Journal of General Virology (1991), **72**, 2283–2287. *Printed in Great Britain*

Sequence characterization of the matrix protein genes of parainfluenza virus types 4A and 4B

Kunio Kondo,[1,2] Masashi Fujii,[1] Taichi Nakamura,[1] Hisanori Bando,[1] Mitsuo Kawano,[1] Masato Tsurudome, Hiroshi Komada,[1] Shigeru Kusakawa,[1] Machiko Nishio[1] and Yasuhiko Ito[1]*

[1] *Department of Microbiology, Mie University School of Medicine, 2-174 Edobashi, Tsu-Shi, Mie Prefecture 514 and* [2] *Fujikura Research Center, Fujikura Kasei Co. Ltd, 3-20-7 Hasune, Itabashi, Tokyo, Japan*

The complete nucleotide sequences of the matrix protein (M) genes of parainfluenza virus types 4A and 4B (PIV-4A and -4B) were determined from cDNA of the mRNA, and found to be 1548 bases in length, exclusive of poly(A) sequences. The sequences contained a large open reading frame of 1146 nucleotides encoding 362 amino acids. A high degree of identity (96·1%) was observed between the amino acid sequences of PIV-4A and PIV-4B M. These M sequences were compared with those of 10 other paramyxoviruses and a phylogenetic tree was constructed.

Human parainfluenza virus type 4 (hPIV-4) is an important causative agent of upper respiratory tract infection in infants and adults. HPIV-4 has been subdivided into two serotypes, A and B. Although a large number of monoclonal antibodies (MAbs) directed against PIV-4A and -4B had been produced (Komada *et al.*, 1989*b*), information was limited until recently, when the structural polypeptides of hPIV-4A were described (Komada *et al.*, 1989*a*), and the nucleotide sequences of the nucleoprotein (NP) (Kondo *et al.*, 1990*a*), haemagglutinin–neuraminidase (HN) (Bando *et al.*, 1990) and phosphoprotein (P) (Kondo *et al.*, 1990*b*) genes of hPIV-4A and -4B were determined.

Paramyxovirus matrix proteins (M) appear to play important roles in virion assembly, the establishment of persistent infection and viral pathogenesis. The antigenicity of hPIV-4A M differs significantly from that of hPIV-4B (Komada *et al.*, 1989*b*). Furthermore, antigenic diversity was observed within the same serotype (Komada *et al.*, 1989*b*). In this study, the nucleotide sequences of the M genes of hPIV-4A and -4B were determined and compared with those of other paramyxovirus M genes.

HPIV-4A (Toshiba strain), hPIV-4B (68-333 strain), Vero cells, COS cells and primary monkey kidney (PMK) cells were used in this study. Virus propagation and purification of virus mRNA and nucleocapsid RNA were performed as reported previously (Kondo *et al.*, 1990*a*, *b*). The cDNA libraries derived from mRNA were constructed according to the method of Okayama & Berg (1982), and genomic cDNA libraries were constructed from hPIV-4 NP RNA by the method of Gubler & Hoffmann (1983) using synthetic hexanucleotides as primers (Kondo *et al.*, 1990*a*). Sequencing was performed by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977), using T7 DNA polymerase (United States Biochemical) and [α-^{35}S]dATP (NEN). The analysis of the sequence data was carried out using the SDC GENETYX program.

We constructed cDNA libraries for poly(A)$^+$ mRNA and genomic RNA obtained from hPIV-4A- or -4B-infected PMK cells. HPIV-4 M gene-specific clones were screened by colony hybridization using the inserted DNA from hPIV-4A or -4B P gene-specific cDNA clones (Kondo *et al.*, 1990*b*) as probes. The M gene specificity of the cDNA clones was confirmed by immunofluorescent staining of COS cells transfected with plasmid (pDS-4aM2) with anti-hPIV-4A M MAbs (data not shown). The complete nucleotide sequences of the hPIV-4A and -4B M genes, and the deduced amino acid sequences are shown in Fig. 1. The M mRNA cDNAs were found to be 1548 bases in length, exclusive of poly(A) sequences. The first AUG codon was found 31 bases form the 5′ end, the flanking sequences of which were favourable for it to be the actual translation initiation site (Kozak, 1986); the termination codon, UAA, was located at nucleotides 1178 to 1180. Thus the sequences contained a large open reading frame of 1146 nucleotides encoding a polypeptide of 382 amino acids

The nucleotide sequence data in this paper have been submitted to the DDBJ, EMBL and GenBank nucleotide sequence databases and assigned the accession numbers DO1050 (hPIV-4A) and DO1051 (hPIV-4B).

