Infection, Genetics and Evolution 37 (2016) 8–13

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Deep sequencing reveals the viral adaptation process of environment-derived H10N8 in mice

Kun Mei [a,c,d,1], Guang Liu [b,1], Zhenzhen Chen [a,c], Zhimin Gao [a,d], Lihua Zhao [a,d], Tao Jin [b], Xiaolan Yu [c,*], Quanjiao Chen [b,d,*]

[a] CAS Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences, Hubei 430071, China
[b] BGI-Shenzhen, Shenzhen 518083, China
[c] Hubei Collaborative Innovation Center for Green Transformation of Bio-resources, Hubei Key Laboratory of Industrial Biotechnology, College of Life Sciences, Hubei University, Hubei 430062, China
[d] Center for Influenza Research and Early-warning (CASCIRE), Chinese Academy of Sciences, Beijing 100101, China

ARTICLE INFO

Article history:
Received 16 June 2015
Received in revised form 28 September 2015
Accepted 13 October 2015
Available online 23 October 2015

Keywords:
H10N8 virus
Mouse lung adaption
Next-generation sequencing (NGS)

ABSTRACT

The H10N8 virus was isolated from the water of Dongting Lake, China. Mice were infected while showing no obvious symptoms and replication was restricted to the lungs. When the wild-type virus was serially passaged in the lungs of mice, the resulting viruses became lethal and capable of replication in many other organs. This offered an applicable model for the exploration of viral genome gradual mutation during adaptation in mice. The different passage viruses from mice lung lavage were named P1, P3, P5, and P7, respectively. We sequenced the four viruses using next-generation sequencing (NGS) to analyze the dynamics of the H10N8 viral genome, polymorphism, and amino acid mutation of related proteins. We aimed to demonstrate how a mutant strain of low pathogenicity could become lethal to mice. Using Illumina high-throughput data, we detected the gradual mutations of F277S, C278Q, F611S and L653P in the polymerase acidic (PA) protein, and of L207V and E627K in the PB2 protein during adaptation. Interestingly, many amino acid sites mutated quickly; the others did so more slowly and remained in a heterozygous state for several generations. The PA amino acids S277 and Q278 have previously been found in clinical wild-type strains, including the human-H10N8 isolate in 2013. This demonstrates that the wild-type H10N8 virus had mutated to adapt to mammalian hosts. These data provide important reference information for influenza virus research.

1. Introduction

RNA viruses exploit all known mechanisms of genetic variation to ensure their survival, and rapidly evolving viruses are ideal systems for exploring evolutionary mechanisms such as the influenza virus (Drummond et al., 2003). In our previous study in 2007, an H10N8 strain, named A/environment/Dongting Lake/Hunan/3-9/2007(H10N8) (EN/DT/HN/3-9/07), was isolated from Dongting Lake in Hunan Province, Southern China. Experimental infection showed that wild-type EN/DT/HN/3-9/07 showed no obvious pathogenicity in BALB/c mice, and the virus was restricted to the lungs. However, the virulence of EN/DT/HN/3-9/07 increased rapidly along with increased mouse lung passage. This process is briefly described as follows: the death of mice inoculated with the P3 (passage 2) virus occurred on day 7 p.i., and all mice died within 11 days p.i.; furthermore, viruses could be detected in multiple organs, including the brain. The mice inoculated with the P5 (passage 4) or P7 (passage 6) virus demonstrated more rapid and serious symptom onset than P3-infected mice. The mice inoculated with the P5 virus all died within 5 days p.i., whereas those inoculated with the P7 virus all died within 4 days p.i. (Zhang et al., 2011).

