



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

# 2 *Applications of Oligonucleotide Fingerprinting to the Identification of Viruses*

Olen M. Kew, Baldev K. Nottay, and  
John F. Obijeski

I. Introduction.....	41
II. Principles and Mechanisms of Fingerprint Analysis.....	43
A. Overview of Fingerprint Analysis.....	43
B. Mechanisms of Two-Dimensional Separation of Oligonucleotides.....	44
C. Sensitivity of Fingerprinting in Detecting Genetic Relatedness among RNA Viruses.....	48
D. Fingerprinting Compared with Other Methods of Virus Identification.....	49
III. Experimental Procedures.....	49
A. Preparation of Radiolabeled RNA.....	49
B. Digestion of RNA.....	54
C. Fingerprinting Methods.....	55
D. Autoradiography.....	63
E. Analysis of Isolated Oligonucleotides.....	63
F. Isolation of Individual RNA Molecules for Fingerprinting....	66
G. Oligonucleotide Mapping.....	67
H. Genome Complexity Analysis.....	71
I. Oligonucleotide Mapping in Genome Sequence Analysis.....	72
IV. Applications.....	72
A. Fingerprinting as a Means of Studying Large Virus Populations.....	72
B. Analysis of Natural Virus Isolates.....	73
C. Other Applications Relating to Virus Identification.....	78
V. Summary and Conclusions.....	79
References.....	80

## *I. Introduction*

Fingerprinting is a technique by which oligonucleotides, produced by cleavage of RNA molecules with specific ribonucleases, are separated in two dimensions. Originally developed as a preparatory method, fingerprinting

is currently most frequently applied as an analytical method for comparing the genomes of RNA viruses. Comparisons are based on the principle that the large, structurally unique oligonucleotides separate into patterns, or "fingerprints," that are highly characteristic of the RNA sequence from which they derive. Accordingly, the sequence, and thus genetic, relationships can be readily assessed. Since the characteristic oligonucleotides originate from all regions of the RNA molecule, the distributions of sequence similarities and differences may be surveyed over the entire viral genome. Although not providing specific base sequence information, fingerprinting does offer a simple and comparatively rapid means of sequence comparison, with the results obtained in a pictorial form.

The applications of fingerprint analysis in virology are very broad. A few examples of the applications include analysis of RNA genome size and complexities, identification of subgenomic RNA molecules, analysis of the messenger RNA species produced by both DNA and RNA viruses, characterization of defective-interfering viral genomes, biochemical description of genetic reassortants and recombinants, and the estimation of the degree of sequence divergence between closely related virus strains.

Oligonucleotide fingerprinting has emerged as an important tool for the identification of RNA virus strains (Brown, 1982; Clewley and Bishop, 1982). Fingerprinting provides a reliable means of following the natural distribution and transmission of a specific viral genotype over time. This capability is of considerable value in human and veterinary epidemiology, in which genetically distinct but antigenically similar epidemic and zoonotic strains in coetaneous circulation can be distinguished, and their transmission separately followed. Because viral antigens tend to be more conserved than the overall genomic RNA sequence, fingerprinting is generally much more sensitive than serology in detecting divergence between strains. On the other hand, fingerprint analysis is not complicated by antigenic drift. Although the availability of monoclonal antibodies promises to greatly improve the precision and discrimination of viral serodiagnosis (Yewdell and Gerhard, 1981), fingerprinting is, in a sense, a more open technique, not requiring the production of new reagents in order to monitor the circulation of a given virus genotype. Further, unlike serology, fingerprint comparisons are not restricted to the small set of genomic sequences specifying the amino acid residues on a polypeptide surface.

In this chapter, we describe the principles and methodologies of oligonucleotide fingerprinting and briefly review the applications of this technique to the identification of RNA viruses. Particular emphasis will be placed on the use of this method to follow the distribution and transmission of viral agents of infectious disease. Several excellent reviews on oligonucleotide fingerprinting have appeared. Those of Beemon (1978) and Wang (1978) concentrate on the retrovirus group. Clewley and Bishop (1982) also

describe methodology and the broad applications of fingerprinting in virology. Pedersen and Haseltine (1980b) and De Wachter and Fiers (1982) describe some current fingerprinting methods.

## II. Principles and Mechanisms of Fingerprint Analysis

### A. OVERVIEW OF FINGERPRINT ANALYSIS

Oligonucleotides for fingerprinting studies are produced by complete digestion of viral RNA with the specific ribonucleases (RNases) T1 and A. RNase T1, from the fungus *Aspergillus*, specifically cleaves phosphodiester bonds adjacent to guanosine (G) residues, producing guanosine 3'-phosphate (Gp) and a set of oligonucleotides of varying length terminating with Gp. Because of its high specificity under moderate digestion conditions, RNase T1 is the enzyme of choice for routine fingerprint studies. The other known single base-specific RNases also cleave at G (Takahashi and Moore, 1982) and thus confer no significant advantage over RNase T1. Pancreatic RNase A, which preferentially cleaves at pyrimidine residues, is infrequently used in routine comparative studies because of the reduced complexity of the fingerprint patterns of its cleavage products. RNase A was used in conjunction with RNase T1 in classical studies of RNA structure (Barrell, 1971), and has been applied more recently in detailed fingerprint comparisons between strains (Nomoto *et al.*, 1981) as well as in studies on total genome sequences (Kitamura and Wimmer, 1980; Kitamura *et al.*, 1981). Many other RNases with different base specificities are known. Some are used in RNA sequence analysis to produce specific partial fragments of end-labeled RNA molecules (Simoncsits *et al.*, 1977; Donis-Keller, 1980; Donis-Keller *et al.*, 1977). However, fingerprinting requires complete digestion of the RNA, conditions for which the specificities of the other known RNases are too broad for them to be applicable.

The number of RNase T1 oligonucleotides in a complete digest is one more than the number of G residues in the RNA. For a random polynucleotide, the size distribution of oligonucleotide fragments is predicted from its composition by the relationship (Beemon, 1978)

$$F_n = (G)^2(1-G)^{n-1}$$

where  $F_n$  is the fraction of oligonucleotides of chain length  $n$ , and  $G$  is the mole fraction of guanosine residues in the RNA. For an RNA of 8000 bases containing 25% G, this relationship predicts approximately 500 Gp monomers ( $n=1$ ), 160 5-mers, 38 10-mers, 9 15-mers, 2 20-mers, and 6 oligonucleotides for which  $n>20$ , to occur among the 2001 oligonucleotides produced by complete digestion. Oligonucleotides with chain lengths  $n \geq 10$

have a high statistical probability of containing unique sequences that occur only once in a haploid viral genome. For the hypothetical 8000-base viral genome described in the example above, oligonucleotides containing 10 or more base residues constitute approximately 25% of the total sequences. Although the structurally unique oligonucleotides of a typical RNA molecule are too numerous to be resolved effectively by a single mechanism of separation, the two-dimensional fingerprinting technique, in which a different separation mechanism is employed for each dimension, allows good resolution to be achieved in practice. The proportion of the genome that can be resolved experimentally into separate oligonucleotide spots is primarily dependent upon the size and complexity of the RNA. This is because, for polynucleotides of a given guanosine content, the number of unique oligonucleotides increases with chain length, but the mean oligonucleotide chain length remains constant (Aaronson *et al.*, 1982). For large RNA molecules, such as viral genomes, an important fraction of the potentially informational oligonucleotides is incompletely resolved, and is thus effectively unavailable for analysis. For nonsegmented genomes, from 5 to 15% of the unique oligonucleotides can be resolved and individually analyzed. With viruses having segmented genomes it is possible to examine higher percentages of the total genome by separate analysis of each segment.

#### B. MECHANISMS OF TWO-DIMENSIONAL SEPARATION OF OLIGONUCLEOTIDES

Two different fingerprinting methods have been developed. The earlier method, developed by Brownlee and Sanger (1969), utilized high-voltage electrophoresis in the first dimension and chromatography in the second. In the second method, that of De Wachter and Fiers (1972), separation is achieved in both dimensions by electrophoresis through polyacrylamide gels. In both methods, oligonucleotides are separated in the first dimension according to their net charges and in the second dimension primarily according to chain length.

First-dimension electrophoresis for both is generally performed near pH 3.5 in the presence of 6–7 *M* urea. At this pH, the greatest charge differences among the bases occur (Markham and Smith, 1952), such that the net negative charges, and thus electrophoretic mobilities, of the nucleoside monophosphates increase in the order Cp < Ap < Gp < Up. In the absence of significant frictional forces, oligonucleotides migrate according to the total charge contributions of their constituent base residues. The combination of acidic pH and high urea concentrations disrupts the secondary structure of the oligonucleotides and reduces interchain aggregation, largely

excluding effects of conformation on mobilities. Further, for the polyacrylamide gel system, the use of large-pore gels eliminates molecular sieving effects on all but the largest oligonucleotides (Frisby *et al.*, 1976), thus maximizing the contributions of charge to electrophoretic mobilities.

Second-dimension separation in the two fingerprinting systems occurs by fundamentally different mechanisms. The method of Brownlee and Sanger (1969) utilizes homochromatography, in which oligonucleotides are competitively displaced from a cationic surface during ascending chromatography. The displacement rate, and thus migration rate, is approximately inversely related to chain length, but base composition and secondary structure significantly influence mobilities. De Wachter and Fiers (1972) employed a highly cross-linked 20% polyacrylamide gel buffered at pH 8 ("neutral gel"). In the neutral pH range, the four bases are largely uncharged. All significant charges on the oligonucleotides therefore derive from ionization at the phosphodiester bonds (and at the terminal phosphate), and the charge:mass ratios for all long oligonucleotides are nearly constant. Thus the polyacrylamide gel acts strictly as a molecular sieve permitting fractionation according to size and shape. To a first approximation, electrophoretic mobilities in the second dimension decrease with the logarithm of the oligonucleotide chain length (De Wachter and Fiers, 1982). Deviations from this relationship, arising from contributions of secondary structure and base composition, are usually small (De Wachter and Fiers, 1982; Lee and Wimmer, 1976; Pedersen and Haseltine, 1980a).

A typical two-dimensional polyacrylamide gel fingerprint of a virus RNA is shown in Fig. 1. In this example, the smaller oligonucleotides have been run off the second-dimension gel in order to improve resolution. The oligonucleotides are arranged in a graticulated pattern formed by a series of bands curving diagonally away from the quadrant nearest the origin of first-dimension electrophoresis (lower left, Fig. 1). This pattern is most evident in the region of the more numerous shorter-chain oligonucleotides. The bands form according to uridine (U) content. Oligonucleotides of the band closest to the origin contain no U residues, those of the next band contain one U, those of the next, two, and so on. Within each band, the cytidine (C)-rich oligonucleotides are diagonally opposite (above and behind in the orientation of the Fig. 1 pattern) the adenosine (A)-rich oligonucleotides of equivalent chain length (Pedersen and Haseltine, 1980a; De Wachter and Fiers, 1982). The lower part of each fingerprint contains the large, RNase T1-resistant oligonucleotides, whose mobilities are diagnostic for a virus genotype. Fingerprints of related and unrelated polioviruses are shown in Fig. 2. The diagnostic oligonucleotides are distributed in a roughly triangular region, a pattern that occurs because with increasing chain length oligonucleotides generally approach the mean base composition, and hence first-dimension mobility, of the entire RNA molecule.

