JOURNAL OF VIROLOGY, Jan. 1980, p. 69-77
0022-538X/80/01-0069/09$02.00/0

Vol. 33, No. 1

Comparisons of Nucleotide Sequences in the Genomes of the New Jersey and Indiana Serotypes of Vesicular Stomatitis Virus

DUNCAN J. McGEOCH,* AIDAN DOLAN, AND CRAIG R. PRINGLE

Medical Research Council Virology Unit, Institute of Virology, Glasgow G11 5JR, Scotland

Nucleotide sequences of around 200 residues were determined adjacent to the 3′ terminus of the genome RNA of vesicular stomatitis virus, New Jersey serotype, and adjacent to the 3′-terminal polyadenylic acid tract of the N protein mRNA of the same virus. These sequences were compared with the corresponding sequences previously determined for the Indiana serotype of vesicular stomatitis virus. The sequences obtained for the two strains were readily aligned, showing 70.8% homology overall. Examination of the sequences allowed identification of the translation initiation and termination codons for the N mRNA of each serotype. The deduced N-terminal and C-terminal amino acid sequences of the two N polypeptides were each similar, and most of the differences between them consisted of substitution by a clearly homologous amino acid. It was proposed that these nucleotide sequences, within limits imposed by their functions, comprise reasonably representative measures of the extent of sequence homology between the genomes of the two serotypes, and that this is higher than previously estimated, but with little exact homology over extended regions.

Vesicular stomatitis virus (VSV) specifies five polypeptides: the external glycoprotein G, the matrix protein M, the nucleocapsid protein N (which forms a ribonucleoprotein complex with the genome RNA), and two proteins constituting the RNA-dependent RNA polymerase, L and NS (1, 21). There are two classical types of VSV, designated the Indiana and the New Jersey serotypes (3, 21) (and hereafter referred to as VSV-Ind and VSV-NJ). Antisera to the external G proteins of these viruses do not show heterotypic activity, but antisera directed against internal components (N and M) do exhibit heterotypic activity; that is, these latter antigens contain group-specific sites (3, 10). Apart from serological studies, the proteins of these viruses have also been compared by making fingerprints of their tryptic digests (7). These agree with the serological results in that the maps of the major internal proteins (N and M) are similar for the two serotypes, whereas the maps of the G proteins are quite distinct.

Two types of comparative study of RNA components of the viruses have been published. First, measurements have been made of the extent to which (+)-strand RNA species of one virus anneal to (−)-strand genome RNA of the other (18). By this assay, the sequences present in the genomes of the two viruses are only 4 to 20% homologous. Second, fingerprints of oligonucleotides generated by T_1 RNase digestion of genome RNA also show large differences between the virus types (4). Thus, the RNA studies appear to be in some conflict with the results obtained by comparison of protein species.

We have previously determined the sequence of 200 nucleotides extending from the 3′ terminus of VSV-Ind genome RNA (14). This includes 47 nucleotides adjacent to the 3′ terminus, which are complementary to the (+)-strand leader RNA (5), and 150 nucleotides complementary to the 5′ terminus of N-protein mRNA (19). We have also published the sequence of 205 nucleotides adjacent to the 3′-terminal polyadenylic acid [poly(A)] tract of N mRNA of VSV-Ind (15). Thus, for VSV-Ind, the sequences of the two terminal regions of N mRNA are known. In this paper we describe the determination of the corresponding sequences for VSV-NJ and employ the data in two ways. First, we analyze N mRNA structure and function, and in particular identify the position of translation termination in each mRNA. Second, we use the sequences from each virus type, each of about 400 nucleotides, as paradigms for the whole genomes (about 11,000 nucleotides; 2) to examine the relations between VSV-Ind and VSV-NJ.

MATERIALS AND METHODS

Materials. Avian myeloblastosis virus reverse transcriptase was a gift of J. W. Beard. Rat liver cytosol RNase inhibitor was a gift of G. D. Searle Co. Nucleo-

side triphosphates, deoxy- and 2',3'-dideoxynucleoside triphosphates (dNTP's and ddNTP's, respectively), $p(dT)_{11}$-rA, and $p(dT)_8$-dC were from PL Biochemicals Inc. Radiochemicals were from the Radiochemical Centre, Amersham.

Preparation of RNAs of VSV-NJ. The VSV-NJ strain used was an isolate of the Missouri strain (16). N mRNA of VSV-NJ was synthesized in vitro and purified as described in reference 15. Virion RNA (vRNA) of VSV-NJ was prepared, and a 3'-terminal poly(A) tract was added by following the procedure described in reference 14.

Sequencing with reverse transcriptase and ddNTP's. The 3'-terminal sequence of VSV-NJ vRNA was determined by copying the polyadenylated vRNA with reverse transcriptase from a $p(dT)_{11}$-rA primer in the presence, one at a time, of each ddNTP species, as described for VSV-Ind (14). The sequence of N mRNA of VSV-NJ adjacent to the 3'-terminal poly(A) tract was similarly analyzed using $p(dT)_8$-dC as the primer (15).

