Characterization of five conserved genotypes of the mumps virus small hydrophobic (SH) protein gene

Claes Örvell, Mina Kalantari and Bo Johansson

Huddinge University Hospital, Department of Clinical Virology, Karolinska Institutet, S-141 86 Huddinge, Sweden

Twenty-one different mumps virus isolates from Sweden and Japan collected over 25 years were compared by nucleotide sequence analysis of the small hydrophobic (SH) protein gene, and the deduced 57 amino acid sequences of the coding part of the gene were aligned with published sequences of viral isolates from the USA, the UK, Sweden and Japan. Five genotypes were found which, in accordance with previously used nomenclature, were named A to E. Genotypes A, C, D and E were found in Europe and genotype B was found in Japan. Amino acid signature sequence motifs specific for each genotype were identified. A triplet of three amino acids at positions 28–30 was the most characteristic. Different genotypes can circulate simultaneously in a given geographical location. In Stockholm, genotypes A and D or C and D were found over different time periods. In contrast, only genotype B was found in Japan.

Mumps virus belongs to the family *Paramyxoviridae*, genus *Rubulavirus* (Rima *et al.*, 1995). About 5 years ago the gene order of six mumps virus genes, and their nucleotide sequences except for that encoding the large (L) protein, were determined (Waxham *et al.*, 1987, 1988; Elango, 1989; Elango *et al.*, 1988, 1989; Elliott *et al.*, 1989, 1990; Yamada *et al.*, 1989; Tanabayashi *et al.*, 1990). Strain variations in the nucleotide sequence have been described for the phospho (P) protein (Yamada *et al.*, 1989), fusion (F) protein (Takeuchi *et al.*, 1989; Forsey *et al.*, 1990) and haemagglutinin–neuraminidase (HN) protein (Waxham *et al.*, 1988; Kövamees *et al.*, 1990) genes. The small hydrophobic (SH) protein gene is located between the F and HN gene on the genome map (Elango *et al.*, 1988, 1989; Elliott *et al.*, 1989). The gene order on the genomic RNA is 3′-N-P-M-F-SH-HN-L-5′, where N, M and L denote the nucleocapsid, membrane and large protein genes, respectively. The SH gene of mumps virus encodes a presumed protein 57 amino acids in length, but the protein has not yet been identified immunologically in virions or in infected cells. Comparisons of the nucleotide sequences from different mumps virus strains have shown that there exist sequence differences in the SH gene (Yeo *et al.*, 1993). The heterogeneity in this part of the genome was more pronounced than that of the P, F and HN genes analysed by other groups (Yamada *et al.*, 1989; Forsey *et al.*, 1990; Brown *et al.*, 1991). Recent studies of the nucleotide sequences of the SH gene of different strains of mumps virus have shown the existence of four genotypes, which have been named A to D (Yeo *et al.*, 1993; Künkel *et al.*, 1994, 1995). In Japan only genotype B has been found, but this genotype has not yet been found in Europe or the USA (Takeuchi *et al.*, 1991; Yeo *et al.*, 1993). Recently isolated wild-type mumps virus strains in Europe belong to the A or C genotype (Yeo *et al.*, 1993; Künkel *et al.*, 1995).

In the present investigation, mumps virus isolates from the cerebrospinal fluid (CSF) of 15 individuals were studied (Table 1). In addition, a recent sample (VS 4098) diagnosed from the serum sample of a patient with mumps by using nested PCR was examined, together with five isolates of mumps virus provided by Japanese colleagues and kindly collected and sent to Sweden by Yoshinobu Kimura, Department of Microbiology, Fukui Medical School, 910-911 Fukui, Japan. Isolation of mumps virus RNA and PCR amplification of a 415 bp fragment encompassing the entire SH gene were performed essentially as has been described for hepatitis C virus by Yun *et al.* (1993). Primers for nested PCR flanking the SH gene were selected using the Oligo version 5.0 program (NBI, Plymouth, MN 55447-5434, USA). The outer primers (BJSHPR3, 5′ CGATGATCTCATCAGGTACT 3′; BJSHPR4, 5′ TCCTAAGTCTGTTCTGGCTT 3′) amplified a 430 bp fragment and the inner (nested) primers (BJSHPR1, 5′ TGTCGATGATCTCATCAG 3′; BJSHPR2, 5′ GTTGTGTTGTGATCCTAAGT 3′) amplified a 415 bp fragment. One of the nested primers (BJSHPR2) was 5′-biotinylated and the other (BJSHPR1) was coupled at the 5′ end with a universal M13 primer sequence (CGACGTTGTAAAACGGCCAGT).

