish Enteric Virome Catalog. *mSystems.* 2021;6(5):e0038221. <https://doi.org/10.1128/mSystems.00382-21> PMID: 34665009
19. Yinda CK, Rector A, Zeller M, Conceição-Neto N, Heylen E, Maes P, et al. A single bat species in Cameroon harbors multiple highly divergent papillomaviruses in stool identified by metagenomics analysis. *Viral Rep.* 2016;6:74-80. <https://doi.org/10.1016/j.virep.2016.08.001> PMID: 32289018
20. Shi C, Beller L, Deboutte W, Yinda KC, Delang L, Vega-Rúa A, et al. Stable distinct core eukaryotic viromes in different mosquito species from Guadeloupe, using single mosquito viral metagenomics. *Microbiome.* 2019;7(1):121. <https://doi.org/10.1186/s40168-019-0734-2> PMID: 31462331
21. Deboutte W, Beller L, Yinda CK, Maes P, de Graaf DC, Matthijssens J. Honey-bee-associated prokaryotic viral communities reveal wide viral diversity and a profound metabolic coding potential. *Proc Natl Acad Sci USA.* 2020;117(19):10511-9. <https://doi.org/10.1073/pnas.1921859117> PMID: 32341166
22. Ortega-Parra N, Zisi Z, Vogel E, Vos C, Hanssen I. First report of infections with isolates of the Asian genotype of cucumber green mottle mosaic virus in cucumber crops in Belgium. *Plant Dis.* 2024;108(4):1121. <https://doi.org/10.1094/PDIS-11-23-2532-PDN>
23. Gauthier NPG, Chorlton SD, Kraiden M, Manges AR. Agnostic Sequencing for Detection of Viral Pathogens. *Clin Microbiol Rev.* 2023;36(1):e0011922. <https://doi.org/10.1128/cmr.00119-22> PMID: 36847515
24. Prussin AJ 2nd, Marr LC, Bibby KJ. Challenges of studying viral aerosol metagenomics and communities in comparison with bacterial and fungal aerosols. *FEMS Microbiol Lett.* 2014;357(1):1-9. <https://doi.org/10.1111/1574-6968.12487> PMID: 24891293
25. Minor NR, Ramuta MD, Stauss MR, Harwood OE, Brakefield SF, Alberts A, et al. Metagenomic sequencing detects human respiratory and enteric viruses in air samples collected from congregate settings. *Sci Rep.* 2023;13(1):21398. <https://doi.org/10.1038/s41598-023-48352-6> PMID: 38049453
26. Hanne I, Engsbro AL, Pareja J, Schneider UV, Lisby JG, Pružinec-Popović B, et al. Multicenter evaluation of the new QIAstat Gastrointestinal Panel for the rapid syndromic testing of acute gastroenteritis. *Eur J Clin Microbiol Infect Dis.* 2019;38(11):2103-12. <https://doi.org/10.1007/s10096-019-03646-4> PMID: 31352670
27. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, et al. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep.* 2015;5(1):16532. <https://doi.org/10.1038/srep16532> PMID: 26559140
28. Federale Overheidsdienst Sociale Zekerheid. Koninklijk besluit tot vaststelling van de financieringsvoorwaarden van de referentiecentra voor humane microbiologie. [Royal Decree establishing the financing conditions of the reference centres for human microbiology]. Brussels: Federale Overheidsdienst JUSTITIE; Feb 2011. Available from: https://www.ejustice.just.fgov.be/cgi/article_body.pl?language=nl&pub_date=2011-03-01&numac=2011022071
29. Tisza M, Javornik Cregeen S, Avadhanula V, Zhang P, Ayvaz T, Feliz K, et al. Wastewater sequencing reveals community and variant dynamics of the collective human virome. *Nat Commun.* 2023;14(1):6878. <https://doi.org/10.1038/s41467-023-42064-1> PMID: 37898601
30. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience.* 2021;10(2):giab008. <https://doi.org/10.1093/gigascience/giab008> PMID: 33590861
31. De Coninck L. ViPER. Zenodo; 2023. [Accessed 26 Sep 2023]. Available from: <https://zenodo.org/record/5502203>
32. Prijbelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. *Curr Protoc Bioinformatics.* 2020;70(1):e102. <https://doi.org/10.1002/cpbi.102> PMID: 32559359
33. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods.* 2015;12(1):59-60. <https://doi.org/10.1038/nmeth.3176> PMID: 25402007
34. Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, et al. You can move, but you can't hide: identification of mobile genetic elements with geNomad. *Bioinformatics.* 2023. <https://doi.org/10.1101/2023.03.05.531206>
35. Tregoning JS, Schwarze J. Respiratory viral infections in infants: causes, clinical symptoms, virology, and immunology. *Clin Microbiol Rev.* 2010;23(1):74-98. <https://doi.org/10.1128/CMR.00032-09> PMID: 20065326
36. Biscaro V, Piccinelli G, Gargiulo F, Ianaro G, Caruso A, Caccuri F, et al. Detection and molecular characterization of enteric viruses in children with acute gastroenteritis in Northern Italy. *Infect Genet Evol.* 2018;60:35-41. <https://doi.org/10.1016/j.meegid.2018.02.011> PMID: 29438743
37. Siebrasse EA, Reyes A, Lim ES, Zhao G, Mkakosya RS, Manary MJ, et al. Identification of MW polyomavirus, a novel polyomavirus in human stool. *J Virol.* 2012;86(19):10321-6. <https://doi.org/10.1128/JVI.01210-12> PMID: 22740408
38. Vanchiere JA, Abudayyeh S, Copeland CM, Lu LB, Graham DY, Butel JS. Polyomavirus shedding in the stool of healthy adults. *J Clin Microbiol.* 2009;47(8):2388-91. <https://doi.org/10.1128/JCM.02472-08> PMID: 19494079
39. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, et al. Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. *Virology.* 2013;436(2):295-303. <https://doi.org/10.1016/j.virol.2012.12.005> PMID: 23276405
40. Ong DSY, Schuurman R, Heikens E. Human bocavirus in stool: A true pathogen or an innocent bystander? *J Clin Virol.* 2016;74:45-9. <https://doi.org/10.1016/j.jcv.2015.11.027> PMID: 2665268
41. Kapusinszky B, Minor P, Delwart E. Nearly constant shedding of diverse enteric viruses by two healthy infants. *J Clin Microbiol.* 2012;50(11):3427-34. <https://doi.org/10.1128/JCM.01589-12> PMID: 22875894
42. Beller L, Deboutte W, Vieira-Silva S, Falony G, Tito RY, Rymenans L, et al. The virota and its transkingdom interactions in the healthy infant gut. *Proc Natl Acad Sci USA.* 2022;119(13):e2114619119. <https://doi.org/10.1073/pnas.2114619119> PMID: 35320047
43. Araujo NM. A dual role for adeno-associated virus in human health. *Viral J.* 2023;20(1):228. <https://doi.org/10.1186/s12985-023-02196-8> PMID: 37817259
44. Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. *Cell Host Microbe.* 2010;7(6):509-15. <https://doi.org/10.1016/j.chom.2010.05.006> PMID: 20542254
45. Bopp L, Wieland U, Hellmich M, Kreuter A, Pfister H, Silling S. Natural History of Cutaneous Human Polyomavirus Infection in Healthy Individuals. *Front Microbiol.* 2021;12:740947. <https://doi.org/10.3389/fmicb.2021.740947> PMID: 34733257

46. Aldridge RW, Lewer D, Beale S, Johnson AM, Zambon M, Hayward AC, et al. Flu Watch Group. Seasonality and immunity to laboratory-confirmed seasonal coronaviruses (HCoV-NL63, HCoV-OC43, and HCoV-229E): results from the Flu Watch cohort study. *Wellcome Open Res.* 2020;5:52. <https://doi.org/10.12688/wellcomeopenres.15812.2> PMID: 33447664
47. Ramaekers K, Keyaerts E, Rector A, Borremans A, Beuselinck K, Lagrou K, et al. Prevalence and seasonality of six respiratory viruses during five consecutive epidemic seasons in Belgium. *J Clin Virol.* 2017;94:72-8. <https://doi.org/10.1016/j.jcv.2017.07.011> PMID: 28772168
48. Horvat C, Casalegno JS, Masson E, Benveniste C, Haesebaert J, Paget J, et al. Contribution of Infant Rhinovirus Bronchiolitis to Hospital Bed and Ventilation Use. *JAMA Netw Open.* 2024;7(2):e2355033. <https://doi.org/10.1001/jamanetworkopen.2023.55033> PMID: 38324316
