

Direct genome sequencing of respiratory viruses from low viral load clinical specimens using the target capture sequencing technology

Nobuhiro Takemae,¹ Yumani Kuba,¹ Kunihiro Oba,² Tsutomu Kageyama¹

AUTHOR AFFILIATIONS See affiliation list on p. 12.

ABSTRACT The use of metagenomic next-generation sequencing technology to obtain complete viral genome sequences directly from clinical samples with low viral load remains challenging—especially in the case of respiratory viruses—due to the low copy number of viral versus host genomes. To overcome this limitation, target capture sequencing for the enrichment of specific genomes has been developed and applied for direct genome sequencing of viruses. However, as the efficiency of enrichment varies depending on the probes, the type of clinical sample, etc., validation is essential before target capture sequencing can be applied to clinical diagnostics. In this study, we evaluated the utility of target capture sequencing with a comprehensive viral probe panel for clinical respiratory specimens collected from patients diagnosed with SARS-CoV-2 or influenza type A. We focused on clinical specimens containing low copy numbers of viral genomes. Target capture sequencing yielded approximately 180- and 2,000-fold higher read counts of SARS-CoV-2 and influenza A virus, respectively, than metagenomic sequencing when the RNA extracted from specimens contained 59.3 copies/ μ L of SARS-CoV-2 or 625.1 copies/ μ L of influenza A virus. In addition, the target capture sequencing identified sequence reads in all SARS-CoV-2- or influenza type A-positive specimens with <26 RNA copies/ μ L, some of which also yielded $>70\%$ of the full-length genomes of SARS-CoV-2 or influenza A virus. Furthermore, the target capture sequencing using comprehensive probes identified co-infections with viruses other than SARS-CoV-2, suggesting that this approach will not only detect a wide range of viruses but also contribute to epidemiological studies.

IMPORTANCE Target capture sequencing has been developed and applied for direct genome sequencing of viruses in clinical specimens to overcome the low detection sensitivity of metagenomic next-generation sequencing. In this study, we evaluated the utility of target capture sequencing with a comprehensive viral probe panel for clinical respiratory specimens collected from patients diagnosed with SARS-CoV-2 or influenza type A, focusing on clinical specimens containing low copy numbers of viral genomes. Our results showed that the target capture sequencing yielded dramatically higher read counts than metagenomic sequencing for both viruses. Furthermore, the target capture sequencing using comprehensive probes identified co-infections with other viruses, suggesting that this approach will not only detect a wide range of viruses but also contribute to epidemiological studies.

KEYWORDS target capture sequencing, next-generation sequencing, direct genome sequencing, respiratory virus, influenza A virus, SARS-CoV-2

Next-generation sequencing (NGS) techniques have proven to be an indispensable tool for monitoring and controlling emerging pathogens, as demonstrated by

Editor Benjamin M. Liu, Children's National Hospital, George Washington University, Washington, DC, USA

Address correspondence to Tsutomu Kageyama, tkage@niid.go.jp.

Nobuhiro Takemae and Yumani Kuba contributed equally to this article. Author order was determined on the basis of seniority.

The authors declare no conflict of interest.

See the funding table on p. 12.

Received 18 April 2024

Accepted 19 August 2024

Published 14 October 2024

Copyright © 2024 Takemae et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

the SARS-CoV-2 pandemic (1). In addition, high-throughput and parallel nucleotide sequence analysis by NGS has enabled metagenomic sequencing, which could provide comprehensive genome sequences present in specimens (2). However, metagenomic sequencing has a lower detection sensitivity for viral genomes than traditional diagnostic methods such as polymerase chain reaction (PCR) because the copy number of viral genomes is much lower than that of host and bacterial genomes in clinical specimens (3). Therefore, obtaining full viral genome sequences directly from clinical specimens with low viral load is still challenging, especially for respiratory viruses, and full viral genome sequencing is often performed after virus isolation.

To overcome the weaknesses of NGS, an enrichment of certain genomes using amplicon-based sequencing (or the tiling amplicon method) or target capture sequencing (or hybridization capture sequencing) has been developed and applied for direct genome sequencing of viruses in clinical or environmental samples (3–5). In the amplicon-based sequencing, single or multiple regions of the target genomes can be amplified using from one set up to many sets of specific primers, providing a highly sensitive and economical method (4). This approach has therefore been used to obtain whole-genome sequences of SARS-CoV-2 from clinical specimens worldwide (6–8). However, because whole viral genome sequences are no longer available if unexpected mutations occur in the primer regions, the primers must be updated frequently, especially for RNA viruses with very fast evolutionary rates.

Target capture sequencing, as the name suggests, is a method for enriching target genomes in the NGS library using complementary probes to increase the sensitivity of NGS analysis (3–5). Typically, multiple 80- or 120-mer biotinylated DNA or RNA probes are used and are designed to overlap so that the entire target genome is covered. The designed probes hybridize to specific target genomes, which is expected to reduce the sequence reads derived from non-target genomes such as bacterial and host-derived genomes, allowing highly sensitive detection of the target genomes. Another advantage of target capture sequencing is its high tolerance for mismatches between the probe and the target nucleotide region (9). This advantage has been useful in the search for unidentified pathogens and/or RNA viruses that frequently mutate, such as SARS-CoV-2 and influenza A viruses. To date, multiple studies have applied target capture sequencing for the genome analysis of SARS-CoV-2 (9–16). More recently, a group attempted to sequence the monkey pox virus genome using target capture sequencing (17).

Target capture sequencing also enables the detection of multiple pathogens by simultaneously incorporating a large number of probes derived from various pathogen genomes into a single assay. Target capture sequencing for viral pathogens was first applied to specific viral genome analyses (18–20), and then commercial multipathogen DNA or RNA probe panels, or proprietary oligonucleotide panels customizable for target pathogens, were introduced and evaluated using clinical and/or nonclinical specimens (5, 9, 11, 12, 14, 21–25). For example, ViroCap, a custom probe panel designed from genome sequences of 34 families of vertebrate DNA or RNA viruses, dramatically increased the number of sequence reads derived from the viruses, as well as their breadth of coverage and average coverage, and detected a total of 32 viruses from the clinical specimens of 22 patients (23). Thus, target capture sequencing has been proven useful for both research and clinical diagnostics. However, as the efficiency of enrichment varies depending on the number of probes and targets included in the probe panel and the type of clinical specimen, as well as the library preparation method, validation is essential before application to clinical diagnostics. Moreover, it is critical to determine the limitations or detection sensitivity of each probe panel in clinical specimens with low viral load.

In this study, we first investigate the enrichment efficiency of target capture sequencing compared to metagenomic sequencing using clinical respiratory specimens collected from patients diagnosed with SARS-CoV-2 or type A influenza with sufficient copy numbers of viral genomes. We then evaluate the utility of target capture sequencing in clinical respiratory specimens, with a focus on clinical specimens

containing low copy numbers of viral genomes. The probe panel for target capture sequencing was selected from the commercially available panels as the panel containing the maximum number of constituent viral pathogen sequences in order to prepare for future pandemics related to novel viruses. We therefore chose the Twist Comprehensive Viral Research Panel (Twist Bioscience, San Francisco, CA) as a representative comprehensive panel containing reference sequences for a total of 15,488 different strains of 3,153 viruses, including zoonotic and human epizootic pathogens (<https://www.twistbioscience.com/products/ngs/fixe-d-panels/comprehensive-viral-research-panel?tab=overview>). Our results will provide important insights into the clinical diagnostic applications of target capture sequencing for various viral pathogens.

