Viral Load and Sequence Analysis Reveal the Symptom Severity, Diversity, and Transmission Clusters of Rhinovirus Infections

Kim Tien Ng, Xiang Yong Oong, Sin How Lim, Jack Bee Chook, Yutaka Takebe, Yoke Fun Chan, Kok Gan Chan, Nik Sherina Hanafi, Yong Kek Pang, Adeeba Kamarulzaman, and Kok Keng Tee

Background. Rhinovirus (RV) is one of the main viral etiologic agents of acute respiratory illnesses. Despite the heightened disease burden caused by RV, the viral factors that increase the severity of RV infection, the transmission pattern, and seasonality of RV infections remain unclear.

Methods. An observational study was conducted among 3935 patients presenting with acute upper respiratory illnesses in the ambulatory settings between 2012 and 2014.

Results. The VP4/VP2 gene was genotyped from all 976 RV-positive specimens, where the predominance of RV-A (49%) was observed, followed by RV-C (38%) and RV-B (13%). A significant regression in median nasopharyngeal viral load (VL) \( (P < .001) \) was observed, from 883 viral copies/µL at 1–2 days after symptom onset to 312 viral copies/µL at 3–4 days and 158 viral copies/µL at 5–7 days, before declining to 35 viral copies/µL at ≥8 days. In comparison with RV-A (median VL, 217 copies/µL) and RV-B (median VL, 275 copies/µL), RV-C–infected subjects produced higher VL (505 copies/µL; \( P < .001 \)). Importantly, higher RV VL (median, 348 copies/µL) was associated with more severe respiratory symptoms (Total Symptom Severity Score ≥17, \( P = .017 \)). A total of 83 phylogenetic-based transmission clusters were identified in the population. It was observed that the relative humidity was the strongest environmental predictor of RV seasonality in the tropical climate.

Conclusions. Our findings underline the role of VL in increasing disease attributed to RV-C infection, and unravel the factors that fuel the population transmission dynamics of RV.

Keywords. rhinovirus; acute respiratory tract infections; symptom severity; viral load; transmission clusters.

Rhinovirus (RV) is a predominant and ubiquitous airborne viral pathogen. With the improvement in viral detection methods, the involvement of RV in the lower respiratory compartment leading to severe and potentially fatal respiratory conditions is increasingly evident [1–3]. Furthermore, individuals with predisposing respiratory conditions such as asthma, chronic obstructive pulmonary disease, and cystic fibrosis may experience increased risk of severe RV-associated complications [4, 5].

There are 3 confirmed RV species (denoted as RV-A, RV-B, and RV-C) circulating worldwide. Despite the clinical burden of RV infections, large-scale molecular epidemiological data of RV are not extensively reported. Although the recent discovery of RV-C has incited renewed interest in investigating the epidemiology of RV infections [6], the effects of RV species on severity of respiratory illness remain insufficiently addressed. In an attempt to identify factors associated with RV morbidity and severity, studies have shown a possible correlation between high viral load (VL) and increased severity of RV infections [7]. However, different observations have also been reported elsewhere [8], potentially due to the variation in sample sizes and the inclusion of study subjects with predisposing conditions (asthma and pneumonia). Furthermore, the use of different VL quantification methods that inherit certain technical limitations (limited RV type coverage) may also affect the quantification efficiency [9].

Spatiotemporal analyses based on viral sequence data and evolutionary history of human immunodeficiency virus type 1 have shown that the emergence of transmission clusters is responsible for the spread of infections, highlighting the role of transmission clusters in escalating viral transmission and disease expansion [10]. The importance of such phylogenetically inferred transmission clusters in fueling the onward disease transmission has also been observed in other viral infections, such as in the recent Ebola virus outbreaks [11]. Despite the
high disease burden caused by RV, the evolutionary history and the dynamic of RV infections remains largely unexplored.

Climatological factors have also been implicated in the incidence of respiratory infections. For instance, findings from studies conducted in the temperate region have shown an association between high relative humidity and increased RV incidence [12]. However, studies of the meteorological factors and air pollutant on RV seasonality remain insufficiently explored in the regions with tropical climate [13].

