Genetic analysis of the M gene of equine influenza virus strains isolated in Poland, in the context of the Asian-like group formation

Małgorzata Kwaśni, Ilona M. Góra, Jan F. Żmudziński, Jerzy Rola, Mirosław P. Polak, Wojciech Rożek

Department of Virology, National Veterinary Research Institute, 24-100 Pulawy, Poland
malgorzata.kwasnik@piwet.pulawy.pl

Received: July 17, 2018 Accepted: October 15, 2018

Abstract

Introduction: Traditionally, evolutionary analysis of equine influenza virus (EIV) is based on the HA gene. However, the specificity of the influenza virus enables the classification of viral strains into different phylogenetic groups, depending on the gene being analysed. The aim of the study was to analyse phylogenetic paths of EIV based on M gene with reference to the HA gene. Material and Methods: M gene of Polish isolates has been sequenced and analysed along with all M sequences of EIV available in GenBank database. Phylogenetic analysis was performed using BioEdit, ClustalW, and MEGA7 softwares. Results: The clustering of the strains isolated not only from Asia but also from Europe into one common Asian-like group of EIV was observed. Twelve nucleotide substitutions in the M gene of strains from the Asian-like group were crucial for the evolutionary analysis. We also observed homology in the M gene of the Asian-like and H7N7 strains. Conclusions: M gene specific for the Asian-like group is present in strains recently isolated in Europe and Asia, which were classified previously in the Florida 2 clade based on HA. Therefore, Asian-like group does not seem to be assigned to a specific geographical region. Traces of H7N7 strains in more conservative genes like M of some contemporary EIV strains may indicate the link between the old phylogenetic group and recent H3N8 strains. Analysis of conservative genes may be more useful in tracking the direction of virus evolution than in the genes where the high variability rate may blur the original relationships.

Keywords: influenza virus, equine influenza, M protein, phylogenetic analysis.

Introduction

Equine influenza virus (EIV) belongs to the Orthomyxoviridae family, influenza type A. Historically two subtypes of EIV have been identified: H7N7 and H3N8. H7N7 EIV has not been isolated from horses since 1979 (39), but serological evidence of its circulation in unvaccinated horses was recorded at the end of the eighties in India (34) and at the beginning of the nineties in Croatia and the USA (24), while EIV H3N8 circulates throughout the world. The H3N8 influenza virus currently circulating has been isolated from pigs and camels and has become established in dogs as canine influenza virus (CIV) (27). The genome of equine influenza viruses consists of eight segments of negative sense, single-stranded RNA. Each RNA segment is complexed into ribonucleoproteins, which make the basic unit of the virion, important for replication and transcription. The seventh segment of EIV is coding for matrix proteins (M). The M segment is considered conservative and mutations within it may influence the virulence, pathogenic properties, and host range of the virus. Originally, it was thought that segment 7 encoded two proteins, M1 and M2, but recently, M42, a variant of M2, has been documented (40). The importance of the M1 and M2 proteins is their role during virus assembly and budding (31). M1 protein is made up of 252 amino acid residues and consists of N-terminal (1–66 aa), linker (67–86 aa), middle (87–164 aa), and C-terminal (165–252 aa) domains (26). M2 is type III membrane protein of 97 amino acids and it possesses N-terminal ectodomain (1–23 aa), transmembrane (24–46 aa), and a cytoplasmic domain (47–97 aa) (23). Seeing that the sequence of M2 ectodomain is conserved has frequently been explored for the development of
a universal influenza A vaccine. One of the anti-influenza therapeutics directed against M2 protein, amantadine, inhibits virus replication by blocking the acid-activated M2 ion channel. Novel vaccines targeting M1 or M2 proteins to induce cross-subtype protection have been proposed (11).

Phylogenetic analysis of EIV is generally based on the haemagglutinin (HA) sequences. Accordingly, the H3N8 subtype evolved as a single lineage until the end of the eighties, and then split into two lineages: American and European. The American lineage diverged into the Kentucky, South American, and Florida sublineages. The Florida sublineage evolved into two clades, and most viruses isolated all over the world recently have been found to belong to the Florida clade 2 (6, 7, 9, 12, 16, 19-21, 29). However, the Florida clade 1 was also responsible for several outbreaks of the disease (1, 38).

