EVOLUTIONARY STASIS OF M1 GENE OF HUMAN INFLUENZA A VIRUSES AND THE POSSIBILITY OF THEIR SUBTYPING BY RESTRICTION ANALYSIS OF M1 GENE POLYMERASE CHAIN REACTION PRODUCT

K.Y. PARK[1*], M.G. LEE[2], J.C. RYU[3]

[1*]WHO National Influenza Center, NIH Korea, 5 Nokbun-Dong, Seoul 122-701; [2]Department of Pharmacology, Yonsei University College of Medicine, Seoul; [3]Doping Control Center, Korea Institute of Science and Technology, Seoul, South Korea

Received July 3, 1997; revised July 18, 1997

Summary. – Nucleotide (nt) and amino acid (aa) sequences of the M1 protein in 36 human influenza A viruses were analyzed. The neighbor joining tree of the nt sequences revealed several lineages associated with past epidemics of human influenza. However, the tree of aa sequences revealed only few specific lineages. This discrepancy in phylogeny between nt and aa sequences indicates that the M1 protein of human influenza A virus nearly reached an evolutionary stasis. A simple subtyping method of human influenza A viruses by restriction fragment length polymorphism (RFLP) analysis of M1 gene polymerase chain reaction (PCR) products is discussed.

Key words: human influenza A virus; matrix protein; phylogeny

Introduction

In the past decade, many molecular epidemiological studies of influenza viruses have focused on haemagglutinin (HA) or neuraminidase (NA) genes since they code major surface antigens, are related to human pandemics, and show rapid sequence changes. The genes coding surface proteins, however, are not expected to have a long evolutionary history within hosts that subject the virus to considerable immune selection pressure such as human. The RNA segment 7 of the influenza A virus is bicistronic, encoding both M1 (matrix) and M2 proteins. M1, the most abundant protein in the influenza virion, forms a shell surrounding the nucleocapsid underneath the virion envelope (Webster *et al.*, 1992). It is known that the same M gene has been retained throughout the antigenic shift of HA and NA in human pandemics (Hall and Air, 1981). Although at very slow rate, M1 gene is constantly changing (Ito *et al.*, 1991). It is possible that there are evolutionary remnants in the nt sequence of M1 gene. Therefore, the phylogeny of M1 gene may yield a valuable information on the epidemiological history of human influenza A viruses.

Because newly isolated strains of influenza virus often grow to relatively low titers in embryonated chicken eggs, high yield strains for vaccine production have often been obtained by laboratory manipulations such as sequential passage in the eggs or reassortment with the high yield laboratory strain A/Puerto Rico/8/34. It has been suggested that the high yield properties are related to M1 gene (Klimov *et al.*, 1991; Xu *et al.*, 1994). To reevaluate this possibility, we included 6 high yield reassortants in this study.

M1 gene has a highly conserved nt sequence and is one of the most reliable targets in PCR diagnosis of influenza A and B viruses (Zhang and Evans, 1991). This, how-

*Corresponding author.

Abbreviations: aa = amino acid; FPV = fowl plague virus; HA = haemagglutinin; MoAb = monoclonal antibody; MoMLV = Moloney murine leukemia virus; NA = neuraminidase; nt = nucleotide; PCR = polymerase chain reaction; RFLP = restriction fragment length polymorphism; RT = reverse transcription

ever, cannot identify the subtypes of human influenza A viruses. To detect the subtypes, another PCR targeting HA or NA gene is required. Besides being time and labour consuming, the method may fail to detect the subtypes due to a high rate of genetic drift or shift in HA and NA genes. Therefore, if a simple restriction analysis of PCR products of M1 gene can identify the subtypes, it will aid world-wide survey centers in rapid analysis of current epidemics of influenza.

Materials and Methods

Virus strains and sequence sources. The nt and aa sequences of 36 human influenza A viruses were analyzed phylogenetically using the sequence of fowl plague virus (FPV) as the outgroup. Among the more than 20 sequences of influenza viruses isolated in South Korea in 1995, 8 viruses which represented different areas or showed different sequences were chosen for the analysis. A PCR band originating from a throat swab of a patient from 1994 (virus not isolated) was also included in this study. Drs. N.J. Cox, Center for Disease Control and Prevention, Atlanta, GA, USA, and J.M. Wood, National Institute for Biological Standards and Control, Potters Bar, UK, kindly donated 4 wild type viruses which represent the recent world-wide epidemics and 6 reassortant strains which have the high yield properties. The sequences of 17 additional human influenza A viruses isolated from 1933 to 1989 and of an FPV were taken from the literature and data bank sources (Table 1). The sequences discussed here are available from the GenBank data base (accession No. U52926-U52944).

The viral isolates were grown in 11-day-old embryonated chicken eggs at 35°C for 3 days and the allantoic fluids were then separately harvested. After HA titers were determined using 0.5% chicken red blood cells, the fluids were stored at -70°C prior to RNA preparation.

Oligonucleotide primers. Two oligonucleotide primer pairs complementary to the conserved region of M1 gene of influenza A virus (Zhang and Evans, 1991) were synthesized on an Applied Biosystems 381A DNA synthesizer and purified when used in PCR (the outer upstream primer: nt 71-90, 5'-CCGTCA GGC-CCCCTCAAAGC-3'; the outer downstream primer: nt 710-691, 5'-GACCAGCACTGGAGCTAGGA-3'; the inner upstream primer: nt 101-120, 5'-CAGAGACTTGAAGATGTCTT-3'; the inner downstream primer: nt 501-482, 5'-TGCTGGGAGTCA GCAATCTG-3'). However, the primers were purified by electrophoresis in a denaturing polyacrylamide gel when used in sequencing.

