

2018

# Detection of viruses in clinical samples by use of metagenomic sequencing and targeted sequence capture

Kristine M. Wylie

Todd N. Wylie

Richard Buller

Brandi Herter

Maria T. Cannella

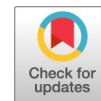
*See next page for additional authors*

---


**Authors**

Kristine M. Wylie, Todd N. Wylie, Richard Buller, Brandi Herter, Maria T. Cannella, and Gregory A. Storch

---



# Detection of Viruses in Clinical Samples by Use of Metagenomic Sequencing and Targeted Sequence Capture

 Kristine M. Wylie,<sup>a,b</sup> Todd N. Wylie,<sup>a,b</sup> Richard Buller,<sup>a</sup> Brandi Herter,<sup>a</sup> Maria T. Cannella,<sup>a</sup> Gregory A. Storch<sup>a</sup>

<sup>a</sup>Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri, USA

<sup>b</sup>The McDonnell Genome Institute, Washington University School of Medicine, St. Louis, Missouri, USA

**ABSTRACT** Metagenomic shotgun sequencing (MSS) is a revolutionary approach to viral diagnostic testing that allows simultaneous detection of a broad range of viruses, detailed taxonomic assignment, and detection of mutations associated with antiviral drug resistance. To enhance sensitivity for virus detection, we previously developed ViroCap, a targeted sequence capture panel designed to enrich nucleic acid from a comprehensive set of eukaryotic viruses prior to sequencing. To demonstrate the utility of MSS with targeted sequence capture for detecting clinically important viruses and characterizing clinically important viral features, we used ViroCap to analyze clinical samples from a diagnostic virology laboratory containing a broad range of medically relevant viruses. From 26 samples, MSS with ViroCap detected all of the expected viruses and 30 additional viruses. Comparing sequencing after capture enrichment with standard MSS, we detected 13 viruses only with capture enrichment and observed a consistent increase in the number and percentage of viral sequence reads as well as the breadth and depth of coverage of the viral genomes. Compared with clinical testing, MSS enhanced taxonomic assignment for 15 viruses, and codons associated with antiviral drug resistance in influenza A virus, herpes simplex virus (HSV), human immunodeficiency virus (HIV), and hepatitis C virus (HCV) could be analyzed. Overall, in clinical samples, MSS with targeted sequence capture provides enhanced virus detection and information of clinical and epidemiologic relevance compared with clinical testing and MSS without targeted sequence capture.

**KEYWORDS** diagnostic testing, sequencing, targeted sequence capture, virology

Metagenomic shotgun sequencing (MSS) has the power to transform diagnostic testing by creating a platform to simultaneously detect and characterize all potential pathogens present in a sample. MSS is not constrained by *a priori* knowledge of which pathogens might be present and is tolerant to sequence variation that could cause targeted molecular tests to fail (1, 2). In addition to pathogen detection, MSS provides information about clinically relevant genomic features, such as taxonomic classification, drug resistance, and virulence factors (1–6).

MSS has been used to characterize the “virome,” the viral component of the microbiome, in a growing number of studies, demonstrating clinical applications in which viral diagnostic testing is expanded or enhanced. In a previous study, we defined specific viruses associated with unexplained fever in young children (7, 8), and MSS allowed us to detect viruses that were not included in the extensive PCR testing that had been carried out on the samples (7–9). MSS has also been used to detect numerous unsuspected viruses from individuals with a variety of clinical syndromes, including encephalitis and fever of unknown origin (10–15). In addition, MSS has been used to demonstrate that anellovirus blooms are associated with immunosuppression in transplant patients and that lower levels of anelloviruses are associated with graft rejection (6, 16, 17). These experiences suggest that expanding the repertoire of current clinical

Received 9 July 2018 Returned for  
modification 6 August 2018 Accepted 14  
September 2018

Accepted manuscript posted online 19  
September 2018

**Citation** Wylie KM, Wylie TN, Buller R, Herter B, Cannella MT, Storch GA. 2018. Detection of viruses in clinical samples by use of metagenomic sequencing and targeted sequence capture. *J Clin Microbiol* 56:e01123-18. <https://doi.org/10.1128/JCM.01123-18>.

**Editor** Yi-Wei Tang, Memorial Sloan Kettering Cancer Center

**Copyright** © 2018 American Society for Microbiology. All Rights Reserved.

Address correspondence to Kristine M. Wylie, [kwylie@wustl.edu](mailto:kwylie@wustl.edu).

tests for viruses could be useful and that sequencing could be a powerful, comprehensive diagnostic tool. A limitation to using MSS for viral diagnostics, however, has been the relatively low sensitivity of detection compared with standard PCR-based molecular tests (5, 7, 18).

We previously reported our development of ViroCap (19), a panel of targeted sequence capture probes that enriches viral nucleic acid from the complete genomes of all vertebrate viruses prior to MSS and increases the sensitivity of virus detection (19, 20). Important advantages compared to PCR are that the viruses detected are not limited by primer and probe sets, that sequence variation of primer and probe sites does not affect detection, and that there is the potential to provide more highly granular information about detected viruses (19). Our aim in this study was to show the impact of using MSS with targeted sequence capture enrichment on viral detection and characterization compared to previous clinical testing on samples submitted to a diagnostic virology laboratory. The samples were selected to represent a broad range of viruses detected in a diagnostic laboratory based in an academic referral hospital. All samples were analyzed by MSS with and without ViroCap to allow us to evaluate the additive value of MSS with targeted sequence capture enrichment for improving virus detection and generating additional relevant clinical information.

## MATERIALS AND METHODS

Samples were tested under a protocol approved by the Washington University School of Medicine Human Research Protection Office for testing of residual material from specimens that had been submitted for routine testing in the Diagnostic Virology Laboratory at St. Louis Children's Hospital. For this study, we selected 26 samples submitted for testing in the Diagnostic Virology Laboratory at St. Louis Children's Hospital, including 24 positive samples that each had one virus detected by that laboratory and 2 with no viruses detected (Table 1). Positive samples were selected to represent a broad range of viruses detectable using routine tests, mostly PCR based, including herpesviruses, polyomaviruses, and common respiratory and gastrointestinal viruses. DNA and RNA viruses were both represented. Sample types included whole blood, plasma, cerebrospinal fluid (CSF), nasopharyngeal (NP) swabs, tracheal aspirates, skin swabs, and stool.

