JOURNAL OF VIROLOGY, Oct. 1990, p. 4893–4902
0022-538X/90/104893-10$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 64, No. 10

Evolution of Influenza A Virus PB2 Genes: Implications for Evolution of the Ribonucleoprotein Complex and Origin of Human Influenza A Virus

OWEN T. GORMAN,* RUBEN O. DONIS,† YOSHIHIRO KAWAOKA, AND ROBERT G. WEBSTER

Department of Virology and Molecular Biology, St. Jude Children's Research Hospital, 332 North Lauderdale, P.O. Box 318, Memphis, Tennessee 38101-0318

Received 19 April 1990/Accepted 6 July 1990

Phylogenetic analysis of 20 influenza A virus PB2 genes showed that PB2 genes have evolved into the following four major lineages: (i) equine/Prague/56 (EQPR56); (ii and iii) two distinct avian PB2 lineages, one containing FPV/34 and H13 gull virus strains and the other containing North American avian and recent equine strains; and (iv) human virus strains joined with classic swine virus strains (i.e., H1N1 swine virus strains related to swine/Iowa/15/30). The human virus lineage showed the greatest divergence from its root relative to other lineages. The estimated nucleotide evolutionary rate for the human PB2 lineage was 1.82×10^{-3} changes per nucleotide per year, which is within the range of published estimates for NP and NS genes of human influenza A viruses. At the amino acid level, PB2s of human viruses have accumulated 34 amino acid changes over the past 55 years. In contrast, the avian PB2 lineages showed much less evolution, e.g., recent avian PB2s showed as few as three amino acid changes relative to the avian root. The completion of evolutionary analyses of the PB1, PB2, PA, and NP genes of the ribonucleoprotein (RNP) complex permits comparison of evolutionary pathways. Different patterns of evolution among the RNP genes indicate that the genes of the complex are not coevolving as a unit. Evolution of the PB1 and PB2 genes is less correlated with host-specific factors, and their proteins appear to be evolving more slowly than NP and PA. This suggests that protein functional constraints are limiting the evolutionary divergence of PB1 and PB2 genes. The parallel host-specific evolutionary pathways of the NP and PA genes suggest that these proteins are coevolving in response to host-specific factors. PB2s of human influenza A viruses share a common ancestor with classic swine virus PB2s, and the pattern of evolution suggests that the ancestor was an avian virus PB2. This same pattern of evolution appears in the other genes of the RNP complex. Antigenic studies of HA and NA proteins and sequence comparisons of NS and M genes also suggest a close ancestry for these genes in human and classic swine viruses. From our review of the evolutionary patterns of influenza A virus genes, we propose the following hypothesis: the common ancestor to current strains of human and classic swine influenza viruses predated the 1918 human pandemic virus and was recently derived from the avian host reservoir.

Three polymerases (PB1, PB2, and PA), the nucleoprotein (NP), and the viral RNAs compose the ribonucleoprotein (RNP) complex of influenza A viruses. The polymerases of this complex are responsible for the transcription and replication of viral RNAs, and the NP encapsidates the negative-strand viral RNAs (26). RNP complexes must be transported from and to the host cell membrane and operate efficiently in the host cell nucleus. Since these proteins form a functional complex, they may also be coevolving as a unit. Furthermore, since these are internal proteins of the virion, they are probably not subjected to selective pressure by host immune systems as strongly as the antigenic surface proteins hemagglutinin (HA) and neuraminidase (NA). Instead, evolution of these internal proteins may be slower and reflect host-specific adaptation.

Because of the potential host-specific nature of the proteins of the RNP complex, their genes make excellent candidates for the investigation of evolutionary relationships among influenza A viruses of different species. A number of questions may be addressed by comparing the evolutionary pathways of the RNP genes. (i) Do some RNP genes show parallel host-specific evolutionary pathways? Congruent patterns of evolution among some genes would indicate that these genes do not reassort independently and that the proteins of these genes have coadapted to specific hosts. (ii) Alternatively, do some RNP genes show evolutionary pathways that are unrelated to host adaptation? Evidence of host-independent evolution would include reduced levels of lineage divergence, a lack of well-defined host-specific lineages, and the ability to reassort independently. For these genes, virus protein function may be constraining gene evolution to a greater extent than host adaptation. (iii) Are there common evolutionary patterns among influenza A virus genes that suggest when and how present strains of viruses originated? If some or all of the genes of a host-specific virus show the same evolutionary pathway, this indicates a common origin for the genes. The relationship of this common evolutionary pathway to those for the genes of other influenza A viruses can provide evidence of derivation.

We present the results of a phylogenetic analysis of 14 influenza A virus PB2 gene sequences along with the results for 6 previously published PB2 sequences (Table 1). These PB2 genes represent a wide variety of influenza A virus serotypes, hosts, and geographic regions. From our analysis of the evolution of PB2 genes together with previous evolutionary studies of the NP (16), PA (31), and PB1 (23), we address the questions described above concerning the evo-

* Corresponding author.
† Present address: Department of Veterinary Science, University of Nebraska, Lincoln, NE 68583.

TABLE 1. Influenza virus strains used in phylogenetic analyses

Strain	Abbreviation	Source
1. B/Ann Arbor/66	B-AA66	DeBorde et al. (13)
2. A/Equine/Prague/1/56 (H7N7)	EQPR56	This report
3. A/Swine/Tennessee/24/77 (H1N1)	SWTN77	This report
4. A/Swine/Iowa/15/30(H1N1)	SWIA30	This report
5. A/WSN/33 (H1N1)	WSN33	Kaptein and Nayak (22)
6. A/PR/8/34 (H1N1)	PR8-34	Fields and Winter (14)
7. A/Chile/1/83 (H1N1)	CHILE83	Schreier et al. (37)
8. A/Singapore/1/57 (H2N2)	SING57	This report
9. A/Korea/426/68 (H2N2)	KOREA68	This report
10. A/NT/60/68 (H3N2)	NT60-68	Jones et al. (21)
11. A/Memphis/8/88 (H3N2)	MEM88	This report
12. A/Equine/London/1416/73 (H7N7)	EQLON73	This report
13. A/Equine/Kentucky/2/86 (H3N8)	EQKY86	This report
14. A/Mallard/NY/6750/78 (H2N2)	MLRDNY78	Treanor et al. (38)
15. A/Turkey/Minnesota/833/80 (H4N2)	TYMN80	This report
16. A/Seal/Massachusetts/133/82 (H4N5)	SEAL82	This report
17. A/FPV/34 (H7N7)	FPV34	Roditi and Robertson (32)
18. A/Ruddy Turnstone/NJ/47/85 (H4N6)	RTNJ85	This report
19. A/Budgerigar/Hokkaido/1/77 (H4N6)	BUDHOK77	This report
20. A/Gull/Maryland/704/77 (H13N6)	GULMD77	This report
21. A/Gull/Astrakhan/227/84 (H13N6)	GULAST84	This report

lution of influenza A virus RNP genes and propose a hypothesis for the origin of influenza A viruses in humans and swine.

MATERIALS AND METHODS

Virus strains. We selected 14 viral isolates from the repository at St. Jude Children's Research Hospital (Memphis, Tenn.). Isolates were selected to represent a wide spectrum of geographic locations, hosts, and dates of isolation and to complement seven gene sequences from literature and data bank sources (Table 1).