Reverse transcription PCR (RT-PCR) was performed using a modification of Zhang and Evans' procedure (1991). Five hundred µl of infected allantoic fluids or throat swabs of patients was centrifuged at 15,000 x g for 30 mins at 4°C. The pellet was used for RNA extraction by a guanadine thiocyanate method (Promega, Total RNA Isolation System). DNA complementary to viral RNA was prepared by using a Moloney murine leukemia virus (MoMLV) reverse transcriptase (Gibco) with a random hexanucleotide (Perkin-Elmer). cDNA was first subjected to 25 cycles (94°C for 60 secs, 45°C for 90 secs, and 72°C for 40 secs) of a reaction containing the outer primers with *Taq* polymerase (Promega) in a Perkin-Elmer DNA Thermal Cycler 480. An aliquot of this reaction was then diluted 1:50 for a second reaction containing the inner primers and performed in further 25 cycles (94°C for 60 secs, 55°C for 60 secs, and 72°C for 40 secs). The PCR products were verified by electrophoresis in 1.5% agarose gel in Tris-acetate buffer containing 0.5 µg/ml ethidium bromide.

Nucleotide sequencing. To resolve the sequence ambiguity, 4 trials of sequencing were performed on each strain, two kinds of direct sequencing using both the upstream and downstream primers. After exonuclease and phosphatase treatments (USB, PCR Products Sequencing Kit), a PCR product was first subjected to nt sequencing using T7 polymerase with unlabelled primers and [alpha^{35}S]dATP. Another nt sequencing was performed on the same PCR product after its purification with Gene Clean II Kit (Bio 101), using *Taq* polymerase with 5'-end ^{32}P-labelled primer (Promega, fmol DNA Sequencing System). With wild type viruses, nt sequencing was done only on nested PCR products (401 bp, nt 101-501 from the 3'-end of v RNA segment 7) using the inner primers. For reassortant strains, an additional sequencing on the first PCR products (640 bp, nt 71-710) using the outer primers was performed for determination of the aa 218 of M1 protein.

RFLP subtyping. Restriction sites of the nested PCR products were investigated using the DNASIS version 7.08 program (Hitachi Software, 1991). *Mse*I and *Alu*I enzymes (NEB) were chosen for RFLP subtyping of influenza A viruses. The PCR products were purified using a QIAquick PCR Purification Kit (Qiagene) and eluted in distilled water. Restriction digestions were performed as recommended by the manufacturer. The digests were subjected to electrophoresis in 3.5% agarose gels and then photographed. Four PCR bands obtained from throat swabs from Kwangju area (South Korea) in late fall of 1995 were used as test samples in RFLP analysis.

Phylogenetic analysis was performed on the sequences of nt 131-470 from the 5'-end of cDNA (3'-end of vRNA) of RNA segment 7 and on the sequences of aa 36-148 from the N-terminus of M1 protein. The aa sequences were deduced from the nt sequences using the DNASIS program. The sequences of each strain were aligned using the CLUSTAL V software (Higgins *et al.*, 1991; EMBL Data Library). Phylogenetic trees, rooted by taking FPV27 for the outgroup, were constructed by the MEGA version 1.02 program (Kumar *et al.*, 1993) using the neighbor joining method (Saitou and Nei, 1987). The evolutionary distances were estimated by the Kimura 2 parameter. Bootstrap probabilities, based on 1000 resamplings, were calculated for each internal branch of the neighbor joining trees.

Results

Variation of nt sequence in 1993 – 1995

Fig. 1A shows the nt sequence differences of each strain in 340 nt analyzed in this study. Compared to WS33(H1N1),

Table 1. Influenza virus strains analyzed in this study

Strain	Abbreviation	GenBank accession No.[a]	Source or reference
A/FPV/27(H7N7)	FPV27(H7N7)		Markusin et al. (1988)
A/WS/33(H1N1)	WS33(H1N1)		Zebedee and Lamb (1989)
A/WSN/33(H1N1)	WSN33(H1N1)		Zebedee and Lamb (1989)
A/Puerto Rico/8/34(H1N1)	R34(H1N1)		Winter and Fields (1980)
A/Fort Monmouth/1/47(H1N1)	FM47(H1N1)		Smeen and Brown (1994)
A/Fort Warren/1/50(H1N1)	FW50(H1N1)		Zabedee and Lamb (1989)
A/USSR/90/77(H1N1)	USSR77(H1N1)		Samokhavalov et al.
A/Taiwan/1/86(H1N1)	TAIW86(H1N1)	U52939	NIBSC[b]
A/Yamagata/120/86(H1N1)	YAMA86(H1N1)	U52944	CDC[c]
A/Wisconsin/3523/88(H1N1)	WISC88(H1N1)		Ito et al. (1991)
A/Texas/36/91(H1N1)	TX91(H1N1)	U52940	NIBSC
A/Seoul/24/95(H1N1)	S24/95(H1N1)	U52936	NIH Korea
A/Kwangju/27/95(H1N1)	K27/95(H1N1)	U52931	NIH Korea
A/Singapore/1/57(H2N2)	SING57(H2N2)		Zabedee and Lamb (1989)
A/Leningrad/134/57(H2N2	LENIN57(H2N2)		Klimov et al. (1985)
A/Ann Arbor/6/60(H2N2)	AA60(H2N2)		Cox et al. (1988)
A/Korea/426/68(H2N2)	KOREA68(H2N2)		Ito et al. (1991)
A/Aichi/2/68(H3N2)	AICHI68(H2N2)		Ito et al. (1991)
A/Udorn/302/72(H3N2)	UDOR72(H3N2)		Lamb and Lai (1981)
A/Port Chalmers/1/73(H3N2)	PC73(H3N2)		Zabedee and Lamb (1989)
A/Bangkok/1/79/(H3N2)	BANG79(H3N2)		Ortin et al. (1983)
A/Memphis/8/88(H3N2)	MEMP88(H3N2)		Ito et al. (1991)
A/Guangdong/39/89(H3N2)	GUAN89(H3N2)		Xu et al. (1993)
A/Johanesburg/33/94(H3N2)	JOHA94(H3N2	U52929	NIBSC
A/Busan/7/95(H3N2)	B7/95(H3N2)	U52926	NIH Korea
A/Kwangju/1/95(H3N2)	K1/95(H3N2)	U52930	NIH Korea
A/Seoul/2/95(H3N2)	S2/95(H3N2)	U52935	NIH Korea
A/Changwon/9/95(H3N2)	C9/95(H3N2)	U52928	NIH Korea
A/Changwon/11/95(H3N2)	C11/95(H3N2)	U52927	NIH Korea
A/Seoul/8/95(H3N2)	S8/95(H3N2)	U52937	NIH Korea
A/Seoul/94[d]	S94	U52938	NIH Korea
X113(H1N1) reassortant	X113(H1N1)	U52941	NIBSC
X117(H3N2) reassortant	X117(H3N2)	U52942	CDC
X121(H3N2) reassortant	X121(H3N2)	U52943	NIBSC
NIB27(H1N1) reassortant	NIB27(H1N1)	U52932	NIBSC
NIB34(H3N2) reassortant	NIB34(H3N2)	U52933	NIBSC
RESVIR-8(H3N2) reassortant	RES8(H3N2)	U52934	NIBSC

