



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Direct multiplexed whole genome sequencing of respiratory tract samples reveals full viral genomic information



Jan Zoll<sup>a,\*</sup>, Janette Rahamat-Langendoen<sup>a</sup>, Inge Ahout<sup>b</sup>, Marien I.de Jonge<sup>b</sup>, Jop Jans<sup>b</sup>, Martijn A. Huijnen<sup>c</sup>, Gerben Ferwerda<sup>b</sup>, Adilia Warris<sup>b,1</sup>, Willem J.G. Melchers<sup>a</sup>

<sup>a</sup> Department of Medical Microbiology, Radboudumc, Nijmegen, The Netherlands

<sup>b</sup> Laboratory of Pediatric Infectious Diseases, Department of Pediatrics, Radboudumc, Nijmegen, The Netherlands

<sup>c</sup> Center for Molecular and Biomolecular Informatics, Radboudumc, Nijmegen, The Netherlands

### ARTICLE INFO

#### Article history:

Received 27 October 2014

Received in revised form 3 February 2015

Accepted 16 February 2015

#### Keywords:

Virus

Respiratory tract

Infection

Whole genome sequencing

### ABSTRACT

**Background:** Acute respiratory tract infections (RTI) cause substantial morbidity during childhood, and are responsible for the majority of pediatric infectious diseases. Although most acute RTI are thought to be of viral origin, viral etiology is still unknown in a significant number of cases.

**Objectives:** Multiplexed whole genome sequencing (WGS) was used for virome determination directly on clinical samples as proof of principle for the use of deep sequencing techniques in clinical diagnosis of viral infections.

**Study design:** WGS was performed with nucleic acids from sputum and nasopharyngeal aspirates from four pediatric patients with known respiratory tract infections (two patients with human rhinovirus, one patient with human metapneumovirus and one patient with respiratory syncytial virus), and from four pediatric patients with PCR-negative RTI, and two control samples.

**Results:** Viral infections detected by routine molecular diagnostic methods were confirmed by WGS; in addition, typing information of the different viruses was generated. In three out of four samples from pediatric patients with PCR-negative respiratory tract infections and the two control samples, no causative viral pathogens could be detected. In one sample from a patient with PCR-negative RTI, rhinovirus type-C was detected. Almost complete viral genomes could be assembled and in all cases virus species could be determined.

**Conclusions:** Our study shows that, in a single run, viral pathogens can be detected and characterized, providing information for clinical assessment and epidemiological studies. We conclude that WGS is a powerful tool in clinical virology that delivers comprehensive information on the viral content of clinical samples.

© 2015 Elsevier B.V. All rights reserved.

## 1. Background

Acute respiratory tract infections (RTI) are a major cause of hospitalization of children below 5 years of age [1]. Respiratory tract infections in these young children are associated with high morbidity and mortality [2]. Over 65% of respiratory infections are caused by viruses [3]. In the last decades, improvement of molecular techniques has resulted in the discovery of new viruses. However, the viral etiology of RTI is still unknown in a significant number of cases [3]. Current diagnostic methods for virus detection are

mainly based on molecular amplification techniques such as PCR, using large panels of the currently known pathogens and frequently detected viruses. Therefore, new viral pathogens might be missed. Moreover, it was estimated that more than 50 percent of the human virus species still await discovery [4].

Next generation sequencing (NGS) techniques have been applied in public health microbiology for outbreak monitoring and for metagenomic studies [5,6]. Additionally, NGS has become a powerful approach in the discovery of new viruses [7–9]. The increasing performance of bench top sequencers such as the Illumina MiSeq and the Ion Torrent PGM is associated with an ongoing reduction in costs. Translation of deep sequencing techniques into routine virus diagnostics on clinical samples seems a logical next step, thereby not only broadening the range of viruses that can be detected but also providing additional characterization of the detected viruses.

\* Corresponding author. Tel.: +31 24 3668259.

E-mail address: [jan.zoll@radboudumc.nl](mailto:jan.zoll@radboudumc.nl) (J. Zoll).

<sup>1</sup> Department of Medical Microbiology, Radboudumc, PO Box 9101, Internal Post: 7770,6500HB Nijmegen, The Netherlands.

**Table 1**

Overview of the clinical samples and patient characteristics used in this study.

