Characterization of Quasispecies of Pandemic 2009 Influenza A Virus (A/H1N1/2009) by De Novo Sequencing Using a Next-Generation DNA Sequencer

Makoto Kuroda1*, Harutaka Katano2, Noriko Nakajima2, Minoru Tobiume2, Akira Ainai2, Tsuyoshi Sekizuka1, Hideki Hasegawa3, Masato Tashiro3, Yuko Sasaki4, Yoshichika Arakawa4, Satoru Hata5, Masahide Watanabe6, Tetsutaro Sata3

1 Laboratory of Bacterial Genomics, Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan, 2 Department of Pathology, National Institute of Infectious Diseases, Tokyo, Japan, 3 Influenza Virus Research Center, National Institute of Infectious Diseases, Tokyo, Japan, 4 Department of Bacterial Pathogenesis and Infection Control, National Institute of Infectious Diseases, Tokyo, Japan, 5 Department of Clinical Laboratory, Nagano Red Cross Hospital, Nagano, Japan, 6 Department of Pathology, Nagano Red Cross Hospital, Nagano, Japan

Abstract

Pandemic 2009 influenza A virus (A/H1N1/2009) has emerged globally. In this study, we performed a comprehensive detection of potential pathogens by de novo sequencing using a next-generation DNA sequencer on total RNAs extracted from an autopsy lung of a patient who died of viral pneumonia with A/H1N1/2009. Among a total of 9.4 x 10^6 40-mer short reads, more than 98% appeared to be human, while 0.85% were identified as A/H1N1/2009 (A/Nagano/RC1-L/2009(H1N1)). Suspected bacterial reads such as Streptococcus pneumoniae and other oral bacteria flora were very low at 0.005%, and a significant bacterial infection was not histologically observed. De novo assembly and read mapping analysis of A/Nagano/RC1-L/2009(H1N1) showed more than >200 coverage on average, and revealed nucleotide heterogeneity on hemagglutinin as quasispecies, specifically at two amino acids (Gly172Glu and Gly239Asn of HA) located on the Sa and Ca antigenic sites, respectively. Gly239 and Asn239 on antigenic site Ca2 appeared to be minor amino acids compared with the highly distributed Asp239 in H1N1 HAs. This study demonstrated that de novo sequencing can comprehensively detect pathogens, and such in-depth investigation facilitates the identification of influenza A viral heterogeneity. To better characterize the A/H1N1/2009 virus, unbiased comprehensive techniques will be indispensable for the primary investigations of emerging infectious diseases.

Introduction

In April 2009, an H1N1 triple-reassortant swine influenza virus (A/H1N1/2009) was detected in humans with febrile respiratory illness in North America [1], and the virus has rapidly spread worldwide by human-to-human transmission. According to the disease outbreak news from the World Health Organization, at least 14,711 people died from A/H1N1/2009 between April 2009 and January 2010 (http://www.who.int/csr/don/en/). Fatal cases from A/H1N1/2009 viral infection were summarized in a report by Gill et al. [2].

The genome of influenza A virus (family Orthomyxoviridae) consists of 8 single-stranded negative sense RNA molecules spanning approximately 13.5 kb. The segments range in length from 890 to 2341 nucleotides (nt) and encode a total of 11 proteins [3]. Genetic diversity in influenza virus results from a high mutation rate associated with replication using a low-fidelity RNA polymerase and the reshuffling (reassortment) of segments among coinfected strains. Multiple-reassortant influenza viruses from avian, human, and swine origins emerged as major pandemic influenza viruses (i.e., 1918 H1N1, 1957 H2N2, and 1968 H3N2) causing significant mortality in humans in the 20th century [4]. Such an “antigenic shift” by multiple reassortant drives the emergence of pandemic influenza viruses, with their severity and clinical outcome always unpredictable [5].

