Proc. Natl. Acad. Sci. USA
Vol. 76, No. 12, pp. 6547–6551, December 1979
Genetics

Evolution of human influenza A viruses in nature: Recombination contributes to genetic variation of H1N1 strains

(mutations/viral RNA/oligonucleotide mapping)

JAMES F. YOUNG AND PETER PALESE

Department of Microbiology, Mount Sinai School of Medicine of the City University of New York, Fifth Avenue and 100th Street, New York, New York 10029

Communicated by Edwin D. Kilbourne, September 24, 1979

ABSTRACT **In June of 1977, a new influenza A pandemic was started by strains of the H1N1 serotype. Oligonucleotide fingerprint analysis of the RNA from viruses isolated during the early stage of this pandemic demonstrated that genetic variation among these 1977 strains could be attributed to sequential mutation [Young, J. F., Desselberger, U. & Palese, P. (1979) *Cell*, 18, 73–83]. Examination of more recent strains revealed that the H1N1 variants that were isolated in the winter of 1978–1979 differed considerably from the H1N1 viruses isolated the previous year. Oligonucleotide and peptide map analysis of the new prototype strain (A/Cal/10/78) suggested that it arose by recombination. It appears that only the *HA, NA, M,* and *NS* genes of this virus are derived from the earlier H1N1 viruses and that the *P1, P2, P3,* and *NP* genes most likely originate from an H3N2 parent. These data suggest that genetic variation in influenza virus strains of the same serotype is not restricted to mutation alone, but can also involve recombination (reassortment).**

Influenza viruses display antigenic variation while circulating in the population. This property has most likely been an important factor contributing to the difficulties of controlling influenza by vaccination. Initial studies of this variation used serological techniques to analyze changes in the surface proteins (for review, see ref. 1). The development of sensitive biochemical techniques in recent years has permitted analysis of the entire genome and has provided tools for determining the underlying molecular mechanisms responsible for these changes. For example, peptide mapping and hybridization studies of the H3N2 pandemic viruses indicate that they may have arisen from earlier H2N2 strains by a recombinational event (refs. 2 and 3; for review, see ref. 4). Evidence has also been presented suggesting that the H2N2 pandemic strain of 1956 arose by recombination (3). However, recombination did not appear to be responsible for the emergence of the most recent pandemic strain in 1977. All eight genes of these viruses were remarkably similar to those of strains isolated in 1950 (5, 6). These data suggested that mechanisms other than recombination can result in the emergence of new pandemic viruses.

More recently, we have focused our efforts on examining the changes occurring in the viral genome of currently circulating H1N1 viruses. Using RNA fingerprinting and sequencing techniques, we found that mutations in these new variants were not restricted to the genes coding for the surface proteins but were scattered throughout the genome. In addition, our analysis of the mutations allowed us to propose a scheme of divergent evolution of these viruses.

In the present report we describe further examination of the evolution of influenza A virus occurring in H1N1 strains circulating in the winter of 1978–1979. Evidence is presented that suggests that these variants arose by recombination of an earlier H1N1 strain with an H3N2 virus. This finding suggests that recombination as well as mutation may be involved in genetic variation of influenza viruses during an interpandemic period.

MATERIALS AND METHODS

Viruses and Cells. Influenza strains A/Berk/40/78 (H1N1)* (isolated in Berkeley, CA in Nov. 1978), A/Cal/10/78 (H1N1), A/Cal/11/78 (H1N1), and A/Cal/20/78 (H1N1) (all isolated in Los Angeles, CA in Dec. 1978), A/Ga/64/79 (H1N1) and A/Vt/14/79 (both isolated in Feb. 1979, in Georgia and Vermont, respectively), A/Wis/7/79 (H1N1) and A/Wis/8/79 (H1N1) (both isolated in March 1979, in Wisconsin), A/USSR/90/77 (H1N1) (precise date and location of isolation unknown), and A/Tex/1/77 (H3N2) (isolated in 1977 in Texas) were kindly provided by A. P. Kendal (Center for Disease Control, Atlanta, GA). Because the gene coding for the H1 hemagglutinin of A/USSR/90/77 and A/Cal/10/78 viruses migrates as a broad band under our conditions used for RNA separation on polyacrylamide gels, recombinants of these viruses were prepared in which this hemagglutinin was replaced with H3 hemagglutinin from recombinant D. This latter recombinant virus and conditions for isolation of these recombinant strains have been described (7, 8). Seed stocks of the above viruses were prepared by inoculation of 11-day-old embryonated chicken eggs as described (9).

