Comparison of the Mutation Rates of Human Influenza A and B Viruses

Eri Nobusawa* and Katsuhiko Sato

Department of Microbiology and Infection, Nagoya City University Graduate School of Medical Science, and Department of Microbiology, Aichi Prefectural Institute of Public Health, Nagoya City, Japan

Received 28 September 2005/Accepted 13 January 2006

Abstract

Human influenza A viruses evolve more rapidly than influenza B viruses. To clarify the cause of this difference, we have evaluated the mutation rate of the nonstructural gene as revealed by the genetic diversity observed during the growth of individual plaques in Madin-Darby canine kidney (MDCK) cells. Six plaques were studied, representing two strains each of type A and B viruses. A total of 813,663 nucleotides were sequenced, giving rates of $2.0 \times 10^{-6}$ and $0.6 \times 10^{-6}$ mutations per site per infectious cycle, which, when extended to 1 year, agree well with the published annual evolutionary rates.

Human influenza A and B viruses (A and B viruses) are equally prevalent among human beings, yet most A virus genes evolve two to three times faster than the corresponding genes in B viruses (11, 13, 15). The high evolutionary rate of A viruses has been attributed to positive selection by the human immune system (2–4, 8, 12), and the difference between A and B to an absence of such selection in B viruses (2, 27). However, a possible alternative explanation of the difference in evolutionary rates, a difference in mutation rates at the molecular level, has not been investigated (27). To correct this lack, we have compared the mutation rates of the nonstructural (NS) genes of several circulating A and B viruses during replication in Madin-Darby canine kidney (MDCK) cells, which are unaffected by the human immune system.

Previously, Parvin et al. reported a mutation rate for the NS gene of influenza A virus A/WSN/33(H1N1) (WSN) of $1.5 \times 10^{-5}$ mutations per nucleotide per infectious cycle during the growth of a single plaque in MDCK cells (23). In their study, Parvin et al. demonstrated that most of the NS mutants they detected replicated with fitness similar to that of the parent virus. This suggests that the MDCK cells did not encourage or discourage the appearance of these mutations, and hence are a suitable medium for the study of mutation rates. We have adopted their method (23), with some modifications.

Since the 1991–1992 and 1992–1993 influenza seasons, the receptor-binding specificity of the A viruses, both H1N1 and H3N2, has changed (19, 20, 21). In particular, the binding affinity of H3N2 viruses to MDCK cells has become weaker (22). To avoid this effect, we have used viruses isolated before this change occurred. All viruses used in this study, A/Aichi/1/87(H1N1) (A/1/87), A/Aichi/12/92(H3N2) (A/12/92), B/Aichi/29/99 (B/29/99), and B/Aichi/44/01 (B/44/01), were directly isolated from clinical samples in MDCK cells.

The mutation rates for the NS genes of A (ANS) and B (BNS) viruses were measured using the procedure illustrated in Fig. 1. In brief, well-isolated parent plaques were picked when they had reached a diameter of approximately 2 mm at 48 and 71 h postinfection for A and B viruses, respectively. The parent plaques were used to produce a series of descendant plaques grown to similar diameters.

The nucleotide sequences of the NS genes in each plaque were determined directly using the reverse transcription-PCR (RT-PCR) technique (Table 1) (21). Mutations were identified by comparing the sequence in randomly selected descendant plaques to the sequence in the corresponding parent plaque using sequence analysis software, GENETYX-MAC ver.11.0 (Genetyx Corporation, Japan). The direct sequencing of RT-PCR products produces an average sequence of many independent copies of the viral cDNA. Although some molecules will be miscopied during the PCR process, they represent a small fraction of the total and do not affect the average sequence (10, 24). All mutations were confirmed by sequencing with the different sense primers listed in Table 1 a total of at least three times.

Most experiments counting the frequency of mutants in viral plaques find a high variability (16, 25, 26). Most of this variability arises from replication errors occurring early in the growth of the plaque. Such random events, when they occur, are amplified to produce large clone lines of identical mutants (Fig. 1), greatly increasing the variance of the number of mutants detected (18, 25). With direct sequencing, it is possible to distinguish each replication error. Counts of the number of unique mutant sequences are related to the number of mutation events, as opposed to the number of progeny resulting from them. Such counts vary only as expected from Poisson statistics. Although no mutant clone lines were detected among the samples taken during the main experiment, in a preliminary test with a B/44/01 parent plaque grown for 96 h, a sample of 169 descendants revealed three mutants. Of these, two exhibited the same mutant sequence, indicating they were very likely clones descended from the same mutation event. By contrast, an A/12/92 parent plaque grown for 72 h contained three different mutant sequences in a sample of 162 descendants.

