Genetic diversity of human rhinoviruses in Cambodia during a three-year period reveals novel genetic types

Monica Naughtin [a], Sareth Rith [a], Anne-Charlotte Sentilhes [a], Sirenda Vong [a], Marie-Line Joffret [b,c], Emmanuel Cornillot [d], Vincent Deubel [a], Francis Delpeyroux [b,c], Roger Frutos [d,e,*], Philippe Buchy [a,*]

[a] Institut Pasteur in Cambodia, Virology Unit, 5 Monivong Blvd, PO Box 983, Phnom Penh, Cambodia
[b] Institut Pasteur, Biologie des Virus Entériques, 25-28 rue du Dr Roux, 75724 Paris Cedex 15, France
[c] INSERM U994, Institut National de Santé et de La Recherche Médicale, Paris, France
[d] CEMIPAI, CPBS, UMR 5236, CNRS-UM1-UM2, 1919 route de Mende, 34293 Montpellier Cedex 5, France
[e] Cirad, UMR 17, Cirad-Ird, TA-A17/G, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France

ARTICLE INFO

ABSTRACT

Acute respiratory viral infections are a major cause of morbidity during early childhood in developing countries. Human rhinoviruses are the most frequent cause of upper respiratory tract infections in humans, which can range in severity from asymptomatic to clinically severe disease. In this study we collected 4170 nasopharyngeal swabs from patients hospitalised with influenza-like illness in two Cambodian provincial hospitals between 2007 and 2010. Samples were screened for 18 respiratory viruses using 5 multiplex PCRs. A total of 11.2% of samples tested positive for human rhinoviruses (HRV). VP4/2 and VP1 regions were amplified and sequenced to study the distribution of rhinoviruses genotypes and species in Cambodia during this three-year period. Five novel genotypes, 2 species A, 2 species B and 1 species C were identified based on VP1 sequences. Co-infections with other viruses were demonstrated.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Human rhinoviruses are highly prevalent respiratory picornaviruses of the *Enterovirus* genus. They display a single stranded positive sense RNA genome of approximately 7200 nucleotides encoding a single polyprotein which is cleaved by viral proteases into 11 individual proteins. Human rhinoviruses are amongst the most frequent causes of upper respiratory tract infections, and are significantly associated with several lower respiratory tract infections including acute wheezing, pneumonia, and bronchiolitis, and cause about one half of asthma exacerbations (Busse et al., 2010). One hundred classical serotypes have been defined by 20-fold or greater decrease in antigen cross-reactivity and form two genetically distinct species, A and B. There has recently been a renewed interest in rhinoviruses following the discovery of the highly prevalent and diverse species C group, using novel molecular-based approaches (McErlean et al., 2007; Briese et al., 2008; Kiang et al., 2008; Lee et al., 2007; Lau et al., 2007).

Species B viruses are generally considered to be rare, with an average prevalence rate of around 7% of total HRV strains (Briese et al., 2008; Miller et al., 2009a). Several studies suggest that species C viruses are a major cause of wheezing and exacerbation of asthma in children, and may have different clinical outcomes to other species [Lau et al., 2007; Miller et al., 2009b), however, other studies find no significant differences (Piotrowska et al., 2009; McCulloch et al., 2014).

The recent whole-genome sequencing of all species A and B, and several species C prototype strains provides a very useful framework for the genetic analysis and classification of clinical specimens, from partial or whole sequences (Palmenberg et al., 2009). Recently, new criteria have been proposed for the classification of species C viruses, which has simplified the identification of novel genotypes, or "genetic types." These criteria classify 33 genetic types, and 28 provisional types which require confirmation by additional VP1 gene sequencing (Simmonds et al., 2010). Within A and B species, there is generally a strong adherence to phylogenetic positioning (branching order) between the different structural protein genes, with a few exceptions (Savolainen et al., 2004; McIntyre et al., 2010). Species C viruses cluster in two phylogenetic groups on the 5' untranslated region (5UTR), one with species A (designated in some studies as Ca) and the other as a

separate C cluster (designated Cc) (Palmenberg et al., 2010; Huang et al., 2009). There is a consensus view that partial genome sequencing of at least two regions (VP4/2, VP1 or 3CD genes) of the HRV genome is sufficient for genetic type classification based on nucleotide homology (McIntyre et al., 2010; Arden and Mackay, 2010). Several recent studies support classification guidelines based on pairwise nucleotide homology distance thresholds, based on studies across the genome, and these data are supported by serotypical distinctions in A and B species. The proposed inter-serotype pairwise nucleotide divergence threshold for A and B species is 13% in the VP1 gene and 10% in the VP4/2 gene (Simmonds et al., 2010; McIntyre et al., 2010, 2013; Savolainen et al., 2002a; Wisdom et al., 2009a). These criteria also apply to the inter-genotype distance for species C viruses, which cannot be classified by serology due to their inability to grow in tissue culture (Simmonds et al., 2010; McIntyre et al., 2010). As a consequence, novel HRV strains were described in the past few years, particularly novel species C viruses (Briese et al., 2008; Renwick et al., 2007; Miller et al., 2009b; McIntyre et al., 2010; Wisdom et al., 2009b; Kistler et al., 2007). Potentially novel species A viruses are rarely reported, however, recently the strain A101 strain has been observed in several studies worldwide (Rathe et al., 2009).

Recombination, which occurs frequently in enteroviruses, was presumed to be uncommon in rhinoviruses, however, several studies have identified that recombination frequently occurs in the 5UTR region (McIntyre et al., 2010; Huang et al., 2009; Tapparel et al., 2009). Several studies have recently compared the location of "breakpoints" of recombination, to reveal typical recombination hotspots in the 5UTR (McIntyre et al., 2010; Huang et al., 2009). Recombination events in the coding regions seem to be rare (McIntyre et al., 2010; Tapparel et al., 2009).

In this work we have evaluated the incidence of acute lower respiratory tract illness (ALRI)-associated rhinoviruses infection during a 3-year hospital-based survey in Cambodia in two distinct geographical locations, and examined the genetic diversity of these viruses.