Swabs were collected in 3 ml of universal transport medium and were vortexed to free sample material from the swab. One hundred microliters was removed and loaded on the NucliSENS easyMAG instrument (bioMérieux, Inc., Marcy l'Etoile, France). Total nucleic acid was eluted with 100  $\mu$ l of elution buffer. Stool (100  $\mu$ l of liquid stool or 100  $\mu$ g of formed stool) was added to 900  $\mu$ l of lysis buffer, and the sample was vortexed on the high setting for 10 min at room temperature. Lysis was continued for an additional 10 min at room temperature without vortexing. Samples were centrifuged for 2 min at 14,000 rpm, and 200  $\mu$ l of the lysate was loaded onto the easyMAG instrument. Total nucleic acid was eluted in 70  $\mu$ l of elution buffer. For whole blood, 200  $\mu$ l of blood was combined with 2 ml of warm lysis buffer and vortexed for 30 s to prevent clotting; 100  $\mu$ l of silica from the easyMAG instrument was then added to the sample, followed by vortexing for 15 s. The sample and silica were then added to the easyMAG instrument for extraction, and total nucleic acid was eluted in 100  $\mu$ l of elution buffer. CSF, plasma, and urine samples were loaded directly onto the easyMAG instrument following an onboard lysis protocol. For CSF, 1 ml was extracted and total nucleic acid was eluted in 50  $\mu$ l elution buffer; for plasma and urine, 100  $\mu$ l was extracted and eluted in 100  $\mu$ l elution buffer.

Samples were prepared so that both DNA and RNA viruses could be detected in the same sample, as previously described (19). In brief, 9  $\mu$ l of total nucleic acid from each sample was used as input material. RNA was reverse transcribed with Superscript IV (Thermo Fisher Scientific, Waltham, MA) using random nonamers tagged with a conserved sequence (5'-GTTTCCCAGTCACGATA-3') followed by second-strand synthesis with Sequenase (Thermo Fisher Scientific, Waltham, MA). DNA and cDNA were subsequently amplified using primers targeting the conserved sequence using AccuPrime Taq DNA polymerase (Thermo Fisher Scientific, Waltham, MA) with the following cycling conditions: 94°C for 2 min; 15 cycles of 94°C for 30 s, 40°C for 30 s, 50°C for 30 s, and 72°C for 1 min; and hold at 4°C. The DNA-cDNA mixture was sheared to an average length of 500 bp using the Qsonica Q800R instrument (Qsonica, Newtown, CT). Dual-indexed sequencing libraries were constructed with the Swift Biosciences Accel NGS-2S library kit (Swift Biosciences, Ann Arbor, MI). Sequencing libraries were pooled in two groups consisting of 13 or 14 samples each, and then each pool was divided into two portions. The first portion was sequenced directly on one lane on the Illumina HiSeq 2500 instrument (Illumina, San Diego, CA), generating 125-base paired-end sequences. The second portion was subjected to targeted sequence capture with ViroCap (19) prior to sequencing on a second lane on the Illumina instrument. ViroCap probes were synthesized by Roche NimbleGen (Madison, WI), and capture was carried out according to the manufacturer's instructions. Viral sequences were identified in the sequence data based on nucleotide and translated sequence alignments against all viral reference sequences in the NCBI nucleotide and nonredundant-protein databases, followed by confirmation alignments against the complete NCBI nucleotide and nonredundant-protein databases, as previously described (19, 21). Sequence coverage metrics were determined using RefCov as described previously (19). Due to the clinical focus of this study, we are reporting only vertebrate viruses. To avoid false positives resulting from index swapping during

**TABLE 1** Samples provided by the diagnostic virology laboratory<sup>a</sup>

Sample no.	Sample type	Virus detected in clinical laboratory	Assay <sup>b</sup>	Cycle threshold	Other test(s) performed <sup>c</sup>	Additional virus(es) detected by MSS <sup>d</sup>
1	Dermal swab	HSV-2	LDT	24.5	VZV	<b>HPV 122</b>
2	CSF	HSV-1	Eragen ASR	30.8	Enterovirus, <i>Toxoplasma gondii</i> , EBV	
3	NP	Influenza B virus	RVP	NA	None	
4	Urine	BKPyV	LDT	16.9	None	
5	Plasma	HHV-6	LDT	NA	Adenovirus, CMV, HSV, EBV	<b>HPV 128</b>
6	CSF	JCPyV	LDT	NA	VZV, HSV, CMV, EBV, <i>Toxoplasma gondii</i>	HIV-1
7	NP	Rhinovirus/enterovirus	RVP	NA	None	HHV-7
8	NP	None	NA	NA	None	TTMV, <b>HMPV</b> , HRV-C
9	Plasma	Parvovirus B19	Focus ASR	33.1	None	TTMDV 1
10	NP	RSV	RVP	NA	None	HHV-7, <b>HPV 107</b>
11	CSF	Parechovirus	LDT	36	Enterovirus, HSV	
12	CSF	Enterovirus	GeneXpert	32.7	None	
13	Stool	Norovirus G2	Focus ASR	18.3	Rotavirus antigen	HPV 45
14	Whole blood	None	NA	NA	<i>Ehrlichia/Anaplasma</i>	<b>EBV</b>
15	Whole blood	CMV	Qiagen ASR	27.3	None	<b>HCV 4g</b> , TTV 10
16	NP	Parainfluenza virus 3	RVP	NA	None	
17	NP	Coronavirus NL63	RVP	NA	None	
18	NP	Human metapneumovirus	RVP	NA	None	HHV-6, HHV-7
19	TA	Influenza A virus H3	RVP	NA	None	<b>HPV 23, HPV 48</b> , EBV, TTV
20	Skin scraping	VZV, wild type	LDT	17.9	Unknown	<b>HPV 100</b>
21	Whole blood	CMV	Qiagen ASR	25.2	None	TTV
22	Whole blood	EBV	LDT	34.1	None	<b>HHV-7</b> , TTV 7
23	Whole blood	EBV	LDT	30.7	None	TTMDV
24	NP	<b>Adenovirus</b>	RVP	NA	None	<b>HPV 13, HHV 6</b> , TTMV ALH8
25	Stool	Rotavirus	Antigen assay	NA	Unknown	TTV 13
26	Skin scraping	VZV vaccine strain	LDT	30.2	Unknown	TTMV

<sup>a</sup>Abbreviations: NP, nasopharyngeal; CSF, cerebrospinal fluid; TA, tracheal aspirate; LDT, laboratory-developed test; ASR, analyte-specific reagent; RVP, respiratory virus panel; HSV, herpes simplex virus; VZV, varicella-zoster virus; EBV, Epstein-Barr virus; CMV cytomegalovirus; BKPyV, BK polyomavirus; JCPyV, JC polyomavirus; TTV, torque teno virus; TTMV, torque teno minivirus; TTMDV, torque teno midivirus; NA, not applicable.

