Gene 464 (2010) 11–16

Contents lists available at ScienceDirect

Gene

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

A phylogenetic approach to detecting reassortments in viruses with segmented genomes

Yoshiyuki Suzuki *

Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken 411-8540, Japan

ARTICLE INFO

Article history:
Received 13 April 2010
Accepted 15 May 2010
Available online 28 May 2010

Received by S. Yokoyama

Keywords:
Reassortment
virus
segmented genome
quartet
phylogenetic tree

ABSTRACT

When multiple strains of viruses with segmented genomes co-infect a single cell, strains with novel genomic constellations may be created. This mutational process, called reassortment, has caused pandemics of influenza A virus in 1957 and 1968. Here a phylogenetic approach to detecting reassortments, which can be used even when the phylogenetic tree constructed for all strains analyzed is unreliable, is presented. A quartet of strains is examined at a time, where a phylogenetic tree is constructed for each genomic segment and the topology is compared among segments only when all quartet trees are supported with a statistical significance. The occurrence of reassortment and the segments involved in the reassortment event are inferred according to the pattern of topological difference among segments. The reassortment point for a pattern is inferred by superimposing the exterior branches of relevant quartet trees on the all-strains trees. In the analysis of H1N1 and H3N2 human influenza A viruses, a topological difference was observed for all pairs of genomic segments, suggesting that there is no pair of segments that has always co-segregated in reassortment during the evolutionary history of these viruses. When the reassortment point was inferred for the pattern of topological difference that was supported with the largest number of quartets for each virus, the results appeared to be mostly correct, suggesting that the method was largely reliable.

1. Introduction

Reassortment is a mutational process of viruses bearing segmented genomes, where, upon co-infection of a single cell by multiple strains, complementary sets of genomic segments are drawn from them during replication to create strains with novel genomic constellations (Hutchinson et al., 2010). By replacing genomic segments through reassortment, viral strains can change phenotypes drastically, which may not be feasible through point mutations. In particular, if viral strains receive genomic segments determining the antigenicity from distantly related strains, the reassortants may acquire a potential to cause pandemics among host organisms that are immunologically naïve to new antigens.

The genome of influenza A virus is eight-segmented, encoding the polymerase basic proteins 2 (PB2, segment 1) and 1 (PB1, segment 2), polymerase acidic protein (PA, segment 3), hemagglutinin (HA, segment 4), nucleoprotein (NP, segment 5), neuraminidase (NA, segment 6), matrix protein (M, segment 7), and non-structural protein (NS, segment 8). It is believed that a virion contains one copy of each segment (Noda et al., 2006). HA and NA determine the antigenicity, which is classified into subtypes H1–H16 and N1–N9, respectively. Human influenza A virus has caused pandemics through reassortment, when an H1N1 strain gained segments 2, 4, and 6 from an avian strain to generate an H2N2 strain in 1957 and when an H2N2 strain gained segments 2 and 4 from an avian strain to generate an H3N2 strain in 1968 (Lindstrom et al., 2004). It is therefore important to clarify the mechanism of reassortment to predict future pandemics.

For this purpose, experimental studies have been conducted to examine the pattern of reassortment in influenza A virus *in vitro*. It has been reported that the reassortment is non-random. For example, the likelihood of each segment to be incorporated into a virion was different among co-infecting strains (Varich et al., 2008). In addition, the incorporation did not appear to be independent among segments (Downie, 2004). In particular, segments 1 and 3 and segments 2 and 3 were reported to be linked during reassortment (Lubeck et al., 1979). Since the viability and replicability of reassortants appeared to be positively correlated with the activity of the ribonucleoprotein complex, which is composed of PB2, PB1, PA, and NP together with a genomic segment (Li et al., 2008, 2010), the compatibility among these proteins was considered to form the basis for the linkage observed among these segments. The direct interaction among genomic segments during packaging may also be a cause of linkage among segments in reassortment. For example, mutations in the packaging signals in segment 1 (Marsh et al., 2008) or 7 (Hutchinson

Abbreviations: PB1(2), polymerase basic protein 1(2); PA, polymerase acidic protein; HA, hemagglutinin; NA, neuraminidase; NP, nucleoprotein; M, matrix protein; NS, non-structural protein.

* Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken 411-8540, Japan. Tel.: +81 55 981 6847; fax: +81 55 981 6848.
E-mail address: yossuzuk@lab.nig.ac.jp.

0378-1119/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2010.05.002

et al., 2008) reduced incorporation of all segments. Mutations in segment 4 reduced packaging of not only itself but also segment 2 (Marsh et al., 2007). Similarly, mutations in segment 3 affected incorporation of segment 5 (and segments 1) (Marsh et al., 2008), and *vice versa* (Hutchinson et al., 2009).

The pattern of reassortments that have actually occurred in the evolution of influenza A virus *in vivo* may be examined by analyzing the sequence data of strains isolated in nature. A simple method to detect reassortment is to compare the topology of phylogenetic trees constructed for different genomic segments. By using this method, it has been inferred that the severe epidemics of H1N1 virus in 1947 and 1951 were caused by reassortments involving segments 1, 3, 4, 5, and 8 and segments 2, 3, 5, 6, 7, and 8, respectively (Nelson et al., 2008). In addition, a reassortment involving segments 1, 3, 4, and 6 was inferred to have occurred in the 2007–2008 season (Furuse et al., 2009; Zaraket et al., 2010). Similarly, the severe epidemic of H3N2 virus over the 2002–2003 season appeared to have been caused by a reassortment involving segment 4 (Barr et al., 2005; Holmes et al., 2005). In this approach, it is required that the topology of the phylogenetic trees is well resolved, e.g., with high bootstrap probabilities. However, this is not always the case, especially in the current situation where the number of sequences available for phylogenetic analysis is large.

To avoid this problem, a distance-based method has been developed for detecting reassortments. Under the assumption that the mutation rate is equal among segments, the occurrence of reassortment was inferred between a pair of strains when the sequence divergence at neutrally evolving sites was different among segments. The p distance at the third codon position was used as the measure of sequence divergence in Rabadan et al. (2008). By using this method, many reassortment events, which were claimed to be independent, have been identified in H3N2 human influenza A virus and H1N1, H1N2, H3N2, and H3N1 swine influenza A viruses (Rabadan and Robins, 2007; Rabadan et al., 2008; Khiabanian et al., 2009a,b). It was proposed that reassortants were produced mainly by drawing one segment from one of co-infecting strains and seven segments from the other, and segments 4 and 6 were involved in the reassortment events more frequently than other segments. However, the assumption of equal mutation rate among segments may not always hold. In addition, since the strains are related phylogenetically, the pattern of reassortment inferred between a pair of strains reflects the summary effect of reassortments on the evolutionary pathway connecting these strains. Therefore, reassortments identified from different pairs of strains may not be independent.

The purpose of the present study was to present a phylogenetic approach to detecting reassortments even when the phylogenetic tree constructed for all strains analyzed is unreliable. The method was applied to H1N1 and H3N2 human influenza A viruses to infer the occurrences, patterns, and points of reassortments.

2. Materials and methods

2.1. Methodology

Previously in the phylogenetic approach to detecting reassortments, a phylogenetic tree was constructed for each genomic segment using all strains analyzed (Barr et al., 2005; Holmes et al., 2005; Nelson et al., 2008; Furuse et al., 2009; Zaraket et al., 2010). Reassortants were identified as a cluster of strains whose phylogenetic position varied among segments. To ensure the occurrence of reassortment, it was required that the cluster and its phylogenetic position were supported with a statistical significance, e.g., a bootstrap probability greater than 95%, in each phylogenetic tree. However, as the number of strains available in the phylogenetic analysis increases due to dense samplings and intensive sequencings of strains, e.g., the Influenza Genome Sequencing Project (Ghedin et al., 2005), branches in the phylogenetic tree may become shortened. As a result, the bootstrap probability may be decreased and the topology may be deteriorated, and the evidence for reassortment may be obscured.