The results from the main sequencing experiment are listed in Tables 1 and 2. A total of six experiments were performed, three with B viruses (Table 1, experiments 1 to 3) and three with A viruses (Table 1, experiments 4 to 6). For B viruses, a total of two mutations were detected among 461,304 nucleo-
For ANS in the parent plaques was (5 ± \sqrt{2})/461,304 = 0.4 \times 10^{-5} ± 0.3 \times 10^{-5} mutations/site.

For A viruses, the first two experiments with primary plaques formed at 34°C (Table 1, experiments 4 and 5) indicated a mutation frequency far lower than that reported by Parvin et al. (23), who grew their parent plaque at 37°C. To examine the possible effect of temperature on mutation frequency, we remeasured A/1/87 at 37°C. The number of mutations detected was identical to that at 34°C. Taking all three A viruses experiments together, the average mutation frequency for ANS was calculated as 1.4 \times 10^{-5} mutations/site/infectious cycle, and 0.4 \times 10^{-5} mutations/site/infectious cycle, giving an observed mutation frequency for BNS of (2 ± \sqrt{2})/461,304 = 0.4 \times 10^{-5} ± 0.3 \times 10^{-5} mutations/site.

The higher mutation frequency of WSN(H1N1) compared to A/1/87(H1N1) may reflect the peculiar character of WSN, a neurotropic virus, produced by repeated pulmonary and intracerebral infection of mice (9). There may also have been some problem with the technology available in 1986. In the absence of PCR and simplified sequence technology, Parvin et al. (23) were unable to repeat the sequence analysis to confirm mutations by reading both strands of the cDNA. For this reason, their reported number of mutations might not be accurate.

To compare the mutation rate per infectious cycle of the A and B viruses, we determined the generation time of each virus in MDCK cells. As shown in Fig. 2, the mean generation time, when the PFU reach 50% of the final value (1, 7), was about 7 and 10 hours for A and B viruses, respectively. As the dynamics of plaque growth are complicated and the budding of progeny viruses could not be synchronized, these generation times were adopted as probable values during plaque formation. Since the 2-mm-diameter parent plaques were picked at 48 h (A viruses) and 71 h (B viruses) postinfection, they correspond to approximately seven infectious cycles in both cases. Using these values, average mutation rates for ANS and BNS were calculated as 1.4 \times 10^{-5} mutations/site/7 infectious cycles = 2.0 \times 10^{-6} mutations/site/infectious cycle, and 0.4 \times 10^{-5} mutations/site/7 infectious cycles = 0.6 \times 10^{-6} mutations/site/infectious cycle, respectively.

It should be noted that some authors interpret “infectious” and “replication” cycle differently (5, 6, 23, 25, 26). The observed difference in mutation rates is (BNS − ANS) = −(1.4 ± 1.0) \times 10^{-6} mutations/site/infectious cycle, giving a z-score of −1.4 and P = 0.08. This difference in mutation rate presumably reflects a more accurate polymerase in the B virus than in the A virus.

Finally, we attempted to estimate evolutionary rates from the observed in vitro mutation rates in MDCK cells. Since the parent plaques for the A and B viruses were grown for 48 and 71 h, and since a year is 8,760 h long, if the viruses propagated continuously all through the year, we estimate evolutionary rates for the A and B viruses of 1.4 \times 10^{-5} mutations/site/48 h \times 8,760 h = 2.6 \times 10^{-3} mutations/site/year and 0.4 \times 10^{-5} mutations/site/71 h \times 8,760 h = 0.5 \times 10^{-3} mutations/year.