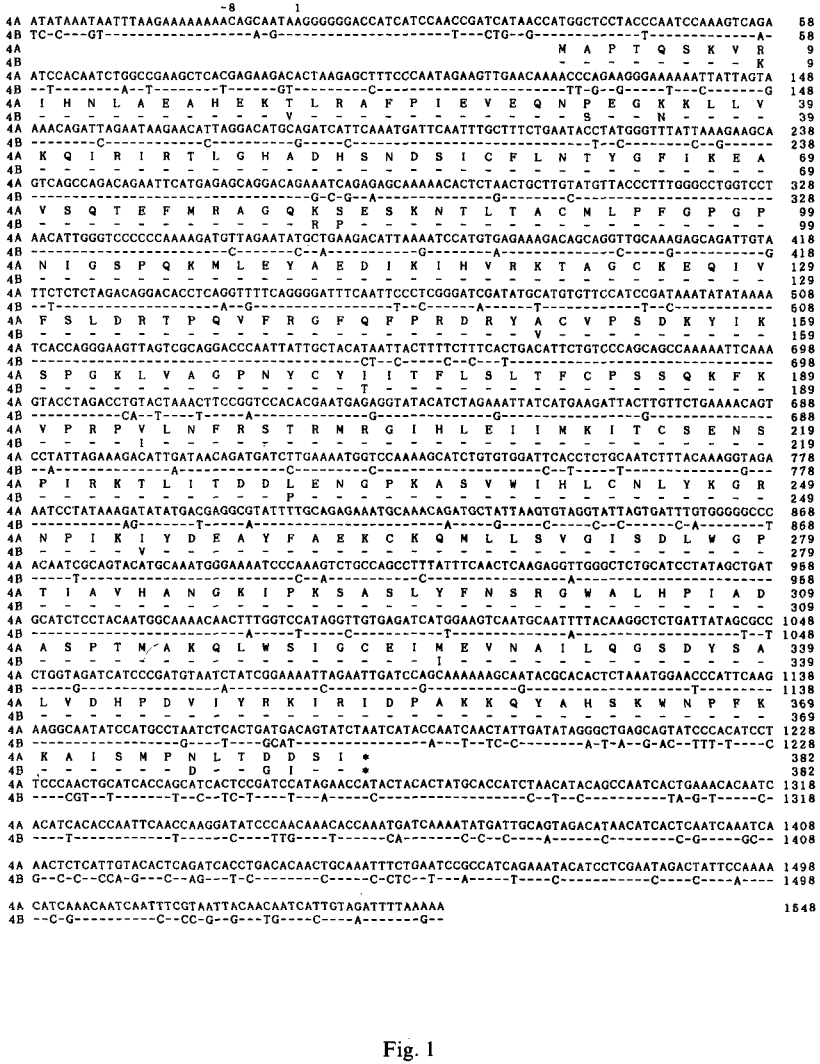

Fig. 1

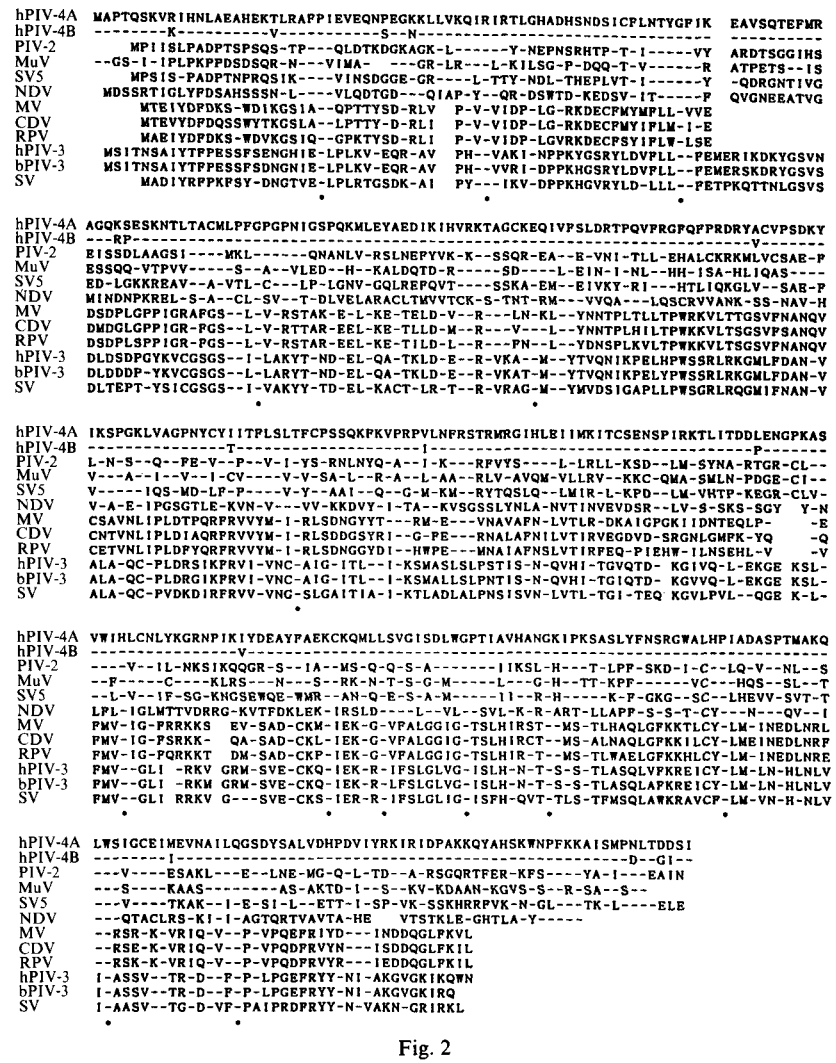

Fig. 2

Fig. 1. Nucleotide sequences of the M genes of hPIV-4A and -4B in the positive sense, and the predicted amino acid sequence. The 5′ end of the M mRNA was determined by primer extension.

Fig. 2. Sequence alignment for the Ms of PIV-2 (Kawano *et al.*, 1990), MuV (Elliott *et al.*, 1989), SV5 (Sheshberadaran & Lamb, 1990), Newcastle disease virus (NDV; Chambers *et al.*, 1986), measles virus (MV) and canine distemper virus (CDV) (Bellini *et al.*, 1986), RPV (Limo & Yima, 1990), hPIV-3 (Galinski *et al.*, 1987), bPIV-3 (Sakai *et al.*, 1987) and Sendai virus (SV) (Blumberg *et al.*, 1984). Asterisks under the sequences indicate residues common to all 12 viruses.

Table 1. *Amino acid sequence identity between paramyxovirus Ms*

Virus*	hPIV-4B	PIV-2	MuV	SV5	NDV	MV	CDV	RPV	hPIV-3	bPIV-3	SV