Computation of nucleotide sequence homology. Nucleotide sequence homology between corresponding regions of VSV-Ind and VSV-NJ RNAs was estimated as follows. The sequences were aligned to maximize coincidence of sequence, allowing for addition/deletion changes where necessary. The number of positions with identical residues in the two sequences were then scored and expressed as a percentage of the total pertinent sequence length, which includes allowance for apparent deletions in either sequence (that is, each nucleotide residue in an addition/deletion region is scored as a change). Residues whose identity was not resolved in one or other strain were excluded from the estimation.

RESULTS

The 3'-terminal sequence of VSV-NJ vRNA. As will be discussed below, the VSV-NJ sequences show considerable homology with previously determined VSV-Ind sequences. Accordingly, for our present purpose of comparison, we have used the VSV-Ind numbering system previously employed, with the 3'-terminal U residue of vRNA as number 1 (14). The type of result obtained is illustrated in Fig. 1, and the deduced vRNA sequence for residues 1 to 201 is listed in Fig. 2 in alignment with the corresponding VSV-Ind sequence. Residue 15 of VSV-NJ was not unambiguously determined. Residues 1 to 47 of VSV-Ind are complementary to the leader RNA, and the N mRNA complementary sequence starts with residue 51, leaving a spacer sequence AAA (14). In VSV-NJ, this spacer is now seen to be AAAA. At VSV-NJ position 106 the identity of the residue was obscured by a species present in all tracks of the gels, including a control with no ddNTP. At VSV-Ind positions 177 and 178, two residues of unknown identity are thought to occur (unpublished data). Apart from these points the sequencing system yielded unambiguous results for residues 1 to 201. Further data, containing some ambiguities, were obtained to about residue 230.

The 3'-terminal sequence of VSV-NJ N mRNA. The numbering system previously used for the 3'-terminal region of VSV-Ind N mRNA was adopted (15). Results obtained with VSV-NJ N mRNA are shown in Fig. 3, and the deduced sequence of VSV-NJ N mRNA adjacent to the 3'-terminal poly(A) tract is shown in Fig. 4, in alignment with the VSV-Ind sequences (15). Between residues 24 and 44, five addition/deletion changes between the two strains had to be postulated to maintain alignment of homologous regions. Following residue 95 of VSV-NJ, there was a strong compression in the spacing of bands for the next few residues, which prevented determination of residues 96 to 98 by the standard protocol. Since it was considered important to obtain the complete sequence in this region, to help deduce the position of translation termination, various attempts were made to resolve this difficulty. The sequence was eventually obtained by use of a sequencing gel containing an increased urea concentration (8.4 M), run at 1,400 to 1,800 V (0.06 W/cm^2 of gel surface). Additionally the sequence in this region was verified by analysis of a DNA copy of the mRNA in the sequencing system of Maxam and Gilbert (12).

DISCUSSION

In this Discussion the sequence data obtained are examined from several viewpoints. Specific functional aspects of the RNAs are first considered. In this category are included identification of the leader RNA sequences, the nature of the junction between leader and N mRNA, the nature of the 5'-terminal sequences of the N mRNA's, the location of the translation termination codons, and the nature of the 3'-noncoding region of each N mRNA. General sequence comparisons are then made, specifically base compositions, homology between the RNAs, and types of change observed between the sequences of the two strains. The N polypeptide amino acid sequences predicted from the nucleotide sequence data are then described and compared.

Sequences in the 3'-terminal regions of the vRNA's. The sequence of the leader RNA of VSV-NJ [that is, of the (+)-strand transcript of 47 to 48 residues complementary to the 3' terminus of the vRNA] has been published by Colonno and Banerjee (6). When this is compared with the complement of our 3'-terminal vRNA sequence, three single base alterations between the two sequences are found, at residues 21, 31, and 46, and also an apparent deletion of