<sup>b</sup>PCR unless otherwise indicated.

<sup>c</sup>These tests were ordered by physicians caring for the patients. Results of indicated assays were negative.

<sup>d</sup>Viruses in bold were detected by MSS with enrichment but not without enrichment. All viruses detected in the clinical lab were also detected by MSS without and with enrichment, except for the adenovirus detected by the clinical laboratory in sample 24 (in bold), which was detected by MSS with enrichment but not without enrichment.

capture (22–24), we established a threshold for each virus, consisting of 0.1% of the total reads for that virus in the appropriate pool. This was based on published studies and is consistent with our experiences dealing with dual-indexed sequencing libraries and the HiSeq 2500 platform (22–24). The threshold number of reads for the virus was subtracted from the raw observed number of reads to establish the reported number of sequence reads. Subsequently, samples were considered positive for a virus if the reported number of sequence reads was  $\geq 1$  after subtraction of the threshold.

Single-nucleotide variants in the HIV-1, influenza A virus, and herpesvirus genomes were identified using VarScan2 (25, 26). HIV-1 resistance mutations were evaluated using the HIV Drug Resistance Database (27). Influenza virus variants were evaluated using the Influenza Database Project (28) through the website at <https://www.fludb.org/>. Specifically, we used the “Identify Sequence Features in Segments” tool ([https://www.fludb.org/brc/influenza\\_batch\\_submission.spg?method=NewSFVTAnnotation&decorator=influenza](https://www.fludb.org/brc/influenza_batch_submission.spg?method=NewSFVTAnnotation&decorator=influenza)). Herpesvirus variants were evaluated using the NIAID Virus Pathogen Database and Analysis Resource (ViPR) (29) through the website at <http://www.viprbrc.org> using the Antiviral Resistance Risk Assessment tool for the *Herpesviridae* (<https://www.viprbrc.org/brc/antiviralResistanceRisk.spg?method=ShowCleanInputPage&decorator=herpes>). The hepatitis C virus (HCV) genome was assembled with IDBA-UD (30). Variants and assemblies were manually reviewed with Tablet (31). The influenza B virus lineage was determined based on the sequence as described previously (32). Polyomavirus BK (BKPyV) subtypes were determined using the algorithm of genomic variants described by Morel et al. (33). The metapneumovirus lineage was determined based on the genomic determinants described by van den Hoogen et al. (34). The varicella-zoster virus (VZV) vaccine strain was distinguished based on genomic features described by Loparev et al. (35).

Sequencing results were confirmed using virus-specific PCR assays (8, 35). Epstein-Barr virus (EBV) DNA was measured using the PCR assay described by Wandinger et al. (36). This assay amplifies a segment of the BamHI W segment of the EBV genome. The presence of human herpesvirus 6 (HHV-6) was confirmed using the PCR assay described by Cone et al. (37). This assay amplifies a segment of the large tegument protein gene. The presence of HHV-7 DNA was confirmed using the PCR assay described by Kidd et al. (38). This assay amplifies a segment of the homolog of the HHV-6 U42 gene. The presence of

VZV DNA and the distinction between the wild-type and OKA vaccine viruses were confirmed or determined using the PCR assay described by Loparev et al. (35). This assay is a LightCycler assay that amplifies a 301-bp segment of VZV open reading frame (ORF) 62, which includes a polymorphic site at position 106262. The distinction between wild-type and vaccine viruses is made on the basis of melt curves. The human rhinovirus (HRV) and human metapneumovirus (HMPV) were confirmed using the GenMark Dx respiratory virus panel (GenMark, Carlsbad, CA).

The percentage of viral reads was calculated for each virus in each sample. This represents the number of viral sequences aligned to the reference genome divided by the total number of sequences generated for that sample. The breadth of coverage represents the percentage of the genome length represented by the sequence data. The depth of coverage is the per-position redundancy in sequence reads that align to the genome, and the mean is the average depth calculated across each base of the reference genome. Statistical analysis of the difference in percentage of viral reads, breadth of coverage (percent), and mean depth of coverage between MSS without and with targeted sequence capture was analyzed using the Wilcoxon signed-rank test. Analyses were limited to the viruses detected in the clinical laboratory.

**Accession number(s).** Sequence data for the 26 samples were submitted to the Sequence Read Archive under accession numbers SAMN05713695 to SAMN05713718 and SAMN05713721 to SAMN05713722.

## RESULTS

All samples were tested by standard MSS and MSS enriched using ViroCap targeted sequence capture. Both sequencing approaches detected every virus that had been detected in the clinical laboratory, with the exception of a single NP swab sample in which adenovirus was detected only using targeted sequence capture. In addition, the sequencing methods detected 30 additional viruses (Table 1), most of which were viruses that are not tested for in the clinical laboratories or were from viral tests that were not requested for these samples. Comparing standard sequencing with sequencing after capture enrichment, we observed consistent and substantial increases in the percentage of viral sequence reads obtained and depth and breadth of sequence coverage for each genome (see Table S1 in the supplemental material). These differences were all highly statistically significant (Wilcoxon test,  $P < 0.001$ ). Data for individual segments of viruses with segmented genomes are shown in Table S2 in the supplemental material.

Of the 30 viruses detected by sequencing that had not been detected in the clinical laboratory, 18 (60%) were detected both by MSS alone and by MSS with ViroCap targeted sequence capture, and 12 (40%) were detected only using capture enrichment (Table 1). The most common virus group detected only by sequencing was anelloviruses ( $n = 10$ ), which are not tested for in clinical laboratories, followed by human papillomavirus (HPV) ( $n = 8$ ). The anelloviruses included 5 alphatorqueviruses (torque teno virus [TTV]), 3 betatorqueviruses (torque teno minivirus [TTMV]), and 2 gammatorqueviruses (torque teno midivirus [TTMDV]). The HPVs included multiple types and were detected in diverse samples. Only one was a high-risk type. Notably, HIV was detected in a single CSF sample and HCV was detected in a single whole-blood sample, in each case from patients previously known to be infected with these viruses. Other viruses detected by sequencing but not initial clinical testing included Epstein-Barr virus (EBV) ( $n = 2$ ), HHV-6 ( $n = 2$ ), HHV-7 ( $n = 4$ ), human metapneumovirus (HMPV) ( $n = 1$ ), and human rhinovirus (HRV) ( $n = 1$ ). Testing for these herpesviruses had not been requested in the clinical laboratory for the samples in which these viruses were detected; however, we subsequently confirmed the presence of these viruses by PCR. HMPV and HRV-C were detected in a single NP sample (sample 8) that was included in the set as a sample that was negative after clinical testing, which was done using the BioFire respiratory multiplex panel that includes assays for these viruses. The signal in the sequence data was strong for the HRV ( $>40,000$  sequence reads, with 80% of the genome covered after enrichment) and weaker for the HMPV ( $>100$  sequence reads, with  $\sim 8\%$  of the genome covered after enrichment), which was detected only using capture enrichment (Tables 1 and S1). These viruses were subsequently confirmed using a different multiplex respiratory panel (GenMark). The other negative sample from the diagnostic virology laboratory (sample 14) had been submitted for molecular testing for *Ehrlichia* and *Anaplasma* and had been negative for those agents but had not been tested for viruses. Sequencing results for this sample revealed no viral

