Genetic analysis of human parainfluenza virus type 3 obtained in Croatia, 2011–2015
Tanja Košutić-Gulija,1,2* Anamarija Slovic,1,2 Sunčanica Ljubin-Sternak,3 Gordana Minarić-Galinović4 and Dubravko Forčić1,2

Abstract
Purpose. This study investigated the HPIV3 circulating strains in Croatia and whether the other parts of HPIV3 genome (F gene and HN 582 nucleotides fragment) could be equally suitable for genetic and phylogenetic analysis.

Methodology. Clinical materials were collected in period 2011–2015 from children suffering from respiratory illnesses. In positive HPIV3 samples viral genome was partially amplified and sequenced for HN and F genes. Obtained sequences were analysed by phylogenetic analysis and genetic characterization was performed.

Results. All samples from this study belonged to subcluster C and over a short period of time, genetic lineage C3a gained prevalence over the other C genetic lineages, from 39% in 2011 to more than 90% in 2013 and 2014. Phylogenetic classifications of HPIV3 based on the entire HN gene, HN 582 nt fragment and entire fusion (F) gene showed identical classification results for Croatian strains and the reference strains. Molecular analysis of the F and HN glycoproteins, showed their similar nucleotide diversity (Fcds \( P=0.0244 \) and HNcds \( P=0.0231 \)) and similar Ka/Ks ratios (F Ka/Ks=0.0553 and HN Ka/Ks=0.0428). Potential N-glycosylation sites, cysteine residues and antigenic sites are generally strongly conserved in HPIV3 glycoproteins from both our and the reference samples.

Conclusion. The HPIV3 subcluster C (genetic lineage C3a) became the most detected circulating HPIV3 strain in Croatia. The results indicated that the HN 582 nt and the entire F gene sequences were as good for phylogenetic analysis as the entire HN gene sequence.

INTRODUCTION
The human parainfluenza virus type 3 (HPIV3) is a causative agent of acute respiratory infection that can proceed to more severe illnesses such as bronchiolitis, pneumonia or croup. Infants and young children are especially susceptible to severe illnesses that frequently require hospital treatment. HPIV3 is a second leading cause of respiratory illness, after respiratory syncytial virus, requiring hospital treatment in infants and young children [1–4]. Although several efforts have been made to produce a vaccine for HPIV3 [5–7], no licensed vaccine is available at present.

HPIV3 is one of the four distinct serotypes of HPIV. All HPIVs are members of the family Paramyxoviridae, the subfamily Paramyxovirinae, genera Respirovirus (HPIV1 and HPIV3) and Rubulavirus (HPIV2 and HPIV4). The HPIV3 genome is a negative single-stranded RNA of 15 462 nucleotides (nt), composed of six genes [nucleo-protein (NP)–phosphoprotein (P)–matrix (M)–fusion (F)–hemagglutinin-neuraminidase (HN)–large (L)] that code for eight proteins (NP, P/D/C, M, F, HN and L proteins) [8]. Viral infection begins with the attachment of viral glycoprotein hemagglutinin-neuraminidase (HN) to the cell-surface receptor with α2,6 or α2,3 sialic acids [9], followed by the
fusion of the viral and cell membranes. The fusion process is directed by a conformational changes of the fusion glycoprotein (F) and requires an interaction of HN and F glycoproteins [10].

Both glycoproteins have been investigated extensively over the past three decades. Therefore, their binding sites [11], functional domains [12, 13], crystal structures [10, 14–16] and antigenic sites [17, 18] are well known.

In general, the molecular epidemiology studies of HPIV3 were based on the HN gene sequences [1, 2, 19–23], while there is limited information regarding the molecular epidemiology of HPIV3 based on the F gene [23, 24]. Consequently, genetic classification of HPIV3 is based on the HN gene [2, 22] and consists of three clusters (A, B, C), with cluster C further divided into subclusters (C1–C5) and genetic lineages in C1 and C3 subclusters. Either the entire HN gene (1719 nt) [2, 22, 23] or a portion of the HN gene (1146 nt) is used in the genetic classification [1].