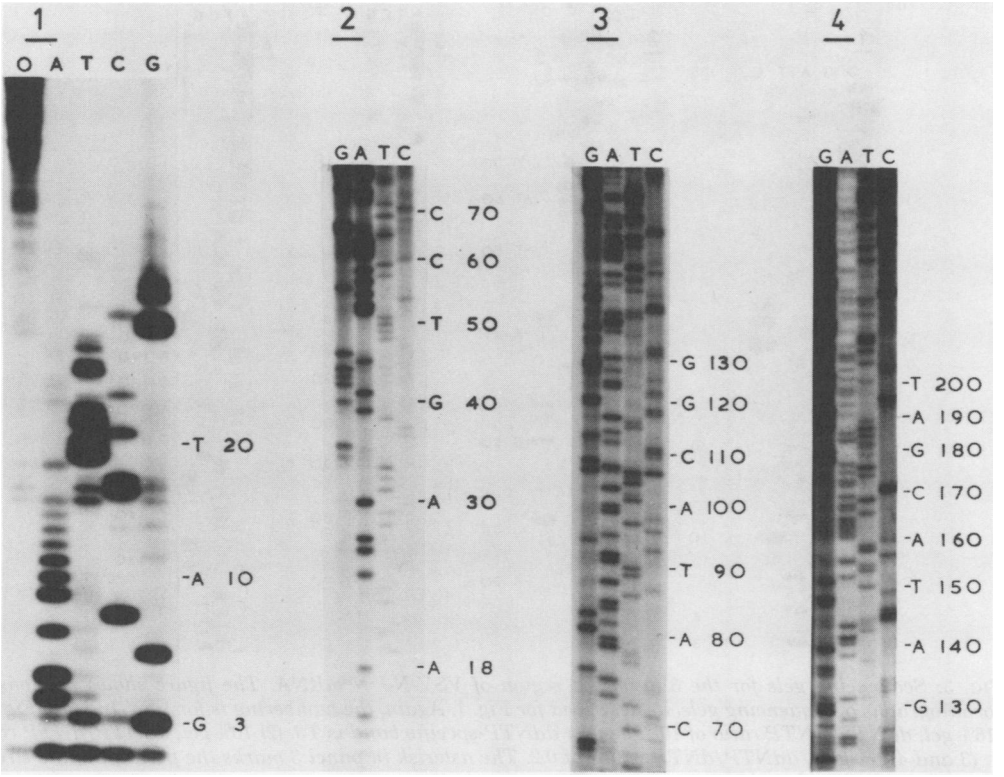

FIG. 1. *Sequencing gels for the 3'-terminal region of VSV-NJ vRNA. The figure shows autoradiographs of portions of sequencing gels. The ddNTP present is indicated at the top of each track. "O" indicates a reaction without any ddNTP. The identity of every tenth residue in the copy DNA is shown, using the VSV-Ind numbering system (see Fig. 2). (1) 16% gel, samples with ddNTP/dNTP ratio of 10. The first ddNTP-specific band is G3. (2) 10% gel, ddNTP/dNTP ratio 0.1. (3 and 4) 6% gels, ddNTP/dNTP ratios 0.1.*

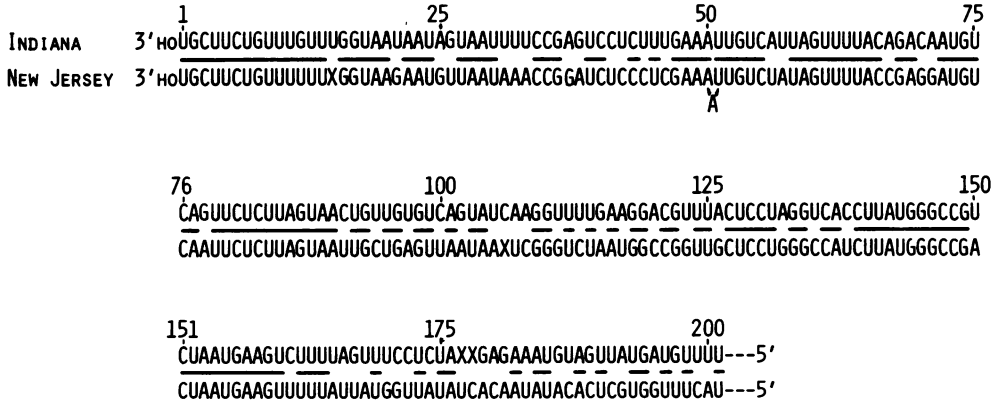

FIG. 2. *Sequences adjacent to the 3' termini of the vRNA's of VSV-Ind and VSV-NJ. Copy DNA sequences were built up from overlapping results of various gel electrophoresis conditions, as illustrated in Fig. 1, and the vRNA sequence was deduced. The sequences are written in the 3' to 5' direction. The numbering is for VSV-Ind. The VSV-Ind sequence is as previously published (14) except that we now consider there are two unidentified residues, at positions 176 and 177 (unpublished data). The VSV-NJ sequence contains an additional A residue, written as lying between residues 50 and 51. Ambiguous residues are written as "X," and regions of identical sequence are denoted by the heavy line between the sequences.*

