Journal of Clinical Virology 66 (2015) 6–11

Contents lists available at ScienceDirect

Journal of Clinical Virology

journal homepage: www.elsevier.com/locate/jcv

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

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

[a] Department of Medical Microbiology, Radboudumc, Nijmegen, The Netherlands
[b] Laboratory of Pediatric Infectious Diseases, Department of Pediatrics, Radboudumc, Nijmegen, The Netherlands
[c] 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.

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 (J. Zoll).
[1] Department of Medical Microbiology, Radboudumc, PO Box 9101,Internal Post: 7770,6500HB Nijmegen, The Netherlands.

http://dx.doi.org/10.1016/j.jcv.2015.02.010
1386-6532/

Table 1
Overview of the clinical samples and patient characteristics used in this study.

Sample material[a] number	Age(days)	Gender	Gest. age[b]	Symptoms	Severity[c] O_2	Diagnostic results[d]	PCR Ct value	DNA/RNA processing[f]
1 Sputum	10	F	25+5	Apnea	O_2	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_2	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_2	RSV[e]		2
7 NPA	406	F	39+3	Moderate dyspnea	O_2	hRV	>40	2
8 NPA	158	F	28+0		O_2	Negative		2
9 NPA	Adult	M				Negative		1
10 NPA	Adult	M				Negative		2

[a] NPA: nasopharyngeal aspirate.
[b] Gestational age in weeks + days.
[c] Severity: O_2 oxygen need; MV mechanical ventilation.
[d] Diagnostic results, hRV: human rhinovirus; hMPV: human metapneumovirus; RSV: respiratory syncytial virus.
[e] Sample 6 was found positive in the RSV rapid test.
[f] 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), influenzavirus-A and B, adenovirus, *Mycoplasma pneumoniae*, *Chlamydophila 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 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).