Cell, Vol. 25, 315–323, August 1981, Copyright © 1981 by MIT

Complete Structure of A/duck/Ukraine/63 Influenza Hemagglutinin Gene: Animal Virus As Progenitor of Human H3 Hong Kong 1968 Influenza Hemagglutinin

Rongxiang Fang,* Willy Min Jou, Danny Huylebroeck, René Devos and Walter Fiers
Laboratory of Molecular Biology
State University of Ghent
Ledeganckstraat, 35
B-9000 Ghent, Belgium

Summary

We have explored the possibility that an animal viral reservoir contained a direct ancestor gene for the H3 hemagglutinin type present in influenza A viruses in humans since 1968. For this purpose, the duck/Ukraine/1/63 hemagglutinin gene was cloned and sequenced. From the comparison of its complete primary structure with that of several human H3 hemagglutinins as well as those of an H2 and an H7 hemagglutinin, we conclude that the duck/Ukraine/63 hemagglutinin sequence fully corroborates its previous identification by immunological and other methods as belonging to the H3 subtype. Moreover, the duck/Ukraine/63 amino acid sequence is more closely related structurally and presumably antigenically to the human Aichi/68 hemagglutinin, which formed the beginning of the H3N2 pandemic in humans, than to that of Victoria/75, which has undergone an additional 7 year drift period in humans. This observation could best be explained by a common ancestor hemagglutinin gene for duck/Ukraine/63 and human Aichi/68. On the basis of silent, accumulated base changes, we estimate that the strain carrying this postulated common progenitor hemagglutinin gene was circulating in the period 1949–1953 in the animal reservoir. This relatively recent divergence, as well as the closer kinship between the duck/Ukraine/63 and the human Aichi/68 hemagglutinin, as compared with the later Victoria/75, strongly suggests that the influenza A virus of the H3N2 subtype circulating in the human population since 1968 has derived its hemagglutinin gene from a strain in the animal reservoir. Undoubtedly, this occurred by reassortment between previously present human H2N2 virus and this animal strain. These results provide support at the molecular level for the general idea that the wide variety of influenza viruses known to be present in animals can serve as a gene reservoir for human influenza A viruses.

Introduction

Influenza A, B and C viruses were isolated for the first time from man in 1933, 1940 and 1947, respectively (Choppin, 1977). They are classified as type A, B or C on the basis of serologically crossreactive internal nucleocapsid protein. The order of isolation of the three orthomyxovirus types parallels their relative importance as human pathogens. Influenza A can undergo radical changes (antigenic shift) in the properties of its two surface antigens, hemagglutinin and neuraminidase, resulting in the appearance of a new human influenza A subtype that is antigenically unrelated to previous strains in one or both of these surface glycoproteins. During this century, antigenic shifts associated with pandemics have occurred in 1918, 1957 and 1968 (Dowdle, 1976). According to the recently revised nomenclature for influenza A virus surface antigens (World Health Organization, 1980) the corresponding human subtypes are referred to as H1N1 (1918–1957), H2N2 (1957–1968) and H3N2 (1968–). Serological evidence suggests that at the turn of the century a virus with an H3 hemagglutinin but with a neuraminidase different from N2 (N8, previously known as Nequine 2) was in circulation (Fedson et al., 1972; Masurel and Marine, 1973). In 1977 the H1N1 subtype reappeared, and since then influenza A viruses of the H3N2 and H1N1 subtypes have been cocirculating in the human population (see below).

Within a particular subtype the influenza A virus is subject to frequent minor antigenic changes (antigenic drift) that cause epidemic outbreaks every 1 to 2 years. Influenza B also undergoes drift and is associated with epidemics, but antigenic shift has not been observed. Influenza C causes isolated and commonly mild infections. It is probably not coincidental (see below) that only influenza A viruses have also been found in pigs, horses and a great variety of birds, both wild and domesticated (Hinshaw et al., 1979; Laver and Webster, 1979).

Although the two surface antigens vary independently, only antibodies against the hemagglutinin are virus-neutralizing (Laver and Kilbourne, 1966), and so changes in the hemagglutinin are considered more important. Consequently, antigenic changes have been studied most intensively on the hemagglutinin molecule. Analyses were initially conducted at the protein level (Waterfield et al., 1979; Dopheide and Ward, 1980; Laver et al., 1980; Ward and Dopheide, 1980), but recently genetic-engineering approaches in combination with fast DNA-sequencing techniques have been applied to study the hemagglutinin gene from several H3-subtype viruses from man (Both and Sleigh, 1980; Min Jou et al., 1980; Sleigh et al., 1980; Verhoeyen et al., 1980), from one H2-subtype human virus (Gething et al., 1980), from one human virus belonging to the H1 subtype (Davis et al., 1980) and from the H7 subtype from fowl (Porter et al., 1979). Such studies of one particular gene of influenza virus are made relatively easier because of its divided genome, consisting of eight negative-stranded RNA segments, most likely each coding for one protein (Palese, 1977), except for segment 8, which specifies

* Present address: Virus Laboratory, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China.

two nonstructural proteins (Inglis et al., 1979; Lamb and Choppin, 1979).

From the comparison of these different H3 hemagglutinins, antigenic drift can be described, in principle, as the result of the stepwise accumulation of single base substitutions. In light of the recently determined three-dimensional structure of an early H3 strain (Wilson et al., 1981; Wiley et al., 1981), it appears that successive major disease-producing viruses had at least one change in four postulated antigenic sites of the hemagglutinin, thus probably allowing the virus to escape neutralization.