FIG. 1. Measuring the mutation rate during the growth of a single plaque. Plaque formation was performed in confluent monolayers of MDCK cells infected with A and B viruses. A well-isolated parent plaque of approximately 2 mm in diameter was picked at an appropriate time postinfection. To clone selected individual descendant viruses in the parent plaque, a second plaque formation was performed using the virus yield directly from the parent plaque. All plaque viruses were eluted from the agar plug into 300 \mu l of minimal essential medium containing 0.2% bovine serum albumin, made into aliquots, and stored in a deep freezer (−80°C) until used. The harvested viruses were used without further modification for sequential analysis of their NS genes, which reveals the sequence of the major virus in the plaque. The magnified views of the plaques show the viruses that exist within them. Star, viruses carrying NS gene sequences identical to that of the major virus in the parent plaque; triangle, diamond, spade, club, heart, and square, viruses carrying NS gene sequences different from that of the major virus in the parent plaque. The two triangles in the parent plaque represent a clone line. One of the descendant plaques is descended from a mutant of this type.
TABLE 1. Analysis of mutation rate for NS genes of human influenza viruses

| Virus type | Expt no. | Virus | PFU of parent plaque | No. of descendant plaques analyzed | Total no. of nucleotides analyzed | No. of nucleotide changes | Mutation frequency/site | Mutation rate |
|------------|----------|-------|----------------------|-----------------------------------|----------------------------------|---------------------------|-----------------------|---------------|
| A          | 1        | B/Aichi/29/99 | $6.4 \times 10^4$ | 143 | 147,576 | 1 | $0.7 \times 10^{-5}$ | $1.0 \times 10^{-6}$ | $0.8 \times 10^{-3}$ |
|            | 2        | B/Aichi/4/01 | $2.0 \times 10^4$ | 146 | 150,672 | 0 | 0 | 0 | 0 |
| A          | 3        | B/Aichi/4/01 | $4.8 \times 10^6$ | 158 | 163,056 | 1 | $0.6 \times 10^{-5}$ | $0.9 \times 10^{-6}$ | $0.8 \times 10^{-3}$ |
| B          | 4        | A/Aichi/12/92 | $3.0 \times 10^4$ | 149 | 124,117
|            | 5        | A/Aichi/1/87 | $2.4 \times 10^4$ | 136 | 113,288 | 2 | $1.8 \times 10^{-5}$ | $2.6 \times 10^{-6}$ | $3.2 \times 10^{-3}$ |
| B          | 6        | A/Aichi/1/87 | $1.2 \times 10^6$ | 138 | 114,954 | 2 | $1.7 \times 10^{-5}$ | $2.5 \times 10^{-6}$ | $3.2 \times 10^{-3}$ |
| Avg        |          |            |                    | 143 | 147,576 | 1 | $0.7 \times 10^{-5}$ | $1.0 \times 10^{-6}$ | $0.8 \times 10^{-3}$ |

To avoid a change in character of the viruses by adaptation to MDCK cells, we limited the number of virus amplifications in MDCK cells, prior to the experiment, to five. The overlay for plaquing consisted of minimal essential medium with 0.6% agar (Diffco), 0.25% bovine serum albumin, 0.01% DEAE-dextran, 0.1% glucose, and 2.0 μg of acetyl trypsin per milliliter. Viruses were plaque purified twice in MDCK cells before preparing parent plaques. Plaque formation was performed in MDCK cells at 34°C unless otherwise noted. Parent plaques were grown for 48 h for A viruses and 71 h for B viruses.

The nucleotide sequence analysis, the viral RNA was directly extracted from 100 μl of each descendant plaque virus (300 μl eluant) using LipoSpin (Nippon Gene, Japan). Half of this viral RNA was used to synthesize cDNA of the NS genes by RT-PCR as described previously (21) using primers ANSF (+) (5′-AGCAGAAGCAGAGCATTTAGTTTATCATCAA-3′, positions 586 to 890) for A viruses and BNSF (+) (5′-AGCAGAAGCAGAGCATTTAGTTTATCATCAA-3′, positions 1 to 27) and ANSR (-) (5′-AGCAGAAGCAGAGCATTTAGTTTATCATCAA-3′, positions 1073 to 1098) for B viruses. RT-PCR products were used directly for sequence analysis on an ABI 3100 DNA sequencer operated according to the manufacturer’s protocols. For sequence reactions, in addition to the primers indicated above, we used primers 153(-) 5′-GTGTCATTCTCTTTACTGTC-3′, positions 637 to 660 and 153(-) 5′-AGGCTTTTCACCAGAAGGGGAGCAT-3′, positions 467 to 492) for A viruses and 153(-) 5′-CCAACCTCTCAGAAGAACACAT-3′, positions 586 to 609) and 172(-) 5′-GACATGAAAACAAAGAGTGAAG-3′, positions 489 to 511) for B viruses. AAGCAGGGTGACAAAGACATAA-3′, positions 637 to 660) and 153(-) 5′-AGGCTTTTCACCAGAAGGGGAGCAT-3′, positions 467 to

 rate of number of nucleotide changes detected to number of nucleotides analyzed.