hPIV-4A	96·6† (0·03)	39·0 (1·15)	49·4 (0·81)	38·7 (1·17)	28·2 (1·72)	19·3 (2·77)	19·3 (2·77)	19·0 (2·83)	19·9 (2·65)	19·9 (2·65)	19·6 (2·71)
hPIV-4B		39·0 (1·15)	49·1 (0·82)	38·3 (1·18)	27·3 (1·79)	19·9 (2·65)	19·0 (2·83)	18·7 (2·90)	19·6 (2·71)	19·6 (2·71)	19·3 (2·77)
PIV-2			42·6 (1·02)	50·6 (0·78)	27·3 (1·79)	20·9 (2·48)	19·0 (2·83)	18·1 (3·06)	20·2 (2·60)	20·2 (2·60)	17·2 (3·36)
MuV				43·6 (0·99)	25·5 (1·94)	19·6 (2·71)	21·2 (2·43)	19·0 (2·83)	21·8 (2·35)	21·2 (2·43)	17·8 (3·15)
SV5					27·3 (1·79)	17·5 (3·25)	18·7 (2·90)	16·6 (3·62)	18·4 (2·98)	19·3 (2·77)	16·0 (3·87)
NDV						18·7 (2·90)	20·6 (2·53)	19·3 (2·77)	21·2 (2·43)	21·5 (2·39)	18·4 (2·98)
MV							76·7 (0·28)	78·8 (0·25)	36·8 (1·24)	37·4 (1·22)	33·7 (1·39)
CDV								78·8 (0·25)	35·6 (1·30)	36·2 (1·27)	35·0 (1·33)
RPV									36·5 (1·26)	37·1 (1·23)	36·2 (1·27)
hPIV-3										93·6 (0·07)	63·2 (0·50)
bPIV-3											65·3 (0·46)

