PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Genetic Structure and Genetic Variation of Influenza Viruses [and Discussion]

P. Palese, V. R. Racaniello, U. Desselberger, J. Young, M. Baez and R. J. C. Hart

Phil. Trans. R. Soc. Lond. B 1980 **288**, 299-305
doi: 10.1098/rstb.1980.0005

Phil. Trans. R. Soc. Lond. B **288**, 299–305 (1980)
Printed in Great Britain

STRUCTURE

Genetic structure and genetic variation of influenza viruses

By P. Palese, V. R. Racaniello, U. Desselberger, J. Young and M. Baez
Department of Microbiology, Mount Sinai School of Medicine, Fifth Avenue and 100th Street, New York, N.Y. 10029, U.S.A.

The following contribution summarizes our most recent results concerning analysis of the influenza A, B and C virus genomes. In addition, we present data on the extent of genetic variation of H1N1 influenza viruses isolated during the months following the 1977 outbreak of H1N1 influenza in China and Russia. A detailed description of these results is published elsewhere.

Genome structure of influenza A and B viruses

It is now well established that influenza A viruses possess a segmented genome of eight single-stranded RNAs which can be separated on polyacrylamide gels. Each of these RNA segments most probably codes for a single translational product and genetic maps correlating genes with specific gene products have been established by several investigators. Mahy's group used cytoplasmic mRNA extracts of influenza A virus-infected cells to direct virus-specific protein synthesis in a cell free translation system. When individual vRNA segments were hybridized to the mRNA extract, translation of the resulting double stranded form was blocked; however, translation of the remaining single stranded RNAs continued. Thus, it was possible to correlate specific genes with viral polypeptides (Inglis *et al.* 1977). A different approach to establishing genetic maps of influenza viruses involved the formation of non-temperature sensitive (non-*ts*) recombinants derived from *ts* mutants and a non-*ts* parent virus. Analysis of the inheritance of genes in several recombinants derived from a single *ts* parent permitted identification of the *ts* gene and correlation of a gene with its gene function (for review, see Scholtissek 1978). The influenza virus genome has also been mapped by analysing recombinants derived from different influenza A viruses. Recombinants were obtained that derive most genes from one parent and only one gene or a few genes from the other parent. 'Segregation' of single genes is recombinants and subsequent analysis of the proteins in the recombinants permitted correlation of all eight influenza A virus genes with their gene products (Palese & Schulman 1976; Ritchey *et al.* 1976*b*; Palese *et al.* 1977).

More recently such an analysis was made for two influenza B viruses (Racaniello & Palese, 1978, 1979*a*). It was established that all influenza B viruses contain eight RNA segments and that these segments are most probably monocistronic, as are the RNA segments of influenza A viruses.

The gene order of different influenza A and B viruses

Analysis of the genomes of two influenza B viruses (B/Lee and B/Maryland viruses) on polyacrylamide gels (26 g/l) containing 6 M urea showed that the genes of influenza B viruses migrated in a different order from the genes of influenza A viruses (Racaniello & Palese 1979*a*). Since we have previously shown that migration of RNA segments in urea-containing polyacrylamide gels is influenced by secondary structure (Desselberger & Palese 1978), we isolated individual genes of the two influenza B viruses and of different influenza A viruses and electrophoresed them on polyacrylamide gels after treatment with glyoxal. Under these conditions, RNAs appear to be completely denatured and thus migration of the RNAs is primarily a function of their molecular masses (McMaster & Carmichael 1977). The result of these experiments provided evidence that the gene order for the genomes of influenza A and B viruses is probably identical (Racaniello & Palese 1979*a*). In other words, in all influenza A and B viruses that were tested, the fourth largest RNA codes for the haemagglutinin, the sixth largest RNA for the viral neuraminidase, and so on. It should also be noted that corresponding genes of all the influenza B virus isolates tested had identical molecular masses when analysed after glyoxalation, and the same was observed for the genes of influenza A viruses. The only exception was the finding that the size of the haemagglutinin and the neuraminidase genes of different influenza A viruses varied slightly (Desselberger & Palese 1978). However, this small variation in gene size of the haemagglutinin and neuraminidase genes did not alter the common gene order of all the human, avian, equine and swine influenza A viruses that were tested. Thus, gel electrophoresis of glyoxalated RNAs provides a fast method for establishing genetic maps of different influenza viruses. In the past, different gel systems and unpredictable RNA patterns have made it difficult to compare the gene maps of different strains (Palese & Schulman 1976; Almond & Barry 1979). Gel electrophoresis of individual genes after glyoxalation now permits rapid identification of the genes of any 'unknown' influenza A or B virus strain regardless of unusual migration characteristics of the RNAs in polyacrylamide gels in the absence of glyoxal. A new technique was developed (Racaniello & Palese 1979*b*) in which a gel slice containing an RNA band from any influenza virus is cut from a gel after electrophoresis. The gel slice is then treated with glyoxal and placed directly into the slot of a second polyacrylamide gel. Subsequent electrophoresis in the presence of glyoxalated marker RNAs permits rapid identification of the particular gene band.