Following Parvin et al. (23), the mutation rate was determined by dividing the observed frequency of nucleotide mutations per site by the number of infectious cycles. Calculated by correcting observed mutation frequencies over 71 h (B viruses) and 48 h (A viruses) to 365 days (8,760 h).

For each descendant plaque, the 833 nucleotides at positions 28 to 860 of the NS gene were analyzed.

The numbers in parentheses show the evolutionary rates estimated from the nucleotide substitutions/site/year, respectively. Evolutionary rates for the NS genes of human influenza A and B viruses have been reported in the range of $1.8 \times 10^{-3}$ to $2.2 \times 10^{-3}$ and $0.45 \times 10^{-3}$ to $0.8 \times 10^{-3}$ nucleotide substitutions/site/year, respectively, on the ba-

TABLE 2. Mutations observed in NS genes of descendant plaque viruses compared to the parent plaque

| Expt no. | Virus | Clone no. | Mutation (position) | Amino acid change (position) | No. of nucleotide differences between parent and reference virus |
|----------|-------|-----------|---------------------|-------------------------------|---------------------------------------------------------------|
| 1        | B/29/99 | 3 G       | T→G (238)           | S→I (NS1, 65)                 | 26 (0.75 $\times 10^{-3}$)                                     |
| 2        | B/44/01 | 8 T       | T→A (678f)          | Y→N (NS1, 212)                | 42 (1.1 $\times 10^{-3}$)                                     |
| 3        | A/12/92 | 53 C      | C→A (339)           | L→I (NS1, 105)                | 45 (2.25 $\times 10^{-3}$)                                    |
| 4        | A/1/87  | 93 T→C    | (121)               | F→A (NS1, 32)                 | 45 (2.8 $\times 10^{-3}$)                                     |
| 5        | A/1/87  | 136 G→A   | (849)               | (NS1, 100)                    | (NS1, 100)                                                    |
| 6        | A/1/87  | 52 A→G    | (355)               | K→R (NS1, 110)                | (NS1, 110)                                                    |
|          |        | 125 A→G   | (665)               | H→R (NSF, 56)                 | (NSF, 56)                                                     |

The numbers of the experiments are according to those in Table 1.

The nucleotide number is according to A/Aichi/2/87 (M34829) for A viruses and B/Ann Arbor/1/66 (M20225) for B viruses (the reference viruses). The numbers in parentheses are GenBank accession numbers.

The NS gene codes for two overlapping proteins, NS1 and NEP (nuclear export protein). The amino acids are numbered according to the proteins in the reference viruses listed in footnote b.

The numbers in parentheses show the evolutionary rates estimated from the differences between the reference virus and parent plaque, expressed in nucleotide differences per site per year. These rates are close to those reported by others (see text).

B/44/01, like several other B viruses, is missing the noncoding nucleotide at position 40 in B/Ann Arbor/1/66. The remaining nucleotides are numbered as if the omission had not occurred.

FIG. 2. Comparison of the generation times of A and B viruses in MDCK cells. Confluent monolayers of MDCK cells (5 × 10^5 cells) were infected with A/Aichi/1/87 and B/Aichi/29/99 at a multiplicity of infection of 8, kept at room temperature for 1 h, washed three times with minimal essential medium, and then incubated in 500 μl of overlay medium (Table 1, footnote d) excluding agar and DEAE-dextran, at 34 or 37°C. At 0 and 5 h postinfection and at each subsequent hour, a sample of the supernatant was removed and analyzed by plaque titration in MDCK cells. The PFU per 500 μl are shown versus the hours postinfection. (Inset) The generation times were determined as the point of maximum release of PFU per hour in each one-step growth experiment (1, 7). This is generally equivalent to the point at which the PFU reaches 50% of its final value, estimated here by averaging the total PFU at 11, 12, and 14 h. The width of the distribution about the mean was likewise estimated by noting the points at which the PFU reached 16 and 84% of its final value, corresponding to plus and minus 1 standard deviation (1). The change in temperature from 34 to 37°C, tested for A/1/87 only, appeared to have no significant effect on the mean generation time.
sis of nucleotide sequencing of circulating viruses (13, 15, 17). Although the generation time of viruses and the selection factors on NS gene products in vivo may not be the same as those in MDCK cells, our estimated evolutionary rates are in good agreement with the reported actual rates. Whether this agreement is a consequence of the random fixation of neutral mutations (14), of some other origin, or purely an accident remains to be solved.