* NDV, Newcastle disease virus; MV, measles virus; CDV, canine distemper virus.

† Percentage identity between Ms of the paramyxoviruses. Figures in parentheses are the number of amino acid substitutions calculated using the formula of Kimura (1983).

with a predicted M_r of 43134 (hPIV-4A) or 43026 (hPIV-4B). This calculated M_r agreed with that estimated (40K) by SDS–PAGE (Komada *et al.*, 1989*a*).

Nucleotide sequence identity in the coding region, 5′ non-coding region, 3′ non-coding region and overall was 90·9, 83·9, 75·5 and 87·1%, respectively. In the coding region, there were 104 nucleotide changes, only 19 of which altered the encoded amino acid residue.

The net charge of hPIV-4 M, calculated for simplicity as Lys, +1, Arg, +1, His, +0·5, Asp, −1 and Glu, −1, was +22·5 at neutral pH. A relative abundance (47·9%, hPIV-4A and 48·4%, hPIV-4B) of hydrophobic residues (Gly, Ala, Val, Leu, Ile, Met, Phe, Try and Pro) was found in M. These results indicate that hPIV-4 M is highly positively charged and is a hydrophobic protein.

A high degree of identity was observed between the deduced hPIV-4A and -4B M amino acid sequences (96·1% overall; Fig. 1). This degree of identity was greater than the immunological relatedness determined using MAbs (70% cross-reactivity; Komada *et al.*, 1989*b*). The predicted secondary structures (Chou–Fasman) were almost identical except for those at the C termini (data not shown).

Dot-matrix comparisons of the amino acid sequence of M of different paramyxoviruses were performed. HPIV-4 showed the greatest similarity with mumps virus (MuV) (Elliott *et al.*, 1989), moderate similarity with PIV-2 (Kawano *et al.*, 1990) and simian virus 5 (SV5) (Sheshberadaran & Lamb, 1990), and little similarity with rinderpest virus (RPV) (Limo & Yilma, 1990) and Sendai virus (Blumberg *et al.*, 1990) (data not shown). For further analysis the predicted amino acid sequence of hPIV-4 M was aligned with those of 10 other paramyxoviruses (Fig. 2) and the degree of similarity was calculated (Table 1). The number of amino acid substitutions was then calculated from the proportion of different amino acids in the sequences using the formula of Kimura (1983). Subsequently a phylogenetic tree was constructed using the unweighted pairwise grouping (UPG) method (Fig. 3). From this analysis of the similarities of the Ms, paramyxoviruses were divided into two groups; one group consisted of hPIV-4A, -4B and -2, and MuV, the other of Sendai virus (PIV-1), hPIV-3 and bovine PIV-3 (bPIV-3). This grouping was consistent with our previous studies (Ito *et al.*, 1987; Kawano *et al.*, 1990; Kondo *et al.*, 1990*b*; Bando *et al.*, 1990); morbilliviruses clearly belonged to the latter group.

Fourteen amino acids were conserved in all the paramyxovirus Ms compared, but there was no homologous region. Furthermore, the Ms of hPIV-4A and -4B contained 10 cysteines, none of which was conserved in all the viruses examined, but five were conserved in hPIV-4 and -2, and MuV, indicating that similar structures were, to a certain extent, conserved in the same group.