RESULTS

Comparison of metagenomic and target capture sequencing

To evaluate the enrichment efficiency of target capture sequencing, we first selected two clinical specimens with cycle threshold (Ct) values below 30 for either SARS-CoV-2 or influenza A virus by real-time PCR: the copy number of SARS-CoV-2 in the RNA solution extracted from CS2022-0121 was 59.3 copies/ μ L, and that of influenza A virus in the F16-31-UTM RNA solution was 625.1 copies/ μ L (Table 1).

In the CS2022-0121 library, metagenomic sequencing revealed 40 viral reads mapping to SARS-CoV-2 and only 53 viral reads even after the human rRNA removal treatment. Their consensus lengths were 5,024 and 5,160 nucleotide (nt) covering approximately 17% of the SARS-CoV-2 genome, and their average coverages were 0.2 and 0.3, respectively. On the other hand, target capture sequencing showed dramatic improvements in all the metrics. The numbers of reads mapped to SARS-CoV-2 in target capture sequencing without and with rRNA removal treatment was 7,331 and 2,101, respectively. These reads resulted in consensus sequences of 28,573 and 22,826 nt covering 95.6% and 76.4% of the SARS-CoV-2 genome, respectively. The enrichment efficiency of SARS-CoV-2 in the target capture sequencing assay without rRNA removal treatment against metagenomic sequencing was approximately 183.

In the F16-31-UTM library, the data obtained in each assay were analyzed in each of the eight segments of the influenza A virus genome (the PB2, PB1, PA, HA, NP, NA, M, and NS gene segments) (Table 1). Metagenomic sequencing without rRNA removal treatment and that with rRNA removal treatment revealed only 10 and six reads mapped to the influenza A virus genome, covering 7.8% and 4.9% of the total genomes, respectively. No reads derived from the NA gene were obtained in either assay, suggesting that subtype identification was not possible with metagenomic sequencing. The number of reads mapped to the influenza A virus by target capture sequencing without rRNA removal treatment and that with rRNA removal treatment was 19,459 and 12,642, respectively. These reads resulted in consensus sequences of 13,387 and 13,008 nt covering 98.2% and 95.4% of the influenza A virus genome, respectively. In both assays, the consensus lengths covering >94% were obtained for all gene segments, except for the PB2 gene in the target capture sequencing with rRNA removal treatment. The enrichment efficiency of influenza A virus in the target capture sequencing assay without rRNA removal treatment compared to metagenomic sequencing without rRNA removal treatment was approximately 1,950.

rRNA removal treatment had a more limited effect in the target capture sequencing compared to the metagenomic sequencing (Table 1). In metagenomic sequencing in the CS2022-0121 specimen, the percentage of reads derived from the human genome declined from 98.9% without rRNA removal treatment to 96.9% with rRNA removal treatment. A slight decline was observed in the F16-31-UTM specimen, from 98.5% to 98.1%. However, target capture sequencing using the same specimens did not yield a reduction in the percentage of human genomes; the percentages were approximately 95%–96% for CS2022-0121 and F16-31-UTM, respectively, irrespective of rRNA removal. Accordingly, the non-rRNA removal treatment assay was adopted for all subsequent target capture sequencing procedures. All sets of normalized coverages

TABLE 1 Comparison of metagenomic and target capture sequencing using specimens containing SARS-CoV-2 and influenza A virus per million reads^a

Specimen name	Cp value	Conc. cp/ μ L	Assay	Segment	Number of reads mapped to the human genome (%)		Mapped reads to the reference viral genome ^b (%)		Average coverage	Consensus length (nt)	Breadth of coverage (%)
CS2022-0121 (SARS-CoV-2)	28.85	59.33	Metagenomic	N.A. ^c	9,88,592	(98.86)	40	(0.00)	0.2	5,024	16.8
			Metagenomic with depletion of rRNA	N.A.	9,68,871	(96.89)	53	(0.01)	0.3	5,160	17.3
			Target capture	N.A.	9,53,497	(95.35)	7,331	(0.73)	35.0	28,573	95.6
			Target capture with depletion of rRNA	N.A.	9,59,712	(95.97)	2,101	(0.21)	9.8	22,826	76.4
F16-31-UTM (A(H1N1)pdm09)	28.54	625.14	Metagenomics	PB2			2	(0.00)	0.1	197	8.4
				PB1			0	(0.00)	0.0	0	0.0
				PA			3	(0.00)	0.2	366	16.4
				HA			2	(0.00)	0.2	179	10.1
				NP			2	(0.00)	0.2	256	16.4
				NA			0	(0.00)	0.0	0	0.0
				MP			1	(0.00)	0.1	69	6.7
				NS			0	(0.00)	0.0	0	0.0
				Total	9,85,241	(98.52)	10	(0.00)	0.1	1,067	7.8
				Metagenomics with depletion of rRNA	PB2			0	(0.00)	0.0	0
			PB1				2	(0.00)	0.1	123	5.3
			PA				2	(0.00)	0.1	245	11.0
			HA				0	(0.00)	0.0	0	0.0
			NP				2	(0.00)	0.2	300	19.2
			NA				0	(0.00)	0.0	0	0.0
			MP				0	(0.00)	0.0	0	0.0
			NS				0	(0.00)	0.0	0	0.0
			Total		9,80,868	(98.09)	6	(0.00)	0.1	668	4.9
			Target capture		PB2			4,944	(0.49)	306.1	2,323
				PB1			2,851	(0.29)	177.8	2,276	97.22
PA				3,510	(0.35)	228.1	2,185	97.85			
HA				2,546	(0.25)	208.0	1,758	98.93			
NP				2,332	(0.23)	216.6	1,553	99.23			
NA				1,547	(0.15)	154.3	1,444	99.04			
MP				1,312	(0.13)	186.4	972	94.64			
NS				417	(0.04)	66.5	876	98.43			
Total	9,56,973	(95.70)		19,459	(1.95)	207.3	13,387	98.2			
Target capture with depletion of rRNA	PB2				2,166	(0.22)	134.7	2,029	86.67		
	PB1			3,105	(0.31)	203.5	2,162	96.82			
	PA			283	(0.03)	44.0	866	97.30			
	HA			1,212	(0.12)	120.8	1,434	98.35			
	NP			831	(0.08)	118.3	974	94.84			
	NA			1,379	(0.14)	113.1	1,753	98.65			
	MP			2,939	(0.29)	184.0	2,261	96.58			
	NS			727	(0.07)	66.8	1,529	97.70			
	Total	9,63,240	(96.32)	12,642	(1.26)	135.2	13,008	95.4			

^aAll data were standardized and calculated per million reads to compare differences between assays.

^b hCoV-19_Wuhan_WIV04_2019 (EPI_ISL_402124) and A/California/04/2009 (H1N1) (EPI_ISL_376192) were used as the reference sequences of CS2022-0121 and F16-31-UTM, respectively.

^cN.A., Not applicable.

across SARS-CoV-2 (Fig. S1) or influenza A virus (Fig. S2) in CS2022-0121 and F16-31-UTM, respectively, downsampled to 1,000,000, are shown.

