Analysis of Nearest Neighbour Base Frequencies in the RNA of a Mammalian Virus: Encephalomyocarditis Virus

(Accepted 21 February 1968)

The technique of frequency analysis of nearest neighbour base sequences in DNA (Josse, Kaiser & Kornberg, 1961) allows the characterization and description of a DNA in terms of the average frequency of occurrence of its sixteen possible doublet sequences. This technique has recently been used to study the patterns of doublet frequencies before and after normalization in the DNA of nine mammalian viruses (Subak-Sharpe, et al. 1966; Morrison et al. 1967; Subak-Sharpe, 1967). It was found that the doublet pattern in the DNA of the four small oncogenic viruses studied resembled the pattern in the host DNA to a remarkable extent, whereas the five large viruses investigated showed virtually no similarity to the host pattern.

The authors reasoned a priori that a cell's translation apparatus, being the result of natural selection, would have a population of transfer RNAs adapted optimally to translate the polypeptide-specifying sequences of the cell DNA. The composition of the population of transfer RNAs was expected to be reflected in the doublet pattern of the cell DNA. Thus, viruses which do not have genetic information for altering the composition of the host transfer RNA population were expected to have doublet patterns closely resembling those of their hosts. Small viruses which carry little genetic information should fall into this group, whereas viruses with genetic information capable of altering the composition of the transfer RNA population should not. Large viruses are most likely to belong to this latter group.

Support for these conjectures derived from the following findings. (1) The four small DNA viruses (SV 40, polyoma, Shope papilloma and human papilloma) had doublet patterns similar to that of host DNA. (2) The five larger DNA viruses (herpes, pseudorabies, equine rhinopneumonitis, vaccinia and adenovirus type 2) had quite different doublet patterns from those of their host DNA. (3) Herpes virus infection alters the composition of the transfer RNA population in the host cell (Subak-Sharpe, Shepherd & Hay, 1966).

The above findings further imply that small mammalian RNA viruses should resemble host DNA in their doublet pattern, taking into account the respective single and double-stranded nature of these genomes.

We report here the first analysis of a mammalian RNA virus, that of the murine encephalomyocarditis virus (EMC). EMC virus is a small mammalian virus with a particle molecular weight of about $10^8$ daltons. The genetic material is single-stranded RNA ($3 \times 10^8$ daltons, Burness, Vizoso & Clothier, 1963), which is sufficient to specify about ten average size polypeptides.

The virus concentrate (kindly provided by Dr A. T. H. Burness) was further purified by treatment with DNase and RNase, followed by centrifugation through a linear sucrose gradient. Viral RNA was isolated using the technique of Montagnier & Sanders (1963), analysed in the ultracentrifuge and found to be homogeneous, with an $S$ value of 34. As a check on the possibility of contamination of the viral RNA with
host cell DNA, the activity of the RNA preparations as templates for *Escherichia coli* DNA polymerase was tested. No detectable synthesis of acid-precipitable material occurred. The presence of RNA did not affect the template activity of calf thymus DNA in the same system.

RNA polymerase (Fraction V) was prepared from *Micrococcus lysodeikticus* using a modification (A. Davidson, 1967, personal communication) of the method of Nakamoto, Fox & Weiss (1964). Preparations were assayed for polynucleotide phosphorylase activity in the $^{32}$P-ADP exchange reaction (Grunberg-Manago, Ortiz & Ochoa, 1956): no activity was detected.

Table 1. Experimentally obtained and normalized values from nearest neighbour frequency analysis of EMC virus RNA

<table>
<thead>
<tr>
<th>Doublet</th>
<th>Experimental</th>
<th>Normalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApU, UpA</td>
<td>560</td>
<td>464</td>
</tr>
<tr>
<td>ApA, UpU</td>
<td>102.8</td>
<td>85.1</td>
</tr>
<tr>
<td>GpU, ApC</td>
<td>59.2</td>
<td>52.5</td>
</tr>
<tr>
<td>UpG, CpA</td>
<td>64.2</td>
<td>72.6</td>
</tr>
<tr>
<td>GpA, UpC</td>
<td>57.5</td>
<td>66.8</td>
</tr>
<tr>
<td>ApG, CpU</td>
<td>68.1</td>
<td>61.3</td>
</tr>
<tr>
<td>GpG, CpC</td>
<td>80.8</td>
<td>56.4</td>
</tr>
<tr>
<td>GpC, CpG</td>
<td>41.2</td>
<td>26.6</td>
</tr>
</tbody>
</table>