Sample material <sup>a</sup> number	Age(days)	Gender	Gest. age <sup>b</sup>	Symptoms	Severity <sup>c</sup> O <sub>2</sub>	Diagnostic results <sup>d</sup>	PCR Ct value	DNA/RNA processing <sup>f</sup>
1 Sputum	10	F	25+5	Apnea	O <sub>2</sub>	Negative		1
2 Sputum	42	M	36+6	Apnea and bradycardia	MV	hRV	25	1
3 Sputum	122	M	25+5	Respiratory failure in infant with underlying disease	O <sub>2</sub>	hMPV	23	1
4 Sputum	4	M	35+5	Respiratory difficulties		Negative		1
5 NPA	329	F	39+6	Mild dyspnea	-	Negative		2
6 NPA	827	F	40+0	Moderate dyspnea	O <sub>2</sub>	RSV <sup>e</sup>		2
7 NPA	406	F	39+3	Moderate dyspnea	O <sub>2</sub>	hRV	>40	2
8 NPA	158	F	28+0		O <sub>2</sub>	Negative		2
9 NPA	Adult	M				Negative		1
10 NPA	Adult	M				Negative		2

<sup>a</sup> NPA: nasopharyngeal aspirate.<sup>b</sup> Gestational age in weeks + days.<sup>c</sup> Severity: O<sub>2</sub> oxygen need; MV mechanical ventilation.<sup>d</sup> Diagnostic results, hRV: human rhinovirus; hMPV: human metapneumovirus; RSV: respiratory syncytial virus.<sup>e</sup> Sample 6 was found positive in the RSV rapid test.<sup>f</sup> Sample processing method 1: Reverse transcription and random PCR according to Zoll et al., 2009 (10). Sample processing method 2: Human ribosomal RNA depletion and subsequent whole genome amplification (WGA) en whole transcriptome amplification (WTA).

## 2. Objectives

In this report, we describe the virome determination directly on clinical samples by multiplexed whole-genome sequencing as a proof of principle for the use of deep sequencing techniques in viral diagnostics.

### 2.1. Study design

#### 2.1.1. Patient samples

Nasopharyngeal-aspirate and sputum samples were obtained from 8 pediatric patients with moderate to severe RTI. Additionally, two nasopharyngeal-aspirate samples from healthy adults were used as control samples (Table 1). All samples were tested with a multiplex-PCR assay for a panel of respiratory pathogens including respiratory syncytial virus (RSV), influenza virus-A and B, adenovirus, *Mycoplasma pneumoniae*, *Chlamydomphila pneumoniae*, bocavirus, parainfluenzavirus 1–4 (PIV), coronavirus OC43 and 229 E, enterovirus, human metapneumovirus (hMPV), and human rhinovirus (hRV). Samples were processed according to the scheme depicted in Fig. 1.

#### 2.1.2. Nucleic acid isolation

DNA and RNA were isolated directly from clinical samples. Either 200 µl virus transport medium or sputum was centrifuged for 10 min. at 10,000xg in order to remove cellular debris. The supernatant was treated with RNase-One (Promega) and Turbo DNase-One (Life Technologies) according to the manufacturer's protocols. Total nucleic acid was purified using Qiagen DNeasy Blood kit according to the manufacturer's protocol. Nucleic acids were eluted with 100 µl elution buffer. Two different approaches were used for the pre-amplification of the viral genomes. Half of the samples were processed using a random amplification method that was successfully used in previous experiments [10]. The limited volume of the clinical samples may result in low total virus content and therefore, a second approach for pre-amplification was used. Nucleic acid fractions from four clinical samples from pediatric patients and a control sample were used as input for the amplification using commercially available kits, originally developed for single-cell whole-genome or whole-transcriptome amplification.

#### 2.1.3. Ribosomal RNA depletion and reverse transcription reaction

Virus identification by NGS requires enrichment of viral particles. Presence of human and bacterial cells will cause an overrepresentation of non-viral reads, especially rRNA.

Therefore, human rRNA from the purified nucleic acids fractions was depleted. Human rRNA was removed from 50 µl of the purified nucleic acid fraction using the GeneRead rRNA-depletion kit (Qiagen, Valencia (CA), USA) according to the manufacturer's protocol. rRNA-depleted RNA was purified with the RNeasy minikit (Qiagen, Valencia (CA), USA). RNA was eluted in 15 µl elution buffer. For the reverse transcription reaction, 13 µl RNA was mixed with 8 µl lysis-buffer supplied with the Qiagen Repli-G Cell WGA and WTA kit and incubated for 5 min. at 24 °C, 3 min at 95 °C and cooled on ice. Reverse transcription was performed on 10 µl RNA using the Qiagen Repli-G Cell WGA and WTA kit according to the manufacturer's protocol.