Understanding the changing dynamics of viral genomes is important in exploring the evolutionary mechanism of viruses and for making clinical and public health decisions. Mice are not natural hosts for influenza viruses. When the newly isolated influenza A virus is used to infect mice, it produces an asymptomatic infection of the respiratory tract with no signs of lethal pathogenic effects; however, the unadapted

Abbreviation: P1, EN/DT/HN/3-9/07, NGS data of P1; P3, EN/DT/HN/3-9/07, NGS sequencing data of P3; P5, EN/DT/HN/3-9/07, NGS sequencing data of P5; P7, EN/DT/HN/3-9/07, NGS sequencing data of P7; Strain1, EN/DT/HN/3-9/07, Sanger sequencing data of P1; Strain2, EN/DT/HN/3-9/07, Sanger sequencing data of P3; Strain3, EN/DT/HN/3-9/07, Sanger sequencing data of P5; Strain4, EN/DT/HN/3-9/07, Sanger sequencing data of P7; Strain5, A/mallard/Korea/1203/2010; Strain6, A/northern shoveler/California/9235/2008; Strain7, A/mallard/Interior Alaska/6MP0758/2006; Strain8, A/longtail duck/Maryland/295/2005; Strain9, A/common scoter/Maryland/297/2005; Strain10, A/mallard/Sweden/7/2003; Strain11, A/chicken/Jiangxi/JXA132717/2014; Strain12, A/duck/Jiangxi/JXA132727/2014; Strain13, A/Jiangxi/IPB13c/2013; Strain14, A/Jiangxi/IPB13a/2013; Strain15, A/Jiangxi/IPB13b/2013; Strain16, A/Jiangxi/IPB13/2013; Strain17, A/Americangreen-wingedteal/Ohio/13OS1869/2013; Strain18, A/duck/Jiangxi/JXA132712/2014; Strain19, A/duck/Guangdong/E1/2012; Strain20, A/quail/Italy/1117/1965; Strain21, A/shearwater/Australia/2/1972.

* Corresponding authors.
E-mail addresses: yxiaolan@163.com (X. Yu), chenqj@wh.iov.cn (Q. Chen).

[1] KM and GL contributed equally to this article.

http://dx.doi.org/10.1016/j.meegid.2015.10.016
1567-1348/

virus can replicate in the lungs, bronchioles, and trachea, and is able to reach a high titer following serial mouse lung passage (Tong et al., 2012; Webster et al., 1992). For this reason, the viral genome of different mouse lung passages is an eminently suitable model for studying the adaptation of the influenza virus to a host, especially mammalian hosts (Bouvier and Lowen, 2010). In this study, the high-throughput sequencing data of P1, P3, P5, and P7 offered a clue to understanding the molecular mechanism of virus–host co-adaptation. NGS provides detailed insight into the variation present within viral populations. There are many successful applications of NGS to influenza virus studies including drug resistance and epidemiology (Ramakrishnan et al., 2009; Téllez-Sosa et al., 2013; Wu et al., 2013; Yasugi et al., 2012), but there has been little study into the evolutionary aspects of influenza virus infection (Bourret et al., 2013). The serially passaged mouse lung viruses were selected as the objective, and NGS was used to explore the gradual molecular change, along with the enhanced virulence. The results of this study can help in further exploration of the evolution of the influenza virus in mammals.

2. Material and methods

2.1. Virus

A/environment/Dongting Lake/Hunan/3-9/2007(H10N8) (EN/DT/HN/3-9/07) was defined as a wild-type virus. Lung passage experiments using the H10N8 strain in SPF mice were carried out as described previously (Zhang et al., 2011). In brief, three female BALB/c mice aged 6 weeks were anesthetized and inoculated intranasally with $10^{6.5}$ EID_{50} purified wild-type virus (50 μl volume). The mice were sacrificed on day 3 p.i., and their lungs and trachea were collected and washed with a total of 2 ml PBS that contained 0.1% BSA and antibiotics, as described previously (Chen et al., 2009). The lung washes were centrifuged at 4 °C at 4 000 ×*g* for 10 min, and the supernatant was harvested, aliquoted, and stored at −80 °C; this was labeled P1. Subsequently, the other three mice were infected with P1, and their lungs and trachea were recovered and washed as described above; the lung washes were labeled P. The same procedure was performed from passage (P1) to passage (P7). The viruses were frozen at −80 °C until use.