Influenza A virus can evade antibodies specific to its attachment protein, hemagglutinin (HA), by the accumulation of amino acid substitutions in HA epitopes [6]. This “antigenic drift” in HA epitopes [7] affects recognition by antibodies that neutralize viral infectivity by blocking the interaction of HA with sialic acid residues on host-cell membranes. The H1 subtype HA has four antigenic sites recognized by monoclonal antibodies with high neutralizing activity, designated Sa, Sb, Ca, and Cb [8]. In addition, 8 continuous B cell/antibody epitopes for human H1N1 HA proteins have been experimentally defined by the Immune Epitope Database and Analysis Resource (IEDB: http://www.immuneepitope.org/) [9]. Immune epitope analysis of HA epitopes in A/H1N1/2009 is also summarized in the Influenza Research Database (http://www.fludb.org/brc/homeExtraPage.do?decorator=influenza&extraPage=separate) [10].
To better predict a future pandemic of influenza A virus, the characterization of possible antigenic drift will be indispensable. Igarashi et al. and Shen et al. reported that a structural comparison of HAs could predict probable future antigenic changes during the evolution of A/H1N1/2009 in the human population [11,12].

In addition to this prediction, extensive investigations on viral quasispecies will be required to reveal the actual appearance of those antigenic changes. Nakamura et al. demonstrated the direct detection of potential pathogens, including influenza virus, using de novo pyrosequencing [13], but the detection appeared to have insufficient redundant sequencing reads to reveal the genetic variation of the viruses. Ramakrishnan et al. demonstrated the discrimination of quasispecies in mixed HA subtype infections of influenza A virus using the same pyrosequencing approach [14]. However, it was shown that the influenza viral RNA sample should be enriched through sequence-specific oligonucleotide capturing prior to pyrosequencing, indicating that such enrichment might represent a biased result.

Here, we performed de novo sequencing using total RNAs extracted from an autopsy lung of a patient infected with A/H1N1/2009, and detected potential pathogens such as bacteria in addition to A/H1N1/2009. Extensive DNA sequencing using the Illumina Genome Analyzer II (GA II) revealed that quasispecies for the HA sequence were generated in single patient. Such heterogeneity demonstrated the antigenic drift of HA, implying the existence of a mechanism to escape pre-existing human immunity to the virus.

Results
Summary of sequencing reads and detection of potential pathogens
To determine the influenza A virus sequence and other potential pathogens, we performed de novo sequencing of double-stranded cDNA from total RNA extracted from the autopsy lung of a patient infected with A/H1N1/2009 virus (A/Nagano/RC1/2009[H1N1]) in August 2009 in Japan. The patient was found to be positive for A/H1N1/2009 by real-time reverse transcriptase-polymerase chain reaction (RT-PCR); histopathological findings were also reported [15]. GA II produced 9.4 x 10⁶ 40-mer reads from the cDNA library (Fig. 1B). To exclude the human-derived read sequences, we performed an analysis pipeline as follows (Fig. 1A). Initially, all 9,475,890 reads were aligned to a reference sequence of human genomic DNA, followed by quality trimming to remove low-quality reads and excluding reads with similarities to ambiguous human sequences, resulting in 9,309,538 reads (98.24%) with a possible human source (Fig. 1B). The remaining 166,352 reads (1.75%) were further analyzed using a BLAST search against non-redundant databases, revealing 80,827 reads (98.24%) with a possible human source (Fig. 1B). The reads remaining 166,352 reads (1.75%) were further analyzed using a BLAST search against non-redundant databases, revealing 80,827 reads (98.24%) with a possible human source (Fig. 1B).

To obtain whole sequences and identify single nucleotide polymorphisms (SNPs) for the segments, the 40-mer short reads were aligned to the sequence of A/Tronto/T0106/2009(H1N1), which was found to be the most similar to the A/H1N1/2009 virus using a BLASTN search. Figure 2 shows dot plot images of the coverage at every nucleotide of the segments. Read coverage was observed at ~×200 on average for all segments, indicating a sufficient redundancy to identify the viral sequences and SNPs. The obtained viral sequences, designated as A/Nagano/RC1/1-2009[H1N1], were consistent with those from A/Nagano/RC1/2009[H1N1] passaged in the Madin-Darby canine kidney (MDCK) cell line, except for 3 possible heterogeneous nucleotides in HA.