Availability of target capture sequencing in clinical specimens with low viral loads

To investigate the availability of target capture sequencing in clinical specimens with low viral loads, we selected clinical specimens with Ct values > 30 against either SARS-CoV-2 or influenza A viruses. The SARS-CoV-2-positive specimens (CS2022-0099, -0110, -0108, -0083, -0090, and -0057) contained 0.52–9.79 RNA copies/μL of the SARS-CoV-2 genome (Table 2), and the influenza A virus (A(H1N1)pdm09 or H3N2 subtypes)-positive specimens (F16-51-UTM, F16-36-UTM, F16-60-UTM, F15-10-UTM, F16-62-UTM, and F14-66-UTM) contained 2.06–21.14 RNA copies/μL of the influenza A virus genome (Table 3).

For SARS-CoV-2-positive specimens (Fig. 1A; Table 3), target capture sequencing identified viral reads covering approximately 80% of the SARS-CoV-2 genome with an average coverage of >12 in CS2022-0099 and CS2022-0110, which contained >6 RNA copies/μL. In contrast, in the case of the specimens containing <5 RNA copies/μL, the viral reads covered 0.7% to 15.7% of the SARS-CoV-2 genome with 6–14 viral reads derived from SARS-CoV-2, with the exception of CS2022-0083, which yielded 1,826 viral reads.

For the influenza A virus-positive specimens (Fig. 1B; Table 3), i.e., F16-51-UTM, F16-36-UTM, and F15-10-UTM, which contained >7 RNA copies of the influenza A virus, viral reads covering >74% of the influenza A virus genome were obtained with an average coverage >18. Except for the NS gene in F16-36-UTM, >67% breadth of coverage was obtained in all segments in these specimens, suggesting efficient genomic analysis by targeted capture sequencing. Surprisingly, in the specimens containing less than ~5 RNA copies/μL, i.e., F16-60-UTM, F16-62-UTM, and F14-66-UTM, consensus sequences covering about 26%–30% of the influenza A virus genome were obtained. The breadth of coverage obtained in the HA and NA genes of these specimens exceeded 10%, which was sufficient for identification of their viral subtypes.

From these results, approximately at least six RNA copies/μL were likely needed to obtain more than 70% breadth of coverage by target capture sequencing under our experimental settings. There was a sharp decrease in the coverage obtained below this threshold concentration. In fact, all specimens containing less than six RNA copies/μL covered less than 31% of both virus genomes.

Detection of simultaneous infections with multiple viral pathogens

Viral pathogen genomes other than SARS-CoV-2 were detected in two of the SARS-CoV-2-positive specimens, CS2022-0083 and CS2022-0108. No viral pathogen genomes other than influenza A virus were detected in influenza A virus-positive specimens. In CS2022-0083, 1,553 of the obtained reads were mapped to circovirus-like genome DCCV-4 (accession number [NC_030470.1](#)) with 53.2% breadth of coverage and an average coverage of 63.4 (Table 4). Unfortunately, for the circovirus-like genome DCCV-4, only two genome sequences obtained directly from environmental samples from a freshwater lake in China have been published, so it was not possible to estimate whether the viruses with similar genomes affected respiratory symptoms. In the CS2022-0108 specimen, which was positive for adenovirus (AdV), human metapneumovirus (MNV), parainfluenza virus 3 (PIV3), and human rhinovirus (RV)/enterovirus (EV) in addition to SARS-CoV-2 by diagnosis with the BioFire FilmArray Respiratory Panel 2.1 (Table 5), 4,354 of the obtained reads were mapped to the reference MN173594.1, which was the enterovirus D68 strain USA/2018/CA-RGDS-1056 polyprotein gene with 93.2% breadth of coverage and an average coverage of 88.5 (Table 4). BLAST analysis revealed that a consensus sequence of 7,335 nt had the highest identity (99.4%) with JH-EV-50/2022 (accession number [OP572066.1](#)) isolated in the USA in 2022. Although it was below the threshold for taxonomy analysis, 30 reads mapped to PIV3 (accession number [NC_038270.1](#)) and 28 reads mapped to ADV (accession number [NC_001405.1](#)) were found. However, no reads mapped to MNV, so the copy number of MNV may have been below the detection limit of the target capture sequencing used in this study.

DISCUSSION

Here, we have demonstrated the effectiveness of target capture sequencing directly from respiratory clinical specimens with low viral load. The capture panel, the Twist Comprehensive Viral Research Panel, was selected for its ability to capture a very wide range of viral pathogens; to our knowledge, this is the first evaluation of this comprehensive panel using human respiratory clinical specimens containing small amounts of SARS-CoV-2 and influenza A virus genomes, although the panel has been used on human cerebrospinal fluid (26), saliva, blood, and feces from wild bats (27) and mosquito samples (28). Target capture sequencing with this panel yielded approximately 180- and 2,000-fold higher read counts of SARS-CoV-2 or influenza A virus, respectively, than metagenomic sequencing when RNA extracts from specimens containing 59.3 or 625.1 RNA copies/ μL of SARS-CoV-2 or influenza A virus were used, respectively. Despite differences in library preparation conditions, sequencers, and capture panels, previous studies also found that the target capture sequencing had high sensitivity for many pathogens in various clinical specimens, including animal specimens (3, 4, 9, 11–13, 16, 22, 23, 29, 30), reinforcing the results of our present study.

The detection sensitivity of the target capture sequencing may not differ substantially from that of RT-PCR methods, although a variety of experimental conditions and positivity thresholds, such as NGS metrics (breadth of coverage, average coverage, etc.), have been adopted in the target capture sequencing studies (15, 16, 29). In fact, all specimens positive for SARS-CoV-2 or influenza type A yielded the corresponding virus-derived reads in our target capture sequencing, although SARS-CoV-2-positive specimens with Ct values > 35 were not available in this study. Nagy-Szakal *et al.* reported that the positive and negative percentage concordances between the RT-PCR method and target capture sequencing using nasopharyngeal swab specimens were 96.7% and 100%, respectively (15). On the other hand, the enrichment efficiency was clearly higher for influenza A virus-positive specimens than for SARS-CoV-2-positive specimens. The Twist Comprehensive Viral Research Panel contains probes derived from 8,050 strains of influenza A viruses and only one strain of SARS-CoV-2 (strain name unavailable). Therefore, there may be a greater number of probes hybridizing to the influenza A virus genome than to the SARS-CoV-2 genome. Unfortunately, the detailed design of the panel, which could be crucial for enrichment (11, 21, 31), was not available, and thus it was not clear how the panel affected our NGS results.

There was no clear correlation between the RNA copy number of the viral genome and the number of corresponding reads obtained from the clinical specimens with low viral load (less than ~ 5 RNA copies/ μL). Likewise, it has been reported that the clinical specimens with Ct values > 30 showed a variety of enrichment efficiencies after target capture sequencing, even among the specimens with the same Ct values (11, 12, 15). On the other hand, in NGS libraries prepared from the dilution of cultured viruses rather than clinical specimens (13), there seems to be a clearer correlation between the number of viral reads and genome copies. Thus, it may be difficult to obtain a stable number of reads from NGS libraries prepared from clinical specimens with low viral load, as

TABLE 2 Overview of the genome reads per million reads obtained from clinical samples showing high Ct values against SARS-CoV-2 (> 30) using target capture sequencing

Specimen name	Ct values	Conc. cp/ μL	Mapped reads to the reference viral genome ^a	Average coverage	Consensus length (nt)	Breadth of coverage (%)
CS2022-0099	31.66	9.79	2,562 (0.26)	12.49	23,904	80.0
CS2022-0110	32.14	6.88	7,845 (0.78)	26	23,980	80.2
CS2022-0108	33.95	0.52	7 (0.00)	0.02	395	1.3
CS2022-0083	34.52	1.55	1,826 (0.18)	8.23	4,704	15.7
CS2022-0090	34.59	0.86	6 (0.00)	0.02	198	0.7
CS2022-0057	34.65	0.52	14 (0.00)	0.04	305	1.0

^ahCoV-19_Wuhan_WIV04_2019 (EPI_ISL_402124) was used as a reference sequence.