The aim of this study was to investigate the molecular characteristics of HPIV3 that circulated in Croatia, Europe, during a period from 2011 to 2015. The investigation was conducted through genetic and phylogenetic analysis of the HN and F genes.

METHODS

Clinical samples
This study included 275 nasopharyngeal aspirates (NPA) that were positive for HPIV3 antigen in direct fluorescent antibody test (DFA) (DFA Light Diagnostics, Chemicon International, Temecula, CA). Clinical samples were taken from paediatric patients with acute respiratory infection, in hospitals from the Zagreb area over five consecutive years (2011–2015). Samples were stored in sterile minimal medium at −20 or −80 °C until molecular analysis. This study was approved by the ethics committees of University Hospital for Infectious Diseases 'Dr Fran Mihaljevic', as part of a scientific project number 6255, fully supported by the Croatian Science Foundation.

RNA extraction
Of 275 samples collected, 166 were chosen for molecular analysis based on their collection date and storage history. Viral RNA was extracted from 250 µl of NPA specimen as reported previously by Chomczynski and Sacchi [25]. RNA pellets were resuspended in 10 µl sterile distilled water and subjected to reverse transcription.

RT-PCR of the partial HN gene
The cDNA was prepared from total RNA using random hexamers and MuLV reverse transcriptase as reported previously [26]. Amplification was carried out with 10 µl cDNA, using OneTaq DNA Polymerase (New England Biolabs) according to manufacturer’s instructions. To amplify a region of the HN gene (7494–8697 nt), primers HPIV3 HN2 7494+ and HPIV3 HN2 8697− were used. The cycling conditions were as follows: initial denaturation at 94 °C for 5 min, followed by 10 cycles of amplification (94 °C for 30 s, 54 °C for 30 s and 68 °C for 1 min 20 s), followed by 30 cycles of amplification (94 °C for 30 s, 51 °C for 30 s and 68 °C for 1 min 20 s) and a final extension of 7 min at 68 °C. The PCR products were analysed on 1.0 % agarose gel and amplified products were excised and purified. All primers used in this study are specified in Table S1 (available in the online Supplementary Material).

RT-nested PCR of entire HN and F genes
The entire HN and F genes were amplified for 22 NPA samples that were selected according to their distribution on the phylogenetic tree of the partial HN gene. cDNA was prepared from total RNA using random hexamers and MuLV reverse transcriptase as reported previously [26]. Both genes were amplified from the same cDNA mixture, which was split in two with OneTaq DNA polymerase used in both reactions. Primer pairs for the HN gene used for the first PCR were HPIV3 HN 6531+ and HPIV3 HN 8702−, with the following PCR conditions: initial denaturation at 94 °C for 5 min, followed by 10 cycles of amplification (94 °C for 30 s, 52 °C for 30 s and 68 °C for 2 min 20 s), followed by 30 cycles of amplification (94 °C for 30 s, 50 °C for 30 s and 68 °C for 1 min 20 s) and a final extension of 7 min at 68 °C.

Five microlitres of the first PCR amplicon was used in two nested reactions, with primer pairs HPIV3 HN1 6662mut+/HPIV3 HN1 7665− and HPIV3 HN2 7494+/HPIV3 HN2 8697−. The cycling conditions for the second-round PCR were identical to conditions used to amplify the partial HN gene.

For amplification of the entire F gene, the HPIV3 F 4737+ and HPIV3 F 4747+/HPIV3 F 5922− were used in two nested reactions with primer pairs: HPIV3 F 4747+/HPIV3 F 5641+/HPIV3 F 5687−. The PCR conditions for the second-round PCR were as follows: initial denaturation at 94 °C for 5 min, followed by 40 cycles of amplification (94 °C for 30 s, 54 °C for 30 s and 68 °C for 2 min 20 s) and a final extension of 7 min at 68 °C. For the second round PCR, 5 µl of amplicon was used in two nested reactions with primer pairs: HPIV3 F 4747+/HPIV3 F 5922− and HPIV3 F 5641+/HPIV3 F 5687−. Cycling conditions for the second round PCR were: initial denaturation at 94 °C for 5 min, followed by 40 cycles of amplification (94 °C for 30 s, 55 °C for 30 s and 68 °C for 1 min 20 s) and a final extension of 7 min at 68 °C.