The coverage plot curves were not flat throughout the segments. Intriguingly, both ends of segment 1 (encoding PB2), the 3’-end of segment 3 (encoding PA), and approximately 700 nt of segment 8 (encoding NS) showed significant abundant coverage greater than ~1000.

Genetic population analysis of the A/H1N1/2009 virus
To identify heterogeneous populations, alignment results were screened using MapView software (Fig. 3B). Three potential heterogeneous genetic populations were found in segment 4 (encoding HA) at the 515, 715, and 716 nt positions (Fig. 3A), but not in other segments. The read alignments shown in Fig. 3B indicate that either the GGT or AAT sequence appeared at the 715–717 nt position, but not the GAT or AGT sequence. In addition, the read coverage implied that the major (GG; HA-Major) or minor (AA; HA-Minor) nucleotides were detected at the frequencies of 75% and 25%, respectively. To validate these variations, HA-Major- or HA-Minor-specific qualitative RT-PCR (qRT-PCR) was performed for the preparation of specific PCR products between the 434 and 738 nt in the HA coding sequence (Fig. 3C). qRT-PCR demonstrated that the expression of HA was ~40,000-fold greater than that of human β-actin, and the expression ratio of HA-Major/HA-Minor was 4.05, suggesting that it corresponds to the read mapping shown in Fig. 3A. Furthermore, HA-Major and HA-Minor sequences were verified by Sanger DNA sequencing of the specific PCR products (Fig. 3D). Taken together, these results suggest the following amino acid substitutions of HA: one nucleotide alteration causes Gly172Glu and the other two alterations cause Gly298Asn (Fig. 3D).
Epitope analysis of heterogeneous HA

To elucidate whether the Gly172Glu and Gly239Asn amino acid substitutions in the HA sequence could be associated with antigenic drift, they were compared to known potential epitopes [8,9]. Representative HA amino acid sequences of the H1N1 influenza A virus were aligned with the heterogeneous HA-Major and HA-Minor sequences. The Gly 172Glu substitution (corresponding to Gly158 in the mature HA lacking a signal peptide) was located on the Sa antigenic site (Fig. 4A).

The HA Gly172Glu substitution is likely to be rare thus far because a BLASTP search against the non-redundant nt database revealed only two identical hits, A/Bayern/62/2009(H1N1) in Germany and A/Catalonia/S1207/2009(H1N1) in Spain (data not shown). One intriguing hit was to A/Pennsylvania/14/2009(H1N1) isolated in the US, whose HA sequence has an Xaa amino acid at position 172 due to the presence of the heterogeneous nucleotide R (A or G) (Fig. 5A). This deposited sequence completely coincides with our observation, suggesting that two variants of HAs are likely to coexist in the human lung, further implying that such heterogeneous populations might frequently be generated during infection.

Furthermore, HA Gly239Asn was located on the Ca2 antigenic site that contributes to binding with the host’s sialic acid receptor [17]. Asp239 (corresponding to Asp225 in the mature HA lacking a signal peptide) was frequently distributed in H1N1 HAs (Fig. 4B), but Gly239 and Asn239 were found to be minor amino acids among HAs; a BLASTN search found 18 and 5 hit entries on the nt database, respectively. As was observed for Gly172, Xaa239
The read coverage profile generated by mapping was very indicative for segment 1 encoding PB2 (Fig. 2). Both ends were highly redundant with up to 3000 coverage. The coverage is reflected by the amounts of vRNA, cRNA, and mRNA of influenza A virus, implying that the coverage bias may detect a more stable region as it is dependent on the expression level or manner of replication.

Contrary to the viral sequences obtained for A/Nagano/RC1-L/2009(H1N1) isolated from passaging in MDCK cells, de novo sequencing revealed the presence of A/Nagano/RC1-L/2009(H1N1) HA quasispecies in the autopsy sample (Fig. 3).

Despite the fact that immunity to A/H1N1/2009 viruses is supposed to be limited among the general human population [19], we detected the amino acid substitution Gly172Glu in the HA Sa antigenic site in A/Nagano/RC1-L/2009(H1N1), and this is not supposed to be limited among the general human population [19].