2. Materials and methods

2.1. Study population and samples

4170 nasopharyngeal respiratory samples were collected from patients of all ages hospitalised in Cambodia's Kampong Cham and Takeo provinces referral hospitals during the period between June 2007 and December 2009. The case definition of ALRI was as follows; for patients aged <5 years, ALRI was defined as an illness of <10 days duration that consisted of cough or difficult breathing, plus tachypnea. For patients aged between 5 and 14 years, case definition included the above symptoms as well as a fever >38 °C. For patients 15 years or older ALRI was defined as a fever >38 °C, plus chest pain or tachypnea or auscultatory crackles (Mardy et al., 2009). A severe case for patients <5 years of age was defined as ALRI in addition to: respiratory rate ⩾60/mn (age <2 months), respiratory rate ⩾50/mn (age <11 months), respiratory rate ⩾40/mn (aged between 1 and 5 years), saturation <93%, cardiac frequency >180 bpm (age <1 year), cardiac frequency >140 bpm (aged between 1 and 5 years), clinical respiratory distress, according to World Health Organisation (WHO) guidelines (WHO, 2005).

2.2. RT-PCR and sequencing

Samples were screened for the presence of 18 common and novel viral respiratory pathogens, including HRSV, human metapneumovirus (HMPV), human bocavirus (HBoV), influenza A and B viruses, human coronaviruses OC43, 229E, HKU1, and NL63, severe acute respiratory syndrome-associated coronavirus (SARS-CoV), parainfluenza viruses 1 to 4, adenoviruses, human rhinoviruses, and enteroviruses, by multiplex reverse transcriptase PCR (RT-PCR)/PCR, as previously reported (Buecher et al., 2010; Arnott et al., 2011, 2013a, 2013b). The RT-PCR step was amplifying an approximately 549 bp fragment of the human rhinoviruses VP4/VP2 and 5′-non-coding region (Savolainen et al., 2002b) and the subsequent semi-nested PCR was producing a 450 bp amplicon (Bellau-Pujol et al., 2005).

Samples positive for rhinoviruses were amplified in the 5UTR and VP4/2 regions using the primers VP4/2F and VP4/2R (Supplementary Table 1) (Lau et al., 2007; Huang et al., 2009). Certain samples could not be amplified using these primers and were amplified separately in the 5UTR and VP4/2 region (Supplementary Table 1). Amplification was performed with the Qiagen one-step RT-PCR kit, as described previously (Huang et al., 2009). For VP1 amplification, RNA was reverse transcribed with random hexamer primers (Promega, United Kingdom) using the superscript III reverse transcriptase (Invitrogen). PCR was then conducted using Gotaq polymerase (Promega), and species-specific primers which have been previously described (Supplementary Table 1) (Tapparel et al., 2009). For VP1 region amplification we used primers VP1 1.2F and VP1 1.18R (and nested PCR VP1 1.55F and VP1.18R) for species A, VP1 1.57F and VP1 2.19R (and nested PCR VP1 1.57F and VP1 1.18R for species B), VP1 P1.160F and VP1 2.68R (and nested VP1 P1.161F and VP1 2.69R) for species C. A touch-down program (annealing temp of 55–1 °C per cycle for 10 cycles, then 30 cycles at 45 °C) was used for both primary and nested PCRs. DNA products were separated by agarose gel electrophoresis and extracted using the QIAquick Gel Extraction Kit (Qiagen, Germany). Products were ligated into TOPO-TA vectors (Invitrogen, Carlsbad, CA). At least two plasmids were sequenced for each clinical sample. At least two recombinant clones were sequenced on an ABI3730 XL automatic DNA sequencer by Macrogen (Seoul, South Korea).

2.3. Sequence alignment and phylogenetic analysis

Alignments were conducted with Muscle, in SeaView 4.2.6 (Gouy et al., 2010). Owing to the very high level of genome saturation detected, no Maximum-Likelihood analysis was conducted for building trees. Trees were therefore built using a neighbour-joining approach. The robustness of nodes was assessed with 500 bootstrap replicates.

2.4. Genetic and sequence polymorphism analyses

Genetic similarity and difference matrices were constructed using Muscle in SeaView 4.2.6 (Gouy et al., 2010). Sequence polymorphism, subsequent tests and statistical tests were performed using functions from the DnaSP 5.10.01 package (Librado and Rozas, 2009). The numbers of haplotype (alleles) were analysed. Sequence polymorphism was investigated by calculation of several measures. Nucleotide diversity, Pi (π), the average number of nucleotide differences per site between two sequences was calculated, using the Jukes and Cantor correction. Theta (Watterson's mutation parameter) was calculated for the whole sequence from S. Eta (η) is the total number of mutations, and S is the number of polymorphic sites. Ka (the number of non-synonymous substitutions per non-synonymous site), and Ks (the number of synonymous substitutions per synonymous site) for any pair of sequences were calculated. Tajima's *D* test was used for statistical testing the hypothesis that all mutations are selectively neutral.

Statistical tests D^* and F^* were used for testing the hypothesis that all mutations are selectively neutral.

2.5. Ethics statement

The project (SISEA) was approved by the National Ethics Committee of Cambodia. All patients/parents of sick children who participated provided written informed consent.

2.6. Nucleotide sequence accession numbers

The sequences described in this study were deposited in Genbank under the following accession numbers:

VP1: KF034075, KF034076, KF034078 to KF034081, KF034083, LN623979 to LN623997, KM986111, KM986112
VP 2/4: KF034009 to KF034074
5′UTR sequences use for recombination were deposited under accession numbers KF033904 to KF034008.

3. Results

3.1. HRV sample screening and detection in two provincial hospital sites

The SISEA project (Surveillance of Infection in South East Asia) was established for surveillance of respiratory viruses in Cambodia, Vietnam, and Laos. Over a three-year period (2007–2009), we collected a total of 4170 respiratory samples from in-patients hospitalised with acute lower respiratory tract illness (ALRI) at two provincial hospitals in Cambodia (Takeo, southern Cambodia and Kampong Cham, central-north Cambodia).

Amongst the samples taken from Takeo hospital there were 977 children and 2279 adults. From the Kampong Cham hospital there was 584 children and 227 adults. RNA was extracted from nasopharyngeal (NPS) samples, and amplified by real time PCR to detect the presence of 18 viruses, as described in previous studies (Buecher et al., 2010). From 4170 patient samples, 1105 samples tested positive for one or more virus. HRV was detected in 455 samples, representing the most frequently detected virus (10.9% of all patients, and 41.2% of patients positive for viral detections), consistent with evidence that rhinoviruses are usually highly prevalent (Miller et al., 2009a; Huang et al., 2009; Buecher et al., 2010; Follin et al., 2009; Loens et al., 2006; Vong et al., 2013; Guerrier et al., 2013). Two hundred sixty-one of these patients were male, 194 female. Two hundred thirty-seven patients were under the age of 2 years (52%), with 273 (60%) under the age of 5, and 80 patients over the age of 50 (17.6%). As previously reported, ALRI in general and HRV infections in particular were observed throughout the year, without clear seasonality (Guerrier et al., 2013). In addition to typical symptoms of ALRI recorded for all patients, supplementary comprehensive clinical data was available from 94 patients, which allowed us report on typical symptoms associated with disease (Table 1), and also to form a case definition to identify severe rhinoviral infection. This definition included the following criteria; no bacterial infection (sputum-negative) or viral co-infection. For children >5 years, or adults, was the inclusion of at least two of the following; a systolic blood pressure of ⩽90 mmHg, a cardiac frequency of ⩾120 bpm, a respiratory rate of ⩾30/mn, saturation of ⩽93%, or a temperature of ⩽35 °C or ⩾40 °C. The criteria for children ⩽5 years was a respiratory rate of either ⩾60/mn (aged < 2 months), ⩾50/mn (aged between 2 and 12 months), ⩾40/mn (aged between 1 and 5 years), oxygen saturation of <93%, a cardiac frequency of >180 bpm (aged <1 year), or >140 bpm (aged between 1 and 5 years), and signs of