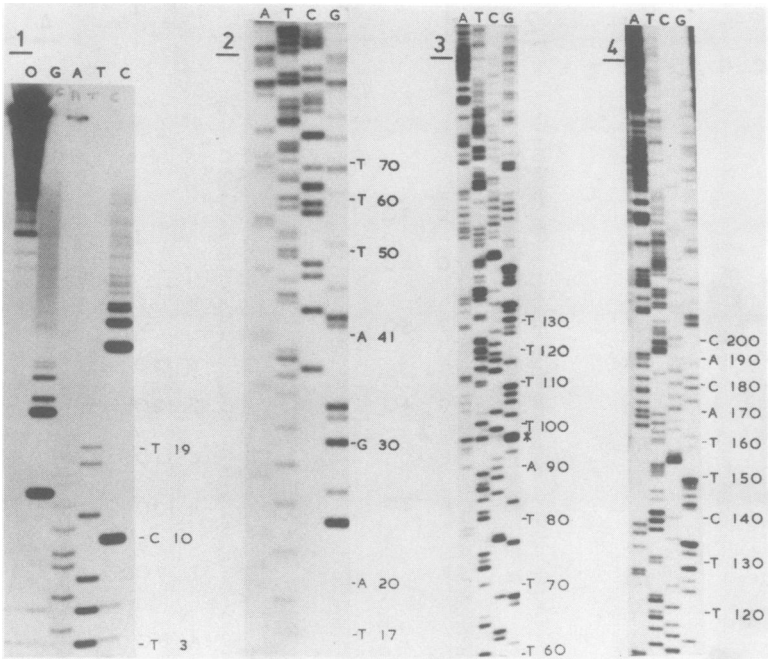

FIG. 3. *Sequencing gels for the 3'-terminal region of VSV-NJ N mRNA. The figure shows portions of autoradiographs of sequencing gels, as described for Fig. 1. Again, the numbering is for VSV-Ind (see Fig. 4). (1) 16% gel, ddNTP/dNTP ratio of 10. The first ddNTP-specific band is T3. (2) 10% gel, ddNTP/dNTP ratio of 1. (3 and 4) 6% gels, ddNTP/dNTP ratios of 0.2. The asterisk in panel 3 marks the position of the strong compression mentioned in the text.*

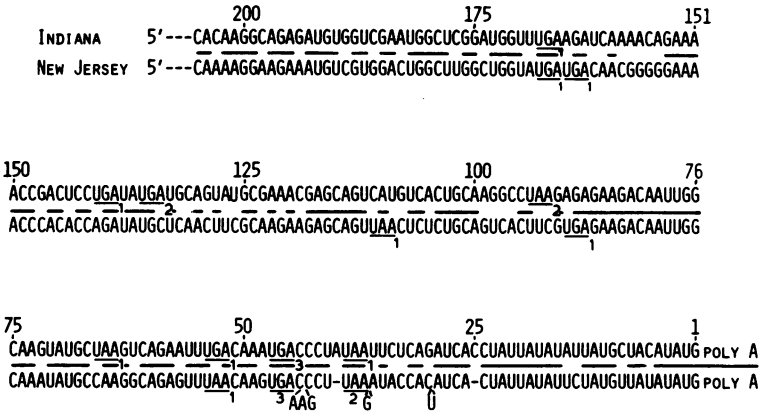

FIG. 4. *Sequences adjacent to the poly(A) tracts of N mRNA's. The sequences are written in the 5' to 3' direction. The VSV-Ind sequence and the numbering are as previously published (15). The VSV-NJ sequence is obtained as illustrated by Fig. 3. In the VSV-NJ sequence a dash is written opposite the position of each extra residue in VSV-Ind. Extra residues in VSV-NJ are written as subscripts. Potential translation termination codons are underlined and their reading frames are numbered, as previously described (15). Regions of identical sequence are indicated by the heavy line between the sequences.*

one A residue in our sequence [(+) strand] at residues 45 or 46. It has been shown that two distinct subtypes of VSV-NJ exist, the genome sequences of which are distinguishable by oligonucleotide fingerprinting (17). The results reported here are for VSV-NJ (Missouri), and the differences between our data and those of Colonno and Banerjee presumably result from their using the more commonly studied strain VSV-NJ (Concan).

As previously described (14), in VSV-Ind vRNA the sequences encoding the leader RNA and those encoding the start of N mRNA are separated by a "spacer," AAA. Examination of the homologous VSV-NJ region (Fig. 2) shows that 3'-UUGUC-5', complementary to the 5'-terminus of N mRNA (8), occurs as expected, but is separated from the leader complement by AAAA in this case. In another intertranscript junction, the N mRNA/NS mRNA junction of VSV-Ind, the spacer is 3'-GA-5' (13). It thus appears from these three examples that the transcription mechanism requires a spacer of a small, but not unique, number of purine residues.

Since gene splicing is thought not to occur in VSV (9), we can deduce the 5'-terminal sequences of the N mRNA's from the 3'-terminal vRNA sequences, as shown in Fig. 5. As previously deduced for VSV-Ind (14, 19), the sequence AUG at residues 64 to 66 must be the translation initiation codon, and in both strains the resulting reading frame is then clear of termination codons for the remainder of determined sequence, unlike both other possible frames (see Fig. 5).