Phylogenetic analysis of equine influenza viruses isolated in Asia was performed by Virmani et al. (36, 37). They observed changes of the M gene sequence of some Indian and Chinese isolates and grouped them in a separate cluster, which has been provisionally designed as an Asian clade. The aim of our study was to analyse phylogenetic paths of EIV based on M gene with reference to the HA gene. We studied M gene sequences of EIV available in the GenBank database, referring them to the sequences of strains isolated in Poland.

Material and Methods

Samples, RT PCR, and sequencing. To obtain M gene sequences of three Polish isolates (A/equine/Pulawy/1/2005, A/equine/Pulawy/1/2006, and A/equine/Pulawy/1/2008), RNA was extracted directly from nasal swabs from horses showing respiratory symptoms, and a two-step protocol was used, as previously described (14). The RNA was transcribed using UNI-12 primer (5'-AGC GAAAGCAGG-3') followed by PCR with M-specific primers: forward 5'-TATTCGTCCTCAGGGAGCAA AAGCAGGTG-3' and reverse 5'-ATATCGTCTCGT ATTAAAGTAGAAAACAGGTAGTCTTTT-3'. Amplified products were visualised by agarose gel electrophoresis and purified using a QIAquick PCR Purification Kit (Qiagen, Germany). Amplicons were sequenced using an ABI PRISM 310 Genetic Analyser and Gene Scan Analysis Software (Applied Biosystems, USA). HA1 sequences of Polish isolates used in this study were previously published (18, 32).

Phylogenetic analysis. M gene sequences encoding M1 and M2 proteins of A/equine/Pulawy/1/2005, A/equine/Pulawy/1/2006, A/equine/Pulawy/1/2008, and 63 other M gene sequences obtained from the GenBank database representative of the main evolutionary lineages of EIV were used to construct a phylogenetic tree. We constructed also phylogenetic tree of HA1 coding sequences of the same set of EIV strains. Phylogenetic trees were generated applying the maximum-likelihood algorithm with bootstrap analysis of 1,000 replications using MEGA 7.0 software (17). For detailed analysis of a particular nucleotide or amino acid position, three alignments were carried out: with M gene sequences, predicted amino acid sequences of M1, and the counterparts for M2 proteins (carried out for 178 EIV sequences each, all available in GenBank database on 1st May, 2018). Each of these alignments is available in Supplementary materials. Accession numbers of all M and HA gene sequences of EIV strains mentioned in the phylogenetic trees or manuscript are included in Tables S1 and S2 (Supplementary materials). The full-length M segment sequences of Polish isolates were submitted to GenBank (accession numbers: KY855445 for A/equine/Pulawy/1/2005, KY855447 for A/equine/Pulawy/1/2006 and KY855446 for A/equine/Pulawy/1/2008). Phylogenetic analysis was carried out using BioEdit v 7.2.5 (13) and Clustal W (35) software.