Direct DNA sequencing was performed from a biotinylated nested PCR product that was immobilized onto streptavidin-coated magnetic beads (Dynal) and single-stranded DNA was generated by denaturation with 0·1 M NaOH. The sequencing was carried out using FITC-labelled M13 universal primer and

Author for correspondence: Claes Örvell.
Fax +46 8 746 53 60. e-mail clor@vird01.hs.sll.se

0001-4047 © 1997 SGM

Table 1. Virus strains arranged by year of isolation

Isolate	Year of isolation	Country of isolation	Genotype	Source or reference
F 24630 V1	1970	Sweden	D	Stockholm
F 24638 V3	1970	Sweden	D	Stockholm
F 26068	1970	Sweden	D	Stockholm
F 27631 V4	1971	Sweden	D	Stockholm
F 27806 V5	1971	Sweden	D	Stockholm
F 32247 V10	1972	Sweden	D	Stockholm
F 33609 V11	1973	Sweden	D	Stockholm
F 50647	1978	Sweden	D	Stockholm
F 72075 V27	1983	Sweden	C	Stockholm
F 72341 V28	1983	Sweden	D	Stockholm
F 74584 V29	1983	Sweden	C	Stockholm
F 75526 V31	1984	Sweden	C	Stockholm
F 77139 V33	1984	Sweden	D	Stockholm
F 80289 V34	1984	Sweden	C	Stockholm
F 81132 V35	1985	Sweden	D	Stockholm
WV3	1987	Japan	B	A. Yamada, NIH, Tokyo, Japan
WS2	1989	Japan	B	A. Yamada, NIH, Tokyo, Japan
FM93	1993	Japan	B	K. Matsumoto, Fukui Prefectural Institute of Public Health, Fukui, Japan
VS 4098	1993	Sweden	A	Stockholm
MK1	1993	Japan	B	T. Togashi, Sapporo City General Hospital, Sapporo, Japan
YS5	1995	Japan	B	T. Togashi, Sapporo City General Hospital, Sapporo, Japan

the Autoread (Pharmacia) kit on an automated laser fluorescent sequencer (ALF, Pharmacia). The phylogenetic analysis performed on the strains sequenced in this study and on previously sequenced strains confirmed the clustering of mumps virus SH gene sequences into distinct clades (Fig. 1). Five different clusters, named A to E were identified. The five Japanese virus strains isolated between 1987 and 1995 grouped together with the previously described B genotype (Yeo *et al*., 1993), containing all the Japanese reference strains together with MAT, MIY, TAK and URwt (Fig. 1, Table 1). According to previously used nomenclature (Yeo *et al*., 1993; Künkel *et al*., 1995), eleven of the sixteen isolates from Sweden belonged to genotype D, four belonged to genotype C and one was genotype A. The previously described UK lineage, designated C, was noticeably split into two groups when analysis was performed together with the Swedish isolates (Fig. 1). The group containing four Swedish isolates together with the reference isolates BS1 and BF was named C. The other UK group with virus isolates from Scotland containing ED2 (Edinburgh 2), ED4 and ED6 isolates was named E. The branches leading to these two clusters showed bootstrap values of 99 and 100 when 100 trees were sampled.

The deduced amino acid sequences of the coding region of the SH protein gene were aligned for the different strains of mumps virus (Fig. 2). Alignment of the 57 amino acids of the protein revealed that in some positions the amino acids were conserved in all genotypes. Nineteen of the 57 amino acids, at positions 6, 11, 13, 15, 18, 23–25, 27, 31, 32, 36, 43, 45–47, 51, 56 and 57 were conserved in all virus strains. The hydrophobic amino acids proline, tryptophan and leucine were well-conserved in the protein.

Clear signature motifs at positions 28–30 differentiated the genotypes from each other. Amino acid signatures for type A, B, C, D and E were TIL, IIS, VVS, IIL and SLS, respectively (marked in bold text in Fig. 2). In addition, single, specific amino acids in certain positions were characteristic for four of the genotypes. These were: for genotype A, tyrosine, glycine and glutamine at positions 40, 52 and 55, respectively; for genotype B, proline and serine at position 10 and 19, respectively; for genotype C, leucine at position 48; and for genotype E, methionine at position 16. Previous studies by Yeo *et al*. (1993) and Künkel *et al*. (1994, 1995) have also identified certain combinations of amino acids, but the characteristic amino acid triplet was not clearly recognized for all genotypes.