TABLE 3 Overview of the genome reads per million reads obtained from clinical specimens showing high Ct values against influenza A virus (> 30) using target capture sequencing

Specimen name	Ct values	Conc. cp/ μ L	Segment	Mapped reads to the reference viral genome ^a (%)	Average coverage	Consensus length (nt)	Breadth of coverage (%)
F16-51-UTM (A(H1N1)pdm09)	33.82	13.23	PB2	372 (0.04)	23.3	2,120	90.6
			PB1	222 (0.02)	13.9	1,743	74.5
			PA	489 (0.05)	32.1	2,107	94.4
			HA	152 (0.02)	12.4	1,427	80.3
			NP	267 (0.03)	24.4	1,474	94.2
			NA ^c	100 (0.01)	10.3	1,330	91.2
			MP	74 (0.01)	10.4	691	67.3
			NS	35 (0.00)	5.8	800	89.9
		Total	1,711 (0.17)	18.3	11,692	85.8	
F16-36-UTM (A(H1N1)pdm09)	33.31	21.14	PB2	766 (0.08)	45.9	2,017	86.2
			PB1	434 (0.04)	26.8	1,680	71.8
			PA	422 (0.04)	28.1	1,680	75.2
			HA	273 (0.03)	22.4	1,303	73.3
			NP	318 (0.03)	29.6	1,113	71.1
			NA	408 (0.04)	38.7	1,043	71.5
			MP	394 (0.04)	53.3	824	80.2
			NS	72 (0.01)	11.1	431	48.4
		Total	3087 (0.31)	32.3	10,091	74.0	
F16-60-UTM (A(H1N1)pdm09)	35.7	4.45	PB2	46 (0.00)	2.9	712	30.4
			PB1	48 (0.00)	3.1	798	34.1
			PA	42 (0.00)	2.8	724	32.4
			HA	26 (0.00)	2.1	601	33.8
			NP	24 (0.00)	2.3	414	26.5
			NA	6 (0.00)	0.6	316	21.7
			MP	4 (0.00)	0.3	70	6.8
			NS	0 (0.00)	0.0	0	0.0
		Total	196 (0.02)	2.1	3,635	26.7	
F15-10-UTM (H3N2)	35.23	7.11	PB2	546 (0.05)	34.9	2,182	93.2
			PB1	334 (0.03)	20.8	1,917	81.9
			PA	780 (0.08)	49.8	2,140	95.8
			HA	156 (0.02)	13.0	1,531	86.9
			NP	290 (0.03)	27.3	1,289	82.3
			NA	358 (0.04)	34.1	1,026	69.9
			MP	132 (0.01)	22.3	813	91.3
			NS	319 (0.03)	46.3	711	69.2
		Total	2915 (0.29)	31.2	11,609	85.2	
F16-62-UTM (H3N2)	35.47	2.06	PB2	214 (0.02)	12.9	1,219	52.1
			PB1	16 (0.00)	1.0	313	13.4
			PA	128 (0.01)	8.2	408	18.3
			HA	24 (0.00)	2.0	180	10.2
			NP	156 (0.02)	14.2	1,017	64.9
			NA	88 (0.01)	8.9	690	47.0
			MP	113 (0.01)	12.5	377	36.7
			NS	0 (0.00)	0.0	0	0.0
		Total	739 (0.07)	7.5	4,204	30.8	
F14-66-UTM (H3N2) ^b	37.01	4.45	PB2	8188 (0.82)	504.5	1255	53.6
			PB1	1867 (0.19)	104.3	570	24.3
			PA	4275 (0.43)	277.3	916	41.0
			HA	486 (0.05)	41.5	181	10.3
			NP	380 (0.04)	34.6	285	18.2

(Continued on next page)

TABLE 3 Overview of the genome reads per million reads obtained from clinical specimens showing high Ct values against influenza A virus (> 30) using target capture sequencing (*Continued*)

Specimen name	Ct values	Conc. cp/ μ L	Segment	Mapped reads to the reference viral genome ^a (%)	Average coverage	Consensus length (nt)	Breadth of coverage (%)
			NA	1569 (0.16)	137.7	279	19.0
			MP	929 (0.09)	121.8	307	29.9
			NS	38 (0.00)	5.6	132	14.8
			Total	17732 (1.77)	183.8	3,925	28.8

^aA/California/04/2009 (H1N1) (EPI_ISL_37619) or A/Nagasaki/14N024/2015(H3N2) (EPI_ISL_176857) was used as a reference sequence.

^bNo downsampling was performed in this specimen because the total number of reads was 406,012, which was less than 1 million reads.

^cNA, Not applicable.

the methods of specimen collection and the proportion of host- or bacteria-derived genomes are quite different in each specimen.

As shown in our study and in previous studies (16, 23, 29), the target capture sequencing using probes for multiple pathogens is also very useful for the identification of co-infections. For example, Kim *et al.* successfully demonstrated that 8% of cases out of 92 SARS-CoV-2-positive nasopharyngeal swabs showed co-infection with rhinovirus (6%) or influenza virus (2%) (16). On the other hand, in our study, co-infection of EV-D68 was detected in the CS2022-0108 specimen. EV-D68 is an emerging viral pathogen first identified in 1962 in children hospitalized due to respiratory disease (32). Since 2005, several countries have reported an increase in the number of patients with respiratory diseases caused by EV-D68 (33). In addition, a rapid increase in EV-D68 infections was reported from eight European countries in 2021 (34). In Japan, EV-D68 outbreaks have been reported several times (35–37). Interestingly, EV-D68 infections were reported in 25 (14%) of 197 specimens positive for HRV/EV by BIOFIRE Respiratory 2.1 collected from 1 September to 13 October 2022 at a hospital in Tokyo, Japan (different from the hospital in this study) (IASR Vol. 43 pp. 290–291: 2022, Dec. <https://www.niid.go.jp/niid/ja/diseases/a/ev-d68/2335-idsc/iasr-news/11650-514p01.html>) (in Japanese). These hospitals are located close to each other in Tokyo, and the specimens were collected during the same season, suggesting that an epidemic of EV-D68 occurred in Tokyo along with the SARS-CoV-2 epidemic. These results suggest that target capture sequencing with comprehensive probe panels could not only detect a wide range of causative viruses but also contribute to epidemiological studies.