It is hoped that the glyoxalation technique will make it possible to compare influenza virus gene maps that have been established in different laboratories.

Structure of the influenza C virus genome

Analysis of the genomes of several influenza C virus strains by using different polyacrylamide gel conditions for electrophoresis suggests the presence of seven specific genes. An analysis of the RNA pattern of influenza C virus is shown in lane 3 of figure 1. (On this gel only six specific influenza C virus genes can be clearly distinguished.) Additional experiments have shown that influenza C viruses may also undergo free reassortment of their genes. MDCK cells were co-infected with a 'plaquing' and a 'non-plaquing' influenza C virus. The former virus (strain C/JJ/50) was inactivated by ultraviolet radiation before infection. Recombinants were obtained after rescue with the second 'non-plaquing' virus, C/Johannesburg/66 (C/JHG/66).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

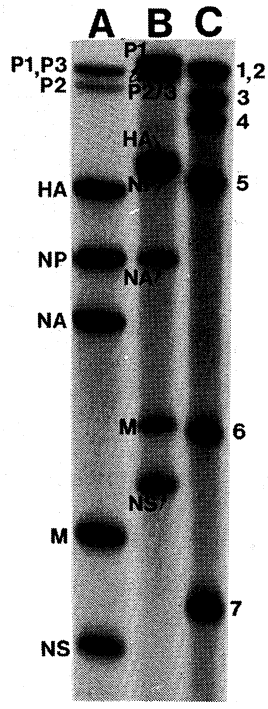

FIGURE 1. Separation on polyacrylamide gels of glyoxalated RNAs of influenza A, B and C viruses. The ^{32}P-labelled RNAs of influenza A/PR/8/34 (H0N1) virus (lane 1) and B/Lee virus (lane 2) were purified from viruses propagated in MDCK cells and ^{32}P-labelled RNA of influenza C/JHG/66 virus (lane 3) derived from egg-grown virus (Racaniello & Palese 1979*a*; Ritchey *et al.* 1976*a*). RNAs were treated with glyoxal for complete denaturation and electrophoresed on polyacrylamide gels (Desselberger & Palese 1978). Identification of the genes in the gel was made by electrophoresis of isolated RNA segments of the previously mapped A/PR/8/34 and B/Lee viruses in adjacent lanes (experiment not shown). It should be noted that the gene order for the influenza A and the influenza B virus is most probably identical. (There is some ambiguity concerning the identification of the P1 and P2 gene in influenza B virus (Racaniello & Palese 1979*a*).) It appears that the genome of influenza C virus consists of seven separate genes. A genetic map for influenza C viruses has not yet been established.

The ribonuclease T_1 resistant oligonucleotide pattern of one such recombinant, which derives specific oligonucleotide spots from both parental strains, provides biochemical evidence for recombination of influenza C viruses (figure 2) (Racaniello & Palase 1979*b*).