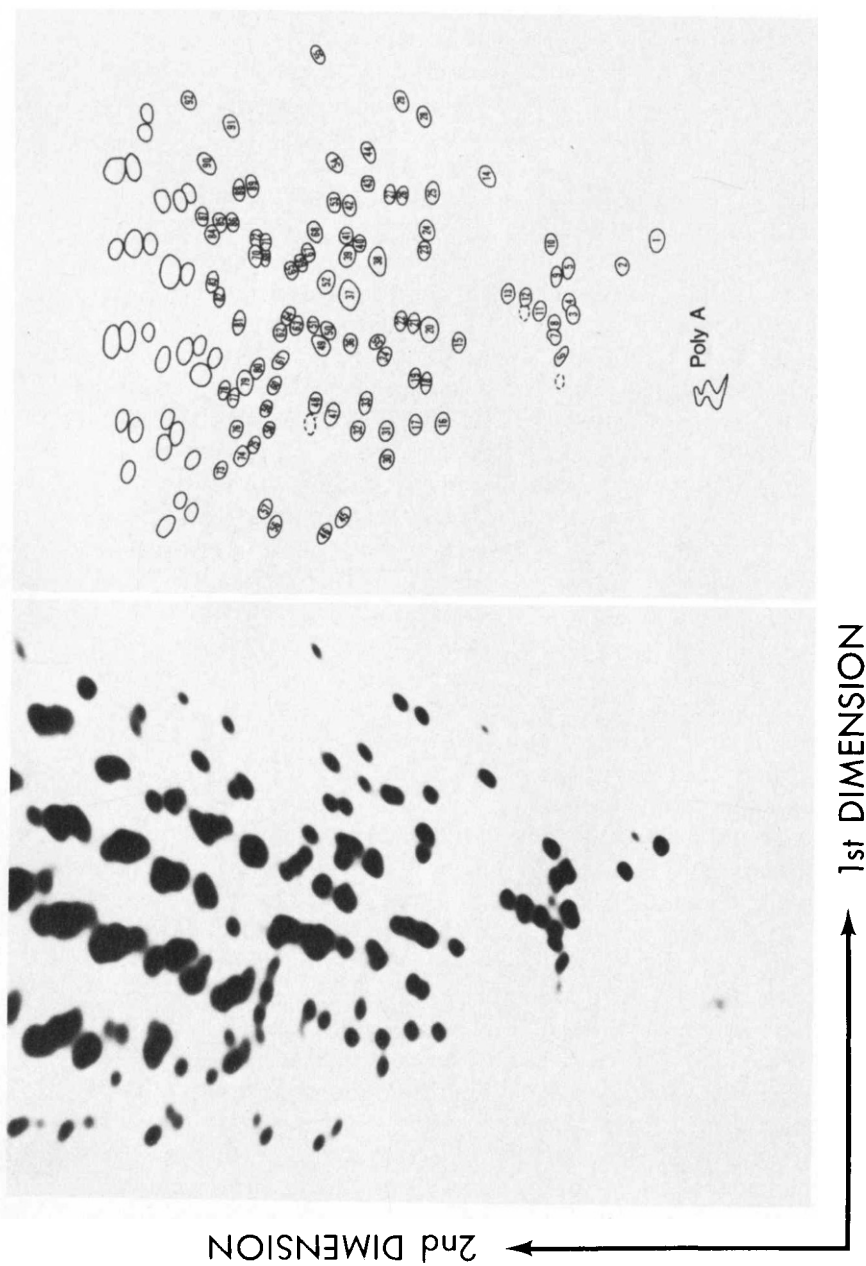


FIG. 1. (Left) Fingerprint of the Leon strain of type 3 poliovirus. The upper half of the fingerprint is not shown. For this fingerprint, second-dimension electrophoresis was terminated when the bromphenol blue dye migrated 30 cm. (Right) Tracing of the fingerprint. Oligonucleotide spots are numbered arbitrarily. Identification of the less well resolved spots is based upon replicate fingerprints and oligonucleotide mapping studies (Fig. 5). Some spots, for example 37 and 52, may contain more than one oligonucleotide. Oligonucleotides that could not be accurately mapped because their spots did not give reproducible intensities are indicated by dashed lines.

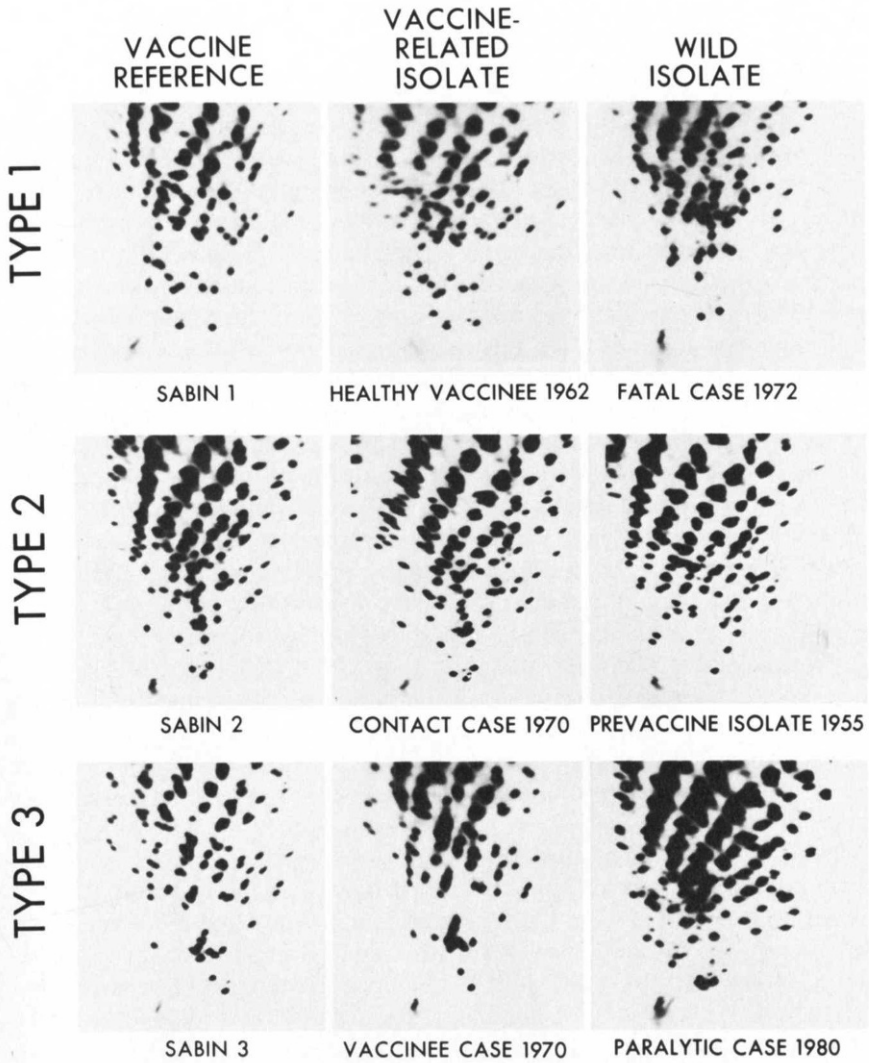


FIG. 2. Fingerprint comparisons of closely and distantly related poliovirus strains. Note that the vaccine-related isolates (center column) have fingerprints that are very similar to those of the respective oral vaccine reference strains (left column). Wild isolates from each of the three serotypes have fingerprints (right column) that are entirely distinct from those of the vaccine-related group, indicating a lack of close genetic relatedness. Note also the strong similarities in the fingerprints of Sabin type 3 (Leon 12 a,b) and its nonattenuated parent, Leon (Fig. 1).



### C. SENSITIVITY OF FINGERPRINTING IN DETECTING GENETIC RELATEDNESS AMONG RNA VIRUSES

Oligonucleotide fingerprinting is a tool of extraordinary power and precision for demonstrating close genetic relationships among viruses and, therefore, for determining the natural distribution and transmission of any given viral genotype. The typical fingerprint contains 30–60 diagnostic spots, each having an independent mobility and, as a result, they represent many independent points of reference for comparison. Because the potential number of patterns formed by the spots of the diagnostic oligonucleotides is very large, equivalent migration of a majority of the spots of two viral RNA samples implies a close genetic relationship between the viruses. Frequently fingerprints of different isolates are compared by simple visual inspection, and similarities in patterns can be perceived when as few as 50% of the diagnostic oligonucleotides spots are shared. The simplest method for confirming genetic relatedness is codigestion and co-electrophoresis of equivalent isotopic amounts [as measured in disintegrations per minute (dpm)] from the two viral RNAs. This approach appears to be sufficient when 50% or more of the oligonucleotide spots comigrate. When fewer spots are shared, additional evidence, such as composition or sequence analysis of corresponding spots, should be obtained to establish genetic relatedness clearly. Moreover, when less than about 25% of the oligonucleotide spots of two different strains comigrate, the remaining pattern similarities are usually difficult to detect, and further comparisons may not be performed unless genetic relatedness is suspected on other grounds, such as epidemiological data or the existence of strains with intermediate fingerprint patterns.

The number of oligonucleotide spots common to two RNA molecules decreases rapidly with differences in their overall sequence homology. For example, Aaronson *et al.* (1982) used computer-simulated mutation of the genome segment encoding the influenza hemagglutinin to examine the relationship between the number of large oligonucleotides shared and the extent of overall base sequence homology. According to their model, two RNA molecules that diverge at 1, 5, or 10% of their total sequences share 85, 50, and 25%, respectively, of their diagnostic oligonucleotides. These estimates are in good agreement with experimental comparisons of foot-and-mouth disease virus genomes by fingerprinting and competition hybridization (Robson *et al.*, 1979). From these studies it is clear that the lower threshold for recognition of relatedness by fingerprinting is near the level of 90% base sequence homology. Thus, the evolutionary range of fingerprinting is quite short.

The basis for the pronounced sensitivity of oligonucleotide fingerprinting to genetic divergence is evident in view of the mechanisms of fractionation.