#### 2.1.4. Amplification and sequencing of the DNA and cDNA samples

The DNA and cDNA fractions obtained after rRNA-depletion and reverse transcription were used for whole-genome amplification (WGA) and whole-transcriptome amplification (WTA) in which the WTA approach was applied for the detection of RNA viruses, using the Qiagen Repli-G Cell WGA and WTA kit respectively. WGA and WTA fractions from individual samples were used for sequencing separately. Approximately 100 ng DNA per sample was used for sequencing on IonTorrent PGM system with an Ion 318 sequencing chip (Life Technologies, Waltham (MA), USA). Sequencing results were analyzed as depicted in Fig. 1.

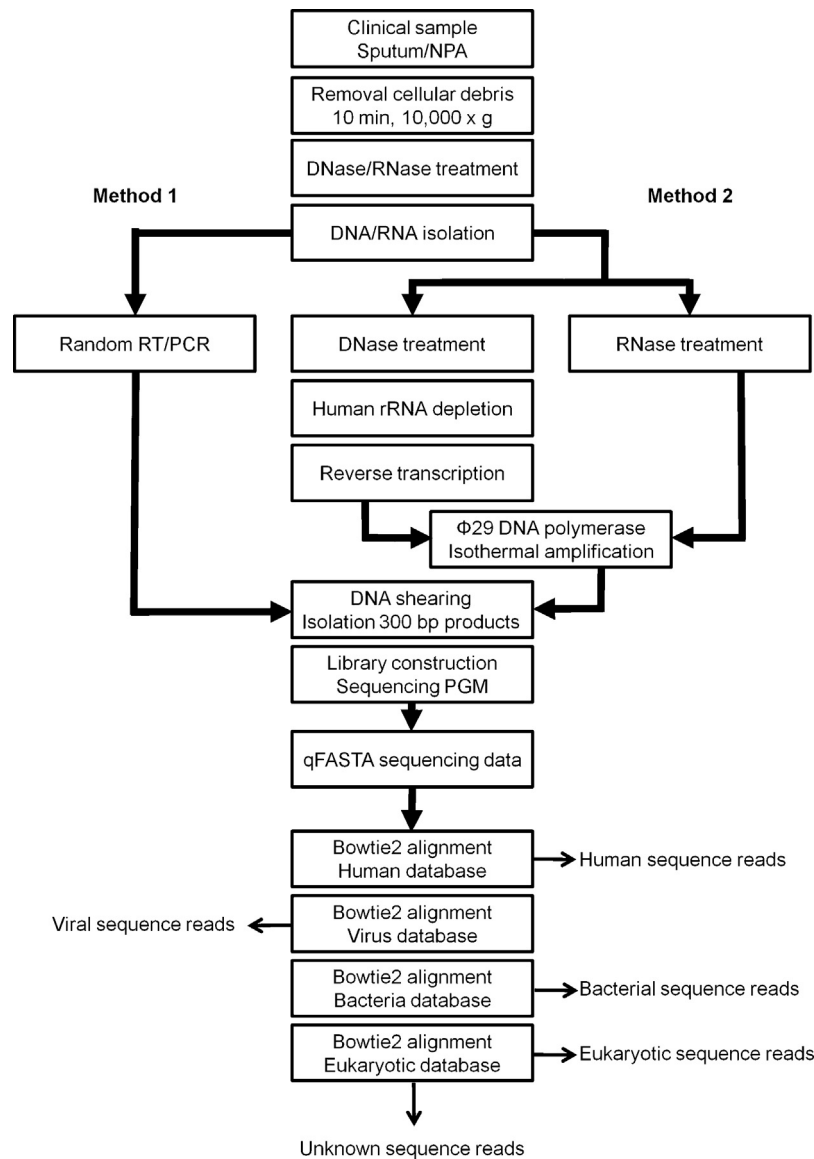
#### 2.1.5. Assembly of virus genomes

Partial or full virus genomes were assembled from selected reads covering the most significant reference sequences found during the virus identification analysis. Assembly and determination of sequence coverage was done with Bowtie2 using the reference sequence as template and the default setting "local-sensitive" switched on [11]. Sequence depth was calculated using Samtools available at the Galaxy web server [12,13].

## 3. Results

### 3.1. Molecular diagnosis

Eight respiratory tract samples from pediatric patients with acute RTI were tested by multiplex-PCR of which two samples were found positive for hRV, one sample for RSV, and one sample was found positive for hMPV. Four samples remained negative in PCR. The two control samples from healthy adults were also found negative (Table 1).