[a]The strains with an accession numbers were sequenced in this study.
[b]Dr. J.M. Wood, National Institute for Biological Standards and Control Potters Bar. UK.
[c]Dr. N.J. Cox, Center for Disease Control and Prevention, Atlanta, GA, USA.
[d]S94 is not an isolated strain, but its M1 gene was detected by PCR in a throat swab.

current H3N2 viruses have maximally 21 different nt with exception of WISC88(H1N1) which is assumed to be a swine virus (Rota *et al.*, 1989). The H1N1 viruses after 1977 have fewer sequence changes than the recent H3N2 viruses. The rate of change per nt for the M1 gene was 0.65×10^{-3} substitutions per year (0.22 nt per year per 340 nt), excluding the H1N1 viruses after 1977 (Fig. 2). There were virtually no changes in the aa sequence of M1 over the past 60 years (Fig.1B and Fig. 2). KOREA68 has a maximum change of 4 of 113 aa analyzed in this study in comparison with WS33.

RFLP subtyping

We found several nt positions specific for each subtype. Among them, nt 277 and 331 of RNA segment 7 were recognizing sites for restriction enzymes *Mse*I and *Alu*I, respectively (Fig. 1A). There are 1 or 2 more restriction sites for *Mse*I on the PCR products. The nt sequence of WISC 88(H1N1) was quite different from that of other viruses. Table 2 shows the expected digestion patterns of the M1 gene PCR products using *Mse*I or *Alu*I according to each subtype. The RFLP method is expected to discriminate sub-