Genotypes A, C and E appeared to be more well-conserved than genotypes B and D. Conservation of amino acids over time was also observed. Only one amino acid differed between the Enders strain of genotype A isolated 50 years ago (Enders *et al*., 1946) and the SBL strain isolated in 1971, and the latter

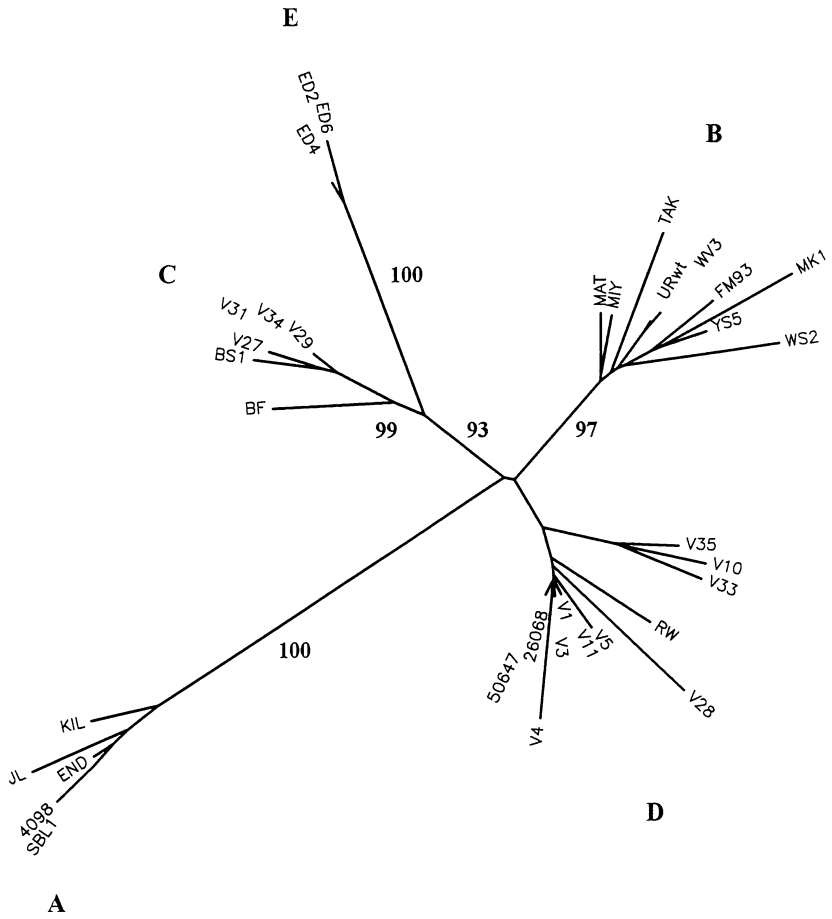

Fig. 1. Phylogenetic tree of mumps virus SH protein sequences from Europe, the USA and Japan. The nucleotide sequence data (300 bp) from 16 Swedish (GenBank accession numbers U50181–U50182 and U50281–U50294) and 5 Japanese (GenBank accession numbers U50295–U50299) isolates were aligned, together with SH sequences of 14 mumps reference strains; SBL1, END, KIL, JL, RW, BF, MAT, MIY, TAK, URwt, ED4, ED2, ED6 and BS1 (GenBank accession numbers X63704–X63709, D90233–D90236 and X63710–X63713, respectively), using the PILEUP program of the GCG Wisconsin Sequence Analysis Package. Some of the reference strains have previously been genotyped by phylogenetic analysis of SH sequences (Yeo *et al*., 1993). Sequence distances were calculated from the alignments using the DNADIST program of the PHYLIP package (version 3.5c) for phylogenetic analysis (Felsenstein, 1988). The NEIGHBOR and DRAWTREE programs of the same package were used for clustering of the distance data and production of the trees. Evaluation of the robustness of the trees was performed by bootstrap analysis (100 trees) using the SEQBOOT and CONSENCE programs. The 16 new Swedish strains were V3, V4, V35, V5, 4098, 50647, V31, V34, V29, 26068, V1, V10, V11, V27, V28 and V33, and the new Japanese strains were FM93, MK1, YS5, WS2 and WV3.