Table 1
Clinical characteristics and coinfections detected in HRV-positive samples.

Total clinical samples: 4170	Number (percentage)
Virus-positive cases	1105
HRV-positive cases	455 (41.2%)
Clinical symptoms (data available from 94 clinical samples)	
Pneumonia	33 (29.5%)
Bronchitis	11 (9.8%)
Bronchiotitis	14 (12.5%)
Pleurisy	8 (7.14%)
Asthma	6 (5.35%)
Rhinovirus + coinfection	34 (7.47%)
Coronovirus	5 (14.7%)
Influenza virus	0 (0%)
Bocavirus	4 (11.76%)
hMPV	1 (2.94%)
RSV	11 (32.35%)
PIV	8 (23.52%)
Adenovirus	5 (14.7%)
Enterovirus	0 (0%)

severe respiratory distress. From this analysis 16 samples fulfilled the criteria for severe rhinoviral infection (approximately 17%).

3.2. Genetic typing of the clinical detections

In order to evaluate the genetic diversity of HRV circulating in the considered human population and the potential involvement of variants or novel types in the occurrence of the severe forms recorded, samples were analysed by PCR. PCR amplification was targeted on the VP1 gene and the VP4/2 gene region. A total of 28 VP1 sequences were obtained and genotyped (Table 2). The VP1 PCR failed on a number of samples, suggesting sequence divergence from the primer sequences used. Out of these 28 viruses 21 belonged to species A, 5 were members of species B and 2 corresponded to species C. Altogether, 5 viruses met the pairwise distance thresholds defined by McIntyre et al. (2013) to be recognised as novel types. 2 belonged to species A (i.e. D386 and K590 with 19.7% and 16.6% distance, respectively) and 2 others belonged to species B (i.e. K606 and P128 with 17% and 22.9% distance, respectively). The last sample corresponding to a new type, i.e. R005, belonged to species C and displayed a pairwise distance to the closest relative of 30.9%.

The VP4/2 sequence was also considered for genotyping the viruses. Sixty VP4/2 sequences were obtained which were distributed into 31, 11 and 18 sequences for species A, B and C, respectively (Table 3). We also detected one enterovirus (EV68), which was previously recognised as HRV87 until it was reclassified as a variant of EV68 (Ishiko et al., 2002). Viruses which were defined as clinically severe are highlighted in Table 3. Sixteen viruses in species A displayed pairwise distances at nucleic acid level to the closest relative ranging from 10% to 19.9%. With respect to species B, 10 viruses out of 11 displayed pairwise distances ranging from 15.8% to 18.8%. The only virus from species C showing a VP4/2 pairwise distance above the 10% threshold to the closest relative was R005 (29.1%) which was identified as a novel type based on VP1 sequence distances. Interestingly, the VP4/2 closest relative, i.e. HRV-C15, was different than the VP1 closest relative, i.e. HRV-C23. However, when considering the pairwise distances at the protein level, a different feature was seen. Several viruses for which the nucleotide pairwise distance with the closest relative was clearly above the 10% threshold displayed a protein sequence 100% identical to that of the same closest virus. Interestingly, several of these putative new types based on nucleotide distance, ,displayed amino-acid distances lower than sample

Table 2
Species classification of clinical detections based on VP1 sequences.

Sample	VP1 nucleic acid		VP1 protein	
	Closest (Sim/Dist)	Species	Closest (Sim/Dist)	Species
D123	HRV-A21 (0.894/0.106)	A	HRV-A21 (0.951/0.0.49)	A
D386	*HRV-A88 (0.803/0.197)*	*A*	*HRV-A88 (0.869/0.131)*	*A*
D502	HRV-A1 (0.922/0.078)	A	HRV-A1 (0.956/0.044)	A
D609	HRV-A20 (0.906/0.094)	A	HRV-A20 (0.980/0.020)	A
D722	HRV-A82 (0.902/0.098)	A	HRV-A82 (0.976/0.024)	A
E208	HRV-A58 (0.920/0.080)	A	HRV-A58 (0.972/0.0.028)	A
E214	HRV-A58 (0.920/0.080)	A	HRV-A58 (0.972/0.0.028)	A
E304	HRV-A60 (0.910/0.090)	A	HRV-A60 (0.974/0.026)	A
E332	HRV-A88 (0.932/0.068)	A	HRV-A88 (0.982/0.018)	A
E485	HRV-A54 (0.905/0.095)	A	HRV-A54 (0.974/0.026)	A
E5061	HRV-A49 (0.955/0.045)	A	HRV-A49 (0.982/0.018)	A
E5062	HRV-A49 (0.950/0.050)	A	HRV-A49 (0.982/0.018)	A
K169	HRV-A78 (0.934/0.066)	A	HRV-A78 (0.986/0.014)	A
K312	HRV-A80 (0.916/0.084)	A	HRV-A80 (0.976/0.024)	A
K315	HRV-A103 (0.989/0.011)	A	HRV-A103 (1.000/0.000)	A
K522	HRV-A85 (0.902/0.098)	A	HRV-A85 (0.969/0.031)	A
K535	HRV-A20 (0.906/0.094)	A	HRV-A20 (0.980/0.020)	A
K590	*HRV-A88 (0.834/0.166)*	*A*	*HRV-A88 (0.968/0.032)*	*A*
M121	HRV-A21 (0.902/0.0.098)	A	HRV-A21 (0.948/0.0.52)	A
R090	HRV-A85 (0.892/0.0108)	A	HRV-A85 (0.937/0.0.63)	A
R120	HRV-A73 (0.907/0.0.093)	A	HRV-A73 (0.962/0.038)	A
K531	HRV-B35 (0.982/0.018)	B	HRV-B35 (0.997/0.003)	B
K589	HRV-B35 (0.982/0.018)	B	HRV-B35 (1.000/0.000)	B
K606	*HRV-B48 (0.830/0.170)*	*B*	*HRV-B48 (0.915/0.085)*	*B*
P128	*HRV-B97 (0.771/0.229)*	*B*	*HRV-B97 (0.868/0.132)*	*B*
D782	HRV-B83 (0.913/0.087)	B	HRV-B83 (0.967/0.033)	B
R005	*HRV-C30 (0.691/0.309)*	*C*	*HRV-C30 (0.778/0.222)*	*C*
E463	HRV-C8 (0.939/0.061)	C	HRV-C8 (0.992/0.008)	C

Novel types are shown in bold italic.