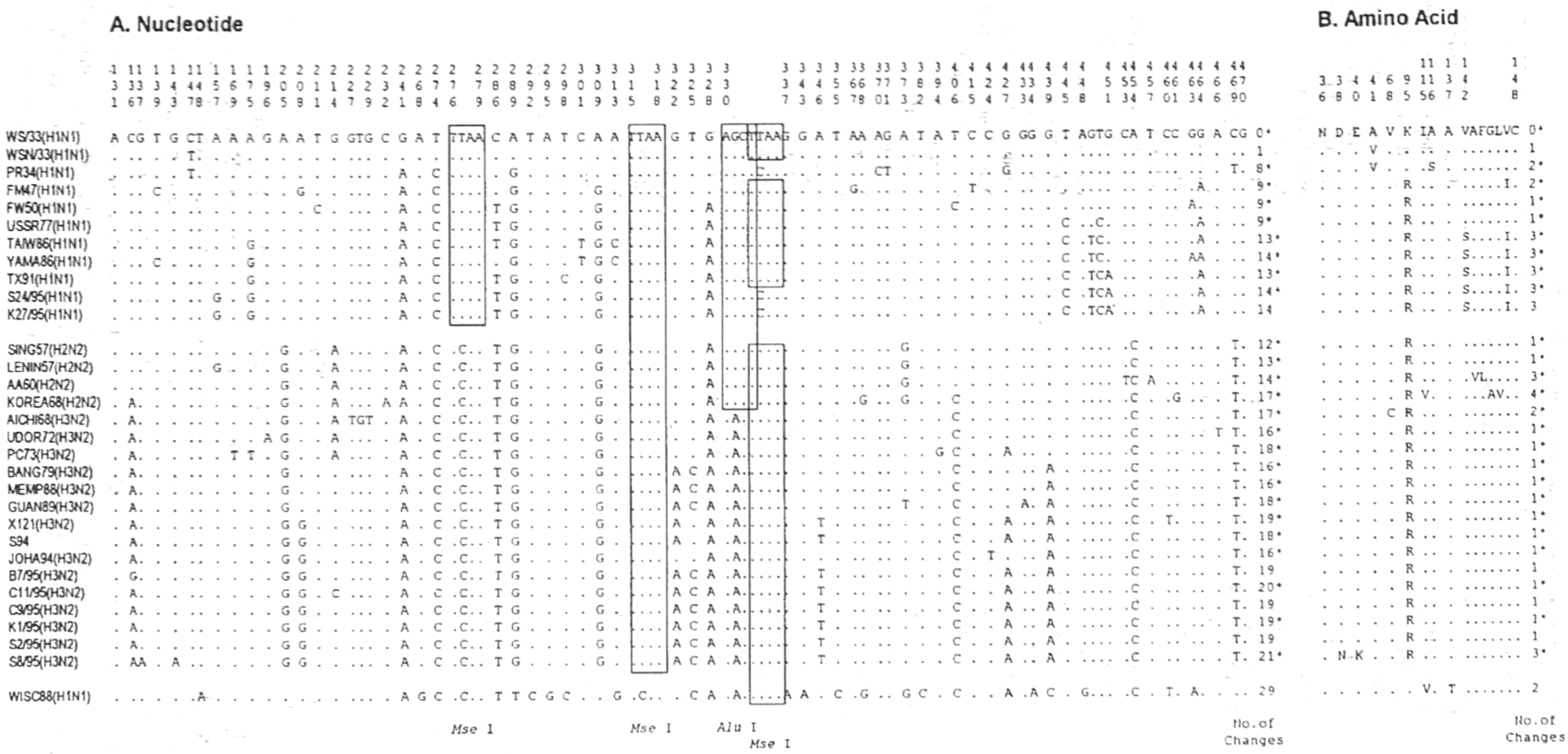

Fig. 1

Variations in nt and aa sequences of the M1 gene in human influenza A viruses from 1933 – 1995

A: Sequences of nt 131-470 from the 5'-end of cDNA (3'-end of vRNA) of RNA segment 7 were compared. The number of difference from WS33 is listed at the end of each sequence. A 1995 H3N2 virus has maximally 21 different nt with exception of WISC88 which is assumed to be a swine virus (Rota *et al.*, 1989). The recognition sites of *Mse*I and *Alu*I are boxed. The *Mse*I site (nt 276-279) is specific for human H1N1 strains, and the *Alu*I site (nt 330-333) is specific for human H1N1 and H2N2 strains. The sequences of the reassortants are not shown with exception of X121 because X113, X117, NIB34, and RES8 have the same sequences as RP34, and NIB27 has the same sequence as TX91 in the region of nt 131-470. The sequence of X121 is thought to represent that of a wild strain, A/Shangdong/9/93 (H3N2). B: Sequences of aa 36-148 from the N-terminus of M1 protein were compared. *The strains used in Fig. 2.

types of human viruses and to detect an accidental infection of humans with non-human viruses. We applied this method to the 4 PCR bands obtained from throat swab samples originating from the Kwangju area of South Korea during late fall of 1995 (Fig. 3). All 4 samples were assumed to contain H1N1 viruses, similar especially to S24/95(H1N1) isolated in Seoul in early fall of 1995. One month later, 4 influenza strains were isolated from throat swabs and confirmed as H1N1 viruses by haemagglutination inhibition test (A/Kwangju/27/95 through A/Kwangju/30/95). In the *Mse*I digestion pattern, the bands of H1N1 strains isolated in 1995 were slightly larger than those of TX91, which was due to a mutation at nt 334. PR34 was expected to show the same pattern. However, the critical difference between H1N1 and H3N2 was easily detectable by a routine agarose gel electrophoresis.

Phylogenetic analysis of nt and aa sequences

Phylogenetic trees based on partial M1 gene sequences of human influenza A viruses were constructed using the neighbor joining method rooted by taking FPV27 for the outgroup (Figs. 4A, 4B). M1 gene of WISC88 is genetically distinct from that of other human viruses. In the neighbor joining tree based on 340 nt of M1 (Fig. 4A), the H1N1 viruses form a lineage differing from the H2N2 and H3N2 viruses (bootstrap probability 81%). Among the H1N1 viruses, the viruses isolated after 1977 form a cluster (bootstrap probability 61%). The H2N2 and H3N2 viruses isolated before 1973 are thought to have a somewhat different M1 gene from those isolated after 1979 (bootstrap probability 64%). On the contrary, the phylogenetic tree based on 113 aa of M1 shows no significant lineage among the subtypes and low bootstrap probabilities in most forks (Fig. 4B).

Relationship between M1 and high yield properties

Among the 6 reassortants, NIB27 has the same sequence as TX91, and X121 has a sequence similar to that of current H3N2 viruses such as S94. The other 4 strains have a nt sequence very similar to that of PR34 reported in 1980. X113, X117, and NIB34 have differences in nt 535 (C→T) and nt 655 (G->A), and RES8 has another difference in nt 560 (A->C), all in comparison with PR34 (data not shown). The phylogeny based on the nt sequences revealed that the M genes of 2 strains originated from wild strains and those of the other 4 strains originated from PR34.

We found no specific sequences for the high yield reassortants and PR34 in the sequence of aa 36-148 when compared with all other strains (Fig. 1B), and of aa 149–220 when compared with the strains in previously reported studies (data not shown). Klimov *et al.* (1991) claimed that Thr^{218} of M1 protein was specific for high yield strains, and Ala^{218}

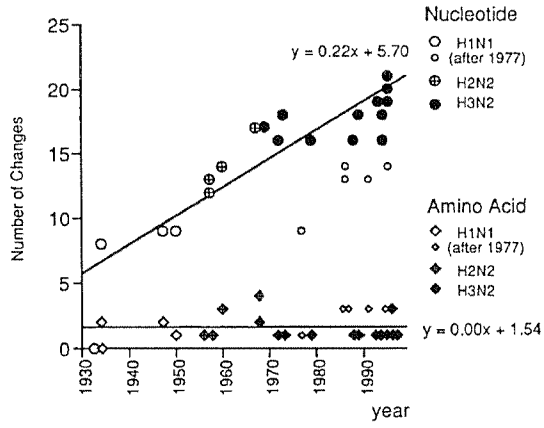

Fig. 2
The evolutionary rates for the M1 gene in human influenza A viruses from 1933 – 1995

The slopes were estimated by regression of the year of isolation against the nt or aa changes in the present study. The strains used in this graph are marked (*) in Fig. 1. Among the strains isolated in South Korea in 1995, 3 H3N2 and 1 H1N1 strains which show different number of changes were included. The H1N1 viruses isolated after 1977 were not included in estimating the evolutionary rates, because their nt sequences showed a significant latency in evolution. The rate of change per nt for the M1 gene was 0.65 x 10^{-3} substitutions per year (0.22 per year per 340 nt). There were virtually no changes in the aa sequence of M1 over the past 60 years

Table 2. Expected digestion patterns of the PCR products of M1 gene

Subtype	*Alu*I		*Mse*I	
	digestion at nt 331	bp	digestion at nt 277	bp
H1N1	yes	231,170	yes	177,166 (194,177)[a]
H2N2	yes	231,170	no	217,166
H3N2	no	401	no	217,166
WISC88	no	401	no	235[b],166

[a]PR34, S24/95, and K27/95 contain C^{334} instead of T^{334}. [b]WISC88(H1N1) of swine origin is supposed to show an abnormal *Mse*I digestion pattern.

Table 3. Origin of M gene and aa^{218} of M1 protein of the 6 high yield reassortants

Strain	History	Origin	aa^{218}
X113(H1N1)	TX91(H1N1) x X31(H3N2)[a]	PR34	Thr
NIB27(H1N1)	TX91(H1N1) x X31(H3N2)[a]	TX91	Thr
X117(H3N2)	BEIJ92(H3N2)[b] x PR34(H1N1)	PR34	Thr
X121(H3N2)	SHAN93(H3N2)[c] x PR34(H1N1)	SHAN93[c]	Thr
NIB34(H3N2)	JOHA94(H3N2) x PR34(H1N1)	PR34	Thr
RES8(H3N2)	JOHA94(H3N2) x PR34(H1N1)	PR34	Thr

[a]A reassortant (AICHI68(H3N2) x PR34(H1N1) developed for reassortment of H1N1 viruses (Kilbourne, 1969). [b]A/Beijing/32/92(H3N2). [c]A/Shangdong/9/93(H3N2).

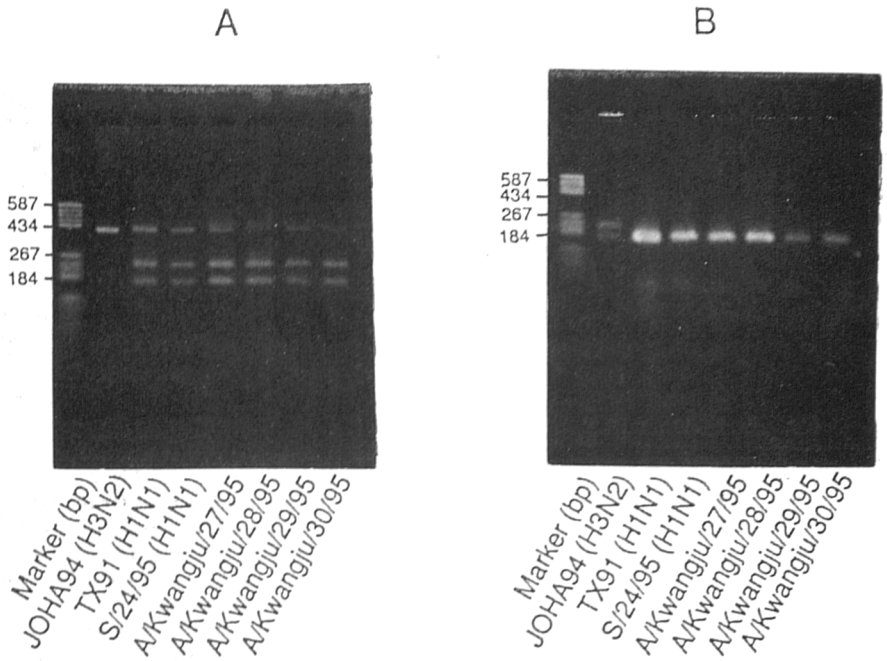

Fig. 3

Rapid estimation of the subtypes of human influenza A viruses by RFLP analysis of M1 gene products

JOHA94(H3N2), TX9(H1N1), and S/24/95(H1N1) strains were used as standards. Four PCR bands obtained from throat swabs in the Kwangju area of South Korea during late fall of 1995 were used as test samples. A: *Alu*I digestion. The digested bands of 231 bp and 170 bp as well as the original substrate band of 401 bp are seen on the *Mse*I digestion. B: JOHA94(H3N2) showed 2 bands of 217 bp and 166 bp as expected in Table 2. TX91(H1N1) showed fused bands of 177 bp and 166 bp, and S24/95 (H1N1) showed fused bands of 194 bp and 177 bp as expected in Table 2. All the 4 test samples showed patterns similar to S24/95.

or Val[218] was specific for low yield strains. However, our study of 6 high yield reassortants showed that either Thr or Ala can occur at this position in high yield strains (Table 3).

Discussion

The phylogeny data using nt sequences of M1 gene were correlated with many known past epidemics of human influenza. The H1N1 viruses appeared in 1918 and have been circulating in humans since then with an intermittent period from 1957 to 1977 (Xu *et al.*, 1993). It is reported that the reappearing USSR77 virus is nearly identical with the strains isolated in 1950 (Nakajima *et al.*, 1978). One possible explanation is that the virus has been preserved in a frozen state without any accumulation of mutations. Our data on the phylogeny and nt changes also indicate the reemergence of the same virus between the early 1950s and 1977. Excluding WISC88 of swine origin, the H1N1 viruses form a single lineage from the H2N2 and H3N2 viruses in the neighbor joining tree. The H1N1 strains from 1995 have fewer nt changes than the H3N2 strains from 1955 in comparison to WS33, the mean number of nt changes in H1N1 being 14, and that in H3N2 19.5. It means that the current H1N1 viruses still have 25 (5.5/0.22) years of latency in evolution. In the neighbor joining tree constructed on the basis of nt sequences, the H1N1 viruses isolated after 1977 form a cluster. It indicates that the current H1N1 strains are direct descendants of USSR77.

It has been speculated that the H3N2 viruses after 1977 such as A/Texas/77 are not succeeding strains to those before 1975 such as A/Victoria/75 (Webster *et al.*, 1992). Our data strongly support this possibility. The neighbor joining tree shows that the H2N2 and H3N2 viruses are divided into two major lineages, those before 1973 and after 1979. In the graph of nt changes (Fig. 2), the H2N2 and H3N2 viruses from 1957 – 1973 are located above the regression line, and those after 1979 are mainly located below the line. Therefore, it can be inferred that a somewhat different H3N2 virus, different at least in the M gene, newly appeared during mid 1970s.

In the present study, we found that there were few changes in the aa sequences and few significant lineages in the phylogeny of the M1 protein sequence. Though we analyzed a partial sequence of M1 gene (aa 36-148) only, the majority of known important sites were located within this sequence. The lipid binding sites have been mapped to aa 62-86 and

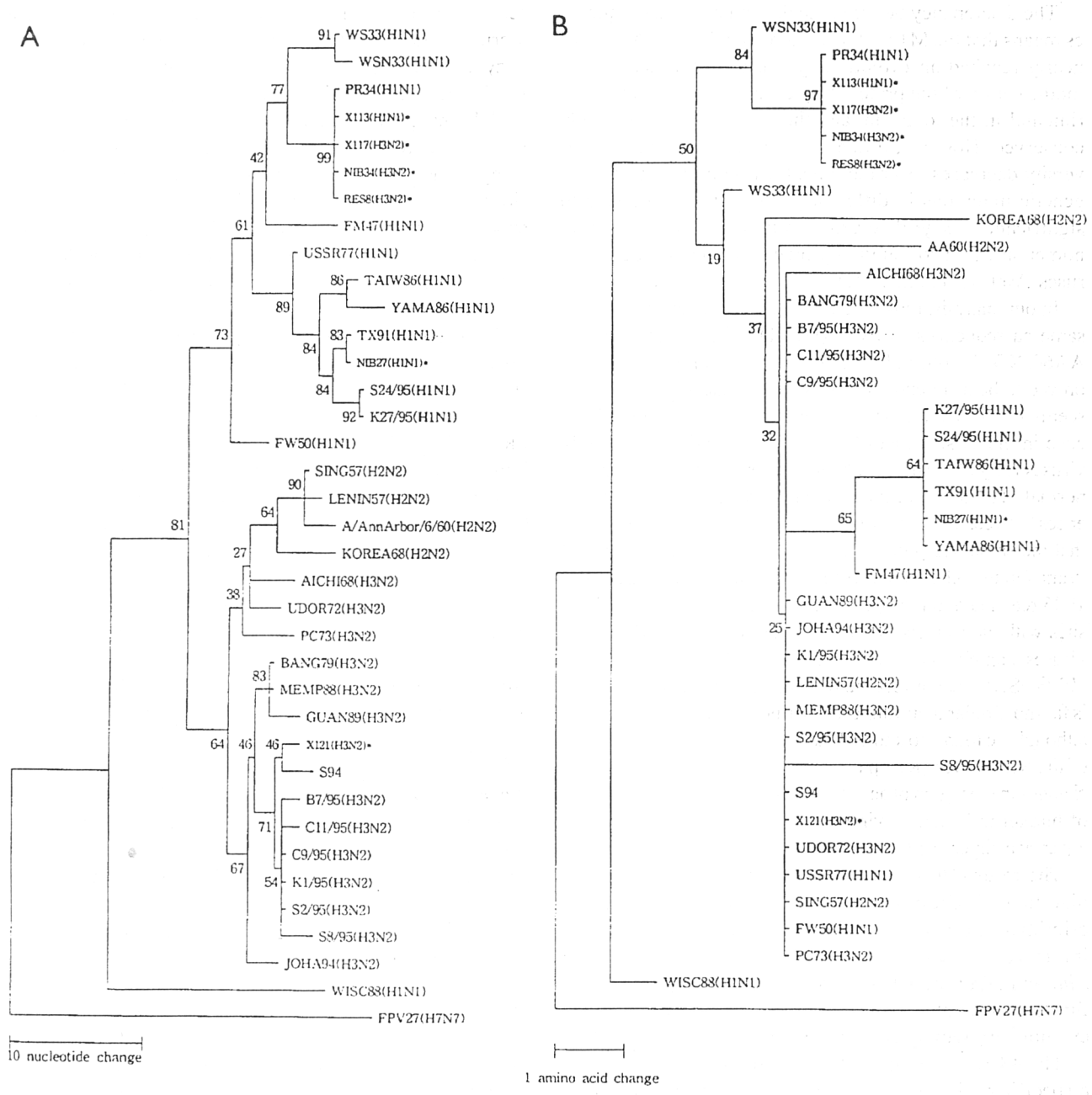

Fig. 4

The neighbor joining tree based on partial nt and aa sequences of M1 gene in human influenza A viruses from 1933 – 1995

The trees were rooted by taking FPV27 for the outgroup. The bootstrap probabilities, as determined for 1000 resamplings, are given in percentage beside the internal branches. *The reassortants. A: Phylogenetic tree based on the nt sequence. M1 gene of WISC88 was genetically distinct from those of other human influenza A viruses. The H1N1 viruses form a lineage differing from the H2N2 and H3N2 viruses. Among the H1N1 viruses, the viruses isolated after 1977 form a cluster. The H2N2 and H3N2 viruses isolated before 1973 form a lineage in contrast to those isolated after 1979. B: Phylogenetic tree based on the aa sequence. The tree shows no significant lineage among the subtypes and low bootstrap probabilities in most forks.