Consider a typical diagnostic oligonucleotide of 15 nucleotide residues. A change in only one base (7% of the total) involving C, A, or U residues will result in an altered first-dimension mobility, and the spot will appear to have shifted laterally in the fingerprint of the mutant RNA. The greatest mobility shifts are expected from C→U transitions, the smallest from C→A transversions. Changes involving G residues will either eliminate or create a cleavage site for RNase T1, and spots may appear to move both laterally and vertically within the diagnostic region, or disappear from it altogether. Similarly, new diagnostic spots may arise from other regions of the genome upon mutation of a G residue.

#### D. FINGERPRINTING COMPARED WITH OTHER METHODS OF VIRUS IDENTIFICATION

With the exception of genome sequencing, other approaches for measuring genetic relatedness, both molecular and serological, are less sensitive than fingerprinting in detecting small mutational changes, but, instead are able to detect more distant relationships among viruses. For example, competition nucleic acid hybridization can detect homologies between strains that share only 25% of their overall sequences (Minson and Darby, 1982). Restriction endonuclease analysis of DNA genomes may be used for quantitative estimates of sequence divergence as great as 20–30% (Brown *et al.*, 1979). Serological methods, especially those involving polyclonal antibodies directed against internal or nonstructural viral protein antigens, are often able to detect conserved homologies in polypeptide structure that are beyond the limits of sensitivity of standard nucleic acid hybridization methods. Genome sequencing has the ultimate sensitivity to detecting mutational differences among viruses, and routine comparative sequence analysis of selected genome regions is possible. However, unlike fingerprinting, current sequencing methods are not well suited for providing a survey of the total sequence divergence among many strains.

### *III. Experimental Procedures*

#### A. PREPARATION OF RADIOLABELED RNA

##### 1. *In Vivo Labeling*

The genomic RNAs of most animal RNA viruses can be labeled to activities suitable for production of high-quality fingerprints by biosynthetic incorporation of ortho[<sup>32</sup>P]phosphate. The general strategies for isotopic labeling of viral macromolecules have been reviewed by Henry (1967). Con-

ditions for optimal incorporation of radiophosphate vary for different virus and cell culture systems. Important factors to be considered first when developing a labeling strategy are duration of the replicative cycle and the multiplicity of infection of input virus.

For viruses having short infectious cycles and for which high-titer inocula are available, infection occurs as a one-step cycle. Labeling periods may precede or coincide with infection and extend throughout the cycle, or be adjusted to correspond to an optimal interval within the cycle. Efficiencies of incorporation may be improved by maintenance of the cells in phosphate-deficient medium for several hours prior to infection, thus depleting intracellular pools of inorganic phosphate. Because exogenous phosphate equilibrates slowly with intracellular nucleotide pools (Henry, 1967), label should be added well before the expected period of maximal RNA synthesis. Chilling (0–4°C, 60–120 min) infected suspension cultures immediately prior to addition of  $^{32}\text{P}$  has been used to improve incorporation of label into poliovirus RNA (Wimmer, 1972). For high-specific-activity labeling, ortho[ $^{32}\text{P}$ ]phosphate is added without carrier to a synthetic medium lacking phosphate. If serum is to be included in the medium, it should be extensively dialyzed against 0.15 M NaCl. The volumes of the labeling medium should be kept to a minimum, and frequent agitation of the medium is recommended for uniform distribution of isotope. The increased cell densities ( $1\text{--}5 \times 10^6$  cells/ml medium) generally used during radiolabeling may necessitate inclusion of supplemental buffers, such as 25 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES) or 0.3% sodium citrate, pH 7.2 (Ogra *et al.*, 1968). These buffers are occasionally inhibitory, and their effects on total yield should be determined before routine use.

For systems in which virus production occurs over an extended period or that require multiple rounds of infections, optimal labeling conditions should balance total virus yield against the specific activity of the final RNA product. Cell viability is prolonged by inclusion of some carrier phosphate in the culture medium, and multiple harvests may be required to assure recovery of a high proportion of intact virions. In all systems, virus should be harvested and purified as early as possible to minimize radiation damage from unincorporated isotope.

Typical inputs of ortho[ $^{32}\text{P}$ ]phosphate are 100–1000  $\mu\text{Ci}$ /ml culture. The total isotope input should be adjusted according to expected yield of label incorporated into intact viral RNA, which may vary even among closely related virus strains, from nearly 1% to less than 0.01%. In some systems, inclusion of actinomycin D (1–5  $\mu\text{g}$ /ml) in the labeling medium may increase incorporation of  $^{32}\text{P}$  into virus RNA by selectively inhibiting host transcription. However, the effects of this antibiotic on labeling efficiencies should be examined carefully before routine use, especially when many dif-

ferent agents are to be compared. For example, some strains of poliovirus, a group generally regarded as actinomycin D resistant, incorporate  $^{32}\text{P}$  poorly in the presence of this antibiotic (Nomoto *et al.*, 1979a; Nottay *et al.*, 1981).

The highest quality fingerprints are obtained with RNA labeled *in vivo* and extracted from purified virions. Some rapid diagnostic fingerprinting procedures utilize labeled intracellular viral RNA (La Torre *et al.*, 1982). RNA labeled *in vivo* is frequently described as "uniformly labeled," implying that each phosphodiester bond is labeled to an equivalent specific activity. Uniformity of label is demonstrated by equivalence of base composition whether the RNA is degraded to 3'-nucleotides (by RNase T2 digestion) or to 5'-nucleotides (by digestion with nuclease P1 or snake venom exonuclease). Although demonstrated in some systems (Billeter *et al.*, 1974), deviations from uniformity have been observed in others, complicating determinations of oligonucleotide base composition (Lee *et al.*, 1979). With long-chain oligonucleotides, the specific activity differences among residues tend to average out, so that the radioactivity in a large oligonucleotide labeled *in vivo* is in close proportion to its chain length.

## 2. *In Vitro* Labeling

*In vitro* labeling of viral RNA presents several advantages over labeling *in vivo*. For some viruses, *in vitro* methods offer the only effective means of labeling genomic RNA to high specific activities (Frisby, 1977). Moreover, *in vitro* isotopic labeling can be performed rapidly and requires substantially lower inputs of radioisotope. Only small quantities (1  $\mu\text{g}$  or less) of chemically pure viral RNA are needed for analysis. Repeated *in vivo* labelings are avoided, as it is generally necessary to prepare a viral RNA sample only once, with small amounts being withdrawn for labeling whenever needed. Finally, 5'- and 3'-radiophosphate end-labeled oligonucleotides can be recovered from the fingerprints for sequence analysis.

Three basic approaches are applied for *in vitro* labeling of RNA for fingerprinting.

*a. 5'-End Labeling.* The most widely used *in vitro* labeling procedure utilizes T4 polynucleotide kinase (PNK) to transfer radiophosphate from  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$  to the free 5'-hydroxyl groups of RNase-generated oligonucleotides (Frisby, 1977). Alkaline phosphatase is added to the RNase T1 digestion mixture to remove all 3'-terminal phosphates from the oligonucleotides. The combined activities of alkaline phosphatase and PNK catalyze a net transfer of phosphate from the 3'- to the 5'-ends, a process that does not significantly alter the electrophoretic mobilities of the treated oligonucleotides (Brownlee and Sanger, 1967). Also, because some PNK preparations contain a contaminating 3'-phosphatase activity (Szekely and

Sanger, 1969), prior removal of the 3'-terminal phosphates eliminates potential charge and migration heterogeneity brought about by partial dephosphorylation of an oligonucleotide. Several approaches have been used to eliminate alkaline phosphatase activity from the subsequent PNK reaction. These include removal by phenol extraction (Frisby, 1977), inactivation with 0.1 *N* HNO<sub>3</sub> (Lee and Fowlks, 1982), or inhibition by inclusion of inorganic phosphate in the PNK reaction (Chaconas and van de Sande, 1980). The last approach, included in the labeling method of Pedersen and Haseltine (1980b), is operationally very simple. We have used their method, described below, to obtain high-quality fingerprints of viral RNA.