The PCR products from the second PCR for HN and F were analysed on 1.0 % agarose gel and amplified products were excised and purified.

Nucleotide sequencing
Purified PCR products were sequenced with a Big Dye Terminator v3.1 Cycle Sequencing kit using primer HPIV3 HN2 7494+ for the partial HN gene, while the complete HN and F nested products were sequenced with primers used in the second round PCR. The nucleotide sequences were determined with an automated DNA sequencer ABI PRISM 3130 Genetic Analyser (Applied Biosystems, Foster City, CA).
Phylogenetic analysis

The F or HN genes sequences obtained in this study were aligned with HPIV3 sequences downloaded from the GenBank database which were previously used in molecular epidemiology studies [1, 2, 22]. All alignments were performed using CLUSTAL X 2.1 software [27].

Selection of the most suitable substitution model was determined with jModelTest 2.1.4 software [28]. Maximum likelihood phylogenetic trees were generated with Molecular Evolutionary Genetics Analyses (MEGA) software, version 6.06 [29]. Bootstrap probabilities for 1000 iterations were calculated to evaluate confidence estimates. Bayesian Markov Chain Monte Carlo (MCMC) inference was performed with BEAST v1.8.2 [30]. Convergence was assessed based on the effective sample size using Tracer v1.5 (http://beast.bio.ed.ac.uk/Tracer) after a 10% burn-in, and only values above 200 were accepted. Maximum clade credibility trees were generated with TreeAnnotator v1.8.2 and visualized with FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree).

Molecular characterization of glycoproteins HN and F

Software DnaSP version 5.0 [31] was applied for analysis of nucleotide sequence data variability. Deduced amino acid sequences of the HN and F genes were generated by translating nucleotide sequences with the standard genetic code using MEGA 6.06 software. Potential N-glycosylated sites in the F and HN protein sequences were predicted using the NetNGlyc 1.0 Server [32].

Nucleotide sequences accession numbers

Twenty-two sequences of the complete F and HN genes and 60 partial sequences of the HN gene have been deposited in GenBank under accession numbers KX467885 to KX467906 and KX467907 to KX467966, respectively. The HPIV3 reference strains used in this study are specified in Table S2.

RESULTS

Patient characteristics

275 NPA samples from the Zagreb region positive for HPIV3 antigen using the DFA test were analysed in this study. The patients were infants or young children that attended hospital because of respiratory illnesses.

The majority (79.2%) of the children were younger than 2 years of age, with 14.9% aged 2–5 years and 5.8% older than 5 years of age.

Among the enrolled children, the most common clinical diagnosis was upper respiratory tract infections (URIs) (59.6%), followed by bronchiolitis (13.8%), pneumonia (12.4%) and laryngitis (croup) (8.0%).

The collection period was from January 2011 to December 2015. During this period, the peaks of the epidemic seasons were in late spring (May) and autumn (October) (Fig. 1).

![Fig. 1. The number of total HPIV3 samples collected from January 2011 to December 2015 in Croatia and subclusters detected in this study (C1, C2, C3 and C5). Results are shown for each month of five calendar years [from left to right, January (J), February (F), March (M), April (A), May (M), June (J), July (J), August (A), September (S), November (N) and December (D)].](image-url)
For 2012, 2014 and 2015, two epidemic seasons per year were noticed, while in 2011 and 2013 only one epidemic season occurred, in spring and autumn, respectively (Fig. 1).

**Phylogenetic analysis of the partial and entire HN gene**

Of the 166 samples analysed for the partial 582 nt HN fragment, 14 were PCR-negative and of the remaining 152 positive samples (91.56%) 141 were used for further analysis. The partial HN sequences showed great diversity, with 81 sequences being unique and submitted to GenBank. The remaining 60 samples had sequences identical to some of the sequences submitted to GenBank, as annotated in brackets in the HN 582 nt phylogenetic tree (Fig. 2).