114-133 (Grigoriades and Frangione, 1981), and critical anti-transcriptase sites to aa 70 and 140 (Hankins *et al.*, 1989). It has also been hypothesized that an Arg/Lys face on an α-helix containing aa 95-105 might bind to the phosphate backbone of RNA (Winter and Fields, 1980). Pro^{54} may be phosphorylated by proline kinase (Whittaker *et al.*, 1995), and the region of aa 58-66 is important for binding to a human MHC class I molecule, HLA-A2 (Falk *et al.*, 1994).

The discrepancy between the data on nt and aa sequences means that the M1 protein of human influenza A viruses nearly reached an evolutionary stasis. In that stasis, any modification of the protein sequence is likely to prove detrimental in the long run, and the aa sequences are highly conserved. However, it does not mean that the genetic diversity decreases; over the time, the accumulation of silent genetic mutations in different populations can be lead to significant lineage divergence in nt sequences. This phenomenon was also found in proteins of avian influenza viruses (Webster *et al.*, 1992).

In our analysis, most of H2N2 and H3N2 viruses have the same aa sequences as FW50(H1N1). The aa sequences of AA60, KOREA68, and AICHI68 are a little different, but they proved to be dead-end mutations. The mutation of S8/95 also seems to end without any succeding strain. The only difference in the aa sequences between WS33 and current H3N2 viruses is Lys^{95}->Arg^{95}. Both these aa are positively charged at normal pH, and it is expected that there are nearly no differences between the two protein structures. We therefore assume that the same M1 protein has been retained over the past 60 years. On the other hand, it seems that the H1N1 viruses after 1986 reached another optimum state. They have two distinct sites with aa changes in the aa sequences compared to other viruses, namely Val^{142}->Ser^{142} and Val^{147}->Ile^{147}. The change of Val^{142}->Ser^{142} may affect the protein structure, because serine is far more polar than valine and serine can be phosphorylated, although we found no known kinases which can recognize Ser^{142} of the M1 protein (Kemp and Pearson, 1990). There were no changes in the region of aa 58-66 of M1 (HLA-A2) in all strains of human influenza A viruses. It seems that this epitope as a part of an internal protein is not subjected to immunoselection.

The results of the RFLP show that a simple restriction digestion of PCR products of M1 gene can effectively discriminate the subtypes of currently circulating human influenza A viruses. The mutation sites recognizable by *Alu*I and *Mse*I are silent. It seems that these sites are not functionally related to the protein of each subtype but to the evolutionary remnants of the M1 gene.

The PCR targeting of HA and/or NA genes may be a specific method for rapid identification of virus subtypes. In an earlier study, Zhang and Evans (1991) devised various primer sets for PCR detection of influenza viruses. They found that PCR targeting of these genes failed to detect certain type of virus such as A/New Jersey/8/76 (H1N1), and concluded that by considering the high rate of sequence variations of HA and NA it may never be possible to define a single set of primers capable of achieving this goal. We therefore highly recommend our method of simple RFLP analysis for rapid estimation of subtypes of epidemic influenza viruses to the world-wide centers.

Although at slow rate, the nt sequence of M1 gene is constantly changing. Even in a single epidemic, different sequences were found e.g. in C9/95 and C11/95. If there is a critical mutation at the recognition site of the restriction enzymes, the digestion pattern will be distored. The 1995 H1N1 strains and PR34 have C instead of T at nt 335 which slightly changes their patterns of *Mse*I digestion. In this case, however, there were no problems in identifying the virus subtypes using a routine agarose gel electrophoresis. Considering the very slow rate of the M1 gene mutation, we expect that our RFLP method will be useful in a long run.

The experiments with the reassortants suggest that M1 may not be associated with the virus high yield properties. If M1 is closely related to the high yield properties of PR34, all M1 of the 6 reassortants should originate from PR34. However, the M1 of 2 reassortants came from the wild strains in this study. This does not correspond to the data of Xu *et al.* (1994) that M1 of all the 8 reassortants from 1979 –1992 came from PR34. We can hardly explain this discrepancy. One possibility is a difference in the methods used in the two studies in identifying the M1. Xu *et al.* (1994) have analyzed M1 by ELISA using a specific MoAb to PR34. Even though an antibody to PR34 may be very specific, there is a chance of a cross-reaction wiyth wild strains. Another possibility is that the wild strains TX91 and SHAN93 attained the high yield properties asociated with M1 by chance. However, this is probably a very rare possibility, since we found no sequences specific for the reassortants in comparing sequences of aa 36-220 of total 252 aa of M1 in wild strains.

Klimov *et al.* (1991) have reported that Thr^{218} is specific for high yield strains, and Ala^{218} or Val^{218} is specific for low yield strains. However, our study reveals that either Thr or Ala can occur at this position in high yield strains. Recently, Ward (1995) has reported that a change of Ala^{41}->Val^{41} may be related to mouse adaptation and increased growth in eggs and MDCK cells of WSN33. Considering our data, this change is not related to, at least, the increased growth, because the high yield strains NIB27 and X121 contain Ala^{41} like all other human influenza A viruses. The fact that M1 of human influenza A viruses nearly reached the evolutionary stasis also means that an existence of more effective M1 related to high yield properties is hardly possible. A further study of complete sequence of RNA segment 7 may reinforce our suggestion that M1 gene of human influenza A viruses reached an evolutionary stasis.