In the present large-scale, population-based RV molecular epidemiological study, we studied the impact of RV species and the nasopharyngeal VL on the symptom severity of acute upper respiratory tract infections. Next, we investigated the genetic diversity, evolutionary histories, and the spatiotemporal dynamics of RV transmission clusters that drive disease transmission. Finally, we analyzed the potential meteorological predictors that influence the seasonality of RV in the context of the tropical climate in Southeast Asia.

METHODS

Study Subjects and Specimens
This study was approved by the University Malaya Medical Ethics Committee (reference number 890.1). Consenting outpatients who were presenting with symptoms of acute upper respiratory tract infections were recruited at the primary care clinics at the University of Malaya Medical Centre in Kuala Lumpur, Malaysia, between February 2012 and May 2014. Nasopharyngeal swabs were collected in universal transport medium using standardized technique. The presence of symptoms associated with acute respiratory tract infection was determined based on previously published criteria [14]. At the point of patient recruitment, the number of days after the onset of symptoms (symptomatic phase) was recorded. To assess the severity of acute respiratory tract infection associated with RV species, a previously described approach based on the Total Symptom Severity Score (TSSS) system was adopted [15], whereby higher score indicates greater severity of respiratory symptoms [15, 16].

Sequencing and Quantification of RV
Total viral RNA was extracted from 3935 nasopharyngeal specimens and screened for viral pathogens using the xTAG Respiratory Viral Panel FAST Assay (Luminex Molecular, Toronto, Canada). Specific enteroviruses were further confirmed through nested polymerase chain reaction amplification and direct sequencing of the VP4/VP2 gene [17]. The RV VL was quantified using a newly developed 1-step TaqMan assay, and VL was expressed in RV viral copies/μL of extracted RNA [18] (Supplementary Methods).

The categorical variables were compared using χ² test, while the differences and association between RV VL and disease severity (based on TSSS) were investigated through the non-parametric Mann-Whitney U test, Kruskal-Wallis test, linear regression, and multivariate analysis using SPSS. To improve clarity, the recorded number of days after the onset of symptoms (symptomatic phase) was grouped into subcategories, namely days 1–2, 3–4, 5–7, and ≥8, based on a previously described method [3].

Phylogenetic and Phylodynamic Analysis of RV
To determine the genetic types and to identify the possible transmission clusters of RV in the present study, neighbor-joining and Bayesian maximum clade credibility trees were reconstructed based on an updated and comprehensive list of global VP4/VP2 sequence data (3397 sequences). The time of most recent common ancestor of the respective transmission clusters observed in RV-A, RV-B, and RV-C was then estimated by the Bayesian coalescent-based relaxed molecular clock model, performed in BEAST 1.7 software (Supplementary Methods).

Meteorological Parameters and Their Associations With RV Cases
To understand the seasonality of RV infections, meteorological data collected from a weather station located within a 5-km radius from the hospital were obtained from the Malaysian Meteorological Department and were analyzed using Statistical Package for Social Sciences version 22.0 (SPSS Inc, Chicago, Illinois) (Supplementary Methods).

RESULTS

Distribution of RV Types in Patients With Acute Respiratory Tract Infection
A total of 3935 consenting outpatients (median age, 38 years; range, 7–95 years) with symptoms of acute respiratory tract infection were recruited, of whom 51% (2009/3935) were positive for at least 1 viral pathogen in the multiplex respiratory virus panel screening assay. Among 2009 subjects, 976 (49%) tested positive for RV, highlighting its high prevalence in individuals with symptoms of acute upper respiratory tract infection (Supplementary Table 1). The species and genetic types of the infecting RV were determined through neighbor-joining phylogenetic reconstruction (Figure 1A–C). Phylogenetic analysis of the VP4/VP2 gene revealed the predominance of RV-A, infecting 49% (473/976) of the subjects, followed by RV-C (38% [372/976]) and RV-B (13% [131/976]). The prevalence of RV-A and RV-C infections were consistently higher than RV-B throughout the study period (Figure 1D). In total, 111 distinct RV types (RV-A: 54 types, RV-B: 16 types, and RV-C: 41 types) were identified by phylogenetic analysis.