The nature and cause of antigenic shift are less clear. One conclusion, however, is evident from a comparison of the H2 and H3 hemagglutinins (Gething et al., 1980; Verhoeyen et al., 1980): the differences in sequence are much too extensive to have arisen by stepwise mutation. This was already realized soon after the appearance of the H3 hemagglutinin in humans in 1968, from peptide mapping data (Laver and Webster, 1972). So a major dilemma remains—namely, what is the origin of new influenza A hemagglutinin subtypes in man?

Soon after the (re-)appearance of the H3 hemagglutinin in man in 1968, it was found that the hemagglutinins from two strains of animal influenza virus—A/equine/Miami/1/63 (at the time antigenically defined as Hequine 2 Nequine 2) and A/duck/Ukraine/1/63 (Havian 7 Nequine 2) (both isolated in 1963, 5 years before the Hong Kong influenza appeared)—were both antigenically related to H3 (Coleman et al., 1968; Zakstelskaja et al., 1969; Laver and Webster, 1973). This observation implicated animal influenza viruses similar to equine/Miami and duck/Ukraine as possible progenitors of the Hong Kong strain, by reassortment of the hemagglutinin (and possibly other genes) between an animal and a human virus. This led to the hypothesis of an "animal reservoir" to explain the origin of new influenza A subtypes in man (Webster and Laver, 1975; Laver and Webster, 1979). In fact, all the human hemagglutinin and neuraminidase subtypes have also been isolated from lower mammals and birds, in the same combination as found in man as well as in other combinations (Hinshaw et al., 1979). In addition, nine other hemagglutinin subtypes and seven neuraminidase subtypes have been described from nonhuman sources (Schild et al., 1980), which could eventually serve as a reservoir for new influenza A strains in humans.

The relationship between Heq2, Hav7 and H3 hemagglutinins was further documented by other criteria: peptide mapping (Laver and Webster, 1973), the presence of a blocked N terminus as compared with other strains (Laver and Webster, 1977) and a high percentage of RNA crosshybridization (Scholtissek et al., 1978a). In fact, the latter studies suggest that the gene constellation of the human H2N2 and H3N2 subtypes differs only in the hemagglutinin acquired by reassortment from an unidentified second strain.

In 1977 a reemergence of viruses of the H1N1 subtype was observed. These viruses not only belonged to a subtype previously found in the human population, but moreover, their surface antigens (Kendal et al., 1978) and even their complete gene constellation (Nakajima et al., 1978; Scholtissek et al., 1978b) closely resembled those of 1950 strains. The reappearance of these H1N1 viruses in the human population after 27 years in essentially unchanged form is in marked contrast with the variability seen in the genes of successive human influenza A isolates, not only in those coding for the surface antigens (Blok and Air, 1980; Both and Sleigh, 1980; Dopheide and Ward, 1980; Laver et al., 1980; Min Jou et al., 1980; Sleigh et al., 1980; Verhoeyen et al., 1980; Ward and Dopheide, 1980), but in all the genes (Young et al., 1979; Ortin et al., 1980). This indicates another mechanism for antigenic shift, by which a virus can reappear in the human population after it has been sequestered for a considerable time by some unknown mechanism in or outside that population.

To test directly the idea of an animal viral reservoir as an ancestor for the H3 hemagglutinins present in influenza A viruses in humans since 1968, we have determined the structure of the duck/Ukraine hemagglutinin gene. This strain is considered to be more closely related to the human species, according to the immunological and hybridization criteria cited above. This study allows a direct comparison with the hemagglutinin gene from the virus that occurred at the beginning of the Hong Kong (H3N2) period in humans, A/Aichi/2/68 (Verhoeyen et al., 1980), as well as with the hemagglutinin genes from viruses that have been circulating for 4 (Sleigh et al., 1980) or 7 years (Min Jou et al., 1980) in the human population. Our results very strongly suggest a recent common ancestry for the human H3 and the animal duck/Ukraine/63 hemagglutinin genes.

Results and Discussion

Cloning and Nucleotide Sequencing

For cloning the duck/Ukraine hemagglutinin gene, the same procedure as for the A/Victoria/3/75 and A/Alchi/2/68 hemagglutinin genes was followed (Min Jou et al., 1980; Verhoeyen et al., 1980). In essence, this involved the synthesis of a double-stranded DNA copy of the hemagglutinin RNA and cloning of the DNA copy in the Pst I site of pBR322 with poly(dA)–poly(dT) tails. Plasmid DNA from clones specifically hybridizing with ^{32}P-labeled hemagglutinin RNA (Grunstein and Hogness, 1975) was prepared (Birnboim and Doly, 1979; Kahn et al., 1979), and the length of the insert was determined by restriction analysis. The longest insert was found in plasmid pDHA18; it contained about 1700 bp including the tails. This indicated that we were probably missing

somewhat more than the 15–30 nucleotides corresponding to the 5′ end of the RNA typical for this way of cloning, involving removal of the terminal hairpin loop with S1 nuclease. Restriction sites on the insert were mapped according to the method of Smith and Birnstiel (1976) and the sequence was determined by the method of Maxam and Gilbert (1980). The length of the cloned hemagglutinin information in plasmid pDHA18 turned out to be 1569 bp. The structure was completed by sequencing a reverse transcript, obtained by extension of a primer close to the end (see Experimental Procedures).

Figure 1 is a restriction map of the A/duck/Ukraine/1/63 hemagglutinin gene, which shows the unique cleavage sites (above) and the sites of the restriction enzymes used for sequencing (below). All the sites used for labeling have also been read through from adjacent sites. In total, about 66% of the sequence has been read from both strands.

