Analysis of the Evolution and Variation of the Human Influenza A Virus Nucleoprotein Gene from 1933 to 1990

L. L. SHU, W. J. BEAN, AND R. G. WEBSTER*

*Department of Virology and Molecular Biology, St. Jude Children's Research Hospital,
332 North Lauderdale, P.O. Box 318, Memphis, Tennessee 38101*

Received 23 November 1992/Accepted 27 January 1993

This study examined the evolution and variation of the human influenza virus nucleoprotein gene from the earliest isolates to the present. Phylogenetic reconstruction of the most parsimonious evolutionary path connecting 49 nucleoprotein sequences yielded a single lineage. The average calculated rate of mutation was 3.6 nucleotide substitutions per year (2.3×10^{-3} substitutions per site per year). Thirty-two percent of these mutations resulted in amino acid substitutions, and the remainder were silent mutations. Analysis of virus isolates from China and elsewhere showed no significant differences in their rate of evolution, genetic diversity, or mean survival time. The nearly constant rate of change was maintained through the two antigenic shifts, and there were no obvious changes in the number or types of mutations associated with the changes in the surface proteins. A detailed comparison of the changes that have occurred on the main evolutionary path with those that have occurred on the side branches of the phylogenetic tree was made. This showed that while 35% of the mutations on the side branches resulted in amino acid changes, only 21% of those on the main path affected the protein sequence. These results suggest that although the rate of change of the human influenza virus nucleoprotein is much higher than that previously described for avian influenza viruses, there are measurable constraints on the evolution of the surviving virus lineage. Comparison of the nucleoproteins of virus isolates adapted to chicken embryos with the nucleoproteins of those grown only in MDCK cells revealed no consistent differences between the virus pairs. Thus, although the nucleoprotein is known to be critical for host specificity, its adaptation to growth in eggs apparently involves no immediate selective pressures, such as are found with hemagglutinin.

The influenza virus nucleoprotein (NP) gene codes for a major structural protein of the virus and, in conjunction with other viral genes, has been shown to be a major determinant of host specificity. Scholtissek et al. (21) showed a critical host-specific role of the NP by demonstrating that the pathogenicity of a virulent avian virus depended on the presence of an NP of avian virus origin and could not be replaced with an NP from a human virus. Bean (2) showed by RNA-RNA hybridization that the NP genes from influenza viruses infecting different species could be divided into five host-specific groups: two equine, two avian, and one containing all human and swine virus isolates. Subsequent sequence analysis of the NPs of numerous influenza virus isolates from various host species allowed detailed analysis of the interrelationships of the five lineages (1, 14) and has shown, for example, that the avian viruses can be subdivided by geographic region and that the human and swine viruses both show a continuous evolutionary progression with time. Most significantly, the evidence now suggests that all of the lineages were derived from viruses similar to those now circulating in waterfowl, which appear to be the original host for the influenza A viruses. In waterfowl, the amino acid sequence of the NP has been shown to be very well conserved (1, 14). In this host, the NP of the virus appears to have achieved an adaptive optimum, and deviation from this optimum is selected against. Transfer of the virus to another host leads to the accumulation of mutations at a much higher rate (1, 13, 14).

Analysis of the rate of accumulation of mutations in the human and swine viruses has suggested that their common

ancestor was an avian virus introduced into either swine or humans around the beginning of this century. Three different subtypes of influenza A viruses (H1N1, H2N2, and H3N2) have circulated in humans since 1933. In 1957 the H1N1 subtype was replaced by H2N2, and in 1968 the H3N2 subtype replaced H2N2. Each of the new antigenic subtypes appears to have originated in Southern China (23, 24), and each has maintained the same NP gene despite changes in the surface proteins (2, 14, 22).

This study was designed to follow the evolution of the human influenza virus NP gene from the earliest isolates to the present and to determine the rate of accumulation of mutations and the nature of these changes. A series of virus isolates from China and Memphis, Tenn., were included in this analysis to determine whether the Chinese virus isolates showed greater diversity or survival time and whether differences in the rate of accumulation of mutations could be detected.