To avoid these problems in the phylogenetic approach, it is necessary to focus on strains whose phylogenetic relationship can be compared among different segments unambiguously. For this purpose, a quartet of strains is chosen and examined at a time in the present study, because a quartet tree supports one of three possible topologies that are mutually inconsistent. The topology of quartet trees is compared among segments only when all of them are supported with a statistical significance. The occurrence of reassortments and the segments involved in the reassortment events are inferred according to the pattern of topological difference. It should be noted that the pattern observed for a quartet reflects the summary effect of reassortments on the evolutionary pathway connecting these strains. The above procedure can be performed for a total of ${}_mC_4$ possible quartets, where m denotes the number of all strains analyzed. Since some quartets share the evolutionary history, the frequency of a pattern observed does not necessarily reflect the number of independent reassortment events. This situation is similar to the case for the distance-based method for detecting reassortments between pairs of strains (Rabadan and Robins, 2007; Rabadan et al., 2008; Khiabanian et al., 2009a,b).

When the occurrence of reassortments is inferred for a quartet, the reassortment points should be located somewhere on the exterior branches of the quartet tree for each segment. Therefore, if a particular pattern of topological difference is observed for multiple quartets, the reassortment points may be inferred by superimposing the exterior branches of relevant quartet trees on the all-strains tree. All relevant quartet trees are expected to be superimposed on the branches containing the reassortment points in the all-strains tree, whether the reassortment point is single or multiple, as long as the pattern commonly observed for these quartets is derived from the same set of reassortment events and the topology of the all-strains tree is correct. However, these conditions may not always hold, and practically the reassortment points may be inferred to be located on the branches where the largest number of quartet trees is superimposed.

2.2. Sequence data

H1N1 and H3N2 human influenza A viruses were used to infer the occurrences, patterns, and points of reassortments by the method presented above. Genomic sequences for 794 strains (6,352 sequences) and 1,699 strains (13,593 sequences) of H1N1 and H3N2 viruses, respectively, excluding laboratory strains and H1N1 pandemic strains in 2009, were downloaded from the Influenza Virus Resource at the National Center for Biotechnology Information (Bao et al., 2008). After excluding redundant strains and sequences, 782 strains (6,256 sequences) and 1,663 strains (13,304 sequences) of H1N1 and H3N2 viruses, respectively, were analyzed. The strain names and the accession numbers in the International Nucleotide Sequence Database for the sequences analyzed in the present study are listed in supplementary Table S1.

2.3. Data analysis

Since the intragenic recombination is considered to be very rare or absent for influenza A virus (Boni et al., 2008; Krasnitz et al., 2008), the phylogenetic analysis is useful for inferring the evolutionary relationship among strains for each genomic segment. A multiple alignment of nucleotide sequences for each segment of 782 H1N1 and 1,663 H3N2 human influenza A viruses was made by using the computer program MUSCLE (version 3.6) (Edgar, 2004). After eliminating the sites with gaps or ambiguous nucleotides, 2,253, 2,263, 2,138, 1,647, 1,478, 1,245, 972, and 816 sites for segments 1–8 of H1N1 virus, respectively, and 2,260, 2,250, 2,129, 1,656, 1,464, 1,348, 954, and 789 sites for segments

1–8 of H3N2 virus, respectively, were used for constructing a phylogenetic tree by the neighbor-joining method (Saitou and Nei, 1987) with the p distance, which is known to produce reliable topologies when a large number of closely related sequences are analyzed (Nei and Kumar, 2000). The reliability of interior branches was assessed by the bootstrap method with 1,000 resamplings (Felsenstein, 1985).