**TABLE 2** Enhanced taxonomic identification provided by MSS<sup>a</sup>

Sample no.	Clinical laboratory Identification	Taxonomic identification by:	
		MSS without capture	MSS with capture
3	Influenza B virus	Influenza B virus, Yamagata lineage	Influenza B virus, Yamagata lineage
4	BKPyV	BKPyV subtype 1a	BKPyV subtype 1a
5	HHV-6	HHV-6B	HHV-6B
7	Rhinovirus/enterovirus	Rhinovirus C11	Rhinovirus C11
10	RSV	RSV-A	RSV-A
11	Parechovirus	HPeV-3	HPeV-3
12	Enterovirus	Enterovirus B	Enterovirus B
13	Norovirus GII	Norovirus GII.4	Norovirus GII.4
18	HMPV	HMPV lineage A	HMPV lineage A
19	Influenza A virus H3	Influenza A virus H3N2	Influenza A virus H3N2
22	EBV	EBV	EBV-1
23	EBV	EBV	EBV-1
24	Adenovirus	Not detected	Human mastadenovirus C
24	HHV-6	Not detected	HHV-6B
25	Rotavirus	Human rotavirus A, G12P[8]	Human rotavirus A, G12P[8]

<sup>a</sup>Includes viruses for which MSS provided enhanced taxonomic identification. Abbreviations: HSV, herpes simplex virus; BKPyV, BK polyomavirus; HHV-6, human herpesvirus 6; RSV, respiratory syncytial virus; HPeV, human parechovirus; HMPV, human metapneumovirus; EBV, Epstein-Barr virus.

sequences with standard MSS and 9 sequence reads corresponding to EBV with ViroCap enrichment (Table S1). EBV was subsequently confirmed with a targeted PCR assay.

Sequence analysis allowed us to characterize the viruses detected in greater detail than was possible in the clinical laboratory (Table 2). We were able to determine the subtype of influenza A virus in sample 19 and the lineage of influenza B virus in sample 3. Other viruses identified to a higher taxonomic level than was done by the clinical virology laboratory were RSV-A, HRV-C11, human parechovirus type 3, enterovirus B, norovirus GII.4, human mastadenovirus C, BKPyV type 1, and rotavirus G12P[8]. For most of these, the increased detail was achieved with or without capture, although the two EBVs were identified as EBV-1 only with capture, and the human mastadenovirus C and HHV-6, both in sample 24, were not detected at all by MSS without capture.

We also examined detection of drug resistance by MSS (Table 3). The sequence of the influenza A virus neuraminidase gene indicated that the virus was susceptible to the neuraminidase inhibitors oseltamivir and zanamivir, and the sequence of the gene encoding the matrix 2 protein indicated that the virus was susceptible to adamantanes. The herpes simplex virus 2 (HSV-2) detected in sample 1 had sequences covering the

**TABLE 3** Detection of antiviral resistance genotypes by MSS without and with capture

Sample no.	Virus	Gene	Relevant drug or drug class	No. of codons queried (codon identification)	No. of codons with sequence available from MSS:	
					Without capture	With capture
19	Influenza A virus	Neuraminidase	Neuraminidase inhibitors	6 (V116, E119, I222, Q136, D198, H274)	6	6
1	Herpes simplex virus 2	Matrix 2	Adamantanes	3 (A30, N31, G34)	3	3
		Thymidine kinase (UL23)	Acyclovir	19 (R34, R51, G59, S66, A72, P85, N100, I101, T131, L158, R177, S182, M183, V192, G201, R217, R221, R223, C337)	0	19
6	HIV-1	Polymerase (UL30)	Polymerase inhibitors	8 (E250, A724, S725, S729, L783, D785, L850, D912)	0	8
		Reverse transcriptase	Reverse transcriptase inhibitors	39 (HIV Drug Resistance Database [27])	10	22
		Protease	Protease inhibitors	33 (HIV Drug Resistance Database [27])	0	33
		Integrase	Integrase inhibitors	23 (HIV Drug Resistance Database [27])	22	22
15	Hepatitis C virus genotype 4	NS5A	NS5A inhibitors	3 (L28, M31, T58)	0	2
		Protease (NS3)	Protease inhibitors	2 (Y56, D168)	0	0

thymidine kinase and polymerase genes, and analysis did not reveal any mutations or polymorphisms known to be associated with resistance to acyclovir, foscarnet, or cidofovir (29). Relevant genome coverage for determining antiviral susceptibility/resistance genotypes was not rendered for the HSV-1 detected in sample 2. The HCV genome detected in sample 15 was typed as subtype 4g but was divergent from the HCV references in the NCBI nucleotide database, with the two largest assembled contigs (1.8 kb and 900 bp) matching at 85 to 87% nucleotide identity over their lengths compared with the most similar virus. After assembly, we had contigs that covered NS5A at positions L28 and M31, allowing us to confirm that this virus did not contain the characterized resistance mutations at those positions (39). We did not have coverage of NS5A position T58 or NS3 positions Y56 and D168, so we were unable to determine whether the virus carried mutations for resistance at those sites. The HIV-1 genome had sufficient sequence coverage of the reverse transcriptase (RT)-, integrase-, and protease-coding sequences to assess resistance to each of these classes (27). Specifically, the virus had sequence variants that were associated with potential low-level resistance to the RT inhibitors abacavir and tenofovir, intermediate resistance to zidovudine, and high-level resistance to the nonnucleoside RT inhibitors efavirenz and nevirapine. The sequence also indicated susceptibility to the nucleoside RT inhibitor lamivudine and the nonnucleoside RT inhibitors emtricitabine, etravirine, and rilpivirine. The sequence of the protease gene showed high-level resistance to the protease inhibitors atazanavir and lopinavir and intermediate resistance to darunavir. Finally, the sequence of the polymerase gene also indicated high-level resistance to the integrase strand transfer inhibitors dolutegravir, elvitegravir, and raltegravir. We did not have sequence coverage of the cytomegalovirus (CMV) UL54 or UL97 genes, so we could not assess drug resistance in those genomes. Determination of the resistance genotype was substantially facilitated using capture for the HSV genes and the HIV polymerase and reverse transcriptase genes.