Greater probe diversity allows the detection of many targeted genomes in target capture sequencing, but there may be a trade-off in terms of the increased number of off-target reads (11). Target capture sequencing dramatically increased the number of target viral reads, but human genome reads still accounted for >95% of the total reads in this study. In addition, unfortunately, human rRNA removal treatment did not

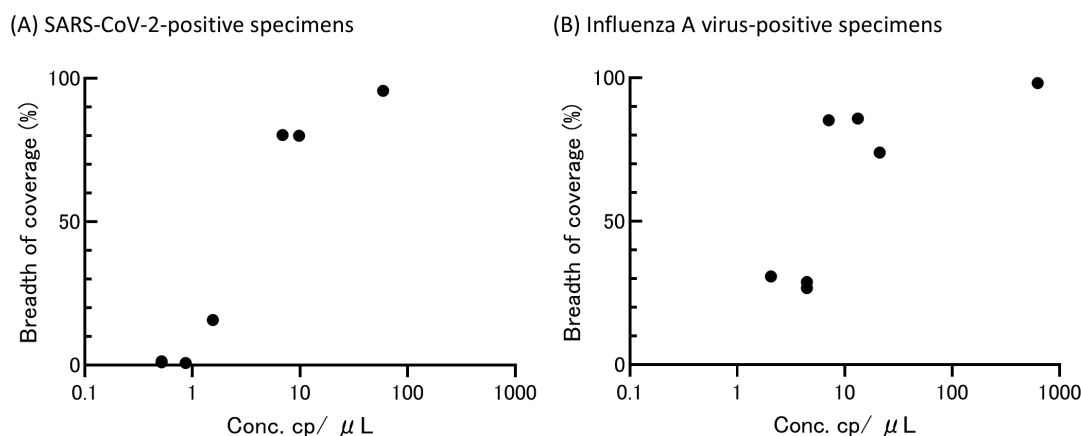
**FIG 1** Correlation between RNA copy number per microliter and breadth of coverage (%) obtained by using target capture sequencing in RNA extracts from clinical specimens containing SARS-CoV-2 (A) and influenza A virus genomes (B), respectively.

TABLE 4 Detection of simultaneous viral infections by the Find Best References tool with the Clustered Reference Viral DataBase^a

Specimen name	Input reads	Number of reads mapped	Breadth of coverage (%)	Average coverage	Accession no. of best match reference	Reference length (nt)	Taxonomy (definition)
CS2022-0083	19,05,260	1,553	53.2	63.4	NC_030470.1	2,985	Circovirus-like genome DCCV-4, complete genome
CS2022-0108 ^b	20,16,746	4,354	93.2	88.5	MN173594.1	7,272	Enterovirus D68 strain USA/2018/CA-RGDS-1056 polyprotein gene, complete cds.

^aThe Clustered Reference Viral DataBase (RVDB, v21.0 June 2021) was used as the reference in the Genomics Workbench software. The thresholds of the minimum number of reads on a reference, minimum coverage, and minimum fraction of reference covered were set to be 100, 10, and 0.5, respectively.

^bThe BioFire FilmArray Respiratory Panel 2.1 detected adenovirus, human metapneumovirus, parainfluenza virus 3, and human rhinovirus/enterovirus in addition to SARS-CoV-2 in this specimen.

significantly reduce the percentage of human genome reads, but instead reduced the number of target viral reads in target captured libraries. Other pretreatments, such as filtering of original specimens and post-extraction DNase treatment, also showed little effect on target capture efficiency (29). Therefore, in our study, especially for specimens with low viral load, we decided to use only high-speed centrifugation prior to RNA extraction. However, the large number of host genomes included in the library after hybridization suggests the possibility of further improvements to increase the enrichment efficiency, even when using a comprehensive probe panel.

Target capture sequencing with a comprehensive panel could be a useful tool to simultaneously identify a variety of viruses in one assay. Although, recently, the genetic detection of pathogens has often been done by PCR, the number of assays must be increased according to the increase in the number of targets. This increases the burden of the assay, number of processes, management of reagents, etc. In addition, PCR methods may miss concurrent infections by other pathogens, as no further diagnosis is done if positivity for a particular pathogen is detected. However, it should be noted that target capture methods are more labor- and cost-intensive when compared to conventional PCR and metagenomic analysis. Target capture sequencing, especially with large probe panels, requires tighter control of the library preparation to decrease the risk of cross-contamination between libraries, contamination from reagents known as “kitome,” airborne contaminants, contamination due to index switching, etc. Target capture sequencing is less likely to miss identifying viruses with genetic diversity, especially respiratory viruses, than traditional Sanger sequencing using universal primers because the Sanger sequencer can only analyze genomes that have been amplified with

TABLE 5 List of clinical specimens collected from patients with respiratory diseases, which are used in this study

Specimen name	Age	Clinical sample	Collection date	Pathogens diagnosed ^a
CS2022-0121	55 years	Nasopharyngeal swab	08-02-2022	SARS-CoV-2
CS2022-0090	10 years and 11 months	Nasopharyngeal swab	10-09-2022	SARS-CoV-2
CS2022-0057	4 years and 3 months	Nasopharyngeal swab	17-06-2022	SARS-CoV-2
CS2022-0083	7 years	Nasopharyngeal swab	26-08-2022	SARS-CoV-2
CS2022-0099	8 years and 1 month	Nasopharyngeal swab	01-10-2022	SARS-CoV-2
CS2022-0110	6 years and 3 months	Nasopharyngeal swab	17-11-2022	SARS-CoV-2
CS2022-0108	3 years and 5 months	Nasopharyngeal swab	15-11-2022	SARS-CoV-2, adenovirus, human metapneumovirus, parainfluenza virus 3, and human rhinovirus/enterovirus
F16-31-UTM	1 year	Nasopharyngeal swab	08-02-2016	Influenza A virus (pdmH1N1)
F16-51-UTM	44 years	Nasopharyngeal swab	07-03-2016	Influenza A virus (pdmH1N1)
F15-10-UTM	7 years	Nasal discharge	26-01-2015	Influenza A virus (H3N2)
F16-36-UTM	9 years	Nasal discharge	01-02-2016	Influenza A virus (pdmH1N1)
F16-60-UTM	10 years	Nasal discharge	23-03-2016	Influenza A virus (pdmH1N1)
F16-62-UTM	16 years	Nasal discharge	05-04-2016	Influenza A virus (H3N2)
F14-66-UTM	13 years	Nasal discharge	28-12-2014	Influenza A virus (H3N2)

^aThe specimens prefixed “CS2022-” were diagnosed using the BioFire FilmArray Respiratory Panel 2.1.

properly designed universal primers, as described in a previous study (38). Unlike the Sanger sequencer, the universal primers specific to each viral family or genus are not required in the target capture sequencing, where many probes tailored to a diversity of viral genomes can be incorporated as a comprehensive probe panel. This may also lead to the discovery of novel genotypes or emerging and re-emerging viruses with potential to cause a pandemic. It is anticipated that this new technology will be further developed and applied to pathogen diagnosis and rapid response to emerging infectious disease threats.

MATERIALS AND METHODS

Sample preparation and RNA extraction

We used 14 clinical specimens (nasopharyngeal swab or nasal discharge) collected from patients with respiratory disease symptoms at Showa General Hospital, Tokyo, to evaluate the target capture sequencing (Table 5). Among the 14 specimens, seven specimens were designated CS2022- and found to be positive for SARS-CoV-2 using a BIOFIRE Respiratory 2.1 panel (BioFire Diagnostics, Salt Lake City, UT). The remaining seven specimens tested positive for type A influenza by real-time RT-PCR at the National Institute of Infectious Diseases, Japan, as described below. All specimens were placed in sterile tubes containing the viral transport medium and stored at -80°C until use. SARS-CoV-2-positive specimens were freeze-thawed one or two times, while influenza A virus-positive specimens were freeze-thawed no more than three times prior to RNA extraction to avoid negative effects on the detection limit of the viral RNA copy number in clinical specimens. In addition, all clinical specimens were centrifuged at 16,000 g for 2 minutes to reduce the risk of contamination from host and bacterial-derived materials. Sixty microliters of RNA was extracted from 140 μL of each clinical specimen by using a Viral RNA Mini kit (Qiagen, Hilden, Germany) with an automated extraction platform QIAcube (Qiagen).