Molecular cloning of the PB2 genes. Viruses were grown in 11-day-old embryonated chicken eggs, and RNA was extracted as described by Bean et al. (4). Full-length PB2 genes were cloned by the procedures of Jones et al. (21) and Winter et al. (40). To summarize, a 12-mer oligodeoxynucleotide primer (5′AGCAAAAGCAGG) was phosphorylated with T4 polynucleotide kinase, and first-strand cDNA was synthesized from the viral RNA template by using avian myeloblastosis virus reverse transcriptase. Second-strand synthesis was done with a phosphorylated 13-mer oligodeoxynucleotide primer (5′AGTAGAAACAAGC) and *Escherichia coli* DNA polymerase I (Klenow fragment). Full-length double-stranded cDNA was blunt end ligated into the *Pvu*II site of pATX vector (a derivative of pAT153 provided by Clayton Naeve, St. Jude Children's Research Hospital). Plasmids were transfected into competent *E. coli* (strain MC-1061) cells, and clones were detected by hybridization of colony screens with strain-specific, short-length labeled cDNA PB2 probes.

Nucleotide sequence determination. Nucleotide sequences of PB2 cDNAs were determined by using the dideoxynucleotide chain termination method (33). Oligodeoxynucleotide primers were annealed to double-stranded template DNA denatured with NaOH as described by Chen and Seeburg (9) and extended with modified T7 DNA polymerase (Sequenase; United States Biochemical Corp., Cleveland, Ohio). Reaction products were separated on 6% polyacrylamide–7 M urea 0.4-mm gels. Usually only one clone for each strain was sequenced, which is adequate for evolutionary analyses. Ambiguities were resolved by sequencing the complementary strand. The sequences of oligonucleotides used as primers are available upon request.

Sequence analysis. The Intelligenetics (Palo Alto, Calif.) software package was used for analysis and translation of nucleotide sequence data. The B/Ann Arbor/1/66 (B-AA66) PB2 nucleotide sequence (2,396 bases [13]) and amino acid sequence (770 bases [13]) were aligned with the influenza A PB2s by using the Needleman-Wunsch pairwise alignment algorithm. Phylogenetic analysis of sequence data was performed with the PAUP (Phylogenetic Analysis Using Parsimony) software package version 2.4.1 (David Swofford, Illinois Natural History Survey, Champaign, Ill.). PAUP employs the maximum parsimony method to generate phylogenetic trees. The trees with the shortest lengths (most parsimonious) were found by implementing the MULPARS, SWAP=GLOBAL, and HOLD=10 options of PAUP. The PRINTD option of PAUP provided difference matrices for sequence data. Tree length is measured in "steps" which are equivalent to nucleotide changes for nucleotide trees and amino acid changes for amino acid trees. Total tree length is the sum of all branch lengths.

Nucleotide sequence accession numbers. The nucleotide sequence accession numbers (EMBL/GenBank accession numbers) for the 14 cloned PB2 genes studied are as follows: M36036, influenza A/Swine/Tennessee/24/77 (H1N1); M36037, influenza A/Turkey/Minnesota/833/80 (H4N2); M36038, influenza A/Swine/Iowa/15/30 (H1N1); M36039, influenza A/Gull/Astrakhan/227/84 (H13N6); M36040, influenza A/Memphis/8/88 (H3N2); M36041, influenza A/Ruddy Turnstone/New Jersey/47/85 (H4N6); M36042, influenza A/Equine/Prague/1/56 (H7N7); M36043, influenza A/Equine/London/1416/73 (H7N7); M36044, influenza A/Singapore/1/57 (H2N2); M36045, influenza A/Seal/Massachusetts/133/82 (H4N5); M36046, influenza A/Budgerigar/Hokkaido/1/77 (H4N6); M36047, influenza A/Korea/426/68 (H2N2); M36048, influenza A/Gull/Maryland/704/77 (H13N6); and M36049, influenza A/Equine/Kentucky/2/86 (H3N8).

RESULTS

Comparisons of PB2 nucleotide sequences. Each of the 14 cloned influenza A virus PB2 genes (Table 1) comprised

DIFFERENCES IN AMINO ACID SEQUENCES

	VIRUS	B	EP	SW		HUMAN							EQ		AVIAN						H13	
	STRAIN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	B-AA66	0	471	478	472	475	474	477	475	475	473	476	473	475	474	471	470	472	470	470	473	475
2	EQPR56	1262	0	74	67	75	72	76	72	73	74	82	61	60	59	56	54	65	56	60	60	61
3	SWTN77	1252	436	0	35	50	48	51	50	44	47	54	48	54	50	43	43	54	43	49	48	47
4	SWIA30	1263	437	239	0	37	34	40	36	34	39	48	36	42	31	28	29	39	27	36	34	35
5	WSN33	1261	448	326	200	0	26	29	27	31	35	44	37	43	36	32	33	41	30	40	35	38
6	PR8-34	1261	455	320	197	90	0	26	22	24	29	37	30	36	32	25	28	37	25	35	29	32
7	CHILE83	1269	443	345	253	178	151	0	18	22	27	35	37	45	38	33	34	43	31	40	36	39
8	SING57	1264	454	346	240	159	136	143	0	10	14	25	34	40	35	30	31	40	28	36	33	36
9	KOREA68	1265	449	344	257	192	173	182	63	0	7	17	36	43	37	30	31	42	29	39	31	34
10	NT60-68	1274	460	359	282	225	200	209	93	49	0	18	39	46	40	33	34	45	32	41	34	37
11	MEM88	1272	468	372	309	250	232	238	153	111	112	0	48	56	49	42	42	52	41	51	43	44
12	EQLON73	1279	434	357	288	298	283	314	320	317	327	344	0	12	21	16	17	27	15	21	21	22
13	EQKY86	1273	434	371	311	318	307	337	335	332	338	365	63	0	26	22	23	33	21	27	26	27
14	MLRDNY78	1265	435	374	308	314	308	357	341	347	357	381	231	261	0	13	14	23	12	21	17	18
15	TYMN80	1269	434	339	281	289	286	333	324	321	338	367	213	236	223	0	7	19	6	16	12	13
16	SEAL82	1281	428	353	292	301	292	332	337	341	354	382	225	259	229	111	0	19	7	17	13	14
17	FPV34	1265	403	397	340	357	341	372	369	368	381	394	336	354	336	325	345	0	18	27	24	26
18	RTNJ85	1262	416	382	336	369	355	387	364	366	366	392	328	343	347	323	322	211	0	15	12	13
19	BUDHOK77	1268	419	406	351	386	373	408	406	407	414	429	365	376	366	347	357	279	221	0	22	23
20	GULMD77	1276	424	412	365	386	379	408	400	400	400	402	368	380	368	356	364	269	223	238	0	5
21	GULAST84	1268	432	410	366	389	380	408	396	397	399	398	372	386	378	366	371	277	220	241	47	0

DIFFERENCES IN NUCLEOTIDE SEQUENCES

FIG. 1. Difference matrix for influenza B virus B-AA66 PB2 and 20 influenza A virus PB2 nucleotide and amino acid sequences. Abbreviations: B, B-AA66; EP, EQPR56; SW, classic H1N1 swine; EQ, recent equine. Numbers along margins of the matrix correspond to virus isolates listed in Table 1.

2,341 nucleotides and a single open reading frame which spanned positions 28 through 2307 and coded for a polypeptide of 759 amino acids. No insertions or deletions were observed in any of the sequences. Nucleotide sequences of the 14 cloned PB2 genes are not presented here but are available from GenBank (see Materials and Methods).

The aligned influenza B virus PB2 (B-AA66) sequence shows 1,252 to 1,281 differences (54.6 to 53.5% difference) compared with that of influenza A virus PB2s (Fig. 1). Over all influenza A virus PB2 genes, the maximum absolute number of nucleotide differences was found between the PB2s of EQPR56 and those of the most recent swine and human virus isolates (468 differences or 20%; Fig. 1). Large differences occurred between PB2s of the swine-human lineage and avian lineages, and the oldest PB2s were more similar to avian PB2s (Fig. 1). This pattern of evolution suggests that PB2s are diverging from an avian virus origin.