Results

The M gene sequences of EIV strains isolated in Poland and around the world for 60 years were analysed. The phylogenetic analysis classified A/equine/Pulawy/1/2005 and A/equine/Pulawy/1/2008 as members of a newly postulated Asian-like group (the “Asian” group was described earlier by Virmani et al. (36)) and A/equine/Pulawy/1/2006 as a member of Florida clade 2 (Fig. 1). EIV strains belonging to the Asian-like group were mostly isolated in Asia (China, Japan, India, Mongolia, Kazakhstan, Turkey), but A/equine/Solihull/1/2007, A/equine/Richmond/1/2007, A/equine/Pulawy/1/2005, A/equine/Pulawy/1/2008, A/equine/Neuville-PresSees/1/2011, A/equine/Cambremer/1/2012, A/equine/Gironde/1/2014, and A/equine/Saone-et-Loire/1/2015 were isolated in Europe. Phylogenetic analysis based on HA1 nucleotide sequences did not show a separate Asian group of EIV. Most strains assigned to the Asian-like group based on M analysis clustered within Florida clade 2, while A/equine/Pulawy/1/2005 was classified into the European lineage (Fig. 2). To identify nucleotides defining the separation of strains into the Asian-like group, M gene sequences (including Polish isolates) were aligned to A/equine/Miami/63. We noted 12 nucleotide substitutions in sequences of strains from Asian-like group as crucial for phylogenetic analysis and classification. Among them, seven substitutions (263 A, 309 A, 406 T, 535 T, 697 T, 966 A, and 978 A) seem to be discriminating for strains from Asian-like group (present in all of Asian, but in less than 10% of other strains) (Table 1.). The substitutions 68 A, 637 A, 856 C, and 967 G are typical for most strains from the Asian-like group, except A/equine/Pulawy/1/2005 (68 G, 637 G, and 856T), A/equine/Hubei/6/2008,
A/equine/Pulawy/1/2008, A/equine/Solihull/1/2007, A/equine/Richmond/1/2007, A/equine/Neувиль-Прессе/1/2011, A/equine/Cambremere/1/2012, A/equine/Gironde/1/2014, A/equine/Saone-et-Loire/1/2015, A/equine/Ankara/1/2013, A/equine/Yokohama/aq13/2010 (856T), and A/equine/Xuzhou/01/2013 (G967A). The substitution 505G is observed in all strains from Asian-like group, but also in 15.5% of other EIV strains. From 12 substitutions listed above, eight were also found in strains from H7N7 group (Table 1).

Some substitutions resulted in amino acid changes within M1 or M2 protein sequences of Polish strains, which are characteristic also for strains from Asian-like group (M1 – 80I, 95K and 15I, 80I, 95K for A/equine/Pulawy/1/2005 and A/equine/Pulawy/1/2008 respectively, M2 – 85S and 89S for both A/equine/Pulawy/1/2005 and A/equine/Pulawy/1/2008) (Table 1).


Table 1. Nucleotide and amino acid positions in M sequences identified as characteristic for strains from Asian-like group

<table>
<thead>
<tr>
<th>Nucleotide substitution</th>
<th>EIV strains with the same substitution</th>
<th>% of EIV strains</th>
<th>Aminoacid changes M1</th>
<th>Aminoacid changes M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>263 A</td>
<td>-</td>
<td>-</td>
<td>80 I</td>
<td>-</td>
</tr>
<tr>
<td>309 A</td>
<td>LaPlata/93, Jilin/1/1989, Guangxi/3/2011, New York/152429/2002, H7N7 group</td>
<td>8.3</td>
<td>95 K</td>
<td>-</td>
</tr>
<tr>
<td>406 T</td>
<td>Newmarket/1/1993, Jilin/1/1989, H7N7 group</td>
<td>7.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>535 T</td>
<td>North Carolina/152429/2002, H7N7 group</td>
<td>6.6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>637 A*</td>
<td>Newmarket/1/1993, Guangxi/3/2011, H7N7 group</td>
<td>7.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>697 T</td>
<td>H7N7 group</td>
<td>6.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>856 C **</td>
<td>-</td>
<td>-</td>
<td>48 S</td>
<td>-</td>
</tr>
<tr>
<td>966 A</td>
<td>H7N7 group</td>
<td>6.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>967 G ***</td>
<td>Belfond/6-2/2009, H7N7 group without Prague/56</td>
<td>6.1</td>
<td>85 S</td>
<td>-</td>
</tr>
<tr>
<td>978 A</td>
<td>Moulton/1/1998, New York/1/175, Lexington/1/66</td>
<td>1.6</td>
<td>89 S</td>
<td>-</td>
</tr>
</tbody>
</table>

* not appeared in A/equine/Pulawy/1/2005
*** not appeared in A/equine/Xuzhou/01/2013

Fig. 1. Phylogenetic tree based on M nucleotide sequences generated using the maximum-likelihood algorithm applying bootstrap analysis with 1,000 replications. Bootstrap values were shown next to the branches. Evolutionary analyses were conducted in MEGA 7.0 software.
Fig. 2. Phylogenetic tree based on HA1 nucleotide sequences generated using the maximum-likelihood algorithm applying bootstrap analysis with 1,000 replications. Bootstrap values were shown next to the branches. Evolutionary analyses were conducted in MEGA 7.0 software.