Because of the known importance of the NP in host specificity and the well-established systematic mutations that occur in the hemagglutinin gene when human viruses are adapted to grow in chicken embryos, NP genes from a series of virus isolates adapted to both mammalian cell culture and chicken embryo culture were sequenced to determine whether equivalent host cell-selected mutations also occur in the NP gene.

MATERIALS AND METHODS

Virus strains. A total of 34 viral isolates were selected from the repository at St. Jude Children's Research Hospital, Memphis, Tenn. The isolates were selected to allow a comparison with similar numbers of isolates from China and

* Corresponding author.

TABLE 1. Human influenza virus strains used in phylogenetic analyses

Strain	GenBank accession no.	Abbreviation	Source or reference
A/Wilson-Smith (H1N1)		WS33	Gorman et al. (14)
A/Puerto Rico/8/34 (H1N1)		PR834	Winter and Fields (25)
A/Hickox/40 (H1N1)		HICKOX40	Gorman et al. (13)
A/Ft. Monmouth/47 (H1N1)		FM47	Gorman et al. (13)
A/Ft. Warren/50 (H1N1)		WARREN50	Altmuller et al. (1)
A/England/19/55 (H1N1)		ENG55	Gorman et al. (13)
A/Singapore/1/57 (H2N2)		SING57	Gorman et al. (13)
A/Ann Arbor/6/60 (H2N2)		AARBOR60	Cox et al. (9)
A/Victoria/5/68 (H2N2)		VIC68	Gorman et al. (13)
A/NT/60/68 (H3N2)		NT6068	Huddleston and Brownlee (15)
A/Beijing/1/68 (H3N2)	L07340	BJ168	This report
A/Memphis/1/71 (H3N2)	L07341	M171	This report
A/Memphis/2/72 (H3N2)	L07342	M272	This report
A/Memphis/102/72 (H3N2)	L07343	M10272	This report
A/Guangdong/243/72 (H3N2)	L07344	GD24372	This report
A/Udorn/307/72 (H3N2)		UDORN72	Buckler-White and Murphy (5)
A/Memphis/101/72 (H3N2)	L07345	M10172	This report
A/Memphis/1/73 (H3N2)	L07346	M173	This report
A/Memphis/4/73 (H3N2)	L07347	M473	This report
A/Beijing/39/75 (H3N2)	L07358	BJ3975	This report
A/Memphis/110/76 (H3N2)	L07349	M11076	This report
A/USSR/77 (H1N1)		USSR77	Altmuller et al. (1)
A/Texas/1/77 (H3N2)		TEXAS77	Altmuller et al. (1)
A/Guangdong/38/77 (H3N2)	L07359	GD3877	This report
A/Memphis/2/78 (H3N2)	L07350	M278	This report
A/Memphis/18/78 (H3N2)	L07351	M1878	This report
A/California/10/78 (H1N1)		CALIF78	Altmuller et al. (1)
A/Memphis/4/80 (H3N2)	L07352	M480	This report
A/Memphis/5/80 (H3N2)	L07360	M580	This report
A/Shanghai/31/80 (H3N2)	L07361	SH3180	This report
A/Hong Kong/5/83 (H3N2)		HK583	Gammelin et al. (12)
A/Memphis/2/85 (H3N2)	L07362	M285	This report
A/Memphis/5/85 (H3N2)	L07363	M585	This report
A/Memphis/14/85 (H3N2)	L07364	M1485	This report
A/Memphis/6/86 (H3N2)	L07365	M686	This report
A/Guangdong/9/87 (H3N2)	L07366	GD987	This report
A/Shanghai/11/87 (H3N2)	L07367	SH1187	This report
A/Memphis/2/88 (H3N2)	L07368	M288	This report
A/Memphis/3/88 (H3N2)	L07369	M388	This report
A/Memphis/8/88 (H3N2)	L07370	M888	This report
A/Memphis/12/88 (H3N2)	L07371	M1288	This report
A/Shanghai/16/89 (H3N2)	L07372	SH1689	This report
A/Guangdong/39/89 (H3N2)	L07373	GD3989	This report
A/Beijing/337/89 (H3N2)	L07374	BJ33789	This report
A/Beijing/352/89 (H3N2)	L07353	BJ35289	This report
A/Beijing/353/89 (H3N2)	L07354	BJ35389	This report
A/Memphis/1/90 (H3N2)	L07355	M190	This report
A/Memphis/6/90 (H3N2)	L07356	M690	This report
A/Shanghai/6/90 (H3N2)	L07357	SH690	This report

the rest of the world. Another 15 gene sequences from the literature and from data bank sources were combined for our evolutionary analysis (Table 1).