2.2. RNA extraction and Sanger sequencing

Viral RNA was extracted directly from the lung wash by lyzing the viruses with RNAiso Plus (TaKaRa). The RNA was reverse-transcribed into single-stranded DNA using M-MLV reverse transcriptase (Promega). All segments were amplified using a Phusion™ High-Fidelity PCR Kit (New England Biolabs). PCR was performed using specific primers for eight genes (primer sequences are available on request) to sequence the full-length genome. The PCR products were purified using a Gel Extraction Kit (Omega Bio-Tek). The fragments were cloned into pGEM-T Easy Vector and sequenced by the dideoxy method using an ABI 3730 DNA sequencer (Applied Biosystems).

2.3. NGS sequencing

The viral RNAs were qualified by an Agilent Technologies 2100 Bioanalyzer. Reverse transcription polymerase chain reaction (RT-PCR) was conducted using Superscript II reverse transcriptase (Invitrogen). cDNAs were synthesized using the influenza A-specific primers MBTuni-12 and MBTuni-13 according to the instruction described in Zhou et al. (2009). The sequencing libraries with an insert size of 200 bp were then prepared by end-repairing, dA-tailing, adapter ligation and PCR amplification, all similar to the procedures provided by the manufacturer's instructions (Illumina). The libraries were sequenced on an Illumina HiSeq 2000 Sequencer by 90 bp paired-end sequencing.

2.4. Sequencing data assembly

NGS sequencing reads were processed and assembled using the method described by Yu et al. (2014). The raw NGS reads were processed by filtering out low-quality reads (8 bases with qualities <66), adaptor-contaminated reads (with >15 bp matched to the adapter sequence), poly-Ns (with 8Ns), duplication and host contaminated reads (SOAP (Li et al., 2009), <5 mismatches). The filtered reads were mapped to INFLUENZA database to choose the best-matching reference sequences. We then used MAQ (Li et al., 2008) to perform reference-based assembly. In addition, the remaining filtered reads were executed de novo assembly using SOAPdenovo (version 1.06) (Li et al., 2010) and edena (v3.121122) (Hernandez et al., 2008), respectively. The de novo contigs (>200 bp) were aligned to the reference-based assembly sequences, to correct for indels and mismatches. The second running of MAQ was performed to generate the final sequences based on the improved sequences obtained from the combination of the above-mentioned three assemblies. To assess the accuracy of assembly, we compared the assembly sequences with Sanger sequences which were deposited to NCBI database before by ourselves (Zhang et al., 2011).

2.5. Nucleotide polymorphism analysis

We used the SOAP software to map filtered reads to the final assemblies, and acquired the indels/coverage of each nucleotide by using the SAMtools software. To quantify the nucleotide polymorphism at each position, a score was generated by using a formula modified from the one described in Crooks and Brenner (2004). Pi_n is the ratio of A/T/G/C.

$$S = -100 * \sum_{1}^{N} Pi_n * \log_2 Pi_n$$

3. Results and discussion

3.1. Summary of sequencing data of the influenza virus genome

A total of 2,484,253 paired-end filtered reads were gained for four samples after removing contaminating or low-quality reads. More

Table 1
Summary of NGS data.

Sample	Filtered reads	Influenza reads (%)	Average depth[a]							
			PB2	PB1	PA	HA	NP	NA	M	NS
P1	1,784,920	87.23	13,427	5642	2194	10,937	3246	24,378	36,110	64,244
P3	1,726,484	82.76	5426	2963	7072	20,845	24,264	23,991	19,496	37,919
P5	1,931,171	79.35	8459	5203	6498	25,252	21,860	25,211	23,587	40,118
P7	2,484,253	80.01	13,813	8097	14,537	31,618	22,236	27,539	32,949	46,480

[a] Segment length of PB2, PB1, PA, NP, H10, N8, M and NS is 2280, 2274, 2151, 1686, 1497, 1413, 982 and 838 bp, respectively.