**Fig. 1.** Schematic representation of the sample preparation. For each sample dataset reads were processed in two steps. First, a human genome-database was constructed using NCBI GRCh38 dataset. Human sequences were removed by mapping the reads to the human genome using Bowtie2 version 2.1.0 with the default setting “local-sensitive” switched on (11). A database was constructed with complete genomic sequences from mammalian viruses, collected from NCBI, ViralZone and EMBL. Reads that remained after removal of human sequences were mapped to the viral database using Bowtie2. Distributions of hits matching viral sequences were processed manually. Additionally, the remaining reads were mapped against datasets containing bacterial and eukaryotic sequences using Bowtie2.

### 3.2. Sequencing of clinical samples

From all samples, sufficient amounts of DNA and RNA were obtained to perform whole-genome sequencing for the detection of DNA viruses, or whole-transcriptome sequencing for the detection of RNA viruses. All reads were used as input for the virus identification. Additionally, reads were also mapped to bacterial and eukaryotic sequences (Tables 2 and 3). With an average read length of 150 bp, the settings used for Bowtie2-alignment allows a sequence variation of approximately 26%. In all samples, the vast majority of the reads was of human or bacterial origin. The bacterial sequences included genomic as well as plasmid and phage sequences. A significant amount of bacterial sequences originated from ribosomal-DNA. This is the direct consequence of the fact that the samples were human rRNA-depleted only, whereas bacterial rRNA was still present during the reverse transcription reaction. A significant number of reads could not be assigned to any known

viral sequences. These sequences are classified as unknown and may represent fragments of genomes of unknown viral species.

### 3.3. Virus-classification of reads

The two different approaches for pre-amplification of nucleic acid extracts resulted in different efficiencies in virus detection. Although the result of virus detection in clinical samples by direct-sequencing are consistent with the results obtained with conventional molecular diagnostic methods for all samples, pre-amplification resulted in ten times more reads that could be assigned to viral genomic sequences when using the Repli-g systems instead of random RT/PCR. Consequently, more sequence information was obtained and almost complete virus genomes could be assembled with reads from positive samples. The results of virus classification of sequencing reads are shown in Fig. 2 and Table 4. Sample 2 contained reads covering a major part of the hRV genome that could be classified as hRV-A38. Sample 3

**Table 2**  
Classification of WGS sequence reads.

Sample material <sup>a</sup> number	Number of reads	Human	Bacteria	Viruses	Others	Unknown
1 sputum	851946	657965	9683	11700	137	172461
2 sputum	688295	567152	17078	11989	12	92064
3 sputum	629721	464131	13729	10411	2628	138822
4 sputum	643957	503513	15088	13299	360	111697
5 NPA	300649	260855	1273	29	1629	36863
6 NPA	289961	234659	18962	24	3375	32941
7 NPA	390579	236274	10891	1179	3581	138654
8 NPA	251208	66980	150082	61	930	33155
9 NPA	713577	527817	11494	11594	49	162623
10 NPA	491410	383045	5238	4	8	103115

<sup>a</sup> NPA: nasopharyngeal aspirate.

