HexaPrime: A novel method for detection of coronaviruses

Krzysztof Pyrc [a,b,*], Karol Stożek [a], Wojciech Galan [a], Jan Potempa [a,c]

[a] Microbiology Department, Faculty of Biochemistry Biophysics and Biotechnology, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland
[b] Malopolska Centre of Biotechnology, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland
[c] Oral Health and Systemic Disease Research Group, School of Dentistry, University of Louisville, Louisville, KY, USA

A B S T R A C T

Despite intense efforts to develop novel and better tools to identify known viruses and to discover new viruses, establishing etiological roles for viruses in human disease is challenging. In large part, this may be attributed to the high variability of viral species and the difficulties in developing broad-spectrum, yet specific, diagnostic assays. To overcome this problem, a novel method for the detection of viruses is described in the current manuscript. The technique relies on the addition of synthetic oligonucleotides to both termini of RNA fragments in a sequence-dependent manner during first- and second-strand DNA synthesis; these oligonucleotides are used subsequently for amplification of the viral nucleic acids of interest. The recognition of the target sequence by the oligonucleotides is mediated by short (6–8 nt) conserved regions, which facilitates development of broad–spectrum assays. The method has been tested for coronaviruses, although it may be also adopted for other RNA viruses.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Viral infections impose a great disease burden on the human population. As well as known viral diseases, viral infection is thought to play a role in the pathogenesis of many other diseases. However, because the detection and identification of viruses is difficult, establishing etiological roles for viruses in these diseases is challenging. This is particularly true for chronic conditions (e.g., Kawasaki disease) where the virus load is very low or the presence of the virus at the time of sampling is uncertain (Baker et al., 2006; Christen et al., 2012; Simoes et al., 2012). Surprisingly, it may also be true for some acute infections (e.g., respiratory tract infections), where no etiological agent can be identified in a large proportion of cases (Ali et al., 2011; Bezerra et al., 2011; Hayden et al., 2012; Mahony et al., 2011; Regamey et al., 2008; Sanghavi et al., 2012; Thomazelli et al., 2007). Although the failure to identify an infecting virus may be due to the presence of a novel pathogen, we have shown previously that it can also be due to the imperfect diagnostic system (Pyrc et al., 2012).

Coronaviruses are large, enveloped, single-stranded RNA viruses. The genomic nucleic acids (27–32 kb) are shielded by the nucleocapsid protein and interact with membrane, envelope and spike proteins to form a viable, enveloped viral particle (Fields et al.,

2007). At present, three major groups of coronaviruses are recognized (alpha-, beta- and gamma-coronaviruses), which are clusters of species sharing a similar genome structure and other general characteristics (International Committee on Taxonomy of Viruses. and King, 2012). Generally, alpha- and beta-coronaviruses are considered to be mammalian pathogens, while gamma-coronaviruses mostly infect birds (Fields et al., 2007). Until 2003, only two human coronaviruses were recognized and the research on this group of pathogens was relatively limited (Bradburne et al., 1967; Kapikian et al., 1969; McIntosh et al., 1967b, 1974; Reed, 1984; Tyrrell and Bynoe, 1965). The emergence of the severe acute respiratory syndrome coronavirus (SARS-CoV) in the 2002/2003 winter season highlighted the importance of the *Coronaviridae* family for human health (Drosten et al., 2003; Ksiazek et al., 2003; Peiris et al., 2003). Realization that there was a threat of animal–human transmission of coronaviruses accelerated research into these viruses, and this resulted in the identification of two novel human coronaviruses, HCoV-NL63 and HCoV-HKU1, which appeared to have resided in the human population for many years, undetected by diagnostic methods used at the time (van der Hoek et al., 2004; Woo et al., 2005).

There are several techniques described for the detection of coronaviruses. The most generic and oldest is electron microscopy, which allows visualization of viral particles in an infectious material (Hamre and Procknow, 1966; McIntosh et al., 1967a). Since the identification of the first coronaviruses with electron microscopy, several other methods have been developed, ranging from cell culture to immune detection and molecular detection (Pyrc et al., 2011, 2012; Schildgen et al., 2006). Recently, more sophisticated methods

* Corresponding author at: Microbiology Department, Faculty of Biochemistry Biophysics and Biotechnology, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland. Tel.: +48 12 664 61 21; fax: +48 12 664 69 02.
E-mail address: k.a.pyrc@uj.edu.pl (K. Pyrc).

were developed, including sequence-independent techniques that allow amplification of all nucleic acids, regardless of the sequence. Such approaches result in the amplification of targets originating from the host's genetic material, as well as targets of viral origin. Comparison of non-infected control samples with virus-infected samples allowed researchers to select nucleic acid sequence motifs present exclusively in the latter set of samples (Allander et al., 2001, 2005; Anway et al., 2001; Lu et al., 2004; van der Hoek et al., 2004). An interesting (and commonly used) diagnostic approach employs degenerated universal primers able to detect all coronaviruses that share sequence homology at the selected, conserved site; theoretically, this also includes members of the family unrecognized previously (Adachi et al., 2004; Drosten et al., 2003; Escutenaire et al., 2007; Ksiazek et al., 2003; Moes et al., 2005; Stephensen et al., 1999). This approach is promising, although high variability of viral sequences coupled with limitations to the level of degeneration of primers may lead to false-negative results if viral species are not identical to reference strains. To overcome this problem, the CODEHOP technique was developed, allowing the design of highly degenerated DNA primers based on more conserved protein sequences (Rose et al., 1998; Zlateva et al., 2011).

In the current manuscript, we describe the development of a novel method for the detection of RNA viruses, combining to some extent the selectivity of the universal primer approach with the broad specificity of sequence-independent methods. The method relies on the design of synthetic oligonucleotides that comprise short (6–8 nt) elements that recognize conserved regions of viral nucleic acids and longer (16–18 nt) anchor elements. These oligonucleotides are then added to termini of RNA fragments during first- and second-strand DNA synthesis. The anchoring regions are subsequently used for amplification and analysis of viral nucleic acids. The use of short (6–8 nt) conserved regions to recognize target sequences facilitates the design and development of broad-spectrum assays, and therefore increases the applicability of the method. Although the sensitivity of the method is limited, and thus most suitable for the detection of cultivable pathogens, its potential for recognizing unknown members of the *Coronaviridae* family means that this is a valuable addition to the range of virus detection and identification tools that already exist.

2. Materials and methods

2.1. Viruses

Virus stocks were generated by infecting HeLa cells (ATCC: CCL-2) with adenovirus; RD cells (ATCC: CCL-136) with enterovirus; LLC-MK2 cells (ATCC: CCL-7) with parainfluenza types 1–3, HCoV-NL63, human metapneumovirus (hMPV) and human respiratory syncytial virus (hRSV); and MDCK cells (ATCC: CCL-2936) with influenza A and B. In all cases, cells were lysed by two freeze-thaw cycles after the appearance of the cytopathic effect. The virus-containing fluid was aliquoted and stored at −80 °C. A control cell lysate from mock-infected cells was prepared in the same manner as the virus stock.