The 3'-terminal sequences of the N mRNA's. For interpreting the sequence data of the 3'-terminal regions of the N mRNA's, it is important to locate the position of translation termination. Since there are no C-terminal amino acid sequences available for either N polypeptide, this must be attempted through examination of the nucleotide sequences. Previously, we decided that for VSV-Ind N mRNA the most likely translation termination codon was UGA at residues 45 to 47 (see Fig. 4), but the arguments available were not conclusive (15). In the 3'-terminal regions of the two N mRNA's now determined, the sequence homology is 71.4%. The two species are therefore closely related, and comparisons between them should help to assign the translation termination codon. We have concluded from such exercises that the previous choice of termination codon, UGA at positions 45 to 47, is correct for both strains. The most important features of the arguments to this end are as follows.

(i) In the reading frame designated number 3 (Fig. 4), both strains show a clear frame from the limit of determined sequence to UGA (residues 45 to 47). We previously argued that the termination codon should lie within 128 residues of the poly(A) tract (15).

(ii) When the sequences are divided into codons in frame 3 from residue 48 to residue 205, and the base changes between the two sequences are classified as changes in each codon position, the results are: first residue in codon, 12 changes; second residue, 6 changes; third residue, 28 changes. Thus, the greatest number of changes is in the third codon position, as would be expected for the correct reading frame with mu-

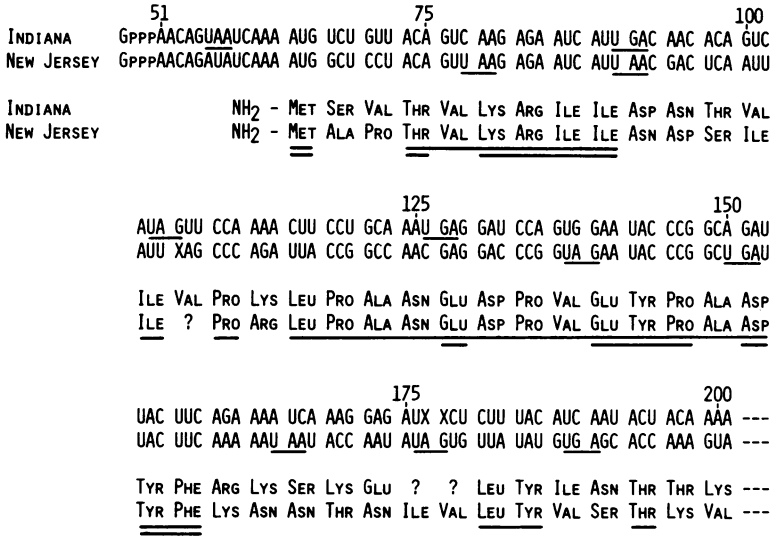

FIG. 5. *5'-terminal sequences of N mRNA's. The 5'-terminal sequences of the N mRNA's are written starting with the known cap structure (8, 19) and as deduced from the vRNA sequences (Fig. 2), with spacing to indicate codons. Numbering is from Fig. 2. Potential termination codons are underlined. The deduced amino acid sequences are written below the mRNA sequences. Codon assignments are from reference 20. Pairs of amino acids identical in the two strains are underlined. Pairs with identical codons are indicated with a double underline.*

tational changes between two closely related mRNA's. Furthermore, if the same exercise is performed for the 5′ regions of the N mRNA's, where the reading frame is identified (see Fig. 5), a similar result is obtained: first residue, 12 changes; second residue, 11 changes; third residue, 23 changes.

(iii) The conservative results of this pattern of codon change are seen when the frame 3 deduced amino acid sequences are considered (Fig. 6). The amino acid residues are identical at 36 of 52 positions. Of these identical pairs, 20 of 36 are coded by nonidentical codons. Of the 16 nonidentical pairs, most are clearly homologous changes (for instance, two occurrences each of Tyr/Phe, Gln/Asn, Glu/Asp, and Asp/Asn changes, and one each of Ser/Ala and Lys/Arg).

(iv) When the same comparisons are made for candidate translation terminations in the other two frames, it is seen that the majority of base changes do not occur in the third positions of predicted codons, that the predicted amino acid sequences show greater differences between the two species than for frame 3, and that the amino acid changes are in general of a less homologous nature. The latter two points are illustrated in Fig. 6.

Thus, we have deduced the reading frame and termination codon for the two 3′-terminal mRNA sequences. Although this is not definitive in the way that availability of C-terminal polypeptide sequence would allow, we do consider that the assignment is strong.

The 3′-terminal noncoding region of heteropolymeric RNA thus comprises residues 1 to 44. The portions immediately distal to the termination codons (20 residues in VSV-Ind, 23 in VSV-NJ) are notable in that they contain four single-residue addition/deletion changes and one such triplet change, in contrast to the other noncoding parts of these mRNA's. This suggests that any sequence-specific function of this region is minimal, and it may act only as a spacer section. This region is followed by an AU-rich tract, residues 11 to 23, which shows symmetry about residue 17 (partial in the case of VSV-NJ). We have previously postulated that this region, together with the remaining sequence to the poly(A), represents part of the control or signaling structure for transcript processing after completion of N mRNA synthesis (13, 15). If this is correct, then the 3′-noncoding region of each of these mRNA's comprises a variable region (of postulated minimal function) followed by transcriptional control regions, followed by the poly(A) tract.

