Nucleotide sequence of the genes encoding the matrix protein of two wild-type measles virus strains

Knut Baczko,1* Ute Brinckmann,1 Iancu Pardowitz,2 Bert K. Rima3 and Volker ter Meulen1

1Institute of Virology, University of Würzburg, D-8700 Würzburg, Germany, 2Max-Planck-Institute for Experimental Medicine, D-3400 Göttingen, Germany and 3Division of Genetic Engineering, The Queen's University of Belfast, Belfast BT9 7BL, U.K.

The nucleotide sequences of the matrix protein (M) genes of two wild-type measles virus (MV) isolates (JM and CM) have been determined and shown to differ in 56 positions; 31 of these differences are located in the non-coding region and 25 in the coding region of the gene. Most (80%) of the mutations in the coding region are changes to the third base of a codon. A maximum parsimony analysis of the available M gene nucleotide sequences allowed the construction of a tree with at least three lineages or subtypes. One wild-type strain (JM) was very similar to a subacute sclerosing panencephalitis virus strain (case B); the second wild-type strain, CM, showed nucleotide sequence similarity with MV from a case of measles inclusion body encephalitis. Both wild-type virus sequences are distinct from those so far determined for vaccine strains.

Measles virus (MV) is an important human pathogen which can cause acute and, less frequently, persistent infections such as subacute sclerosing panencephalitis (SSPE) and measles inclusion body encephalitis (MIBE). SSPE usually develops 5 to 10 years after acute measles virus infection; in contrast, MIBE has a much shorter incubation period, usually in the order of several months, and has been observed only in immunocompromised patients (Roos et al., 1981; Johnson, 1982).

To date, the pathogenetic mechanisms leading to SSPE and MIBE have not been fully understood. However, studies on MV expression in brain tissue (Baczko et al., 1984, 1986, 1988; Schneider-Schaulies et al., 1989, 1990) and sequence analysis (Cattaneo et al., 1986, 1988a, 1989; Enami et al., 1989) have shown that these infections are characterized by a restriction in viral envelope gene expression owing to increased transcriptional attenuation (Baczko et al., 1986, 1988; Cattaneo et al., 1987a), exclusive synthesis of readthrough transcripts (Baczko et al., 1984; Cattaneo et al., 1986, 1987b) and/or translational defects (Baczko et al., 1984, 1986, 1988), which are produced by mutations forming altered initiation codons (Cattaneo et al., 1986, 1988a) or new stop codons (Cattaneo et al., 1989), as well as by missense mutations that greatly reduce the stability of the matrix protein (M) (Cattaneo et al., 1988b). At least some of the mutations (Cattaneo et al., 1989; Wong et al., 1989) may be due to a cellular dsRNA unwinding activity (Bass & Weintraub, 1988; Bass et al., 1989), which results mainly in transitions from U to C in the positive-strand sequence owing to unequal expression of the underlying enzymic nucleotide changes. This phenomenon has been called 'biased hypermutation' and is most obvious in MIBE viruses, in which about 50% of the U residues in the M mRNA are changed to C (Cattaneo et al., 1988a).

Further interpretation of the significance of nucleotide differences between brain-derived (SSPE and MIBE) and other non-defective, tissue culture-adapted MVs has been hampered by the fact that the latter group are all vaccine or vaccine-related viruses (Cattaneo et al., 1989), and that no data are available from wild-type viruses. This affects the nature of the consensus sequence used and thus it is not clear whether any of the differences reported in SSPE viruses are due to prolonged persistence of the virus in the host or the result of infection by different lineages of MV. Therefore, we have started cloning and sequencing wild-type viruses. Here we report the sequences of two wild-type MV M genes which shed new light on the changes observed in one SSPE case (B) and the MIBE case reported earlier (Cattaneo et al., 1988a, 1989).

Both wild-type MV strains (CM and JM) were isolated from patients with acute measles in the late 1970s in the U.S.A. CM was a gift from Dr B. Fields (Boston, Mass., U.S.A.) and JM from Dr J. Milstein (Bethesda, Md., U.S.A.). CM was propagated first two times in CVI cells and then three times in Vero cells; JM was propagated 13
times in Vero cells. RNA was extracted from infected Vero cells as described previously (Baczko et al., 1984, 1986).

Full length cDNA clones of all the genes except the L gene were cloned according to the method described by Schmid et al. (1987), with the exception that pBluescript II SK+ (Stratagene) was used as a vector. The M genes were sequenced by the dideoxynucleotide chain termination method using the T7 kit from Pharmacia and synthetic oligonucleotide primers internal to the M gene. In addition, the previously unknown 3'-non-coding sequences of the M genes of SSPE viruses B and K (Cattaneo et al., 1986, 1989) were determined from newly isolated cDNA clones. In all viruses, 42 nucleotides in the non-coding region could not be determined owing to a limitation of the cloning procedure. Computer sequence analysis was done using the GCG sequence analysis software package version 6.2 (Devereux et al., 1984). The most parsimonious evolutionary tree for the MV M genes was calculated using a DNA binary tree program based on a graph reduction algorithm run on a VAX 11/730 computer (I. Pardowitz, unpublished results).

Comparison of the M gene sequences of the two wild-type viruses showed that they differ at 56 positions (Fig. 1, Table 1). Similar to most SSPE viruses, many of these differences (31 or 55%) are located in the 3'-non-coding region of the M gene; 20 of the 25 (80%) differences in the coding region are changes to the third base of the codon, which is different to the situation in SSPE viruses (Cattaneo et al., 1989). One other change is silent and consequently only four amino acids are altered (Table 1).