The phylogenetic analysis was performed on the 582 nt fragments from 42 reference sequences of the known genotypes [1, 2, 22], along with the 81 sequences obtained in this study. The topology of the tree, as well as calculated genetic distances, were consistent with the regular reference topology for separation of clusters and subclusters, in spite of the fact that a shorter fragment of the HN gene was used (Fig. 2).

All samples from this study belonged to cluster C, and were distributed in subclusters C1, C2, C3 and C5 (Table 1). The majority of samples belonged to subcluster C3. Temporal distribution demonstrated co-circulation of subclusters C1, C3 and C5 during the examined period, while subcluster C2 was detected only in 2011 (one sample) (Table 1). The largest number of different subclusters was recorded in 2011 with C1, C3, C5 and C2 subclusters represented at 44, 39, 11 and 5% respectively. In the following years, C3 subcluster gained prevalence over C1 and C5, with 59% of the samples in 2012 belonging to this subcluster, more than 90% in 2013 and 2014, and 82% in 2015.

In the phylogenetic tree based on the 582 nt of the HN gene, the topology of the samples suggests the existence of genetic lineages in C1 and C3 subclusters (Fig. 2). Goya et al. [1] determined that genetic distances in the range 0.010–0.019 indicated the presence of genetic lineages. Following this criterion, we confirmed that samples from this study belong to C1a, C1b, C3a, C3b, C3d and C3e genetic lineages, with the majority of the samples in lineage C3a (Table S3).

To confirm these results, we selected 22 samples from this study with different topologic placements on the 582 nt HN tree and determined their complete HN gene coding sequences (1716 nt). For the reconstruction of the tree based on the entire HN gene, 32 reference sequences were included, along with Croatian samples (Fig. S1).

The topology of clusters, subclusters and genetic lineages confirmed the existence of genetic lineages present in the partial HN gene tree for the majority of our sequences and all of the reference sequences. These results were also confirmed with the calculated genetic distances ($P$) (Tables S3 and S4a). Phylogenetic reconstructions for the entire HN and partial 582 nt HN were generated with two methods, and similar tree topologies and statistical support were obtained.

One sample obtained in this study, HR/20.12 (2547), KX467892 had a different topology depending on whether the 582 nt or the entire HN gene was used for the reconstruction of the phylogenetic tree. When only the partial 582 nt HN sequence was used, sample HR/20.12 (2547), KX467892 belonged to the C5 subcluster (Fig. 2), while it clustered in genetic lineage C3a when the entire HN sequence was used (Fig. S1).

Phylogenetic and genetic distance comparison of 582 nt and 1716 nt HN gene sequences confirm that the 582 nt HN gene region is appropriate for fast phylogenetic screening of HPIV3 clinical samples.

**Phylogenetic analysis of the F gene coding region**

To address the gap in the knowledge of the F gene evolutionary changes, we have included phylogenetic analysis of HPIV3 based on the entire F gene.

The F gene phylogenetic tree was generated using the same 22 samples used in the amplification of the complete HN gene and 20 sequences previously deposited in GenBank (Fig. 3). Reference sequences were selected according to previously published data for HN phylogenetic analysis [1, 22].

No F reference sequences were available for subclusters C2 and C4, as well as genetic lineages C3c, C3d and C3f.

The topologies of samples and reference sequences in the F and HN gene phylogenetic trees were similar (Figs 3, S1). In the F gene tree, a distinct separation of clusters A, B and C, as well as subclusters C1, C2, C3 and C5, is shown, which was also confirmed with the genetic distances ($P$) between these groups. The same genetic distances as for the previously defined HN based classification were used for the F-gene-based classification [1, 2]. Genetic distances between A:B, A:C and B:C clusters are 0.053, 0.055 and 0.060, respectively. In cluster C, four subclusters are visible in the F gene tree and the genetic distance range was 0.026–0.038. No sequences representative of subcluster C4 were available in the GenBank database or among the sequences obtained in this study.

Sample HR/17.11 (9124) (KX467887) obtained in this study represents the C2 subcluster and could serve as a C2 reference sequence in future phylogenetic analyses.

The genetic lineages in C1 and C3 subclusters of the F gene tree overlapped with genetic lineages in the HN gene tree (Table S4a, b).