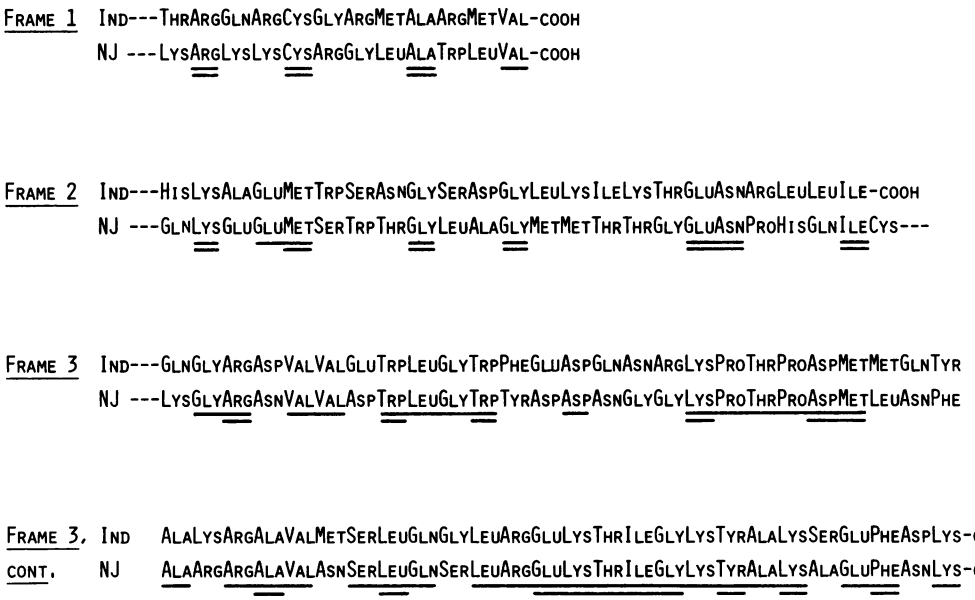

FIG. 6. *Amino acid sequences in each reading frame at the 3′-terminal regions of the mRNA's. For each possible reading frame the amino acid sequences are listed for the following regions of the N mRNA's: Frame 1, from nucleotide residue 204 to the stop codons at 166 to 168; Frame 2, from residue 205 to the stop codon in VSV-Ind at 139 to 141; Frame 3, from residue 203 to the stop codons at 45 to 47. Thus, for Frames 1 and 2 and the first line of Frame 3, amino acid sequences from a given section of each mRNA are in approximate columnar alignment. Pairs of identical acids are underlined, and such pairs with identical codons are marked with a double underline. Codon assignments are from reference 20.*

General comparisons of sequences from the two virus strains. The nucleotide compositions obtained from the sequence data are summarized in Table 1. The values for each (−)-strand RNA region are similar, all having a high U content. With the exception that the 3′-terminal regions of the vRNA's (particularly of VSV-Ind) are richer in U and poorer in C, the compositions obtained by sequencing are also quite close to the nucleotide composition values for vRNA of VSV-Ind, obtained by Bishop et al. (2) by digestion and fractionation of radiolabeled vRNA. Thus, in this sense the sequences obtained are typical of the whole genomes.

As mentioned above, the sequences for corresponding regions of RNA of the two strains can be readily aligned, as indicated in Fig. 2 and 4. The sequence homologies between RNAs of the two strains are: for the 3′-terminal region of the vRNA's, 70.2%; for the 3′-terminal region of the N mRNA's, 71.4%; and for both together, 70.8%. Thus the two strains are closely related in terms of the sequences studied. However, in experiments measuring the extent to which (+) strand of one VSV strain annealed to (−) strand of the other, Repik et al. (18) concluded that VSV-Ind and VSV-NJ possessed only 4 to 20% nucleotide sequence homology. Consideration of this apparent discrepancy has several aspects, as follows.

(i) The two techniques do not measure precisely the same quantity "homology." Calculation of homology from aligned sequences has an effective range of about 25 to 100%, since random alignment of two unrelated sequences will give some scored coincidence of residues, close to 25% for sequences having the nucleotide composition of these VSV RNAs; this is in contrast to the potential range of the annealing assay, essentially 0 to 100%. In further use, the term "homology" is defined as obtained from alignment of sequences, and computed as described in Materials and Methods.

(ii) Examination of the sequences shows that attempted cross-hybridization would result in lack of annealing of the many short homologous regions bounded by changed regions, and in degradation of small annealed fragments to acid-soluble species during "trimming" of the hybrids by RNase, before estimation of annealed material by acid precipitation. Both these effects would result in underestimation of the relatedness of the RNAs.