Madin–Darby canine kidney cells (MDCK) (10) were propagated in minimal essential medium containing Earle's salts (GIBCO) supplemented with 10% fetal bovine serum.

Growth and purification of viruses and the extraction of viral RNAs have been described elsewhere (11, 12).

Labeling and Analysis of Viral RNA Segments. Viral RNAs were labeled with ^{32}P (ICN Pharmaceuticals) in MDCK cells and extracted from purified virus as described (11, 12). The ^{32}P-labeled viral RNAs were separated on 2.8% polyacrylamide gels containing 7 M urea (11–13). Oligonucleotide fingerprint analysis of total RNA and isolated RNA segments has been described (13, 14).

Analysis of Viral Polypeptides. Virus-infected MDCK cells were pulse-labeled with L-[^{35}S]methionine (Amersham) 7 hr after infection as described (15). Labeled viral polypeptides were subjected to electrophoresis on sodium dodecyl sulfate ($NaDodSO_4$)/polyacrylamide slab gels with the discontinuous buffer system of Maizel (16). Running conditions and preparation of samples were as reported (15), except that the separating gel contained a 5–20% exponential gradient of polyacrylamide with a concentration of 0.13% *N,N'*-methylene bisacrylamide throughout.

Peptide maps of individual viral polypeptides from pulse-labeled infected cells were examined by the procedure of

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

Abbreviation: $NaDodSO_4$, sodium dodecyl sulfate.

* For nomenclature of influenza viruses see *Bull. WHO* (1971) **45**, 119.