We also detected some viruses by sequencing that were not confirmed by virus-specific PCR and probably represented laboratory contamination. We did not include these in our results in Table 1 or in the total numbers of viruses detected. We detected <10 enterovirus D68 sequence reads without capture and approximately 2,000 reads with capture in 4 samples (5, 6, 9, and 10 sequences) and a single rotavirus read in 1 sample with capture (11 sequences). We had recently been studying both enterovirus D68 and rotavirus in our laboratory, and the sequences were from the same types that we had been examining. To test the possibility that these were laboratory contaminants, we reextracted nucleic acid from the original samples and tested the new extracts by PCR. The presence of the viruses was not confirmed on the newly extracted nucleic acid. We also detected 4 EBV reads in sample 7 with capture, which were not confirmed by virus-specific PCR. Lastly, we detected VZV in 3 samples, 2 with and without capture and 1 only with capture. In those that were detected with and without capture, read numbers were higher with capture. Although the numbers of sequence reads were substantial with capture, the presence of VZV in these samples was not confirmed by virus-specific PCR.

## DISCUSSION

Application of MSS to diagnostic virology has great appeal based on the capability for (i) broad-range detection and (ii) detailed taxonomic characterization (40–47) and for characterization of antiviral resistance genotypes (48). Our findings in this study illustrate and quantitate these advantages using actual clinical samples and measure the further enhancement achieved by the use of targeted sequence capture to improve the sensitivity of standard metagenomic shotgun sequencing.

To assess the gain of yield resulting from adding targeted sequence capture to MSS, we applied these methods to a group of samples known to contain a broad range of viruses detected in diagnostic virology laboratories. Using paired aliquots from the same sequencing libraries, we directly compared the impact of MSS using our standard viral metagenomic sequencing procedure in parallel with the same procedure with the



addition of enrichment with targeted sequence capture using the ViroCap reagent that we have previously described (19). ViroCap contains approximately 2 million capture probes that span the genomes of all vertebrate viruses known at the time the reagent was synthesized in 2014. In the present study, standard MSS increased the yield of viruses detected by 75% over that obtained by standard clinical testing due to the ability of sequencing to broadly detect viruses not specifically included in requested clinical tests, and the addition of ViroCap increased the yield by an additional 50% over standard clinical testing, more than doubling the yield of viruses that were detected in the diagnostic laboratory. While we chose samples containing common clinical viruses for this study, ViroCap has the capability to enrich genomes from uncommon viruses and those with divergent genome sequences (19). The increased yield of viruses included some with obvious clinical significance, most importantly, HIV, which was detected without and with capture in 1 sample, and HCV, which was detected only with capture in 1 sample. While in these cases the viral infections had been previously diagnosed, sequencing has the capability of detecting unsuspected chronic viral infection, such as infection with HIV, HCV, or HBV, when evaluating samples for acute viral infections, which can have important benefits for the patient. As metagenomic sequencing is increasingly applied as a diagnostic tool, it will be important for providers and patients to understand this capability.

Other viruses detected by sequencing but not in the initial clinical tests were HMPV, HRV-C, HHV-6 and -7, EBV, and HPV. Each of these viruses can be pathogenic for humans but also can be found in the absence of symptoms. HMPV and HRV-C were found in an NP sample that was negative when tested by a multiplex molecular panel in the diagnostic laboratory. The HMPV was detected only using targeted capture enrichment and may have been present at a low level that escaped detection by the multiplex assay and MSS without enrichment. The HRV-C was detected both without and with ViroCap, with more than 40,000 reads detected with capture enrichment. It is possible that HRV-C was not detected by the multiplex panel because of sequence differences in the primer binding sites, as has been reported previously for RT-PCR assays directed at human rhinoviruses, which are very diverse (1), and for human parainfluenza virus (43). Because the primer sequences used in the multiplex assay are not available, we cannot directly test this hypothesis. HHV-7 is commonly detected in human samples, usually in the absence of disease manifestations (49). Of the 4 samples with HHV-7 detected by sequencing, 3 were from the respiratory tract and 1 was whole blood. Three were detected without and with targeted sequence capture enrichment, and 1 was detected only with capture enrichment. HHV-6 was detected in two NP samples. In one, viral sequences were detected only with targeted sequence capture enrichment, and in the other, viral sequences were detected both without and with capture enrichment. The pathogenicity of HHV-6 in these two cases cannot be determined by use of available clinical data from these patients. Finally, EBV was detected in one whole-blood sample only with targeted sequence capture and in 1 tracheal aspirate sample without and with capture enrichment.

The most frequently detected virus group in the MSS data that was not detected in the diagnostic laboratory was anelloviruses, which are not tested for in the clinical laboratory. They were detected in 38% of all samples, including diverse sample types such as whole blood, plasma, nasopharyngeal secretions, and stool. This group of viruses is not currently associated with any disease, but their quantitative level in plasma has been identified as being informative regarding the state of the host immune system (17, 50, 51). The family *Anelloviridae* includes 3 genera that are considered human viruses: *Alphatorquevirus* (e.g., TTV), *Betatorquevirus* (e.g., TTMV), and *Gammatorquevirus* (e.g., TTMDV). We detected members of each of these genera. The use of ViroCap did not increase the number of samples positive for anelloviruses but was uniformly associated with substantial increases in the breadth and depth of genome coverage. The anelloviruses have extensive genome heterogeneity (52) We showed previously that ViroCap can enrich anelloviruses with as little as 58% nucleotide sequence identity to the included capture probes (19), so while ViroCap can enrich

anelloviruses with that level of genetic diversity, it is possible that some more-divergent anelloviruses were not enriched.

The second most common family of viruses detected by MSS was papillomaviruses. These viruses were detected in only one sample by standard MSS and in 8 samples with targeted sequence capture enrichment, including diverse sample types such as whole blood, plasma, NP swab, tracheal aspirate, skin, and stool. Each sample contained a different type of papillomavirus, reflecting the known diversity of human papillomaviruses. Even with use of capture enrichment, the number of sequence reads was typically low (see Table S1 in the supplemental material). In blood or plasma samples, it is likely that these viruses represent contamination of samples from the skin, the most common site where papillomaviruses are present, during sample acquisition. While in these cases papillomaviruses were not associated with a clinical presentation, this work demonstrates the utility of using MSS with ViroCap to enhance detection of papillomaviruses. This could have clinical application for detecting high-risk papillomaviruses associated with developing cancers, including head-and-neck and cervical cancers, possibly even when the viruses are present at very low levels and before dysplasia or other clinical phenotypes emerge.