Ten virus strains isolated in Memphis from 1985 to 1990 for which both chicken embryo and MDCK cell isolates were available were used for the host cell variation analysis (see Table 2) (17).

Molecular cloning of the NP gene. Viruses were grown in 11-day-old embryonated eggs, and RNA was extracted as described by Bean et al. (4). Full-length NP genes were cloned by the procedures of Jones et al. (16) and Winter et al. (26). Briefly, a 12-mer oligodeoxynucleotide primer (5' AGC AAAAGCAGG) was phosphorylated with T4 polynucleotide kinase, and first-strand cDNA was synthesized from a viral RNA template by using avian myeloblastosis virus reverse

transcriptase. Second-strand synthesis was done with a phosphorylated 13-mer oligodeoxynucleotide primer (5' AG TAGAAACAAGC) and *Escherichia coli* DNA polymerase I (Klenow fragment). Full-length double-stranded cDNA was blunt-end ligated into the *Pvu*II site of vector pATX (a derivative of pAT153, kindly provided by Clayton Naeve).

Polymerase chain reaction amplification. Sequence analysis of pairs of egg-grown and MDCK cell-grown viruses was done by polymerase chain reaction amplification of first- or second-passage allantoic fluid diluted 1/1,000 or undiluted cell culture fluid. Viral RNAs were extracted, reverse transcribed into cDNA, and amplified by using 25-mer primers from nucleotide positions 17 to 1503, as described by Katz et al. (17). The amplification proceeded through a total of 35 cycles of denaturation at 94°C (1 min), annealing at 50°C (1

min), and primer extension at 74°C (3 min). All experiments were performed with negative controls in which mock RNA extraction and cDNA synthesis had been carried out on samples of sterile water. Polymerase chain reaction products were gel purified (Geneclean; Bio 101, Inc.) and treated with Klenow fragment at 37°C for 30 min to ensure blunt ends, extracted with phenol-chloroform, and ligated into pATX as described above.

Nucleotide sequence determination. The nucleotide sequences of NP cDNAs were determined by the dideoxynucleotide chain termination method (20). Oligodeoxynucleotide primers were annealed to double-stranded template DNA denatured with NaOH, as described by Chen and Seeburg (8), and extended with modified T7 DNA polymerase (Sequenase; U.S. Biochemical Corp., Cleveland, Ohio). The reaction products were separated on 6% polyacrylamide-7 M urea 0.4-mm-thick gels. For each of the virus strains, two or more clones were sequenced.

Sequence analysis. The Intelligenetics (Palo Alto, Calif.) software package was used for analysis and translation of nucleotide sequence data. Phylogenetic analysis of sequence data was performed with the PAUP software package, version 2.4.1 (David Swofford, Illinois Natural History Survey, Champaign), as described previously (3). The maximum-parsimony method was used in the analyses because it provides a basis for the reconstruction of the ancestral strains. This reconstruction was done with the algorithm described by Fitch (11).

RESULTS

Analysis of nucleotide sequences. Each of the 34 cloned NP genes had a single open reading frame which spanned positions 46 through 1539 and coded for a polypeptide of 498 amino acids. While most of the clones had a total of 1,565 nucleotides, several had an insertion of an extra A in the noncoding region at position 1545. The significance of this is not known.

Figure 1 shows the amino acid sequences for the 49 strains, as deduced from the nucleotide sequence data. Inspection of these sequences reveals two relatively large conserved regions that have not changed in any of the 49 virus isolates. One is a 28-amino-acid region from positions 35 to 62, and the other is a 39-amino-acid region from 147 to 185. There are also two hot spots with multiple mutations. These are at position 101, where four independent substitutions at different times during the historical period of the study have changed Asp to Gly and four have changed Asp to Asn, and at position 353, where multiple changes have introduced five different amino acids.