The number of possible quartets that could be used for detecting reassortments was ${}_{782}C_4 = 15{,}462{,}460{,}585$ and ${}_{1{,}663}C_4 = 317{,}533{,}644{,}065$ for H1N1 and H3N2 viruses, respectively. Since it was time consuming to analyze all of them, 10,000,000 quartets were sampled with replacement using pseudo-random numbers. The effect of replacement was considered to be negligible because of the large numbers of possible quartets. For each quartet, a phylogenetic tree was constructed for each of eight genomic segments by the neighbor-joining method with the p distance. The reliability of the interior branch was assessed by the bootstrap method with 100 resamplings. The topology of quartet tree was compared among segments only when the bootstrap probability was greater than 95% for all of them. The difference in the topology was assessed between each pair of segments. The pattern of topological difference among segments was designated as a string of eight digits, each corresponding to segments 1–8 arranged in this order. For example, the string 11121111 indicates that the topology was identical for all segments except for segment 4. It should be noted that in a string only the topological difference is described and the exact value is uninformative.

For a pattern of topological difference, the reassortment points were inferred by superimposing the exterior branches of relevant quartet trees on the all-strains tree for each segment. The reassortment points were inferred to be located on the branches where the largest number of quartet trees was superimposed.

3. Results and discussion

3.1. Inferring the occurrences of reassortments

When a phylogenetic tree was constructed for each of eight genomic segments in 10,000,000 quartets randomly chosen from 782 strains of H1N1 human influenza A virus, the bootstrap probability was greater than 95% for all segments in 4,615,110 quartets. The topology was identical for all segments in most (4,589,220, 99.4%) of these quartets, and a difference was observed for 25,890 (0.6%) quartets. The number of quartets demonstrating a difference in topology between each pair of segments is presented in Table 1. These numbers do not necessarily reflect the frequency of occurrences of independent reassortment events, as indicated above. A topological difference was observed for all pairs of segments, suggesting that there is no pair of segments that has always co-segregated in reassortment during the evolutionary history of H1N1 virus.

In the quartet analysis of 1,663 strains of H3N2 human influenza A virus, the bootstrap probability was greater than 95% for all segments in 4,217,903 of 10,000,000 quartets. The topology was identical for all segments in most (3,905,895, 92.6%) of these quartets, and a difference was observed for 312,008 (7.4%) quartets. Similarly to the case for H1N1 virus, a topological difference was observed for all pairs of segments in H3N2 virus (Table 1).

These results indicate that the reassortment is prevalent in the evolutionary history of H1N1 and H3N2 human influenza A viruses. It should be noted, however, that in the present method a reassortment was detected only when all of the quartet trees for eight genomic segments were supported with a bootstrap probability greater than 95%, suggesting that only the reassortments occurring between relatively distantly related strains were detectable. Since a reassortment occurs between contemporary strains circulating at the same geographical location, where the strains are likely to be relatively closely related, it is plausible that the reassortment in H1N1 and H3N2 viruses is more prevalent than detected in the present study.

3.2. Inferring the patterns of reassortments

The 25,890 and 312,008 quartets for which a signature of reassortment was observed in H1N1 and H3N2 human influenza A viruses, respectively, were classified according to the pattern of topological difference among eight genomic segments. The number of quartets demonstrating each pattern is presented in Table 2. Again, these numbers do not necessarily reflect the frequency of independent reassortment events. It should also be noted that each pattern does not necessarily indicate the pattern of independent reassortment, because the pattern observed for a quartet reflects the summary effect of reassortments on the evolutionary pathway connecting these strains, as indicated above. In fact, for almost all patterns, there existed a pattern that was different at only one segment, implying that an accumulation of single-segmental reassortments can largely explain the observed patterns. Single-segmental reassortments may therefore be common in H1N1 and H3N2 viruses. Segments 4 and 6 were frequently involved in the difference, similarly to the results obtained from the distance-based method for detecting reassortments between pairs of strains (Rabadan and Robins, 2007; Rabadan et al., 2008; Khiabanian et al., 2009a,b). However, this does not necessarily indicate that reassortments involving these segments occurred frequently, because some quartets as well as pairs of strains are related phylogenetically.