Comparison with Other Influenza A Hemagglutinins

The complete nucleotide sequence of the duck/Ukraine hemagglutinin gene is shown in Figure 2 with a comparison with the Aichi sequence (Verhoeyen et al., 1980). The total length of the hemagglutinin genomic RNA is 1765 nucleotides, exactly the same length as in Aichi and two other H3 strains, A/NT/60/68/29C (Both and Sleigh, 1980) and A/Memphis/102/72 (Sleigh et al., 1980), the only exception so far being a one-triplet insertion in the HA1 part of the hemagglutinin in A/Victoria/3/75 (Min Jou et al., 1980). Apart from this, they are all interrelated by nucleotide substitutions, and the different functional parts of the molecule have a constant length in all H3 strains—29 nucleotides precede the AUG signal for initiating translation, the translated region extends over 1698 nucleotides (specifying 566 amino acids: 16 in the signal peptide, 328 in HA1, 1 residue connecting HA1 and HA2, 221 in HA2) and the 3′-terminal noncoding part extends over 38 nucleotides, UGA being the terminator codon in all cases. The total length of the molecule, as well as the lengths of untranslated and translated regions, differs from those in H2 and H7 hemagglutinin subtype genes (Gething et al., 1980; Porter et al., 1979). So the duck/Ukraine hemagglutinin has the detailed organization found in the human H3 hemagglutinins.

The nucleotide sequence homology between the duck/Ukraine/63 and the human Aichi/68 hemagglutinin amounts to 90.9% and the amino acid sequence homology to 95.8% (Table 1). Based on RNA–RNA hybridization data, Scholtissek et al. (1978a) found a 92% nucleotide sequence homology between the hemagglutinin gene from duck/Ukraine and that from A/Hong Kong/1/68, a strain very closely related to Aichi. The amino acid sequence homology between the hemagglutinins of subtypes H2, H3 and H7 that have been completely sequenced is significantly lower—that is, between 36 and 48% (Gething et al., 1980; Verhoeyen et al., 1980; Porter et al., 1979). For comparison, the amino acid sequence divergence is 5.1% (94.9% homology) as the result of a 7 year drift period in the human population between Aichi/68 and Victoria/75. We conclude that the antigenic identification of the duck/Ukraine hemagglutinin (Hav7) as an H3 hemagglutinin by immunological and other criteria is fully corroborated by its complete primary structure.

It may be noted also that, in the comparison of duck/Ukraine/63 with Aichi/68, the HA2 part has remained relatively constant, presumably because its structure is optimized for an essential function, such as virus fusion (Gething et al., 1978). On the other

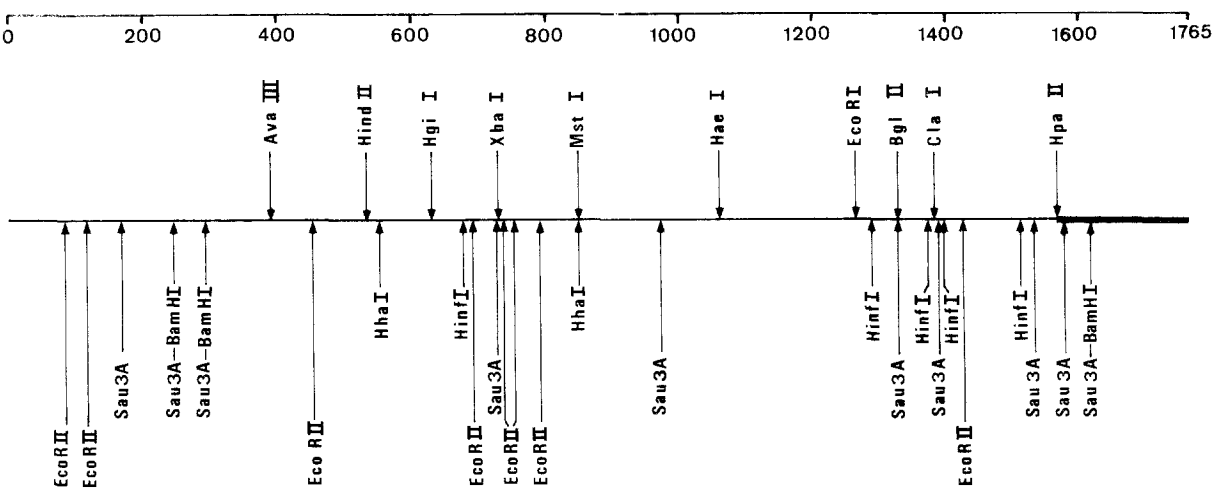

Figure 1. Restriction Map of the duck/Ukraine Hemagglutinin Insert in Plasmid pDHA18

(Above) the unique restriction sites found in the hemagglutinin information, on the basis of a computer search in the total nucleotide sequence and the list of enzymes compiled by Roberts (1980) with a few additions. (Below) the restriction sites of additional enzymes used for sequencing. Heavy line on the right: the part of the sequence that was missing from the insert and that was sequenced from primer-extended cDNA.