(iii) The sequences that we have determined comprise about 4% of each genome RNA, and so results with these may not be representative of the entire genomes. We envisage three primary classes of sequence type differing in their stability during evolutionary variation. First, sequences encoding polypeptide directly subject to evolutionary pressure should be the most labile. In the present instance, sequences encoding the antigenic sites of the G proteins alone fall clearly into this category. Second, sequences representing recognition sites and control regions in the RNA, and sequences encoding crucial functional sites in proteins, may be the most stable. Last, between these extremes, there should be regions of RNA whose sequences drift in a neutral, unselected manner; for instance, with changes in the third positions of codons. The sequences that we have determined apparently fall partly into the second of these classes (for instance, the polymerase initiation site presumed to exist at the 3′-terminus of vRNA [11, 14], the ribosome binding site at the 5′-terminus of N mRNA [19], and sites concerned with processing events at intertranscript junctions [13, 14]) and partly into the third class.

The simplest resolution of these considerations is that the regions studied are, within recognized limits imposed by their function, probably reasonably representative measures of the extent of sequence homology between the genomes of VSV-Ind and VSV-NJ, and that this is higher than was previously estimated, but that previous estimates were qualitatively correct in indicating little "exact homology" (18) over extended regions.

In evolving from a common ancestor, the genomes of VSV-Ind and VSV-NJ have clearly undergone several types of change. Single base substitutions of every possible class are represented in the sequences determined [that is, A-U, A-G, A-C, U-G, U-C, and G-C changes, viewed in either the (+) or (−) strand], although any given change observed could be the end result of two or more other changes. Next, as

TABLE 1. *Nucleotide compositions of (−)-strand RNA regions*

RNA	No. of nucleotides	A (%)	U (%)	G (%)	C (%)	Source of data (reference)
3′ terminus, vRNA, VSV-Ind	199	23.1	41.7	20.1	15.1	Sequence (14)
3′ terminus, vRNA, VSV-NJ	200	24.5	37.5	20.5	17.5	Sequence, this paper
Downstream region of N gene, VSV-Ind	205	23.9	35.1	18.5	22.4	Sequence (15)
Downstream region of N gene, VSV-NJ	208	24.0	34.1	19.7	22.1	Sequence, this paper
Total genome RNA, VSV-Ind	~11,000	25.2	34.4	22.4	18.0	RNA digestion (2)

evidenced by the sequences for the 3'-noncoding regions of the N mRNA's (Fig. 4), changes of an addition or deletion nature can occur, although it is not logically possible to resolve which type gave rise to any given observed difference. We do not detect any evidence of large-scale rearrangements between the two types in the regions sequenced.

The amino acid sequences of the N polypeptides. Knowledge of the nucleotide sequence and reading frame of portions of the mRNA's should allow us to deduce the encoded amino acid sequences. This exercise is critically dependent on the nucleotide sequences being correct and, in particular, on the absence of any frame-shifting errors in the analysis. The following features of our nucleotide sequence data gave confidence in deducing amino acid sequences. First, the nucleotide sequences reported here were all determined with apparent unambiguity, and without using the gel sequencing technique near the limit of its resolution. Second, in the predicted amino acid coding regions, the nucleotide sequences from the two serotypes support each other in that there are no frame-shifting differences between the sequences, and the deduced amino acid sequences are readily relatable to each other in terms of the frequency of identical residues and homologous changes (see section "The 3'-terminal sequences of the N mRNA's" above, and below).

The deduced amino acid sequences for the N-terminal and C-terminal regions of the N polypeptides are presented in, respectively, Fig. 5 and Fig. 6. We estimate that together these account for about 23% of the amino acid complement of each N protein. The N-terminal sequences contain identical amino acids at 27 positions out of the 43 unambiguously determined, and at 8 more positions the residues are obviously homologous (two instances each of Asp/Asn, Val/Ile, and Lys/Arg changes, and one each of Ser/Ala and Ser/Thr changes). The C-terminal sequences are even more similar, with 36 identical residues out of the 52 unambiguously determined, and 10 obviously homologous residues (listed above). Clearly we would expect the two proteins to give similar peptide maps, as observed by Doel and Brown (7) (although our predicted sequences are not long enough to make worthwhile predictions of peptide species).

In both strains, 18 amino acid species are represented in the sequences, with Cys and His absent. Both C-terminal sequences have Lys as the most abundant amino acid, but are not in toto strongly basic, and the N-terminal regions contain a near equality of acidic and basic groups. The N-terminal regions also show a higher content of hydrophobic amino acids and a striking Pro content.

The work presented here comprises a fresh way of looking at relations between VSV strains. Previous studies have analyzed aspects of genome structure, such as ability to form hybrids between RNAs of two strains (18), oligonucleotide fingerprints (4), and fingerprints of peptides (7). All of these present aspects of the structure of the whole, or a large part, of the genome, but in ways which are either of low resolution or are not amenable to numerical treatment. In contrast, our sequence data represent a high-resolution analysis of a small portion of the genome. It should be possible to extend this type of analysis with relative ease (particularly for sequences adjacent to the 3' termini of genome RNAs) to examining relations between other members of the rhabdovirus family.