NUCLEOTIDE SEQUENCES OF THE TERMINAL REGIONS OF THE RNAS OF INFLUENZA A, B AND C VIRUSES

It was first shown that influenza A virus RNAs possess (p)ppA at their 5′ end (Young & Content 1970) and U_{OH} at their 3′ end (Lewandowsky *et al.* 1971). More recently Moss *et al.* (1978) presented evidence that the 5′ terminal of all influenza A and B virus segments is conserved and starts with (p)ppAGU. Data by Skehel & Hay (1978) and Barry & Mahy (1979) suggested that the conservation of sequence at the 5′ end of vRNA segments of influenza A and B viruses extends to position 12 (or 13). Sequence data of the 5′ ends of individual cRNA molecules synthesized *in vitro* suggested the presence of conserved regions at the 3′ ends of virion RNAs (Skehel & Hay 1978). Our efforts were aimed at directly sequencing the 5′ and 3′ ends of vRNA segments. In the present analysis we compared the genomes of viruses belonging to the influenza A, B and C virus groups.The sequencing techniques involved specific

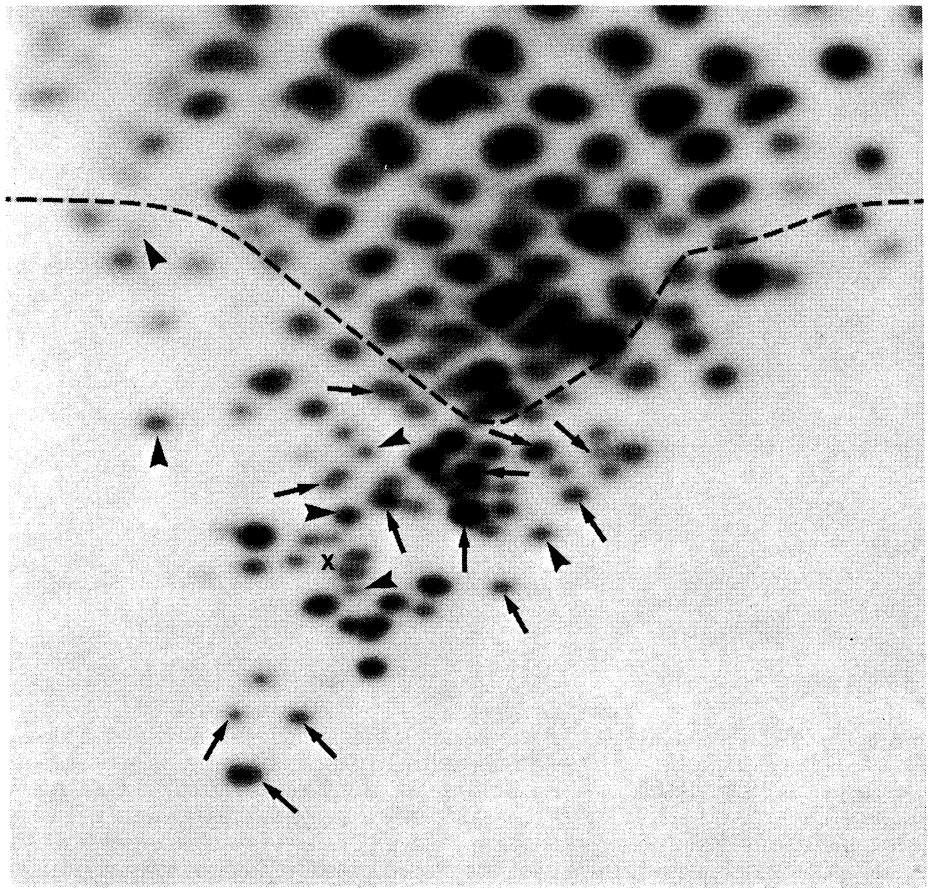

FIGURE 2. Ribonuclease T_1-resistant oligonucleotide pattern of the RNA of a recombinant influenza C virus. Recombinant P1 was obtained after co-infection of MDCK cells with influenza C/JHG/66 and C/JJ/50 viruses. The oligonucleotide pattern was obtained as previously described (Nakajima *et al.* 1978; Pedersen & Haseltine 1978). Influenza C/JHG/66 virus-specific spots are identified by arrows and C/JJ/50 virus-specific spots by arrowheads. Only spots below the broken line were included in the evaluation. Separate maps of the parents, and maps in which a mixture of the recombinant and one of the parents was coelectrophoresed, are not shown.

labelling of the 5′ or 3′ ends of individual vRNA segments, followed by partial enzyme digestion and analysis on polyacrylamide gels (Donis-Keller *et al.* 1977; Simoncsits *et al.* 1977; Efstratiadis *et al.* 1977; England & Uhlenbeck 1978). In addition, sequences were verified by using the 'wandering spot' technique (Semler *et al.* 1978). Table 1 shows the terminal sequences of the