One sample, KF530247, has different topology when F and HN gene phylogenetic trees are compared. In the HN gene tree this sample belongs to cluster B, while in the F gene tree it falls within subcluster C3 (Fig. 3).

**Molecular analysis of the F and HN glycoproteins**

The 22 F and HN gene sequences were combined into 3748 nt long fragments (KX467885 to KX467906) and the nucleotide differences for these genes and their specific regions calculated using DNAsp 5.10 software. All data are
Fig. 2. Phylogenetic tree generated using the HN gene fragment of 582 nt from the 82 sequences obtained in this study and 41 sequences retrieved from GenBank. The GenBank sequences are shown in grey while sequences from this study are shown in black. The number of identical sequences obtained from samples in this study is indicated in brackets. The tree was generated using the maximum-likelihood method; the Tamura–Nei model with gamma distribution rate (TN93+G) was implemented. The scale bar indicates the proportion of nucleotide substitutions; the numbers at branch nodes are bootstrap values determined for 1000 iterations (only values above 65% are shown). The genetic classification is indicated by the brackets to the right of the figure.
summarised in Table S5. Among the analysed coding regions (cds) and non-coding regions (NCRs), the most divergent is 3' NCR F ($\pi=0.0702$) while the most conserved is 5' NCR F ($\pi=0.016$). The nucleotide diversity between F cds and HN cds is similar; Fcds $\pi=0.0244$ and HNcds $\pi=0.0231$, respectively. The ratio of nonsynonymous to synonymous mutations ($Ka/Ks$) for the F cds is 0.0553, whereas a somewhat lower ratio of 0.0428 was calculated for the HN cds. $Ka/Ks$ ratio is reflected in the percentage of amino acid diversity in the F and HN proteins, which is 5.56% for the F and 5.24% for HN protein.

To examine the immutability of the defined HN and F properties, potential N-glycosylation sites, number of cysteines and the positions of disulphide bonds, antigenic sites etc., we analysed the 22 sequences from this study and sequences from the GenBank database. Twenty-seven GenBank sequences for the F protein originated from the

<table>
<thead>
<tr>
<th>Subcluster</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>8</td>
<td>12</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>27</td>
</tr>
<tr>
<td>C2</td>
<td>1</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C3</td>
<td>7</td>
<td>26</td>
<td>21</td>
<td>32</td>
<td>18</td>
<td>104</td>
</tr>
<tr>
<td>C5</td>
<td>2</td>
<td>6</td>
<td>ND</td>
<td>1</td>
<td>ND</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>44</td>
<td>23</td>
<td>34</td>
<td>22</td>
<td>141</td>
</tr>
</tbody>
</table>

**Table 1.** Number of HPIV3 PCR-positive clinical samples collected in 2011–2015 with subcluster determination assigned by phylogenetic analysis and genetic distance calculation; ND not detected

**Fig. 3.** Phylogenetic tree generated using the coding region of the F gene from the 22 sequences obtained in this study and 21 sequences retrieved from GenBank. The GenBank sequences are shown in grey while sequences from this study are shown in black. The tree was generated using the maximum-likelihood method; the Tamura-Nei model with gamma distribution rate (TN93+G) was implemented. The scale bar indicates the proportion of nucleotide substitutions; the numbers at branch nodes are bootstrap values determined for 1000 iterations (only values above 65% are shown). The genetic classification is indicated by the brackets to the right of the figure.
period between the early 1980s to 2014, while 50 HN protein sequences originated from between 1957 and 2014.

Analysis of potential N-glycosylation sites showed that all sites defined in the early 1980s (F 238, 359, 446 and 508/HN 308, 351 and 523) persisted in all samples. Several samples have additional N-glycosylation sites in the HN protein: sample HR/19.12 (2501) KX467891 at position 441; sample HR/44.12 (3108) KX467893 at position 30; KF687317 (USA 2009) at position 252; and EU424062 (CAN), M18762 (USA 1980), M18764 (USA 1983) at position 389. One sample, EU326526, has an additional site in the F protein at position 491.