strain was identical in amino acid sequence to VS 4098 diagnosed in 1993. A high degree of similarity was observed between the genotype C strains, which differed in at most three amino acids. The Swedish strains V29, V31 and V34 were identical in amino acid composition among themselves and also with three group C strains isolated in Switzerland and Germany in 1993 (Künkel *et al*., 1994, 1995). Interestingly, the two British strains differed more between themselves than in comparison with each of the Swedish strains of the same genotype. Group B exhibited more pronounced heterogeneity compared to genotype A and C, varying from one to ten amino acids when the different strains were compared. Both marked conservation and also variation was found in this group. For example, the Urabe strain, isolated more than 25 years ago (Yamanishi *et al*., 1970), differed by four amino acids from the recently isolated strains FM93 and YS5, but only by one amino acid from the WV3 strains isolated in 1987. The WS2 strain isolated in 1989 was the most divergent strain compared to the other members of the group.

In genotype D, four of the strains (V1, F26068, V11 and F50647) were identical in amino acid composition and the remainder of the strains varied at one up to ten amino acid positions at the most. The reference strain RW isolated in the late 1970s (McCarthy & Johnson, 1980) differed by two to eight amino acids compared with the Swedish strains.

Different genotypes were found over different time periods in Stockholm. Of the eight virus strains that were isolated in our laboratory in Stockholm during the 1970s all belonged to

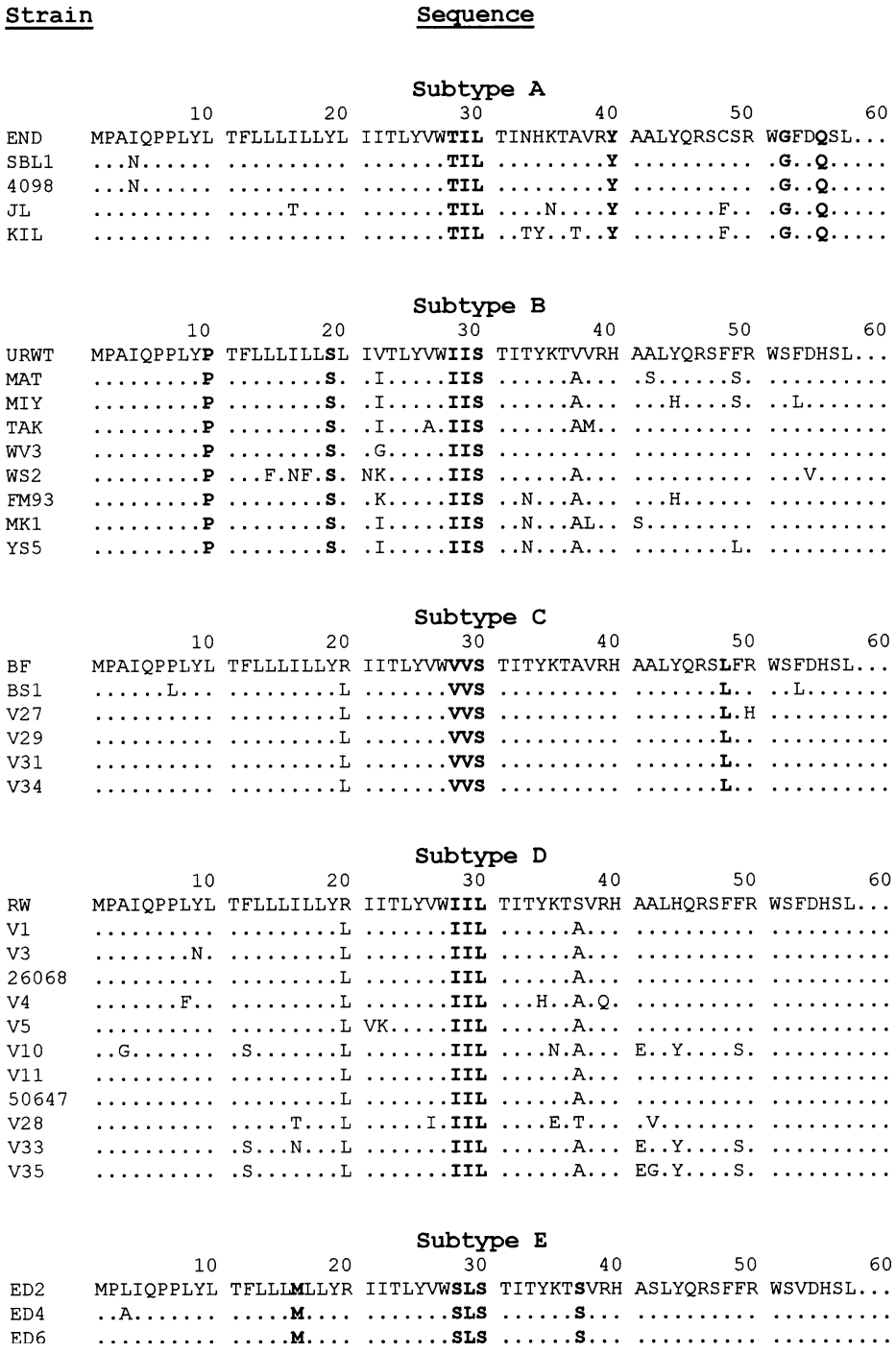

Fig. 2. Alignment of mumps virus SH protein sequence data within the different genotypes. Identified signature amino acids are in bold text. Strain names are as given in the legend to Fig. 1.

genotype D. The seven strains that were isolated about 15 years later belonged both to genotypes C and D. The serum sample collected in 1993 was shown to contain RNA of genotype A.

Recent strains circulating in Europe have been shown to belong to genotypes A, C and E (Yeo *et al.*, 1993; Künkel *et al.*, 1994, 1995). Genotype D has not been found among recently isolated strains in Europe (Yeo *et al.*, 1993; Künkel *et al.*, 1994, 1995). Conversely, the nine genotyped strains from Stockholm in the 1970s and one strain from Berlin in 1977 (Künkel *et al.*, 1995) all belonged to genotype D. These data indicate that there may exist a dynamic flow of different strains waxing and waning in Europe. Co-circulation of different genotypes in a restricted geographical location has also been described for

two subgroups (genotypes) of respiratory syncytial (RS) virus, another member of the paramyxovirus family (Waris, 1990).

In contrast to the situation in Europe, only one genotype was isolated in Japan. Altogether, nine strains originating in Japan over a period of 25 years have been sequenced, and all of them belonged to genotype B. It would be interesting to know if other genotypes exist in Japan.

The SH protein is presumed to be located in the envelope of the virus. The conservation of the protein within different genotypes may be due to the fact that it is not exposed to immunological pressure from antibodies. Both the presumed location of the protein within the envelope of the virus and its hydrophobic character could be factors that prevent such exposure. The biological function of the protein is not known and it has not been identified immunologically. Further biochemical studies may reveal the function of the protein and also answer the question as to why the different genotypes of the SH gene of mumps virus have evolved.

The authors are grateful to their Japanese colleagues Y. Kimura, A. Yamada, K. Matsumoto and T. Togashi for providing the Japanese mumps virus isolates. This study was supported by a grant from the Swedish Medical Research Council no. B93-16X-10389-01A.

References

Brown, E. G., Furesz, J., Dimmock, K., Yarosh, W. & Contreras, G. (1991). Nucleotide sequence analysis of Urabe mumps vaccine strain that caused meningitis in vaccine recipients. *Vaccine* **9**, 840–842.

Elango, N. (1989). Complete nucleotide sequence of the matrix protein mRNA of mumps virus. *Virology* **168**, 426–428.

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.

Elango, N., Kövamees, J., Varsanyi, T. M. & Norrby, E. (1989). mRNA sequences and deduced amino acid sequence of the mumps virus small hydrophobic protein gene. *Journal of Virology* **63**, 1413–1415.

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.

Elliott, G. D., Yeo, R. P., Afzal, M. A., Simpson, E. J. B., Curran, J. A. & Rima, B. K. (1990). Strain-variable editing during transcription of the P gene of mumps virus may lead to the generation of non-structural proteins NS1 (V) and NS2. *Journal of General Virology* **71**, 1555–1560.

Enders, J. F., Levens, J. H., Stokes, J., Jr, Morris, E. P. & Berenberg, W. (1946). Attenuation of virulence with retention of antigenicity of mumps virus after passage in embryonated egg. *Journal of Immunology* **54**, 283–291.

Felsenstein, J. (1988). Phylogenies from molecular sequences: interference and reliability. *Annual Review of Genetics* **22**, 521–565.

Forsey, T., Mawn, J. A., Yates, P. J., Bentley, M. L. & Minor, P. D. (1990). Differentiation of vaccine and wild mumps viruses using the polymerase chain reaction and dideoxynucleotide sequencing. *Journal of General Virology* **71**, 987–990.

Kövamees, J., Rydbeck, R., Örvell, C. & Norrby, E. (1990). Hemagglutinin-neuraminidase (HN) amino acid alterations in neutralization escape mutants of Kilham mumps virus. *Virus Research* **17**, 119–130.

Künkel, U., Schreier, E., Siegl, G. & Schultze, D. (1994). Molecular characterization of mumps virus strains circulating during an epidemic in Eastern Switzerland 1992/1993. *Archives of Virology* **136**, 433–438.

Künkel, U., Driesel, G., Henning, U., Gerike, E., Willers, H. & Schreier, E. (1995). Differentiation of vaccine and wild mumps viruses by polymerase chain reaction and nucleotide sequencing of the SH gene. *Journal of Medical Virology* **45**, 121–126.

McCarthy, M. & Johnson, R. T. (1980). Comparison of the structural polypeptides of five strains of mumps virus. *Journal of General Virology* **46**, 15–27.

Rima, B. K., Alexander, D. J., Billeter, M. A., Collins, P. L., Kingsbury, D. W., Lipkind, M. A., Nagai, Y., Örvell, C., Pringle, C. R. & ter Meulen, V. (1995). *The Paramyxoviridae*. In *Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses*, pp. 268–274. Edited by F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo & M. D. Summers. Vienna & New York: Springer-Verlag.

Takeuchi, K., Tanabayashi, K., Hishiyama, M., Yamada, A. & Sugiura, A. (1989). Cloning and sequencing of the fusion protein gene of mumps virus (Miyahara strain). *Nucleic Acids Research* **17**, 5839.

Takeuchi, K., Tanabayashi, K., Hishiyama, M., Yamada, A. & Sugiura, A. (1991). Variations of nucleotide sequences and transcription of the SH gene among mumps virus strains. *Virology* **181**, 364–366.

Tanabayashi, K., Takeuchi, K., Hishiyama, M., Yamada, A., Tsurudome, M. & Sugiura, A. (1990). Nucleotide sequence of the leader and nucleocapsid protein gene of mumps virus and epitope mapping with the *in vitro* expressed nucleocapsid protein. *Virology* **177**, 124–130.

Waris, M. (1990). Pattern of respiratory syncytial virus epidemics in Finland: two-year cycles with alternating prevalence of groups A and B. *Journal of Infectious Diseases* **163**, 464–469.

Waxham, M. N., Server, A. C., Goodman, H. W. & Wolinsky, J. S. (1987). Cloning and sequencing of the mumps fusion protein gene. *Virology* **159**, 381–388.

Waxham, M. N., Aronowski, J., Server, A. C., Wolinsky, J. S., Smith, J. A. & Goodman, H. W. (1988). Sequence determination of the mumps virus HN gene. *Virology* **164**, 318–325.

Yamada, A., Takeuchi, K., Tanabayashi, K., Hishiyama, M. & Sugiura, A. (1989). Sequence variation of the P gene among mumps virus strains. *Virology* **172**, 374–376.

Yamanishi, K., Takahashi, M., Kurimura, T., Ueda, S., Minekawa, Y., Ogiono, T., Suzuki, N., Baba, K. & Okuno, Y. (1970). Studies on live mumps vaccine. III. Evaluation of newly developed live mumps vaccine. *Biken Journal* **13**, 157–161.

Yeo, R. P., Afzal, M. A., Forsey, T. & Rima, B. K. (1993). Identification of a new mumps virus lineage by nucleotide sequence analysis of the SH gene of ten different strains. *Archives of Virology* **128**, 371–377.

Yun, Z. B., Lindh, G., Weiland, O., Johansson, B. & Sönnerborg, A. (1993). Detection of hepatitis C virus (HCV) RNA by PCR related to HCV antibodies in serum and liver histology in Swedish blood donors. *Journal of Medical Virology* **39**, 57–61.

Received 29 March 1996; Accepted 3 September 1996