Infection, Genetics and Evolution 24 (2014) 183–186

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Short communication

The genome of an influenza virus from a pilot whale: Relation to influenza viruses of gulls and marine mammals

Marco Groth [a], Jeannette Lange [b], Pumaree Kanrai [c], Stephan Pleschka [c], Christoph Scholtissek [c], Andi Krumbholz [d], Matthias Platzer [a], Andreas Sauerbrei [b], Roland Zell [b,*]

[a] Genome Research, Fritz Lipmann Institute, Leibniz Institute of Age Research, Beutenbergstr. 11, D-07745 Jena, Germany
[b] Department of Virology and Antiviral Therapy, Jena University Hospital, Friedrich Schiller University, Hans-Knoell-Str. 2, D-07745 Jena, Germany
[c] Institute of Medical Virology, Justus Liebig University, Schubertstrasse 81, D-35392 Giessen, Germany
[d] Institute for Infection Medicine, Christian Albrecht University of Kiel and University Medical Center Schleswig-Holstein, Campus Kiel, Brunswiker Str. 4, D-24105 Kiel, Germany

ARTICLE INFO

Article history:
Received 3 February 2014
Received in revised form 19 March 2014
Accepted 25 March 2014
Available online 3 April 2014

Keywords:
Avian influenza virus
Influenza virus ecology
Whale
Pinnipeds

ABSTRACT

Influenza virus A/whale/Maine/328B/1984 (H13N2) was isolated from a diseased pilot whale. Since only a partial sequence was available, its complete genome was sequenced and compared to the sequences of subtype H13 influenza viruses from shorebirds and various influenza viruses of marine mammals. The data reveal a rare genotype constellation with all gene segments derived of an influenza virus adapted to gulls, terns and waders. In contrast, the phylogenetic trees indicate that the majority of influenza viruses isolated from marine mammals derived from influenza viruses adapted to geese and ducks. We conclude that A/whale/Maine/328B/1984 is the first record of an infection of a marine mammal from a gull-origin influenza virus.

Influenza A virus (FLUAV) is the single species of the genus *Influenzavirus A* of the family *Orthomyxoviridae* (McCauley et al., 2012). The genome of FLUAV comprises eight RNA segments with negative strand polarity. It has a total length of appr. 13,600 nucleotides and encodes up to 13 proteins. The virion is composed of an envelope with three integral viral membrane proteins (hemagglutinin, HA, neuraminidase, NA, proton channel M2), the matrix protein M1 associated with the inner membrane and eight ribonucleoprotein complexes, each comprising one RNA segment, nucleoproteins (NP) and the RNA-dependent RNA polymerase complex (PB1, PB2, PA). The nonstructural proteins NS1 and NS2/NEP (nuclear export protein) are expressed in the virus-infected cell. Some virus strains may express additional proteins like PB1-F2, PB1-N40 and PA-X (McCauley et al., 2012; Jagger et al., 2012).

Influenza virus ecology is very complex. Sixteen HA and nine NA types were identified in aquatic birds, their main hosts, and can recombine to up to 144 FLUAV subtypes (McCauley et al., 2012). Further subdivision into several distinct genetic lineages of each genome segment was proposed (Webster et al., 1992). For each FLUAV genome segment, accumulation of nucleotide substitutions (genetic drift) and geographic isolation gave rise to several characteristic lineages of the eastern (Eurasia) and western (America) hemispheres (Webster et al., 1992; Olsen et al., 2006). For some segments, host adaptation lead to specific lineages of geese, ducks and swans (order *Anseriformes*) on the one hand, and gulls, terns and waders (*Charadriiformes*) on the other hand (Webster et al., 1992). A total of 184 lineages that reassort to at least 110 FLUAV subtypes (HA/NA combinations) and more than 500 genotypes have been recorded (Lu et al., 2007).

Stable infection chains also exist in humans, pigs and horses. In addition, trans-species infections were observed in many other species of birds and mammals. Influenza virus infections of marine mammals (pinnipeds and cetaceans) are of special interest as these mammals may come in close contact to aquatic birds. First evidence of FLUAV infection of whales was provided by Lvov et al. (1978) who succeeded to isolate 14 H1N3 influenza virus strains from one liver and 13 lung specimens of striped whales (baleen whales, family *Balaenopteridae*) that were hunted in the 1975/76 whaling season in the South Pacific. Later, Hinshaw et al. (1986) described the isolation of H13N2 and H13N9 strains from the hilar node and lungs of a sick pilot whale (*Globicephala melaena*, family *Delphinidae*). Capturing of this diseased animal coincided with two strandings of dead pilot whales at the New England coast in October and November 1984. However, virus isolation from the stranded dead whales was not successful. Various publications reported influenza virus isolates from harbor seals (*Phoca vitulina*) that indicated local influenza epizootics at the New England coast