Clinical Characteristics of RV Infection, Viral Load Dynamics, and Association With Symptom Severity of Acute Respiratory Infections
To investigate the variation in clinical manifestations during acute respiratory infection, the clinical characteristics among subjects positive for RV were compared to those infected with other respiratory viruses (Table 1). Of the 976 RV-infected subjects, 129 subjects were excluded from analysis due to coinfection with at least 1 other respiratory virus or incomplete data. It
Figure 1. Phylogenetic transmission clusters of rhinovirus (RV)–A, RV-B, and RV-C VP4/VP2 gene. Neighbor-joining trees based on global RV VP4/VP2 sequence data (3397 sequences) are shown. Phylogeny reconstructions indicated that the 976 Malaysian strains were classified as RV-A (n = 473; A), RV-B (n = 131; B), and RV-C (n = 372; C). A total of 111 distinct RV types (or serotypes) were identified and indicated at the tips of the tree (marked with parentheses). In addition, potentially novel RV-A and RV-C types (asterisk) that were not classified with any previously defined RV types (cutoff of >10% nucleotide divergence in the partial VP4/VP2) were also observed at 1% (8/976) and 2% (15/976), respectively. Eighty-three transmission clusters (filled circles) were observed across RV-A (28 clusters, size range 2–9), RV-B (15 clusters, size range 2–11), and RV-C (40 clusters, size range 2–11). Newly sequenced RVs that do not form transmission clusters are indicated (hollow circles). The statistical significance of the branching order was validated by bootstrap analysis of 1000 replicates and the scale bar represents the nucleotide substitutions per site.

D. Bar chart illustrating the monthly distribution of RV-A, RV-B, and RV-C infections in Kuala Lumpur, Malaysia, between March 2012 and May 2014. Abbreviations: HRV, human rhinovirus; RV, rhinovirus.

Table 1. Clinical Manifestations in Patients Who Presented With Acute Respiratory Infections

<table>
<thead>
<tr>
<th>Signs and Symptoms</th>
<th>Positive for RV (n = 847)</th>
<th>Positive for Other Viruses (n = 962)</th>
<th>P Valueb</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sneezing</td>
<td>637 (75.2)</td>
<td>670 (69.6)</td>
<td>.008</td>
<td>1.3 (1.1–1.6)</td>
</tr>
<tr>
<td>Nasal discharge</td>
<td>696 (82.2)</td>
<td>732 (76.1)</td>
<td>.001</td>
<td>1.4 (1.2–1.8)</td>
</tr>
<tr>
<td>Nasal congestion</td>
<td>600 (70.8)</td>
<td>614 (63.8)</td>
<td>.001</td>
<td>1.4 (1.1–1.7)</td>
</tr>
<tr>
<td>Headache</td>
<td>538 (63.5)</td>
<td>634 (65.9)</td>
<td>.289</td>
<td>0.9 (0.8–1.1)</td>
</tr>
<tr>
<td>Sore throat</td>
<td>619 (73.1)</td>
<td>689 (71.6)</td>
<td>.489</td>
<td>1.1 (0.9–1.3)</td>
</tr>
<tr>
<td>Hoarse voice</td>
<td>668 (78.9)</td>
<td>790 (82.1)</td>
<td>.081</td>
<td>0.8 (0.6–1.0)</td>
</tr>
<tr>
<td>Muscle ache</td>
<td>488 (57.6)</td>
<td>678 (70.5)</td>
<td>&lt;.001</td>
<td>0.6 (0.5–0.7)</td>
</tr>
<tr>
<td>Cough</td>
<td>808 (95.4)</td>
<td>908 (94.4)</td>
<td>.289</td>
<td>1.3 (0.8–1.9)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; OR, odds ratio; RV, rhinovirus.

aCoinfection and cases with incomplete demographic and clinical data were excluded (n = 129).