Dried RNA (200–500 ng) in a 0.5-ml polypropylene tube (Eppendorf) is resuspended in 1  $\mu$ l of sterile water to which is added 1  $\mu$ l of 40 mM Tris-HCl (pH 8.0) containing for each microgram of RNA 0.2 unit of RNase T1 (Sankyo, distributed by Calbiochem) and  $2 \times 10^{-4}$  units of *Escherichia coli* alkaline phosphatase (BAP) (Worthington BAPF; P/L Biochemicals). It is important to omit EDTA, normally present during RNase T1 digestion of prelabeled RNA, because it is a strong inhibitor of BAP. For less than 100 ng of RNA, addition of 100 ng of polyguanylic acid (Miles Biochemicals) to the digests facilitates cleavage by RNase T1 (Pedersen and Haseltine, 1980b). Incubation is for 30 min at 37°C. The PNK reaction immediately follows by addition of 50  $\mu$ l of PNK reaction mixture and 5 units of PNK (P/L Biochemicals; New England Nuclear). The PNK reaction mixture is prepared by first lyophilizing 50–200  $\mu$ Ci [ $\gamma$ -<sup>32</sup>P]ATP in a 0.5-ml Eppendorf tube, then dissolving the labeled ATP (>2000 Ci/mmol) in 50  $\mu$ l of 10 mM K<sub>2</sub>HPO<sub>4</sub>–K<sub>3</sub>PO<sub>4</sub> (pH 9.5), 10 mM Mg(OAc)<sub>2</sub>, 5 mM dithiothreitol. Incubation is at 37°C for 1–16 hr. Because the rates of phosphorylation by PNK vary widely for different oligonucleotides (Frisby, 1977), prolonged incubation favors more uniform incorporation of label into the oligonucleotides. However, with increased incubation, nuclease contamination of the PNK preparations may become evident, and the optimal conditions may involve a compromise between maximizing uniform incorporation and minimizing the appearance of secondary spots. The reaction is terminated by addition of 50  $\mu$ l of 0.6 *M* NH<sub>4</sub>OAc, and 100  $\mu$ g of carrier yeast tRNA. The oligonucleotides and carrier are precipitated by mixing with 300  $\mu$ l of cold (-20°C) 95% ethanol and chilling in a dry ice-ethanol bath for 20 min. The precipitate is collected by centrifugation in a microcentrifuge (15 min at 4°C), washed with 400  $\mu$ l of cold 95% ethanol, and centrifuged again (5 min at 4°C). After lyophilization, the labeled digest is ready to be resuspended in first dimension electrophoresis sample buffer.

*b. 3'-End Labeling.* Use of the base-specific chemical cleavage method of Peattie (1979) for sequence determination of oligonucleotides requires radiolabeling of the 3'-ends. Currently, this is most frequently accom-

plished by use of the T4 RNA ligase-mediated transfer of cytidine 3'-[5'-<sup>32</sup>P]diphosphate (pCp) to the 3'-terminal hydroxyls of dephosphorylated oligonucleotides (England and Uhlenbeck, 1978). Lee and Fowlks (1982) have combined this method with fingerprinting to produce 3'-labeled oligonucleotides by a simple procedure. Only an outline of the procedures is given here, as detailed descriptions are presented in the original articles. Digestion with RNase T1 and BAP is performed essentially as described above in the subsection on 5'-end labeling. Afterward, BAP is inactivated by making the digest mixture 100 mM in HNO<sub>3</sub> (pH 2; 10 min at room temperature), followed by addition of ZnSO<sub>4</sub> to a final concentration of 1 mM (2 min at room temperature) to inactivate RNase T1. The digest is precipitated with ethanol, washed with 95% ethanol, and lyophilized. The precipitate is resuspended in the ligase reaction system (30 μl) containing 50 mM HEPES (pH 8.3), 5 mM ATP, 10 mM MgCl<sub>2</sub>, 3.3 mM dithiothreitol, 10% (v/v) dimethyl sulfoxide, 15% (v/v) glycerol, 200 μCi [5'-<sup>32</sup>P]pCp, and 6 units of T4 RNA ligase (P/L Biochemicals). Incubation (4°C, 16 hr) is terminated by ethanol precipitation of the labeled oligonucleotides as described for 5'-end labeling.

*c. Radioiodination.* Commerford (1971, 1980) described a method for the radioiodination of cytosine residues of RNA in the presence of thallium ions that is applicable to oligonucleotide fingerprinting (Robertson *et al.*, 1980; Clewley and Bishop, 1982; Lee and Fowlks, 1982). Radiolabeling with <sup>125</sup>I has the advantages of being very rapid, producing labeled RNA with high specific activities and having a longer half-life (60 days) than <sup>32</sup>P-labeled RNA (14.3 days). The factors governing the efficiencies of iodine labeling are discussed in detail elsewhere (Huang and Pagano, 1977; Commerford, 1980). We have used the method of Lee and Fowlks (1982), described below, to produce fingerprints of high quality.

From 1 to 20 μg of viral RNA in water (5 μl in a 1.5-ml Eppendorf tube) is combined with 12.5 μl of 0.3 mM thallic nitrate buffer: 30 μl of 0.1 M thallic nitrate [Tl(NO)<sub>3</sub>; K and K; Alfa] in 1 N HNO<sub>3</sub> mixed with 9.97 ml of 0.2 M NaOAc (pH 4.4). Carrier-free <sup>125</sup>I (0.3–1.5 mCi, pH 7–11) is added, and the total volume is brought to 25 μl with water. After 4 min of incubation at 70°C, 475 μl of sodium dodecyl sulfate (SDS) buffer [10 mM Tris-HCl (pH 7.5), 0.12 M NaCl, 0.1% SDS, 1 mM EDTA] is added, and the mixture is incubated for an additional 10 min at 70°C. Then 1 ml of cold 95% ethanol is added, and the mixture is chilled in dry ice-ethanol for 20 min, then centrifuged (30 min, 4°C). The precipitate is washed four or five times by gently filling the tube with cold 80% ethanol and decanting. The final supernatants are monitored with a Geiger counter; after confirmation that no soluble label remains, the pellet is lyophilized. The labeled RNA is ready for digestion with RNase T1.