Evolutionary tree of PB2 nucleotide sequences. A phylogenetic analysis of 20 influenza A virus PB2 nucleotide sequences (from the 14 cloned genes and 6 published sequences; Table 1) is presented as an evolutionary tree rooted to an aligned influenza B virus PB2 (B-AA66) sequence (Fig. 2, left). With the exclusion of the B-AA66 sequence, the shortest (most parsimonious) tree found by PAUP was 2,425 steps (nucleotide changes) long (Fig. 2, left). The next shortest tree was 13 steps longer and differed in the terminal branch placement of SWTN77 and MLRDNY78 within their respective lineages. Structural similarities found between influenza A and B virus hemagglutinins (25) suggest a close evolutionary relationship between the two viruses and support the choice of an influenza B virus PB2 as the outgroup to root the tree. The B-AA66 PB2 is separated from the root of influenza A virus PB2s by a branch distance of 1,152 nucleotide changes (Fig. 3) which does not include deletions made for sequence alignment. The closest influenza A virus PB2s to B-AA66, EQPR56, and FPV34 are 1,322 and 1,450 steps away and show 43.5 and 38.1% similarity in branch distance, respectively, to the B-AA66 PB2 (Fig. 3).

Divergent evolution of four PB2 lineages. The phylogenetic tree shows that influenza A virus PB2 genes have evolved into four divergent lineages rooted at the deep forks of the tree (Fig. 2, left). The first lineage, EQPR56, represents the sister group (a group sharing an immediate ancestor with another lineage) to all other influenza A virus PB2s. The relative closeness to the B-AA66 root and the distance to other PB2s indicate that EQPR56 is the oldest of the PB2 lineages (branch distances are shown in Fig. 3). For the NP gene, EQPR56 shows a similar sister group relationship to other lineages (16). Of the two recent equine representatives, EQKY86 and EQLON73, the latter shares the same serotype (H7N7) as the EQPR56 strain. Bean (2) has suggested that recent equine viruses of the H7N7 subtype may be reassortants consisting of EQPR56 HA and NA genes and recent H3N8 internal genes. In addition to the PB2 gene, the NP, PB1, and PA genes of EQLON73 have been shown to be derived from recent equine H3N8 viruses (16, 23, 31). The failure to isolate other viruses containing EQPR56-like internal genes since 1956 suggests that lineages for these genes may be extinct.

At the next fork of the tree, the avian lineage containing FPV34 as its most primitive member forms a sister group to the remaining two lineages. Included in this avian lineage are representatives of Old World viruses (FPV34, BUDHOK77, and GULAST84) and North American viruses (RTNJ85 and GULMD77). The two H13 gull viruses (GULMD77 and GULAST84) form a short but distinct branch within this lineage.

Of the remaining two lineages, one containing avian and recent equine PB2s serves as the sister group to the swine-human lineage. The two avian PB2s (MLRDNY78 and TYMN80) of the avian-equine lineage are North American isolates. The seal PB2 (SEAL82) is also a North American isolate and appears closely related to the turkey PB2 (TYMN80). This corroborates earlier findings of Mandler et al. (26a) and others cited therein that influenza viruses of marine mammals are recently derived from avian viruses.

In the swine-human lineage, the classic swine PB2s (i.e., related to SWIA30) form the sister group to human PB2s. In the human branch, the CHILE83 isolate is positioned between the PR8-34 and the SING57 PB2s. This is expected, since the CHILE83 virus is descended from H1N1 viruses that reemerged in 1977 and appear to be derived from the

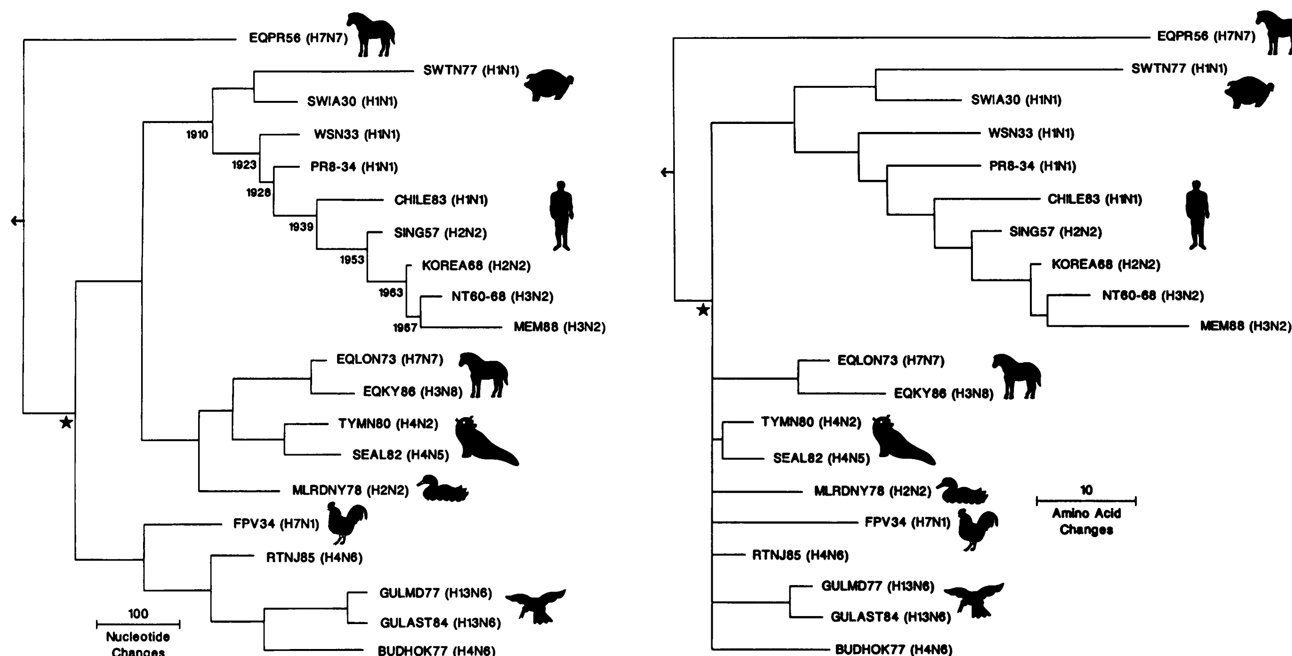


FIG. 2. Phylogenetic trees for 20 influenza A virus PB2 genes rooted to B/Ann Arbor/66 (B-AA66) PB2. Left, nucleotide tree. Sequences were analyzed with PAUP (David Swofford, Illinois Natural History Survey) which uses a maximum parsimony algorithm (see Materials and Methods). The arrow indicates the direction of the B-AA66 PB2 from the root node. With the exclusion of the B-AA66 outgroup, the number of variable characters represented is 1,570, the total tree length is 2,425 nucleotide changes, and the consistency index (proportion of changes due to forward mutations) is 0.487. Horizontal distance is proportional to the minimum number of nucleotide differences needed to join nodes and PB2 sequences. Vertical lines are for spacing branches and labels. Dates for hypothetical ancestor nodes were derived by dividing branch distance by evolutionary rate estimates (Fig. 4). The star denotes the avian ancestor node. Right, amino acid tree. The tree was generated by using the topology option of PAUP to make predicted amino acid sequences conform to the topology of the nucleotide tree (left). Note that the collapse of the avian amino acid lineages is caused by zero-length internal branches (see text). The number of variable characters represented is 172, the total tree length is 271 amino acid changes, and the consistency index is 0.801. Virus isolates represented in these trees are listed in Table 1.