Real-time RT-PCR and digital PCR for identification and quantification of SARS-CoV-2 and influenza A virus

Identification of SARS-CoV-2 or influenza A virus was performed by a one-step real-time RT-PCR, as previously described (39, 40). The absolute number of SARS-CoV-2 RNA copies present in each extracted RNA was determined by a digital PCR using Absolute Q 1-step RT-dPCR Master Mix (Thermo Fisher Scientific, Waltham, MA) in a QuantStudio Absolute Q digital PCR system (Thermo Fisher Scientific). Nine microliters of a reaction mixture containing 2.97 μL of each extracted RNA was loaded. Thermal cycling was performed as follows: reverse transcription at 55°C for 10 minutes, preheating at 96°C for 10 minutes, and 40 cycles of denaturation at 96°C for 5 seconds and annealing/extension at 60°C for 30 seconds. The absolute number of influenza A virus RNA copies present in each extracted RNA was determined by a one-step real-time PCR based on the number of the Twist Synthetic Influenza H3N2 RNA control (Twist Bioscience). Primers and probes targeting N genes of SARS-CoV-2 (N2 assay) (39) or M genes of influenza A virus (40) were used in each PCR assay.

Library preparation

Libraries for metagenomic sequencing of CS2022-0121 and F16-31-UTM were prepared using a NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB, Ipswich, MA) to elucidate the enrichment efficiency of the target capture sequencing. Briefly, 13 μL of each RNA was converted to single-stranded cDNA using random primers after heat fragmentation, and then double-stranded cDNA was synthesized. After end-repairing and dA-tailing reactions, the adapters diluted at a 1:10 ratio were ligated according to the NEB protocol. After size selection performed using AMPure XP beads (Beckman Coulter, Indianapolis, IN), the adapter-ligated DNA was amplified by 17 cycles of PCR.

Libraries for the target capture sequencing were prepared using a Twist Comprehensive Viral Research Panel (Twist Bioscience) as follows. Fifteen microliters of each RNA was converted to cDNA using Protoscript II First-Strand cDNA synthesis (NEB) and random primer 6 (NEB). The NEBNext Ultra II Non-Directional RNA Second-Strand Synthesis Module was subsequently used to convert single-stranded cDNA to double-stranded cDNA. The libraries were then generated using a Twist Library Preparation EF Kit 2.0 and Unique Dual Indices (UDI) (Twist Bioscience). Standard hybridization workflow target capture using the Twist Comprehensive Viral Research Panel was followed by a Twist Standard Target Enrichment workflow with slight modifications. Briefly, 1 μg of each dual-indexed library was dried using a vacuum concentrator with no heat. Hybridization capture was performed by adding 1 μg of the Twist Comprehensive Viral Research Panel to each library for 16 hours at 70°C. After the hybridization was complete, the hybridized libraries were collected using streptavidin beads. The streptavidin binding bead slurry was then amplified according to the following protocol: initialization at 40°C for 45 seconds, followed by 21 cycles of denaturation at 98°C for 15 seconds and annealing at 60°C for 30 seconds, with a final extension at 72°C for 30 seconds. In order to validate our workflow for target capture sequencing, we preliminarily prepared the NGS libraries of the synthetic influenza H3N2 RNA control (Twist Bioscience) with two different dilutions, resulting in 10 and 1,000 RNA copies/ μL , spiked into a background of human reference RNA (Agilent Technologies, Palo Alto, CA) according to the manufacturer's instructions. After 16 hours of hybridization with the Twist Comprehensive Viral Research Panel, we confirmed that sufficient amounts of sequence reads derived from influenza A virus were obtained in both samples (data not shown).

In the preparation of NGS libraries for the metagenomic and target capture sequencings of CS2022-0121 and F16-31-UTM, the effect of human rRNA removal was also assessed using a QIAseq FastSelect rRNA removal kit (Qiagen). The rRNA removal reaction was performed after RNA thermal denaturation at 95°C for 5 minutes according to the instructions.

Sequencing and data analysis

The concentrations of metagenomic or target-enriched libraries were measured using a Qubit 2.0 Fluorometer in combination with a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) and then analyzed on an Agilent 4150 TapeStation in combination with an Agilent D1000 ScreenTape System (Agilent Technologies). The library pool (up to eight samples), which was individually diluted to 1 nM, was sequenced with 150-bp paired-end reads on the Illumina MiSeq platform, using a MiSeq Reagent kit v2 (Illumina, San Diego, CA).

Generated sequence reads were imported into the Genomics Workbench software (version 21.0.4; Qiagen). The data analysis workflow was as follows. Briefly, the low-quality reads and reads with <30 bp were trimmed and downsampled to 1,000,000 reads for performing comparisons between assays or between specimens. Then, the downsampled reads were mapped to the SARS-CoV-2 (hCoV-19_Wuhan_WIV04_2019 (EPI_ISL_402124)) or influenza A viruses (A/California/04/2009 (H1N1) (EPI_ISL_376192) or A/Nagasaki/14N024/2015(H3N2) (EPI_ISL_176857)) using the default parameters in the Map Reads to Reference tool. Average coverages, number of reads, and breadth of coverage against each reference sequence were evaluated. The enrichment efficiency of target capture sequencing was calculated as the ratio of the number of reads mapped to the reference sequence by target capture sequencing without rRNA removal treatment to that by metagenomic sequencing without rRNA removal treatment.

We further performed taxonomic analysis to evaluate the target capture sequencing as a tool for detection of simultaneous infections with multiple viral pathogens. The analysis was carried out using the Find Best References using Read Mapping Tool with Clustered Reference Viral DataBase (RVDB, v21.0 June 2021) (41) as the reference in the Genomics Workbench software. The thresholds for the minimum number of reads on a reference, minimum coverage, and minimum fraction of reference covered (minimum

fraction of the reference sequence to be covered by read for a reference) were set to be 100, 10, and 0.5, respectively, and the other parameters were the default settings. Then, consensus sequences were obtained by using the Map Reads to Reference tool with default settings using appropriate references.

ACKNOWLEDGMENTS

This research was supported by grants from AMED (nos. JP22fk0108543 and JP24fk0108692) and a JSPS KAKENHI grant (no. JP22K10501).

AUTHOR AFFILIATIONS

¹Center for Emergency Preparedness and Response, National Institute of Infectious Diseases, Tokyo, Japan

²Department of Pediatrics, Showa General Hospital, Kodaira, Tokyo, Japan

AUTHOR ORCID*s*

Nobuhiro Takemae  <http://orcid.org/0000-0001-5152-8938>

Tsutomu Kageyama  <http://orcid.org/0009-0000-1829-5124>

FUNDING

Funder	Grant(s)	Author(s)
Japan Agency for Medical Research and Development (AMED)	JP22fk0108543	Nobuhiro Takemae Yumani Kuba Kunihiro Oba Tsutomu Kageyama
Japan Agency for Medical Research and Development (AMED)	JP24fk0108692	Nobuhiro Takemae Yumani Kuba Tsutomu Kageyama
MEXT Japan Society for the Promotion of Science (JSPS)	JP22K10501	Nobuhiro Takemae

ETHICS APPROVAL

This study was approved by the Research and Ethical Committee of the National Institute of Infectious Diseases (NIID), Japan (approval #1544).