3.3. Inferring the points of reassortments

For H1N1 human influenza A virus, the most frequently observed pattern of topological difference among eight genomic segments, supported by 21,804 quartets, was 12112122. This pattern may be generated from a single reassortment event, in which four genomic segments were drawn from each of co-infecting strains. However, it is also possible that multiple reassortments occurred on the evolutionary pathway connecting relevant quartet strains, because single-segmental reassortments appear to be common, as indicated above. For H3N2 human influenza A virus, the most frequently observed pattern was 11121111, which was supported by 291,645 quartets. From the parsimony principle, this pattern was considered to be generated from a single reassortment event involving only segment 4.

Table 1
Numbers of quartets demonstrating a topological difference between pairs of segments in H1N1 (above the diagonal) and H3N2 (below the diagonal) human influenza A viruses.

	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Segment 7	Segment 8
Segment 1		23,412	2,257	3,125	24,221	1,628	23,379	23,475
Segment 2	17,961		23,129	23,508	1,019	24,465	208	298
Segment 3	17,753	244		2,414	22,328	3,719	23,010	23,106
Segment 4	297,109	304,351	304,391		24,258	1,550	23,409	23,505
Segment 5	17,824	322	106	304,468		25,217	900	798
Segment 6	19,657	2,711	2,747	306,032	2,786		24,337	24,433
Segment 7	17,874	369	143	304,499	235	2,878		116
Segment 8	17,784	267	40	304,410	121	2,758	160	

Table 2
Numbers of quartets demonstrating the patterns of topological difference among eight genomic segments in H1N1 and H3N2 human influenza A viruses.

H1N1		H3N2	
Pattern	Number of quartets	Pattern	Number of quartets
11111111	4,589,220	11111111	3,905,895
12112122	21,804	11121111	291,645
11121211	1,259	12212222	12,121
12222122	1,230	12222222	5,193
11212111	744	11111211	2,111
12122222	233	12212122	348
11221111	173	12121211	118
11112112	99	11111121	117
11121111	92	12111111	88
11222222	61	11112111	52
12111111	58	12232122	52
12121211	25	11121311	38
11121222	16	11112311	16
12222111	15	11111112	14
12212222	13	11231122	8
11111121	10	12121311	8
11112111	10	12131311	8
11211111	9	11111212	6
12232222	9	11121211	6
12131111	8	11123211	6
12111113	6	12212232	5
12111211	5	12213122	5
11111211	4	11112112	4
12121222	2	11222222	4
12212111	2	12232222	4
11111112	1	11112211	3
12131333	1	11112212	3
12232111	1	11121121	3
		11211111	3
		12121122	3
		11121131	2
		11231133	2
		12223212	2
		12232322	2
		11112122	1
		11121122	1
		11121212	1
		11211113	1
		12111122	1
		12222122	1
		12222212	1
		12232332	1

Since the numbers of quartets supporting the patterns 12112122 and 11121111 in H1N1 and H3N2 viruses were relatively large, respectively, the reassortment points for these patterns were inferred by superimposing the exterior branches of relevant quartet trees on the all-strains tree for each segment. For the pattern 12112122 of H1N1 virus, all (21,804) relevant quartet trees were superimposed on particular branches in each all-strains tree, suggesting that this pattern was derived from the same set of reassortment events in all quartets (Table 3). In segment 1, all quartet trees were superimposed on two branches that supported the clusters of strains isolated in 2005–2007 and 2008–2009 with bootstrap probabilities 99% and 100%, respectively (Fig. 1). For segment 2, however, only one branch that supported the cluster of strains isolated in 2008–2009 with the bootstrap probability 93% was inferred as containing the reassortment point. Similar results to segment 1 were obtained for segments 3 (bootstrap probabilities 92% and 100%), 4 (96% and 94%), 6 (79% and 71%), and 7 (11% and 38%), and those to segment 2 were obtained for segments 5 (95%) and 8 (97%) (Table 3). The results obtained for different segments varied slightly, probably because of the errors in constructing all-strains trees. However, it was inferred that a reassortment occurred on either one or both of these branches, even though the bootstrap probabilities for these branches were relatively small in some segments.