*d. Comments.* Fingerprints produced from oligonucleotides labeled *in*

*vitro* generally correspond well with fingerprints of the same RNA labeled *in vivo*. However, depending on the method of labeling, small differences may exist. As mentioned, spots of 5'- or 3'-end-labeled oligonucleotides may vary in intensity as a result of preferential activities of PNK and RNA ligase (Frisby, 1977; Lee and Fowlks, 1982). Moreover, the total number of spots may differ among fingerprints labeled *in vivo* and *in vitro*. Spots of 5'-labeled material, for example, may be missing if PNK phosphorylation is very inefficient or blocked by the presence of a 5'-terminal cap structure (Banerjee, 1980) or a protein (Wimmer, 1982). Spots may also be missing from fingerprints of 3'-labeled material (Lee and Fowlks, 1982), possibly as a result of incomplete dephosphorylation. Secondary oligonucleotide fragments, produced by trace nucleases, may also be labeled, thus giving rise to additional spots (Szekely and Sanger, 1969).

Fingerprints of iodinated RNA may also be deficient in certain spots. Because radioiodination of RNA occurs primarily by the production of 5-iodocytosine, 5-iodouracil being a minor product (Commerford, 1971), only cytosine-containing oligonucleotides are effectively labeled. Although the larger diagnostic oligonucleotides generally contain some cytosine, the spots of cytosine-deficient oligonucleotides may be light. The poly(A) streak observed in the fingerprints of *in vivo*-labeled RNA of many positive stranded viruses is absent from fingerprints of iodinated RNA.

Fingerprints produced with *in vitro*-labeled material are generally satisfactory for comparisons of the genomic RNAs of different viruses. Only when the fingerprints of the most closely related strains are being compared does the potential existence of artifactual spots present uncertainties in estimating the extent of sequence divergence.

## B. DIGESTION OF RNA

Digestion of *in vivo*-labeled RNA or iodinated RNA is performed in 1.5-ml Eppendorf tubes. Viral RNA plus 100  $\mu$ g of carrier yeast tRNA are dissolved in 1 ml of 20 mM Tris-HCl (pH 7.8), 150 mM NaCl, transferred to a 1.5-ml Eppendorf tube, and mixed with 200  $\mu$ l of 1% (w/v) cetyltrimethylammonium bromide (CTAB; obtained from Sigma or BDH). The CTAB precipitation step purifies the RNA of uncharged polymeric contaminants and nucleases that are precipitated by ethanol (Ralph and Bergquist, 1967). After 20 min on ice, the insoluble cetyltrimethylammonium ribonucleate salts are collected by centrifugation in a microcentrifuge (5 min at 4°C). The flocculent pellet is washed twice with cold 70% ethanol/30% 150 mM NaCl in water, to convert the RNA back to the sodium salt, then once with 70% ethanol-water, and finally with 95% ethanol. Each wash step is followed by 2 min centrifugation at 4°C. The RNA is dried under vacuum,

and the pellet is resuspended in 20  $\mu$ l of 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, containing 0.2 unit of RNase T1 per  $\mu$ g RNA (enzyme: RNA mass ratio of 1:25). Samples of double-stranded RNA are denatured by boiling for 2 min and quick chilling (Walker *et al.*, 1980) or freezing (Sugiyama *et al.*, 1982) before addition of enzyme. Digestion is at 37°C for 30 min. To 20  $\mu$ l of digest is added 40  $\mu$ l of urea-dye mix (9 M urea/50% glycerol saturated with bromphenol blue/0.3% trypan red).

Digestion of RNA with RNase A (Calbiochem) is at 37°C for 30 min in 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.3 M NaCl, at an enzyme:RNA mass ratio of 1:10. Increased ionic strength suppresses secondary cleavage by RNase A at adenosine residues. However, the salt will impair resolution in the first dimension and should be removed by ethanol precipitation prior to electrophoresis (Nomoto *et al.*, 1981).

Stocks of the RNase T1 and RNase A are stored as small aliquots in their respective digestion buffers at -20°C at concentrations of 1 mg/ml.

## C. FINGERPRINTING METHODS

### 1. Two-Dimensional Gel Electrophoresis

Many variations of the basic technique of De Wachter and Fiers (1972) have been described (Frisby *et al.*, 1976; Lee and Wimmer, 1976; Kennedy, 1976; Clewley *et al.*, 1977b; Pedersen and Haseltine, 1980a,b; Stewart and Crouch, 1981). Our fingerprinting procedures essentially follow the modifications described by Lee and Wimmer (1976) and Lee *et al.* (1979).

*a. Reagents.* All stock solutions (Tables I and II) should be prepared with deionized water using chemical reagents of the highest grades available. Especially critical to the production of high quality fingerprints are the chemical purities of the acrylamide *N,N'*-methylene bisacrylamide (Bis) and urea components.

Highly purified, electrophoresis-grade acrylamide and Bis from commercial sources (e.g., Bio-Rad Laboratories; Eastman Organic Chemicals) can be used without further purification. Reagent-grade acrylamide can be purified by recrystallization in chloroform, and Bis by recrystallization in acetone, as described by Adesnik (1971). Alternatively, monomers can be mixed in the specified ratios (Tables I and II) and purified by stirring (12 hr, room temperature) with mixed-bed resin (20 g/liter; Bio-Rad AG 501-X8), followed by filtration through Whatman No. 1 paper. Purified solutions are passed through a Whatman GF/C filter and stored in the dark at 4°C.

Solutions of ultrapure urea (Research Plus; BDH AnalAR) require only filtration (0.22  $\mu$ m) before use. Reagent grade urea is purified by stirring



TABLE I  
STOCK SOLUTIONS FOR FIRST-DIMENSION GEL SYSTEM

Component	Volume required for one gel
Acrylamide: Bis (40:0.5) <sup>a</sup>	18.0 ml
Urea, 9M	60.0 ml
Deionized H <sub>2</sub> O	8.5 ml
Saturated citric acid	0.7-1.0 ml
FeSO <sub>4</sub> •7H <sub>2</sub> O [0.16% (w/v)]	0.5 ml
Ascorbic acid [1% (w/v)]	2.0 ml
H <sub>2</sub> O <sub>2</sub> (30%)	75.0 μl
Electrode buffer: 25 mM citric acid adjusted to pH 3.3 with 10 N NaOH	
Sample buffer: 20 μl of RNA digest mixed with 40 μl of 9 M urea-50% glycerol, saturated with bromphenol blue-0.3% trypan red	

<sup>a</sup> Acrylamide (400 g/liter) and *N,N'*-methylene bisacrylamide (5 g/liter).

(12 hr) at 9 M solution with mixed-bed resin (20 g/liter). The urea solution is further purified by filtration through a bed of silicic acid (Lee and Wimmer, 1976), prepared by suspending about 100 g of powdered silicic acid (Sigma SIL-R) in 1 liter of deionized H<sub>2</sub>O, and filtering the suspension under suction through a 15-cm filter of Whatman No. 4 paper. Solutions of purified urea may be stored for up to 3 months at 4°C.

Saturated citric acid is stored at room temperature.

TABLE II  
STOCK SOLUTIONS FOR SECOND-DIMENSION GEL SYSTEM

Component	Volume required for one gel	
	Connecting gel	Separating gel
Acrylamide:Bis (23.2:1.55) <sup>a</sup>	—	209.0 ml
Acrylamide:Bis (40:0.5) <sup>b</sup>	2.5 ml	—
Tris-borate, 1 M, adjusted to pH 8.2 <sup>c</sup>	1.0 ml	11.0 ml
Deionized H <sub>2</sub> O	6.4 ml	—
TEMED	25 μl	0.1 ml
Ammonium persulfate (10%)	125 μl	0.4 ml
Electrode buffer: 50 mM Tris-borate (pH 8.2)		

<sup>a</sup> Acrylamide (232) g/liter) and *N,N'*-methylene bisacrylamide (15.5 g/liter)

<sup>b</sup> From Table I.