TABLE 1. 5′ AND 3′ TERMINAL NUCLEOTIDE SEQUENCES OF THE vRNA SEGMENTS OF INFLUENZA A, B AND C VIRUSES

	5′ end	3′ end
A/PR/8/34 virus	AGUAGAAACAAGG.....	.....CCUGCUUU$^{C}_{U}$GCU
B/Lee/40 virus	AGUAG$^{U}_{A}$AACAA.....	.....U$^{U}_{G}$CGCUUCUGCU
C/JHG/66 virus	AGCAGUAGCAA.....	.....CCUGCUU$^{C}_{U}$UGCU

Sequences are conserved in all segments of the viruses except where otherwise indicated.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

vRNA segments of influenza A/PR/8/34 virus, of B/Lee/40 virus and of C/JHG/66 virus. A comparative analysis of the data reveals several characteristic features.

1. A high degree of conservation is observed for the first 11–13 nucleotides at the 5′ and 3′ ends of individual vRNA segments of a particular virus. However, sequence conservation at the ends is not complete. Position 4 at the 3′ ends of A/PR/8/34 virus segments varies. Similarly, nucleotide position 6 at the 5′ end varies in different segments of B/Lee/40 virus and one position at the 3′ ends of C/JHG/66 virus segments is not conserved.

2. The conserved 5′ and 3′ ends of individual RNA segments represent partial inverted repeats, i.e. the 5′ and 3′ terminal sequences are partly complementary.

3. In comparing the conserved terminal sequences of the genomes of an influenza A, B and C virus, a high degree of homology is observed.

It is possible that the conserved terminal sequences function as binding sites for the viral polymerase complex. Partial complementarity of the terminal sequences of vRNA segments results in similar polymerase initiation sites at the 3′ ends of cRNAs (+ strands) and of vRNAs (− strands). Although this mechanism could be the basis for the conservation of terminal regions, there are other possible explanations for the findings. Details of the sequence analysis of the vRNA segments have been published (Desselberger *et al.* 1979).

Generation of new H1N1 influenza virus strains in Nature by sequential mutational events

The unexpected outbreak of H1N1 influenza in 1977 gave us the unique opportunity to study genetic variation among pandemic influenza virus strains. Because of increased surveillance by W.H.O. laboratories and national health authorities in the entire world, we had access to different isolates obtained during the first months of the outbreak. The RNAs of such isolates were compared by using oligonucleotide pattern analysis. Previous work had demonstrated the close relationship of the H1N1 isolates of 1977 with strains that circulated in 1950 (Nakajima *et al.* 1978). This finding was even more difficult to comprehend since we demonstrated that H1N1 viruses that circulated in the beginning (1947) and at the end of the first H1N1 period (1956) differed quite extensively from those obtained in 1950. Several possibilities were discussed that might explain how these strains were genetically conserved over a 27-year period (Nakajima *et al.* 1978). In order to learn more about the 'genetic stability' of these new H1N1 isolates, several 1977 H1N1 strains were examined. In this study we determined the number of detectable mutations among different isolates by using oligonucleotide fingerprints and partial RNA sequence data for analysis and made the following observations:

1. A comparison of several isolates demonstrated changes in most of the genes (changes were not restricted to the genes coding for the surface glycoproteins, haemagglutinin and neuraminidase).

2. The conservation of particular mutations in several isolates and the appearance of new changes suggested a common ancestry of the strains and allowed the construction of an evolutionary tree (figure 3) in which the distance between strains reflects the number of observed mutations.

3. Genetic drift does occur among the new H1N1 viruses. These strains do not appear to possess unusual genetic stability.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

FIGURE 3. Possible scheme for the evolution of recent H1N1 influenza virus isolates. Viruses were compared by oligonucleotide fingerprinting and partial RNA sequencing. Changes shared by several isolates were taken into account to establish groups and possible common ancestors. The lengths of the lines between isolates represent relative genetic distances, with observed base changes as a unit of measurement (Young *et al.* 1979).

These studies were supported by Public Health Service grant AI-11823 from the National Institute of Allergy and Infectious Diseases, by grant PCM-76-11066 from the National Science Foundation, and by grant VC-234 from the American Cancer Society, P.P. is the recipient of an I.T. Hirschl Career Scientist Award.