49. Simpson S, Kaufmann MC, Glozman V, Chakrabarti A. Disease X: accelerating the development of medical countermeasures for the next pandemic. *Lancet Infect Dis.* 2020;20(5):e108-15. [https://doi.org/10.1016/S1473-3099\(20\)30123-7](https://doi.org/10.1016/S1473-3099(20)30123-7) PMID: 32197097
50. Rosario K, Fierer N, Miller S, Luongo J, Breitbart M. Diversity of DNA and RNA Viruses in Indoor Air As Assessed via Metagenomic Sequencing. *Environ Sci Technol.* 2018;52(3):1014-27. <https://doi.org/10.1021/acs.est.7b04203> PMID: 29298386
51. Uda K, Koyama-Wakai C, Shoji K, Iwase N, Motooka D, Nakamura S, et al. WU polyomavirus detected in children with severe respiratory failure. *J Clin Virol.* 2018;107:25-8. <https://doi.org/10.1016/j.jcv.2018.08.003> PMID: 30114678
52. Kuypers J, Campbell AP, Guthrie KA, Wright NL, Englund JA, Corey L, et al. WU and KI polyomaviruses in respiratory samples from allogeneic hematopoietic cell transplant recipients. *Emerg Infect Dis.* 2012;18(10):1580-8. <https://doi.org/10.3201/eid1810.120477> PMID: 23017213
53. Abed Y, Wang D, Boivin G. WU polyomavirus in children, Canada. *Emerg Infect Dis.* 2007;13(12):1939-41. <https://doi.org/10.3201/eid1312.070909> PMID: 18258053
54. Rockett RJ, Bialasiewicz S, Mhango L, Gaydon J, Holding R, Whitley DM, et al. Acquisition of human polyomaviruses in the first 18 months of life. *Emerg Infect Dis.* 2015;21(2):365-7. <https://doi.org/10.3201/eid2102.141429> PMID: 25626138
55. Borges JT, Nakada LYK, Maniero MG, Guimarães JR. SARS-CoV-2: a systematic review of indoor air sampling for virus detection. *Environ Sci Pollut Res Int.* 2021;28(30):40460-73. <https://doi.org/10.1007/s11356-021-13001-w> PMID: 33630259
56. Eren ZB, Vatansever C, Kabadayı B, Haykar B, Kuloğlu ZE, Ay S, et al. Surveillance of respiratory viruses by aerosol screening in indoor air as an early warning system for epidemics. *Environ Microbiol Rep.* 2024;16(4):e13303. <https://doi.org/10.1111/1758-2229.13303> PMID: 38982659
57. Hernaez B, Muñoz-Gómez A, Sanchiz A, Orviz E, Valls-Carbo A, Sagastagoitia I, et al. Monitoring monkeypox virus in saliva and air samples in Spain: a cross-sectional study. *Lancet Microbe.* 2023;4(1):e21-8. [https://doi.org/10.1016/S2666-5247\(22\)00291-9](https://doi.org/10.1016/S2666-5247(22)00291-9) PMID: 36436538
58. Geenen C, Traets S, Gorissen S, Happaerts M, Beuselinck K, Laenen L, et al. Interpretation of indoor air surveillance for respiratory infections: a prospective longitudinal observational study in a childcare setting. *EBioMedicine.* 2025;112:105512. <https://doi.org/10.1016/j.ebiom.2024.105512> PMID: 39884186
59. Kwok KTT, de Rooij MMT, Messink AB, Wouters IM, Smit LAM, Cotten M, et al. Establishing farm dust as a useful viral metagenomic surveillance matrix. *Sci Rep.* 2022;12(1):16308. <https://doi.org/10.1038/s41598-022-20701-x> PMID: 36175536
60. Justen L, Grimm S, Esvelt K, Bradshaw W. Indoor Air Sampling for Detection of Viral Nucleic Acids. *SSRN Electron J.* 2024 [Accessed 18 Jul 2024]. <https://doi.org/10.2139/ssrn.4823882>
61. Prussin AJ 2nd, Torres PJ, Shimashita J, Head SR, Bibby KJ, Kelley ST, et al. Seasonal dynamics of DNA and RNA viral bioaerosol communities in a daycare center. *Microbiome.* 2019;7(1):53. <https://doi.org/10.1186/s40168-019-0672-z> PMID: 30935423

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2025.

License, supplementary material and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.