<sup>c</sup> Tris base (121 g/liter) and boric acid (63 g/liter).

Ferrous sulfate [0.16% (w/v), in 1-ml aliquots] and ascorbic acid [1% (w/v), in 2-ml aliquots] are filtered (0.22  $\mu\text{m}$ ) and stored at  $-20^{\circ}\text{C}$ .

Hydrogen peroxide (30%) is dispensed into 1-ml volumes and stored at  $4^{\circ}\text{C}$ .

Electrophoresis grade *N,N,N',N'*-tetramethylethylenediamine (TEMED; Bio-Rad Laboratories) is stored at  $4^{\circ}\text{C}$ .

Solutions of ammonium persulfate (10%) are prepared fresh before use.

Solutions of Tris-borate (1.0 *M*; 121 g/liter Tris base and 63 g/liter boric acid adjusted to pH 8.2) are autoclaved and stored at room temperature.

*b. Apparatus.* First-dimension plates (12  $\times$  45 cm) are cut from 1/8-in. glass. Second-dimension plates are made from 3/16 in. glass cut to 36  $\times$  43 cm. Spacers are 1/16 in. (0.16 cm)  $\times$  1 cm strips of polyvinyl chloride, cut to lengths of 45 cm for the first-dimension assembly and 43 cm for the second-dimension plates. The first-dimension well former is 1/16 in. Teflon with two or three teeth, 1.25 cm wide  $\times$  4 cm long, separated by 1.25 cm. Plates should be cleaned immediately after use with a mild nonabrasive detergent, rinsed thoroughly with deionized water, and dried. Just before use, plates are washed with 95% ethanol and air dried. Cleaning acid should be avoided, because the glass surface may be altered, causing the first-dimension gels to adhere so tightly to the glass that removal and transfer are prevented.

The plates and spacers are assembled to form a mold for the gel, and the assembly is clamped with medium steel binder clips. Precautions must be taken to prevent leakage of the unpolymerized gel from the mold. We prefer to seal the sides and bottom with 1-in. Teflon tape (3M catalog No. 5490) taking special care to reinforce the corners, where leaks are most prone to occur. Alternatively, spacers may be coated with a thin film of silicone vacuum grease before assembly. Chamber bottoms may also be sealed by use of a temporary spacer (Clewley *et al.*, 1977b) or with plasticine (De Wachter and Fiers, 1972). Problems of leakage of the first-dimension gels are reduced by setting the chamber at a nearly horizontal angle during polymerization. To assure that the first-dimension gel strip is in firm contact with the glass plates during polymerization of the second-dimension gel, the bottom of the assembly is clamped with several steel binder clips, and air pockets are removed by hydrostatic pressure. Good contact can be achieved also by using slightly thinner spacers for the second-dimension chamber than those used for the first dimension (Pedersen and Haseltine, 1980a), or by using spacers of compressible silicone rubber (Clewley *et al.*, 1977b).

The electrophoresis units used in our laboratory, purchased from Lee's Unique Instruments (De Soto, TX), are shown in Fig. 3. The slab gel chambers are placed in the lower buffer reservoirs and connected to the upper reservoirs with filter paper wicks of Whatman 3 MM or equivalent, precut

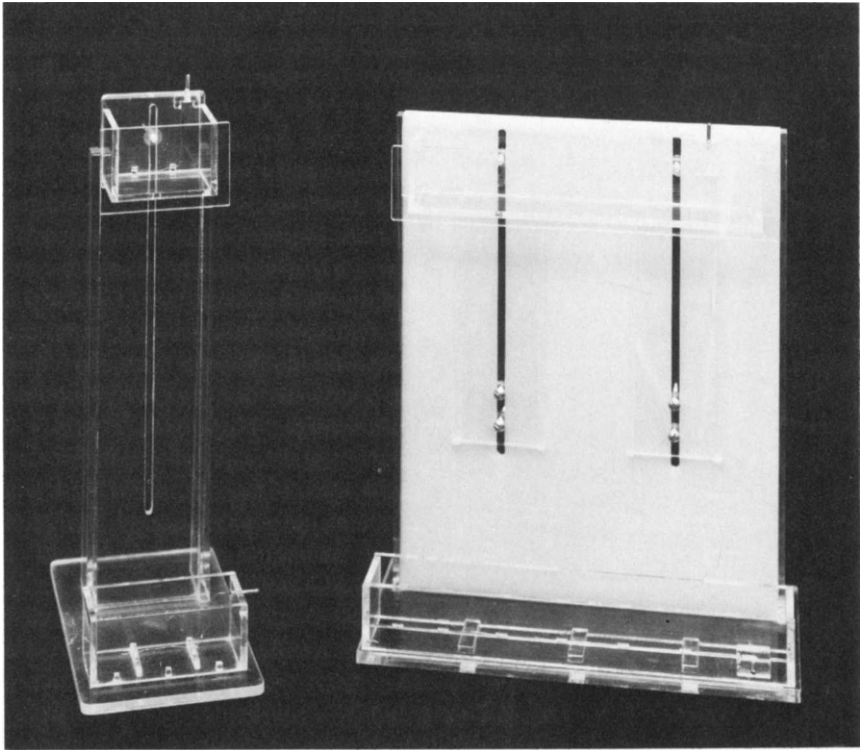


FIG. 3. The electrophoresis buffer chambers used in our laboratory for fingerprint analysis: first-dimension apparatus (left), and second-dimension apparatus (right). The units are constructed of Plexiglas and can be purchased from Lee's Unique Instruments (De Soto, Texas).

to  $9 \times 9$  cm (first dimension) and  $6.5 \times 34$  cm (second dimension). The wicks are saturated with electrode buffer, and the exposed parts are wrapped in Saran wrap or between Mylar sheets to prevent evaporation. Upper and lower reservoir capacities are 375 ml each for the first-dimension unit and 1000 ml each for the second-dimension apparatus. Although the second-dimension unit can accommodate up to three gels, we use only one gel per unit and do not recirculate reservoir buffers.

*c. First-Dimension Electrophoresis.* The first-dimension gel is 8% polyacrylamide (acrylamide:Bis ratio, 80:1), 6 M urea, pH 3.3. The electrode buffer is 25 mM citric acid adjusted to pH 3.3 with 10 N NaOH. Stock solutions and mixing ratios are given in Table I, and the conditions for electrophoresis are shown schematically in Fig. 4.