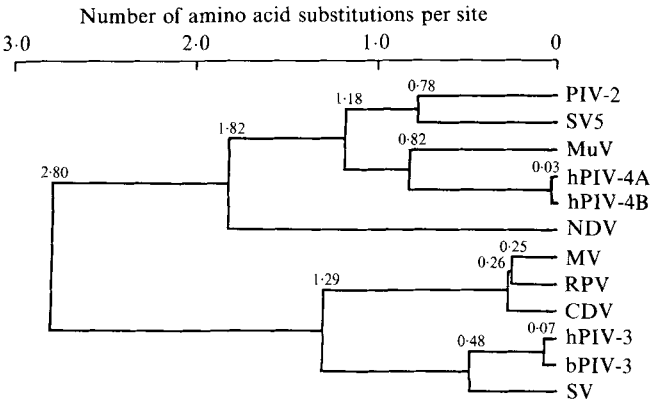

Fig. 3. The phylogenetic tree for the Ms of the paramyxoviruses. On the basis of the alignment (Fig. 2) we estimated the number of amino acid substitutions for each pair (Table 1) and constructed a phylogenetic tree for the 12 paramyxoviruses using the UPG method (Kimura, 1983).

The identity between the amino acid sequences of the Ms of hPIV-4A and -4B was found to be 96·1%, which is greater than that (93·6%) between hPIV-3 and bPIV-3 (Galinski *et al.*, 1987; Sakai *et al.*, 1987), and lower than that (99·5%) between different strains of MuV (Tanabayashi *et al.*, 1990). These data support the subdivision of hPIV-4 into subtypes A and B.

In a previous study (Komada *et al.*, 1989*b*), a large number of MAbs against hPIV-4 were isolated and characterized. Nineteen MAbs against hPIV-4A M show only 57% cross-reactivity with hPIV-4B and 88% with other strains of hPIV-4A. This immunological cross-reactivity between hPIV-4A and -4B Ms is inconsistent with the identity between the primary amino acid sequences of M of hPIV-4. Although the identity (93%) between the amino acid sequences of NPs of hPIV-4 strains was lower than that for M, 12 of 14 MAbs against hPIV-4A NP showed cross-reactivity with four strains of hPIV-4B, and 13 MAbs against hPIV-4B NP reacted with four strains of hPIV-4A. This discrepancy could be explained in part by experimental results showing that the antigenicity of NP is dependent on the primary amino acid sequence, whereas the antigenicity of M is determined by the tertiary structure. All 14 anti-hPIV-4A NP MAbs reacted with the polypeptide in a Western blot assay, suggesting that anti-NP MAbs recognize continuous sequences of the polypeptide (unpublished data). By contrast, only two of the 19 MAbs against hPIV-4A M react in Western blots (unpublished data). These findings show that the tertiary structure of M may play an important role in its antigenicity and function.

The intergenic regions between the P and M genes of hPIV-4A and -4B are eight nucleotides long, the sequences being CAGCAATA and CAGAAGTA, respectively; the M–F and F–HN intergenic regions consist of 17 and 7 bases, respectively (unpublished data). Therefore, the length and sequence of the intergenic regions of hPIV-4 are not conserved. No nucleotide sequence or length conservation has been found in the intergenic regions of MuV (Elango *et al.*, 1988), SV5 (Sheshberadaran & Lamb, 1990) or PIV-2 (unpublished data), all of which belong to the same group of paramyxoviruses. On the other hand, the intergenic regions in members of the other group, Sendai virus and PIV-3, are trinucleotides (Gupta & Kingsbury, 1984).

References

Bando, H., Kondo, K., Kawano, M., Komada, H., Tsurudome, M., Nishio, M. & Ito, Y. (1990). Molecular cloning and sequence analysis of human parainfluenza type 4A virus HN gene: its irregularities in structure and activities. *Virology* **175**, 307–312.

Bellini, W. J., Englund, G., Richardson, C. D., Rozenblatt, S. & Lazzarini, R. A. (1986). Matrix genes of measles virus and canine distemper virus: cloning, nucleotide sequences, and deduced amino acid sequences. *Journal of Virology* **58**, 408–416.

Blumberg, B. M., Rose, K., Sinoma, M. G., Roux, L., Giorgi, C. & Kolakofsky, D. (1984). Analysis of the Sendai virus M gene and protein. *Journal of Virology* **52**, 656–663.

Chambers, P., Millar, N. S., Platt, S. G. & Emmerson, P. T. (1986). Nucleotide sequence of the gene encoding the matrix protein of Newcastle disease virus. *Nucleic Acids Research* **14**, 9051–9061.