The current study suggests that one of the reasons for the lower evolutionary rate of influenza B viruses could be a lower mutation rate.

We are greatly indebted to J. Mosher for helpful suggestions and valuable discussion. K. Nakajima and S. Itamura also provided critical discussions. Finally, we are grateful to N. Murayama for excellent technical assistance.

This work was supported in part by a Grant-in-Aid for Research in Nagoya City University.

REFERENCES

1. Adams, M. H., and F. E. Wassermann. 1956. Frequency distribution of phage release in the one-step growth experiment. Virology 2:96–108.
2. Air, G. M., A. J. Gibbs, W. G. Laver, and R. G. Webster. 1990. Evolutionary changes in influenza B are not primarily governed by antibody selection. Proc. Natl. Acad. Sci. USA 87:3884–3888. (Erratum, 87:6607.)
3. Buonagurio, D. A., S. Nakada, J. D. Parvin, M. Krystal, P. Palese, and W. M. Fitch. 1986. Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene. Science 232:980–982.
4. Cox, N. J., and C. A. Bender. 1995. The molecular epidemiology of influenza viruses. Semin. Virol. 6:359–370.
5. Drake, J. W., and J. J. Holland. 1999. Mutation rates among RNA viruses. Proc. Natl. Acad. Sci. USA 96:13910–13913.
6. Elena, S. F., and A. Moya. 1999. Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. J. Evol. Biol. 12:1078–1088.
7. Ellis, E. L., and M. Delbrück. 1939. The growth of bacteriophages. J. Gen. Physiol. 22:365–384.
8. Fitch, W. M., J. M. Leiter, X. Q. Li, and P. Palese. 1991. Positive Darwinian evolution in human influenza A viruses. Proc. Natl. Acad. Sci. USA 88:4270–4274.
9. Francis, T., Jr., and H. E. Moore. 1940. A study of the neurotropic tendency in strains of the virus of epidemic influenza. J. Exp. Med. 72:717–728.
10. Green, P. M., and F. Giannelli. 1994. Direct sequencing of PCR-amplified DNA. Mol. Biotechnol. 1:117–124.
11. Hiromoto, Y., T. Saito, S. E. Lindstrom, Y. Li, R. Nerome, S. Sugita, M. Shinjoh, and K. Nerome. 2000. Phylogenetic analysis of the three polymerase genes (PB1, PB2 and PA) of influenza B virus. J. Gen. Virol. 81:929–937.
12. Ina, Y., and T. Gojobori. 1994. Statistical analysis of nucleotide sequences of the hemagglutinin gene of human influenza A viruses. Proc. Natl. Acad. Sci. USA 91:8388–8392.
13. Kawaoka, Y., O. T. Gorman, T. Ito, K. Wells, R. O. Donis, M. R. Castrucci, I. Donatielli, and R. G. Webster. 1998. Influence of host species on the evolution of the nonstructural (NS) gene of influenza A viruses. Virus Res. 55:143–156.
14. Kimura, M. 1991. Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics. Proc. Natl. Acad. Sci. USA 88:5969–5973.
15. Lindstrom, S. E., Y. Hiromoto, H. Nishimura, T. Saito, R. Nerome, and K. Nerome. 1999. Comparative analysis of evolutionary mechanisms of the hemagglutinin and three internal protein genes of influenza B virus: multiple cocirculating lineages and frequent reassortment of the NP, M, and NS genes. J. Virol. 73:4413–4426.
16. Lubeck, M. D., J. L. Schulman, and P. Palese. 1980. Antigenic variants of influenza viruses: marked differences in the frequencies of variants selected with different monoclonal antibodies. Virology 102:458–462.
17. Ludvig, S., U. Schultz, J. Mandler, W. M. Fitch, and C. Scholtissek. 1991. Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses. Virology 183:566–576.
18. Murayama, N. 1988. Influenza B virus evolution: cocirculating lineages and comparison of evolutionary pattern with those of influenza A and C viruses. Virology 163:112–122.