Parainfluenza 1 and 2 viruses and influenza B virus were isolated from clinical samples processed in our laboratory. Stock samples containing influenza A virus, enterovirus (human echovirus 9), hRSV, human parainfluenza 3 virus and human adenovirus were provided kindly by Marcel Muller (University of Bonn Medical Center, Germany); HCoV-NL63 was a kind gift from Lia van der Hoek (Academic Medical Center, University of Amsterdam, The Netherlands). hMPV and human coronavirus HKU1 were provided kindly by Oliver Schildgen (University Hospital Witten/Herdecke, Cologne, Germany) and Astrid Vabret (University Hospital of Caen, France), respectively.

2.2. Selection of HexaPrime primers

To identify conserved sites in coronavirus sequences, three approaches were used: (1) dedicated HexaPrime software (described in Section 3.2: HexaPrime software), (2) Bioedit (ver. 7.1.3.0; http://www.mbio.ncsu.edu/bioedit/bioedit.html) automatic conserved site finder, and (3) visual inspection. Application of all these approaches allowed appropriate screening of coronavirus sequence data. GenBank numbers of sequences included in the analysis are as follows: **HCoV-NL63** (DQ445911; DQ445912; AY567487; NC_005831; JQ765566; JQ765564; JQ765568; JQ765570; JQ765572; JQ765574; JQ765565; JQ765563; JQ765567; JQ765571; JQ765569; JQ765575; JQ765573), **HCoV-OC43** (JN129835; JN129834; AY391777; AY585229; AY585228; AY903459; AY903460; DQ011855; NC_007732; NC_005147; FJ415324), **HCoV-HKU1** (HM034837; AY597011; DQ415914; DQ415912; DQ415910; DQ415908; DQ415906; DQ415904; DQ415902; DQ415900; DQ415898; DQ415896; DQ415913; DQ415911; DQ415909; DQ415907; DQ415905; DQ415903; DQ415901; DQ415899), **HCoV-229E** (NC_002645; AF304460), **SARS-CoV** (NC_004718, JX162087, JQ316196, AY559089, AY559095, AY613950, AY545914, AP006558, AP006561, EU371561, EU371564, AY278554, AY338175, AY357076, FJ882945, AY279354, GU553363, AY313906, FJ882930, FJ882936, HQ890529, JF292908, HQ890538, JF292915), **BCoV** (AF391541; AF391542; NC_003045; EF424618; EF424616), **MHV** (JQ173883; NC_006852; GU593319; AY700211; AF208066; AF208067; AB551247; AF201929; FJ884687; FJ884686; NC_001846; FJ647227; FJ647225; FJ647223; FJ647221; FJ647219; FJ647226; FJ647224; FJ647222; FJ647218; FJ647220), **porcine epidemic diarrhea virus** (JQ023162; JQ023161; JN825712; JQ282909; JN547228; AF353511; DQ811787; NC_003436; EF185992; Z25483; Z24733), **canine coronavirus** (JQ404410, JQ404409, GQ477367), **equine coronavirus** (NC_010327; EF446615), **porcine hemagglutinating encephalomyelitis virus** (NC_007732; DQ011855), **feline coronavirus** (JQ408980; GQ152141; JN183882; JN183883; FJ938058; FJ938060; HQ392471; HQ392469; HQ392472; HQ392470; HQ012371; HQ012369; HQ012367; HQ012372; HQ012370; HQ012368; GU553361; GU553362; FJ938062; FJ938059; FJ938054; FJ938052; FJ938061; FJ938057; FJ938055; FJ938053; DQ848678; DQ010921; EU186072; DQ286389; JQ408980; AY994055; NC_002306; JQ408981), **transmissible gastroenteritis virus** (DQ811786; DQ811789; FJ755618; HQ462571; EU074218; DQ201447; DQ811788; DQ811785; DQ443743; AJ271965).

2.3. Nucleic acid extraction

RNA from clinical specimens and viral culture supernatants was extracted using a Total RNA mini kit (A&A Biotechnology, Gdynia, Poland), according to the manufacturer's protocol. DNA was isolated using a Viral DNA/RNA kit (A&A Biotechnology, Gdynia, Poland), according to the manufacturer's instructions. Following extraction, samples were incubated for 30 min at 37 °C with DNase (DNase Turbo, Thermo Scientific, Vilnius, Lithuania) and the RNA was re-isolated. Following isolation, total RNA was precipitated in the presence of 30 μg of glycogen (Life Technologies, Warsaw, Poland) with three volumes of isopropanol (16 h, −20 °C). Samples were centrifuged subsequently ($12,000 \times g$, 45 min) and the resulting pellets were washed with 70% ethanol, dried (2 min at room temperature) and re-suspended in 5 μl of nuclease-free water (Sigma–Aldrich, Poznan, Poland). Purified RNA was stored at −80 °C.

Table 1
Primers used for quantitative PCR.

Target species	Primer	Primer sequence (5′-3′)
HCoV-NL63	Sense primer [63NF2]	AAA CCT CGT TGG AAG CGT GT
	Antisense primer [63NR1]	CTG TGG AAA ACC TTT GGC ATC
	Probe [63NP]	FAM-ATG TTA TTC AGT GCT TTG TCC CTC GTG AT-TAMRA
HCoV-HKU1	Sense primer [HKUqPCR5]	CTG GTA CGA TTT TGC CTC AA
	Antisense primer [HKUqPCR3]	CAA TCA CGT GGA CCC AAT AAT
	Probe [HKUqPCRP]	FAM-TTG AAG GCT CAG GAA GGT CTG CTT CTA A-TAMRA

2.4. Virus detection by reverse transcription and quantitative PCR

Isolated nucleic acids (see Section 2.3) were reverse transcribed with High Capacity cDNA Reverse Transcription Kit (Life Technologies, Warsaw, Poland), according to the manufacturer's instructions. HCoV-NL63 and HKU1 virus yields were determined using real-time PCR with standard curves, as previously described (Golda et al., 2011; Pyrc et al., 2010). Briefly, five microliters of cDNA was amplified in 10 μl reaction mixture, containing 1× TaqMan® Universal PCR Master Mix, No AmpErase®UNG (Life Technologies, Warsaw, Poland), specific probe labeled with FAM (6-carboxyfluorescein), and TAMRA (6-carboxytetramethylrhodamine) (200 nM) and primers (900 nM each). All primers and probes are listed in Table 1. Rox was used as a reference dye. The reaction was monitored on a 7500 fast real-time PCR machine (Life Technologies, Warsaw, Poland) with the following settings: 2 min 50 °C, 10 min at 92 °C, and 40 cycles of 15 s at 92 °C, and 1 min at 60 °C.