K169.2 for which the nucleotide distance was below the 10% threshold (i.e. 8.5%). For instance, sample K169.2 which cannot be considered a new genotype with 8.5% nucleotidic distance displayed a protein distance of 1.7% whereas sample K606, considered a putative new genotype with 18.8% nucleotidic distance showed a protein distance of only 0.8%. Protein distances for all the 10 putative new species B genotypes ranged from 0.8% to 3.3%. Protein distances for putative new species A genotypes ranged from 0% to 5.2% with 15 viruses out of 16 displaying protein distances below 3%. The novel VP1 type R005, although displaying 29.1% distance to the closest relative (HRV-C15) based on the gene sequence, was characterised by a protein sequence 100% identical to that of HRV-C15.

3.3. Transition/transversion ratio and genome saturation

Owing to the variations observed between the nucleic acid and protein distances, nucleotide polymorphism and genome saturation were assessed to determine the solidity of a phylogenetic analysis. The ratio of observed transition (mutation from a purine nucleotide to another or a pyrimidine nucleotide to another) versus transversion (mutation for a purine nucleotide to pyrimidine one or from a pyrimidine nucleotide to purine one) was calculated as an indicator of genome saturation. Data are summarised in Supplementary Table 2. The ratio ranged from 0.714 to 1.086 for the coding sequences depending on the gene and virus species. Owing to the presence of indels in the VP1 sequences, calculation was done on indel-free segments. For the 5'UTR, the transition/transversion ratio was 0.515. For reference, the transition/transversion ratio was calculated for the gene E of the dengue DENV-1 virus (Duong et al., 2013) with a value of 14.49. The rate of polymorphic sites (S) ranged from 44.3% to 71.3 5% depending on the sequence while Θ ranged from 42.8 to 101.8. The DENV-1 E gene used as a reference displayed a rate of polymorphic sites of 12.7% and a Θ value of 40.1. The ratio between parsimony informative sites (Pa) and singletons (Si) was strongly biased towards the former with a rate of Pa ranging from 88.5% to 96.1% for the coding regions while it was 84.3% for the 5'UTR. For reference, the values of Pa and Si were very close for the DENV-1 gene E, although higher for Si, with rates of 48.7% and 51.3%, respectively. A strong bias was also observed in the rate of synonymous replacement (Ns) which ranged from 86.2% to 96.4%, the rate for the DENV-1 E gene being in the same range with 84.9%. Accordingly, a very low value of the Ka/Ks was observed for all sequences. No bias was observed on the individual rate of A, T, G and C bases.

3.4. Phylogenetic analyses and protein polymorphism assessment

Phylogenetic analysis was conducted on VP1 and VP4/2 (Fig. 1a and b). Owing to the high level of genome saturation observed when analysing the transition/transversion ratio, the topology of trees based on protein alignments was also determined for the VP1 and VP4/2 proteins in order to assess congruent topologies between nucleic acids and proteins. The high level of genome saturation could not allow for running relevant evolution models for maximum likelihood analysis and trees were therefore built using the distance neighbour-joining method. For both VP1 and VP4/2 nucleic acids and protein trees were perfectly congruent (Supplementary Figs. 1 and 2).

3.5. Recombination across the 5UTR

Several studies have examined recombination and found that the region between the VP4/2 and 5UTR could be a "hotspot" for recombination events (Huang et al., 2009; Tapparel et al., 2009; McIntyre et al., 2010). A similar analysis was conducted and the putative recombination breakpoints for the viruses from species A and C observed in this study were found to be located in the same area as previously described (Huang et al., 2009) (Data not shown).

Table 3
Species classification of clinical detections based on VP4/2 sequences.

Sample	VP4/2 nucleic acid		VP4/2 protein	
	Closest (Sim/Dist)	Species	Closest (Sim/Dist)	Species
D101	HRV-A10 (0.902/0.098)	A	HRV-A10 (0.991/0.009)	A
D123	HRV-A57 (0.818/0.182)	A	HRV-A57 (1.000/0.000)	A
D214	HRV-A15 (0.921/0.079)	A	HRV-A15 (0.991/0.009)	A
D386	**HRV-A89 (0.862/0.138)**	**A**	**HRV-A89 (0.983/0.017)**	**A**
D609	HRV-A68 (0.856/0.144)	A	HRV-A68 (0.974/0.026)	A
D722	HRV-A82 (0.896/0.104)	A	HRV-A82 (0.991/0.009)	A
E031	HRV-A16 (0.921/0.079)	A	HRV-A16 (0.991/0.009)	A
E186	HRV-A36 (0.915/0.085)	A	HRV-A36 (1.000/0.000)	A
E201	HRV-A78 (0.918/0.082)	A	HRV-A78 (0.991/0.009)	A
E208	HRV-A89 (0.886/0.114)	A	HRV-A89 (0.983/0.017)	A
E214	HRV-A80 (0.890/0.110)	A	HRV-A80 (0.991/0.009)	A
E241	HRV-A78 (0.912/0.088)	A	HRV-A78 (0.991/0.009)	A
E253	HRV-A67 (0.933/0.076)	A	HRV-A67 (1.000/0.000)	A
E301	HRV-A16 (0.921/0.079)	A	HRV-A16 (0.991/0.009)	A
K143	HRV-A10 (0.915/0.085)	A	HRV-A10 (1.000/0.000)	A
K169-1	HRV-A89 (0.862/0.138)	A	HRV-A89 (0.974/0.026)	A
K169-2	HRV-A78 (0.915/0.085)	A	HRV-A78 (0.983/0.017)	A
K312	HRV-A80 (0.887/0.113)	A	HRV-A80 (0.983/0.017)	A
K315	HRV-A65 (0.801/0.199)	A	HRV-A65 (0.948/0.052)	A
K522	HRV-A40 (0.854/0.146)	A	HRV-A40 (1.000/0.000)	A
K535	HRV-A68 (0.849/0.151)	A	HRV-A68 (0.974/0.026)	A
K590	**HRV-A7 (0.863/0.137)**	**A**	**HRV-A7 (0.991/0.009)**	**A**
M121	HRV-A57 (0.814/0.186)	A	HRV-A57 (1.000/0.000)	A
P058	HRV-A67 (0.933/0.076)	A	HRV-A67 (1.000/0.000)	A
P157	HRV-A60 (0.915/0.085)	A	HRV-A60 (0.991/0.009)	A
P282	HRV-A80 (0.900/0.100)	A	HRV-A80 (0.983/0.017)	A
P291	HRV-A47 (0.912/0.088)	A	HRV-A47 (1.000/0.000)	A
R010	HRV-A23 (0.921/0.079)	A	HRV-A23 (1.000/0.000)	A
R056	HRV-A51 (0.921/0.079)	A	HRV-A51 (0.991/0.009)	A
R090	HRV-A40 (0.857/0.143)	A	HRV-A40 (0.991/0.009)	A
R120	HRV-A73 (0.870/0.130)	A	HRV-A73 (0.991/0.009)	A
D218	HRV-B79 (0.832/0.168)	B	HRV-B79 (0.967/0.033)	B
D346	HRV-B27 (0.832/0.168)	B	HRV-B27 (0.975/0.025)	B
D782	HRV-B83 (0.930/0.070)	B	HRV-B83 (0.992/0.008)	B
E192	HRV-B27 (0.832/0.168)	B	HRV-B27 (0.975/0.025)	B
E209	HRV-B27 (0.825/0.175)	B	HRV-B27 (0.967/0.033)	B
E496	HRV-B79 (0.839/0.161)	B	HRV-B79 (0.967/0.033)	B
K148	HRV-B79 (0.842/0.158)	B	HRV-B79 (0.975/0.025)	B
K531	HRV-B79 (0.842/0.158)	B	HRV-B79 (0.967/0.033)	B
K589	HRV-B79 (0.839/0.161)	B	HRV-B79 (0.967/0.033)	B
K606	**HRV-B52 (0.812/0.188)**	**B**	**HRV-B52 (0.992/0.008)**	**B**
P128	**HRV-B27 (0.828/0.172)**	**B**	**HRV-B27 (0.975/0.025)**	**B**
E170	HRVCpat21 (0.966/0.034)	C	HRVCpat21 (0.966/0.034)	C
E175	HRVC-pat5 (0.960/0.040)	C	HRVC-pat5 (0.983/0.017)	C
E179	HRV-C12 (0.948/0.052)	C	HRV-C12 (0.992/0.008)	C
E223	HRV-C23 (0.930/0.070)	C	HRV-C23 (0.992/0.008)	C
E227	HRV-Cpat11 (0.969/0.031)C			
E255	HRV-C6 (0.980/0.020)	C	HRV-C6 (0.992/0.008)	C
E257	HRV-C6 (0.986/0.014)	C	HRV-C6 (1.000/0.000)	C
E284	HRV-C1 (0.974/0.026)	C	HRVC_pat14 (0.992/0.008)C	
E297	HRV-C6 (0.986/0.014)	C	HRV-C6 (1.000/0.000)	C
E463	HRV-C8 (0.918/0.082)	C	HRV-C8 (0.983/0.017)	C
E506	HRV-C32 (0.986/0.014)	C	HRV-C32 (1.000/0.000)	C
E554	HRV-C26 (0.963/0.037)	C	HRV-C26 (0.966/0.034)	C
K237	HRV-C32 (0.986/0.014)	C	HRV-C32 (1.000/0.000)	C
P145	HRV-C1 (0.966/0.034)	C	HRVC_pat14 (0.992/0.008)C	
P225	HRV-C11 (0.933/0.067)	C	HRV-Cpat4 (1.000/0.000)	C
P246	HRV-C8 (0.915/0.085)	C	HRV-C8 (0.983/0.017)	C
P531	HRV-C23 (0.936/0.064)	C	HRV-C23 (0.992/0.008)	C
R005	**HRV-C15 (0.709/0.291)**	**C**	**HRV-C15 (1.000/0.000)**	**C**
R102	HRV-C22 (0.954/0.046)	C	HRV-C22 (0.966/0.034)	C

Strains which were obtained from clinically severe hospitalised patients are highlighted in grey.
Novel types identified based on VP1 distances are shown in bold italic.

4. Discussion

This work is one of the largest respiratory virus surveillance studies, with more than 4000 clinical samples, providing evidence of the major rank occupied by HRV amongst respiratory viruses causing influenza-like illness. Rhinoviruses are frequently concurrently detected with other respiratory viruses due to their high circulation in the community. However, some studies have proposed that despite the frequency of co-detections, rhinoviruses are statistically the least likely to be involved in coinfection and may actually reduce the probability of another virus infecting the host (Duong et al., 2013). HRV/HRV co-detections have
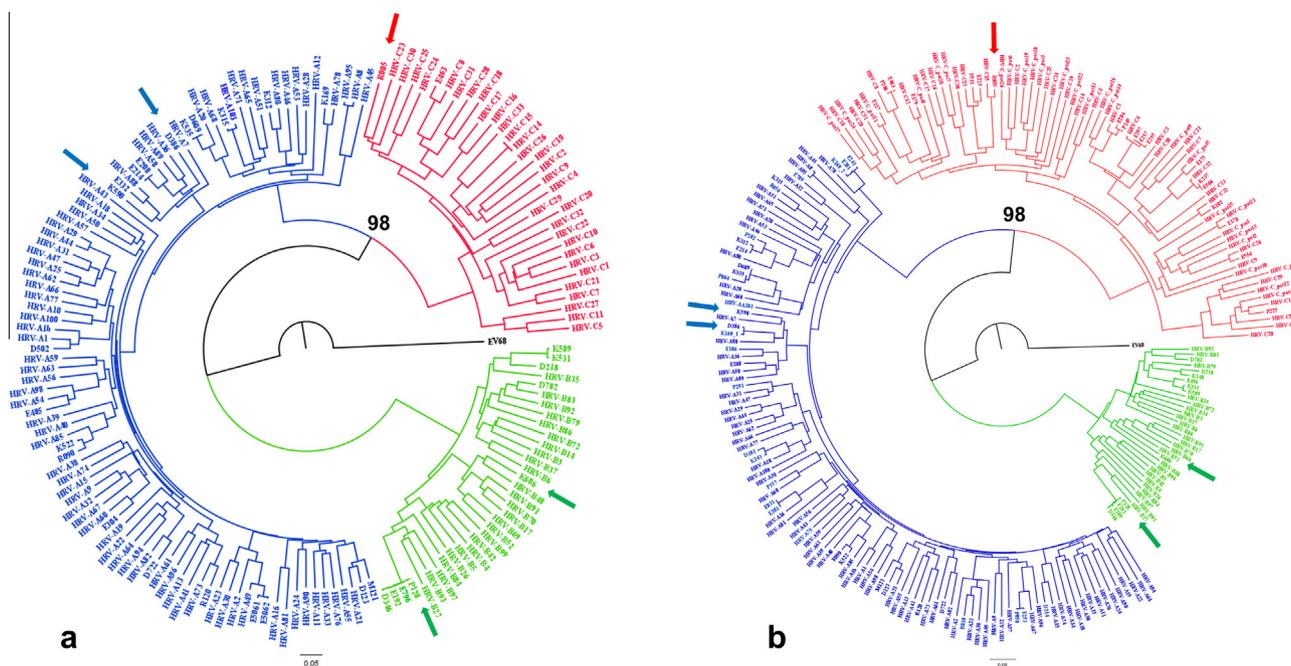
Fig. 1. Neighbour-joining tree of VP1 and VP4/2 sequences. Neighbour-joining tree of VP1 (a) and VP4/2 (b) sequences of HRV clinical detections and reference sequences. The HEV-68 strain was used as an outgroup. Novel types are identified by arrows. Reference sequences are marked by the prefix HRV-.

occasionally been reported (Lee et al., 2007; Renwick et al., 2007; Huang et al., 2009), although, interestingly no co-detections from the same species have been reported. In this work we report one HRV/HRV co-detection, and significantly both viruses detected from one patient were species A (HRV78 and HRV81). The frequency of HRV co-detections is difficult to assess given that often only a few clones are sequenced and therefore if one virus is present in a lesser proportion than the other, it might not be detected. In this work we analysed 455 viruses and we found out that 35 only out of these 455 were co-detected with other viruses (7.7% of total). This percentage is even lower than that reported in previous studies, i.e. 23% of total [40] which were concluding that rhinoviruses were statistically the least likely respiratory virus to be involved in co-infection (Greer et al., 2009). This further highlights the independent role of rhinoviruses as major aetiological contributors to severe respiratory disease.

Few studies to date have been able to link clinical and virological data to identify strains with increased virulence. The viruses from clinically severe patients in this study comprised 6 members of species A, including one sample resembling the A101 novel genetic type (Rathe et al., 2009), 1 virus from species B, 3 viruses from species C and 1 enterovirus (EV-68). Previous studies have suggested that species C viruses may be associated with increased disease severity (Lauinger et al., 2013). However, no identifiable clustering of patterns of the clinically severe viruses could be seen in this work. The reported clinically severe viruses were spread in different genotypic groups, indicating that severity might not be an evolutionary pattern strictly related to viruses or sequences analysed in this work.

Since the characterisation of the prototype strains by serology in the 1960's there have been very few reports of novel viruses from A or B species. A recent study sequenced the entire genome of a virus designated A101, which has a nucleotide identity across the genome of only 76.6% with the nearest species A virus (HRV-53) (Rathe et al., 2009). Another study conducted in Thailand found species A to be the most common species detected regardless of age groups (Fry et al., 2011). Following the standards developed by McIntyre et al. (2013) five novel genotypes were

identified in this work with respect to the VP1 gene. Two viruses from species A (i.e. D386 and K590), two viruses from species B (i.e. K606 and P128) and one virus from species C (i.e. R005). The divergence at the nucleic acid level was high enough above the threshold set by McIntyre et al. (2013) to consider these viruses actual new genotypes. This conclusion is strongly supported for three viruses, i.e. D386, P128 and R005, by the high divergence also observed at the protein level. This protein divergence was however very limited for the last two VP1 variants (i.e. K606 and K590). A similar situation is encountered when using VP4/2 as a target region for genotyping. Most of the putative variants identified on nucleotide distance display a very limited divergence with the closest relative at protein level (less than 3%) and in several cases there is a perfect identity. The best example found in this work is perhaps the novel VP1 type, R005, which displays the highest divergence to the closest relative of all novel types based on the VP1 gene and protein sequences and the highest divergence to the closest relative on the VP4/2 gene sequence but 100% VP4/2 protein sequence identity. In addition, the closest relative to R005 is not the same according to VP1 or VP4/2.

The polymorphism and genome saturation analyses conducted in this work suggest that HRV viruses might be exposed to two opposite trends: a very high rate of mutations generating genome saturation and a strong negative selective pressure resulting in highly conserved proteins. It is therefore difficult when in presence of conflicting nucleic-acid and amino-acid distance data to formally assign these viruses to a novel type. We suggest that the typology standards developed by McIntyre et al. (2013) should be completed with protein thresholds into order to assign unambiguously these conflicting detections to specific types.

The rhinoviruses studied in this work can be clearly discriminated into three genetic distinct taxa, i.e. the already recognised species A, B and C. This is especially true for the VP1 gene which displays indels. However, the relative position of the detected viruses within each taxon and their phylogeny cannot be established. Both the high rate of polymorphism and level of saturation observed can easily result in the occurrence of reverse mutations and blurred phylogeny. The clear genetic clustering and the high

negative selective pressure observed on both the VP1 and VP4/2 genes suggest that the three species described might correspond to virus groups having undergone independent evolution and specialisation.

This work represents the insights into HRV genetic diversity in South East Asia and provides the first comprehensive view of the HRV genetic diversity and the association between disease severity and infecting genotype in Cambodia during a 3-year period. This work also provides the first demonstration of the very high level of genome saturation and nucleotidic polymorphism of HRV indicating that typing and classification standards should comprise protein distance thresholds. Finally, it shows that much more research is needed to understand the complex dynamics of this virus.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.meegid.2015.07.030.

References

Arden, K.E., Mackay, I.M., 2010. Newly identified human rhinoviruses: molecular methods heat up the cold viruses. Rev. Med. Virol. 20, 156–176.

Arnott, A., Vong, S., Mardy, S., Chu, S., Naughtin, M., Sovann, L., Buecher, C., Beaute, J., Rith, S., Borand, L., Asgari, N., Frutos, R., Guillard, B., Touch, S., Deubel, V., Buchy, P., 2011. A study of the genetic variability of human respiratory syncytial virus (HRSV) in Cambodia reveals the existence of a new HRSV group B genotype. J. Clin. Microbiol. 49, 3504–3513.

Arnott, A., Vong, S., Rith, S., Naughtin, M., Ly, S., Guillard, B., Deubel, V., Buchy, P., 2013a. Human bocavirus amongst an all-ages population hospitalised with acute lower respiratory infections in Cambodia. Influenza Other Respir. Viruses 7, 201–210.

Arnott, A., Vong, S., Sek, M., Naughtin, M., Beaute, J., Rith, S., Guillard, B., Deubel, V., Buchy, P., 2013b. Genetic variability of human metapneumovirus amongst an all age population in Cambodia between 2007 and 2009. Infect Genet. Evol. 15, 43–52.

Bellau-Pujol, S., Vabret, A., Legrand, L., Dina, J., Gouarin, S., Petitjean-Lecherbonnier, J., Pozzetto, B., Ginevra, C., Freymuth, F., 2005. Development of three multiplex RT-PCR assays for the detection of 12 respiratory RNA viruses. J. Virol. Methods 126, 53–63.

Briese, T., Renwick, N., Venter, M., Jarman, R.G., Ghosh, D., Koendgen, S., Shrestha, S.K., Hoegh, A.M., Casas, I., Adjogoua, E.V., Akoua-Koffi, C., Myint, K.S., Williams, D.T., Chidlow, G., van den Berg, R., Calvo, C., Koch, O., Palacios, G., Kapoor, V., Villari, J., Dominguez, S.R., Holmes, K.V., Harnett, G., Smith, D., Mackenzie, J.S., Ellerbrok, H., Schweiger, B., Schønning, K., Chadha, M.S., Leendertz, F.H., Mishra, A.C., Gibbons, R.V., Holmes, E.C., Lipkin, W.I., 2008. Global distribution of novel rhinovirus genotype. Emerg. Infect. Dis. 14, 944–947.

Buecher, C., Mardy, S., Wang, W., Duong, V., Vong, S., Naughtin, M., Vabret, A., Freymuth, F., Deubel, V., Buchy, P., 2010. Use of a multiplex PCR/RT-PCR approach to assess the viral causes of influenza-like illnesses in Cambodia during three consecutive dry seasons. J. Med. Virol. 82, 1762–1772.

Busse, W.W., Lemanske Jr., R.F., Gern, J.E., 2010. Role of viral respiratory infections in asthma and asthma exacerbations. Lancet 376, 826–834.

Duong, V., Simmons, C., Gavotte, L., Viari, A., Ong, S., Chantha, N., Lennon, N.J., Birren, B.W., Vong, S., Farrar, J.J., Henn, M.R., Deubel, V., Frutos, R., Buchy, P., 2013. Genetic diversity and lineage dynamic of dengue virus serotype 1 (DENV-1) in Cambodia. Infect. Genet. Evol. 15, 59–68.

Follin, P., Lindqvist, A., Nystrom, K., Lindh, M., 2009. A variety of respiratory viruses found in symptomatic travellers returning from countries with ongoing spread of the new influenza A(H1N1)v virus strain. Euro. Surveill. 14.

Fry, A.M., Lu, X., Olsen, S.J., Chittaganpitch, M., Sawatwong, P., Chantra, S., Baggett, H.C., Erdman, D., 2011. Human rhinovirus infections in rural Thailand: epidemiological evidence for rhinovirus as both pathogen and bystander. PLoS ONE 6, e17780.

Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221–224.

Greer, R.M., McErlean, P., Arden, K.E., Faux, C.E., Nitsche, A., Lambert, S.B., Nissen, M.D., Sloots, T.P., Mackay, I.M., 2009. Do rhinoviruses reduce the probability of viral co-detection during acute respiratory tract infections? J. Clin. Virol. 45, 10–15.

Guerrier, G., Goyet, S., Chheng, E.T., Rammaert, B., Borand, L., Te, V., Try, P.L., Sareth, R., Cavailler, P., Mayaud, C., Guillard, B., Vong, S., Buchy, P., Tarantola, A., 2013. Acute viral lower respiratory tract infections in Cambodian children: clinical and epidemiologic characteristics. Pediatr. Infect. Dis. J. 32, e8–e13.

Huang, T., Wang, W., Bessaud, M., Ren, P., Sheng, J., Yan, H., Zhang, J., Lin, X., Wang, Y., Delpeyroux, F., Deubel, V., 2009. Evidence of recombination and genetic diversity in human rhinoviruses in children with acute respiratory infection. PLoS ONE 4, e6355.

Ishiko, H., Miura, R., Shimada, Y., Hayashi, A., Nakajima, H., Yamazaki, S., Takeda, N., 2002. Human rhinovirus 87 identified as human enterovirus 68 by VP4-based molecular diagnosis. Intervirology 45, 136–141.

Kiang, D., Kalra, I., Yagi, S., Louie, J.K., Boushey, H., Boothby, J., Schnurr, D.P., 2008. Assay for 5′ noncoding region analysis of all human rhinovirus prototype strains. J. Clin. Microbiol. 46, 3736–3745.

Kistler, A., Avila, P.C., Rouskin, S., Wang, D., Ward, T., Yagi, S., Schnurr, D., Ganem, D., DeRisi, L., Boushey, H.A., 2007. Pan-viral screening of respiratory tract infections in adults with and without asthma reveals unexpected human coronavirus and human rhinovirus diversity. J. Infect. Dis. 196, 817–825.

Lau, S.K., Yip, C.C., Tsoi, H.W., Lee, R.A., So, L.Y., Lau, Y.L., Chan, K.H., Woo, P.C., Yuen, K.Y., 2007. Clinical features and complete genome characterization of a distinct human rhinovirus (HRV) genetic cluster, probably representing a previously undetected HRV species, HRV-C, associated with acute respiratory illness in children. J. Clin. Microbiol. 45, 3655–3664.

Lauinger, I.L., Bible, J.M., Halligan, E.P., Bangalore, H., Tosas, O., Aarons, E.J., MacMahon, E., Tong, C.Y., 2013. Patient characteristics and severity of human rhinovirus infections in children. J. Clin. Virol. 58, 216–220.

Lee, W.M., Kiesner, C., Pappas, T., Lee, I., Grindle, K., Jartti, T., Jakiela, B., Lemanske Jr, R.F., Shult, P.A., Gern, J.E., 2007. A diverse group of previously unrecognized human rhinoviruses are common causes of respiratory illnesses in infants. PLoS ONE 2, e966.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

Loens, K., Goossens, H., de Laat, C., Foolen, H., Oudshoorn, P., Pattyn, S., Sillekens, P., Ieven, M., 2006. Detection of rhinoviruses by tissue culture and two independent amplification techniques, nucleic acid sequence-based amplification and reverse transcription-PCR, in children with acute respiratory infections during a winter season. J. Clin. Microbiol. 44, 166–171.

Mardy, S., Ly, S., Heng, S., Vong, S., Huch, C., Nora, C., Asgari, N., Miller, M., Bergeri, I., Rehmet, S., Veasna, D., Zhou, W., Kasai, T., Touch, S., Buchy, P., 2009. Influenza activity in Cambodia during 2006–2008. BMC Infect. Dis. 9, 168.

McCulloch, D.J., Sears, M.H., Jacob, J.T., Lyon, G.M., Burd, E.M., Caliendo, A.M., Hill, C.E., Nix, W.A., Oberste, M.S., Kraft, C.S., 2014. Severity of rhinovirus infection in hospitalized adults is unrelated to genotype. Am. J. Clin. Pathol. 142, 165–172.

McErlean, P., Shackelton, L.A., Lambert, S.B., Nissen, M.D., Sloots, T.P., Mackay, I.M., 2007. Characterisation of a newly identified human rhinovirus, HRV-QPM, discovered in infants with bronchiolitis. J. Clin. Virol. 39, 67–75.

McIntyre, C.L., Knowles, N.J., Simmonds, P., 2013. Proposals for the classification of human rhinovirus species A, B and C into genotypically assigned types. J. Gen. Virol. 94, 1791–1806.

McIntyre, C.L., McWilliam Leitch, E.C., Savolainen-Kopra, C., Hovi, T., Simmonds, P., 2010. Analysis of genetic diversity and sites of recombination in human rhinovirus species C. J. Virol. 84, 10297–10310.

Miller, E.K., Khuri-Bulos, N., Williams, J.V., Shehabi, A.A., Faouri, S., Al, Jundi I., Chen, Q., Heil, L., Mohamed, Y., Morin, L.L., Ali, A., Halasa, N.B., 2009a. Human rhinovirus C associated with wheezing in hospitalised children in the Middle East. J. Clin. Virol. 46, 85–89.

Miller, E.K., Edwards, K.M., Weinberg, G.A., Iwane, M.K., Griffin, M.R., Hall, C.B., Zhu, Y., Szilagyi, P.G., Morin, L.L., Heil, L.H., Lu, X., Williams, J.V., 2009b. A novel group of rhinoviruses is associated with asthma hospitalizations. J. Allergy Clin. Immunol. 123, 98–104.

Palmenberg, A.C., Rathe, J.A., Liggett, S.B., 2010. Analysis of the complete genome sequences of human rhinovirus. J. Allergy Clin. Immunol. 125, 1190–1199.

Palmenberg, A.C., Spiro, D., Kuzmickas, R., Wang, S., Djikeng, A., Rathe, J.A., Fraser-Liggett, C.M., Liggett, S.B., 2009. Sequencing and analyses of all known human rhinovirus genomes reveal structure and evolution. Science 324, 55–59.

Piotrowska, Z., Vázquez, M., Shapiro, E.D., Weibel, C., Ferguson, D., Landry, M.L., Kahn, J.S., 2009. Rhinoviruses are a major cause of wheezing and hospitalization in children less than 2 years of age. Pediatr. Infect. Dis. J. 28, 25–29.

Rathe, J., Liu, X., Tallon, L., Gern, J., Liggett, S., 2009. Full-genome sequence and analysis of a novel human rhinovirus strain within a divergent HRV-A clade. Arch. Virol. 155, 83–87.

Renwick, N., Schweiger, B., Kapoor, V., Liu, Z., Villari, J., Bullmann, R., Miething, R., Briese, T., Lipkin, W.I., 2007. A recently identified rhinovirus genotype is associated with severe respiratory-tract infection in children in Germany. J. Infect. Dis. 196, 754–1760.

Savolainen, C., Laine, P., Mulders, M.N., Hovi, T., 2004. Sequence analysis of human rhinoviruses in the RNA-dependent RNA polymerase coding region reveals large within-species variation. J. Gen. Virol. 85, 2271–2277.

Savolainen, C., Mulders, M.N., Hovi, T., 2002a. Phylogenetic analysis of rhinovirus isolates collected during successive epidemic seasons. Virus Res. 23, 41–46.

Savolainen, C., Blomqvist, S., Mulders, M.N., Hovi, T., 2002b. Genetic clustering of all 102 human rhinovirus prototype strains: serotype 87 is close to human enterovirus 70. J. Gen. Virol. 83, 333–340.

Simmonds, P., McIntyre, C.L., Savolainen-Kopra, C., Tapparel, C., Mackay, I.M., Hovi, T., 2010. Proposals for the classification of human rhinovirus species C into genotypically-assigned types. J. Gen. Virol. 91, 2409–2419.

Tapparel, C., Junier, T., Gerlach, D., Van-Belle, S., Turin, L., Cordey, S., Mühlemann, K., Regamey, N., Aubert, J.D., Soccal, P.M., Eigenmann, P., Zdobnov, E., Kaiser, L., 2009. New respiratory enterovirus and recombinant rhinoviruses among circulating picornaviruses. Emerg. Infect. Dis. 15, 719–726.

Vong, S., Guillard, B., Borand, L., Rammaert, B., Goyet, S., Te, V., Try, P.L., Hem, S., Rith, S., Ly, S., Cavailler, P., Mayaud, C., Buchy, P., 2013. Acute lower respiratory infections in ⩾5 year–old hospitalized patients in Cambodia, a low-income tropical country: clinical characteristics and pathogenic etiology. BMC Infect. Dis. 13, 97.

WHO, 2005. Pocket book of hospital care for children. <http://wwwwhoint/child_adolescent_health/documents/9241546700/en/indexhtml>.

Wisdom, A., Leitch, E.C.M., Gaunt, E., Harvala, H., Simmonds, P., 2009a. Screening respiratory samples for detection of human rhinoviruses (HRVs) and enteroviruses: comprehensive VP4-VP2 typing reveals high incidence and genetic diversity of HRV species C. J. Clin. Microbiol. 47, 3958–3967.

Wisdom, A., Kutkowska, A.E., McWilliam Leitch, E.C., Gaunt, E., Templeton, K., Harvala, H., Simmonds, P., 2009b. Genetics, recombination and clinical features of human rhinovirus species C (HRV-C) infections; interactions of HRV-C with other respiratory viruses. PLoS ONE 4, e8518.