1950 strains of human virus (1, 28, 36). The continuity of the human PB2 lineage shows that the PB2 gene was unaffected by reassortment of the human viral genome with avian genes in 1957 and 1968 (23).

The union of classic H1N1 swine and human PB2 lineages suggests that the human PB2s are derived from a swine ancestor or at least the two lineages share a common ancestor.

Patterns of divergence among PB2 genes. In contrast to absolute differences (Fig. 1), branch distances (Fig. 3) from the phylogenetic tree attempt to account for reverse mutations and thus may provide a more accurate estimation of evolutionary distance. A low consistency index (the proportion of nucleotide changes due to forward mutations in phylogenetic trees) of 0.487 and the much higher number of nucleotide changes predicted by the phylogenetic tree in comparison to the absolute differences (Fig. 1 versus Fig. 3) indicate that reverse mutations are a common feature in the evolution of influenza A virus PB2 genes. These reversals are related to evolutionary constraints to maintain ancestral amino acid sequences in some lineages.

It is readily apparent that the human lineage contains the most divergent PB2s (Fig. 2 and 3). The most divergent PB2s at the terminal branches of the two avian lineages and the swine-human lineage are closer to the avian root than to each other (Fig. 3). Similarly, the oldest PB2s of these lineages are closer to the avian root than to each other, with the exception of the SWIA30 and WSN33 PB2s, which share a recent common ancestor. A comparison of PB2s between the two avian lineages shows that they are closer to the avian root than to PB2s in the other lineage. Finally, the ancestral PB2 of the swine and human lineages (Fig. 2, left; represented by the node that joins the two branches and is dated 1910) is closer to the avian root than all avian PB2s. This pattern of evolution can be explained by divergent evolution of PB2 genes from a common avian virus ancestor.

Comparison of amino acid sequences. To evaluate the product of nucleotide evolution, we translated the open reading frames of the 20 influenza A virus PB2 nucleotide sequences. Absolute differences in deduced amino acid sequences were much less than expected, given the large number of nucleotide differences; this indicates that the PB2 protein is conserved, especially among avian virus PB2s (Fig. 1). Within the human lineage, PB2s show 4 to 8 nucleotide changes for each amino acid change among isolates, but among avian virus PB2s within their respective lineages, the range is 10 to 17 changes for each amino acid. The two recent equine virus PB2s show five changes for each amino acid. This pattern suggests that coding changes among the avian virus PB2 genes are relatively uncommon (see below). The pattern of differences is the same as with the nucleotide data; large differences occur between PB2s of the swine-human and avian lineages, and the oldest PB2s are more similar to avian PB2s.

Evolutionary tree of amino acid sequences. Since amino acid evolution is determined by nucleotide evolution, we used the topology option of PAUP to generate an amino acid sequence tree that uses the branching pattern of the nucle-

AMINO ACID BRANCH DISTANCES

	VIRUS	B	EP	SW		HUMAN							EQ		AVIAN						H13		SH	A
	STRAIN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	B-AA66	0	496	528	511	513	513	519	515	519	524	535	496	502	493	488	489	499	487	494	495	496	493	484
2	EQPR56	1322	0	98	81	83	83	89	85	89	94	105	66	72	63	58	59	69	57	64	65	66	63	54
3	SWTN77	1692	710	0	35	55	55	61	57	61	66	76	56	62	53	48	49	59	47	54	55	56	35	44
4	SWIA30	1557	575	239	0	38	38	44	40	44	49	59	39	45	36	31	32	42	30	37	38	39	18	27
5	WSN33	1552	570	340	205	0	26	32	28	32	37	47	41	47	38	33	34	44	32	39	40	41	20	29
6	PR8-34	1555	573	343	208	91	0	26	22	26	31	41	41	47	38	33	34	44	32	39	40	41	20	29
7	CHILE83	1652	670	440	305	188	151	0	18	22	27	37	47	53	44	39	40	50	38	45	46	47	26	35
8	SING57	1645	663	433	298	181	144	143	0	10	15	25	43	49	40	35	36	46	34	41	42	43	22	31
9	KOREA68	1686	704	474	339	222	185	184	63	0	7	17	47	53	44	39	40	50	38	45	46	47	26	35
10	NT60-68	1721	739	509	374	257	220	219	98	49	0	18	52	58	49	44	45	55	43	50	51	52	31	40
11	MEM88	1785	803	573	438	321	284	283	162	113	112	0	62	68	59	54	55	65	53	60	61	62	41	50
12	EQLON73	1586	604	530	395	390	393	490	497	538	573	637	0	12	21	16	17	27	15	22	23	24	21	12
13	EQKY86	1621	639	565	430	425	428	525	532	573	608	672	63	0	27	22	23	33	21	28	29	30	27	18
14	MLRDNY78	1541	559	485	350	345	348	445	452	493	528	592	245	280	0	13	14	24	12	19	20	21	18	9
15	TYMN80	1589	607	533	398	393	396	493	500	541	576	640	254	289	248	0	7	19	7	14	15	16	13	4
16	SEAL82	1602	620	546	411	406	409	506	513	554	589	553	267	302	261	309	0	20	8	15	16	17	14	5
17	FPV34	1450	468	562	427	422	425	522	529	570	605	669	456	491	411	459	472	0	18	25	26	27	24	15
18	RTNJ85	1488	506	600	465	460	463	560	567	608	643	707	494	529	449	497	510	212	0	13	14	15	12	3
19	BUDHOK77	1613	631	725	590	585	598	685	692	733	768	832	619	654	574	622	635	337	227	0	22	23	21	12
20	GULMD77	1621	639	733	598	593	596	693	700	741	776	840	627	662	582	630	643	345	235	240	0	5	19	10
21	GULAST84	1624	642	736	601	596	599	696	703	744	779	843	630	665	585	633	646	348	338	243	47	0	20	11
22	SW-Human	1452	470	240	105	100	103	200	193	234	269	333	290	325	245	293	306	322	360	485	493	496	0	9
23	Avian	1290	308	402	267	262	265	362	355	396	431	495	296	331	251	299	312	160	198	323	331	334	162	0

NUCLEOTIDE BRANCH DISTANCES

FIG. 3. Branch distance matrix for influenza virus PB2s in the phylogenetic trees (Fig. 2). Abbreviations: B, B-AA66; EP, EQPR56; SW, classic H1N1 swine; EQ, recent equine; SH (SW-Human), ancestor node of human and swine lineages (labeled 1910 on Fig. 2, left); and A (avian), ancestor node of avian lineages (labeled with a star in Fig. 2). Numbers 1 to 21 along the margins of the matrix correspond to virus isolates listed in Table 1.