2.5. HexaPrime reverse transcription and second-strand synthesis

Total RNA that was isolated and purified as described in Section 2.3 was used for the synthesis of the first and second cDNA strands. Five microliters of total RNA were mixed with 1.5 pM (1.5 μl) of the reverse transcription (RT) primer (a list of RT primers used in the current study is given in Table 2), incubated at 65 °C for 5 min, cooled on ice for 2 min, and mixed with the RT MIX (25 U of MultiScribe Reverse Transcriptase (Life Technologies, Warsaw, Poland), 1 μl of 10× DNA Polymerase I buffer (Thermo Scientific, Vilnius, Lithuania), 0.4 μl of 100 mM dNTPs, 0.2 μl of DMSO in a total volume of 3.5 μl). Reactions were carried out for 120 min at 37 °C. Following the incubation, samples were heat-inactivated at 85 °C for 5 min.

The resulting single-stranded cDNA was used for second-strand synthesis, with no purification step in-between. Briefly, the sample (10 μl) was denatured for 1 min at 95 °C and cooled on ice. Subsequently, 5 μl of the second-strand mix was added (0.5 U of RNase H, 0.5 μl of 10× DNA Polymerase I buffer, 4.5 U of DNA Polymerase I (Thermo Scientific, Vilnius, Lithuania), 0.1 μl of DMSO, and 3 pM of second-strand (SS) primer). A list of SS primers used in the current study is given in Table 2. Samples were incubated for 120 min at 15 °C. Subsequently, double-stranded cDNA was purified by means of phenol:chloroform:isoamyl alcohol (pH 8.0) isolation. DNA in resulting samples was precipitated with 300 μl of isopropanol (16 h; −20 °C), centrifuged (12,000 × g, 45 min), washed with 70% ethanol, dried at room temperature for 2 min, and re-suspended in 5 μl of nuclease-free water (Sigma–Aldrich, Poznan, Poland).

2.6. HexaPrime PCR amplification

The resulting double-stranded DNA was used directly for amplification according to the protocol given below. Amplification was carried out in a total volume of 20 μl with 1× DreamTaq PCR Master Mix (Thermo Scientific, Vilnius, Lithuania) in the presence of 1 pM each of forward and reverse primer (RT and SS primers) and template DNA (5 μl). The first PCR cycling conditions included initial denaturation for 3 min at 95 °C, followed by 40 cycles of 10 s at

95 °C, 30 s at 56 °C, 15 s at 72 °C, and then 5 min at 72 °C for the final elongation.

For the second, nested PCR, the reaction was prepared in a similar manner, with 5 μl of the first PCR mixture as a template and 12 pM of each primer (a list of the PCR2 primers used in the current study is given in Table 2). The cycling conditions included initial denaturation for 3 min at 95 °C; followed by 13 cycles of 10 s at 95 °C, 30 s at 68 °C (which was decreased by 1 °C per cycle), and 11 s at 72 °C; a further 30 cycles of 10 s at 95 °C, 30 s at 55 °C, and 11 s at 72 °C; and then 5 min at 72 °C for the final elongation. Following the reaction, 20 μl of the sample was loaded onto a 1.5% agarose gel and visualized with ethidium bromide.

2.7. Evaluation of HexaPrime assay sensitivity

To determine the sensitivity of the assay, serial dilutions of the virus stocks described in Section 2.1 were used in concentrations ranging from 10^9 to 10^5 copies per reaction (as determined using real-time PCR). RT, SS synthesis, PCR and nested PCR were conducted as described above (Sections 2.5 and 2.6). The efficiency of different SS synthesis enzymes, T7 Polymerase (Thermo Scientific, Vilnius, Lithuania), DNA Polymerase I (Thermo Scientific, Vilnius, Lithuania), and Sequenase 2.0 (Affymetrix, United Kingdom), was evaluated by means of densitometry following bands separation on a 1.5% agarose gel.

To test whether it is possible to detect coronaviral RNA not only in cell culture but also in clinical samples, 100 μl aliquots of clinical specimens, including sputum, bronchoalveolar lavage fluid, and nose wash, which had tested negatively for all known pathogens (Pyrc et al., 2012), were spiked with 1 μl of HCoV-NL63 virus stock (final TCID$_{50}$ of 400). Subsequently, the HexaPrime assay was carried out and the resulting products were analyzed on a 1.5% agarose gel.

2.8. Evaluation of HexaPrime assay specificity

To determine whether the coronavirus primers that were chosen in the current study specifically amplified coronavirus nucleic acids, the HexaPrime assay with these primers was also conducted on several other human viruses, including HCoV-HKU1, influenza A and B viruses, parainfluenza 1–3 viruses, hRSV, human adenovirus, human rhinovirus, hMPV and human echovirus 9. Briefly, viral nucleic acids were isolated and concentrated as described above (Section 2.3). Samples were subjected to the HexaPrime assay using primer set 2 (Table 2), as described above (Sections 2.5 and 2.6).

3. Results

3.1. Design of HexaPrime primers

The logic for HexaPrime primer design is similar to the universal primer PCR design and therefore the presence of conserved sites in the nucleic acids of the viral group of interest is a prerequisite. On the other hand, it also shares some characteristics with sequence-independent methods, as anchor oligonucleotides are added to flank the region of interest and serve as templates

Table 2
Evaluated primer sets. SS: second-strand synthesis; RT: reverse transcription; 5PCR2 and 3PCR2: 5′ and 3′ primers for the second amplification step (see Fig. 1 for further details).