bCalculated by χ² test. The difference is significant at the .01 level.
was observed that most of the RV-infected subjects experienced sneezing, nasal discharge, and nasal congestion, but fewer experienced muscle ache. No other conditions such as fever or chronic respiratory diseases (eg, asthma) were recorded. No differences in the clinical manifestations were observed between singly RV-infected and coinfected patients (n = 67) (Supplementary Table 2) [19, 20]. A significant negative correlation between RV VL and the estimated number of days from onset of symptoms was observed, with a correlation coefficient (r) of –0.121 (P < .001). A significant regression in median RV VL (P < .001) was observed, from 883 viral copies/µL at days 1–2 to 312 copies/µL at days 3–4 and 158 copies/µL at days 5–7, before declining further to 35 copies/µL at day ≥8 (Figure 2A and Supplementary Table 3). Similar trend was observed across RV species, though only RV-A (P = .047) and RV-C (P = .021) exhibited significant observation (Supplementary Table 4). Of note, subjects with respiratory symptoms for ≥8 days had detectable RV RNA.

Taking other covariates (eg, patient demographics) into consideration, multiple linear regression analysis was performed to assess the difference in VL between RV species at different symptomatic phases. At the species level, it was generally shown that subjects infected with RV-C had a significantly higher VL compared with those infected with RV-A and RV-B (Figure 2B and Supplementary Table 5). Such difference in VL was only evident at days 1–2 (P = .006). Interestingly, the multiple linear regression analysis revealed that patients with higher VL generally had higher TSSS (P = .017), indicating the increased severity of RV-associated acute respiratory tract infection (Figure 2C and Supplementary Table 6). Such association was profound at days 1–2 of the symptomatic phase (P = .012), which coincided with the peak VL. Together, analysis using multivariate analysis further indicated that patients infected with RV-C recorded higher VL and higher TSSS (Supplementary Table 7), suggesting that the increased symptom severity among RV-C–infected individuals could be attributed to the high VL.

Evolutionary Histories of RV Transmission Clusters
Phylogenetic reconstruction uncovered a total of 28 RV-A transmission clusters of varying sizes (2–9 subjects), predominantly involving RV-A32 (14% [4/28]). Similarly, a total of 15 and 40 transmission clusters (involving 2–11 subjects) was observed among RV-B and RV-C, respectively, with the predominance of RV-B79 (27% [4/15]), RV-B69 (27% [4/15]), and RV-C22 (10% [4/40]) (Figure 1 and Supplementary Figure 1). Based on Bayesian analyses, mean evolutionary rates of 3.21–3.24 (95%
confidence interval $3.04 - 3.41 \times 10^{-3}$ substitutions/site/year (across RV-A, RV-B, and RV-C) were estimated and were used to elucidate the divergence times of RV transmission clusters (Supplementary Figure 2).

**Seasonality of Acute Respiratory Illness Associated With RV Infection in the Tropical Region**

It was observed that the number of RV-infected cases peaked between October and December in 2012 and 2013, where the number of rain days, total amount of rainfall, and relative humidity were high (Figure 3). During the peak detection periods, the ground temperature and air pollutant ($\text{PM}_{10}$) readings were low. Statistical analysis showed that relative humidity seemed likely to be the main predictor for the increased number of RV infections, based on both partial ($r = 0.520, P = .011$) and bivariate ($r = 0.491, P = .009$) correlations as well as regression analyses (standard regression coefficient, $\beta = 1.329, t = 2.79, P = .011$) (Supplementary Table 8). It is important to note that relative humidity showed a significant positive correlation with the number of rain days ($r = 0.886, P < .001$) and total amount of rainfall ($r = 0.833, P < .001$), and a significant negative correlation with mean temperature ($r = -0.715, P < .001$) and $\text{PM}_{10}$ ($r = -0.700, P < .001$), underlining the indirect role of these meteorological factors and $\text{PM}_{10}$ on RV activity and seasonality.