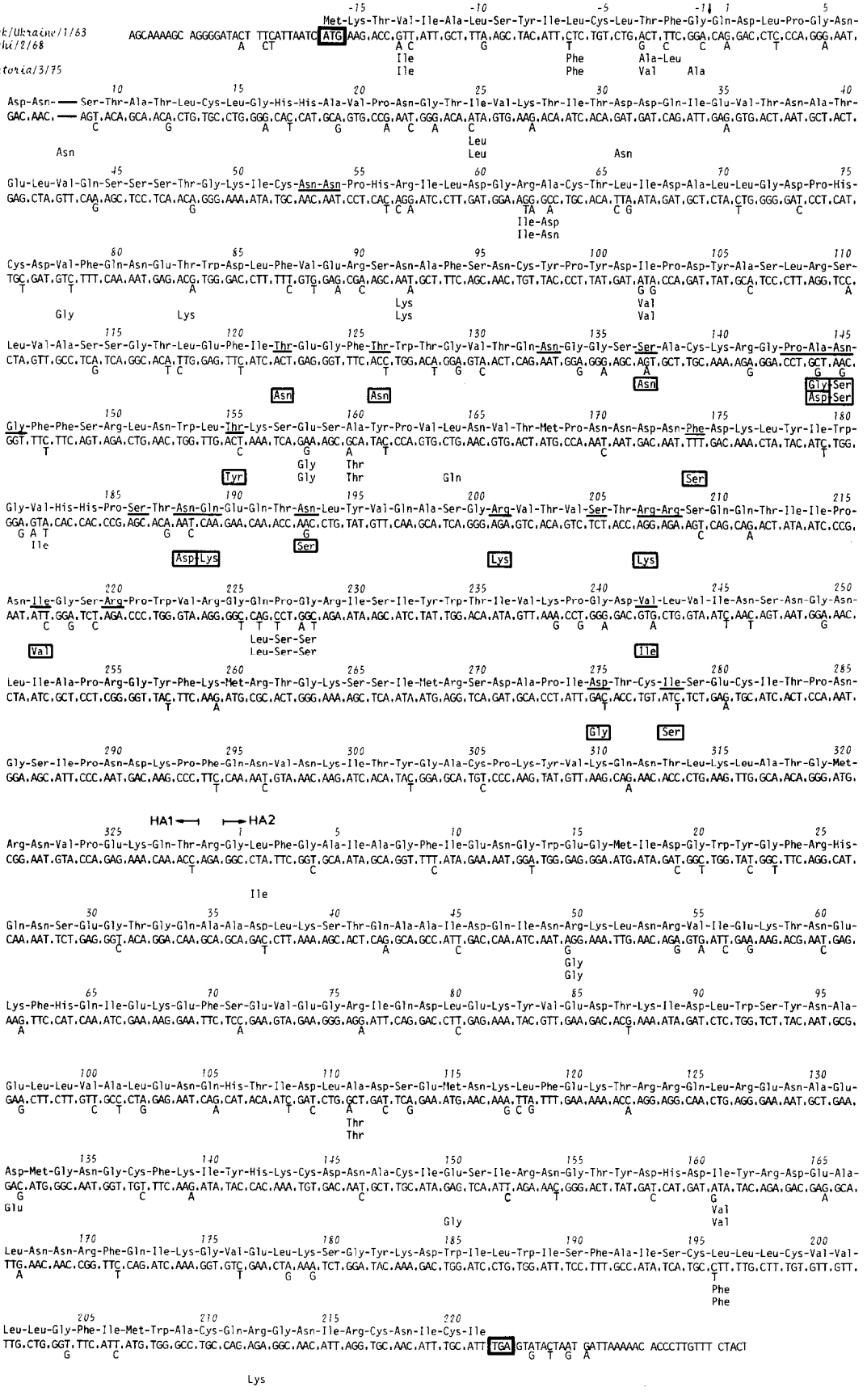
A/Duck/Ukraine/1/63
A/Aichi/2/68
A/Victoria/3/75
HA1
HA2

Table 1. Nucleotide and Amino Acid Sequence Divergence between duck/Ukraine/63 and Aichi/68 Hemagglutinin Genes

	Length (in Nucleotides)	Substitutions	Silent Mutations	Mutations Leading to an Amino Acid Change	Amino Acid Changes
5′ Noncoding region	29	3 (10.3%)			
Signal sequence	48	7 (14.6%)	3 (6.2%)	4 (8.3%)	4 (25.0%)
HA1	984	93 (9.5%)	77 (7.8%)	16 (1.6%)	15 (4.6%)
			23 (2.3%)[a]	*22 (2.2%)*[a]	*22 (6.7%)*[a]
Connecting peptide	3	0	0	0	0
HA2	663	53 (8.0%)	48 (7.2%)	5 (0.8%)	5 (2.3%)
			10 (1.5%)[a]	*4 (0.6%)*[a]	*4 (1.8%)*[a]
3′ Noncoding region	38	4 (10.5%)			
Total coding portion	1698	153 (9.0%)	128 (7.5%)	25 (1.5%)	24 (4.2%)
Total molecule	1765	160 (9.1%)			

[a] For comparison, these values show the divergence between Aichi/68 and Victoria/75 hemagglutinin (from Verhoeyen et al., 1980).

hand, the HA1 part has diverged much more, presumably in the absence of strong immunological pressure (see below), and this reflects the more extensive malleability of the structure–function relationship in this part of the hemagglutinin molecule.

Moreover, the duck/Ukraine amino acid sequence is more closely related to the human Aichi hemagglutinin/68 than to that of Victoria/75, which has undergone an additional 7 year drift period in humans. The total number of amino acid differences amounts to 24 for duck/Ukraine–Aichi and 40 for duck/Ukraine–Victoria, respectively (Figure 2). None of these changes affects the general framework of the molecule as determined by x-ray diffraction analysis of the Aichi hemagglutinin (Wilson et al., 1981; Wiley et al., 1981). While there are a number of changes in antigenically neutral positions in HA1, the amino acid differences in the antigenic sites of the molecule (Wiley et al., 1981) show a closer correlation between duck/Ukraine and Aichi than between duck/Ukraine and Victoria (Table 2). There are only four such changes between the first two strains affecting two of the four antigenic sites, whereas there are 14 amino acid differences between duck/Ukraine and Victoria in all four antigenic sites. Also, the duck/Ukraine sequence predicts an unchanged glycosylation pattern relative to Aichi, while it is known that the pattern in Victoria must be different (Min Jou et al., 1980; Verhoeyen et al., 1980). According to these comparisons, the duck/Ukraine and Aichi hemagglutinins are definitely more closely related structurally, and presumably also antigenically, than duck/Ukraine and Victoria. This could best be explained by a common ancestor gene.