In addition to detecting viruses present in the sample, MSS also produced some sequence reads that were probably the result of laboratory contamination. We suspected that sequence reads represented contamination when the presence of the viral nucleic acid in the original sample was not confirmed by virus-specific PCR. In the cases of enterovirus D68 and rotavirus, the suspicion of contamination was heightened because work with these viruses had been done previously in our laboratory. In the case of VZV, we suspect contamination because VZV was found in 3 different samples and not confirmed by virus-specific PCR. In contrast, in the 2 samples in which VZV was detected in the diagnostic lab and by MSS, the confirmatory virus-specific PCR assay was positive. Finally, the one case of EBV that was not confirmed by virus-specific PCR had a very low number of sequence reads. This discrepancy could reflect contamination but might also result from failure of the virus-specific PCR to detect a very low level of viral nucleic acid or from the presence of a mutation in the region of the genome targeted by the primers. It is also possible that the VZV- and EBV-positive samples that were not confirmed by PCR represent situations in which we detected RNA transcripts from the viruses during sequencing that were not detected with the PCR assays, as they are directed at DNA genomes. Importantly, these findings highlight a current limitation in application of MSS to clinical diagnosis. Defining appropriate procedures and work process controls to control and identify contamination must be a high priority as this technology moves forward toward implementation in clinical medicine.

While targeted capture improves virus detection, there are several limitations to the approach. The first is the time required for the assay. In our experiments, we hybridized the capture probes to the libraries for 72 h and then performed several additional hours of sample handling prior to sequencing and analysis, adding days to the assay. Second, viral genomes that are novel or highly divergent from the capture probe sequences will not be enriched. This does not mean that the viruses could not still be detected in the sequence data, but the signal would not be improved with capture. Finally, the cost of capture probes could be prohibitive if only a few samples were assayed. However, if samples are multiplexed, this cost can be offset.

In summary, this study further demonstrates the potential power of metagenomic sequencing as a viral diagnostic method, confirming its high sensitivity, ability to detect a broad range of viruses, capacity to provide detailed taxonomic identification, and capacity to characterize antiviral resistance genotypes. All of these capabilities were enhanced by including a sequence capture step targeting viruses. Metagenomic sequencing has the capability to bring about a paradigm shift in diagnostic virology. Challenges going forward are to simplify and standardize processes at each step from sample collection to bioinformatics analysis, to control sample cross-contamination, and to lessen cost. A number of challenges related to clinical interpretation of the extensive data provided by MSS also remain, including the significance of viruses that

are not known to be pathogenic or that are detected in unusual body sites and the dilemma raised by detecting unsuspected viruses that may have major clinical implications, such as HIV and HCV. Ideally, future studies should involve the use of sequencing approaches to explore samples in which no pathogen was previously detected and to focus on specific, larger cohorts of patients with similar symptoms or types of infections (e.g., central nervous system infections, gastroenteritis, etc.) who do not have a specific infection diagnosed.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JCM.01123-18>.

**SUPPLEMENTAL FILE 1**, PDF file, 0.1 MB.

## ACKNOWLEDGMENTS

This work was supported by grant R01 AI097213-06 to G.A.S.

The funding agency had no role in study design, interpretation, or the decision to submit the work for publication.

We have no conflicts of interest to report.

We thank the Diagnostic Virology Laboratory at St. Louis Children's Hospital for acquiring samples and performing PCR assays and The McDonnell Genome Institute for generating sequencing data.