Elango, N., Varsanyi, T. M., Kövamees, J. & Norrby, E. (1988). Molecular cloning and characterization of six genes, determination of gene order and intergenic sequences and leader sequence of mumps virus. *Journal of General Virology* **69**, 2893–2900.

Elliott, G. D., Afzal, M. A., Martin, S. J. & Rima, B. K. (1989). Nucleotide sequence of the matrix, fusion and putative SH protein genes of mumps virus and their deduced amino acid sequences. *Virus Research* **12**, 61–75.

Galinski, M. S., Mink, M. A., Lambert, D. M., Wechsler, S. L. & Pons, M. W. (1987). Molecular cloning and sequence analysis of the human parainfluenza 3 virus gene encoding the matrix protein. *Virology* **157**, 24–30.

Gubler, U. & Hoffmann, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.

Gupta, K. C. & Kingsbury, D. W. (1984). Complete sequences of the intergenic and mRNA start signals in the Sendai virus genome: homologies with the genome of vesicular stomatitis virus. *Nucleic Acids Research* **12**, 3829–3841.

Ito, Y., Tsurudome, M., Hishiyama, M. & Yamada, A. (1987). Immunological interrelationships among human and non-human paramyxoviruses revealed by immunoprecipitation. *Journal of General Virology* **68**, 1289–1297.

Kawano, M., Bando, H., Ohgimoto, S., Okamoto, K., Kondo, K., Tsurudome, M., Nishio, M. & Ito, Y. (1990). Complete nucleotide sequence of the matrix gene of human parainfluenza type 2 virus and expression of the M protein in bacteria. *Virology* **179**, 857–861.

Kimura, M. (1983). Rate of evolution at the molecular level. Comparative studies of protein sequences. In *The Neutral Theory of Molecular Evolution*, pp. 65–76. Edited by M. Kimura. Cambridge: Cambridge University Press.

Komada, H., Tsurudome, M., Bando, H., Nishio, M. & Ito, Y. (1989*a*). Virus specific polypeptides of human parainfluenza virus type 4 and their synthesis in infected cells. *Virology* **171**, 254–259.

Komada, H., Tsurudome, M., Ueda, M., Nishio, M., Bando, H. & Ito, Y. (1989*b*). Isolation and characterization of monoclonal

antibodies to human parainfluenza virus type 4 and their use in revealing the antigenic relationship between subtype 4A and 4B. *Virology* **171**, 28–37.

KONDO, K., BANDO, H., KAWANO, M., TSURUDOME, M., KOMADA, H., NISHIO, M. & ITO, Y. (1990*a*). Sequence analyses and comparison of parainfluenza virus type 4A and 4B protein genes. *Virology* **174**, 1–8.

KONDO, K., BANDO, H., TSURUDOME, M., KAWANO, M., NISHIO, M. & ITO, Y. (1990*b*). Sequence analysis of the phosphoprotein (P) genes of human parainfluenza type 4A and 4B viruses and RNA editing at transcript of the P genes: the number of G residues added is imprecise. *Virology* **178**, 321–326.

KOZAK, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292.

LIMO, M. & YILMA, T. (1990). Molecular cloning of the rinderpest virus matrix gene: comparative sequence analysis with other paramyxoviruses. *Virology* **175**, 323–327.

OKAYAMA, H. & BERG, P. (1982). High efficiency cloning of full length cDNAs. *Molecular and Cellular Biology* **2**, 161–170.

SAKAI, Y., SUZU, S., SHIODA, T. & SHIBUTA, H. (1987). Nucleotide sequence of the bovine parainfluenza 3 virus genome: its 3′ end and the genes of NP, P, C and M proteins. *Nucleic Acids Research* **15**, 2927–2944.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5463–5467.

SHESHBERADARAN, H. & LAMB, R. A. (1990). Sequence characterization of the membrane protein gene of paramyxovirus simian virus 5. *Virology* **176**, 234–243.

TANABAYASHI, K., TAKEUCHI, K., HISHIYAMA, M., YAMADA, A. & SUGIURA, A. (1990). Nucleotide sequence of the gene encoding the matrix protein of mumps virus (Miyahara strain). *Virus Genes* **3**, 361–365.

(Received 3 April 1991; Accepted 23 May 1991)