Arch Virol (2013) 158:251–255
DOI 10.1007/s00705-012-1487-4

BRIEF REPORT

No novel coronaviruses identified in a large collection of human nasopharyngeal specimens using family-wide CODEHOP-based primers

Kalina T. Zlateva · Frank E. J. Coenjaerts · Kelly M. Crusio · Christine Lammens · Frank Leus · Marco Viveen · Margareta Ieven · Willy J. M. Spaan · Eric C. J. Claas · Alexander E. Gorbalenya

Received: 26 June 2012 / Accepted: 16 August 2012 / Published online: 28 September 2012

Abstract Novel viruses might be responsible for numerous disease cases with unknown etiology. In this study, we screened 1800 nasopharyngeal samples from adult outpatients with respiratory disease symptoms and healthy individuals. We employed a reverse transcription (RT)-PCR assay and CODEHOP-based primers (CT12-mCODEHOP) previously developed to recognize known and unknown corona- and toroviruses. The CT12-mCODEHOP assay detected 42.0 % (29/69) of samples positive for human coronaviruses (HCoV), including HCoV-229 (1/16), HCoV-NL63 (9/17), and HCoV-OC43 (19/36), and additionally HCoV-HKU1 (3), which was not targeted by the diagnostic real-time PCR assays. No other coronaviruses were identified in the analyzed samples.

K. T. Zlateva · K. M. Crusio · W. J. M. Spaan · E. C. J. Claas · A. E. Gorbalenya (✉)
Department of Medical Microbiology, Leiden University Medical Center, Albinusdreef 2, E4-P, P.O. Box 9600, 2300 RC Leiden, The Netherlands
e-mail: a.e.gorbalenya@lumc.nl

F. E. J. Coenjaerts · M. Viveen
Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands

C. Lammens · M. Ieven
Department of Medical Microbiology, Vaccine and Infectious Disease Institute, Universiteit Antwerpen, University Hospital Antwerp, Antwerp, Belgium

F. Leus
Department of Data Management, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

Keywords Coronavirus · Torovirus · CODEHOP · Virus discovery · RT-PCR · Real-time PCR · Respiratory tract infections

Respiratory tract infections (RTIs) are the most common diseases in humans and are responsible for significant morbidity and mortality worldwide, especially in young children, the elderly, and immunocompromised individuals. During the past decade, a growing variety of new viruses has been implicated in human RTIs, including severe acute respiratory syndrome coronavirus (SARS-HCoV) [1–3], human CoVs (HCoV) NL63 [4, 5] and HKU1 [6], human metapneumovirus (hMPV) [7], human bocavirus (HBoV) [8], WU and KI polyomaviruses (WUPyV and KIPyV, respectively) [9, 10], and viruses that belong to the previously unrecognized species *Human rhinovirus C* (HRV-C) [11, 12]. These discoveries were facilitated by advancements in diagnostic methods. Still, in recently conducted studies, no causative agent could be established in 20 to 45 % of respiratory disease cases [13–16]. These figures indicate that both the improvement of the current diagnostic methods and the search for new respiratory viruses must be continued.

The main challenge to the identification of new viruses is their unknown characteristics. One of the popular approaches that have succeeded is the use of either degenerate or consensus primers that recognize highly conserved genome regions of already known viruses. These primers can detect new pathogens that are closely related to members of established species. For the identification of distantly related unknown viruses, the combined consensus and degenerate primer design strategy implemented in the consensus-degenerate hybrid oligonucleotide primer (CODEHOP) technique proved to be superior [17–20].

Also, other genetic-based techniques that are sequence independent have been developed and used in virus discovery. For instance, the VIDISCA method, based on recognition of restricted enzyme cleavage sites, was introduced [5] and advanced by employing non-rRNA hexamers for random priming, high-throughput genome sequencing [21], and sample pooling using centrifugation [22]. It is highly sensitive and unbiased, but samples with an abundance of background RNA/DNA may not be suitable for VIDISCA, and its processivity remains relatively limited. Thus, CODEHOP- and VIDISCA-based techniques are complementary rather than an alternative for virus discovery at this stage of technology development.