**Discussion**

The phylogenetic tree based on M gene sequences clustered EIV strains into groups: Predivergent, European, American, Asian clade 1, Florida clade 2, Asian-like, and H7N7. However, European and American lineages are less distinct than those observed in the HA1 analysis (Figs 1 and 2). Phylogenetic analysis based on HA1 nucleotide sequences did not show a separate Asian-like group of EIV. Most strains assigned to the Asian-like group based on M analysis clustered within Florida clade 2, while A/equine/Pulawy/1/2005 was classified into the European lineage (Fig. 2). Analysis based on HA1 sequences of recent Chinese, Mongolian, and Indian EIV strains by Qi et al. (28) also classified these strains to Florida sublineage clade 2, suggesting that they derived from strains circulating in Europe as A/equine/Newmarket/1/07 and A/equine/Cheshire/1/07. We observed that not only Indian, Japanese, Chinese, Mongolian, Turkish, and Kazakh isolates belong to the Asian group but also the strains isolated in Europe such as A/equine/Solihull/1/2007, A/equine/Richmond/1/2007, A/equine/Pulawy/1/2005, A/equine/Pulawy/1/2008, A/equine/Neuville-Pres-Sees/1/2011, A/equine/Cambrere/1/2012, A/equine/Gironde/1/2014, and A/equine/Saone-et-Loire/1/2015. The appearance of strains isolated in Europe belonging to the Asian group might be explained either by independent emergence or by the movement of infected horses between Asia and Europe.

We observed 12 nucleotide substitutions in sequences of the M gene, specific for the Asian-like group, which seem to be significant for separation of the new clade (Table 1). Seven substitutions were discriminating for the Asian-like group (present in all of Asian-like strains but in less than 10% of other strains). Some of these nucleotide substitutions resulted in amino acid changes. Amino acid residues at positions: 15I, 80I, 95K of M1 and 85S, 89S of M2 sequences (present also in A/equine/Pulawy/1/2005 and A/equine/Pulawy/1/2008) were described before as characteristic for Asian strains by Virmani et al. (36). They indicated also substitution 48S within M2 sequence as typical for strains from Asian group, this substitution does not occur within sequences of Polish isolates. Substitution 15I occurs within N-terminal domain, 80I within linker, and 95K within middle domain of M1 protein. N-terminal and middle domain of M1 may affect virus budding, morphology, and/or vRNP incorporation into virions (33, 41). Substitutions 85S and 89S are located in cytoplasmic domain of M2 protein. This domain is responsible for virus particle morphology, virus infectivity, and reducing the incorporation of NP and viral RNA into progeny virus particles (5, 30).

Despite A/equine/Pulawy/1/2005 being classified into the Asian-like group, analysis of the M gene showed some similarities to strains of European lineage. Considering positions distinguishing Asian-like from European lineage strains, A/equine/Pulawy/1/2005 has 12 and 2 substitutions characteristic for Asian-like and European lineages, respectively. The occurrence of substitutions in the M gene characteristic for the European lineage together with the classification into European lineage based on HA or NS genes indicates that A/equine/Pulawy/1/2005 can be phylogenetically located between European and Asian-like lineages. The last documented isolation of EIV strains belonging to European lineage was in 2007 (A/equine/Athens/02/2007 and A/equine/Switzerland/P112/2007) when a new Asian-like group appeared (3, 4).