Phylogenetic analysis of the 49 human influenza A virus NP gene nucleotide sequences gave a minimum tree length of 522 nucleotide and 152 amino acid changes. Many trees (>100) of this length were found because of a number of very short and zero-length branches among the closely related strains. All shared a single main lineage which has short side branches. To determine the changes that have occurred on the main lineage, which was assumed to begin at the junction of the A/Wilson-Smith/33 and A/Puerto Rico/8/34 strains and end at the most recent internal node (Fig. 2), the sequences of the internal nodes were reconstructed and the locations of all coding and silent mutations were determined. The main lineage requires 192 nucleotide substitutions, of which 40 (21%) affect the amino acid sequence. These are shown in Fig. 2. In contrast, 35% of the substitutions in the side branches resulted in amino acid changes.

The 40 amino acid changes that have occurred on the main evolutionary path affect 36 positions. Of these amino acid changes, 20 are nonconservative, with 6 affecting the charge at 5 positions and 14 affecting the polarity at 11 positions. Of these 16 positions, 4 have had two changes in the main lineage, each of which was a back mutation to the ancestral type (Fig. 2).

Estimation of evolution rates. On the basis of this phylogenetic reconstruction, the rate of evolution of the NP gene was calculated by plotting the year of isolation for each virus sequence against the total distance (number of nucleotide changes) from the junction of the two earliest virus isolates. This is shown in Fig. 3. The average rate of change is 3.57 nucleotides per year (2.3×10^{-3} substitutions per site per year). The average rate of amino acid substitution is 1.03 per year (2.1×10^{-3} substitutions per site per year).

Variation analysis. Sequence analysis of egg- and MDCK cell-isolated viruses was designed to determine whether any systematic differences in the NP gene equivalent to those found in the hemagglutinin gene could be detected. For each virus pair, three clones from the chicken embryo isolate and three clones from the MDCK isolate were sequenced to make certain that any differences reflected the original population and were not artifacts of the amplification and cloning procedure. When the same viruses were cultured in MDCK cells and eggs, 8 of the 10 pairs of viruses had the same NP gene sequences. Two of the pairs, the A/Memphis/1/90 and A/Memphis/8/88 viruses, had both silent and coding changes in the sequence. These are detailed in Table 2.

DISCUSSION

Previous studies have shown the host-specific nature of NP gene evolution, the very high level of phenotypic conservation of the NP gene in avian hosts, and the divergence of NP genes of the swine and human viruses from their proposed avian ancestor (1, 14). These studies looked at the evolution of the human virus NP gene to determine the nature of the changes that have occurred during nearly six decades of virus evolution and to determine whether any differences could be detected in the rate of mutation or in the pattern of diversity in viruses from the proposed epicenter of virus generation in Southern China.

The rate of accumulation of mutations in the NP gene has been remarkably constant, despite the replacement of the genes for hemagglutinin (twice) and neuraminidase (once) during the emergence of the 1957 Asian and 1968 Hong Kong pandemics. One would expect that the accumulation of mutations would have some biological effect. There are at least three possibilities. The substitutions could be part of a continuing process of selection. Alternatively, the progressive accumulation of mutations could be due to neutral or deleterious mutations that are fixed by genetic drift. The latter may arise through the periodic cloning of the virus in the human population. This scenario was postulated for the nonstructural protein (6) which also showed a linear, but slightly lower, rate of accumulation of nucleotide substitutions (1.94×10^{-3} substitutions per site per year). If the mutations were deleterious, one could expect a decrease in the fitness of the virus (7). Neither the behavior of the virus in the human population nor the characterization of these virus isolates in the laboratory has suggested that fitness is either increasing or decreasing, but this could be a reflection of our inability to measure fitness as it applies to a virus in global circulation. In 1977 an H1N1 virus appeared in humans, and it was shown to be a reintroduction of a virus

FIG. 1. Alignment of 49 influenza virus NP amino acid sequences. The first identified human influenza virus, WS33, was used as a baseline. Only the differences from this baseline are shown. Viruses are arranged according to date of isolation.