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental material ([Spectrum00986-24-s0001.pdf](#)). Fig. S1 and S2.

REFERENCES

1. Quer J, Colomer-Castell S, Campos C, Andrés C, Piñana M, Cortese MF, González-Sánchez A, Garcia-Cehic D, Ibáñez M, Pumarola T, Rodríguez-Frías F, Antón A, Tabernero D. 2022. Next-generation sequencing for confronting virus pandemics. *Viruses* 14:600. <https://doi.org/10.3390/v14030600>
2. Bassi C, Guerriero P, Pierantoni M, Callegari E, Sabbioni S. 2022. Novel virus identification through metagenomics: a systematic review. *Life (Basel)* 12:2048. <https://doi.org/10.3390/life12122048>
3. Ceballos-Garzon A, Comtet-Marre S, Peyret P. 2024. Applying targeted gene hybridization capture to viruses with a focus to SARS-CoV-2. *Virus Res* 340:199293. <https://doi.org/10.1016/j.virusres.2023.199293>
4. Gaudin M, Desnues C. 2018. Hybrid capture-based next generation sequencing and its application to human infectious diseases. *Front Microbiol* 9:2924. <https://doi.org/10.3389/fmicb.2018.02924>
5. Fitzpatrick AH, Rupnik A, O'Shea H, Crispie F, Keaveney S, Cotter P. 2021. High throughput sequencing for the detection and characterization of RNA viruses. *Front Microbiol* 12:621719. <https://doi.org/10.3389/fmicb.2021.621719>
6. Lu J, du Plessis L, Liu Z, Hill V, Kang M, Lin H, Sun J, François S, Kraemer MUG, Faria NR, et al. 2020. Genomic epidemiology of SARS-CoV-2 in Guangdong province, China. *Cell* 181:997–1003. <https://doi.org/10.1016/j.cell.2020.04.023>