REFERENCES (Palese *et al.*)

Almond, J. W. & Barry, R. D. 1979 *Virology* **92**, 407–415.
Barry, R. D. & Mahy, B. W. J. 1979 *Br. med. Bull.* **35**, 39–46.
Desselberger, U. & Palese, P. 1978 *Virology* **88**, 394–399.
Desselberger, U., Racaniello, V. R., Zazra, J. & Palese, P. 1979 *Gene.* (In the press.)
Donis-Keller, H., Maxam, A. M. & Gilbert, W. 1977 *Nucl. Acids Res.* **4**, 2527–2538.
Efstratiadis, A., Vournakis, J. N., Donis-Keller, H., Chaconas, G., Dougall, D. K. & Kafatos, F. C. 1977 *Nucl. Acids Res.* **4**, 4165–4172.
England, T. E. & Uhlenbeck, O. C. 1978 *Nature, Lond.* **275**, 560–561.
Inglis, S. C., McGeogh, D. J. & Mahy, B. W. J. 1977 *Virology* **78**, 552–536.
Lewandowski, L. J., Content, J. & Leppla, S. H. 1971 *J. Virol.* **8**, 701–707.
McMaster, G. K. & Carmichael, G. G. 1977 *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835–4838.
Moss, B., Keith, J. M., Gershowitz, A., Ritchey, M. B. & Palese, P. 1978 *J. Virol.* **25**, 312–318.
Nakajima, K., Desselberger, U. & Palese, P. 1978 *Nature, Lond.* **274**, 334–339.
Palese, P., Ritchey, M. B. & Schulman, J. L. 1977 *Virology* **76**, 114–121.
Palese, P. & Schulman, J. L. 1976 *Proc. natn. Acad. Sci. U.S.A.* **73**, 2142–2146.
Pedersen, F. S. & Haseltine, W. A. 1979 *Methods Enzymol.* **34**. (In the press.)
Racaniello, V. R. & Palese, P. 1978 In *Negative strand viruses and the host cell* (ed. B. W. J. Mahy & R. D. Barry), pp. 4–36. London: Academic Press.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Racaniello, V. R. & Palese, P. 1979*a* *J. Virol.* **29**, 361–373.
Racaniello, V. R. & Palese, P. 1979*b* In preparation.
Ritchey, M. B., Palese, P. & Kilbourne, E. D. 1976*a* *J. Virol.* **18**, 738–744.
Ritchey, M. B., Palese, P. & Schulman, J. L. 1976*b* *J. Virol.* **20**, 307–313.
Scholtissek, C. 1978 *Curr. Topics Microbiol. Immunol.* **80**, 139–169.
Semler, B. L., Perrault, J., Abelson, J. & Holland, J. J. 1978 *Proc. natn. Acad. Sci. U.S.A.* **75**, 4704–4708.
Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. & Guilley, H. 1977 *Nature, Lond.* **269**, 833–836.
Skehel, J. J. & Hay, A. J. 1978 *Nucl. Acids. Res.* **5**, 1207–1220.
Young, J. F., Desselberger, U. & Palese, P. 1979 *Cell* **18**, 73–83.
Young, R. J. & Content, J. 1970 *Nature, new Biol.* **230**, 140–142.

Discussion

R. J. C. Hart (*Public Health Laboratory, Church Lane, Heavitree, Exeter, Devon, U.K.*). In February 1960 an H1N1 virus resembling A/England/1/51 was isolated from a young soldier suffering from pyrexia, headache and sore throat. He appeared to be one case in a small outbreak. The patient's sera showed rising titres of antibody to his own virus and A/Eng/1/51, but not to H2N2 viruses (Isaacs, A., Hart, R. J. C. & Law, V. G. 1962, *Bull. Wld Hlth Org.* **26**, 253–259).