A Recent Ancestor Hemagglutinin Gene for the Duck/Ukraine/63 and the Human H3 Strains

To assess the time of divergence of a common ancestor gene towards duck/Ukraine/63 and Aichi/68 respectively, we took the silent changes as an internal clock, assuming constant mutation rate independent of the host. This assumption requires a similar error frequency during RNA replication and a similar frequency of replication in the two hosts, and also that silent mutations are really neutral, which might not be completely true (Min Jou and Fiers, 1976; Verhoeyen et al., 1980). The estimation is based on the known number of silent mutations between Aichi (1968) and Victoria (1975), 0.334% per year in HA1 and 0.216% per year in HA2. There is a period of 5 years from 1963 to 1968 (Figure 3) from the postulated progenitor strain M to the appearance of Aichi, and an unknown period of time U for both M and duck/Ukraine to have derived from the common ancestor N. The number of silent mutations between Aichi and duck/Ukraine amounts to 7.83% in the HA1 part and 7.24% in the HA2 part. On this basis, the time of divergence (U) of the Aichi/68 from the postulated ancestor N can be estimated as 9.2 years based on the HA1 part ($U = [7.83\% - 5 \times 0.334\%]/2 \times 0.334\%$) or as 14.3 years based on the HA2 part ($U = [7.24\% - 5 \times 0.216\%]/2 \times 0.216\%$).

Figure 2. Nucleotide and Amino Acid Sequence of the duck/Ukraine/63 Hemagglutinin Gene

The nucleotide sequence is shown as the positive strand of the double-stranded DNA copy. It should be remembered that the viral RNA is of negative (opposite) polarity. The amino acid sequence is shown on top. Numbering is separate for the signal peptide (backwards) and for HA1 and HA2, respectively. Differences in nucleotide (Aichi) and amino acid sequence (Aichi and Victoria) are indicated underneath. Underlining: amino acid residues presumably involved in one of the four postulated antigenically important sites (Wiley et al., 1981) in the duck/Ukraine sequence. Boxes: corresponding changes in Aichi and/or Victoria.

Table 2. Amino Acid Sequence Differences among duck/Ukraine/63, Aichi/68 and Victoria/75 in the Four Presumptive Antigenic Sites of Hemagglutinin

Amino Acid Position in HA1	Amino Acid Residue in			Antigenic Site
	duck/Ukraine/63	Aichi/68	Victoria/75	
122	Thr	Thr	Asn	A*
126	Thr	Thr	Asn	A*
137	Ser	Asn	Ser	A*
144	Ala	Gly	Asp	A
145	Asn	Ser	Ser	A
155	Thr	Thr	Tyr	B
174	Phe	Phe	Ser	D*
188	Asn	Asn	Asp	B
189	Gln	Gln	Lys	B
193	Asn	Ser	Asn	B
201	Arg	Arg	Lys	D
207	Arg	Arg	Lys	D
217	Ile	Ile	Val	D
242	Val	Val	Ile	D*
275	Asp	Asp	Gly	C
278	Ile	Ile	Ser	C

There are four postulated antigenic sites, A, B, C and D, on the hemagglutinin (Wiley et al., 1981).
* It is considered uncertain whether the amino acid in this position is part of the site.

The influenza strain carrying the common progenitor hemagglutinin gene can thus be estimated to have been circulating around 9 to 14 years before 1963—that is, in the period 1949–1953. As far as is known, at that time only strains of the H1N1 subtype were actively replicating in the human population (Dowdle, 1976), so the common ancestor was almost certainly present in the animal reservoir (for example, duck). No direct correlation between the 1890 H3 hemagglutinin and that from Aichi/68 is possible, as the first human influenza virus was isolated in 1933 (Smith et al., 1933). In theory it is possible that this virus has been sequestered in some way and that its hemagglutinin simply reappeared in 1968. However, unlike the 1977 event, where the whole virus reappeared essentially unchanged (Nakajima et al., 1978; Scholtissek et al., 1978b), this cannot be the case here. There is good evidence that the H3N2 subtype emerging in 1968 retained seven of the eight genes of the previous H2N2-subtype virus in humans and received only the H3 hemagglutinin from another nonidentified virus (Scholtissek et al., 1978a). Also, the virus circulating at the end of the previous century had an N8 (Neq 2) and not an N2 neuraminidase gene. For both reasons mentioned above, it is unlikely that the H3 appearing in 1968 is the old 1890 human hemagglutinin reappearing, as this would have required an immediate reassortment as well.

Replication of Influenza A Viruses in Ducks

The evolution from about 1950 to 1963 (duck/Ukraine) and from about 1950 to 1968 (Aichi) must have been mostly in the absence of strong immunological selection, such as occurs in at least some animal reservoirs like ducks (see below) but not in man. Indeed, as can be deduced from the results summarized in Table 1, the phenotypic mutant yield (number of amino acid changes per 100 nucleotide substitutions) is 16% in the HA1 part for the comparison of duck/Ukraine/63 with Aichi/68, while it amounts to 46% for Aichi/68 compared with Victoria/75. This is because in the latter situation, where there is immunological selection, each mutant that happens to arise in an antigenic site will have a positive selective value and will spread through the population. The above conclusion is further substantiated by the fact that, out of the 15 amino acid differences between duck/Ukraine/63 and Aichi/68, only 4 are in the postulated antigenic sites as defined by Wiley et al. (1981)—that is, a ratio of 0.27—while out of 22 amino acid differences between Aichi/68 and Victoria/75, 15 are part of such antigenic sites—that is, a ratio of 0.68.