Primer set	Position[a]	Primer	Sequence	Amplicon size[b] (bp)
1	16,427–16,550	1_SS	AGCAAGATCCAATCTAGA**STGATG**	155
		1_5PCR2	CAAGATCCAATCTAGA**STGATGA**	
		1_RT	CCAAGGGATTCCCAACCT**YCCAAC**	
		1_3PCR2	AAGGGATTCCCAACCT**YCCAACA**	
2	16,413–16,550	2_SS	GCAAGATCCAATCTAGA**ATGATSA**	169
		2_5PCR2	AAGATCCAATCTAGA**ATGATSATT**	
		2_RT	CCAAGGGATTCCCAACCT**YCCAAC**	
		2_3PCR2	AAGGGATTCCCAACCT**YCCAACA**	
3	16,545–16,681	3_SS	AGCAAGATCCAATCTAGA**RTGTTG**	168
		3_5PCR2	CAAGATCCAATCTAGA**RTGTTGG**	
		3_RT	AGGGATTCCCAACCT**TCATCNACA**	
		3_3PCR2	GGATTCCCAACCT**TCATCNACAAA**	
4	19,497–19,579	4_SS	AGCAAGATCCAATCTAGA**CARTGG**	114
		4_5PCR2	CAAGATCCAATCTAGA**CARTGGGG**	
		4_RT	CCAAGGGATTCCCAACCT**YNGCAT**	
		4_3PCR2	AAGGGATTCCCAACCT**YNGCATC**	
5	1,949,719,588	5_SS	AGCAAGATCCAATCTAGA**CARTGG**	123
		5_5PCR2	CAAGATCCAATCTAGA**CARTGGGG**	
		5_RT	CCAAGGGATTCCCAACCT**CKWGTCA**	
		5_3PCR2	AAGGGATTCCCAACCT**CKWGTCAT**	
6	19,959–20,032	6_SS	AGCAAGATCCAATCTAGA**TTTGAYA**	105
		6_5PCR2	CAAGATCCAATCTAGA**TTTGAYAC**	
		6_RT	CCAAGGGATTCCCAACCT**AANGCA**	
		6_3PCR2	AAGGGATTCCCAACCT**AANGCATG**	
7	21,507–21,568	7_SS	AGCAAGATCCAATCTAGA**NNATGC**	93
		7_5PCR2	CAAGATCCAATCTAGA**NNATGCC**	
		7_RT	CCAAGGGATTCCCAACCT**CCATART**	
		7_3PCR2	AAGGGATTCCCAACCT**CCATARTT**	
8	21,564–21,631	8_SS	AGCAAGATCCAATCTAGA**AAYTAT**	99
		8_5PCR2	CAAGATCCAATCTAGA**AAYTATGG**	
		8_RT	CCAAGGGATTCCCAACCT**TAYTGA**	
		8_3PCR2	AAGGGATTCCCAACCT**TAYTGACA**	

[a] Position in alignment.
[b] The expected amplicon size for HCoV-NL63.

for subsequent amplification. First, isolated RNA is subjected to RT with primers designed using HexaPrime (or other) software. The 3′ primers (RT primers) are used for both reverse transcription and for the first PCR amplification. They consist of short (6–8 nt) elements homologous to a conserved region present in all targeted genomes and longer (16–18 nt) artificial anchor sequences that are used as a template for PCR amplification (Fig. 1). Following RT, the RNA–DNA duplex is disrupted with RNase H and a mixture of DNA Polymerase I and SS primers (used for both second-strand synthesis and the first PCR) is added to the sample. The primers for SS synthesis are constructed in the same manner as the RT primers. Therefore, the SS synthesis results in the addition of another anchor (16–18 nt) at the 5′ end of the amplified fragment (Fig. 1). The resulting double-stranded DNA, flanked with synthetic oligonucleotides anchored on both termini, is used for the first PCR amplification with RT and SS primers. The product of this amplification is used subsequently for the nested amplification with PCR2 primers that amplify a target within the template DNA. This amplification results in the generation of short amplicons, which can be visualized on a standard agarose gel and sequenced. The general overview of the HexaPrime assay is presented in Fig. 1.

To design appropriate primers for the detection of coronaviruses, a set of coronaviral sequences was obtained from the GenBank database (accession numbers are provided in Section 2). Sequences were aligned and inspected for the presence of short, conserved regions. Inspection was performed with three independent approaches:

(1) HexaPrime software. This Java-based software for the identification of conserved sequences was developed for the current study. The software is available publicly at http://www.virogenetics.info (for more details see Section 3.2: HexaPrime software).
(2) Bioedit. Conserved sequences were identified based on the automatic conserved site finder feature of this software.
(3) All sequences were inspected by eye and conserved sequences were marked.

This analysis led to the identification of several conserved regions (6–8 nt) and allowed the design of multiple primer sets, which are listed in Table 2. All fragments were selected based on sequence identity alone without any pre-selection based on the location or predicted RNA structure.

3.2. HexaPrime software

The HexaPrime software for the identification of conserved sites in multiple viral genomes was prepared in the Java environment. Before the HexaPrime analysis, the collection of all sequences of interest in the FASTA format is prepared. During analysis each sequence is divided into overlapping words of desired length. Subsequently, for each sequence these words are converted into a bit array. Arrays corresponding to each sequence are then multiplied logically and the resulting array contains only words present in all tested sequences. This approach appeared to be more effective than simple comparison of sequences and it provides complete information on the presence of a certain nucleotide stretch in the whole genome sequence, which may be of importance during *in silico* screening of potential primer pairs. The resulting list of words does not contain information on the position of these words in
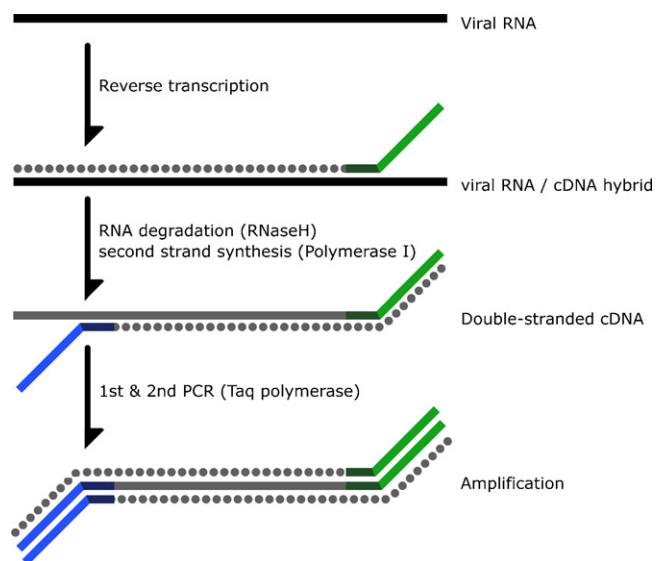
Fig. 1. The HexaPrime assay. Briefly, reverse transcription (RT) and second-strand (SS) synthesis reactions were conducted on total RNA (black) in the presence of primers comprising short (6–8 nt) elements that recognize conserved regions of viral nucleic acids (dark green and dark blue for RT and SS primers, respectively) and longer (16–18 nt) anchors that serve as a template for subsequent PCR amplification (light green and light blue for RT and SS primers, respectively). Products of the PCR amplification of coronaviral RNA may be analyzed further by gel electrophoresis and sequencing.

Fig. 2. The HexaPrime software. The general algorithm for the software is presented. (A) Sequences are divided into six nucleotide-long words. (B) Subsequently, these words are converted into a numerical format and a bit array is formed. Arrays created for each sequence are further logically multiplied and the resulting array is converted back to words. (C) The resulting word list contains information on the genomic position and distance between words in each of the genomes.

each genome; therefore, these are recovered subsequently from sequence data and the distance between each word pair is determined. Results are presented as a list of words with distances between them and their position in the original sequences of each genome. The overview of the method is presented in Fig. 2.