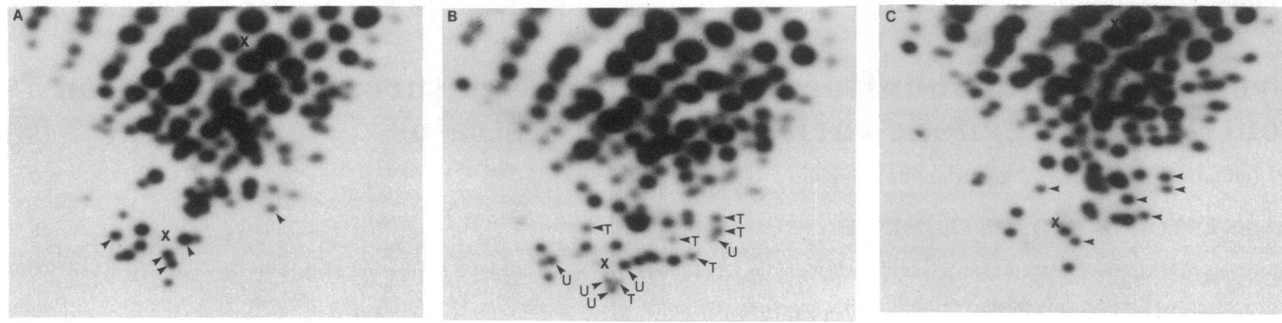

FIG. 1. Two-dimensional oligonucleotide fingerprints of the RNAs of influenza A/USSR/90/77 (*A*), A/Cal/10/78 (*B*), and A/Tex/1/77 (*C*) viruses. The positions of the dye markers, xylene cyanol FF and bromophenol blue, are indicated by Xs. Several A/USSR/90/77 virus-specific large oligonucleotides appear to be present in the map of the A/Cal/10/78 virus (compare *A* and *B*) (e.g., five large oligonucleotides shared by the two viruses are indicated by arrows in the map of A/USSR/90/77 virus RNA and by arrows with a U in the map of A/Cal/10/78 virus RNA). On the other hand, several A/Tex/1/77 virus-specific oligonucleotides (not found in the map of the A/USSR/90/77 virus) are also shared by the RNA of A/Cal/10/78 virus (compare *B* and *C*) (e.g., six large oligonucleotides shared by the A/Cal/10/78 and A/Tex/1/77 viruses are indicated by arrows in the map of A/Tex/1/77 virus RNA and by arrows with a T in the map of A/Cal/10/78 virus RNA). Only a few of the shared oligonucleotides are indicated in this figure. Shared spots were identified by comigration of the oligonucleotides of two viral RNAs in the same system.

Cleveland *et al.* (17) as modified for influenza virus proteins (18) except that the protein gels and peptide maps were autoradiographed directly with Kodak X-Omat XR2 x-ray films.

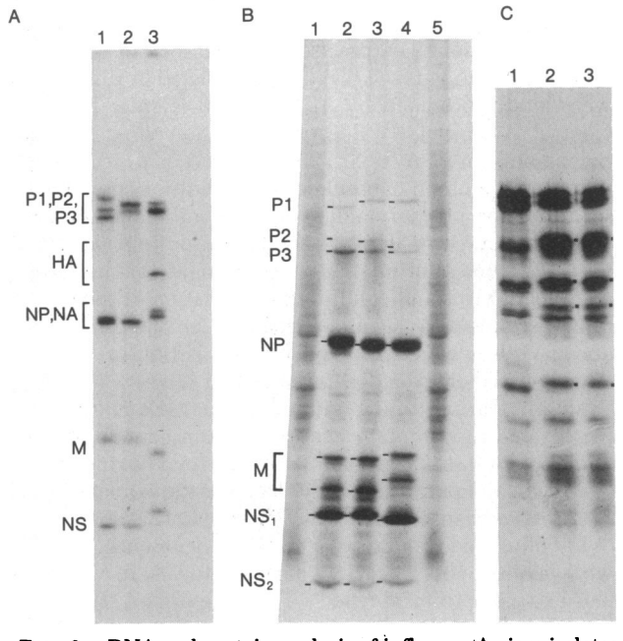

FIG. 2. RNA and protein analysis of influenza A virus isolates. (*A*) Polyacrylamide gel analysis of ^{32}P-labeled RNA of A/USSR/90/77 virus (lane 1), A/Cal/10/78 virus (lane 2), and A/Tex/1/77 virus (lane 3). The polypeptide product specified by each gene is noted to the left on the autoradiogram. (*B*) Comparison of infected-cell proteins by NaDodSO$_4$/polyacrylamide gel electrophoresis. Lanes 1 and 5, uninfected cells; lane 2, A/USSR/90/77 virus-infected cells; lane 3, A/Cal/10/78 virus-infected cells; lane 4, A/Tex/1/77 virus-infected cells. The positions of virus-specific proteins are noted by small dashes to the left of each lane; proteins are identified to the left of the autoradiogram. (*C*) Determination of the NP protein derivation in A/Cal/10/78 virus by using one-dimensional peptide maps. After separation of virus-infected cell proteins on 5–13% gradient polyacrylamide gels (*B*), the NP proteins were excised from the gel and subjected to partial digestion in the stack of a second polyacrylamide gel in the presence of 0.5 μg of *Staphylococcus aureus* V8 protease per slot. Oligopeptides were separated in a 10–15% gradient NaDodSO$_4$/polyacrylamide gel. Lane 1, digestion pattern of A/USSR/90/77 virus NP protein; lane 2, digestion pattern of A/Cal/10/78 virus NP protein; lane 3, digestion pattern of A/Tex/1/77 virus NP protein. A/Tex/1/77 virus-specific oligopeptides shared by the NP polypeptide of A/Cal/10/78 virus are indicated by dots.