Influenza A infections in ducks do not show the epidemic character observed in man, leading to the establishment of one particular strain in large areas. Rather, a number of strains of different subtype cocirculates at any moment in the duck population (Hinshaw et al., 1979). Between viruses of the same subtype the other genes vary widely (Sriram et al., 1980). This genetic diversity in avian influenza viruses strongly suggests reassortment of genes. The presence of two or more antigenically distinguishable influenza viruses in 7% of the cloacal samples from feral ducks has been demonstrated, and reassortment has been directly proven by the study of viruses shed by experimentally infected ducks (Hinshaw et al., 1980). The same authors have shown that juvenile ducks infected with influenza virus shed virus in their feces for 30 days and produce very low levels of circulating antibodies. The prolonged replication of influenza viruses in ducks again increases the chance for coinfection and reassortment.

In the human population we are at present in a previously undocumented situation of cocirculation of two influenza A subtypes, H1N1 and H3N2. Recently a reassortment event between both subtypes was described, leading to a new H1N1 variant carrying four genes coding for internal proteins from the cocirculating H3N2 parent, and thus probably providing a selective advantage to the virus (Young and Palese, 1979). All these data reinforce the idea that genetic reassortment plays an important role in the evolution of influenza A viruses. Although directly proven in experimental conditions (Hinshaw et al., 1980) or in a carefully controlled (human) population with only two types of virus known to be circulating (Young and

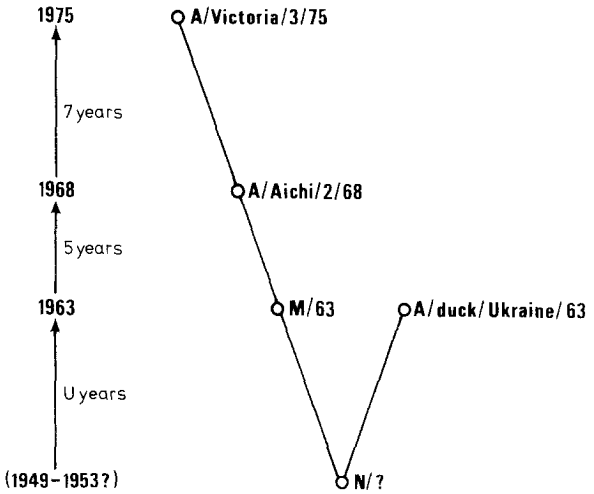

Figure 3. Genealogical Tree of H3 Hemagglutinins from Humans and from an Animal Reservoir

The evolution of Aichi to Victoria corresponds to a 7 year drift period in the human population. M/63 is a postulated strain cocirculating in an animal (for example, duck) reservoir with duck/Ukraine; N is the proposed common ancestor. The time of divergence was estimated (see text) based on the accumulation of silent base substitutions in Victoria, Aichi and duck/Ukraine.

Palese, 1979), similar, solid evidence for the existence and importance of reassortment between an animal reservoir and the human population would be difficult to obtain. We have therefore carried out a direct structural comparison between an H3 hemagglutinin from a human influenza A strain and one from the duck strain, to see whether they had a common progenitor in the recent past.

Duck viruses of the same subtype isolated from one geographical area at the same time are antigenically indistinguishable. Nevertheless, they were shown to vary quite extensively in their hemagglutinin and neuraminidase RNAs (Sriram et al., 1980), presumably by the accumulation of a relatively large number of silent nucleotide substitutions. This again illustrates that the way the virus spreads in ducks is much more local as compared with that in humans, and that it spreads without strong immunological pressure.

In conclusion, the comparison of the hemagglutinin genes from duck/Ukraine with Aichi and other human H3 strains constitutes virtual proof that the Aichi hemagglutinin gene had its origin in animals (for example, ducks) and that it diverged relatively recently from a common ancestor (probably around 1949–1953). Their similarities (in amino acid sequence and probably in antigen specificity) and differences (silent nucleotide changes) are in agreement with the properties of hemagglutinins belonging to the same subtype cocirculating in ducks. These results strongly support the hypothesis of an animal reservoir for human influenza A virus genes, as reviewed by Webster and Laver (1975).

Experimental Procedures

Most of the procedures used in this work have been described previously (Devos et al., 1979; Min Jou et al., 1980). The virus A/duck/Ukraine/1/63 was obtained from J. Skehel, who received it originally from B. Tumova in 1968. After three passages in eggs it was stored frozen; an aliquot was sent to us in 1980. It was grown on a preparative scale in embryonated eggs, and this yielded the virus preparation used for RNA extraction and subsequent cloning. A sample of the same virus batch was sent back to J. Skehel. It was found to be indistinguishable from the original virus stock on the basis of the same reference antisera as those used in the work leading to the new nomenclature system for influenza viruses (Schild et al., 1980). To sequence the information missing at the 5′ end of the viral RNA, a restriction fragment Hinf I–Sau 3A located near the end of the insert in pDHA18 was isolated and used as a primer for reverse transcription. It was obtained from a Hinf I digest of 20 μg plasmid DNA, 5′-^{32}P-labeled and recut with Sau 3A. The Hinf I–Sau 3A fragment (labeled strand, 23 nucleotides; complementary strand, 24 nucleotides) was denatured in 40% formamide, 1 mM EDTA (pH 8.0) at 90°C for 2 min, chilled in dry ice/ethanol and applied to a 10% acrylamide strand-separation gel (acrylamide:bisacrylamide concentration, 50:1) run at 4°C according to the method of Maxam and Gilbert (1980). The 5′-^{32}P-labeled single-stranded DNA was recovered from the gel and precipitated with ethanol. It was mixed with 10–15 μg duck/Ukraine viral RNA in water, heated at 65°C for 3 min and then kept at 41°C for 30 min in 100 mM KCl. Reverse transcription was carried out in 50 mM Tris-HCl (pH 8.3), 10 mM $MgCl_2$, 30 mM β-mercaptoethanol, 0.1 mM of each dNTP and 30–40 U reverse transcriptase for 1 hr at 41°C. The RNA was hydrolyzed in 0.3 N NaOH for 4 hr at 37°C and adjusted to pH 7.0 with HCl, and the cDNA transcripts were precipitated with ethanol. This cDNA was used for sequencing according to the method of Maxam and Gilbert (1980). Alternatively, the full-size cDNA was purified by electrophoresis on a 6% acrylamide gel prior to sequencing.