We also observed another substitution of Gly239Asn in the HA Ca2 antigenic site of A/Nagano/RC1-L/2009(H1N1). This antigenic site plays a crucial role in conferring specificity to galactose of the human α2-6 sialylated glycan receptor [20]. Interestingly, Asp239 (corresponding to Asp225 in the mature HA) acts as a sialic acid receptor [20].

Table 1. BLASTN search results of de novo assembly contigs against database of Influenza virus sequences.

| Euler-SR_contigs | Contig length (bp) | Virus segment | Top hit of accession number using BLASTN search against database of Influenza virus sequences | Length of subject (bp) | Identities | Contig location for A/Toronto/T0106/2009(H1N1) |
|------------------|--------------------|---------------|-----------------------------------------------------------------------------------------------|----------------------|------------|----------------------------------------------|
| >826 183 2835    | 183                | 1             | gb[GQ328865]INFLUENZA A virus (A/Finland/553/2009(H1N1)) segment 1 polymerase PB2 (PB2)         | 2345                 | 167/168    | 5–168                                        |
| >324 1558 136    | 1558               | 1             | gb[GQ365425]INFLUENZA A virus (A/Fukushima/1/2009(H1N1)) segment 1 polymerase PB2 (PB2)       | 2280                 | 1556/1558 | 201–1758                                    |
| >1194 239 112    | 239                | 1             | gb[GQB94926]INFLUENZA A virus (A/Delaware/03/2009(H1N1)) segment 1 polymerase PB2 (PB2)      | 2280                 | 214/214   | 1894–2107                                   |
| >887 174 3294    | 174                | 1             | gb[GQB94833]INFLUENZA A virus (A/Rhode Island/08/2009(H1N1)) segment 1 polymerase PB2 (PB2) | 2280                 | 156/156   | 2145–2300                                   |
| >890 2204 4651    | 2204              | 2             | gb[GQB94924]INFLUENZA A virus (A/New Mexico/04/2009(H1N1)) segment 2 polymerase PB1 (PB1)     | 2274                 | 2200/2204 | 41–2244                                     |
| >696 2198 3968    | 2198              | 3             | gb[GQB66924]INFLUENZA A virus (A/Thailand/CU-H106/2009(H1N1)) segment 3 polymerase PA (PA)   | 2238                 | 2152/2155 | 54–2208                                     |
| >868 1761 3831    | 1761              | 4             | gb[CY045503]INFLUENZA A virus (A/Bayern/66/2009(H1N1)) segment 4 sequence                      | 1754                 | 1750/1754 | 1–1741                                      |
| >897 1514 1710    | 1514              | 5             | gb[GO502907]INFLUENZA A virus (A/Toronto/R8557/2009(H1N1)) segment 5 nucleocapsid protein      | 1558                 | 1511/1514 | 36–1549                                     |
| >224 101 9       | 101               | 6             | gb[GO502908]INFLUENZA A virus (A/Toronto/R8557/2009(H1N1)) segment 6 neuraminidase (NA)       | 1458                 | 101/101   | 3–103                                       |
| >1206 1302 2468    | 1302              | 6             | gb[GO906584]INFLUENZA A virus (A/Stockholm/49/2009(H1N1)) segment 6 neuraminidase (NA)        | 1447                 | 1299/1300 | 124–1423                                   |
| >750 1019 1128    | 1019              | 7             | gb[CY045957]INFLUENZA A virus (A/Toronto/T0106/2009(H1N1)) segment 7 sequence                 | 1026                 | 1017/1017 | 9–1025                                      |
| >809 834 4399     | 834               | 8             | gb[GQ485660]INFLUENZA A virus (A/Ekaterinburg/01/2009(H1N1)) segment 8 nuclear export       | 877                  | 828/830   | 52–881                                      |

Schematic representation of contigs is shown in Fig. 2.
doi:10.1371/journal.pone.0010256.t001
Quasispecies of A/H1N1/2009

Segment 1, PB2 (2346 nt; AC: x368)

Segment 2, PB1 (2321 nt; AC: x208)

Segment 3, PA (2231 nt; AC: x170)

Segment 4, HA (1765 nt; AC: x210)

Segment 5, NP (1562 nt; AC: x236)

Segment 6, NA (1443 nt; AC: x182)