3.3. Evaluation of HexaPrime primer sets

To validate the designed primer sets, all were tested on samples containing HCoV-NL63 RNA. All steps of the procedure were evaluated and optimized carefully to determine the optimal reaction conditions. Briefly, different enzymes and buffers, concentrations of salts and primers, and thermal profiles for amplification were tested. These results are not described in the manuscript due to space constraints, but the final optimal conditions are presented.

RNA was isolated from virus stocks and concentrated by precipitation in the presence of glycogen. The concentration of the RNA samples resulted in a proportional increase in virus yield, as determined by real-time PCR (data not shown). Subsequent analysis showed that this step also enhanced the sensitivity of the HexaPrime assay (data not shown). Isolated and concentrated RNA was processed according to the HexaPrime method as described in Section 2. Briefly, RNA was reverse transcribed using the Multi-Scribe Reverse Transcriptase and the RT primers listed in Table 2. Second-strand synthesis was conducted using DNA Polymerase I and the SS primers listed in Table 2. The first and second PCR amplification was conducted with RT/SS primers and PCR2 primers, respectively. Gel electrophoresis of the resulting products yielded several bands, although only amplification with primer sets 2, 4, 5 and 8 yielded distinct bands in samples containing viral RNA (Fig. 3A). DNA fragments of the appropriate size were then excised from the gel, isolated, and cloned into the pTZ57R plasmid. Clones were sequenced and their identity was confirmed. Based on assay performance (densitometry) and sequencing results, primer set 2 was used for subsequent analyses with alpha- and beta-coronaviruses (Fig. 3B).
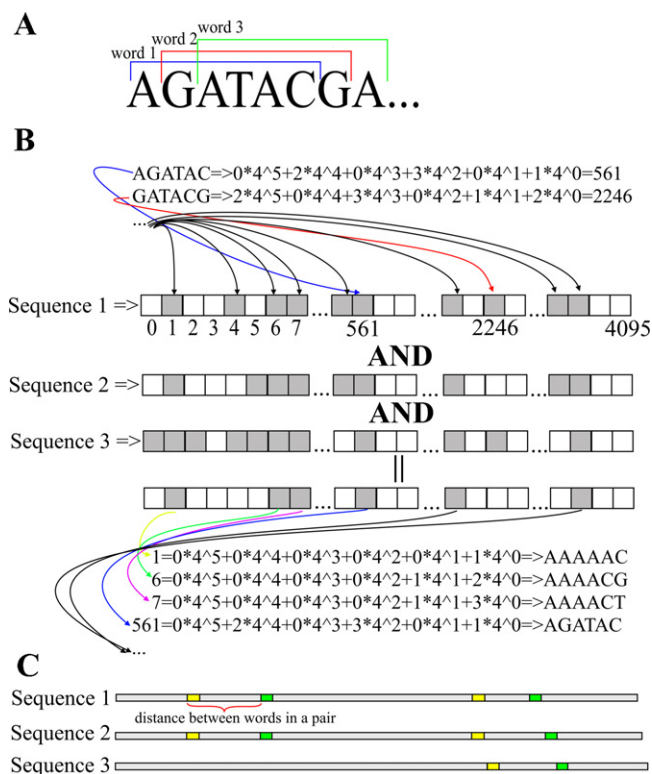
3.4. Sensitivity of HexaPrime assay

To determine the sensitivity of the HexaPrime assay, the concentration of viral RNA in cultured samples was determined using real-time PCR. Following this, the RNA concentration was set to 10^9 copies ml^{-1} and serial, 10-fold dilutions to 10^5 copies ml^{-1} were prepared. These samples were subjected to the HexaPrime assay, and the results presented in Fig. 3C show that the assay is sensitive to 10^6 copies ml^{-1}. To test whether the performance of the vaguest step of the reaction (i.e., SS synthesis) is optimal, an effort was made to evaluate other enzymes able to carry out the SS synthesis. As clearly visible in Fig. 3C, DNA Polymerase I and Polymerase T7 showed similar efficiency, while SS synthesis with Sequenase 2.0 did not result in discrete bands.

3.5. Cross-reactivity of HexaPrime primers

To determine whether the HexaPrime assay is specific toward coronaviruses, primer set 2 for coronaviruses was tested for cross-reactivity with the viral nucleic acids of other respiratory pathogens. Briefly, RNA or DNA of hMPV, adenovirus, rhinovirus, enterovirus, influenza A and B viruses, parainfluenza 1–3 viruses and hRSV was isolated from cell culture and subjected to the HexaPrime assay without dilution. Even though some background amplification was observed for some samples, no cross-reactivity (i.e., no products of similar size or quantity) was detected in any of the assays, showing that the HexaPrime assay is highly specific (Fig. 4). No amplification was observed in negative control samples (mock-infected cell culture) or water.
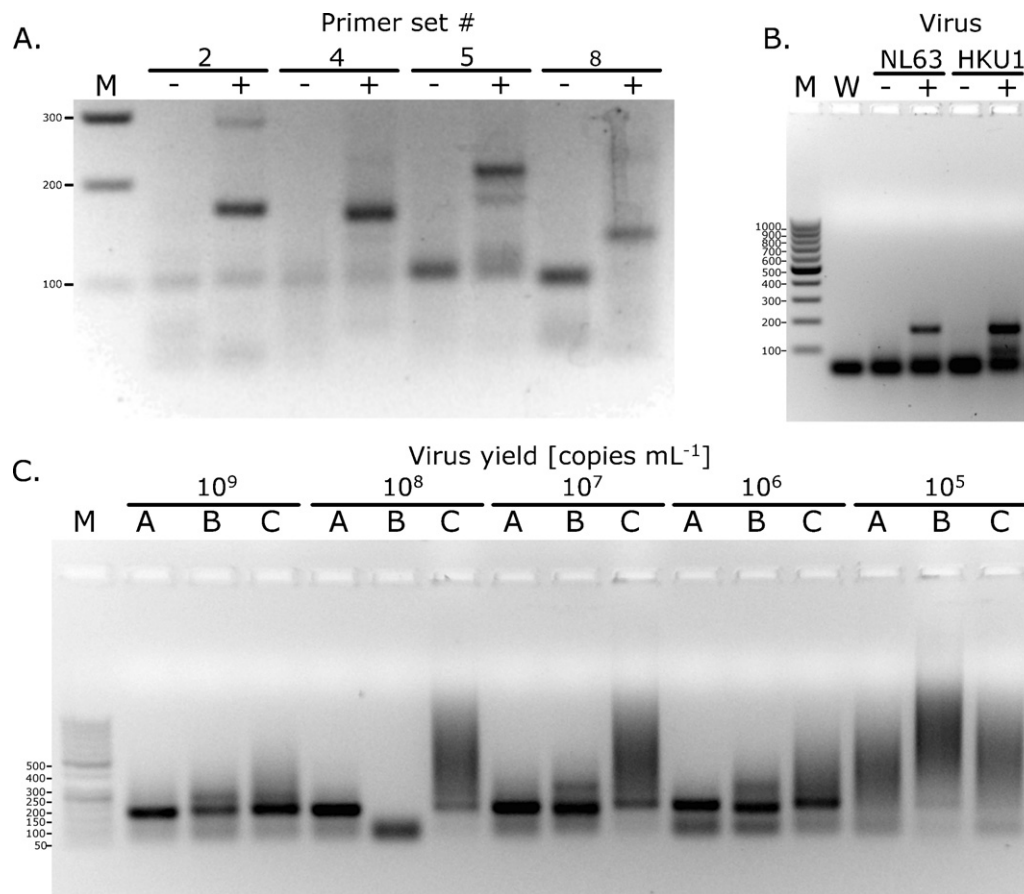
Fig. 3. Evaluation of the HexaPrime assay. (A) Evaluation of different primer pairs for the detection of coronaviruses. Analysis was conducted using the HCoV-NL63 virus and all primer sets given in Table 2 were tested. Only amplification with primer sets 2, 4, 5 and 8 yielded distinct bands. Sequencing of products and analysis of fragment size revealed that only primer set 2 allowed efficient amplification of the desired product. M: size marker; mock-infected (−) or HCoV-NL63-infected (+) cell culture supernatant. (B) Detection of HCoV-NL63 and HCoV-HKU1 with the HexaPrime assay using primer set 2. All experimental procedures were conducted as described in Section 2. M: size marker; W: water; NL63 and HKU1: mock-infected (−) or virus-infected (+) cell culture supernatant. (C) Sensitivity of the HexaPrime assay. Concentrated samples containing viral RNA (10^9 copies ml^{-1}) were subjected to 10-fold serial dilutions in cell culture supernatant and the HexaPrime assay was conducted. For each RNA concentration, three different enzymes for SS DNA synthesis were trialed. A, B and C denote DNA Polymerase I, T7 Polymerase, and Sequenase 2.0, respectively.