* Corresponding author. Tel.: +49 3641 9395711; fax: +49 3641 9395732.
E-mail address: roland.zell@med.uni-jena.de (R. Zell).

http://dx.doi.org/10.1016/j.meegid.2014.03.026
1567-1348/

caused by H7N7 in 1979/80, H4N5 in 1982/83, H4N6 in 1991, H3N3 in 1991/92 and H3N8 in 2011 (Geraci et al., 1982; Hinshaw et al., 1984; Callan et al., 1995; Anthony et al., 2012). All epizootics were associated with high mortality. Finally, two influenza A(H1N1)pdm09 isolates from nasal swabs of two healthy female northern elephant seals (*Mirounga angustirostris*) living at the Californian Pacific coast were published (Goldstein et al., 2013).

Despite the recovery of 14 FLUAV isolates from baleen whales and three isolates from a pilot whale many years ago, only partial sequences of these viruses are available in the GenBank (Supplementary Table S1). This lack of data prompted us to sequence the genome of A/whale/Maine/328B/1984 (H13N2) – the only whale FLUAV isolate available to us – and to compare it with the published sequences of other marine mammals and avian H13 viruses.

A/whale/Maine/2/1984 (H13N2, from hilar node) and A/whale/Maine/2B/1984 (H13N2, from lungs) were isolated by inoculation of embryonated hens' eggs from tissue suspensions of a sick pilot whale captured near Portland, Maine, in October 1984 (Hinshaw et al., 1986). A third virus was isolated from the hilar node of the same animal, designation: A/whale/Maine/1/1984 (H13N9). In later publications, the two former viruses were obviously renamed: Henceforward, the H13N2 isolate from the hilar node was termed A/pilot whale/Maine/328HN/1984 (Chambers et al., 1989) and A/whale/Maine/328/1984, respectively (Mandler et al., 1990; Kawaoka et al., 1998). Accordingly, A/whale/Maine/328B/1984 (H13N2) used in the present study refers to the isolate from the lungs of the pilot whale (original name: A/whale/Maine/2B/1984, Hinshaw et al., 1986). It was stocked at −80 °C for many years in the virus strain collection of the Institute of Virology (Vet. Faculty) of the Justus Liebig University Giessen, Germany. Later, the strain collection was transferred to the Institute of Medical Virology (Med. Faculty). The passage history of this isolate is unknown.

The virus was grown in embryonated hens' eggs. Virions were sedimented by ultracentrifugation from 10 ml allantoic liquid, and RNA was extracted from the pellet. For sequencing, 0.5 μg RNA was used for library preparation. Sequencing was performed on a HiSeq2000 instrument (Illumina) (Bentley et al., 2008). Sequencing resulted in 2,887,709 single-end reads. Mapping of the reads to a set of avian FLUAV reference sequences with low stringency using ssaha2 (Ning et al., 2001) identified 120,172 FLUAV-similar reads. These were used for *de novo* assembly using the CLC Assembly Cell v4.0 (CLC bio, Aarhus, Denmark). Then, the reads were mapped to the sequence of A/herring gull/DE/475/1986 (H13N2) using ELAND (Illumina) to determine the consensus sequences of the segments (average coverage 422×). Both approaches resulted in almost complete sequences of each segment. Only the NS segment had two gaps of 73 and 51 nucleotides that had to be closed with conventional Sanger sequencing as described previously (Zell et al., 2008). Some 400 substitutions were observed, all with low nucleotide frequency (1–4%) compared to the consensus sequence. They indicate a moderate sequence heterogeneity (approx. 1 substitution per 34 nucleotides). The proportion of synonymous/non-synonymous substitutions was >0.75. The sequence of A/whale/Maine/328B/1984 (H13N2) was submitted to the GenBank (Acc. Nos. KJ372717-KJ372724). For phylogenetic analysis, sequences of A/whale/Maine/328B/1984 were manually aligned with FLUAV sequences retrieved from the GenBank. Four Bayesian Metropolis-coupled Markov chains were calculated with MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003) using an optimal substitution model until convergence was reached.

The sequence data of the whale FLUAV reveals a typical avian signature of the deduced virus proteins with no obvious adaptations to the mammalian host. However, such substitutions may have gone lost upon virus passaging in embryonated hens' eggs. The genotype constellation J-G-E-13A-D-2G-C-1C is not unique as it matches the genotypes of A/herring gull/DE/475/1986 and A/herring gull/NJ/782/1986 in a collection of 7135 complete FLUAV genomes comprising 534 different genotypes (www.flugenome.org; retrieval date: 14 December 2013).

Phylogenetic analyses included representative sequences of avian H13 viruses, available sequences of marine mammals (pdmH1N1, H3N3, H3N8, H7N7, H13N2, H13N9) and additional reference sequences. The HA tree (Fig. 1) comprises 24 genotypes of the H1, H3, H4, H7 and H13 subtypes. The HA of lung strain A/whale/Maine/328B/1984 clusters together with that of the hilar node (HN) isolate. It belongs to lineage 13A that co-circulates with lineages 13B and 13C. Whereas lineage 13A includes only American FLUAV isolates, the majority of the 13B and 13C strains are from Eurasia (Fig. 1A). With a few exceptions, most hosts of the H13 subtype belong to the *Charadriiformes* order. In contrast, the HA lineages of the seal viruses belong to those genotypes that mainly infect geese/ducks (i.e., 3C, 4A, 7F).

Fig. 1B presents the NA tree. The NA genes of the H13 isolates comprise the lineages 2D, 2G, 3E, 6A, 6B, 8A, and 9A, but only 3E and 6B are restricted to shorebirds. The NA gene of A/whale/Maine/328B/1984 is part of the 2G lineage. The NA gene of A/whale/Maine/1/1984 belongs to the 9A lineage, and the neuraminidases of the seal isolates cluster with the 3A, 7G and 8A lineages.

Phylogenetic analyses of the whale and seal FLUAVs (Supplementary Figs. S1–S6) reveal 18 lineages of internal genes (Supplementary Table S2). The PB2 gene of the whale virus belongs to the J lineage of shorebirds. The PB1 lineage G is of special interest. The GenBank lists more than 2500 entries including 1400 sequences of the highly pathogenic H5N1 strains that first emerged in Hong Kong in 1997 (Subbarao et al., 1998). This lineage is characteristic of Eurasian birds, but 11 sequences are of American origin. One branch of these sequences comprises strains from gulls and whale that were isolated in the 1980s and display a truncated PB1-F2 with a stop codon in position 58 (Fig. S2). The PB1 of A/whale/328B/1984 has a P_{13} residue which is typical of avian isolates. In addition, the importin-binding domain (amino acid 687–759) exhibits a rare K691R substitution. The PA gene clusters in lineage E. The NP sequence is almost identical to the previously published sequence of A/whale/Maine/328/1984 (M27520, Gorman et al., 1990a) and is part of the D lineage characteristic of American shorebirds. The sequence of the M segment clusters with isolates of gulls (lineage C). Surprisingly, our NS sequence belongs to the 1C lineage (*Charadriiformes*) and does not match the previously published 1D sequence of A/whale/328/1984 (GenBank Acc. No. M80952; compare Kawaoka et al., 1998).

The seal FLUAVs exhibit the following properties: The 3C, 4A and 7F hemagglutinin genes belong to lineages found in American geese and ducks whereas the neuraminidase lineages 3B, 7G and 8A are found in viruses of aquatic birds from both avian orders, *Anseriformes* and *Charadriiformes*. The internal genes mainly belong to the genotypes of American geese and ducks (Figs. S1–S6). The PB2 genes cluster with lineage C that is characteristic of American avian viruses. Three of four PB1 genes of seal FLUAVs belong to lineage F (H3N8, H4N5, H7N7 isolates); the H3N3 isolate, however, exhibits the Eurasian avian lineage G (like the pilot whale isolate). The PA gene also clusters inconsistently: three isolates (H3N3, H3N8, H7N7) have a PA of lineage H (American avian), the H4N5 isolate has a PA gene of lineage E (Eurasian and American avian). The remaining segments of the seal isolates belong to lineages of American geese and ducks (NP: lineage H, M: lineage E, NS: lineage 1D).

There are two studies reporting the isolation of FLUAVs from whales. The first paper describes the isolation of 14 influenza viruses from apparently healthy baleen whales (Lvov et al., 1978), the other characterizes three FLUAV strains isolated from a diseased toothed whale (Hinshaw et al., 1986). FLUAV infection of pinnipeds has also been observed occasionally. Four mass extinctions of harbor seals at the New England coast were