LITERATURE CITED

1. **Banerjee, A. K., G. Abraham, and R. J. Colonno.** 1977. Vesicular stomatitis virus: mode of transcription. J. Gen. Virol. **34**:1-8.
2. **Bishop, D. H. L., H. G. Aaslestad, H. F. Clark, A. Flamand, J. F. Obijeski, P. Repik, and P. Roy.** 1975. Evidence for the sequence homology and genome size of rhabdovirus RNAs, p. 259-292. *In* B. W. J. Mahy and R. D. Barry (ed.), Negative strand viruses, vol. 1. Academic Press, London.
3. **Cartwright, B., and F. Brown.** 1972. Serological relationships between different strains of vesicular stomatitis virus. J. Gen. Virol. **16**:391-398.
4. **Clewley, J. P., D. H. L. Bishop, C.-Y. Kang, J. Coffin, W. M. Schnitzlein, M. E. Reichmann, and R. E. Shope.** 1977. Oligonucleotide fingerprints of RNA species obtained from rhabdoviruses belonging to the vesicular stomatitis virus subgroup. J. Virol. **23**:152-166.
5. **Colonno, R. J., and A. K. Banerjee.** 1978. Complete nucleotide sequence of the leader RNA synthesized in vitro by vesicular stomatitis virus. Cell **15**:93-101.
6. **Colonno, R. J., and A. K. Banerjee.** 1978. Nucleotide sequence of the leader RNA of the New Jersey serotype of vesicular stomatitis virus. Nucleic Acids Res. **5**:4165-4176.
7. **Doel, T. R., and F. Brown.** 1978. Tryptic peptide analysis of the structural proteins of vesicular stomatitis virus. J. Gen. Virol. **38**:351-361.
8. **Franze-Fernandez, M. T., and A. K. Banerjee.** 1978. In vitro RNA transcription by the New Jersey serotype of vesicular stomatitis virus. I. Characterization of the mRNA species. J. Virol. **26**:179-187.
9. **Freeman, G. J., D. D. Rao, and A. S. Huang.** 1979. RNA synthesis of vesicular stomatitis virus. VIII. Oligonucleotides of the structural genes and mRNA. Gene **5**:141-157.
10. **Kang, C.-Y., and L. Prevec.** 1970. Proteins of vesicular stomatitis virus. II. Immunological comparisons of viral antigens. J. Virol. **6**:20-27.
11. **Keene, J. D., M. Schubert, R. A. Lazzarini, and M. Rosenberg.** 1978. Nucleotide sequence homology at the 3' termini of RNA from vesicular stomatitis virus and its defective interfering particles. Proc. Natl. Acad. Sci. U.S.A. **75**:3225-3229.
12. **Maxam, A. M., and W. Gilbert.** 1977. A new method for

sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.

13. **McGeoch, D. J.** 1979. Structure of the gene N: gene NS intercistronic junction in the genome of vesicular stomatitis virus. Cell **17**:673–681.
14. **McGeoch, D. J., and A. Dolan.** 1979. Sequence of 200 nucleotides at the 3'-terminus of the genome RNA of vesicular stomatitis virus. Nucleic Acids Res. **6**:3199–3211.
15. **McGeoch, D. J., and N. T. Turnbull.** 1978. Analysis of the 3'-terminal nucleotide sequence of vesicular stomatitis virus N protein mRNA. Nucleic Acids Res. **5**: 4007–4024.
16. **Pringle, C. R., I. B. Duncan, and M. Stevenson.** 1971. Isolation and characterization of temperature sensitive mutants of vesicular stomatitis virus, New Jersey serotype. J. Virol. **8**:836–841.
17. **Reichmann, M. E., W. M. Schnitzlein, D. H. L. Bishop, R. A. Lazzarini, S. T. Beatrice, and R. R. Wagner.** 1978. Classification of the New Jersey serotype of vesicular stomatitis virus into two subtypes. J. Virol. **25**:446–449.
18. **Repik, P., A. Flamand, H. F. Clark, J. F. Obijeski, P. Roy, and D. H. L. Bishop.** 1974. Detection of homologous RNA sequences among six rhabdovirus genomes. J. Virol. **13**:250–252.
19. **Rose, J. K.** 1978. Complete sequences of the ribosome recognition sites in vesicular stomatitis virus mRNAs: recognition by the 40S and 80S complexes. Cell **14**:345–353.
20. **Sadgopal, A.** 1968. The genetic code after the excitement. Adv. Genet. **14**:325–404.
21. **Wagner, R. R.** 1975. Reproduction of rhabdoviruses, p. 1–92. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. 4. Plenum Press, New York.