Interestingly, the cluster of strains isolated in 2008–2009, a candidate cluster of reassortants identified in the present study, apparently corresponded to "clade 2B", which has been identified as reassortants from a small-scale phylogenetic analysis in Zaraket et al. (2010). The "clade 2B" has been inferred to be generated by drawing segments 2, 5, 7, and 8 from "clade 1" and segments 1, 3, 4, and 6 from "clade 2A" or "clade 2C" in the 2007–2008 season (Zaraket et al., 2010), which was consistent with the pattern of topological difference 12112122. The cluster of strains isolated in 2005–2007, another candidate cluster of reassortants identified in the present study, corresponded to "clade 1" in Zaraket et al. (2010). Although the results of Zaraket et al. (2010) have not been supported with a statistical significance, the results obtained in the present study were consistent with Zaraket et al. (2010) and further provided a statistical support, because the reassortments were identified from the comparison of eight quartet trees that were all supported with the bootstrap probability greater than 95%.

Table 3
Statistics for the reassortment points inferred in the all-strains trees of H1N1 and H3N2 human influenza A viruses.

Subtype	Segment	Number of quartets	Bootstrap probability	Isolation year
H1N1	1	21,804	99	2005–2007
		21,804	100	2008–2009
	2	21,804	93	2008–2009
	3	21,804	92	2005–2007
		21,804	100	2008–2009
	4	21,804	96	2005–2007
		21,804	94	2008–2009
	5	21,804	95	2008–2009
	6	21,804	79	2005–2007
		21,804	71	2008–2009
	7	21,804	11	2005–2007
		21,804	38	2008–2009
	8	21,804	97	2008–2009
H3N2	1	288,647	76	2003
	2	288,648	65	2003
	3	288,648	76	2003
	4	288,647	64	2003
	5	288,648	30	2003
	6	284,216	94	2004–2008
	7	288,648	44	2003
	8	284,227	17	2004–2008

For the pattern 11121111 of H3N2 virus, there existed no branch in the all-strains trees where all (291,645) relevant quartet trees were superimposed (Table 3). This is probably because the all-strains trees contained topological errors or because this pattern was derived from different sets of reassortment events in different quartets. However, the majority (284,216–288,648) of relevant quartet trees were superimposed on the branch that supported the cluster of strains isolated in 2003 or 2004–2008 (Fig. 1). The former branch was identified by segments 1, 2, 3, 4, 5, and 7, whereas the latter branch was identified by segments 6 and 8. The bootstrap probabilities for these branches were 76%, 65%, 76%, 64%, 30%, 94%, 44%, and 17% in the all-strains trees for segments 1–8, respectively (Table 3). It was inferred that a reassortment occurred on either one or both of these branches, even though the bootstrap probabilities of these branches were relatively small in all segments.

Similarly to the case for H1N1 virus, the cluster of strains isolated in 2003, a candidate cluster of reassortants identified in the present study, apparently corresponded to "clade A", which has been identified as reassortants from a small-scale phylogenetic analysis in Holmes et al. (2005). The "clade A" has been inferred to be generated by drawing segment 4 from "clade B" and other segments from another clade in the 2002–2003 season (Holmes et al., 2005), which was consistent with the pattern of topological difference 11121111. The cluster of strains isolated in 2004–2008, another candidate cluster of reassortants identified in the present study, corresponded to "clade B" in Holmes et al. (2005). Although the results of Holmes et al. (2005) sometimes have not been supported with a statistical significance, the