3.6. HexaPrime assay performance in clinical samples

The HexaPrime assay is designed to detect viral RNA in complex clinical specimens. Therefore, assay performance was evaluated using different sample types, including nose wash, bronchoalveolar lavage fluid and sputum, spiked with virus. Briefly, 0.5 μl of culture supernatant from LLC-MK2 cells infected with HCoV-NL63 was inoculated into 100 μl of each clinical sample type (all of which had previously tested negative for coronaviruses). Reactions were performed according to the protocol described above and the resulting
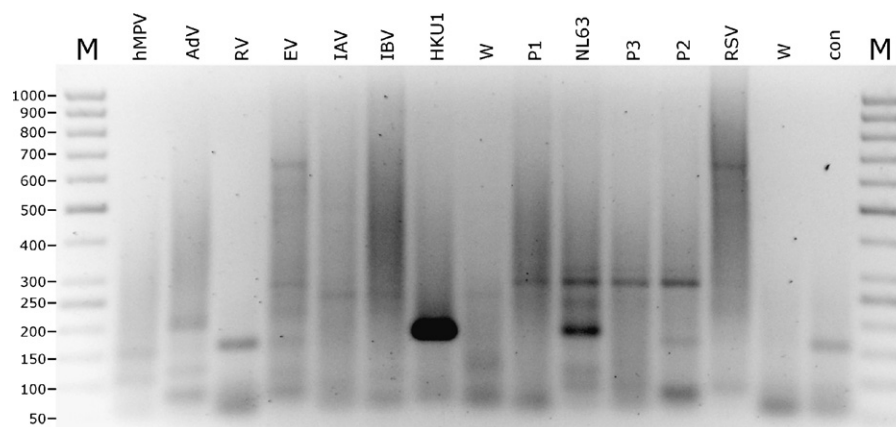

Fig. 4. Specificity of the HexaPrime assay. The HexaPrime assay was conducted on samples containing different human respiratory viruses, including human metapneumovirus (hMPV), adenovirus (AdV), rhinovirus (RV), enterovirus (EV), influenza A and B viruses (IAV and IBV, respectively); parainfluenza 1, 2 and 3 viruses (P1, P2 and P3, respectively); and respiratory syncytial virus (RSV). M: size marker.
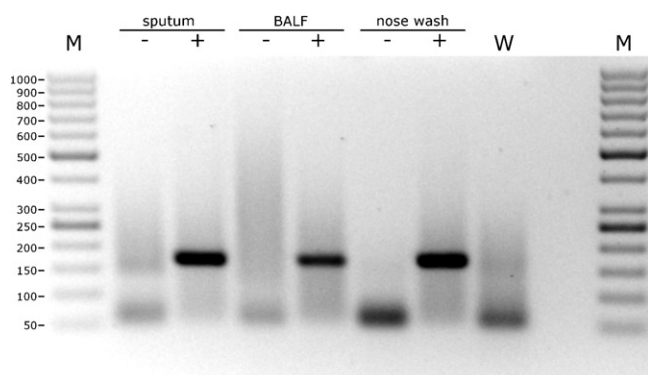
Fig. 5. Detection of HCoV-NL63 in different types of clinical samples. Different sample types (sputum, bronchoalveolar fluid (BALF) and nose wash) were spiked with HCoV-NL63. No inhibition of HexaPrime reactions and no additional bands were observed. W: water; − and + signs: samples not inoculated or inoculated with HCoV-NL63.

products were analyzed by standard gel electrophoresis. No inhibition of the reaction was observed for any of the clinical samples tested (Fig. 5).

4. Discussion

Numerous methods have been developed for the detection of viral pathogens, and the universal primer approach, for all its limitations, is undoubtedly the most sensitive method for semi-generic detection (Moes et al., 2005; Pyrc et al., 2012). On the other hand, it is only appropriate for identifying viruses that are closely related to known viruses. For the detection of novel viruses that are not detected using universal primers, several sequence-independent techniques have been developed. These include random-priming PCR, representational difference analysis (RDA), and virus discovery based on cDNA amplified fragment length polymorphism (VIDISCA) (Adachi et al., 2004; Allander et al., 2001, 2005; Anway et al., 2001; Drosten et al., 2003; Escutenaire et al., 2007; Ksiazek et al., 2003; Lu et al., 2004; Moes et al., 2005; Stephensen et al., 1999; van der Hoek et al., 2004). These methods are named sequence-independent as the amplification is not driven by specific primers, but instead a random (or semi-random) amplification is carried out. For example, a random-priming PCR method utilizes primers containing random nucleotides at their 5′ termini, which are able to attach to and amplify targets, including viral genetic material, indiscriminately (Fouchier et al., 2004). A more specific approach is employed by RDA, where primers selected arbitrarily are used and, following the PCR reaction, infected samples are compared to those of non-infected samples (Liang and Pardee, 1992; Lu et al., 2004). A very different approach is employed in the VIDISCA and SISPA methods. The whole amplification process is based on the presence of common restriction sites in RNA/DNA molecules. Following enzymatic restriction, these restriction sites serve as sites for ligation of synthetic DNA fragments which, in turn, constitute primer attachment sites. In this way, every fragment flanked with certain restriction sites will be amplified. Comparison of the characteristic amplification pattern of infected and control samples may serve to identify nucleic acids of viral origin. Unfortunately, all these approaches are extremely laborious and high-throughput analysis is not possible (Allander et al., 2001; Reyes and Kim, 1991; van der Hoek et al., 2004).