**Table 3**  
Classification of WTA sequence reads.

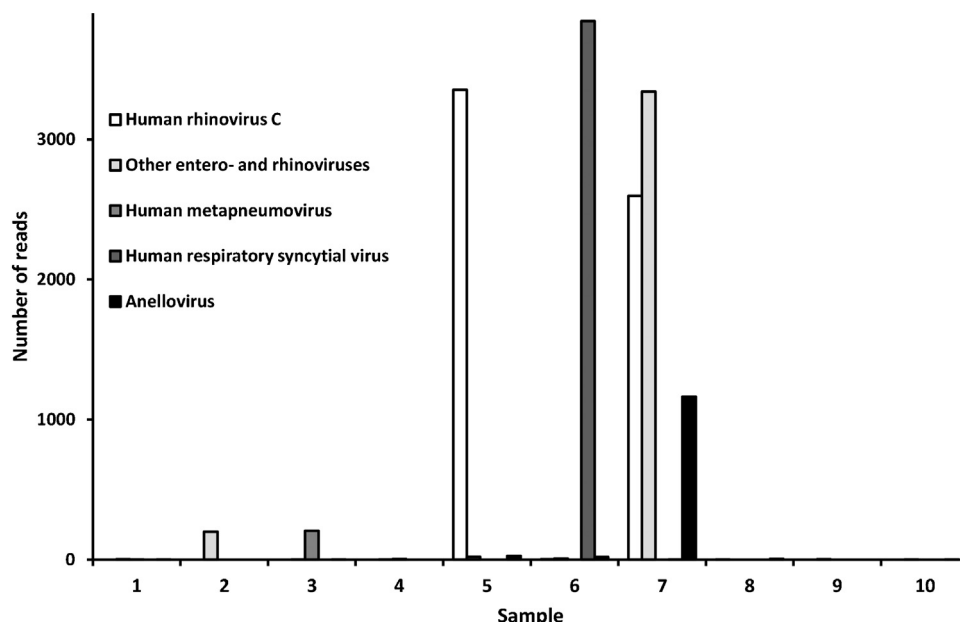
Sample material <sup>a</sup> number	Number of reads	Human	Bacteria	Viruses	Others	Unknown
1 sputum	142052	39785	64914	214	29	37110
2 sputum	137258	27527	38565	233	14	70919
3 sputum	144250	38536	44049	372	1008	60285
4 sputum	107629	13732	48325	491	149	44932
5 NPA	187122	28234	122391	4025	4987	27485
6 NPA	218906	48521	155652	4452	5124	5157
7 NPA	254716	42116	175950	6324	19822	10504
8 NPA	255191	14533	218825	197	2548	19088
9 NPA	142846	52215	69538	90	33	20970
10 NPA	108559	30668	41814	106	29	35942

<sup>a</sup> NPA: nasopharyngeal aspirate.

contained a significant amount of reads that could be mapped to hMPV. Sample 5 contained reads mapping to hRV-C. In the diagnostic multiplex-PCR, this sample tested negative. Analysis of the viral sequence obtained from sample 5 revealed several mismatches between the forward primer used in PCR and the target sequence in the 5'-untranslated region (5'-UTR) of hRV-C, explaining the negative-PCR result. Respiratory syncytial virus sequences were detected in sample 6. The virus found in this sample could be classified as RSV-typeA. Finally, sample 7 contained reads that could be assigned to hRV-C. Whereas in sample 5 all picornavirus associated sequences could be assembled to a hRV-C genome, sample 7 contained significant numbers of reads mapped to other

human enteroviruses and hRV-A genome. Most probably, this patient was infected by enteroviruses and different rhinoviruses as well.

Interestingly, reads assigned to other viral pathogens were found in a number of samples. In three samples reads were found that could be assigned to members of the anelloviridae family. Especially sample 7 contained significant amounts of reads that could be mapped to various anelloviridae or torque-teno virus (TTV) genera, including torque-teno-like mini virus (TTMV) and torque-teno-like midi virus (TTMDV). Almost complete anellovirus genomes could be assembled using the reads from sample 7 and were identified as TTV genotype-19, TTMV genotype-1, and TTMDV genotype-2.



**Fig. 2.** Virus classification of sequence reads.

**Table 4**  
Overview of the diagnostic results obtained in this study.

Sample material <sup>a</sup> number	Diagnostic results <sup>b</sup>	PCR Ct value	DNA/RNA processing <sup>d</sup>	WGS results	% Genome detected	Number of reads	Sequence depth <sup>e</sup>
1 sputum	Negative		1	Negative	NA	NA	NA
2 sputum	hRV	25	1	hRV38	91%	191	4
3 sputum	hMPV	23	1	hMPV	38%	206	5
4 sputum	Negative		1	Negative	NA	NA	NA
5 NPA	Negative		2	hRVC	99%	3354	72
6 NPA	RSV <sup>c</sup>		2	RSV	91%	3844	42
7 NPA	hRV	>40	2	hRVC	87%	2597	63
8 NPA	Negative		2	Negative	NA	NA	NA
9 NPA	Negative		1	Negative	NA	NA	NA
10 NPA	Negative		2	Negative	NA	NA	NA

<sup>a</sup> NPA: nasopharyngeal aspirate.

<sup>b</sup> Diagnostic results, hRV: human rhinovirus; hMPV: human metapneumovirus; RSV: respiratory syncytial virus.

<sup>c</sup> Sample 6 was found positive in the RSV rapid test.

<sup>d</sup> Sample processing method 1: Reverse transcription and random PCR according to Zoll et al., 2009 [10]. Sample processing method 2: Human ribosomal RNA depletion and subsequent whole genome amplification (WGA) or whole transcriptome amplification (WTA).