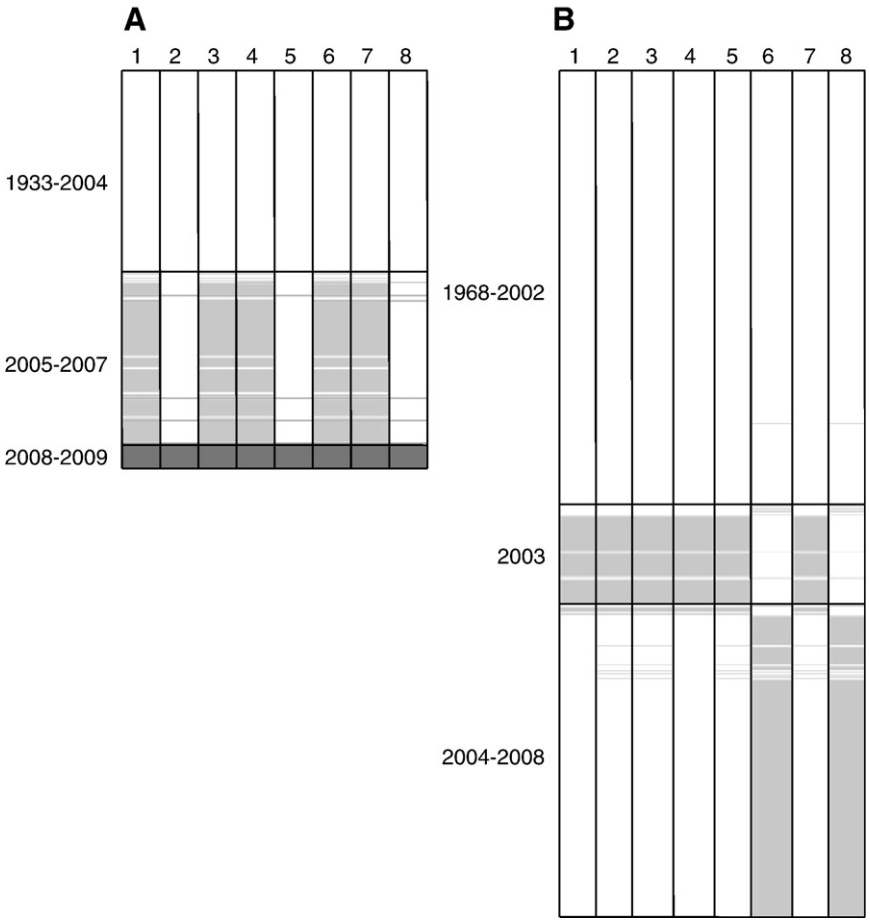

Fig. 1. Clusters of strains inferred as reassortants, shaded in light-gray or dark-gray, for each segment of (A) H1N1 and (B) H3N2 human influenza A viruses. Strains are arranged according to the isolation year, with the earliest strain placed at the top and the latest strain at the bottom. The segment numbers and the isolation years are described at the top and the left-hand side of each panel.

results obtained in the present study were consistent with Holmes et al. (2005) and further provided a statistical support.

The reassortment points inferred for H1N1 and H3N2 viruses appeared to be mostly correct, suggesting that the method presented in this paper was largely reliable. Since the reassortment points can be roughly identified with a statistical significance even when the all-strains trees are unreliable, this method may be useful for detecting reassortments systematically in a large-scale analysis of viruses with segmented genomes.

Acknowledgments

The author thanks Yuki Kobayashi and two anonymous reviewers for valuable comments. The present study was supported by KAKENHI 20570008.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2010.05.002.