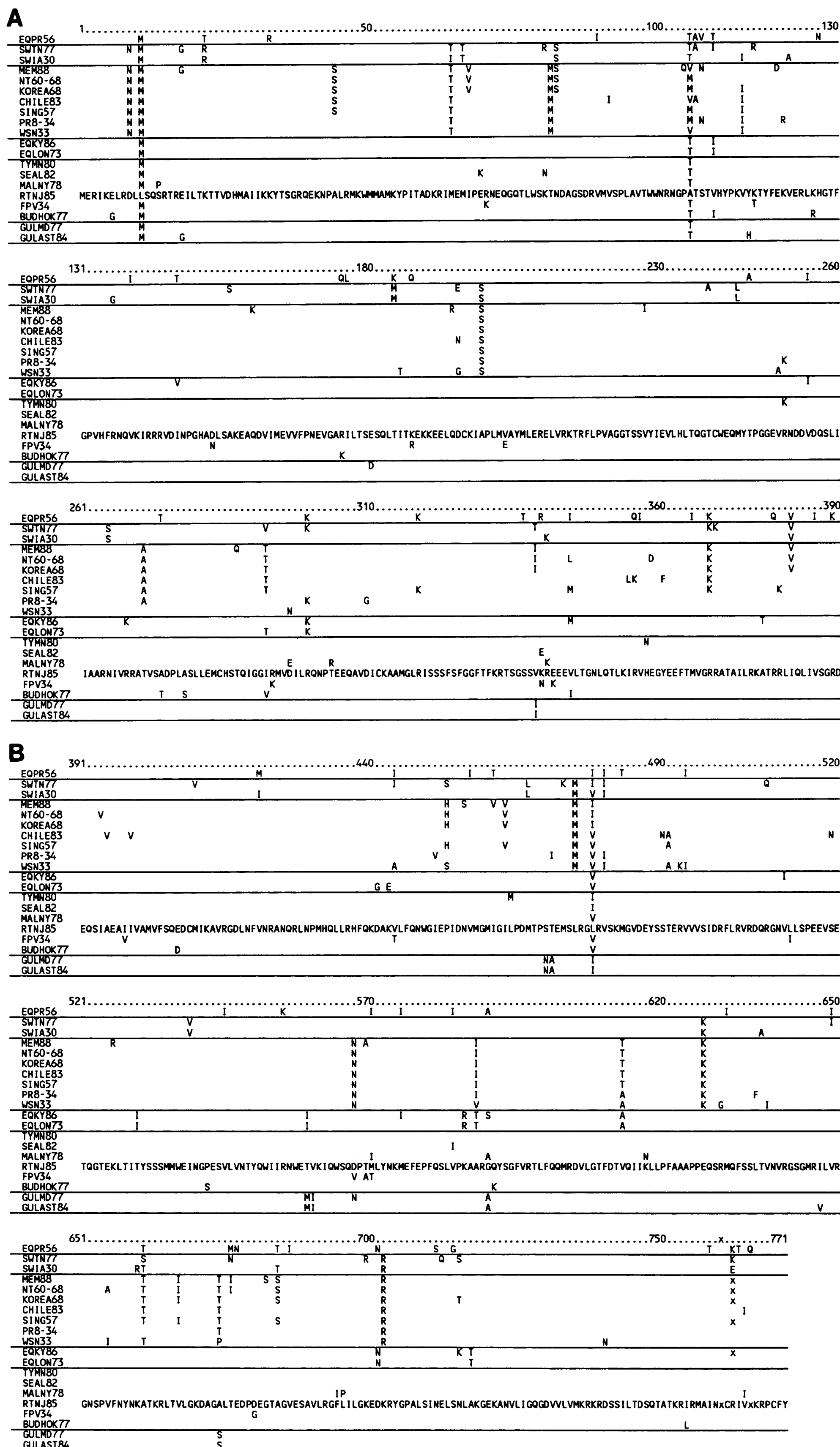
otide tree (Fig. 2, right). In essence, this amino acid tree represents one in which all silent nucleotide changes have been subtracted. This approach allows direct comparison of branches in the nucleotide and amino acid trees for the effect of genetic changes on the evolution of PB2 proteins. The amino acid sequence tree shows that at the protein level, the various PB2 lineages are divergent and evolving at different rates. An obvious difference between the two trees is the collapse of the internal branches above the root of the avian lineage; these internal branches have zero lengths in the amino acid tree, because the steps represented in the nucleotide tree are silent coding changes.

Patterns of divergence among PB2 proteins. The avian virus PB2 amino acid lineage is the most primitive, because it is closest to the root of the tree; several avian PB2s (RTNJ85, TYMN80, and SEAL82) are separated from the root by fewer than eight amino acid differences, and others are also relatively close. There is no correlation with year of isolation and distance from the avian root, suggesting that protein evolution in the avian lineage is virtually static. For example, the 1985 isolate RTNJ85 is the closest PB2 to the avian root (three steps; Fig. 2 and 3), while the PB2 of the 1934 isolate FPV34 is more distant (15 steps). The same evolutionary stasis has been found among avian virus NPs (16). The EQPR56, swine, and human PB2 lineages show significant evolution at the protein level. For example, the recent human PB2 MEM88 has accumulated 41 amino acid changes relative to the root of the swine-human lineage. Similarly, the SWTN77 PB2 has accumulated 35 amino acid changes.

The PB2s at the terminal branches of the EQPR56 and swine-human lineages are at least 32 steps closer to the avian root than to each other (Fig. 3). Similarly, the oldest nonavian PB2s, WSN33 and SWIA30, are at least nine steps closer to the avian root than to each other. The ancestral PB2 node of the swine-human lineage is only nine steps from the avian root; only three of eight avian PB2s are closer. This suggests that the ancestral PB2 protein of the swine-human lineage was not distinct from avian PB2s. As with the nucleotide tree, this pattern of divergent evolution away from avian-like PB2s suggests an avian origin for all PB2s.

The generation of an amino acid tree by using the topology of the nucleotide tree permits a more accurate estimate of the ratio of nucleotide changes to amino acid changes within lineages. For each PB2, these ratios are calculated by dividing the branch distance to the ancestor node of the nucleotide lineage by the analogous branch distance in the amino acid tree (Fig. 2 and 3). The ratios for PB2s of the swine-human lineage range from 5 to 9, which is similar to ratios derived from absolute differences (4 to 8; Fig. 1). However, avian PB2s have much higher ratios, ranging from 11 to more than 70. This disparity is the result of the reconstruction of nucleotide characters lost by reverse mutations in the phylogenetic analysis (PAUP) of nucleotide sequences. These reversals are not detectable by direct comparison of sequence data (Fig. 1).

Consensus amino acid sequences. A comparison of amino acid sequences among the major PB2 groups shows amino acids unique to each lineage (Fig. 4). To detect patterns of unique (derived) amino acids among the lineages, we used the most primitive PB2 sequence, RTNJ85 (the PB2 closest to the root of the tree), as a baseline. The RTNJ85 PB2, which is only three steps from the avian root (Fig. 2, right), shows three unique amino acids at positions 11, 105, and 478 (Fig. 4). Despite more than 370 nucleotide changes and branch distances exceeding 600 nucleotide changes among avian virus PB2 genes, only five amino acids show a pattern correlated with an avian subgroup; H13 gull PB2s share unique amino acids at sites 470, 471, 559, 560, and 674. There are 8 to more than 30 amino acid differences that are unique for each of the following remaining groups: EQPR56, classic swine, human, and recent equine. In addition, there are at least seven amino acids that are unique to both swine and human lineages; these amino acids are shared by the oldest isolates (SWIA30, WSN33, and PR8-34). There are two amino acids (sites 575 and 701) that appear in the EQPR56 PB2 and in at least one recent equine PB2. These amino acids may represent convergence due to adaptation to a common host. There are no amino acids that are correlated specifically with mammalian hosts, i.e., uniquely shared among swine, human, and equine hosts.


A
B
EQPR56
SWTN77
SWIA30
MEM88
NT60-68
KOREA68
CHILE83
SING57
PR8-34
WSN33
EQKY86
EQLON73
TYMN80
SEAL82
MALNY78
RTNJ85
FPV34
BUDHOK77
GULMD77
GULAST84

Estimation of evolutionary rates. Mammalian and avian lineages include PB2s from virus isolates spanning more than 50 years. The dating of isolates over this long period provided an opportunity to estimate evolutionary rates for lineages of PB2 genes. Evolutionary rate was estimated by plotting year of isolation against branch distance to the ancestor node of the lineage; the slope of the regression line for the plotted points equals the number of nucleotide changes per year (Fig. 5). A regression was not possible with avian virus PB2s because of a lack of correlation between dates of virus isolation and lineage position. This situation can arise if a lineage is composed of a subset of descendants sharing a common ancestor. Except for the placement of the CHILE83 PB2, data for the human lineage appears to be in order. For the purpose of estimation of evolutionary rates, we treated the CHILE83 isolate as having appeared in 1956, because the virus appeared 6 years after the reappearance of the H1N1 viruses in 1977, and the first of these viruses were found to be very similar to 1950 human virus strains (1, 28, 36).

Regression of data for the seven human virus PB2s yielded a slope of 4.27 nucleotide changes per year (1.82×10^{-3} changes per nucleotide per year), with a correlation coefficient of 0.99 (Fig. 5, top). Our values are within the range of estimates reported for human NP genes (1.62×10^{-3} to 2.2×10^{-3} substitutions per site per year [1, 16]) and human NS genes (2×10^{-3} substitutions per site per year [7]). For human PB2 amino acid sequences, the evolutionary rate is estimated to be 0.35 base changes per year or 4.61×10^{-4} changes per base per year (Fig. 5, bottom). This value is less than half the per-base rate for the human NP protein (1.14×10^{-3} changes per base per year [16]). This suggests that there may be greater constraints on the evolution of the human PB2 protein relative to the NP protein. Although not included in the regression, data points for the two swine PB2s, SWIA30 and SWTN77, lie just outside the scatter of human data points, suggesting that the swine PB2 evolutionary rate is roughly similar to that for human PB2s. Even though the avian PB2 data could not be subjected to regression, a comparison of nucleotide and amino acid trees indicates that avian PB2 proteins have not evolved appreciably over the past 50 years. For the NP gene, Gorman et al. (16) did not detect evolution of avian NP amino acid sequences over the past 50 years. The lack of evolution of some internal proteins of avian influenza viruses may be attributed to stabilizing selection; avian viral proteins may be well adapted to their hosts, and genes with mutations that result in changes in the adapted amino acid sequence are apparently eliminated.

Estimation of dates for ancestral PB2s. Estimates of evolutionary rates were used to calculate dates for hypothetical ancestor nodes of the human lineage (Fig. 2, left). This was accomplished by dividing the branch-internodal distance by the evolutionary rate. The year estimate for an ancestor node represents the mean of estimates calculated from all descendants. The same year estimate may also be derived directly from the regression of nucleotide data (Fig. 5); the year of isolation corresponds to branch distance from the ancestor of the human lineage. The 1910 estimate for the human PB2 lineage is close to the 1914 estimate for the ancestor of the human NP lineage (16). A conservative extrapolation of evolutionary rates for human influenza viruses places estimates for the oldest ancestor nodes of the tree in the range of 1870 to 1890. This estimate for the age of the avian ancestor nodes is likely to be too recent, since our analysis suggests that PB2s of human influenza viruses are evolving more rapidly than those of avian viruses. Also, because of long-term evolutionary stasis of avian PB2 proteins, the deepest (oldest) internal branches of the avian nucleotide lineages are likely to be longer than those shown in Fig. 2 (left).

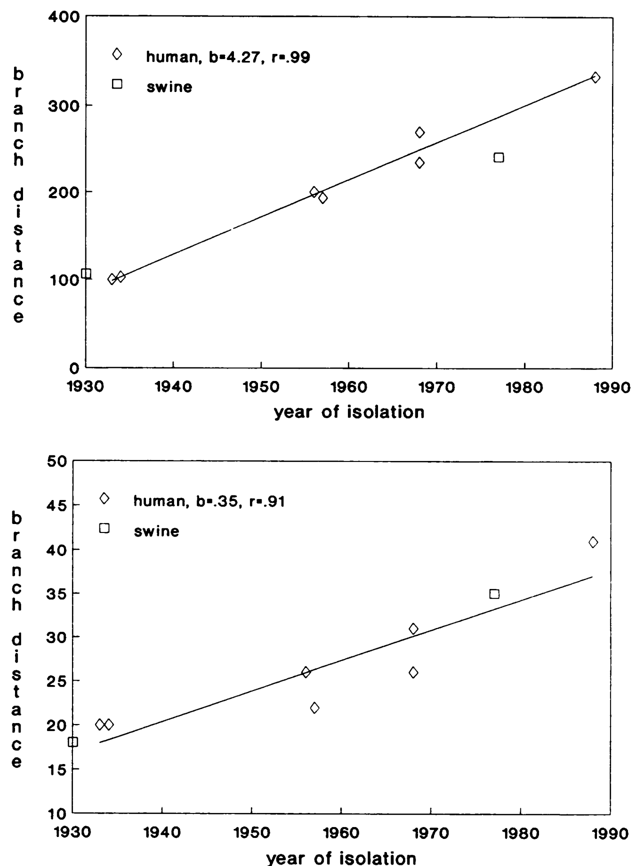


FIG. 5. Evolutionary rates for human influenza A virus PB2 genes and proteins. Upper panel, PB2 nucleotide evolutionary rate. Lower panel, PB2 amino acid evolutionary rate. Evolutionary rate is estimated by regression of year of isolation against branch distance from the swine-human common ancestor node of the nucleotide and amino acid phylogenetic trees (Fig. 2). Regression statistic b provides rate estimates. The common ancestor node corresponds to the branch point dated 1910 in the nucleotide tree (Fig. 2, left). Two swine data points are shown for reference and are not included in the regression of human data.

DISCUSSION

Evolutionary pathways of RNP proteins. The combination of our evolutionary analysis of influenza A virus PB2 genes and previous analyses of the PA, PB1, and NP genes

FIG. 4. Predicted amino acid sequences of the 20 influenza A virus PB2 genes. The amino acid sequence of RTNJ85 is written in full (A and B); for other sequences, only differences from this baseline are shown. PB2 groups recognized in Fig. 1 are separated by lines in the following order (top to bottom): EQPR56, classic H1N1 swine, human, recent equine, avian, and H13 gull. Virus isolates represented are listed in Table 1.

provides an opportunity to compare patterns of evolutionary pathways among proteins of the RNP complex. In an analysis of 41 sequences, Gorman et al. (16) showed that influenza A virus NP genes have evolved into the following five divergent, host-specific lineages: (i) equine/Prague/56, (ii) recent equine, (iii) swine and human, (iv) H13 gull, and (v) avian. Although only 16 isolates were investigated, the pattern of evolution for influenza A virus PA genes as presented by Okazaki et al. (31) is congruent with that for the NP gene. In contrast, an analysis of 16 influenza A virus PB1 genes by Kawaoka et al. (23) shows little congruence with the NP-PA pattern of evolution. Furthermore, PB1 gene evolution shows much less host specificity; avian virus PB1 genes have been introduced into the human virus gene pool most recently during the antigenic shifts of 1957 and 1968 (23). However, because the objective of the PB1 study was to resolve the origin of PB1 genes in human influenza A viruses, only three avian virus genes are represented in the analysis.

The evolutionary patterns we found for influenza A virus PB2 genes differ from those found for the NP, PA, and PB1 genes. Only the following four major PB2 lineages are evident: EQPR56; swine-human; and two distinct avian lineages, one including recent equine PB2s and the other including H13 gull PB2s. The EQPR56 and swine-human lineages are comparable with those of the NP and PA genes, but homologs of the distinct H13 gull and recent equine lineages of the NP and PA genes are absent. Host specificity is demonstrated for the human-swine lineage, but the short branches for recent equine and H13 gull PB2s within avian PB2 lineages suggest a relatively recent avian origin. Unlike the NP gene, evolutionary patterns related to geographic separation of viruses and hosts were not evident among PB2 genes.