<sup>e</sup> Sequence depth: average number of times a base is read in the detected viral genomic region.

#### 4. Discussion

Here, we present a proof of principle study with clinical samples obtained from pediatric patients with moderate to severe RTI. Using standard molecular diagnostics, viral pathogens were detected in only half of the samples. Direct multiplexed whole-genome sequencing of nucleic acids extracted from these samples confirmed the results obtained by PCR. Thus, sensitivity of the WGS approach was at least similar to the sensitivity of the diagnostic PCR in this set of samples. Moreover, of the 4 patient samples tested negative in PCR, one sample revealed hRV-typeC using NGS. Due to the heterogeneous character of the hRV target region for detection by PCR, positive samples containing hRV will be missed using routine molecular diagnostics. Therefore, sequence-independent detection methods like WGS will enhance a causative diagnosis of viral RTI.

Recently, the use of NGS in diagnostics of viral respiratory tract infections was reported [14]. The described NGS based method was at least as sensitive as RT/PCR, confirming our results. The method for pre-amplification of input nucleic acid was similar to our method of random-amplification by PCR. Similar to our results, this method gave relatively small numbers of virus reads.

Two different methods of sample-processing were tested in this study. The amplification of DNA or cDNA with the high-fidelity enzyme Phi29 DNA-polymerase combined with DNA ligation provided optimal results and high sequence coverages were reached. Almost complete viral genomic sequences could be retrieved from clinical samples that tested positive in PCR with high Ct values. Despite the low viral load in these samples, direct-sequencing provided enough information to classify the detected viruses into viral strains. For small RNA viruses like picornaviruses and RSV, direct-sequencing of nucleic acid prepared from clinical samples can result in nearly complete viral-genomic data. Further improvements can be achieved by depletion of both human and bacterial rRNA.

Using WGS directly on clinical samples broadens the range of virus recovery (also viruses unable to grow in cell cultures or unknown viruses can be detected), provides the ability to characterize the detected viruses at the same time, and will speed-up the diagnostic process. A major drawback of the current viral molecular diagnostic tools is that they employ a stepwise approach for detection and typing, and that only known specific targets can be detected resulting in false negative results due to the heterogeneous character of viral genomes. Based on the sequencing-data, we were able to assemble significant parts of the viral-genomes

present in the samples and in all cases virus species could be determined.

Sequence reads of viral origin were detected using Bowtie2. The sensitivity settings for sequence alignment correspond with values for penalty-scores for mismatches and gaps, optimal for virus detection. Further decrease of penalty-scores will cause increased misalignments and over-interpretation of the results. The sensitivity settings used in our approach were also appropriate for the detection of more variable capsid encoding regions of several RNA viruses. However, the strategy of alignment of sequence reads to a database of all known viral sequences might be a limitation in the detection of unknown, more distantly related viruses.

The direct-sequencing approach delivers more comprehensive information on the microbial content of clinical samples. In almost all samples, reads were found from anelloviruses, which are widespread in the population and most people are infected in the first month after birth [15]. Three genera are associated with human infections: TTV or alphatorquevirus, TTMV or betatorquevirus, and TTMDV or gammatorquevirus. A study in pediatric patients with RTI showed an association with high levels of human anellovirus and fever in children [15]. The observed high prevalence of anellovirus in the present study is in line with a detection rate of 50% in respiratory samples by NGS as described by Prachayangprecha et al. [14]. Although associations between TTV and RTI have been extensively studied, no correlation could be determined [16–18].

Significant numbers of reads could not be assigned to sequences of known organisms. These pools of unassigned reads might contain relevant information about potential pathogens that remained undiscovered until now. It is an intriguing challenge to identify these potential pathogens.

Sample preparation and sequencing combined with bioinformatic-analysis at a basic level can be completed within 48 h. Although the costs of the WGS approach is an obstacle for the implementation of this method for routine diagnosis, costs of sequencing platforms and consumables are expected to decrease significantly in the coming years. Application of WGS and direct-sequencing in clinical microbiology will become affordable for most laboratories. However, a number of other obstacles need to be addressed as well.

Due to the materials used in sample preparation a number of contaminants are introduced. Most enzymes like nucleases, proteases and polymerases are produced in recombinant-protein expression systems. Sequences from protein expression-vectors were detected in all samples used in this study. A well-known source of contaminants is the silica-based spin-filter columns used



in DNA purification kits. The use of these columns might introduce parvovirus-like sequences and provide false positive diagnostic results [19] and should be avoided in NGS based methods. Similar to current molecular diagnostic methods, WGS is sensitive to cross-contamination. Therefore, the same procedures used for prevention of cross contamination in assays like PCR should be applied to WGS. Extraction of relevant information from sequencing data requires proper software tools. In recent years, a number of research groups working in the fields of microbiology or bioinformatics developed programs for the processing of sequence-data [14,20–22]. Processing pipelines have to be capable of processing millions of sequence reads. Pipelines based on NCBI Blast are time-consuming and therefore less suitable for first-pass analyses. In our study we used a rapid procedure based on the short-read sequence assembling program Bowtie2 [11]. Assembly of reads against a dataset of all known viruses was sufficient to extract all significant virological genomic data that match the diagnostic results of the PCR-positive samples. Implementation of this kind of pipelines in Python-scripts were recently published by several groups [20,21]. However, standardization of sample preparation and data processing is necessary for a proper implementation of multiplexed whole-genome sequencing in diagnostics.

The results presented in this study demonstrate that multiplexed WGS is a powerful diagnostic tool in clinical virology. The direct sequencing approach delivers more comprehensive information on the viral content of clinical samples compared to current routine molecular diagnostics. The use of sequence independent detection methods, like WGS, will not only increase the chance to detect the causative agent of viral RTI, but the rapid availability of highly detailed typing information makes it also possible to trace transmission patterns of viruses leading to timely installment of proper infection control measures. In our opinion, direct-sequencing will become a serious alternative for current molecular diagnostic tools.

## Funding

This research was funded by Agentschap NL, LSH, grant FES0908 (VIRGO project), Technology Foundation STW (FES0901 and FES HTSM), Netherlands Genomics Initiative Project No. 050-060-452, and EU (FP7).

## Competing interests

There were no competing interests.

## Ethical approval

Ethical approval by Committee on research involving human subjects of Radboud University Medical Center.