FIG. 2. Proposed phylogenetic tree for the 49 nucleotide sequences analyzed in this study. The branch length is proportional to the number of nucleotide substitutions. Amino acid substitutions affecting the main lineage are indicated to the right of the main lineage on which the mutations occurred. The earlier amino acids, the amino acid position numbers, and the later amino acids are shown. Those followed by an asterisk have changed twice in the main lineage.

that had circulated in 1950 (19). Both the H3N2 and the H1N1 viruses have survived in the human population since that time. As noted by Donis (10), the earlier virus did not replace the later one, suggesting that the accumulated substitutions were either neutral or nearly neutral, giving neither virus a selective advantage. If we make the assumption that the human virus NP can accumulate functionally neutral amino acid substitutions, the lack of change in the avian virus NP becomes even more puzzling. If these substitutions can be tolerated in human viruses, why can they not be tolerated in avian viruses?

The comparison of the virus isolates from China with those from elsewhere detected no differences in the rate of

FIG. 3. Accumulation of mutations in nucleotide and amino acid sequences with time. The total number of mutations from the junction of the two earliest strains and their positions on the phylogenetic tree are plotted against the year of virus isolation. Strain A/USSR/77 has been plotted at 1950 because of earlier evidence of its close antigenicity and genetic relationship to viruses from that year (18, 19). Open circles, virus from elsewhere; closed circles, virus from China.

mutation, in genetic diversity, or in survival time of the viruses on the terminal branches. Within our sample, all of the virus lineages found in China also contained isolates from elsewhere. We cannot rule out the existence of other lineages surviving in isolated populations that have not been sampled or were not included in this analysis, but it is apparent that the predominant strains circulate worldwide and there are no obvious geographical differences in the distribution of these strains.

Comparison of the NPs of virus isolates adapted to chicken embryos with the NPs of those grown only in MDCK cells revealed no consistent differences between the virus pairs. Thus, although the NP is known to be critical for host specificity, its adaptation to growth in eggs apparently involves no immediate selective pressures, such as are found with hemagglutinin. Although the role of the NP in determining host specificity was established in earlier studies (21), we do not know at what level the constraint on species specificity acts. The adaptation of the NP gene to favor growth or pathogenicity in a particular host could be an adaptation to the host cell environment, or it may be an adaptation of the NP gene to other components of the virion that have been selected for a particular host. All of the studies on the evolution and variation of influenza viruses have been restricted to analysis of a single gene or gene

TABLE 2. Detailed differences of A/Memphis/1/90 and A/Memphis/8/88 viruses grown in MDCK cells and eggs^a

Virus	Nucleotide position	Nucleotide		Amino acid position	Amino acid	
		MDCK cells	Eggs		MDCK cells	Eggs
A/Memphis/1/90	417	T	C	230	Leu	Phe
	733	C	T			
	1371	C	T			
A/Memphis/8/88	232	A	G	63	Ile	Val
	347	A	G	101	Asp	Gly
	1338	A	G			

^a Eight other virus isolate pairs show no differences: A/Memphis/6/86, A/Memphis/2/88, A/Memphis/12/88, A/Memphis/6/90, A/Memphis/2/85, A/Memphis/5/85, A/Memphis/3/88, and A/Memphis/14/85.

product. This has been sufficient for some viral characteristics. However, analysis of the complex biological properties of the virus-host interaction may require a coordinated study of multiple virus genes.

ACKNOWLEDGMENTS

We thank William Meyer and Jacqueline Katz for productive influenza viruses isolated in MDCK cells and Dayna Anderson for manuscript preparation.

This work was supported by Public Health Service grants AI-20591, AI-08831, and AI-27497 from the National Institute of Allergy and Infectious Diseases, by Cancer Center Support grant (CORE) CA-21765, and by the American Lebanese Syrian Associated Charities.