U. Desselberger. In this context, it might be worth mentioning the kind of variation that we found for influenza viruses isolated from different animal species. When comparing groups of avain viruses of the same subtype (Hav6Nav4 and Hav1N2), which were isolated by Dr Hannoun from the same flock of birds in 1976 and 1977, it was found that their RNAs differed considerably: all eight RNA segments were found to show differences, as evidenced by changes in their oligonucleotide maps. A more extensive study of equine influenza virus strains (Heq1-Neq1) isolated between 1956 and 1977 showed that three different groups (1956, 1964–6, 1969–77) could be differentiated with respect to their oligonucleotide maps. Among isolates of the most recent group, much less variation was observed than among human or avian influenza viruses isolated over a comparable period of time. Thus, the extent of genetic variation in different influenza A viruses may depend on the strain and/or on the animal host.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

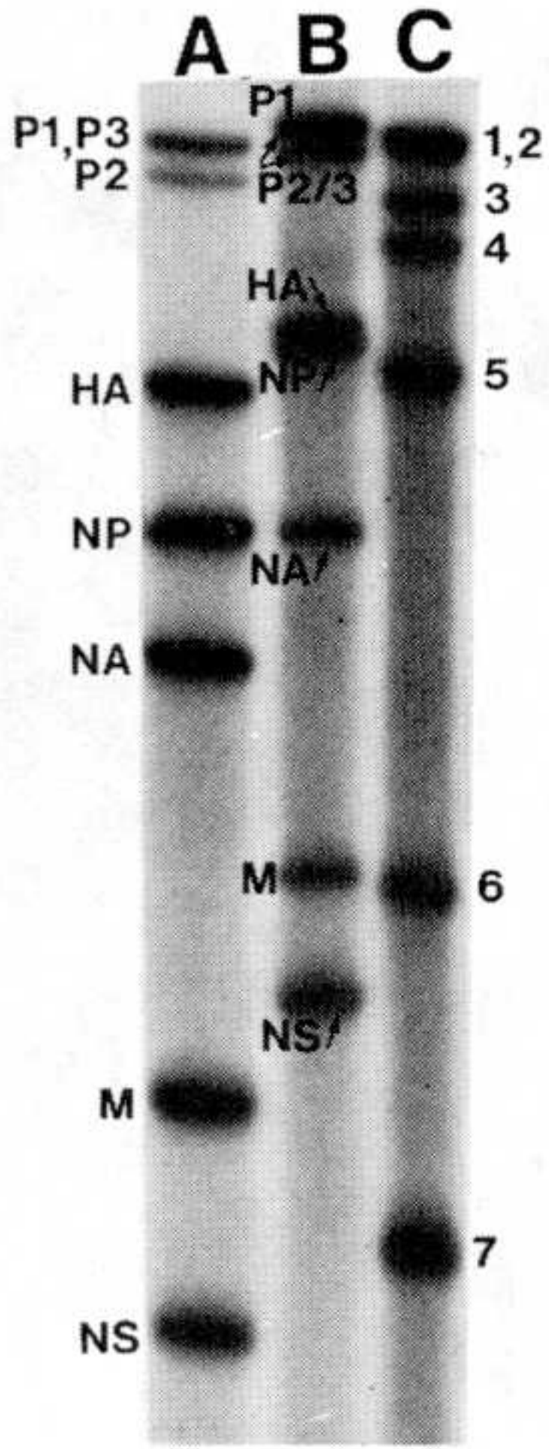

FIGURE 1. Separation on polyacrylamide gels of glyoxalated RNAs of influenza A, B and C viruses. The ^{32}P-labelled RNAs of influenza A/PR/8/34 (H0N1) virus (lane 1) and B/Lee virus (lane 2) were purified from viruses propagated in MDCK cells and ^{32}P-labelled RNA of influenza C/JHG/66 virus (lane 3) derived from egg-grown virus (Racaniello & Palese 1979*a*; Ritchey *et al.* 1976*a*). RNAs were treated with glyoxal for complete denaturation and electrophoresed on polyacrylamide gels (Desselberger & Palese 1978). Identification of the genes in the gel was made by electrophoresis of isolated RNA segments of the previously mapped A/PR/8/34 and B/Lee viruses in adjacent lanes (experiment not shown). It should be noted that the gene order for the influenza A and the influenza B virus is most probably identical. (There is some ambiguity concerning the identification of the P1 and P2 gene in influenza B virus (Racaniello & Palese 1979*a*).) It appears that the genome of influenza C virus consists of seven separate genes. A genetic map for influenza C viruses has not yet been established.