All 14 cysteine residues in the HN protein [14] were conserved in all samples, except for EU814626 (IND 2005), where cysteines at positions 363 and 571 were replaced with serine. Eleven cysteine residues in the F protein [17] were preserved in all samples, except in the sample HR/17.11 (9124) KX467887, where cysteine on position 18 was also replaced with serine.

Antigenic sites for the HN and F proteins were determined in the late 1980s [17]. Despite the wide temporal distance (more than 20 years), 22 samples from this study maintained the same amino acid composition at all sites in both glycoproteins.

Two binding sites that serve for the attachment of the HPIV3 to the host cell were identified in the HN glycoprotein: bifunctional site I at T193 and D216 and site II at H552 and N551 [11, 33]. Three samples from this study, HR/21.14 (1014) KX467904, HR/40.14 (1406) KX467905, HR/46.13 (1650) KX467896, differed in the amino acid positions adjacent to binding site I (Table 2).

More recently, amino acid residues at positions 275, 277, 372 and 426 were identified as contributors in binding to α2,6-linked sialic acid [9]. These positions were completely conserved in the 260 strains isolated over the past 10 years from our study and that of Fukushima et al. [9].

Data published by Streltsov et al. [16] suggested a 216-loop region (residues 210–221 forming a loop around D216), which is important for the HN receptor binding and may help the HN to balance the opposing functions. Only one sample from this study, HR46.13 (1650) KX467896, has an amino acid change at position 217 (Table 2).

**DISCUSSION**

This study presents molecular epidemiology and genetic analysis of the HPIV3 F and HN glycoproteins obtained from clinical samples collected from paediatric patients in the Zagreb region, Croatia, over a period from 2011 to 2015. Although HPIV3 is a common respiratory agent which can cause serious respiratory illnesses in a paediatric population, little information about its molecular epidemiology is available for European countries [23].

The patient population analysed in this study was representative of previous HPIV studies [1, 19, 21, 23, 34]. Namely, the majority of clinical samples originated from very young children (79.2 % were under 2 years old) who suffered from respiratory illness which required some kind of medical assistance in hospital. Few patients were immunocompromised.

Patients from this study mainly suffered from URI (59.4 %), as was also reported by Mizuta et al. [19], contrary to other studies reporting pneumonia as the most common manifested illness among the HPIV3 infected and hospitalised patients [2, 35].

The reason for such a discrepancy could be in an imprecise meaning of the phrases ‘hospital treatment’ or ‘hospitalized patients’. In our study, ‘hospital treatment’ or ‘hospital attendance’ referred to every kind of medical treatment in hospital: hospital ambulance, day care hospital or hospital ward (hospitalization). We believe that the collection of clinical samples

---

**Table 2.** Summary of samples obtained in this study with amino acid substitution in a previously reported functional place. N-glycosylation site or cysteine number

<table>
<thead>
<tr>
<th>Gene</th>
<th>Functional place (reference)</th>
<th>Amino acid position</th>
<th>Sample no.</th>
<th>Amino acid substitution/position</th>
</tr>
</thead>
<tbody>
<tr>
<td>HN</td>
<td>N-glycosylation*</td>
<td>+1 additional site</td>
<td>KX467891</td>
<td>/441</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>KX467893</td>
<td>/30</td>
</tr>
<tr>
<td></td>
<td>Binding site †</td>
<td>T193, D216</td>
<td>KX467904</td>
<td>V→I/191</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>KX467905</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Binding pocket for α2,6 sialic acid receptor‡</td>
<td>275,277,372,426</td>
<td>KX467896</td>
<td>D→N/279</td>
</tr>
<tr>
<td></td>
<td>'216 loop§</td>
<td>210–221</td>
<td>KX467896</td>
<td>I→T/217</td>
</tr>
<tr>
<td>F</td>
<td>Cysteine number</td>
<td></td>
<td></td>
<td>+1 additional site</td>
</tr>
</tbody>
</table>

*Coelingh et al. [18].
†Porotto et al. [11].
‡Fukushima et al. [9].
§Streltsov et al. [16].
||Coelingh and Tierney[17].
limited to children in a hospital ward would have increased the percentage of illnesses such as pneumonia, croup or bronchiolitis in the HPIV3-infected population.