The PB1 and PB2 genes share the feature of fewer host-specific lineages, but with the exception of the H1N1 classic swine and H1N1 human lineages, none of the other lineages in the two genes share parallel evolutionary pathways. Also, the PB1 genes of influenza A and B viruses are more similar than the respective PB2 genes (59.5 to 60.7% [23] versus 45.4 to 46.5% similarity for aligned sequences). The greater similarity of influenza A and B virus PB1s suggests that this protein is evolving more slowly than PB2. In contrast, we found an aligned B/Lee/40 NP gene to show only 33.6 to 35.5% similarity with influenza A virus NP genes (16). The lower sequence similarity of influenza A and B virus NP genes relative to the polymerase genes and differences in evolutionary pathways among PA, PB1, PB2, and NP genes suggest that RNP proteins are evolving in response to different factors.

The RNP complex as an evolutionary unit. The different evolutionary pathways of RNP proteins suggest that the complex does not evolve as a unit. Parallel host-specific evolution of the NP and PA genes indicates that they do not reassort with genes of other lineages and that they are coevolving in response to host-specific factors. Also, parallel evolution raises the possibility of a functional linkage between NP and PA. In contrast, evolution of the PB1 and PB2 genes appears more independent of host factors, i.e., there are fewer host-specific lineages and overall their lineages are less divergent. Greater similarity of PB1 and PB2 with influenza B virus homologs relative to NP and much slower protein evolution of PB2 in human influenza A viruses relative to NP suggest that virus-specific constraints on protein function may be limiting the evolutionary divergence of these polymerase proteins. PB1 has been implicated in both initiation and elongation of viral mRNA (6), while PB2 plays a role in cap recognition of mRNAs (5, 6, 30, 39). The conserved evolution of PB1 and PB2 explains their relative lack of host specificity and predicts that there may not be significant host adaptation barriers to prevent these genes from forming reassortants and becoming integrated into virus gene pools of alternate hosts. Unlike the NP gene (and possibly the PA gene), the PB1 and PB2 genes are not likely to be involved in the maintenance of host-specific virus gene pools.

Avian origin for RNP genes. Our phylogenetic analyses make no assumptions about relationships among influenza A virus PB2s. Independent evidence supporting an avian (waterfowl) origin for all influenza viruses includes the proliferation and maintenance of numerous hemagglutinin and neuraminidase subtypes among migratory waterfowl populations, asymptomatic infection indicative of coadaptation of virus and host, and a propensity for dissemination of viruses to other hosts because of migratory habits and shedding of prodigious quantities of virus (20). In support of this hypothesis, most of the new influenza virus genes that have appeared in mammalian gene pools (especially swine and human) over the past 30 years have been shown to have an avian origin (Hinshaw and Webster [20], Kida et al. [24], Chambers et al. [8], Kawaoka et al. [23], Gorman et al. [16], and Mandler et al. [26a]). In the analysis presented here, we have identified the PB2s of a seal (SEAL82) and two equine isolates (EQLON73 and EQKY86) as having a recent avian origin. The results of our analysis of PB2 gene evolution are consistent with the hypothesis of an avian origin for all influenza A viruses. Our summary evidence includes the following: (i) older nonavian PB2 genes are more avian-like than recent PB2s; (ii) recent PB2 genes in the four lineages are more distant from recent PB2s in other lineages than from the common ancestor of the avian lineages; and (iii) avian PB2 proteins appear to be in evolutionary stasis, while nonavian lineages have diverged significantly from ancestors indistinguishable from current avian PB2s. Although the evolutionary pathways are quite different from those of the PB2 gene, virtually the same lines of evidence supporting an avian origin for influenza A viruses have been found for the NP gene (16).

Origin of classic swine and human viruses. The close phylogenetic relationship of the PB2s from the oldest swine (SWIA30) and human (WSN33 and PR8-34) isolates suggests a common origin for the PB2s of human viruses and classic swine viruses (i.e., swine viruses related to SWIA30). The same close phylogenetic relationship occurs for the NP (16), PA (31), and H1N1 lineage PB1 genes (23). For the other genes of human and classic swine viruses, a close phylogenetic relationship is suggested by the similarity of nucleotide sequences (M and NS genes [29]; Y. Kawaoka, unpublished data) and by the antigenic similarity of surface proteins (HA and NA genes [11, 12, 17, 18, 27]). Estimates provided in this paper for the PB2 gene, by Gorman et al. (16) for the NP gene, and by Nakajima et al. (29) for the M and NS genes predict a date between 1910 and 1914 for the common ancestor of human and classic swine viruses. Phylogenetic analyses of PB1, PB2, PA, and NP genes indicate that the genes of this ancestor were avian-like. In cases in which genes of avian origin have appeared recently in virus gene pools of mammalian hosts (e.g., avian-like PB1 genes in human virus in 1957 and 1968 [23]), the early isolates of these viruses contain genes virtually indistinguishable from contemporary avian genes, but later isolates in the new lineages show a rapid accumulation of unique mutations. Retrospec-

tive serological studies by Mulder and Masurel (27) indicate that the appearance of an H1N1 virus in the human population was correlated with the 1918 pandemic; previous strains in 1890 and 1900 appeared to have H2N8 and H3N8 serotypes, respectively. Taken together, these results suggest the following hypothesis: a new H1N1 virus of recent avian origin (not a reassortant) appeared in human and swine populations prior to the 1918 influenza pandemic, previously existing human virus strains became extinct, and the new virus evolved rapidly and independently in human and swine host reservoirs.

Phylogenetic analysis cannot determine in which host this hypothetical virus first appeared. It is possible that this avian-derived virus was first adapted in swine hosts and then was secondarily introduced into the human population by contact with infected pigs. Humans are susceptible to classic swine viruses (10, 19, 34), and there are now numerous examples of recent introductions into and persistence of H1N1 and H3N2 avian viruses in swine (15, 16, 24, 35). Alternatively, evidence of direct introduction of avian viral genes into the human viral gene pool is weak; persons handling chickens and wild ducks that are shedding virus show no detectable antibodies to avian viruses (3; V. Hinshaw and R. G. Webster, unpublished data), although replication of avian viruses in humans has been demonstrated (A. S. Beare, unpublished data). The appearance and maintenance of avian influenza viruses in the swine host reservoir is probably not a new phenomenon; it is likely that similar conditions preceded the appearance of the ancestor of the highly virulent 1918 virus. Understanding the role of the swine host reservoir in the ongoing evolution of human influenza A viruses may provide the key to understanding the origin of human and classic swine viruses. Also, this knowledge may provide the capability to intervene and prevent periodic pandemics of human influenza viruses.

ACKNOWLEDGMENTS

We thank Raphael Onwuzuruigbo and Evelyn Stigger for technical assistance, Clayton Naeve and the St. Jude Children's Research Hospital Molecular Resource Center for preparation of oligonucleotides, Patricia Eddy and the St. Jude Children's Research Hospital Molecular Biology Computer Facility for computer support, William J. Bean for assistance with sequence analysis, and Thomas Chambers for reviewing earlier versions of the manuscript.

This work was supported by Public Health Service research grant AI-08831 from the National Institutes of Health, National Institute of Allergy and Infectious Diseases grants AI-29599 and AI-29680, Cancer Center Support (CORE) grant CA-21765, National Research Service Award 5T32-CA09346 to O.T.G., and American Lebanese Syrian Associated Charities.