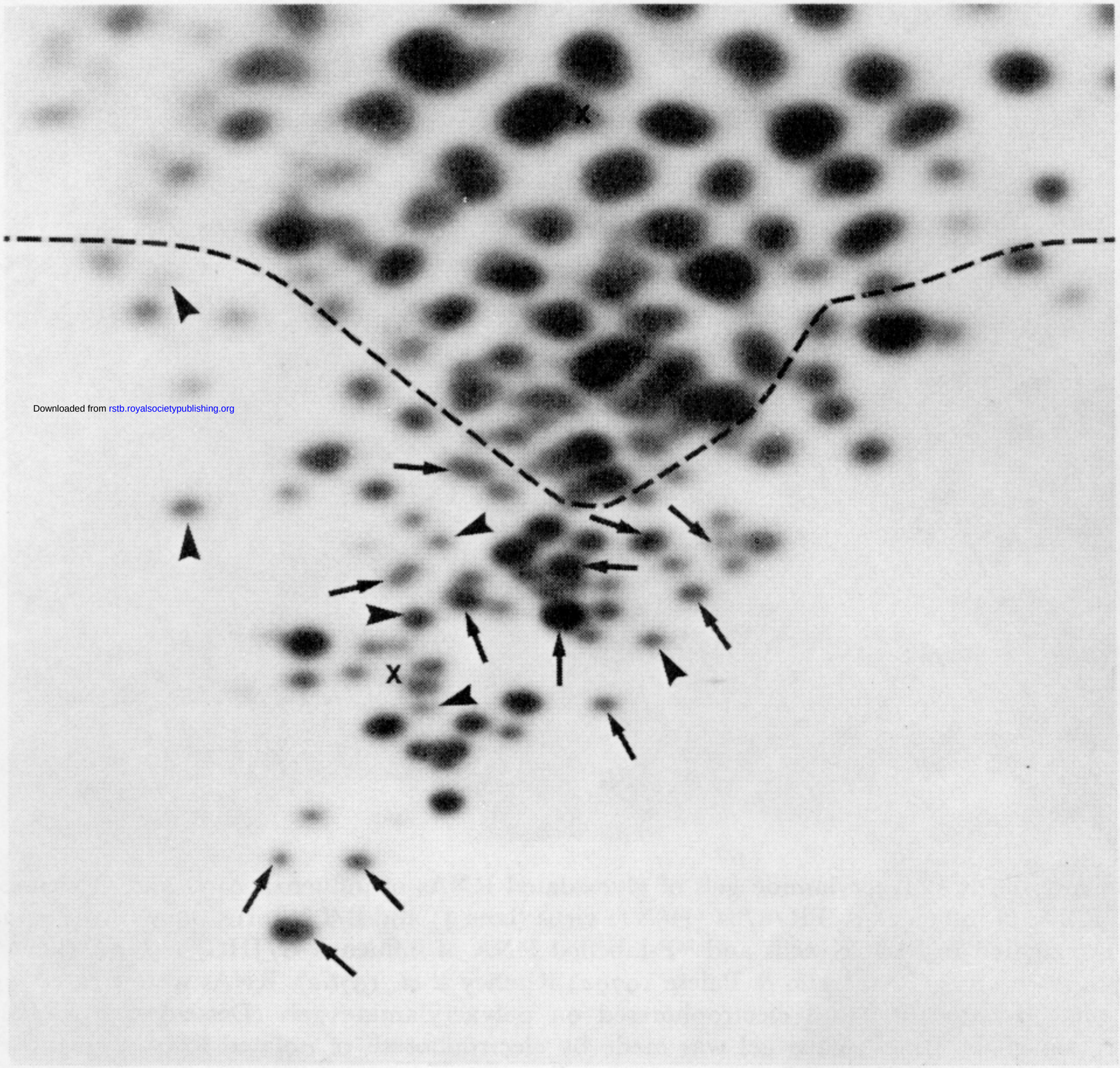

FIGURE 2. Ribonuclease T_1-resistant oligonucleotide pattern of the RNA of a recombinant influenza C virus. Recombinant P1 was obtained after co-infection of MDCK cells with influenza C/JHG/66 and C/JJ/50 viruses. The oligonucleotide pattern was obtained as previously described (Nakajima *et al.* 1978; Pedersen & Haseltine 1978). Influenza C/JHG/66 virus-specific spots are identified by arrows and C/JJ/50 virus-specific spots by arrowheads. Only spots below the broken line were included in the evaluation. Separate maps of the parents, and maps in which a mixture of the recombinant and one of the parents was coelectrophoresed, are not shown.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES