Infection, Genetics and Evolution 39 (2016) 85–91

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Phylogenetic and molecular analyses of human parainfluenza type 3 virus in Buenos Aires, Argentina, between 2009 and 2013: The emergence of new genetic lineages

Stephanie Goya [a,b], Alicia Susana Mistchenko [a,c], Mariana Viegas [a,b,*]

[a] Virology Laboratory, Ricardo Gutierrez Children's Hospital, Buenos Aires City, Argentina, Gallo 1330 2°, (1425) Ciudad Autónoma Buenos Aires, Argentina
[b] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina
[c] Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC), Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 5 October 2015
Received in revised form 18 December 2015
Accepted 4 January 2016
Available online 9 January 2016

Keywords:
Human parainfluenza virus 3
Genotypes
Seasonality
Emerging lineages
Molecular characterization
B-cell epitopes

ABSTRACT

Despite that human parainfluenza type 3 viruses (HPIV3) are one of the leading causes of acute lower respiratory tract infections in children under five, there is no licensed vaccine and there is limited current information on the molecular characteristics of regional and global circulating strains. The aim of this study was to describe the molecular characterization of HPIV3 circulating in Buenos Aires. We performed a genetic and phylogenetic analysis of the HN glycoprotein gene. Between 2009 and 2013, 124 HPIV3-positive samples taken from hospitalized pediatric patients were analyzed. Four new genetic lineages were described. Among them, C1c and C3d lineages showed local circulation patterns, whereas C3e and C3f comprised sequences from very distant countries. Despite the diversity of the described genotypes, C3a and C3d predominated over the others, the latter was present during the first years of the study and it was progressively replaced by C3a. Molecular analyses showed 28 non-synonymous substitutions; of these, 13 were located in potentially predicted B-cell epitopes. Taken together, the emergence of genetic lineages and the information of the molecular characteristics of HN protein may contribute to the general knowledge of HPIV3 molecular epidemiology for future vaccine development and antiviral therapies.

1. Introduction

Human parainfluenza viruses (HPIV) are one of the leading causes of acute lower respiratory tract infections (LRTI) in children less than 5 years old, immunocompromised or chronically ill patients, and the elderly; their incidence is surpassed only by human respiratory syncytial virus (RSV) (Henrickson, 2003).

HPIVs belong to the family *Paramyxoviridae* and they have been divided into types 1 to 4, among which HPIV1 and HPIV3 are classified as members of the genus *Respirovirus*, while HPIV2 and HPIV4 are members of the genus *Rubulavirus* (Karron and Collins, 2013).

HPIV3 is the most frequent type of HPIV (Liu et al., 2013) and is associated with the most severe clinical presentations (bronchiolitis and pneumonia) (Henrickson, 2003; Karron and Collins, 2013); it is one of the leading causes of hospitalization in pediatric patients during spring and summer, with a high impact on public health (Ministerio de Salud. Boletín Integrado de Vigilancia N199. Argentina, 2013; Mao et al., 2012; Viegas et al., 2004). HPIV3 is an enveloped, single-stranded negative sense RNA virus. Its genome has 15,462 nucleotides and it encodes eight proteins, among them the fusion protein (F) and the hemagglutinin–neuraminidase protein (HN) are associated with the viral envelope (Henrickson, 2003; Karron and Collins, 2013).

The HN protein is a transmembrane glycoprotein of 572 aminoacids. It consists of a cytoplasmic domain, a membrane spanning region, a stalk region, and a globular head, and it is responsible for the binding and cleavage to the host cell sialic acid, stabilization, and activation of the F protein.The stalk region and the globular head protrude the virus envelope (ectodomain) (Karron and Collins, 2013; Palermo et al., 2009).