Depending on the nucleotide sequences of the M or HA1 genes, EIV strains isolated in Poland were classified as Asian-like, Florida clade 2, or European lineage. For several other EIV strains, analysis of different genes gave discordant results for the strain's phylogenetic group. A/equine/Moulton/1/1998 was classified as American lineage by HA gene analysis, but as European lineage by M gene analysis; A/equine/Tennessee/5/1986, A/equine/Kentucky/1/1986, and A/equine/Johannesburg/1/1986 were classified as Predivergent lineage by HA gene, but as American lineage by M gene analysis (Figs 1 and 2). A/equine/Tennessee/5/86 was classified by Lindstrom et al. (22) as an ancestral strain of more recent European isolates, based on the topology of phylogenetic tree constructed using neighbour-joining method. Murcia et al. (25) also showed dual phylogenetic classifications of several EIV strains depending on the analysed genes and explained it as a result of intrasubtype reassortation. The study of EIV genes indicated that the M gene has not evolved in
parallel with the HA gene (10). We hypothesise that changes within the M gene are important in the determination of a new EIV group, and other genes may follow the evolution of the M gene.

Among 12 nucleotide substitutions in sequences of the M gene, specific for the Asian-like group described above, some were also observed in strains from the H7N7 group (Table 1), despite the distant relationship between M genes. Evolutionary analysis of the M gene carried out by Ito et al. (15) showed at least several major host-related lineages. The most divergent of them was formed by equine H7 strain A/equine/Praha/56, subsequently H13 gull viruses, human and classical swine viruses, and an avian lineage (subdivided into North American avian viruses, including recent equine viruses, and Old World avian viruses, including avian-like swine strains). Our previous analysis based on the NS gene showed similarities between H7N7 strains and a group comprising A/equine/Pulawy/1/2005, A/equine/Aboyne/2005, and A/equine/Lincolnshire/1/2005 representing European lineage (18). The similarities of the Asian-like and H7N7 groups seem not to be purely coincidental. Considering the ability of the influenza virus to circulate in the horse population over a long period without much genetic changes, described as the “frozen evolution” phenomenon (2, 8), it seems possible that contemporary EIV strains show similarities to strains isolated a few decades ago. Traces of H7N7 strains in more conservative genes like M or NS of some contemporary H3N8 strains may indicate the link between the old phylogenetic group (H7N7) and recent H3N8 EIV strains.

In summary, our phylogenetic analysis of the M gene showed that strains from Asia and Europe clustered in the Asian-like group. Despite the fact that according to HA1 analysis they belong to two lineages, European and Florida 2, the two strains A/equine/Pulawy/1/2005 and A/equine/Pulawy/1/2008 showed a phylogenetic relationship based on the M gene, and both were closely related to strains from the Asian-like group of EIV. M gene typical for strains from Asian-like group is present in more recently isolated strains that clustered into the Florida clade 2 for the HA gene. Therefore, Asian-like group does not seem to be assigned to a specific geographical region. Analysis of the M gene of A/equine/Pulawy/1/2005 classified into Asian-like group showed some similarities to strains from European lineage. We observed also similarities in the M gene of the Asian-like and H7N7 EIV strains. Analysis of conservative M genes might reveal some traces of evolutionary changes hardly seen in more variable genes like HA.

**Conflict of Interests Statement:** The authors declare that there is no conflict of interests regarding the publication of this article.

**Financial Disclosure Statement:** The study was supported by the project, KNOW (Leading National Research Centre) Scientific Consortium “Healthy Animal - Safe Food”, a decision of the Ministry of Science and Higher Education, No. 05-1/KNOW2/2015 (K/02/1.0).

**Animal Rights Statement:** None required.

The online version of this article (DOI:10.2478/jvetres-2018-0057) offers the following supplementary material: 1) Supplementary Figure S1. Multiple alignment for 178 nucleotide sequences of M gene of EIV (M gene sequences available in GenBank database on 1st May 2018); 2) Supplementary Figure S2. Multiple alignment for 178 amino acid sequences of M1 protein of EIV (M1 protein sequences available in GenBank database on 1st May 2018); 3) Supplementary Figure S3. Multiple alignment for 178 amino acid sequences of M2 protein of EIV (M2 protein sequences available in GenBank database on 1st May 2018); 4) Supplementary Table S1. The detailed information of M gene sequences used in this study; 5) Supplementary Table S2. The detailed information of HA gene sequences used in this study.

**References**