Acknowledgments

We thank Dr. J. Skehel for kindly providing the A/duck/Ukraine/1/63 strain and for serological tests; M. Verhoeyen for advice; and Dr. J. W. Beard and the Office of Program and Logistics of the NIH for a generous gift of AMV reverse transcriptase. This research was supported by grants from the Fonds voor Kollektief Fundamenteel Onderzoek and the Gekoncerteerde Akties of the Belgian Ministry of Science.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 20, 1981; revised May 7, 1981

References

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. *7*, 1513–1523.

Blok, J. and Air, G. M. (1980). Comparative nucleotide sequences at the 3′ end of the neuraminidase gene from eleven influenza type A viruses. Virology *107*, 50–60.

Both, G. W. and Sleigh, M. J. (1980). Complete nucleotide sequence of the haemagglutinin gene from a human influenza virus of the Hong Kong subtype. Nucl. Acids Res. *8*, 2561–2575.

Choppin, P. W. (1977). Influenza: the agent. In Hospital Practice: Status Report on Influenza, W. Dial, ed. (New York: HP Publishing Co.), pp. 15–23.

Coleman, M. T., Dowdle, W. R., Pereira, H. G., Schild, G. C. and Chang, W. K. (1968). The Hong Kong/68 influenza A2 variants. Lancet *2*, 1384–1386.

Davis, A. R., Hiti, A. L. and Nayak, D. P. (1980). Construction and characterization of a bacterial clone containing the hemagglutinin gene of the WSN strain (H0N1) of influenza virus. Gene *10,* 205–218.

Devos, R., van Emmelo, J., Contreras, R. and Fiers, W. (1979). Construction and characterization of a plasmid containing a nearly full-size DNA copy of bacteriophage MS2 RNA. J. Mol. Biol. *128,* 595–619.

Dopheide, T. A., and Ward, C. W. (1980). Structural studies on a Hong Kong influenza hemagglutinin. The structure of the light chain and the arrangements of the disulphide bonds. In Structure and Variation in Influenza Virus, G. Laver and G. Air, eds. (New York: Elsevier North-Holland), pp. 21–26.

Dowdle, W. R. (1976). Influenza: epidemic patterns and antigenic variation. In Influenza Virus, Vaccines, and Strategy, P. Selby, ed. (London: Academic Press), pp. 17–21.

Fedson, D. S., Huber, M. A., Kasel, J. A. and Webster, R. G. (1972). Presence of A/Equi-2 hemagglutinin and neuraminidase antibodies in man. Proc. Soc. Exp. Biol. Med. *139,* 825–826.

Gething, M. J., White, J. M. and Waterfield, M. D. (1978). Purification of the fusion protein of Sendai virus: analysis of the NH_2-terminal sequence generated during precursor activation. Proc. Nat. Acad. Sci. USA *75,* 2737–2740.

Gething, M.-J., Bye, J., Skehel, J. and Waterfield, M. (1980). Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. Nature *287,* 301–306.

Grunstein, M. and Hogness, D. S. (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Nat. Acad. Sci. USA *72,* 3961–3965.

Hinshaw, V. S., Webster, R. G. and Rodriguez, R. J. (1979). Influenza A viruses: combinations of hemagglutinin and neuraminidase subtypes isolated from animals and other sources. Arch. Virol. *62,* 281–290.

Hinshaw, V. S., Bean, W. J., Webster, R. G. and Sriram, G. (1980). Genetic reassortment of influenza A viruses in the intestinal tract of ducks. Virology *102,* 412–419.

Inglis, S. C., Barrett, T., Brown, C. M. and Almond, J. W. (1979). The smallest genome RNA segment of influenza virus contains two genes that may overlap. Proc. Nat. Acad. Sci. USA *76,* 3790–3794.

Kahn, M., Kolter, R., Thomas, C., Figurski, D., Meyer, R., Remaut, E. and Helinski, D. R. (1979). Plasmid cloning vehicles derived from plasmids Col E1, F, R6K and RK2. Meth. Enzymol. *68,* 268–280.

Kendal, A. P., Noble, G. R., Skehel, J. J. and Dopheide, W. R. (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.

Lamb, R. A. and Choppin, P. W. (1979). Segment 8 of the influenza virus genome is unique in coding for two polypeptides. Proc. Nat. Acad. Sci. USA *76,* 4908–4912.

Laver, W. G. and Kilbourne, E. D. (1966) Identification in a recombinant influenza virus of structural proteins derived from both parents. Virology *30,* 493–501.

Laver, W. G. and Webster, R. G. (1972). Studies on the origin of pandemic influenza. II. Peptide maps of the light and heavy polypeptide chains from the hemagglutinin subunits of A2 influenza viruses isolated before and after the appearance of Hong Kong influenza. Virology *48,* 445–455.

Laver, W. G. and Webster, R. G. (1973). Studies on the origin of pandemic influenza. III. Evidence implicating duck and equine influenza viruses as possible progenitors of the Hong Kong strain of human influenza. Virology *51,* 383–391.

Laver, W. G. and Webster, R. G. (1977). Hemagglutinin molecules of Hong Kong, equine-2, and duck/ukraine influenza viruses lack N-terminal aspartic acid. Virology *81,* 482–485.