Segment 7, M (1026 nt; AC: x247)

Segment 8, NS (892 nt; AC: x463)
in the human upper airway [22], while it apparently shows minimal contact with α2-3 sialylated glycans present in the human lower respiratory tract [23]. Indeed, the recombinant A/H1N1/2009 HA has been characterized to exhibit lower binding to the alveolar tissue of the lower respiratory tract [17]. However, we previously detected abundant viral nucleoprotein of A/Nagano/2009 HA has been characterized to exhibit lower binding to the alveolar tissue of the lower respiratory tract [17]. However, we previously detected abundant viral nucleoprotein of A/Nagano/

Figure 2. Dot plot of short read coverage (Cov.) at every nucleotide for the 8 segments of A/Nagano/RC1-L/2009(H1N1). To obtain the consensus sequences for the respective 8 segments, 40-mer short reads were aligned to the complete segment sequences of A/Tronto/T0106/2009(H1N1) (gb|CY045951.1 – 8). Short read sequencing was performed using total RNA including human RNA, and also vRNA, cRNA, and mRNA from influenza A virus; thus, coverage bias was detected throughout the segments, but the average coverage (AC) is likely to be similar at approximately >200 or more. The horizontal red arrows show the location of the contigs obtained by de novo assembly, as shown in Table 1.

doi:10.1371/journal.pone.0010256.g002

Figure 3. Genetic variations of the HA nucleotide sequence. (A) Schematic representation of 3 nucleotide variations (positions 515, 715, and 716 nt) in the HA coding nucleotide sequence. Three variations were classified as Major (75% appearance) or Minor (25% appearance) by read coverage (>), and the coding amino acids are also shown. (B) Arrows indicate positions 715 and 716 nt of the HA sequence, and the alignment image of the 40-mer reads. Nucleotides shown in red are the mismatches to the reference sequence of A/Tronto/T0106/2009(H1N1). Every read suggested that either the GGT or AAT sequence was present, but not the GAT or AGT sequence. (C) An amplification plot for HA-specific qRT-PCR. (D) Validation of genetic variation by Sanger capillary sequencing. HA-Major or HA-Minor PCR products were obtained by qRT-PCR using HA-Major- or HA-Minor-specific PCR primers. HA-Major PCR product shows G nucleotides at positions 515, 715, and 716 nt, while HA-Minor shows A nucleotides.

doi:10.1371/journal.pone.0010256.g003
RC1-L/2009(H1N1) in pneumocytes expressing α2-3 sialylated glycans in autopsy lung tissue sections [15], suggesting that the above substitutions could affect the binding affinity to both types of sialylated glycans.

Very suggestive reports predicted the possible future antigenic drift of A/H1N1/2009 viruses from viral sequence and structural comparative analyses [11,12]. Prior to the initiation of the current study (September 2009), Igarashi et al. predicted possible substitutions and these included the two amino acid substitutions presented here (Gly172Glu and Asp239Gly) [11]. Furthermore, Shen et al. suggested that host-driven antigenic drift based on evolutionary trends appeared to favor Asp239 (corresponding to Asp204 in the mature HA) in swine HAs and the 1918 pandemic, while Asp204 (corresponding to Asp190 in the mature HA) was favored in seasonal H1N1 HAs [12]. These predictions are very attractive and our experiments demonstrated one of them a posteriori. Furthermore, recent study has shown that receptor-binding avidity can influence antigenic drift [24]. HA antigenic sites Sa is the membrane proximal region, therefore, the identified variations on both Sa and Ca2 might contribute to the alteration of antigenicity and receptor-binding avidity by synergistic effect. The newly identified Asn239 substitution could be a probable candidate for further investigation of the manner of viral binding to sialic acid on the host receptors.