## References

- [1] S. Kouni, P. Karatkitos, A. Chranoti, M. Theodoridou, G. Chrousos, A. Michos, Evaluation of viral co-infections in hospitalized and non-hospitalized children with respiratory infections using microarrays, *Clin. Microbiol. Infect.* 19 (2013) 772–777.
- [2] G.L. Armstrong, L.A. Conn, R.W. Pinner, Trends in infectious disease mortality in the United States during the 20th century, *J. Am. Med. Assoc.* 281 (1999) 61–66.
- [3] M.J. Lodes, D. Suci, J.L. Wilmoth, M. Ross, S. Munro, K. Dix, K. Bernards, A.G. Stöver, M. Quintana, N. Lihoshi, W.J. Lyon, D.L. Danley, A. McShea, Identification of upper respiratory tract pathogens using electrochemical detection of an oligonucleotide microarray, *PLoS One* 2 (2007) e924.
- [4] M. Woolhouse, F. Scott, Z. Hudson, R. Howey, M. Chase-Topping, Human viruses: discovery and emergence, *Philos. Trans. R. Soc. B* 367 (2012) 2864–3871.
- [5] J.M. Norman, S.A. Handley, H.B. Virgin, Kingdom-agnostic metagenomics and the importance of complete characterization of enteric microbial communities, *Gastroenterology* 146 (2014) 1459–1469.
- [6] A.P. Underwood, T. Dallman, N.R. Thomson, M. Williams, K. Harker, N. Perry, B. Adak, G. Willshaw, T. Cheasty, J. Green, G. Douglas, J. Parkhill, J. Wain, Public health value of next-generation DNA sequencing of enterohemorrhagic *Escherichia coli* isolates from an outbreak, *J. Clin. Microbiol.* 51 (2013) 232–237.
- [7] G.J.N. Grard Fair, D. Lee, E. Slikas, I. Steffen, J. Muyembe, T. Sittler, N. Veeraraghavan, J.G. Ruby, C. Wang, M. Makuwa, P. Mulembakani, R.B. Tesh, J. Mazet, A.W. Rimoin, T. Taylor, B.S. Schneider, G. Simmons, E. Delwart, N.D. Wolfe, C.Y. Chiu, E.M. Leroy, A novel rhabdovirus associated with acute hemorrhagic fever in Central Africa, *PLoS Pathog.* 8 (2012) e1002924.
- [8] M. De Vries, B.B. Oude Munnink, B.B.M. Deijns, M.M.M. Canuti, S.M.S.M. Koekkoek, R.R. Molenkamp, M. Bakker, S. Jurriaans, B.D.C. van Schaik, A.C. Luyf, S.D. Olabarriaga, A.H.C. van Kampen, L. van der Hoek, Performance of VIDISCA-454 in feces-suspensions and serum, *Viruses* 4 (2012) 1328–1334.
- [9] Q. Wang, P. Jia, Zhao Z. VirusFinder, Software for efficient and accurate detection of viruses and their integration sites in host genomes through next generation sequencing data, *PLoS One* 8 (2012) e64465.
- [10] J. Zoll, S. Erkens-Hulshof, K. Lanke, F. Verduyn-Lunel, W.J.G. Melchers, E. Schoondermark-van der Ven, M. Roivainen, J.M.D. Galama, F.J.M. van Kuppeveld, Saffold virus, a human Theiler's-Like coronavirus, is ubiquitous and causes infections early in life, *PLoS Pathog.* 5 (2009) e1000416.
- [11] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, *Nat. Methods* 9 (2012) 357–359.
- [12] D. Blankenberg, G. Von Kuster, N. Coraor, G. Ananda, R. Lazarus, Mangan, M.A. Nekrutenko, J. Taylor, Galaxy: a web-based genome analysis tool for experimentalists. *Current Protocols in Molecular Biology*, *Curr. Protoc. Mol. Biol.* (January) (2010) 1–21, Chapter 19, Unit 19.
- [13] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug 15;25(16):2078-9.
- [14] S. Prachayangprecha, C.M.E. Schapendonk, M.P. Koopmans, A.D.M.E. Osterhaus, A.C. Schürch, S.D. Pas, A.A. van der Eijk, Y. Poovorawan, B.L. Haagmans, S.L. Smits, Exploring the potential of next-generation sequencing in detection of respiratory viruses, *J. Clin. Microbiol.* 52 (2014) 3722–3730.
- [15] E. McElvania-Tekippe, K.M. Wylie, E. Deych, E. Sodergren, G. Weinstock, G.A. Storch, Increased prevalence of anellovirus in pediatric patients with fever, *PLoS One* 7 (2012) e50937.
- [16] J.-Y. Chung, T.H. Han, J.W. Koo, S.W. Kim, J.K. Seo, E.S. Hwang, Small Anellovirus infections in Korean children, *Emerg. Infect. Dis.* 13 (2007) 791–793.
- [17] M. Giménez-Barcons, X. Forn, S. Ampurdanes, M. Guilera, M. Soler, C. Soguero, A. Sánchez-Fueyo, A. Mas, J. Bruix, J.M. Sánchez-Tapias, J. Rodés, J.C. Saiz, Infection with a novel human DNA virus (TTV) has no pathogenic significance in patients with liver diseases, *J. Hepatol.* 30 (1999) 1028–1034.
- [18] F. Maggi, M. Pifferi, C. Fornai, E. Andreoli, E. Tempestini, M. Vatteroni, S. Presciutti, S. Marchi, A. Pietrobello, A. Boner, M. Pistello, M. Bendinelli, TT virus in the nasal secretions of children with acute respiratory diseases: relations to viremia and disease severity, *J. Virol.* 77 (2003) 2418–2425.
- [19] S.N. Naccache, A.L. Greninger, D. Lee, L.L. Coffey, T. Phan, A. Rein-Weston, A. Aronsohn, J. Hackett-jr, E.L. Delwart, C.Y. Chiu, The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns, *J. Virol.* 87 (2013) 11966–11977.
- [20] M. Cotton, B. Oude Munnink, M. Canuti, M. Deijns, S.J. Watson, P. Kellam, L. van der Hoek, Full genome virus detection in fecal samples using sensitive nucleic acid preparation, deep sequencing, and a novel iterative sequence classification algorithm, *PLoS One* 9 (2014) e93269.
- [21] T.J. Petty, S. Cordey, I. Padioleau, M. Docquier, L. Turin, O. Preynat-Seauve, E.M. Zdobnov, L. Kaiser, Comprehensive human virus screening using high-throughput sequencing with a user-friendly representation of bioinformatics analysis: a pilot study, *J. Clin. Microbiol.* 52 (2014) 3351–3361.
- [22] G. Zhao, S. Krishnamurthy, Z. Cai, V.L. Popov, A.P. Travassos da Rosa, H. Guzman, S. Cao, H.W. Virgin, R.B. Tesh, D. Wang, Identification of novel viruses using VirusHunter—an automated data analysis pipeline, *PLoS One* 8 (2013) e78470.