**EMC RNA base composition from the above data**

<table>
<thead>
<tr>
<th>Base</th>
<th>Observed %</th>
<th>Expected % (from Faulkner et al. 1961)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>27.9</td>
<td>25.6</td>
</tr>
<tr>
<td>C</td>
<td>21.7</td>
<td>23.6</td>
</tr>
<tr>
<td>U</td>
<td>26.4</td>
<td>27.3</td>
</tr>
<tr>
<td>G</td>
<td>23.9</td>
<td>23.5</td>
</tr>
</tbody>
</table>

'Expected' values were calculated on the assumption that only representative sequences in the complementary strand were synthesized. The symbols used are as follows: A = adenine, C = cytosine, U = uracil, T = thymine, G = guanine and p = phosphate.

$\alpha$-$^{32}$P-Labelled ribonucleoside triphosphates (Schwarz BioResearch) were routinely examined by paper chromatography in the ammonium isobutyrate solvent and purified where necessary.

Fifty $\mu$g. of EMC RNA was incubated at 30° for 10 min. in the presence of 150 units of RNA polymerase and all four ribonucleoside triphosphates (one labelled), using the conditions of Fox et al. (1964). The acid-insoluble fraction was washed (Weiss & Nakamoto, 1961), hydrolysed with KOH and the four mononucleotides were separated by electrophoresis (Sebring & Salzman, 1964). The extent of synthesis encountered was in the range 2 to 4% of the RNA primer.

Table 1 lists the result of the nearest neighbour analysis of EMC RNA and also the respective normalized (Subak-Sharpe et al. 1966) values for each doublet. The analysis was made in duplicate on three occasions (once in the presence of Actinomycin D) and a representative result is given here. The base composition obtained in this analysis is also given and fits well with the composition expected on the basis of chemical analysis (Faulkner et al. 1961).
To compare host and viral doublet patterns, the results obtained for EMC have been superimposed upon those reported for mouse thymus (Swartz, Trautner & Kornberg, 1962) (Fig. 1, top). As a further comparison, the results from a polyoma virus analysis (Subak-Sharpe et al. 1966) have been similarly superimposed (Fig. 1, bottom). Both viruses resemble the host to a considerable degree.

Fig. 1. Nearest neighbour doublet frequencies (top) of EMC virus RNA and (bottom) of polyoma virus DNA (Subak-Sharpe et al. 1966) respectively superimposed on mouse thymus DNA (Swartz et al. 1962). Excess of viral doublet over mouse doublet is indicated by solid top of column: excess of mouse doublet over viral doublet by hatched top of column. * In the EMC RNA, T is replaced by U.

Fig. 2. Doublet frequency patterns of nucleic acids normalized (Subak-Sharpe et al. 1966) to 50% G + C content and expressed as deviation from random expectation (62.5) in doublets/1000. * In EMC RNA, T is replaced by U. Polyoma virus and mouse thymus data are respectively from Subak-Sharpe et al. 1966 and Swartz et al. 1962.

Additionally, the observed doublet frequencies have been normalized (Subak-Sharpe et al. 1966) and plotted as deviations from random expectation. Figure 2 shows the general design for mouse thymus, polyoma and EMC. Again, there is a striking resemblance.

It should be noted that, under conditions of limited synthesis, the pattern for EMC RNA (Fig. 1, 2) represents the doublet pattern in one strand only, the strand complementary to the viral RNA.

The small RNA virus EMC has thus been shown to have a doublet frequency pattern very similar to that of the DNA of its natural host cell. This is in agreement with the stated expectation for a small RNA virus incapable of changing the population of transfer RNAs.
Short communications

In addition we have obtained preliminary evidence (Frearson, Hay & Subak-Sharpe, 1967) that poliomyelitis virus RNA has a doublet pattern very similar to that of EMC. Several other viral RNAs are being studied.

We should like to thank Dr A. T. H. Burness for a gift of EMC virus and Dr M. G. Burdon for the DNA polymerase assays.

Institute of Biochemistry

M.R.C. Unit for Experimental Virus Research

Institute of Virology

University of Glasgow, Scotland

REFERENCES


(Received 12 February 1968)

* At present on leave at the Department of Biochemistry, College of Physicians and Surgeons of Columbia University, New York, New York 10032.