We have recently described the development and validation of modified CODEHOP-based primers that were designed to recognize distantly related corona- and toroviruses of the family *Coronaviridae* (CT12-mCODEHOP) in a reverse transcription (RT)-PCR assay [23]. These primers detected four known circulating human coronaviruses and two bovine toroviruses directly from biological specimens. Additionally, *in silico* predictions indicated that the designed primers may recognize coronaviruses belonging to most recently discovered lineages. These results demonstrated that the CT12-mCODEHOP one-step RT-PCR assay can be a suitable tool for identification of novel members of the family *Coronaviridae*. In the current study, we present results of the first large-scale application of this method to a human respiratory specimen collection. We detected only the four currently known coronaviruses, providing further support for the prevalence of these respiratory coronaviruses in the human population.

A total of 1800 nasopharyngeal swabs collected from 10 primary-care networks (based in Antwerp, Gent, Rotenburg, Utrecht, Lodz, Barcelona, Mataro, Jonkoping, Cardiff and Southampton) in seven European countries (Belgium, Germany, the Netherlands, Poland, Spain, Sweden and the United Kingdom), as part of the EU Framework 6 supported GRACE (Genomics to Combat Resistance Against Antibiotics in Community-Acquired LRTI in Europe) Network were included in this study [21, 24]. The specimens were obtained from immunocompetent adult outpatients of $\geq$18 years of age with respiratory illness and asymptomatic matched controls over three consecutive winter seasons during the period 2007–2010. The inclusion criteria for symptomatic patients were acute cough with duration of up to and including 28 days or individuals in whom the general practitioner suspected the presence of acute lower respiratory tract infection. Control samples collected from the same practices were obtained from immunocompetent adults with no signs of acute respiratory illness of the same gender and age within 5 years of the patient's age (younger or older). One thousand sixty-four samples were collected from symptomatic adults at their first visit to the general practitioner (V1 samples), 484 were follow-up samples (V2 samples) obtained 4 weeks later, and 252 specimens were from subjects without respiratory disease symptoms. Total DNA/RNA nucleic acids were extracted using a NucliSENS EasyMAG (bioMerieux, Grenoble, France) and stored at −80 °C until further use. The samples were screened by in-house monoplex or multiplex real-time RT-PCR assays for the following respiratory pathogens: HCoV-229E, -OC43, and -NL63, human rhinovirus (HRV) A/B/C, parainfluenza virus (PIV) 1-4, hMPV, adenovirus (HAdV), HBoV, respiratory syncytial virus (RSV), influenza (Inf) A/B viruses, WUPyV, and KIPyV. In addition, bacterial infections were detected as well, partly by conventional culture, but also by molecular methods, for *Chlamydophila pneumoniae*, *Legionella pneumophila*, *Mycoplasma pneumoniae*, *Bordetella sp.*, *Streptococcus pneumoniae* and *Haemophilus spp.*

For real-time PCR detection of HCoV, viral RNA was reverse-transcribed into cDNA by using a MultiScribe reverse transcriptase kit and random hexamers (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's guidelines, followed by inactivation of reverse transcriptase for 5 min at 95 °C. HCoV type-specific primers and probes were based on the conserved regions of the nucleocapsid gene (Table 1). Their suitability for use in a multiplex assay was confirmed in a separate test that excluded that they cross-interact. Samples were assayed in a 50-µl reaction mixture containing 20 µl of cDNA, TaqMan universal PCR master mix (PE Applied Biosystems), primers, and fluorogenic probes labeled with the 5′ reporter dye 6-carboxy-fluorescein (FAM) and the 3′ quencher dye 6-carboxy-tetramethyl-rhodamine (TAMRA). Amplification and detection were performed with the Taqman 7900HT system, essentially as described elsewhere [25]. The presence of possible inhibitors of the amplification reaction was controlled by using a murine encephalomyocarditis virus (RNA) or phocine herpes virus (DNA) as internal controls. Coronavirus identification was performed by re-testing of the multiplex real-time PCR positive samples with the separate type-specific (monoplex) primer and probe sets. Interquartile median (IQM) and interquartile range (IQR) estimates were obtained with SPSS v17.

For virus discovery, GRACE samples were analyzed blindly in a pool of two as reported previously [23]. In brief, 8 µl of pooled RNA/DNA extractions containing 4 µl of each sample were digested with DNase using a DNase I Amplification Grade Kit, Life Technologies, Breda, The Netherlands) in a 10-µl final reaction volume following the manufacturer's instructions. The CT12-mCODEHOP RT-PCR assay was performed using a QIAGEN OneStep RT-PCR Kit (QIAGEN Benelux B.V., Venlo, The

Table 1 Species-specific primers and probes used for coronavirus detection in multiplex and monoplex PCR assays

Target	Orientation	Primer or probe sequence (5' → 3')	Concentration
HCoV-229E	S	CAG TCA AAT GGG CTG ATG CA	300 nM
	AS	CAA AGG GCT ATA AAG AGA ATA AGG TAT TCT	300 nM
		FAM-CCC TGA CGA CCA CGT TGT GGT TCA**-TAMRA**	100 nM
HCoV-NL63	S	GCG TGT TCC TAC CAG AGA GGA	50 nM
	AS	GCT GTG GAA AAC CTT TGG CA	300 nM
		FAM-ATG TTA TTC AGT GCT TTG GTC CTC GTG AT**-TAMRA**	100 nM
HCoV-OC43	S	CGA TGA GGC TAT TCC GAC TAG GT	900 nM
	AS	CCT TCC TGA GCC TTC AAT ATA GTA ACC	900 nM
		FAM-TCC GCC TGG CAC GGT ACT CCC T**-TAMRA**	125 nM

Netherlands) in a 50-µl final volume containing 2 µM A2 forward (5'-GGGIWGGGATTACMCIaartgygaymg-3') and 4 µM B1 reverse (5'-CCCASAIGWTGTIccnccnggytt-3') primers and 10 µl of the RNA sample. Amplification was carried out in a Biorad MyCycler (BioRad, Veenendaal, The Netherlands) using a touch-down thermocycling protocol as described elsewhere [23]. PCR products were analyzed in 2 % TAE agarose gels stained with ethidium bromide. Amplification products of about the expected 200-bp size were confirmed by individual re-testing of the pooled samples using 8 µl of RNA/DNA extract. The origin of the bands was established by sequence analysis of PCR products.

Employing a multiplex real-time PCR assay designed to detect HCoV-OC43, -229E and -NL63, 69 out of 1800 nasopharyngeal specimens were found to be positive and were subsequently typed using species-specific monoplex assays (Table 2). The viral load in specimens containing viruses of different HCoV species was comparable, corresponding to median Ct values ranging between 26.8 and 27.9. Parallel screening of all specimens was conducted using the CT12-mCODEHOP RT-PCR assay in the search for new members of the family *Coronaviridae*. The assay amplified a virus sequence in 1.8 % (32/1800) of the tested samples and in 42.0 % (29/69) of the samples found to be HCoV positive in the real-time PCR assays. One specimen was positive for HCoV-229E, 9 for HCoV-NL63, and 19 for HCoV-OC43; viral loads were relatively high and characterized by a median interquartile Ct value of 25.32 (interquartile range, 22.03-27.52). Additionally, the CT12-mCODEHOP RT-PCR assay detected three samples that were positive for HCoV-HUK1, which was not targeted by the real-time PCR assays, bringing the total number of coronavirus-positives to 72 out of 1800 samples (4 %). The majority of coronaviruses 5.6 % (59/1064) were identified in V1 samples obtained from symptomatic patients during their first visit to the general practitioner, 1.9 % (9/484) were detected in V2 follow-up samples, and 1.6 % (4/252) in asymptomatic controls.

The CT12-mCODEHOP RT-PCR assay showed the highest sensitivity towards HCoV-OC43, closely followed by HCoV-NL63, and then HCoV-229E. HCoV-OC43 infection was confirmed in 52.8 % (19/36) of the samples typed by real-time PCR, with viral loads corresponding to a Ct interquartile range of 22.44-27.56 (median 26.21). The detection rate for HCoV-NL63 was 53 % (9/17), with a higher median interquartile viral load than for HCoV-OC43, corresponding to Ct 23.6. Only one (6.7 %) HCoV-229E sample with a high virus load (Ct 21.14) was confirmed. The viral load in positive samples not detected by the CT12-mCODEHOP RT-PCR assay (false negatives) was relatively low (interquartile Ct value range, 27.75-34.51, median, 30.68). False-negative HCoV-229E and HCoV-NL63 samples had a lower median interquartile Ct value, 28.3 and 28.9, respectively, in comparison to 32.32 for HCoV-OC43, indicating higher viral loads and a lower sensitivity of the assay towards HCoV-229E and HCoV-NL63. These results are in accordance with our previous findings using quantified copy RNA transcripts and HCoV-positive respiratory specimens [23].

Our goal was to search for unknown respiratory human coronaviruses that are distantly related to known coronaviruses and are prototypes of new genera that might be circulating in the human population. No new coronaviruses were discovered in this study. Although this observation is negative and thus belongs to those rarely reported, its documentation is important for unbiased knowledge-building [26]. This study was the first, to our knowledge, in which a large collection of samples was screened using a primer set designed to recognize both corona- and toroviruses rather than targeting (human) coronaviruses exclusively (e.g., see ref. [27]). Nonspecific amplification products were not observed for any of the analyzed specimens, including numerous samples that were positive for other virus pathogens (data in preparation), indicating that the CT12-mCODEHOP RT-PCR assay is highly selective with respect to coronaviruses in nasopharyngeal samples. The lack of false positives contributed to a relatively high

Table 2 Detection of human coronaviruses using the CT12-mCODEHOP one-step RT-PCR assay in nasopharyngeal swabs that were positive for coronavirus by real-time PCR[a]

Coronavirus	N[b]	Detected		Not detected	
		N (%)	Ct IQM (IQR)[c]	N (%)	Ct IQM (IQR)[c]
HCoV-229E	16	1 (6.7 %)	21.14	15 (93.7 %)	28.28 (22.79-31.87)
HCoV-NL63	17	9 (53 %)	23.63 (22.0-27.52)	8 (47 %)	28.88 (27.10-34.28)
HCoV-OC43	36	19 (52.8 %)	26.21 (22.44-27.56)	17 (47.2 %)	33.04 (30.68-34.85)
All HCoVs	69	29 (42.0 %)	25.32 (22.03-27.52)	40 (58.0 %)	30.68 (27.75-34.51)

[a] HCoVs were initially detected using a multiplex coronavirus real-time RT-PCR and then typed by monoplex real-time PCR assays. Cycle threshold number (Ct) values obtained by both methods were similar, and the monoplex-based results are presented in the table

[b] Total number of HCoV-positives

[c] IQM, interquartile median and IQR, interquartile range, for Ct values

processivity and low cost of the assay that allowed us to analyze a large number of samples. In contrast, the rate of false negatives was considerable ($\sim$60 %) and varied for different HCoVs but was close to that established in our pilot study in which this assay was introduced [23]. If the false-negative rate of detection, especially for HCoV-229E, had been lower, even at the expense of an increase in false positives, our confidence about the lack of new coronaviruses in the analyzed samples probably would have been higher. Unfortunately, increasing the primer degeneracy in our assay, which controls sensitivity, was not practical. It is already high due to an enormous evolutionary distance between coronaviruses and toroviruses, which these primers were designed to cover [23]. The primer design was done according to an approach that minimizes bias with respect to any particular subset of coronaviruses, e.g., HCoVs, and thus covers the genetic diversity evenly. The validity of these primers was confirmed in *in silico* analysis of coronaviruses that were not used in the design of the CT12-mCODEHOP primers [23]. Thus, the observed false-negative rate, obviously high by virus diagnostic standards, is the current state of the art for this method. Because of the characteristics of this method, the negative conclusion about new respiratory human coronaviruses reported in this study is restricted to pathogens that are comparable in their prevalence and load to the known circulating respiratory coronaviruses. Furthermore, we acknowledge that this conclusion may also depend on a number of other variables, including, but not limited to, the time frame, population and geographic location surveyed. Even with these limitations, we believe that the results we have obtained are a step forward in characterizing the natural diversity of human coronaviruses and how we study it. Improving the method and expanding the application of this assay to larger datasets that also include stool and cerebrospinal fluid clinical specimens obtained from acute disease cases of unknown etiology would be worth exploring. This line of research could also benefit from application of virus-discovery techniques that involve deep next-generation sequencing (e.g., see ref. [28]).

Acknowledgments We thank our GRACE project colleagues Herman Goossens, Anton M. van Loon and all involved in the GRACE Primary Care Network for the opportunity to conduct this study, and Lia van der Hoek, Igor Sidorov, Chris Lauber and Michael Rozanov for advice and discussions. This study was partially supported by an EU FP6 GRACE grant (LSHM-CT12-2005-518226).

References

1. Drosten C, Preiser W, Gunther S, Schmitz H, Doerr HW (2003) Severe acute respiratory syndrome: identification of the etiological agent. Trends Mol Med 9:325–327
2. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong SX, Urbani C, Comer JA, Lim W, Rollin PE, Dowell SF, Ling AE, Humphrey CD, Shieh WJ, Guarner J, Paddock CD, Rota P, Fields B, DeRisi J, Yang JY, Cox N, Hughes JM, Leduc JW, Bellini WJ, Anderson LJ (2003) A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med 348:1953–1966
3. Peiris JS, Yuen KY, Osterhaus AD, Stohr K (2003) The severe acute respiratory syndrome. N Engl J Med 349:2431–2441
4. Fouchier RAM, Hartwig NG, Bestebroer TM, Niemeyer B, de Jong JC, Simon JH, Osterhaus ADME (2004) A previously undescribed coronavirus associated with respiratory disease in humans. Proc Natl Acad Sci USA 101:6212–6216
5. van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJM, Wolthers KC, van Wertheim-Dillen PME, Kaandorp J, Spaargaren J, Berkhout B (2004) Identification of a new human coronavirus. Nat Med 10:368–373
6. Woo PCY, Lau SKP, Chu CM, Chan KH, Tsoi HW, Huang Y, Wong BHL, Wong HL, Poon RWS, Cai JJ, Luk WK, Poon LLM, Wong SSY, Guan Y, Peiris JSM, Yuen KY (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 79:884–895
7. van den Hoogen BG, de Jong J, Groen J, Kuiken T, De Groot R, Fouchier RAM, Osterhaus ADME (2001) A newly discovered human pneumovirus isolated from young children with respiratory tract disease. Nat Med 7:719–724
8. Allander T, Tammi MT, 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 USA 102:12891–12896

9. Allander T, Andreasson K, Gupta S, Bjerkner A, Bogdanovic G, Persson MAA, Dalianis T, Ramqvist T, Andersson B (2007) Identification of a third human polyomavirus. J Virol 81:4130–4136
10. Gaynor AM, Nissen MD, Whiley DM, Mackay IM, Lambert SB, Wu G, Brennan DC, Storch GA, Sloots TP, Wang D (2007) Identification of a novel polyomavirus from patients with acute respiratory tract infections. PLoS Pathog 3:595–604
11. Lamson D, Renwick N, Kapoor V, Liu Z, Palacios G, Ju J, Dean A, George KS, Briese T, Lipkin WI (2006) MassTag polymerase-chain-reaction detection of respiratory pathogens, including a new rhinovirus genotype, that caused influenza-like illness in New York State during 2004–2005. J Infect Dis 194:1398–1402
12. Lau SKP, Yip CCY, Tsoi HW, Lee RA, So LY, Lau YL, Chan KH, Woo PCY, Yuen KY (2007) Clinical features and complete genome characterization of a distinct human rhinovirus (HRV) genetic previously undetected HRV cluster, probably representing a species, HRV-C, associated with acute respiratory illness in children. J Clin Microbiol 45:3655–3664
13. Ali SA, Gern JE, Hartert TV, Edwards KM, Griffin MR, Miller EK, Gebretsadik T, Pappas T, Lee WM, Williams JV (2011) Real-world comparison of two molecular methods for detection of respiratory viruses. Virol J 8:332
14. Arden KE, McErlean P, Nissen MD, Sloots TP, Mackay IM (2006) Frequent detection of human rhinoviruses, paramyxoviruses, coronaviruses, and bocavirus during acute respiratory tract infections. J Med Virol 78:1232–1240
15. Brittain-Long R, Andersson LM, Olofsson S, Lindh M, Westin J (2012) Seasonal variations of 15 respiratory agents illustrated by the application of a multiplex polymerase chain reaction assay. Scand J Infect Dis 44:9–17
16. Rand KH, Rampersaud H, Houck HJ (2011) Comparison of two multiplex methods for detection of respiratory viruses: FilmArray RP and xTAG RVP. J Clin Microbiol 49:2449–2453
17. Baines JE, McGovern RM, Persing D, Gostout BS (2005) Consensus-degenerate hybrid oligonucleotide primers (CODEHOP) for the detection of novel papillomaviruses and their application to esophageal and tonsillar carcinomas. J Virol Methods 123:81–87
18. Rose TM, Henikoff JG, Henikoff S (2003) CODEHOP (COnsensus-DEgenerate hybrid oligonucleotide primer) PCR primer design. Nucleic Acids Res 31:3763–3766
19. Rose TM (2005) CODEHOP-mediated PCR—a powerful technique for the identification and characterization of viral genomes. Virol J 2:20
20. Staheli JP, Ryan JT, Bruce AG, Boyce R, Rose TM (2009) Consensus-degenerate hybrid oligonucleotide primers (CODEHOPs) for the detection of novel viruses in non-human primates. Methods 49:32–41
21. de Vries M, Deijs M, Canuti M, van Schaik BDC, Faria NR, van de Garde MDB, Jachimowski LCM, Jebbink MF, Jakobs M, Luyf ACM, Coenjaerts FEJ, Claas ECJ, Molenkamp R, Koekkoek SM, Lammens C, Leus F, Goossens H, Ieven M, Baas F, van der Hoek L (2011) A sensitive assay for virus discovery in respiratory clinical samples. PLoS One 6:e16118
22. Tan LV, van Doorn HR, van der Hoek L, Hien VM, Jebbink MF, Ha DQ, Farrar J, Nguyen VVC, de Jong MD (2011) Random PCR and ultracentrifugation increases sensitivity and throughput of VIDISCA for screening of pathogens in clinical specimens. J Infect Dev Ctries 5:142–148
23. Zlateva KT, Crusio KM, Leontovich AM, Lauber C, Claas E, Kravchenko AA, Spaan WJM, Gorbalenya AE (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
24. Butler CC, Hood K, Verheij T, Little P, Melbye H, Nuttall J, Statistician MJK, Molstad S, Godycki-Cwirko M, Almirall J, Torres A, Gillespie D, Rautakorpi U, Coenen S, Goossens H (2009) Variation in antibiotic prescribing and its impact on recovery in patients with acute cough in primary care: prospective study in 13 countries. British Med J 338:b2242
25. Meerhoff TJ, Houben ML, Coenjaerts FEJ, Kimpen JLL, Hofland RW, Schellevis F, Bont LJ (2010) Detection of multiple respiratory pathogens during primary respiratory infection: nasal swab versus nasopharyngeal aspirate using real-time polymerase chain reaction. Eur J Clin Microbiol Infect Dis 29:365–371
26. Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2:696–701
27. 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
28. Daly GM, Bexfield N, Heaney J, Stubbs S, Mayer AP, Palser A, Kellam P, Drou N, Caccamo M, Tiley L, Alexander GJM, Bernal W, Heeney JL (2011) A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing. PLoS One 6:e28879