References

Bao, Y., et al., 2008. The influenza virus resource at the National Center for Biotechnology Information. J. Virol. 82, 596–601.
Barr, I.G., et al., 2005. An influenza A(H3) reassortant was epidemic in Australia and New Zealand in 2003. J. Med. Virol. 76, 391–397.
Boni, M.F., Zhou, Y., Taubenberger, J.K., Holmes, E.C., 2008. Homologous recombination is very rare or absent in human influenza A virus. J. Virol. 82, 4807–4811.
Downie, J.C., 2004. Reassortment of influenza A virus genes linked to PB1 polymerase gene. Int. Congr. Ser. 1263, 714–718.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.
Furuse, Y., et al., 2009. Reassortment between amantadine-resistant and -sensitive H1N1 influenza A viruses generated an amantadine-sensitive virus during the 2007–2008 season. J. Infect. Dis. 200, 1766–1773.
Ghedin, E., et al., 2005. Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. Nature 437, 1162–1166.
Holmes, E.C., et al., 2005. Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. PLoS Biol. 3, e300.
Hutchinson, E.C., Curran, M.D., Read, E.K., Gog, J.R., Digard, P., 2008. Mutational analysis of *cis*-acting RNA signals in segment 7 of influenza A virus. J. Virol. 82, 11869–11879.
Hutchinson, E.C., Wise, H.M., Kudryavtseva, K., Curran, M.D., Digard, P., 2009. Characterisation of influenza A viruses with mutations in segment 5 packaging signals. Vaccine 27, 6270–6275.
Hutchinson, E.C., von Kirchbach, J.C., Gog, J.R., Digard, P., 2010. Genome packaging in influenza A virus. J. Gen. Virol. 91, 313–328.
Khiabanian, H., Trifonov, V., Rabadan, R., 2009a. Reassortment patterns in swine influenza viruses. PLoS Curr. Influenza RRN1008.
Khiabanian, H., Trifonov, V., Rabadan, R., 2009b. Reassortment patterns in swine influenza viruses. PLoS ONE 4, e7366.
Krasnitz, M., Levine, A.J., Rabadan, R., 2008. Anomalies in the influenza virus genome database: new biology or laboratory errors? J. Virol. 82, 8947–8950.
Li, C., Hatta, M., Watanabe, S., Neumann, G., Kawaoka, Y., 2008. Compatibility among polymerase subunit proteins is a restricting factor in reassortment between equine H7N7 and human H3N2 influenza viruses. J. Virol. 82, 11880–11888.

Li, C., et al., 2010. Reassortment between avian H5N1 and human H3N2 influenza viruses creates hybrid viruses with substantial virulence. Proc. Natl Acad. Sci. USA 107, 4687–4692.
Lindstrom, S.E., Cox, N.J., Klimov, A., 2004. Genetic analysis of human H2N2 and early H3N2 influenza viruses, 1957–1972: evidence for genetic divergence and multiple reassortment events. Virology 328, 101–119.
Lubeck, M.D., Palese, P., Schulman, J.L., 1979. Nonrandom association of parental genes in influenza A virus recombinants. Virology 95, 269–274.
Marsh, G.A., Hatami, R., Palese, P., 2007. Specific residues of the influenza A virus hemagglutinin viral RNA are important for efficient packaging into budding virions. J. Virol. 81, 9727–9736.
Marsh, G.A., Rabadan, R., Levine, A.J., Palese, P., 2008. Highly conserved regions of influenza A virus polymerase gene segments are critical for efficient viral RNA packaging. J. Virol. 82, 2295–2304.
Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, Oxford, New York.
Nelson, M.I., et al., 2008. Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. PLoS Pathog. 4, e1000012.
Noda, T., et al., 2006. Architecture of ribonucleoprotein complexes in influenza A virus particles. Nature 439, 490–492.
Rabadan, R., Robins, H., 2007. Evolution of the influenza A virus: some new advances. Evol. Bioinform. 3, 299–307.
Rabadan, R., Levine, A.J., Krasnitz, M., 2008. Non-random reassortment in human influenza A viruses. Influenza Other Respi. Viruses 2, 9–22.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Varich, N.L., Gitelman, A.K., Shilov, A.A., Smirnov, Y.A., Kaverin, N.V., 2008. Deviation from the random distribution pattern of influenza A virus gene segments in reassortants produced under non-selective conditions. Arch. Virol. 153, 1149–1154.
Zaraket, H., et al., 2010. Genetic makeup of amantadine-resistant and oseltamivir-resistant human influenza A/H1N1 viruses. J. Clin. Microbiol. 48, 1085–1092.