In conclusion, this study demonstrated that de novo sequencing can comprehensively detect pathogens, and such in-depth

Figure 4. Alignment of HA amino acid sequences in influenza A virus around the identified mutations in A/Nagano/RC1-L/2009(H1N1). (A) Genetic variation at position 515 nt causes the amino acid substitution Gly172Glu; HA-Major: Gly172, HA-Minor: Glu172. (B) Genetic variation at position 715 and 716 nt causes the amino acid substitution Gly239Asn; HA-Major: Gly239, HA-Minor: Asn239.

doi:10.1371/journal.pone.0010256.g004

Quasispecies of A/H1N1/2009
A study investigating influenza A virus heterogeneity during infection facilitated the identification of influenza A viral quasispecies. The possibility of mixed infections with variants remains to be elucidated in this case, but worldwide sequencing efforts suggest that quasispecies of the A/H1N1/2009 virus evidently appear and are observed. To better characterize the currently emerging A/H1N1/2009 virus and prevent worse pandemics in the near future, unbiased de novo sequencing techniques will be indispensable for the primary investigations of emerging infectious diseases.

**Materials and Methods**

**Ethics Statement**

The study protocol was approved by the institutional medical ethical committee, National Institute of Infectious Diseases, Japan (Approval No.236), and the study was conducted according to the Declaration of Helsinki Principles. In the autopsy case, a written consent for autopsy was obtained from relatives.

**Ethics Statement**

The study protocol was approved by the institutional medical ethical committee, National Institute of Infectious Diseases, Japan (Approval No.236), and the study was conducted according to the Declaration of Helsinki Principles. In the autopsy case, a written consent for autopsy was obtained from relatives.

**Total RNA and cDNA preparation from autopsy human lung**

Information for the patient was previously reported [15]. Briefly, in August 2009, a 33-year-old male patient with chronic heart failure due to dilated cardiomyopathy, mild diabetes mellitus, atopic dermatitis, asthma, and obesity (BMI: 38) died from respiratory failure and multiple organ dysfunction syndrome.

**qRT-PCR analysis**

qRT-PCR was performed using 100 ng of total RNA, HA variant-specific primers (forward common primer: pdmFlu09-HA-F, 5’-CGAACAAAGGTGTAACGGCAGCAT–3’; HA-Major reverse primer: pdmFlu-HA-R_Major, 5’-ATAGTTCTATTCTCCGTTGAGCAT–3’; HA-Minor reverse primer: 5’-ATAGTTCTATTCTCCGTTGAGCAT–3’), and
the SuperScript III Platinum SYBR Green One-Step qRT-PCR kit with ROX (Invitrogen), and analyzed using the ABI PRISM 7900HT Real-time PCR System (Applied Biosystems, Foster City, CA, USA). The following qRT-PCR program was used: RT reaction, 50°C for 3 min; initial denaturation, 95°C for 5 min; 2 steps of amplification (x40 cycles); 95°C for 15 sec and 60°C for 30 sec. The human β-actin gene was used as the internal control. PCR products were resolved by 5% polyacrylamide gel electrophoresis, followed by Sanger sequencing using the BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems).

Virus isolation

The A/H1N1/2009 virus was isolated from MDCK cells passaged once with trypsin.

Supporting Information

Text S1  Fastq file of the 40-mer short reads with similarity to influenza A virus extracted from whole obtained short reads.