RESULTS

The RNAs from several H1N1 viruses (A/Cal/10/78, A/Cal/11/78, A/Cal/20/78, A/Berk/40/78, A/Ga/64/79, A/Wis/7/79, A/Wis/8/79, and A/Vt/14/79), isolated in late 1978 and early 1979, all had similar oligonucleotide fingerprints, which were different from those of H1N1 strains isolated in 1977. Fig. 1 *A* and *B* shows the maps of the RNAs from a 1977 strain (A/USSR/90/77) and from a later strain (A/Cal/10/78). Although several large oligonucleotides were common to the RNAs of these two viruses, the number of differences seen in the A/Cal/10/78 virus map when compared to that of the A/USSR/90/77 virus map appeared to be greater than expected for strains of the same subtype. Comparison of the A/Cal/10/78 map with the fingerprints of RNAs from other influenza virus subtypes revealed that several large oligonucleotides in the A/Cal/10/78 map were also found in the map of A/Tex/1/77 virus (Figs. 1 *B* and *C*). This observation prompted us to examine the A/Cal/10/78 H1N1 isolate in more detail to determine if its dissimilarity from typical H1N1 viruses could be due to the exchange of gene(s) that are derived from an H3N2 strain.

RNA Patterns. Analysis of the RNAs of A/USSR/90/77, A/Cal/10/78, and A/Tex/1/77 on polyacrylamide gels is shown in Fig. 2*A*. The RNAs coding for the HA, M, and NS proteins from the A/Cal/10/79 virus appear to be migrating identically to corresponding genes from the H1N1 strain A/USSR/90/77. The diffuse *HA* band is consistent with that seen for other recent H1N1 strains and quite different from that seen for the *HA* gene of H3N2 viruses (lane 3). The migration characteristic of the *H1* RNA segment is most likely due to secondary structure and not size heterogeneity; electrophoresis after glyoxalation of the RNA reveals a single band (data not shown). The remaining genes of the A/Cal/10/78 virus did not appear to comigrate with those of the A/USSR/90/77 virus. The overall pattern of the three *P* genes of A/Cal/10/79 virus more greatly resembled that observed for the A/Tex/1/77 virus than that of the A/USSR/90/77 strain.

Oligonucleotide Patterns of Isolated RNA Segments. In order to establish the origin of the *P* genes and the *M* and *NS* genes, oligonucleotide maps of the isolated RNAs were prepared. Fig. 3 shows the fingerprints of isolated genes or gene mixtures from A/USSR/90/77, A/Cal/10/78, and A/Tex/1/77 viruses. Comparisons of the individual gene maps and coelectrophoresis of the oligonucleotides (not shown) resulted in the determination of the A/Cal/10/78 gene-specific spots present in the corresponding genes of the H1N1 and H3N2 viruses, as