**DISCUSSION**

Recent studies have investigated the potential role of VL in molding the dynamics of viral-associated acute respiratory infection [21, 22]. However, limitations that may hamper accurate VL quantification remain evident in many existing assays, such as the potential risk of VL misestimation and the

![Figure 3](https://academic.oup.com/cid/article-abstract/67/2/261/4828127/109x136)

**Figure 3.** Seasonality of rhinovirus (RV) infections and meteorological profiles in Kuala Lumpur, Malaysia, between March 2012 and May 2014. Bar and line charts illustrating the trends between RV incidence and meteorological factors are depicted. Meteorological data from February 2012 were omitted due to the incomplete sampling period. The number of RV-infected cases was higher during the months where the number of rain days, total amount of rainfall, and relative humidity were high, whereas the ground temperature and air pollutant ($\text{PM}_{10}$) readings were low. The increase or decrease (arrows) in meteorological readings that coincided with the annual peak of RV detection between October and December are indicated.
suboptimal sensitivity to detect broad array of RV types [23]. Here, a newly established TaqMan assay with a broader coverage for improved detection and quantification of RV VL was used [18]. From 847 RV-infected patients, RV VL peak was observed in 1–2 days after symptoms onset, in line with previously reported observations [24]. A significant regression in RV VL was observed thereafter, an indication of viral clearance by the immune system. Importantly, some patients had detectable VL 1 week after the onset of symptoms. Such observation clearly suggests the prolonged RV shedding in the respiratory tract [25], which may facilitate viral transmission during the second week of infection.

Several studies have attempted to examine the influence of RV species on VL during infection. For instance, study has demonstrated that in RV-C–infected patients with pneumonia, higher mean VL was reported [3]. However, insignificant difference in median peak VL between RV species was also reported in hospitalized patients elsewhere [26]. In the present analysis, which was established based on the large-scale RV molecular epidemiology and VL data, a significant association between RV species and VL was found, of whom RV-C–infected subjects exhibited higher nasopharyngeal VL in comparison to subjects infected with RV-A and RV-B. Such notable difference in VL between RV species could potentially due to, among others, the utilization of different cellular receptors for virus entry [27]. In comparison to RV-A and RV-B that utilize the intercellular adhesion molecule 1, RV-C uses the highly expressed cadherin-related family member 3 (CDHR3) as cellular receptor, in which a single-nucleotide polymorphism (CS29Y) in CDHR3 is associated with the upregulation of receptors on cell surface, promoting viral replication with a consequent increase in VL [28].

Several other studies have investigated the impact of VL on the virulence and severity of respiratory tract infection. For instance, it has been shown that higher VL is a risk factor for the development of respiratory complications such as lower respiratory infection, bronchial hyperreactivity, and respiratory failure, leading to prolonged hospitalization [21]. Here, statistical analysis revealed a significant correlation between higher VL and increased symptom severity, a manifestation that is associated with increased vascular permeability and stimulation of mucus hypersecretion during RV infection [29]. Several studies have also demonstrated the immunomechanism whereby an increased production of interleukin 10 following heightened RV replication leads to an attenuated type 1 T helper (Th1) immune response, resulting in an increased symptom severity of acute respiratory infection [29]. However, such finding should be interpreted with caution as the present study focused primarily on nasopharyngeal VL in outpatients (median age, 38 years) with upper respiratory tract symptoms who sought medical care, and may not be reflected in those with lower respiratory illnesses that required hospitalization. Importantly, the inclusion of age and time matched asymptomatic controls are necessary to avoid potential biases.

Studies have shown the ability of viral sequence data in defining transmission clusters, highlighting the importance and advantage of viral genetic information in assessing the epidemic linkages [30, 31]. Here, multiple transmission clusters across RV-A, RV-B, and RV-C species were observed (33% [320/976] of RV VP4/VP2 sequences), suggesting that the observed RV disease burden was largely linked to multiple subepidemics. To the best of our knowledge, the transmission clusters of RV were mapped for the first time at the population level, providing significant insights in understanding the dynamics of RV transmission. However, it is important to acknowledge that the actual number of transmission clusters circulating in the population could have been underestimated due to sampling bias from a single study site. Genealogical analysis estimated that most of the RV transmission clusters originated around 2010, highlighting their recent ancestral origin, though RV-C seemed to have a more diverse and older origin. Such observation suggested that RV-C might have emerged earlier, but went undetected due to the lack of a reliable detection system [6].

