Infection, Genetics and Evolution 45 (2016) 311–319

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research Paper

Comparative genome and evolutionary analysis of naturally occurring Beilong virus in brown and black rats

Patrick C.Y. Woo [a,b,c,d,e,*,1], Annette Y.P. Wong [b,1], Beatrice H.L. Wong [b], Carol S.F. Lam [b], Rachel Y.Y. Fan [b], Susanna K.P. Lau [a,b,c,d,e,*], Kwok-Yung Yuen [a,b,c,d,e]

[a] State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong, China
[b] Department of Microbiology, The University of Hong Kong, Hong Kong, China
[c] Research Centre of Infection and Immunology, The University of Hong Kong, Hong Kong, China
[d] Carol Yu Centre for Infection, The University of Hong Kong, Hong Kong, China
[e] Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Zhejiang University, Hangzhou 310006, China

ARTICLE INFO

Article history:
Received 31 March 2016
Received in revised form 13 September 2016
Accepted 18 September 2016
Available online 20 September 2016

Keywords:
Animal RNA virus
Paramyxovirus
Rodent

ABSTRACT

Recently, we reported the presence of Beilong virus in spleen and kidney samples of brown rats and black rats, suggesting that these rodents could be natural reservoirs of Beilong virus. In this study, four genomes of Beilong virus from brown rats and black rats were sequenced. Similar to the Beilong virus genome sequenced from kidney mesangial cell line culture, those of J-virus from house mouse and Tailam virus from Sikkim rats, these four genomes from naturally occurring Beilong virus also contain the eight genes (3′-N-P/V/C-M-F-SH-TM-G-L-5′). In these four genomes, the attachment glycoprotein encoded by the G gene consists of 1046 amino acids; but for the original Beilong virus genome sequenced from kidney mesangial cell line, the G CDS was predicted to be prematurely terminated at position 2205 (TGG → TAG), resulting in a 734-amino-acid truncated G protein. This phenomenon of a lack of nonsense mutation in naturally occurring Beilong viruses was confirmed by sequencing this region of 15 additional rodent samples. Phylogenetic analyses showed that the cell line and naturally occurring Beilong viruses were closely clustered, without separation into subgroups. In addition, these viruses were further clustered with J-virus and Tailam virus, with high bootstrap supports of >90%, forming a distinct group in *Paramyxoviridae*. Brown rats and black rats are natural reservoirs of Beilong virus. Our results also supports that the recently proposed genus, *Jeilongvirus*, should encompass Beilong virus, J-virus and Tailam virus as members.

1. Introduction

Paramyxoviruses are enveloped, negative-stranded RNA viruses that are traditionally divided into two subfamilies, *Paramyxovirinae* and *Pneumovirinae*. In December 2015, these two subfamilies *Paramyxovirinae* and *Pneumovirinae* were elevated to family status as *Paramyxoviridae* and *Pneumoviridae* by the International Committee for Taxonomy of Viruses (http://talk.ictvonline.org/files/proposals/animal_dsrna_and_ssrna-_viruses/m/animal_rna_minus_ec_approved/5668.aspx). In the past decade, a number of paramyxoviruses in the original subfamily *Paramyxovirinae*, as well as their natural hosts, have been discovered. Examples include tuhoko virus 1, 2, and 3 from fruit bats (Lau et al. 2010), feline morbillivirus from domestic cats (Woo et al. 2012a), Sunshine virus from snakes (Hyndman et al. 2012), porcine parainfluenza virus 1 from swine (Lau et al. 2013), Mojiang paramyxovirus from rats (Wu et al. 2014) and Hawaiian cetacean morbillivirus from beaked whale (Jacob et al. 2016). Traditionally, there are five genera within the original subfamily *Paramyxovirinae*, namely *Respirovirus*, *Rubulavirus*, *Morbillivirus*, *Henipavirus* and *Avulavirus*. Recently, two new genera, *Aquaparamyxovirus*, which consists of Atlantic salmon paramyxovirus, and *Ferlavirus*, which consists of Fer-de-Lance virus discovered in the common lancehead snake and anaconda paramyxovirus discovered in green anaconda (Clark et al. 1979; Falk et al. 2008; ICTV 2012; Woo et al. 2014), were proposed. In addition, there are a number of paramyxoviruses which are still not officially classified into specific genera. These include Beilong virus, J-virus and Tailam virus, all discovered in rodents and probably belong to the same genus (Jack et al. 2005; Magoffin et al. 2007; Woo et al. 2011);

* Corresponding authors at: State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Hong Kong, China.
E-mail addresses: pcywoo@hku.hk (P.C.Y. Woo), skplau@hku.hk (S.K.P. Lau).
[1] PCY Woo and AYP Wong contributed the same to the manuscript.

http://dx.doi.org/10.1016/j.meegid.2016.09.016
1567-1348/

and Tupaia paramyxovirus discovered in the tree shrew (*Tupaia belangeri*) (Tidona et al. 1999).

In 2003, two putative novel "human cDNAs", named Angrem 104 and Angrem 52, were identified in an experiment screening for genes upregulated by angiotensin II, using a human kidney mesangial cell line (Li et al. 2006; Liang et al. 2003). Sequence analysis revealed that they were homologous to the matrix, fusion and phosphoprotein genes of paramyxoviruses, suggesting the possibility of a novel paramyxovirus (Basler et al. 2005; Schomacker et al. 2004). Interestingly, it was later found that similar sequences were not detectable in human kidney mesangial cell lines from other sources or in human kidney specimens, but were detectable in a rat kidney mesangial cell line used in the same laboratory prior to the acquisition of the human kidney mesangial cell line (Basler et al. 2005; Li et al. 2006). Isolation and characterization of the virus confirmed that it was a novel paramyxovirus, named Beilong virus, most closely related to J-virus discovered in kidney autoculture of moribund house mouse and our recently discovered Tailam virus from Sikkim rats (Li et al. 2006; Woo et al. 2011).