LITERATURE CITED

1. **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.
2. **Bean, W. J.** 1984. Correlation of influenza A virus nucleoprotein genes with host species. Virology **133:**438–442.
3. **Bean, W. J., Y. Kawaoka, J. M. Wood, J. E. Pearson, and R. G. Webster.** 1985. Characterization of virulent and avirulent A/chicken/Pennsylvania/83 influenza A viruses: potential role of defective interfering RNAs in nature. J. Virol. **54:**151–160.
4. **Bean, W. J., G. Sriram, and R. G. Webster.** 1980. Electrophoretic analysis of iodine-labeled influenza virus RNA segments. Anal. Biochem. **102:**228–232.
5. **Blass, D., E. Patzelt, and E. Kuechler.** 1982. Cap-recognizing protein of influenza virus. Virology **116:**339–348.
6. **Braam, J., I. Ulmanen, and R. M. Krug.** 1983. Molecular model of a eucaryotic transcription complex: functions and movements of the influenza P proteins during capped RNA-primed transcription. Cell **34:**609–618.
7. **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 and uniform rate of change in the NS gene. Science **232:**980–982.
8. **Chambers, T. M., S. Yamnikova, Y. Kawaoka, D. K. Lvov, and R. G. Webster.** 1989. Antigenic and molecular characterization of subtype H13 hemagglutinin of influenza virus. Virology **172:**180–188.
9. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4:**165–170.
10. **Dasco, C. C., R. B. Couch, H. R. Six, J. F. Young, J. M. Quarles, and J. A. Kasel.** 1984. Sporadic occurrence of zoonotic swine influenza virus infections. J. Clin. Microbiol. **20:**833–835.
11. **Davenport, F. M., A. V. Hennessy, J. Drescher, J. Mulder, and T. Francis.** 1964. Further observations on the relevance of serologic recapitulation of human infection with influenza viruses. J. Exp. Med. **120:**1087–1097.
12. **Davenport, F. M., A. V. Hennessy, and T. Francis.** 1953. Epidemiologic and immunologic significance of age distribution of antibody to antigenic variants of influenza virus. J. Exp. Med. **98:**641–656.
13. **DeBorde, D. C., A. M. Donabedian, M. L. Herlocher, C. W. Naeve, and H. F. Maassab.** 1988. Sequence comparison of wild-type and cold-adapted B/Ann Arbor/1/66 influenza virus genes. Virology **163:**429–443.
14. **Fields, S., and G. Winter.** 1982. Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. Cell **28:**303–313.
15. **Gammelin, M., J. Mandler, and C. Scholtissek.** 1989. Two subtypes of nucleoproteins (NP) of influenza A viruses. Virology **170:**71–80.
16. **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.
17. **Hennessy, A. V., and F. M. Davenport.** 1958. Epidemiologic implications of the distribution by age of antibody response to experimental influenza virus vaccines. J. Immunol. **80:**114–121.
18. **Hennessy, A. V., F. M. Davenport, and T. Francis.** 1955. Studies of antibodies to strains of influenza virus in persons of different ages in sera collected in a postepidemic period. J. Immunol. **75:**401–405.
19. **Hinshaw, V. S., W. J. Bean, R. G. Webster, and B. C. Easterday.** 1978. The prevalence of influenza viruses in swine and the antigenic and genetic relatedness of influenza viruses from man and swine. Virology **84:**51–62.
20. **Hinshaw, V. S., and R. G. Webster.** 1982. The natural history of influenza A viruses, p. 79–104. *In* A. S. Beare (ed.), Basic and applied influenza research. CRC Press, Inc., Boca Raton, Fla.
21. **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 segments of strains A/PR/8/34 and A/WSN/33. Nucleic Acids Res. **11:**1555–1566.
22. **Kaptein, J. S., and D. P. Nayak.** 1982. Complete nucleotide sequence of the polymerase 3 gene of human influenza virus A/WSN/33. J. Virol. **42:**55–63.
23. **Kawaoka, Y., S. Krauss, and R. G. Webster.** 1989. Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. J. Virol. **63:**4603–4608.
24. **Kida, H., K. F. Shortridge, and R. G. Webster.** 1988. Origin of the hemagglutinin gene of H3N2 influenza viruses from pigs in China. Virology **162:**160–166.
25. **Krystal, M., R. M. Elliot, E. W. Benz, Jr., J. F. Young, and P. Palese.** 1982. Evolution of influenza A and B viruses: conservation of structural features in the hemagglutinin genes. Proc. Natl. Acad. Sci. USA **79:**4800–4804.
26. **Lamb, R. A.** 1989. The genes and proteins of the influenza viruses, p. 1–87. *In* R. M. Krug (ed.), The influenza viruses. Plenum Publishing Corp., New York.

26a. **Mandler, J., O. T. Gorman, S. Ludwig, E. Schroeder, W. M. Fitch, R. G. Webster, and C. Scholtissek.** 1990. Derivation of the

nucleoproteins (NP) of influenza A viruses isolated from marine mammals. Virology **176**:255–261.

27. **Mulder, J., and N. Masurel.** 1958. Pre-epidemic antibody against 1957 strain of Asiatic influenza in serum of older people living in The Netherlands. Lancet **i**:810–814.
28. **Nakajima, K., U. Desselberger, and P. Palese.** 1978. Recent human influenza viruses are closely related genetically to strains isolated in 1950. Nature (London) **274**:334–339.
29. **Nakajima, K., E. Nobusawa, and S. Nakajima.** 1984. Genetic relatedness between A/swine/Iowa/15/30 (H1N1) and human influenza viruses. Virology **139**:194–198.
30. **Nichol, S. T., C. R. Penn, and B. W. J. Mahy.** 1981. Evidence for the involvement of influenza A (Fowl plague Rostock) virus protein P2 in ApG and mRNA primed in vitro RNA synthesis. J. Gen. Virol. **57**:407–413.
31. **Okazaki, K., Y. Kawaoka, and R. G. Webster.** 1989. Evolutionary pathways of the PA genes of influenza A viruses. Virology **172**:601–608.
32. **Roditi, I. J., and J. S. Robertson.** 1984. Nucleotide sequence of the avian influenza virus A/Fowl Plague/Rostock/34 segment 1 encoding the PB2 polypeptide. Virus Res. **1**:65–71.
33. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
34. **Schnurrenberger, P. R., G. T. Woods, and R. J. Martin.** 1970. Serologic evidence of human infection with swine influenza virus. Am. Rev. Respir. Dis. **102**:356–361.
35. **Scholtissek, C., H. Burger, P. A. Bachmann, and C. Hannoun.** 1983. Genetic relatedness of hemagglutinins of the H1 subtype of influenza A viruses isolated from swine and birds. Virology **129**:521–523.
36. **Scholtissek, C., V. Von Hoyningen, and R. Rott.** 1978. Genetic relatedness between the new 1977 epidemic strain (H1N1) of influenza and human influenza strains isolated between 1947 and 1957 (H1N1). Virology **89**:613–617.
37. **Schreier, E., D. R. Petzold, S. Michel, and S. Dittmann.** 1988. Evolution of influenza polymerase: nucleotide sequence of the PB2 gene of A/Chile/1/83 (H1N1). Arch. Virol. **103**:179–187.
38. **Treanor, J., Y. Kawaoka, R. Miller, R. G. Webster, and B. Murphy.** 1989. Nucleotide sequence of the avian influenza A/Mallard/NY/6750/78 virus polymerase genes. Virus Res. **14**:257–270.
39. **Ulmanen, I., B. A. Broni, and R. M. Krug.** 1981. Role of two of the influenza virus core P proteins in recognizing cap 1 structures (m^7GpppNm) on RNAs and in initiating viral RNA transcription. Proc. Natl. Acad. Sci. USA **78**:7355–7359.
40. **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.