Laver, W. G. and Webster, R. G. (1979). Ecology of influenza viruses in lower mammals and birds. Brit. Med. Bull. *35,* 29–33.

Laver, W. G., Air, G. M., Dopheide, T. A. and Ward, C. W. (1980). Amino acid sequence changes in the haemagglutinin of A/Hong Kong (H3N2) influenza virus during the period 1968–1977. Nature *283,* 454–457.

Masurel, N. and Marine, W. M. (1973). Recycling of asian and Hong Kong influenza A virus hemagglutinin in man. Am. J. Epidemiol. *97,* 44–49.

Maxam, A. M. and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. *65,* 499–560.

Min Jou, W. and Fiers, W. (1976). Studies on the bacteriophage MS2. XXXIII. Comparison of the nucleotide sequences in related bacteriophage RNAs. J. Mol. Biol. *106,* 1047–1060.

Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N. and Emtage, S. (1980). Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. Cell *19,* 683–696.

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

Ortin, J., Najera, R., Lopez, C., Davila, M. and Domingo, E. (1980). Genetic variability of Hong Kong (H3N2) influenza viruses: spontaneous mutations and their location in the viral genome. Gene *11,* 319–331.

Palese, P. (1977). The genes of influenza virus. Cell *10,* 1–10.

Porter, A. G., Barber, C., Carey, N. H., Hallewell, R. A., Threlfall, G. and Emtage, J. S. (1979). Complete nucleotide sequence of an influenza virus haemagglutinin gene from cloned DNA. Nature *282,* 471–477.

Roberts, R. J. (1980). Restriction and modification enzymes and their recognition sequences. Nucl. Acids Res. *8,* r63–r80.

Schild, G. C., Newman, R. W., Webster, R. G., Major, D. and Hinshaw, V. S. (1980). Antigenic analysis of influenza A virus surface antigens: considerations for the nomenclature of influenza virus. Arch. Virol. *63,* 171–184.

Scholtissek, C., Rohde, W., von Hoyningen, V. and Rott, R. (1978a). On the origin of the human influenza virus subtypes H2N2 and H3N2. Virology *87,* 13–30.

Scholtissek, C., von Hoyningen, V. and Rott, R. (1978b). Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza strains isolated between 1947 and 1957 (H1N1). Virology *89,* 613–617.

Sleigh, M. J., Both, G. W., Brownlee, G. G., Bender, V. J. and Moss, B. A. (1980). The haemagglutinin gene of influenza A virus: nucleotide sequence analysis of cloned DNA copies. In Structure and Variation in Influenza Virus, G. Laver and G. Air, eds. (New York: Elsevier North-Holland), pp. 69–79.

Smith, H. O. and Birnstiel, M. L. (1976). A simple method for DNA-restriction site mapping. Nucl. Acids Res. *3,* 2387–2398.

Smith, W., Andrewes, C. H. and Laidlaw, P. P. (1933). A virus obtained from influenza patients. Lancet *2,* 66–68.

Sriram, G., Bean, W. J., Jr., Hinshaw, V. S. and Webster, R. G. (1980). Genetic diversity among avian influenza viruses. Virology *105,* 592–599.

Verhoeyen, M., Fang, R.-X., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E. and Fiers, W. (1980). Antigenic drift between the hemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75. Nature *286,* 771–776.

Ward, C. W. and Dopheide, T. A. (1980). The Hong Kong (H3) hemagglutinin. Complete amino acid sequence and oligosaccharide distribution for the heavy chain of A/Memphis/102/72. In Structure and Variation in Influenza Virus, G. Laver and G. Air, eds. (New York: Elsevier North-Holland), pp. 27–38.

Waterfield, M. D., Espelie, K., Elder, K. and Skehel, J. J. (1979). Structure of the haemagglutinin of influenza virus. Brit. Med. Bull. *35,* 57–63.

Webster, R. G. and Laver, W. G. (1975). Antigenic variation of influenza viruses. In The Influenza Viruses and Influenza, E. D. Kil-

bourne, ed. (New York: Academic Press), pp. 269–314.

Wiley, D. C., Wilson, I. A. and Skehel, J. J. (1981). Structural identification of the antibody binding sites of the Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature *289*, 373–378.

Wilson, I. A., Skehel, J. J. and Wiley, D. C. (1981). The haemagglutinin membrane glycoprotein of influenza virus: structure at 3 Å resolution. Nature *289*, 366–373.

World Health Organization. (1980). A revision of the system of nomenclature for influenza viruses—a WHO memorandum. Bull WHO *58*, 294–295.

Young, J. F. and Palese, P. (1979). Evolution of human influenza A viruses in nature. Recombination contributes to genetic variation of H1N1 strains. Proc. Nat. Acad. Sci. USA *76*, 6547–6551.

Young, J. F., Desselberger, U. and Palese, P. (1979). Evolution of human influenza A viruses in nature: sequential mutations in the genomes of new H1N1 isolates. Cell *18*, 73–83.

Zakstelskaja, L. J., Evstigneeva, N. A., Isachenko, V. A., Shenderovith, S. P. and Efimova, V. A. (1969). Influenza in the USSR: new antigenic variant A2/Hong Kong/1/68 and its possible precursors. Am. J. Epidemiol. *90*, 400–405.

Note Added in Proof

After this paper had been accepted, information on the protein structure of A/duck/Ukraine/1/63 hemagglutinin as present in a recombinant virus was reported by Ward and Dopheide (Biochem. J. *195*, 337–340, 1981). When compared with our data, there are amino acid differences at positions 4, 63, 81, 102, 135, 145, 158, 160, 186, 227 and 309 in HA1 and at positions 2, 50, 67, 71, 106, 111, 133 and 154 in HA2. These rather extensive differences presumably have to be explained on the basis of the origins and histories of the virus stocks used in these studies.