Acknowledgements. We wish to thank Prof. K.J. Song, Department of Microbiology, Korea University College of Medicine, Seoul, for helpful discussions. We thank also Ms H.J. Yoon for technical assistance.

References

Cox NJ, Kitame F, Kendal AP, Maassab HF, Naeve C (1988): Identification sequence changes in the cold adapted, live at-

tenuated influenza vaccine strain, A/Ann Arbor/6/60 (H2N2). *Virology* **167**, 554–567.

Falk K, Rotzschke O, Stevanovic S, Gnau V, Sparbier K, Jung G, Rammensee HG, Walden P (1994): Analysis of a naturally occurring HLA class I-restricted viral epitope. *Immunology* **82**, 337–342.

Grigoriades A, Frangione B (1981): Insertion of influenza M protein into the viral lipid bilayer and localisation of site of insertion. *J. Virol.* **40**, 323–328.

Hall RM, Air GM (1981): Variation in nucleotide sequence coding for the N-terminal regions of matrix and nonstructural proteins of influenza A virus. *J. Virol.* **38**, 1–7.

Hankins RW, Nagata K, Bucher DJ, Popple S, Ishihama A (1989): Monoclonal antibody analysis of influenza virus matrix protein epitopes involved in transcription inhibition. *Virus Genes* **3**, 111–126.

Ito T, Gorman OT, Kawaoka Y, Bean WJ, Webster RG (1991): Evolutionary analysis of the influenza A virus M gene with comparison of the M1 and M2 proteins. *J. Virol.* **65**, 5491–5498.

Kemp BE, Pearson RB (1990): Protein kinase recognition sequence motifs. *Trends Biochem. Sci.* **15**, 342–346.

Kilbourne ED (1969): Future influenza vaccines and the use of genetic recombinants. *Bull. WHO* **41**, 643–645.

Kimura M (1980): A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequence. *J. Mol. Evol.* **16**, 111–120.

Klimov AI, Cox NJ, Yotov WV, Rocha E, Alexandrova GI, Kendal AP (1992): Sequence changes in the live attenuated, cold-adapted variants of influenza A/Leningrad/134/57 (H2N2). *Virology* **186**, 795–797.

Klimov AI, Sokolov NI, Orlova NG, Ginzburg VP (1991): Correlation of amino acid residues in the M1 and M2 proteins of influenza virus with high yielding properties. *Virus Res.* **19**, 105–114.

Lamb RA, Lai C-J (1981): Conversion of influenza virus membrane protein (M1) amino acid sequence and open reading frame of RNA segment 7 encoding a second protein (M2) in H1N1 and H3N2 strains. *Virology* **112**, 746–751.

Markusin S, Ghiasi H, Sokolov N, Shilov A, Sinitsin B, Brown D, Klimov A, Nayak D (1988): Nucleotide sequence of RNA segment 7 and the predicted amino acid sequence of M1 and M2 proteins of FPV/Weybridge (H7N7) and WSN (H1N1) influenza virus. *Virus. Res.* **10**, 263–272.

Nakajima K, Desselberger U, Palese P (1978): Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. *Nature* **274**, 334–339.

Ortin J, Martinez C, Del Rio L, Davila M, Lopez-Galindez C, Villanueva N, Domingo E (1983): Evolution of the nucleotide sequence of the influenza virus RNA segment 7 during drift of the H3N2 subtype. *Gene* **23**, 233–239.

Rota RA, Rocha EP, Harmon MW, Hinshaw VS, Sheerar MG, Kawaoka Y, Cox NJ, Smith TF (1989): Laboratory characterization of a swine influenza virus isolated from a fatal case of human influenza. *J. Clin. Microbiol.* **27**, 1413–1416.

Saitou N, Nei M (1987): The neighbor-joining method: A new method for reconstrudcting phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

Samokhavalov EI, Karginov VA, Chizhikov VE, Blinov VM, Yuferov VP, Vasilenko SK, Uryvavev LV, Zhdanov VM (1985): Primary structure of RNA segment 7 of A/USSR/90/77 (H1N1) influenza virus. *Bioorg. Khim.* **11**, 1080–1085.

Smeenk CA, Brown EG (1994): The influenza virus variant A/FM/1/47-MA possesses single amino acid replacements in the hemagglutinin, controlling virulence, and in the matrix protein, controlling virulence as well as growth. *J. Virol.* **68**, 530–534.

Ward AC (1995): Specific changes in the M_1 protein during adaptation of influenza virus to mouse. *Arch. Virol.* **140**, 383–389.

Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992): Evolution and ecology of influenza A viruses. *Microbiol. Rev.* **56**, 152–179.

Whittaker G, Kemler I, Helenius A (1995): Hyperphosphorylation of mutant influenza virus matrix protein, M1, causes its retention in the nucleus. *J. Virol.* **69**, 439–445.

Winter G, Fields S (1980): Cloning of influenza cDNA into M13: the sequence of the RNA segment encoding the A/PR/8/34 matrix protein. *Nucleic Acids Res.* **8**, 1965–1974.

Xu K, Kilbourne ED, Hall HE, Cox NJ (1994): Nonimmunoselected intrastrain genetic variation detected in pairs of high-yielding influenza A (H3N2) vaccine and parental viruses. *J. Infect. Dis.* **170**, 1432–1438.

Xu X, Rocha EP, Regnery HL, Kendal AP, Cox NJ (1993): Genetic and antigenic analyses of influenza A (H1N1) viruses, 1986–1991. *Virus Res.* **28**, 37–55.

Zebedee S, Lamb RA (1989): Nucleotide sequence of influenza virus RNA segment 7: a comparison of five isolates. *Nucleic Acids Res.* **17**, 2870.

Zhang W, Evans DH (1991): Detection and identification of human influenza viruses by the polymerase chain reaction. *J. Virol. Methods* **33**, 165–189.