The HN and F proteins are the major target antigens of humoral immune response, inducing neutralizing antibodies, and the HN glycoprotein has the largest antigenic and genetic variation (Schmidt et al., 2011; Spriggs et al., 1987). Accordingly, the HN gene has been used for genotyping HPIV3 strains in molecular epidemiological studies (Mao et al., 2012; van Wyke Coelingh et al., 1987; Villaran et al., 2014; Almajhdi, 2015). In 2012, Mao et al. carried out the first study in order to unify and define a genetic classification for HPIV3, and it was recently reclassified by Almajhdi (Mao et al., 2012; Almajhdi, 2015). In these phylogenetic studies using HN gene and based on evolutionary divergence values (calculated as genetic distances), they described the existence of clusters (A, B, and C), subclusters (C1–C5), and genetic lineages (C1a–C1b, C3a–C3c) among HPIV3 strains. The authors

* Corresponding author at: Gallo 1330 2°, (1425) Ciudad Autónoma Buenos Aires, Argentina.
E-mail addresses: goyastephanie@gmail.com (S. Goya), asmistchenko@hotmail.com (A.S. Mistchenko), viegasmariana@conicet.gov.ar (M. Viegas).

http://dx.doi.org/10.1016/j.meegid.2016.01.002
1567-1348/

determined that the minimum genetic distance that must exist between two sequences of HPIV3 for classification in clusters is 0.045 and differences in the range of 0.019–0.045 define subclusters. Sequences with genetic distance under 0.010 belong to the same genetic lineage, and those with genetic distance between 0.010 and 0.019 belong to the same subcluster but to different genetic lineages (Almajhdi, 2015).

In Argentina, HPIV3 infections are the second leading cause of hospitalization for acute LRTI in patients less than 2 years old (Ministerio de Salud. Boletín Integrado de Vigilancia N199. Argentina, 2013). However, while the universal influenza vaccination in children is recommended, there is no a licensed vaccine against HPIV3 (Weinberg et al., 2009). Nevertheless, in recent years, intranasal live attenuated vaccines were generated by a reverse genetic system; they are well tolerated and immunogenic, and Phase I clinical trials with these vaccines are underway (Schmidt et al., 2011).

Taking into account that there is limited current information on the molecular characteristics of regional and global circulating HPIV3 strains, the aim of this study is to describe the molecular characterization of HPIV3 circulating in Buenos Aires, in a period of 5 years (2009–2013) through the genetic and phylogenetic analysis of the HN glycoprotein gene.

2. Materials and methods

2.1. Clinical samples

Nasopharyngeal aspirates (NPA) were taken from pediatric patients hospitalized in neonatology and pediatric wards of Public Hospitals from Buenos Aires city and Greater Buenos Aires and they were referred to the Virology Laboratory of the Ricardo Gutiérrez Children's Hospital (HNRG) for viral diagnosis. The period studied included 5 consecutive years (2009–2013).

A rapid detection of respiratory viruses by immunofluorescence assay (IFA) with monoclonal antibodies against HPIV 1, 2, and 3, human RSV, human adenovirus (AdV), and influenza A and B (Flu A and B, respectively) was performed according to manufacturer's instructions (EMD Millipore, Darmstadt, Germany) (Gardner and McQuilin, 1968).

Samples that were positive for HPIV3 by IFA were preserved at −70 °C until the molecular study was performed. All samples molecularly analyzed in this study belonged to children hospitalized at the HNRG, and they were randomly selected from the total HNRG samples. The samples and sequences were coded using the acronym HNRG, a serial number, country of origin and the year in which they were obtained.

2.2. RNA extraction and RT-PCR amplification of partial HN gene

Total RNA was extracted from 200 μl of NPA samples using PureLink® Viral RNA/DNA Mini Kit (Invitrogen, Massachusetts, USA) according to the manufacturer's instructions. RNA was eluted in 50 μl of RNase and DNase-free distilled water, and stored at −70 °C.

The partial HN gene (globular head) was amplified with the OneStep RT-PCR Kit (Qiagen, Venlo, Netherlands) according to the manufacturer's instructions and using the following previously published primers: HPIV3-HN.f1: 5′-ATGATCTAATACARTCAGGAGTRAATACAAG-3′ (nt 7,068–7,099, sense) and HPIV3-HN.r1: 5′-TATCTCGAGTTATGATTAACTG CAGC-3′ (nt 8,514–8,540, anti-sense) at a final concentration of 0.6 μM (Mao et al., 2012; Mizuta et al., 2014). The reaction mix was carried out in a 25 μl of reaction volume containing 7 μl of total RNA. Thermal cycling conditions were 50 °C for 30 min and then 95 °C for 15 min followed by 35 cycles of 94 °C for 30 seg, 50 °C for 30 seg, and 72 °C for 96 seg, and a final extension step at 72 °C for 10 min. The amplification products were 1.5% agarose gel purified with Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Irvine, California), and the product was eluted in 15 μl of RNase and DNase-free distilled water.