## REFERENCES

- Graf EH, Simmon KE, Tardif KD, Hymas W, Flygare S, Eilbeck K, Yandell M, Schlaberg R. 2016. Unbiased detection of respiratory viruses by use of RNA sequencing-based metagenomics: a systematic comparison to a commercial PCR panel. *J Clin Microbiol* 54:1000–1007. <https://doi.org/10.1128/JCM.03060-15>.
- Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, Salamat SM, Somasekar S, Federman S, Miller S, Sokolic R, Garabedian E, Candotti F, Buckley RH, Reed KD, Meyer TL, Seroogy CM, Galloway R, Henderson SL, Gern JE, DeRisi JL, Chiu CY. 2014. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N Engl J Med* 370:2408–2417. <https://doi.org/10.1056/NEJMoa1401268>.
- Zoll J, Rahamat-Langendoen J, Ahout I, Jonge MI de, Jans J, Huijnen MA, Ferwerda G, Warris A, Melchers WJG. 2015. Direct multiplexed whole genome sequencing of respiratory tract samples reveals full viral genomic information. *J Clin Virol* 66:6–11. <https://doi.org/10.1016/j.jcv.2015.02.010>.
- Kothari A, Burgess MJ, Crescencio JCR, Kennedy JL, Denson JL, Schwalm KC, Stoner AN, Kincaid JC, Davies FE, Dinwiddie DL. 2017. The role of next generation sequencing in infection prevention in human parainfluenza virus 3 infections in immunocompromised patients. *J Clin Virol* 92:53–55. <https://doi.org/10.1016/j.jcv.2017.05.010>.
- Thorburn F, Bennett S, Modha S, Murdoch D, Gunson R, Murcia PR. 2015. The use of next generation sequencing in the diagnosis and typing of respiratory infections. *J Clin Virol* 69:96–100. <https://doi.org/10.1016/j.jcv.2015.06.082>.
- Lewandowska DW, Zagordi O, Zbinden A, Schuurmans MM, Schreiber P, Geissberger F-D, Huder JB, Böni J, Benden C, Mueller NJ, Trkola A, Huber M. 2015. Unbiased metagenomic sequencing complements specific routine diagnostic methods and increases chances to detect rare viral strains. *Diagn Microbiol Infect Dis* 83:133–138. <https://doi.org/10.1016/j.diagmicrobio.2015.06.017>.
- Wylie KM, Mihindukulasuriya KA, Sodergren E, Weinstock GM, Storch GA. 2012. Sequence analysis of the human virome in febrile and afebrile children. *PLoS One* 7:e27735. <https://doi.org/10.1371/journal.pone.0027735>.
- Colvin JM, Muenzer JT, Jaffe DM, Smason A, Deych E, Shannon WD, Arens MQ, Buller RS, Lee WM, Weinstock EJS, Weinstock GM, Storch GA. 2012. Detection of viruses in young children with fever without an apparent source. *Pediatrics* 130:e1455–e1462. <https://doi.org/10.1542/peds.2012-1391>.
- Holtz LR, Wylie KM, Sodergren E, Jiang Y, Franz CJ, Weinstock GM, Storch GA, Wang D. 2011. Astrovirus MLB2 viremia in febrile child. *Emerg Infect Dis* 17:2050–2052. <https://doi.org/10.3201/eid1711.110496>.
- Naccache SN, Peggs KS, Mattes FM, Phadke R, Garson JA, Grant P, Samayoa E, Federman S, Miller S, Lunn MP, Gant V, Chiu CY. 2015. Diagnosis of neuroinvasive astrovirus infection in an immunocompromised adult with encephalitis by unbiased next-generation sequencing. *Clin Infect Dis* 60:919–923. <https://doi.org/10.1093/cid/ciu912>.
- Hoffmann B, Tappe D, Höper D, Herden C, Boldt A, Mawrin C, Niederstraßer O, Müller T, Jenckel M, van der Grinten E, Lutter C, Abendroth B, Teifke JP, Cadar D, Schmidt-Chanasit J, Ulrich RG, Beer M. 2015. A variegated squirrel bornavirus associated with fatal human encephalitis. *N Engl J Med* 373:154–162. <https://doi.org/10.1056/NEJMoa1415627>.
- Wilson MR, Suan D, Duggins A, Schubert RD, Khan LM, Sample HA, Zorn KC, Rodrigues Hoffman A, Blick A, Shingde M, DeRisi JL. 2017. A novel cause of chronic viral meningoencephalitis: Cache Valley virus. *Ann Neurol* 82:105–114. <https://doi.org/10.1002/ana.24982>.
- Grard G, Fair JN, Lee D, Slikas E, Steffen I, Muyembe J-J, Sittler T, Veeraraghavan N, Ruby JG, Wang C, Makuwa M, Mulembakani P, Tesh RB, Mazet J, Rimoin AW, Taylor T, Schneider BS, Simmons G, Delwart E, Wolfe ND, Chiu CY, Leroy EM. 2012. A novel rhabdovirus associated with acute hemorrhagic fever in central Africa. *PLoS Pathog* 8:e1002924. <https://doi.org/10.1371/journal.ppat.1002924>.
- Chiu CY, Coffey LL, Murkey J, Symmes K, Sample HA, Wilson MR, Naccache SN, Arevalo S, Somasekar S, Federman S, Stryke D, Vespa P, Schiller G, Messenger S, Humphries R, Miller S, Klausner JD. 2017. Diagnosis of fatal human case of St. Louis encephalitis virus infection by metagenomic sequencing, California, 2016. *Emerg Infect Dis* 23:1964–1968.
- Murkey JA, Chew KW, Carlson M, Shannon CL, Sirohi D, Sample HA, Wilson MR, Vespa P, Humphries RM, Miller S, Klausner JD, Chiu CY. 2017. Hepatitis E virus-associated meningoencephalitis in a lung transplant recipient diagnosed by clinical metagenomic sequencing. *Open Forum Infect Dis* 4:ofx121. <https://doi.org/10.1093/ofid/ofx121>.
- Young JC, Chehoud C, Bittinger K, Bailey A, Diamond JM, Cantu E, Haas AR, Abbas A, Frye L, Christie JD, Bushman FD, Collman RG. 2015. Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients. *Am J Transplant* 15:200–209. <https://doi.org/10.1111/ajt.13031>.
- De Vlaminck I, Khush KK, Strehl C, Kohli B, Luikart H, Neff NF, Okamoto J, Snyder TM, Cornfield DN, Nicolls MR, Weill D, Bernstein D, Valentine HA, Quake SR. 2013. Temporal response of the human virome to immunosuppression and antiviral therapy. *Cell* 155:1178–1187. <https://doi.org/10.1016/j.cell.2013.10.034>.
- Lim ES, Zhou Y, Zhao G, Bauer IK, Droit L, Ndao IM, Warner BB, Tarr PI, Wang D, Holtz LR. 2015. Early life dynamics of the human gut virome and