In the current manuscript, a novel assay for the detection of coronaviruses is presented, the HexaPrime assay. Although the sensitivity of the assay with the primer sets tested in the current work is indisputably lower than that described for universal primers, it is similar to that described for sequence-independent methods.

Because a shorter conserved nucleotide stretch is required for successful priming, the major advantage of the HexaPrime assay is that it provides broader specificity when compared to the standard universal primer approach. It is also worth noting that methods for sequence-independent detection of viral pathogens (e.g., VIDISCA or RDA) require highly-trained personnel and appropriate infrastructure, while the HexaPrime assay is relatively simple. The only additional step, compared to a typical nested PCR, is the ligation of anchors to 5′ and 3′ ends of the RNA fragments, which is a single-tube reaction.

To increase the sensitivity of the method, an additional step was added to the protocol that includes precipitation of RNA in the presence of glycogen. This concentration method was confirmed to be effective by quantitative PCR (data not shown). The addition of such a step may therefore also be beneficial for other virus detection methods, such as VIDISCA or SISPA (Allander et al., 2001; van der Hoek et al., 2004).

The analyses performed in the current study show clearly that the HexaPrime assay is highly specific for coronaviral species and it may therefore be useful for routine screening of clinical samples. Due to the limited sensitivity of the method, the in vitro culture of pathogens may be a prerequisite for successful amplification. To this end, fully-differentiated airway epithelium cultures may be a useful pre-amplification tool to increase the titers of airway viruses (Pyrc et al., 2012). Obviously, due to technical limitations, this method may be used mostly in research facilities focused on discovery of new pathogens, rather than in clinical virology laboratory settings. The method has been tested for the detection of viruses that belong to a single viral family thus far, although it may also be adopted for other RNA viruses.

Acknowledgments

References

Adachi, D., Johnson, G., Draker, R., Ayers, M., Mazzulli, T., Talbot, P.J., Tellier, R., 2004. Comprehensive detection and identification of human coronaviruses, including the SARS-associated coronavirus, with a single RT-PCR assay. J. Virol. Methods 122, 29–36.

Ali, S.A., Gern, J.E., Hartert, T.V., Edwards, K.M., Griffin, M.R., Miller, E.K., Gebretsadik, T., Pappas, T., Lee, W.M., Williams, J.V., 2011. Real-world comparison of two molecular methods for detection of respiratory viruses. Virol. J. 8, 332.

Allander, T., Emerson, S.U., Engle, R.E., Purcell, R.H., Bukh, J., 2001. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. Proc. Natl. Acad. Sci. U. S. A. 98, 11609–11614.

Allander, T., Tammi, M.T., Eriksson, M., Bjerkner, A., Tiveljung-Lindell, A., Andersson, B., 2005. Cloning of a human parvovirus by molecular screening of respiratory tract samples. Proc. Natl. Acad. Sci. U. S. A. 102, 12891–12896.

Anway, M.D., Johnston, D.S., Crawford, D., Griswold, M.D., 2001. Identification of a novel retrovirus expressed in rat Sertoli cells and granulosa cells. Biol. Reprod. 65, 1289–1296.

Baker, S.C., Shimizu, C., Shike, H., Garcia, F., van der Hoek, L., Kuijper, T.W., Reed, S.L., Rowley, A.H., Shulman, S.T., Talbot, H.K., Williams, J.V., Burns, J.C., 2006. Human coronavirus-NL63 infection is not associated with acute Kawasaki disease. Adv. Exp. Med. Biol. 581, 523–526.

Bezerra, P.G., Britto, M.C., Correia, J.B., Duarte Mdo, C., Fonceca, A.M., Rose, K., Hopkins, M.J., Cuevas, L.E., McNamara, P.S., 2011. Viral and atypical bacterial detection in acute respiratory infection in children under five years. PloS ONE 6, e18928.

Bradburne, A.F., Bynoe, M.L., Tyrrell, D.A., 1967. Effects of a new human respiratory virus in volunteers. Br. Med. J. 3, 767–769.

Christen, U., Bender, C., von Herrath, M.G., 2012. Infection as a cause of type 1 diabetes? Curr. Opin. Rheumatol. 24, 417–423.

Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H.R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R.A., Berger, A., Burguiere, A.M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J.C., Muller, S., Rickerts, V., Sturmer, M., Vieth, S., Klenk, H.D., Osterhaus, A.D., Schmitz, H., Doerr, H.W., 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. 348, 1967–1976.

Escutenaire, S., Mohamed, N., Isaksson, M., Thoren, P., Klingeborn, B., Belak, S., Berg, M., Blomberg, J., 2007. SYBR Green real-time reverse transcription-polymerase chain reaction assay for the generic detection of coronaviruses. Arch. Virol. 152, 41–58.

Fields, B.N., Knipe, D.M., Howley, P.M., 2007. Fields Virology, 5th ed. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia.

Fouchier, R.A., Hartwig, N.G., Bestebroer, T.M., Niemeyer, B., de Jong, J.C., Simon, J.H., Osterhaus, A.D., 2004. A previously undescribed coronavirus associated with respiratory disease in humans. Proc. Natl. Acad. Sci. U. S. A. 101, 6212–6216.

Golda, A., Malek, N., Dudek, B., Zeglen, S., Wojarski, J., Ochman, M., Kucewicz, E., Zembala, M., Potempa, J., Pyrc, K., 2011. Infection with human coronavirus NL63 enhances streptococcal adherence to epithelial cells. J. Gen. Virol. 92, 1358–1368.

Hamre, D., Procknow, J.J., 1966. A new virus isolated from the human respiratory tract. Proc. Soc. Exp. Biol. Med. 121, 190–193.

Hayden, R.T., Gu, Z., Rodriguez, A., Tanioka, L., Ying, C., Morgenstern, M., Bankowski, M.J., 2012. Comparison of two broadly multiplexed PCR systems for viral detection in clinical respiratory tract specimens from immunocompromised children. J. Clin. Virol. 53, 308–313.

International Committee on Taxonomy of Viruses, King, A.M.Q., 2012. Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses. Academic Press, London; Waltham, MA.