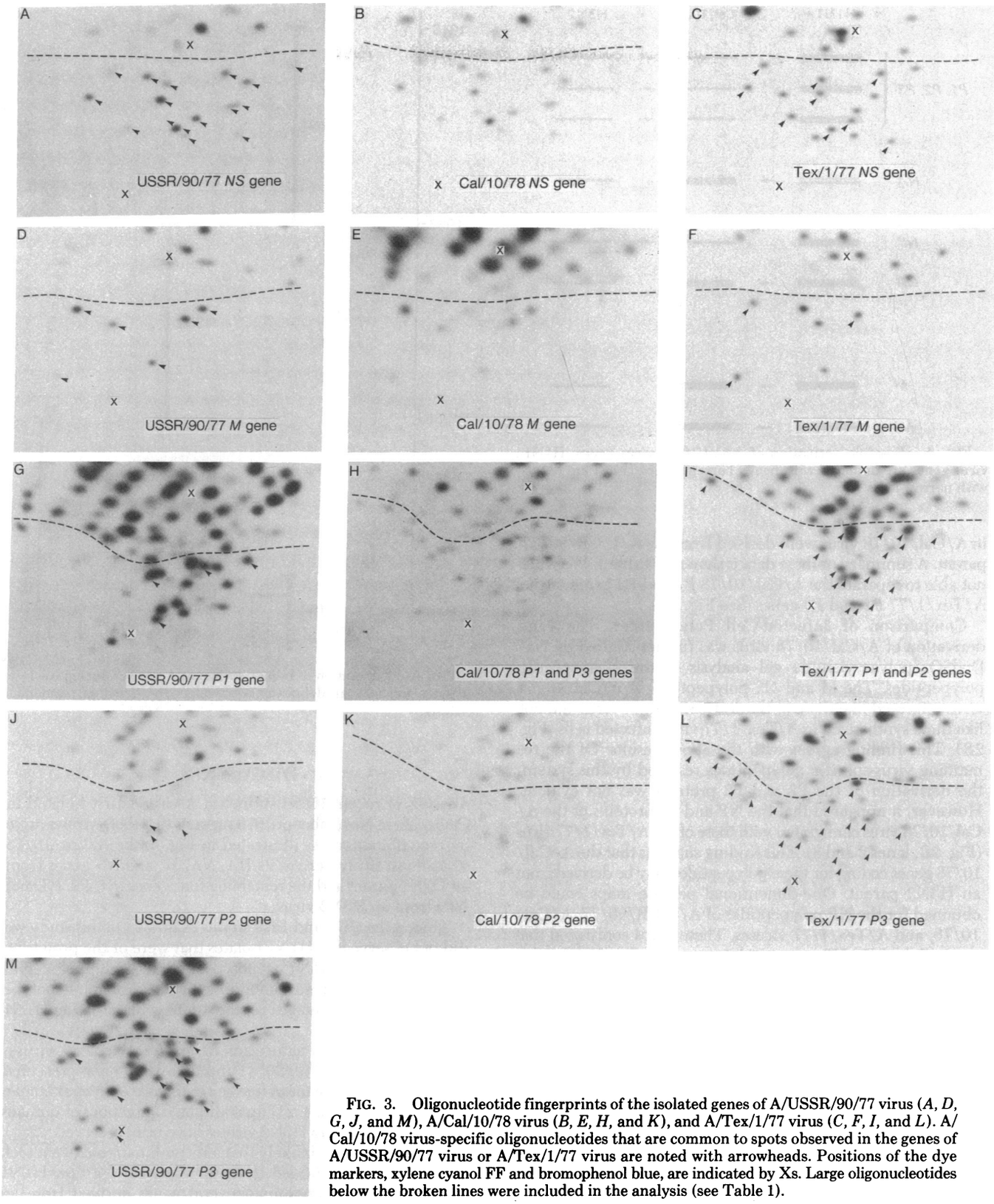

FIG. 3. Oligonucleotide fingerprints of the isolated genes of A/USSR/90/77 virus (*A*, *D*, *G*, *J*, and *M*), A/Cal/10/78 virus (*B*, *E*, *H*, and *K*), and A/Tex/1/77 virus (*C*, *F*, *I*, and *L*). A/Cal/10/78 virus-specific oligonucleotides that are common to spots observed in the genes of A/USSR/90/77 virus or A/Tex/1/77 virus are noted with arrowheads. Positions of the dye markers, xylene cyanol FF and bromophenol blue, are indicated by Xs. Large oligonucleotides below the broken lines were included in the analysis (see Table 1).

indicated by the arrows. The apparent identity of the *NS* and *M* genes of A/Cal/10/78 and A/USSR/90/77 viruses by this technique (Fig. 3 *A* and *B* and Fig. 3 *D* and *E*) confirms that these two A/Cal/10/78 genes are most likely derived from an H1N1 parent. Analysis of the fingerprints of the three *P* genes from the A/Cal/10/78 virus indicated that they are more closely related to those from the A/Tex/1/77 virus than to those of the A/USSR/90/77 virus. Although the three *P* genes from the A/Cal/10/78 virus are not identical to those from the A/Tex/1/77 virus, the similarity of these genes suggests that they are derived from an H3N2 parent. Furthermore, because the *M* and *NS* genes of the A/Cal/10/78 and the A/USSR/90/77 viruses appear to be indistinguishable, differences in the *P*-gene maps of these two viruses would not be expected if the *P* genes

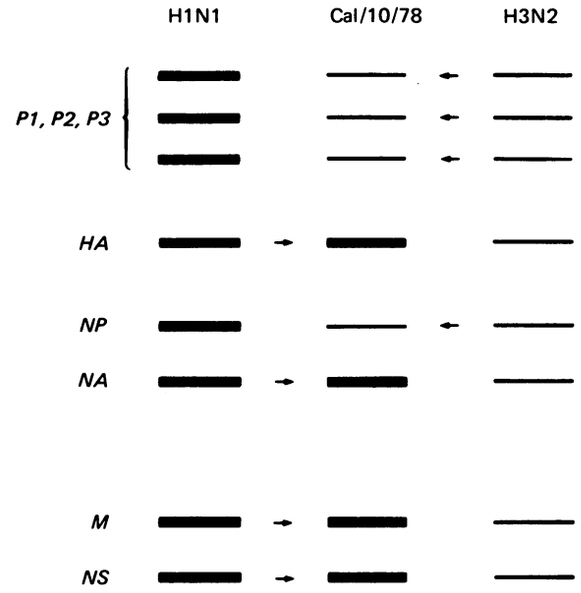

FIG. 4. Possible derivation of A/Cal/10/78 virus genes. H1N1 virus genes are depicted with heavy bars; genes of H3N2 are noted with narrow lines.

in A/Cal/10/78 virus were derived from the A/USSR/90/77 parent. A summary of these data is shown in Table 1. We were not able to separate the A/Cal/10/78 *P1* and *P3* genes or the A/Tex/1/77 *P1* and *P2* genes (See Fig. 2*A*).

Comparison of Infected-Cell Polypeptides. The gene derivation of A/Cal/10/78 virus was further studied by Na-DodSO$_4$/polyacrylamide gel analysis of the virus-specific polypeptides. The M and NS polypeptides of A/Cal/10/78 virus migrated like those of the A/USSR/90/77 virus and not like those synthesized in A/Tex/1/77 virus-infected cells (Fig. 2*B*). This finding agrees with the above results. Of the remaining virus-specific polypeptides resolved by this system, the derivation of the P2 and P3 proteins was not evident. However, it appeared that the NP and P1 proteins of the A/Cal/10/78 virus comigrated with those of the A/Tex/1/77 virus (Fig. 2*B*, lanes 2 and 3). This finding suggests that the A/Cal/10/78 genes coding for these polypeptides may be derived from an H3N2 parent. One-dimensional peptide maps could be obtained for the NP polypeptides of A/USSR/90/77, A/Cal/10/78, and A/Tex/1/77 viruses. These maps confirmed that the A/Cal/10/78 NP polypeptide is more closely related to that of the A/Tex/1/77 virus than to that of the A/USSR/90/77 virus (Fig. 2*C*).

Table 1. Gene-specific oligonucleotides of A/USSR/90/77 and A/Tex/1/77 viruses in common with oligonucleotides of the A/Cal/10/78 virus

Genes	A/USSR/90/77	A/Tex/1/77
P1	4/15	17/21
P2	3/12	
P3	9/26	18/23
M	6/6	3/6
NS	16/16	9/15

The number of spots in common with A/Cal/10/78 virus genes are shown in the numerator as compared to the total number (denominator) of A/USSR/90/77 virus and A/Tex/1/77 virus gene-specific oligonucleotides. For the analysis, spots below the broken lines in Fig. 3 were used. Boxed figures indicate the genes that show the greatest relatedness to corresponding genes in the A/Cal/10/78 virus. The *P1* and *P2* genes of the A/Tex/1/77 virus could not be separated.

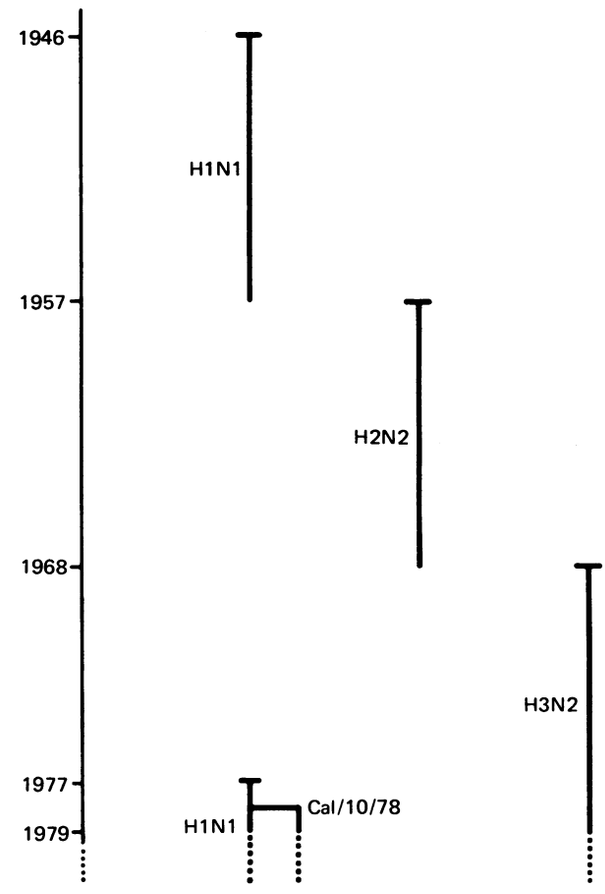

FIG. 5. Circulation of human influenza A viruses during the last 33 years. Vertical bars delineate the prevalence of individual pandemic strains.

DISCUSSION

Analysis of recent H1N1 influenza A viruses first isolated in California in November of 1978 suggests that these viruses arose by recombination. As illustrated in Fig. 4, the A/Cal/10/78 strain most likely derives its *HA*, *NA*, *M*, and *NS* genes from an H1N1 parent and the remaining four genes (*P1*, *P2*, *P3*, and *NP*) from an H3N2 virus.

Among the 1978 and 1979 strains examined in this study we did not observe any H1N1 isolates that were of the prototype A/USSR/90/77 H1N1 genotype (nonrecombinant). This suggests that the viruses containing the recombinant genotype have a survival advantage over the strains with the earlier H1N1 gene combination.

Recombination in nature among avian influenza A viruses has been observed (19, 20, †) and data have been presented suggesting this mechanism for the emergence of several human pandemic strains (3, 4, 21). In addition, coinfections of humans with both H1N1 and H3N2 viruses have been observed (22, ‡). Therefore, it is not unlikely that a recombinational event took place in humans and led to the emergence of a new H1N1 prototype strain containing several genes derived from an H3N2 virus.

† Hinshaw, V. S. & Sriram, G. (1979) *Abstracts of the Annual Meeting of the American Society for Microbiology*, Los Angeles/Honolulu, p. 260.

‡ Kendal, A. P., Beare, A. S., Cox, N. J. & Scholtissek, C. (1979) *Abstracts of the Annual Meeting of the American Society for Microbiology*, Los Angeles/Honolulu, p. 304.

The human influenza A viruses circulating during the four most recent pandemic periods are shown in Fig. 5. The "shift" from one pandemic period to another has been proposed to be the result of emerging recombinant viruses with surface protein(s) different from the preceding virus. However, this mechanism cannot explain the reemergence in 1977 of H1N1 pandemic strains, which were similar to those circulating in 1950 (5) (Fig. 5). Thus, several mechanisms may be responsible for the appearance of new pandemic strains.

During the circulation of a new pandemic virus, genetic changes or "drift" are observed in the genomes of these strains. In the past, these changes occurring in strains of one serotype were thought to involve mutational changes only. In the present paper we provide evidence that, in addition to mutation, recombination among strains can contribute to genetic variation of influenza viruses belonging to the same serotype.

We were helped by discussions with U. Desselberger, V. R. Racaniello, J. L. Schulman, and E. D. Kilbourne. We thank A. P. Kendal from the Center for Disease Control for providing us with the virus isolates. P.P. is the recipient of an I. T. Hirschl Career Research Award. This work was supported by Grant AI-11823 from the National Institutes of Health, Grant PCM78-07844 from the National Science Foundation, and Grant MV-23A from the American Cancer Society.

1. Kilbourne, E. D. (1975) in *Influenza Viruses and Influenza*, ed. Kilbourne, E. D. (Academic, New York), pp. 1–14 and pp. 483–538.
2. Laver, W. G. & Webster, R. G. (1972) *Virology* **48,** 445–455.
3. Scholtissek, C., Rohde, W., von Hoyningen, V. & Rott, R. (1978) *Virology* **87,** 13–20.
4. Laver, W. G. & Webster, R. G. (1979) *Br. Med. Bull.* **35,** 29–33.
5. Nakajima, K., Desselberger, U. & Palese, P. (1978) *Nature (London)* **274,** 334–339.
6. Scholtissek, C., von Hoyningen, V. & Rott, R. (1978) *Virology* **89,** 613–617.
7. Schulman, J. L. & Palese, P. (1976) *J. Virol.* **20,** 248–254.
8. Palese, P., Ritchey, M. & Schulman, J. L. (1977) *J. Virol.* **21,** 1187–1195.
9. Ritchey, M., Palese, P. & Kilbourne, E. D. (1976) *J. Virol.* **18,** 738–744.
10. Gaush, C. R. & Smith, T. F. (1968) *Appl. Microbiol.* **16,** 588–594.
11. Palese, P. & Schulman, J. L. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 2142–2146.
12. Ritchey, M., Palese, P. & Schulman, J. L. (1976) *J. Virol.* **20,** 307–313.
13. Young, J. F., Desselberger, U. & Palese, P. (1979) *Cell* **18,** 73–83.
14. Pedersen, F. S. & Haseltine, W. A. (1980) *Methods Enzymol.* **65,** 680–687.
15. Ritchey, M., Palese, P. & Schulman, J. L. (1977) *Virology* **76,** 122–128.
16. Maizel, J. V. (1971) in *Methods in Virology*, eds. Maramorosch, K. & Koprowski, H. (Academic, New York), Vol. 5, pp. 179–246.
17. Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. (1977) *J. Biol. Chem.* **252,** 1102–1106.
18. Racaniello, V. R. & Palese, P. (1979) *J. Virol.* **29,** 361–373.
19. Desselberger, U., Nakajima, K., Alfino, P., Pedersen, F. S., Haseltine, W. A., Hannoun, C. & Palese, P. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 3341–3345.
20. Webster, R. G., Isachenko, V. A. & Carter, M. (1974) *Bull. WHO* **51,** 325–332.
21. Laver, W. G. (1978) *Virology* **86,** 78–87.
22. Yamane, N., Arikawa, J., Odagiri, T., Sukeno, N. & Ishida, N. (1978) *Jpn. J. Med. Sci. Biol.* **31,** 431–434.