- bacterial microbiome in infants. *Nat Med* 21:1228–1234. <https://doi.org/10.1038/nm.3950>.
19. 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>.
  20. Briese T, Kapoor A, Mishra N, Jain K, Kumar A, Jabado 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>.
  21. Wylie KM, Mihindukulasuriya KA, Zhou Y, Sodergren E, Storch GA, Weinstock GM. 2014. Metagenomic analysis of double-stranded DNA viruses in healthy adults. *BMC Biol* 12:71. <https://doi.org/10.1186/s12915-014-0071-7>.
  22. Kircher M, Heyn P, Kelso J. 2011. Addressing challenges in the production and analysis of illumina sequencing data. *BMC Genomics* 12:382. <https://doi.org/10.1186/1471-2164-12-382>.
  23. Kircher M, Sawyer S, Meyer M. 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the illumina platform. *Nucleic Acids Res* 40:e3. <https://doi.org/10.1093/nar/gkr771>.
  24. Sinha R, Stanley G, Gulati GS, Ezran C, Travaglini KJ, Wei E, Chan CKF, Nabhan AN, Su T, Morganti RM, Conley SD, Chaib H, Red-Horse K, Longaker MT, Snyder MP, Krasnow MA, Weissman IL. 2017. Index switching causes “spreading-of-signal” among multiplexed samples in illumina HiSeq 4000 DNA sequencing. *bioRxiv* <https://doi.org/10.1101/125724>.
  25. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, Weinstock GM, Wilson RK, Ding L. 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25:2283–2285. <https://doi.org/10.1093/bioinformatics/btp373>.
  26. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 22:568–576. <https://doi.org/10.1101/gr.129684.111>.
  27. Tang MW, Liu TF, Shafer RW. 2012. The HIVdb system for HIV-1 genotypic resistance interpretation. *Intervirology* 55:98–101. <https://doi.org/10.1159/000331998>.
  28. Squires RB, Noronha J, Hunt V, Garcia-Sastre A, Macken C, Baumgarth N, Suarez D, Pickett BE, Zhang Y, Larsen CN, Ramsey A, Zhou L, Zaremba S, Kumar S, Deitrich J, Klem E, Scheuermann RH. 2012. Influenza research database: an integrated bioinformatics resource for influenza research and surveillance. *Influenza Other Respir Viruses* 6:404–416. <https://doi.org/10.1111/j.1750-2659.2011.00331.x>.
  29. Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, Liu M, Kumar S, Zaremba S, Gu Z, Zhou L, Larson CN, Dietrich J, Klem EB, Scheuermann RH. 2012. ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res* 40:D593–D598. <https://doi.org/10.1093/nar/gkr859>.
  30. Peng Y, Leung HCM, Yiu SM, Chin FYL. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28:1420–1428. <https://doi.org/10.1093/bioinformatics/bts174>.
  31. Milne I, Bayer M, Stephen G, Cardle L, Marshall D. 2016. Tablet: visualizing next-generation sequence assemblies and mappings. *Methods Mol Biol* 1374:253–268. [https://doi.org/10.1007/978-1-4939-3167-5\\_14](https://doi.org/10.1007/978-1-4939-3167-5_14).
  32. Arvia R, Corcioli F, Pierucci F, Azzi A. 2014. Molecular markers of influenza B lineages and clades. *Viruses* 6:4437–4446. <https://doi.org/10.3390/v6114437>.
  33. Morel V, Martin E, François C, Helle F, Faucher J, Mourez T, Choukroun G, Duverlie G, Castelain S, Brochot E. 2017. A simple and reliable strategy for BK virus subtyping and subgrouping. *J Clin Microbiol* 55:1177–1185. <https://doi.org/10.1128/JCM.01180-16>.
  34. van den Hoogen BG, Herfst S, Sprong L, Cane PA, Forleo-Neto E, de Swart RL, Osterhaus ADME, Fouchier RAM. 2004. Antigenic and genetic variability of human metapneumoviruses. *Emerg Infect Dis* 10:658–666. <https://doi.org/10.3201/eid1004.030393>.
  35. Loparev VN, McCaustland K, Holloway BP, Krause PR, Takayama M, Schmid DS. 2000. Rapid genotyping of varicella-zoster virus vaccine and wild-type strains with fluorophore-labeled hybridization probes. *J Clin Microbiol* 38:4315–4319.
  36. Wandinger K, Jabs W, Siekhaus A, Bubel S, Trillenberger P, Wagner H, Wessel K, Kirchner H, Hennig H. 2000. Association between clinical disease activity and Epstein-Barr virus reactivation in MS. *Neurology* 55:178–184. <https://doi.org/10.1212/WNL.55.2.178>.
  37. Cone RW, Huang ML, Hackman RC, Corey L. 1996. Coinfection with human herpesvirus 6 variants A and B in lung tissue. *J Clin Microbiol* 34:877–881.
  38. Kidd IM, Clark DA, Ait-Khaled M, Griffiths PD, Emery VC. 1996. Measurement of human herpesvirus 7 load in peripheral blood and saliva of healthy subjects by quantitative polymerase chain reaction. *J Infect Dis* 174:396–401. <https://doi.org/10.1093/infdis/174.2.396>.
  39. Schnell G, Tripathi R, Beyer J, Reisch T, Krishnan P, Lu L, Dekhtyar T, Hall C, Vilchez RA, Pilot-Matias T, Collins C. 2015. Hepatitis C virus genotype 4 resistance and subtype demographic characterization of patients treated with ombitasvir plus paritaprevir/ritonavir. *Antimicrob Agents Chemother* 59:6807–6815. <https://doi.org/10.1128/AAC.01229-15>.
  40. Lecuit M, Eloit M. 2015. The potential of whole genome NGS for infectious disease diagnosis. *Expert Rev Mol Diagn* 15:1517–1519. <https://doi.org/10.1586/14737159.2015.1111140>.
  41. Barzon L, Lavezzo E, Costanzi G, Franchin E, Toppo S, Palù G. 2013. Next-generation sequencing technologies in diagnostic virology. *J Clin Virol* 58:346–350. <https://doi.org/10.1016/j.jcv.2013.03.003>.
  42. Lefterova MI, Suarez CJ, Banaei N, Pinsky BA. 2015. Next-generation sequencing for infectious disease diagnosis and management: a report of the Association for Molecular Pathology. *J Mol Diagn* 17:623–634. <https://doi.org/10.1016/j.jmoldx.2015.07.004>.
  43. Xu L, Zhu Y, Ren L, Xu B, Liu C, Xie Z, Shen K. 2017. Characterization of the nasopharyngeal viral microbiome from children with community-acquired pneumonia but negative for Luminex xTAG respiratory viral panel assay detection. *J Med Virol* 89:2098–2107. <https://doi.org/10.1002/jmv.24895>.
  44. Parker J, Chen J. 2017. Application of next generation sequencing for the detection of human viral pathogens in clinical specimens. *J Clin Virol* 86:20–26. <https://doi.org/10.1016/j.jcv.2016.11.010>.
  45. Yozwiak NL, Skewes-Cox P, Stenglein MD, Balmaseda A, Harris E, DeRisi JL. 2012. Virus identification in unknown tropical febrile illness cases using deep sequencing. *PLoS Negl Trop Dis* 6:e1485. <https://doi.org/10.1371/journal.pntd.0001485>.
  46. Lewandowska DW, Schreiber PW, Schuurmans MM, Ruehe B, Zagordi O, Bayard C, Greiner M, Geissberger FD, Capaul R, Zbinden A, Böni J, Benden C, Mueller NJ, Trkola A, Huber M. 2017. Metagenomic sequencing complements routine diagnostics in identifying viral pathogens in lung transplant recipients with unknown etiology of respiratory infection. *PLoS One* 12:e0177340. <https://doi.org/10.1371/journal.pone.0177340>.
  47. Li L, Deng X, Linsuwanon P, Bangsberg D, Bwana MB, Hunt P, Martin JN, Deeks SG, Delwart E. 2013. AIDS alters the commensal plasma virome. *J Virol* 87:10912–10915. <https://doi.org/10.1128/JVI.01839-13>.
  48. Wang CH, Nie K, Zhang Y, Wang J, Zhou SF, Li XN, Zhou HY, Qi SX, Ma XJ. 2017. An improved barcoded oligonucleotide primers-based next-generation sequencing approach for direct identification of viral pathogens in clinical specimens. *Biomed Environ Sci* 30:22–34. <https://doi.org/10.3967/bes2017.003>.
  49. Caserta MT, Krug LT, Pellett PE. 2014. Roseoloviruses: unmet needs and research priorities: perspective. *Curr Opin Virol* 9:167–169. <https://doi.org/10.1016/j.coviro.2014.10.005>.
  50. Görzer I, Jaksch P, Kundi M, Seitz T, Klepetko W, Puchhammer-Stöckl E. 2015. Pre-transplant plasma Torque Teno virus load and increase dynamics after lung transplantation. *PLoS One* 10:e0122975. <https://doi.org/10.1371/journal.pone.0122975>.
  51. Focosi D, Antonelli G, Pistello M, Maggi F. 2016. Torquetenovirus: the human virome from bench to bedside. *Clin Microbiol Infect* 22:589–593. <https://doi.org/10.1016/j.cmi.2016.04.007>.
  52. Hino S. 2002. TTV, a new human virus with single stranded circular DNA genome. *Rev Med Virol* 12:151–158. <https://doi.org/10.1002/rmv.351>.