2.3. DNA sequencing

The purified PCR products were labeled with the BigDye Terminator v3.1 sequencing kit (Applied Biosystems, Massachusetts, USA) and electrophoresis was performed in an ABI3500 genetic analyzer (Applied Biosystems, Massachusetts, USA). The sequencing reactions were performed with the PCR primers and the internal primers: HPIV-HN.f2: 5′-AACTGTGTTCAACTCCHAAAG-3′ (nt 7,605–7,625, sense), HPIV-HN.f3: 5′-CAAGTTGGCAYAGCAAGTTAC-3′ (nt 8,082-8,102 sense), HPIV-HN.r2: 5′-CTGAATTGTAAGAAGCCTTGT-3′ (nt 7,094–7,114, anti-sense) (Mao et al., 2012), and HPIV-HN.r3: 5′-ATCTTGTT GTYGAGATTGAGCCA-3′ (nt 7,715–7,738, anti-sense), HPIV-HN.r4: 5′-TCAATTATTCCTAATTGTAACTTGCT-3′ (nt 8,093–8,119, anti-sense) designed for this study.

The SeqScape Software v2.7 (Applied Biosystems, Massachusetts, USA) was used to analyze, assemble, and generate the consensus nucleotide sequences obtained for each of the analyzed viral strains.

2.4. Phylogenetic analyses

The HNRG sequences were aligned with 50 previously published HPIV3 sequences downloaded from GenBank with Clustal W v2 software (Thompson et al., 1994). The jModelTest v0.1.1 software was used to determine the most suitable evolutionary model for the set of the analyzed sequences (Darriba et al., 2012).

The phylogenetic analyses were performed by ML (PhyML v.3.1 software, Guindon et al., 2010), Bayesian criteria (MrBayes v.3.2.1 software, Ronquist and Huelsenbeck, 2003), and distance methods (Neighbor-Joining, MEGA v6 software, Tamura et al., 2013). Initial random trees were used to infer the ML tree. The branch support for ML and distance methods were evaluated by non-parametric bootstrapping with 1000 pseudo-replica. The convergence of the Monte Carlo Markov Chains (MCMC) implemented in the Bayesian criteria was evaluated in TRACER v.1.5 with an effective sample size (ESS) >200; the initial 10% of the run length was discarded as burn-in. Consensus trees were visualized with FigTree v.1.4.0.

2.5. Molecular characterization and adaptive evolutionary analyses

The MEGA v6 software was used to define genetic lineages by estimating genetic distances (p-distances) within and among sequences in phylogenetic clades. Standard error (SE) estimates were obtained by the bootstrap method (1000 replicates) (Tamura et al., 2013).

The DNAsp v5 software was used for the analysis of the polymorphisms and divergences between the set of HNRG sequences and the prototype HPIV3 strain Washington/1957 (GenBank accession number: JN089924) (Librado and Rozas, 2009). The amino acid sequences were inferred using the universal genetic code.

Natural selection on the HNRG HN sequences was estimated from the ratio of non-synonymous (dN) to synonymous (dS) substitutions per site (dN/dS) at every codon in the alignment and the overall $\Omega = dN/dS$. The analysis was performed using the procedures available in the HyPhy package and accessed through the Datamonkey web server (Kosakovsky Pond and Frost, 2005).

N-glycosylation sites were predicted using N-Glycosite server (Zhang et al., 2004).

Potential conformational B-cell epitopes were mapped from the molecular model of the HN ectodomain region of the reference strain (PDB: 4MZA) with “DiscoTope: Structure-based antibody prediction” tool (IEDB Analysis Resource server) with a specificity of 75% and a sensitivity of 47% (Haste Andersen et al., 2006). The PDBePISA service of the EMBL-EBI website was used to identify sites on the surface of the protein (Krissinel and Henrick, 2007). Potential linear B-cell epitopes were predicted with “Bepipred: Linear Epitope Prediction” tool (IEDB Analysis Resource server) with a specificity of 75% and a sensitivity of 49% (Larsen et al., 2006).