References

1. Dawood FS, Jain S, Finelli L, Shaw MW, Lindstrom S, et al. (2009) Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N Engl J Med 360: 2605–2615.
2. Gill JR, Sheng ZM, Ely SF, Guiner DG, Beasley MB, et al. (2010) Pulmonary pathologic findings of fatal 2009 pandemic influenza A/H1N1 viral infections. Arch Pathol Lab Med 134: 235–243.
3. Gherlin E, Sengamanay NA, Shumway M, Zaborsky J, FeldibyHum T, et al. (2005) Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. Nature 437: 1162–1166.
4. Kilbourne ED (2006) Influenza pandemics of the 20th century. Emerg Infect Dis 12: 9–14.
5. Nelson MI, Vihoud C, Simonsen L, Bennett RT, Griesemer SB, et al. (2008) Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. PLoS Pathog 4: e1000122.
6. Knossow M, Szebeli J (2006) Variation and infectivity neutralization in influenza. Immunology 119: 1–7.
7. Air GM, Laver WG, Webster RG (1987) Antigenic variation in influenza viruses. Contrib Microbiol Immunol 8: 20–59.
8. Caton AJ, Brownlee GG, Yewdell JW, Gerhard W (1982) The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31: 417–427.
9. Bui HH, Peters B, Assarson E, Mhasawu E, Sette A (2007) Ab and T cell epitopes of influenza A virus, knowledge and opportunities. Proc Natl Acad Sci U S A 104: 246–251.
10. Squires B, Macken C, Garcia-Sastre A, Godbole S, Nonomura J, et al. (2008) BioHealthBase: informatics support in the elucidation of influenza virus host pathogen interactions and virulence. Nucleic Acids Res 36: D97–505.
11. Igarashi M, Ito K, Yoshida R, Tomabechi D, Kida H, et al. (2010) Predicting the antigenic structure of the pandemic (H1N1) 2009 influenza virus hemagglutinin. PLoS One 5: e8553.
12. Shen J, Ma J, Wang Q (2009) Evolutionary trends of A(H1N1) influenza virus hemagglutinin since 1918. PLoS One 4: e7789.
13. Nakamura S, Yang CS, Sakon N, Ueda M, Tougan T, et al. (2009) Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. PLoS One 4: e7129.
14. Ramakrishnan MA, Tu ZJ, Singh S, Chockalingam AK, Guiner DR, et al. (2009) The feasibility of using high resolution genome sequencing of influenza A viruses to detect mixed infections and quasispecies. PLoS One 4: e7105.
15. Nakajima N, Hata S, Sato Y, Tobiume M, Kato H, et al. (2010) The first autopsy case of pandemic influenza (A/H1N1pdm) virus infection in Japan: Detection of high copy number of the virus in type II alveolar epithelial cells by pathological and virological examination. Jpn J Infect Dis 63: 67–71.
16. Qiu W, Hata S, Sato Y, Tobiume M, Katano H, et al. (2010) The first case report of A/H1N1/2009 virus infection in humans. Jpn J Infect Dis 63: 67–71.
17. Palacios G, Hornig M, Cisterna D, Savji N, Bussetti AV, et al. (2009) Staphylococcus pneumoniae coinfection is correlated with the severity of H1N1 pandemic influenza. PLoS One 4: e5549.
18. Maines TR, Jayaraman A, Belser JA, Wadford DA, Pappas C, et al. (2009) Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. Science 325: 484–487.
19. Louie J, Jean C, Chen T-H, Park S, Ueki R, et al. (2009) Bacterial coinfections in lung tissue specimens from fatal cases of 2009 pandemic influenza A (H1N1) - United States, May-August 2009. MMWR Morb Mortal Wkly Rep 58: 1071–1074.
20. Greenbaum JA, Kotturi MF, Kim Y, Oseroff C, Vaughan K, et al. (2009) Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population. Proc Natl Acad Sci U S A 106: 20365–20370.
21. Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, et al. (2009) The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science 305: 1830–1842.
22. Soundararajan V, Tharakaraman K, Raman R, Raguram S, Shriver Z, et al. (2009) Extrapolating from sequence—the 2009 H1N1 ‘swine’ influenza virus. Nat Biotechnol 27: 310–313.
23. Chandraenakaran A, Sinivasan A, Raman R, Viswanathan K, Raguram S, et al. (2008) Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. Nat Biotechnol 26: 107–113.
24. Shinya K, Ebina M, Yamada S, Ono M, Kasai N, et al. (2006) Avian flu: influenza virus receptors in the human airway. Nature 440: 435–436.
25. Henley SE, Das SK, Bailey AL, Schmidt LM, Hickman HD, et al. (2009) Hemagglutinin receptor binding avidity drives influenza virus antigenic drift. Science 326: 734–736.
26. Miura S, Klar B, Hauser DH (2009) Visual and statistical comparison of metagenomes. Bioinformatics 25: 1849–1853.
27. Chaisson MJ, Pevzner PA (2008) Short read fragment assembly of bacterial genomes. Genome Res 18: 324–330.
28. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18: 821–829.
29. Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18: 1851–1858.
30. Bao H, Guo H, Wang J, Zhou R, Lu X, et al. (2009) MapVirc: visualization of short reads alignment on a desktop computer. Bioinformatics 25: 1534–1535.