Although the HPIV3-positive samples could be detected during the whole analysed period, in 2012, 2014 and 2015 the seasonal pattern included two epidemics per calendar year. The epidemics were in spring and autumn, with epidemic peaks in May and October. In 2011 and 2013 one epidemic occurred per year, in the spring and autumn, respectively. Our results are not in full accordance with published epidemic patterns of HPIV3, which reported one epidemic per year in the May–August period [1, 33, 34]. An epidemiological study of South American countries reported seasonal behaviour of HPIV3 with an epidemic peak between May and June, and also more than one epidemic peak per calendar year [21]. This study suggested that HPIV appearance could be connected with some climatic characteristics.

Previous phylogenetic studies [1, 2, 22] were based on HN coding sequences and their authors determined the ranges of genetic distances required for genetic classification of HPIV3. The phylogenetic studies based on the HPIV3 F gene were conducted in the early 1990s in the USA [24] and, recently, in 2015 in Spain [23]. Using rules for HN classification, we generated phylogenetic trees based on the F gene cds, HN gene cds and partial HN of 582 nt. Phylogenetic analysis confirmed that the F gene and the 582 nt fragment of the HN gene were suitable for HPIV3 classification. Because the 582 nt HN sequence is an appropriate length for processing in one PCR/sequencing reaction, we believe that this finding will be useful for fast screening from a limited amount of clinical samples in large molecular epidemiology studies.

The topology of the samples from this study, as well as for the chosen reference samples, remained the same for all reconstructed trees regarding classification into clusters and subclusters. The only exception was reference sample KF530247 that grouped within cluster C in the F gene tree and in cluster B in the HN gene tree. The literature data also confirmed that sample KF530247 belonged to cluster B in the HN gene tree [22].

As recombination seems to be generally rare or even absent in most negative-sense RNA viruses, our assumption is that the KF530247 genome sequence is an artificial chimera of two or more different HPIV3 viruses, either due to laboratory contamination or informatics error.

As this sample was not available for new isolation and resequencing, the reason for such a discrepancy in the phylogenetic results was not further investigated.

All samples from this study were classified in cluster C, which is the most frequent cluster to be detected in recent HPIV3 molecular studies [1, 2, 22, 23]. During the study period, subcluster C3 became prevalent over other C subclusters with more than 90% positive samples in 2012 and 2013, which is consistent with recent results [1, 23]. The domination of C3a genetic lineage in our sample pool is also in line with recent HPIV3 studies.

Nucleotide and amino acid diversity of HN and F glycoproteins could indicate changes in their functions regarding the attachment to receptors, fusion and stimulation of the immune system to produce neutralization antibodies.

The nucleotide substitutions in HPIV3 F and HN are distributed evenly throughout entire genes. However, amino acid substitutions are restricted to the cytoplasmic and transmembrane regions, while changes in the stalk and globular head were limited and found only in individual samples. Although Godoy [23] revealed some amino acid substitutions which could be specific marks for phylogenetic subgroups, we did not detect such specificity in the sample pool analysed in this study.

Extensive analysis of the functional characteristics of the F and HN glycoproteins from the samples obtained in this study and reference sequences showed unmodified functional sites regardless of the temporal and geographical distance of the samples. These data indicated the very conservative nature of HPIV3 glycoproteins.

Only a few samples have amino acid substitutions that nominally could influence the functionality of F and HN glycoproteins, and the importance of these mutations remains to be explored in further investigations of HPIV3.

We believe that the data from this study contribute to knowledge of the molecular epidemiology and circulation patterns of the HPIV3 in Europe. It also presents an additional proof of the conserved nature of the HPIV3 glycoproteins, which could help in the design and usage of an efficient HPIV3 vaccine.

Funding information
Funding for this research was received from the Croatian Science Foundation (project number 6255).

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
This study is approved by the ethics committees of the University Hospital for Infectious Diseases ‘Dr Fran Mihaljevic’, as part of scientific project number 6255 supported by the Croatian Science Foundation.

References
4. Ljubin-Sternak S, Marijan T, Ivkovic-Jurekovic I, Čepin-Bogović J, Gagro A et al. Etiology and clinical characteristics of single and...