7. Sekizuka T, Itokawa K, Hashino M, Kawano-Sugaya T, Tanaka R, Yatsu K, Ohnishi A, Goto K, Tsukagoshi H, Ehara H, Sadamasu K, Taira M, Shibata S, Nomoto R, Hiroi S, Toho M, Shimada T, Matsui T, Sunagawa T, Kamiya H, Yahata Y, Yamagishi T, Suzuki M, Wakita T, Kuroda M. 2020. A genome epidemiological study of SARS-CoV-2 introduction into Japan. *mSphere* 5:e00786-20. <https://doi.org/10.1128/mSphere.00786-20>
8. Ulhuq FR, Barge M, Falconer K, Wild J, Fernandes G, Gallagher A, McGinley S, Sugadol A, Tariq M, Maloney D, Kenicer J, Dewar R, Templeton K, McHugh MP. 2023. Analysis of the ARTIC V4 and V4.1 SARS-CoV-2 primers and their impact on the detection of Omicron BA.1 and BA.2 lineage-defining mutations. *Microb Genom* 9:mgen000991. <https://doi.org/10.1099/mgen.0.000991>
9. Wylezich C, Calvelage S, Schlottau K, Ziegler U, Pohlmann A, Höper D, Beer M. 2021. Next-generation diagnostics: virus capture facilitates a sensitive viral diagnosis for epizootic and zoonotic pathogens including SARS-CoV-2. *Microbiome* 9:51. <https://doi.org/10.1186/s40168-020-00973-z>
10. Chiara M, D'Erchia AM, Gissi C, Manzari C, Parisi A, Resta N, Zambelli F, Picardi E, Pavesi G, Horner DS, Pesole G. 2021. Next generation sequencing of SARS-CoV-2 genomes: challenges, applications and opportunities. *Brief Bioinform* 22:616–630. <https://doi.org/10.1093/bib/bbaa297>
11. Pipoli da Fonseca J, Kornobis E, Turc E, Enouf V, Lemée L, Cokelaer T, Monot M. 2022. Capturing SARS-CoV-2 from patient samples with low viral abundance: a comparative analysis. *Sci Rep* 12:19274. <https://doi.org/10.1038/s41598-022-23422-3>
12. Doddapaneni H, Cregeen SJ, Sucgang R, Meng Q, Qin X, Avadhanula V, Chao H, Menon V, Nicholson E, Henke D, Piedra FA, Rajan A, Momin Z, Kottapalli K, Hoffman KL, Sedlaczek FJ, Metcalf G, Piedra PA, Muzny DM, Petrosino JF, Gibbs RA. 2021. Oligonucleotide capture sequencing of the SARS-CoV-2 genome and subgenomic fragments from COVID-19 individuals. *PLoS One* 16:e0244468. <https://doi.org/10.1371/journal.pone.0244468>
13. Xiao M, Liu X, Ji J, Li M, Li J, Yang L, Sun W, Ren P, Yang G, Zhao J, et al. 2020. Multiple approaches for massively parallel sequencing of SARS-CoV-2 genomes directly from clinical samples. *Genome Med* 12:57. <https://doi.org/10.1186/s13073-020-00751-4>
14. Rehn A, Braun P, Knüpfer M, Wölfel R, Antwerpen MH, Walter MC. 2021. Catching SARS-CoV-2 by sequence hybridization: a comparative analysis. *mSystems* 6:e0039221. <https://doi.org/10.1128/mSystems.00392-21>
15. Nagy-Szakal D, Couto-Rodríguez M, Wells HL, Barrows JE, Debieu M, Butcher K, Chen S, Berki A, Hager C, Boorstein RJ, Taylor MK, Jonsson CB, Mason CE, O'Hara NB. 2021. Targeted hybridization capture of SARS-CoV-2 and metagenomics enables genetic variant discovery and nasal microbiome insights. *Microbiol Spectr* 9:e0019721. <https://doi.org/10.1128/Spectrum.00197-21>
16. Kim KW, Deveson IW, Pang CNI, Yeang M, Naing Z, Adikari T, Hammond JM, Stevanovski I, Beukers AG, Verich A, Yin S, McFarlane D, Wilkins MR, Stelzer-Braid S, Bull RA, Craig ME, van Hal SJ, Rawlinson WD. 2021. Respiratory viral co-infections among SARS-CoV-2 cases confirmed by virome capture sequencing. *Sci Rep* 11:3934. <https://doi.org/10.1038/s41598-021-83642-x>
17. Roychoudhury P, Sereewit J, Xie H, Nunley E, Bakhash SM, Lieberman NAP, Greninger AL. 2023. Genomic analysis of early monkeypox virus outbreak strains, Washington, USA. *Emerg Infect Dis* 29:644–646. <https://doi.org/10.3201/eid2903.221446>
18. Duncavage EJ, Magrini V, Becker N, Armstrong JR, Demeter RT, Wylie T, Abel HJ, Pfeifer JD. 2011. Hybrid capture and next-generation sequencing identify viral integration sites from formalin-fixed, paraffin-embedded tissue. *J Mol Diagn* 13:325–333. <https://doi.org/10.1016/j.jmoldx.2011.01.006>
19. Depledge DP, Palser AL, Watson SJ, Lai IY-C, Gray ER, Grant P, Kanda RK, Leproust E, Kellam P, Breuer J. 2011. Specific capture and whole-genome sequencing of viruses from clinical samples. *PLoS One* 6:e27805. <https://doi.org/10.1371/journal.pone.0027805>
20. Koehler JW, Hall AT, Rolfe PA, Honko AN, Palacios GF, Fair JN, Muyembe JJ, Mulembekani P, Schoepp RJ, Adesokan A, Minogue TD. 2014. Development and evaluation of a panel of filovirus sequence capture probes for pathogen detection by next-generation sequencing. *PLoS One* 9:e107007. <https://doi.org/10.1371/journal.pone.0107007>
21. Metsky HC, Siddle KJ, Gladden-Young A, Qu J, Yang DK, Brehio P, Goldfarb A, Piantadosi A, Wohl S, Carter A, et al. 2019. Capturing sequence diversity in metagenomes with comprehensive and scalable probe design. *Nat Biotechnol* 37:160–168. <https://doi.org/10.1038/s41587-018-0006-x>
22. Nieuwenhuijse DF, van der Linden A, Kohl RHG, Sikkema RS, Koopmans MPG, Oude Munnink BB. 2022. Towards reliable whole genome sequencing for outbreak preparedness and response. *BMC Genomics* 23:569. <https://doi.org/10.1186/s12864-022-08749-5>
23. Wylie TN, Wylie KM, Herter BN, Storch GA. 2015. Enhanced virome sequencing using targeted sequence capture. *Genome Res* 25:1910–1920. <https://doi.org/10.1101/gr.191049.115>
24. Chalkias S, Gorham JM, Mazaika E, Parfenov M, Dang X, DePalma S, McKean D, Seidman CE, Seidman JG, Koralnik IJ. 2018. ViroFind: a novel target-enrichment deep-sequencing platform reveals a complex JC virus population in the brain of PML patients. *PLoS One* 13:e0186945. <https://doi.org/10.1371/journal.pone.0186945>
25. Zhan SH, Alamouti SM, Daneshpajouh H, Kwok BS, Lee MH, Khattra J, Houck HJ, Rand KH. 2021. Target capture sequencing of SARS-CoV-2 genomes using the ONETest coronaviruses plus. *Diagn Microbiol Infect Dis* 101:115508. <https://doi.org/10.1016/j.diagmicrobio.2021.115508>
26. Castellet A, Camacho J, Fernández-García MD, Tarragó D. 2023. Shotgun metagenomics to investigate unknown viral etiologies of pediatric meningoencephalitis. *PLoS One* 18:e0296036. <https://doi.org/10.1371/journal.pone.0296036>
27. Lwande OW, Thalín T, de Jong J, Sjödin A, Näslund J, Evander M, Ecke F. 2022. Alphacoronavirus in a daubenton's myotis bat (*Myotis daubentonii*) in Sweden. *Viruses* 14:556. <https://doi.org/10.3390/v14030556>
28. Crispell G, Williams K, Zielinski E, Iwami A, Homas Z, Thomas K. 2022. Method comparison for Japanese encephalitis virus detection in samples collected from the Indo-Pacific region. *Front Public Health* 10:1051754. <https://doi.org/10.3389/fpubh.2022.1051754>
29. Briese T, Kapoor A, Mishra N, Jain K, Kumar A, Jhabdo OJ, Lipkin WI. 2015. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *MBio* 6:e01491-15. <https://doi.org/10.1128/mBio.01491-15>
30. Wen S, Sun C, Zheng H, Wang L, Zhang H, Zou L, Liu Z, Du P, Xu X, Liang L, Peng X, Zhang W, Wu J, Yang J, Lei B, Zeng G, Ke C, Chen F, Zhang X. 2020. High-coverage SARS-CoV-2 genome sequences acquired by target capture sequencing. *J Med Virol* 92:2221–2226. <https://doi.org/10.1002/jmv.26116>
31. Dickson ZW, Hackenberger D, Kuch M, Marzok A, Banerjee A, Rossi L, Klowak JA, Fox-Robichaud A, Mossmann K, Miller MS, Surette MG, Golding GB, Poinar H. 2021. Probe design for simultaneous, targeted capture of diverse metagenomic targets. *Cell Rep Methods* 1:100069. <https://doi.org/10.1016/j.crmeth.2021.100069>
32. Schieble JH, Fox VL, Lennette EH. 1967. A probable new human picornavirus associated with respiratory diseases. *Am J Epidemiol* 85:297–310. <https://doi.org/10.1093/oxfordjournals.aje.a120693>
33. Esposito S, Bosis S, Niesters H, Principi N. 2015. Enterovirus D68 infection. *Viruses* 7:6043–6050. <https://doi.org/10.3390/v7112925>
34. Benschop KS, Albert J, Anton A, Andrés C, Aranzamendi M, Armannsdóttir B, Bailly J-L, Baldanti F, Baldvinsdóttir GE, Beard S, et al. 2021. Re-emergence of enterovirus D68 in Europe after easing the COVID-19 lockdown, September 2021. *Euro Surveill* 26:2100998. <https://doi.org/10.2807/1560-7917.ES.2021.26.45.2100998>
35. Korematsu S, Nagashima K, Sato Y, Nagao M, Hasegawa S, Nakamura H, Sugiura S, Miura K, Okada K, Fujisawa T. 2018. "Spike" in acute asthma exacerbations during enterovirus D68 epidemic in Japan: a nation-wide survey. *Allergol Int* 67:55–60. <https://doi.org/10.1016/j.alit.2017.04.003>
36. Takeuchi S, Kawada JI, Horiba K, Okuno Y, Okumura T, Suzuki T, Torii Y, Kawabe S, Wada S, Ikeyama T, Ito Y. 2019. Metagenomic analysis using next-generation sequencing of pathogens in bronchoalveolar lavage fluid from pediatric patients with respiratory failure. *Sci Rep* 9:12909. <https://doi.org/10.1038/s41598-019-49372-x>
37. Kaida A, Iritani N, Yamamoto SP, Kanbayashi D, Hirai Y, Togawa M, Amo K, Kohdera U, Nishigaki T, Shiomi M, Asai S, Kageyama T, Kubo H. 2017. Distinct genetic clades of enterovirus D68 detected in 2010, 2013, and 2015 in Osaka City, Japan. *PLoS ONE* 12:e0184335. <https://doi.org/10.1371/journal.pone.0184335>

38. Liu B. 2017. Universal PCR primers are critical for direct sequencing-based enterovirus genotyping. *J Clin Microbiol* 55:339–340. <https://doi.org/10.1128/JCM.01801-16>
39. Shirato K, Matsuyama S, Takeda M. 2022. Less frequent sequence mismatches in variants of concern (VOCs) of SARS-CoV-2 in the real-time RT-PCR assays developed by the national institute of infectious diseases, Japan. *Jpn J Infect Dis* 75:96–101. <https://doi.org/10.7883/yoken.JJID.2021.213>
40. WHO. 2021. WHO information for the molecular detection of influenza viruses. Available from: https://cdn.who.int/media/docs/default-source/influenza/molecular-detection-of-influenza-viruses/protocols_influenza_virus_detection_feb_2021.pdf?sfvrsn=df7d268a_5. Retrieved 15 Aug 2024.
41. Goodacre N, Aljanahi A, Nandakumar S, Mikailov M, Khan AS. 2018. A reference viral database (RVDB) to enhance bioinformatics analysis of high-throughput sequencing for novel virus detection. *mSphere* 3:e00069-18. <https://doi.org/10.1128/mSphereDirect.00069-18>