REFERENCES

- Altmuller, A., W. M. Fitch, and C. Scholtissek. 1989. Biological and genetic evolution of the nucleoprotein gene of human influenza A viruses. *J. Gen. Virol.* **70**:2111-2119.
- Bean, W. J. 1984. Correlation of influenza A virus nucleoprotein genes with host species. *Virology* **133**:438-442.
- Bean, W. J., M. Schell, J. Katz, Y. Kawaoka, C. Naeve, O. Gorman, and R. G. Webster. 1992. Evolution of the H3 influenza virus hemagglutinin from human and nonhuman hosts. *J. Virol.* **66**:1129-1138.
- Bean, W. J., G. Sriram, and R. G. Webster. 1980. Electrophoretic analysis of iodine-labeled influenza virus RNA segments. *Anal. Biochem.* **102**:228-232.
- Buckler-White, A. J., and B. R. Murphy. 1986. Nucleotide sequence analysis of nucleoprotein genes of an avian and a human virus strain identifies two classes of nucleoproteins. *Virology* **155**:345-355.
- 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.
- Chao, L. 1990. Fitness of RNA virus decreased by Muller's ratchet. *Nature (London)* **348**:454-455.
- Chen, E. Y., and P. H. Seeburg. 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. *DNA* **4**:165-170.
- Cox, N. J., F. Kitame, A. P. Kendal, H. F. Maassab, and C. Naeve. 1989. Identification of sequence changes in the cold-adapted live attenuated influenza vaccine strain A/Ann Arbor/6/60 (H2N2). *Virology* **167**:554-567.
- Donis, R. O. 1991. Muller's ratchet and flu virus. *Nature (London)* **353**:308-309.
- Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* **20**:406-416.
- Gammelin, M., J. Mandler, and C. Scholtissek. 1989. Two subtypes of nucleoproteins (NP) of influenza A viruses. *Virology* **170**:71-80.
- Gorman, O. T., W. J. Bean, Y. Kawaoka, I. Donatelli, Y. Guo, and R. G. Webster. 1991. Evolution of influenza A virus nucleoprotein genes: implications for the origins of H1N1 human and classical swine viruses. *J. Virol.* **65**:3704-3714.
- Gorman, O. T., W. J. Bean, Y. Kawaoka, and R. G. Webster. 1990. Evolution of the nucleoprotein gene of influenza A virus. *J. Virol.* **64**:1487-1497.
- Huddleston, J. A., and G. G. Brownlee. 1982. The sequence of the nucleoprotein gene of human influenza A virus, strain A/NT/60/68. *Nucleic Acids Res.* **10**:1029-1038.
- Jones, K. L., J. A. Huddleston, and G. G. Brownlee. 1983. The sequence of RNA segment 1 of influenza virus A/NT/60/68 and its comparison with the corresponding segment of strains A/PR/8/34 and A/WSN/33. *Nucleic Acids Res.* **11**:1555-1566.
- Katz, J. M., M. Wang, and R. G. Webster. 1990. Direct sequencing of the HA gene of influenza (H3N2) virus in original clinical samples reveals sequence identity with mammalian cell-grown virus. *J. Virol.* **64**:1808-1811.
- Kendal, A. P., G. R. Noble, J. J. Skehel, and W. R. Dowdle. 1978. Antigenic similarity of influenza A (H1N1) viruses from epidemics in 1977-1978 to "Scandinavian" strains isolated in epidemics of 1950-1951. *Virology* **89**:632-636.
- Nakajima, K., U. Desselberger, and P. Palese. 1978. Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. *Nature (London)* **274**:334-339.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
- Scholtissek, C., H. Burger, O. Kistner, and K. F. Shortridge. 1985. The nucleoprotein as a possible factor in determining host specificity of influenza H3N2 viruses. *Virology* **147**:287-294.
- Scholtissek, C., W. Rohde, V. Von Hoyningen, and R. Rott. 1978. On the origin of the human influenza virus subtype H2N2 and H3N2. *Virology* **87**:13-20.
- Shortridge, K. F. 1982. Avian influenza A viruses of Southern China and Hong Kong: ecological aspects and implications for man. *Bull. W.H.O.* **60**:129-135.
- Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers, and Y. Kawaoka. 1992. Evolution and ecology of influenza A viruses. *Microbiol. Rev.* **56**:152-179.
- Winter, G., and S. Fields. 1981. The structure of the gene encoding the nucleoprotein of human influenza virus A/PR/8/34. *Virology* **114**:423-428.
- Winter, G., S. Fields, M. J. Gait, and G. G. Brownlee. 1981. The use of synthetic oligodeoxynucleotide primers in cloning and sequencing segment 8 of influenza virus (A/PR/8/34). *Nucleic Acids Res.* **9**:237-245.