Since J-virus and Tailam virus were both from rodents and Beilong virus was amplifiable from a rat kidney mesangial cell line developed from Sprague Dawley rats (an outbred breed of albino rats originated from brown rats) (Jack et al. 2005; Li et al. 2006; Woo et al. 2011), we previously carried out a territory-wide molecular epidemiology study in rats to look for Beilong virus and its possible relatives (Woo et al. 2012b). Beilong viruses were detected in 40 kidney and nine spleen samples from 40 brown rats (*Rattus norvegicus*) and three black rats (*Rattus rattus*) (Woo et al. 2012b). Although the genome of Beilong virus isolated from kidney mesangial cell line has been sequenced and two unique additional open reading frames (ORFs) in the 3′ half of the attachment glycoprotein (G) gene, named X_1 and X_2, were identified, several important questions remain unanswered. First, are the viruses from brown rats and black rats the same virus infecting two different rat species or are they two different viruses adapting to different hosts? Second, are there differences between these naturally occurring viruses and the Beilong virus isolated from the kidney mesangial cell line? To answer these questions, we sequenced the complete or near complete genomes of four naturally occurring Beilong viruses from brown rats and black rats and performed comparative genomic analysis. Our results also support that the recently proposed genus, *Jeilongvirus*, should encompass Beilong virus, J-virus and Tailam virus as members.

2. Materials and methods

2.1. RNA extraction

Viral RNA was extracted from the spleen and kidney specimens using RNeasy Mini Spin Column (QIAgen, Hilden, Germany) (Lau et al. 2009; Lau et al. 2005; Lau et al. 2013). The RNA was eluted in 50 μl of RNase-free water and was used as the template for RT-PCR.

2.2. Genome sequencing

Four complete or near complete genomes of the Beilong viruses discovered from brown rats and black rats were amplified and sequenced using the RNA extracted directly from the spleen (strains ERN081008_1S and YLRR120908_3S) and kidney (strains HKIARN060309_1K and YLRR120908_1K) samples as templates. The RNA was converted to cDNA by a combined random-priming and sequence-specific priming strategy. The cDNA was amplified by degenerate primers designed by multiple alignments of the genomes of Beilong virus from kidney mesangial cell line culture and other closely related paramyxoviruses with complete genomes available, using strategies described in our previous publications (Lau et al. 2010; Woo et al. 2012a; Woo et al. 2011). Additional primers were designed from the results of the first and subsequent rounds of sequencing.

For each PCR, the PCR mixture (25 μl) contained cDNA, 5 μl 5× iProof™ HPLC HF Buffer, 200 μM of each dNTPs and 0.5 U iProof™ High-Fidelity DNA Polymerase (Bio-Rad, Hercules, CA, USA). The mixtures were amplified in 35 cycles of 98 °C for 10 s, 55 °C for 30 s and 72 °C for 2 min and a final extension at 72 °C for 10 min in an automated thermal cycler (Applied Biosystem, Foster City, CA, USA). Standard precautions were taken to avoid PCR contamination and no false-positive was observed in negative controls.

The PCR products were gel-purified using the QIAquick gel extraction kit (QIAgen, Hilden, Germany). Both strands of the PCR products were sequenced twice with Prism 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA), using the two PCR primers.

The 5′ ends of the viral genomes were confirmed by rapid amplification of cDNA ends using the SMARTer™ RACE cDNA Amplification Kit (Clontech, USA). Sequences were assembled and manually edited to produce final sequences of the viral genomes.

2.3. Genome analysis

The genome organizations of the four Beilong viruses were predicted based on the annotated Beilong virus genome from kidney mesangial cell line culture (Li et al. 2006) and were verified with ORF finder on NCBI server. Comparison of amino acid identities between the genomes of Beilong viruses from brown and black rats and those of other paramyxoviruses was performed using MatGAT (Campanella et al. 2003). Comparison of the nucleotide and amino acid sequences of the various ORFs between the Beilong viruses from brown and black rats and Beilong virus from kidney mesangial cell line culture was performed by multiple alignment using MUSCLE embedded in Mega 6.06 (Edgar 2004; Tamura et al. 2013).

2.4. Phylogenetic analysis

Phylogenetic trees were constructed using maximum likelihood method by PhyML 3.0 (Guindon et al. 2010) with substitution models selected according to the results of ModelGenerator (Keane et al. 2006). 1755, 2748, 1179, 1785, 3885 and 7224 bases of the nucleocapsid (N), phophoprotein (P), matrix protein (M), fusion protein (F), attachment protein (A) and large polymearase (L) genes were used for phylogenetic analysis. Nucleotide sequences were aligned by codon using MUSCLE embedded in Mega 6.06.

2.5. Genetic distance analysis

The SimPlot program (version 3.5.1) (Lole et al. 1999) was used to analyze the genetic distance of genes in Beilong virus, Tailam virus and J-virus, using the sequence of ERN081008_1S as reference. Concatenated alignment of N, P, M, F, small hydrophobic protein (SH), transmembrane protein (TM), G and L genes was used. The large gap region at the end of the G gene alignment was excluded from the analysis. The similarity plot was configured with window size: 200 nucleotides; step: 20; F84 (maximum likelihood); GapStrip: on (50%).

2.6. Partial G gene sequencing

To ascertain the difference between the naturally occurring Beilong viruses from brown and black rats and the one from kidney mesangial cell line, 521 bases of the G genes in the Beilong viruses from 15 additional positive kidney and spleen samples of brown rats and black rats were amplified and sequenced. Reverse transcription, PCR and DNA