To show the relationship between different SSPE, vaccine and wild-type viruses, the most parsimonious evolutionary tree was calculated (Fig. 1) as previously described (Cattaneo et al., 1989). For the calculation, U to C transitions in the MIBE virus sequence (C) were not taken into account. Compared to the published tree (Cattaneo et al., 1989) this changed the position of SSPE virus K only, because K and C shared a large number of U to C changes. Too few sequences are available to determine whether this is coincidence or whether U residues in certain positions are more prone to hypermutation than others. When U to C transitions in other sequences were not considered, no change of their position in the tree was observed (results not shown).

Three main branches can be distinguished (Fig. 1), which then split into smaller ones. The first branch contains the vaccine viruses, which are very related, differing at a maximum 10 positions. Only the Hallé (L) SSPE virus (Buckland et al., 1990) is in this group. This may indicate that vaccine viruses are only rarely associated with these diseases, a conjecture supported by epidemiological evidence, because after the introduction of MV vaccination the incidence of SSPE was greatly reduced (Halsey et al., 1978, 1980).

The second branch of the tree splits into at least two subclasses. The first contains SSPE virus B and the wild-type strain JM, and the second, rather diffuse, subclass contains SSPE viruses N, I, A and M but no wild-type virus. However, with the accumulation of more data, wild-type viruses may be found that are members of this subclass. The placement of the SSPE K sequence in the first subclass (B and JM) is based on only two common mutations, G to A at position 1131 and C to T at 1139.

### Table 1. Comparison of the differences between related MVs

<table>
<thead>
<tr>
<th>Virus*</th>
<th>Coding region†</th>
<th>Non-coding regions‡</th>
<th>Amino acid differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM/JM</td>
<td>1007/25 (2.5%)</td>
<td>415/31 (7.5%)</td>
<td>4</td>
</tr>
<tr>
<td>CM/C</td>
<td>887/9 (0.9%)</td>
<td>406/8 (2.0%)</td>
<td>–</td>
</tr>
<tr>
<td>JM/B</td>
<td>1007/12 (1.2%)</td>
<td>417/24 (5.8%)</td>
<td>6</td>
</tr>
<tr>
<td>E/Q</td>
<td>1007/6 (0.6%)</td>
<td>417/1 (0.2%)</td>
<td>5</td>
</tr>
</tbody>
</table>

* CM and JM are wild-type viruses; C is the MIBE virus; B is an SSPE virus; E and Q are vaccine-related viruses; K, B, N, I, A and L are SSPE viruses; C is the MIBE virus. CM and JM are wild-type viruses (boxed). The length of a horizontal line is proportional to the number of different nucleotides between the sequence observed and those calculated for common ancestors. The dotted lines indicate U to C transitions.

† Results are expressed as the total number of nucleotides/number of nucleotides different between the two sequences.

‡ The secondary U to C transitions in MIBE strain C virus were not considered.
Whether these two mutations are due to chance or reflect a common ancestry can only be proven by sequencing more wild-type strains.

The third main branch contains wild-type CM virus and the MIBE virus (C), but no SSPE virus. From these results we conclude that different MV types exist and that they cocirculate in the population. A similar analysis will be necessary to distinguish whether these two mutations are due to chance or reflect a common ancestry. A similar approach was taken by comparison of the N gene sequences of these two viruses. For the N gene an even lower mutation rate of 0.2% (K. Baczko, unpublished results) was observed and the two viruses are very similar in the hypervariable region of this gene (Taylor et al., 1991).

Comparison of the coding regions of the M genes of the MIBE virus and CM shows many nucleotide exchanges, but most of these are U to C transitions due to biased hypermutation. After correction for these, there are only nine mutations in the coding region, which gives a mutation rate of about 0.9%. We feel confident that this treatment of the data is correct, because it agrees well with that obtained from the comparison of the N gene sequences of these two viruses. For the N gene an even lower mutation rate of 0.2% (K. Baczko, unpublished results) was observed and the two viruses are very similar in the hypervariable region of this gene (Taylor et al., 1991). This result suggests that a virus from this lineage may have infected the MIBE patient. These observations taken together imply that the biased hypermutation event may be important in determining the destructiveness of the virus in this persistent infection.

The nucleotide sequences of the M genes of wild-type JM virus and SSPE virus B are very similar, especially in the coding region. There are only 12 differences (Table 1), which result in six amino acid changes. This figure is almost as low as that observed between Edmonston and CAM, two MV vaccine strains (Table 1). Therefore, SSPE patient B appears to have been infected with a virus from the JM lineage. The small number of differences between JM and B virus M gene sequences is interesting when one considers that the virus probably persisted in this patient for 5 to 10 years. The apparent lack of mutation probably indicates that despite the rapid mutation of RNA viruses (Holland et al., 1982) the M gene is under selective pressure during most of the period of persistence in some SSPE cases, such as case B. Whether this selection pressure is removed once the virus has entered the brain is not clear. The absence of M from the brain and the ability of RNA extracted from brain to direct the synthesis of M in vitro could indicate that in such cases the alteration of transcription attenuation plays an important role in determining the destructiveness of the virus.

In conclusion, the determination of nucleotide sequences of the M genes of two wild-type MV strains has identified strains which may have been involved in an SSPE and an MIBE case. Biased hypermutation appears to be an important determinant of destructiveness in the case of the MIBE strain. The number of differences between the sequences of an SSPE virus and a wild-type strain of the same lineage is so small that it probably represents a divergence no greater than that observed between virus strains within one group.

We thank Susanne Müller for excellent technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft, the Wellcome Trust, and the Multiple Sclerosis Society of Great Britain and Northern Ireland.

References


(Received 4 March 1991; Accepted 20 May 1991)