In a recent molecular surveillance study of 29 months, it was reported that >100 RV types were found circulating during the study period [32], and that the circulating RV types could change over time [33]. In the present study that spanned a period of 27 months, a total of 111 distinct RV types were identified, in which up to 14 distinct RV types were seen circulating concurrently in the study population within a given week (Supplementary Figure 3). However, it is important to note that the study subjects were recruited from a single study site, potentially leading to the underestimation of prevalence and distribution of circulating RV. In comparison to a study from Malaysia that detected 26 RV types in children presented with respiratory infections [34], a more diverse population of RV types was detected in adults in the present study. Although this may suggest that the adult population may play an important role in sustaining viral transmission and persistence in the general population, further investigation is necessary to test the hypothesis.

The impact of meteorological factors has been shown to correlate with incidence and seasonality of respiratory viruses [12]. For instance, temperate winters appeared to boost viral transmission by increasing the viral survivability in aerosols and on surfaces. However, the effects of air pollutants such as PM_{10} remain insufficiently reported. Malaysia has a tropical equatorial climate accompanied by the southwest monsoon (spans between May and September) and northeast monsoon (November–March) rainy seasons, of which the northeast monsoon brings more rainfall than does the southwest monsoon. As anticipated, the peaks of total rainfall, number of rain days, and relative humidity coincided with the Malaysian northeast monsoon. Statistical analysis revealed that relative humidity was the strongest predictor for RV infections, in
congruence with finding reported elsewhere [12]. It has been reported that RV is more stable and viable in conditions with high humidity, which extends the protective effect of droplets on viruses trapped on fomites or aerosols [12, 35]. Although the direct effects of other meteorological parameters and PM$_{10}$ on RV prevalence were not observed, it has been shown that relative humidity has a positive correlation with the number of rainy days and amount of rainfall, while exhibiting negative correlation with mean ground temperature and PM$_{10}$, suggesting a multifactorial contribution to the RV seasonality and incidence. To have a thorough assessment, other factors such as fine particulate matter and oxidant pollutant levels (nitrogen dioxide and trioxgen) [36], upon availability, as well as climate-induced changes in human behaviors (eg, staying indoors during the rainy season), should also be taken into consideration [37, 38].

Given the fact that RV is one of the most prevalent respiratory viruses, we believe that the burden of RV infections in Kuala Lumpur could be higher than documented. Also, since the proviso in analyzing evolutionary history more accurately relies on the depth of population-based sampling, a study of such nature should be continued and expanded to more recruitment centers in different countries to improve the resolution of RV genomic diversity and transmission dynamics in the region.

Nevertheless, our data reveal that RV contributed to nearly half of acute viral respiratory tract infections in adult outpatients. Remarkably, RV-C and high VL were shown to be the important determinants of the severity of acute respiratory illnesses. The phylogeny-based transmission clusters of RV were mapped for the first time, suggesting that the high RV disease burden in the population was largely linked to multiple subepidemics involving RV-A and RV-C. The detection of diverse RV types highlights the enormous genetic complexity and rapid evolution of circulating RV's that warrant continuous molecular surveillance at the population level. Finally, the seasonality of RV in the tropical Southeast Asia region was largely influenced by the relative humidity in the environment.

**Supplementary Data**

Supplementary materials are available at Clinical Infectious Diseases online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyrighted and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

**Notes**


*Acknowledgments.* We thank Ann C. Palmenberg (University of Wisconsin) for advice, and Yeon Kang Yong, Nyoke Pin Wong, Nur Ezreen Syafina, See We Teoh, and the Malaysian Meteorological Department for assistance and support. We also extend our gratitude to all individuals who agreed to participate in the study.

**Financial support.** This work was supported in part by grants from the Ministry of Higher Education, Malaysia: High Impact Research (grant number UML.C/625/1/HR/MOE/CHAN/02/02 to K. K. T.) and the Postgraduate Research Fund (PG097-2015A to K. T. N.).

**Potential conflicts of interest.** All authors: No potential conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

**References**