Kapikian, A.Z., James Jr., H.D., Kelly, S.J., Dees, J.H., Turner, H.C., McIntosh, K., Kim, H.W., Parrott, R.H., Vincent, M.M., Chanock, R.M., 1969. Isolation from man of avian infectious bronchitis virus-like viruses (coronaviruses) similar to 229E virus, with some epidemiological observations. J. Infect. Dis. 119, 282–290.

Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W., Rollin, P.E., Dowell, S.F., Ling, A.E., Humphrey, C.D., Shieh, W.J., Guarner, J., Paddock, C.D., Rota, P., Fields, B., DeRisi, J., Yang, J.Y., Cox, N., Hughes, J.M., LeDuc, J.W., Bellini, W.J., Anderson, L.J., 2003. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. 348, 1953–1966.

Liang, P., Pardee, A.B., 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257, 967–971.

Lu, Y., Wang, S.Y., Lotz, J.M., 2004. The use of differential display to isolate viral genomic sequence for rapid development of PCR-based detection methods. A test case using Taura syndrome virus. J. Virol. Methods 121, 107–114.

Mahony, J.B., Petrich, A., Smieja, M., 2011. Molecular diagnosis of respiratory virus infections. Crit. Rev. Clin. Lab. Sci. 48, 217–249.

McIntosh, K., Becker, W.B., Chanock, R.M., 1967a. Growth in suckling-mouse brain of IBV-like viruses from patients with upper respiratory tract disease. Proc. Natl. Acad. Sci. U. S. A. 58, 2268–2273.

McIntosh, K., Chao, R.K., Krause, H.E., Wasil, R., Mocega, H.E., Mufson, M.A., 1974. Coronavirus infection in acute lower respiratory tract disease of infants. J. Infect. Dis. 130, 502–507.

McIntosh, K., Dees, J.H., Becker, W.B., Kapikian, A.Z., Chanock, R.M., 1967b. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. Proc. Natl. Acad. Sci. U. S. A. 57, 933–940.

Moes, E., Vijgen, L., Keyaerts, E., Zlateva, K., Li, S., Maes, P., Pyrc, K., Berkhout, B., van der Hoek, L., Van Ranst, M., 2005. A novel pancoronavirus RT-PCR assay: frequent detection of human coronavirus NL63 in children hospitalized with respiratory tract infections in Belgium. BMC Infect. Dis. 5, 6.

Peiris, J.S., Lai, S.T., Poon, L.L., Guan, Y., Yam, L.Y., Lim, W., Nicholls, J., Yee, W.K., Yan, W.W., Cheung, M.T., Cheng, V.C., Chan, K.H., Tsang, D.N., Yung, R.W., Ng, T.K., Yuen, K.Y., 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 361, 1319–1325.

Pyrc, K., Milewska, A., Potempa, J., 2011. Development of loop-mediated isothermal amplification assay for detection of human coronavirus-NL63. J. Virol. Methods 175, 133–136.

Pyrc, K., Sims, A.C., Dijkman, R., Jebbink, M., Long, C., Deming, D., Donaldson, E., Vabret, A., Baric, R., van der Hoek, L., Pickles, R., 2010. Culturing the unculturable: human coronavirus HKU1 infects, replicates, and produces progeny virions in human ciliated airway epithelial cell cultures. J. Virol. 84, 11255–11263.

Pyrc, K., Stozek, K., Wojcik, K., Gawron, K., Zeglen, S., Karolak, W., Wojarski, J., Ochman, M., Hubalewska-Mazgaj, M., Bochenek, G., Sanak, M., Zembala, M., Szczeklik, A., Potempa, J., 2012. Use of sensitive, broad-spectrum molecular assays and human airway epithelium cultures for detection of respiratory pathogens. PloS ONE 7, e32582.

Reed, S.E., 1984. The behaviour of recent isolates of human respiratory coronavirus in vitro and in volunteers: evidence of heterogeneity among 229E-related strains. J. Med. Virol. 13, 179–192.

Regamey, N., Kaiser, L., Roiha, H.L., Deffernez, C., Kuehni, C.E., Latzin, P., Aebi, C., Frey, U., 2008. Viral etiology of acute respiratory infections with cough in infancy: a community-based birth cohort study. Pediatr. Infect. Dis. J. 27, 100–105.

Reyes, G.R., Kim, J.P., 1991. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. Mol. Cell. Probes 5, 473–481.

Rose, T.M., Schultz, E.R., Henikoff, J.G., Pietrokovski, S., McCallum, C.M., Henikoff, S., 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Res. 26, 1628–1635.

Sanghavi, S.K., Bullotta, A., Husain, S., Rinaldo, C.R., 2012. Clinical evaluation of multiplex real-time PCR panels for rapid detection of respiratory viral infections. J. Med. Virol. 84, 162–169.

Schildgen, O., Jebbink, M.F., de Vries, M., Pyrc, K., Dijkman, R., Simon, A., Muller, A., Kupfer, B., van der Hoek, L., 2006. Identification of cell lines permissive for human coronavirus NL63. J. Virol. Methods 138, 207–210.

Simoes, P.W., Medeiros, L.R., Simoes Pires, P.D., Edelweiss, M.I., Rosa, D.D., Silva, F.R., Silva, B.R., Rosa, M.I., 2012. Prevalence of human papillomavirus in breast cancer: a systematic review. Int. J. Gynecol. Cancer 22, 343–347.

Stephensen, C.B., Casebolt, D.B., Gangopadhyay, N.N., 1999. Phylogenetic analysis of a highly conserved region of the polymerase gene from 11 coronaviruses and development of a consensus polymerase chain reaction assay. Virus Res. 60, 181–189.

Thomazelli, L.M., Vieira, S., Leal, A.L., Sousa, T.S., Oliveira, D.B., Golono, M.A., Gillio, A.E., Stwien, K.E., Erdman, D.D., Durigon, E.L., 2007. Surveillance of eight respiratory viruses in clinical samples of pediatric patients in southeast Brazil. J. Pediatr. (Rio. J) 83, 422–428.

Tyrrell, D.A., Bynoe, M.L., 1965. Cultivation of a novel type of common-cold virus in organ cultures. Br. Med. J. 1, 1467–1470.

van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J., Berkhout, B., 2004. Identification of a new human coronavirus. Nat. Med. 10, 368–373.

Woo, P.C., Lau, S.K., Chu, C.M., Chan, K.H., Tsoi, H.W., Huang, Y., Wong, B.H., Poon, R.W., Cai, J.J., Luk, W.K., Poon, L.L., Wong, S.S., Guan, Y., Peiris, J.S., Yuen, K.Y., 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol. 79, 884–895.

Zlateva, K.T., Crusio, K.M., Leontovich, A.M., Lauber, C., Claas, E., Kravchenko, A.A., Spaan, W.J., Gorbalenya, A.E., 2011. Design and validation of consensus-degenerate hybrid oligonucleotide primers for broad and sensitive detection of corona- and toroviruses. J. Virol. Methods 177, 174–183.