The acrylamide-urea solution is titrated to pH 3.3 with saturated citric

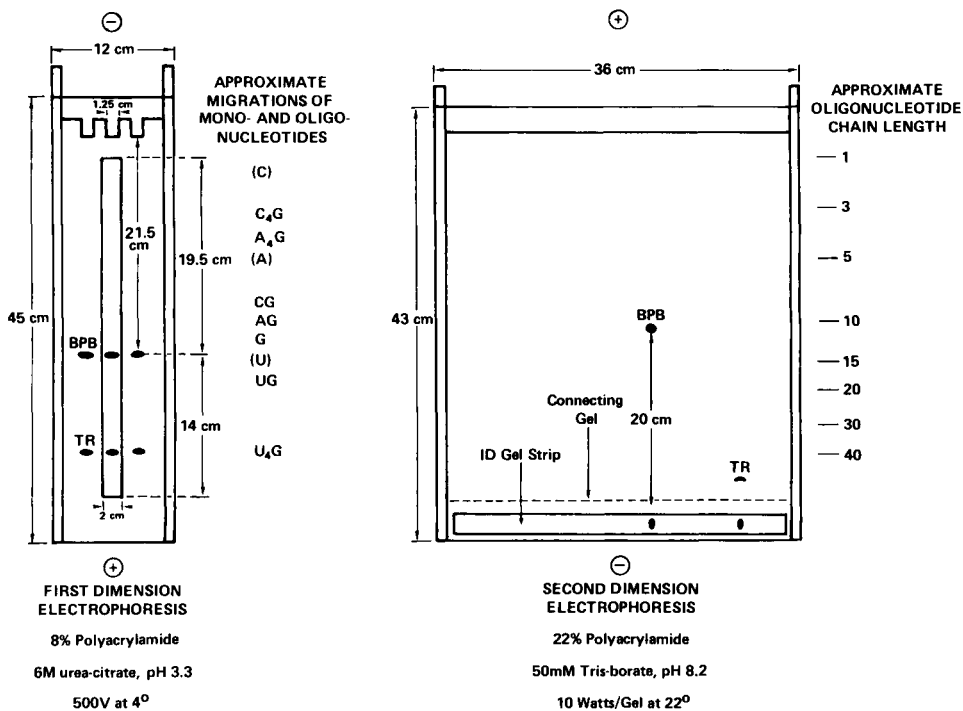


FIG. 4. Schematic representation of the two-dimensional polyacrylamide gel electrophoresis fingerprinting method. Separation in the first dimension is according to oligonucleotide charge, which follows from composition. The approximate positions relative to the dye bands of the 3'-mononucleotides and some small oligonucleotides are shown. The 3'-mononucleotides of cytosine (C), adenosine (A), and guanosine (G) are not produced by RNase T1 digestion, as indicated by the enclosure of their symbols in parentheses. Second-dimension mobilities decrease with the logarithm of the oligonucleotide chain lengths.

acid (usually less than 1 ml is required; substantially greater amounts indicate the presence of impurities in the urea). The solution is warmed to 37°C, then mixed with the catalysts of Jordan and Raymond (1969) in the order  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , ascorbic acid, and hydrogen peroxide, and poured to within about 1.5 cm from the top of the prewarmed (37°C) mold. A 9 × 10 cm sheet of Parafilm, inserted about 1 cm into the mold and supported and curved with one hand, serves as a convenient pouring guide. Bubbles are removed by briefly holding the mold upright. The mold is then set at an angle of about 15°, and the well former is inserted about 2 cm into the gel.

After polymerization (30 min), the bottom tape seal and the well former

are removed, and the top of the gel is washed with and then overlaid with electrode buffer. The sample (up to 60  $\mu$ l and 100  $\mu$ g of RNA digest), in loading buffer, is briefly heated (60°C for 3 min) to disrupt aggregates, and layered in the sample well with a microsyringe (Hamilton). The loaded gel is transferred to the electrophoresis unit in a cold room, and the wick is inserted, saturated with electrode buffer, and wrapped. Electrophoresis is performed at 4°C at a constant potential of 450 V (about 7 mA). Total power (watts = volts  $\times$  amperes), a measure of heat production, should not exceed 3.2 W. The run is terminated when the bromphenol blue dye marker (green at pH 3.3) migrates 21.5 cm from the bottom of the sample well (about 16 hr). The trypan red marker migrates ahead of the bromphenol blue (Fig. 4), but behind the leading oligonucleotide fragments.

*d. Second-Dimension Electrophoresis.* After first-dimension electrophoresis, a glass plate is removed, and, using a Plexiglas template as a guide, the gel is cut with a rotary pizza cutter into 2.5 cm-wide strips containing the migration lanes of oligonucleotides. The strips are cut at their ends 19.5 cm above and 14 cm below the bromphenol blue dye mark (Fig. 2). A piece of Parafilm, about 4  $\times$  45 cm, is centered and then pressed gently on top of a gel strip, with care taken to remove air pockets. One end of the Parafilm and gel strip is gently lifted with forceps and pulled directly back to transfer the gel, without stretching, from the glass plate to the Parafilm strip. At this point the gel strips may be frozen at -70°C if desired. Otherwise, each strip is soaked for 5 min in 500 ml of second-dimension electrode buffer (50 mM Tris-borate, pH 8.2) to remove urea, the presence of which will otherwise inhibit polymerization at the junction between the first- and second-dimension gels. The lower reservoir of the second-dimension electrophoresis apparatus is a convenient vessel for this operation, after which the buffer is discarded. The strip is drained of buffer and placed, gel side up, along the edge of a laboratory bench. A second-dimension glass plate is placed over the gel, and contact is made when the strip is centered, parallel to, and about 1-2 mm from, the short (36 cm) edge of the plate. The ends of the Parafilm are folded over the plate to hold the gel in place, and then the plate is turned over, and the Parafilm is removed. After assembling, sealing, and clamping the chamber, about 10 ml of second-dimension electrode buffer is dispensed into the chamber, and the chamber is placed upright in a tank of cold tap water (40 cm deep) to allow the hydrostatic pressure to express bubbles between the gel and the glass plate. The chamber is removed from the tank, and the buffer is allowed to overlay the gel for 10 min, then poured out.

The second-dimension gel is 22% polyacrylamide (acrylamide:Bis ratio, 15:1) in 50 mM Tris-borate, pH 8.2. Components and mixing volumes for the second-dimension gel are given in Table II. A 10% low-cross-linked

polyacrylamide connecting gel is cast first to form a sealant and junction between the first- and second-dimension gels. After addition of catalysts, about 5 ml of connecting gel solution is pipetted down a spacer; the chamber is quickly tilted from side to side to assure good coverage of the gel strip and then set on a level surface for polymerization (1–2 min). The second-dimension separating gel solution is transferred (via tubing attached to a glass reservoir) until the chamber is about one-third full. The chamber is tilted as before to mix any remaining unpolymerized connecting gel, placed upright in the water tank, filled to about 2 cm from the top, and then overlaid with water. Immersion at this stage serves two purposes. The cold water acts as a heat sink to reduce convection during polymerization, and the balance in hydrostatic pressure retards leakage of material from the chamber (Lee and Wimmer, 1976). Polymerization is complete in about 30–40 min.

Electrophoresis is performed at room temperature at a constant voltage of 600 V (about 16 mA/gel), or at a constant power of 10–12 W/gel. Regulation at constant power is preferred because as the resistance of the second-dimension gel system increases during the run, the voltage may be allowed to rise (up to 1000 V) under conditions of constant Joule heating. For recovery of all size classes of oligonucleotides, electrophoresis is terminated when the bromphenol blue dye migrates 20 cm (about 16 hr). Although separation in the first dimension is limited by the size of the gel strip that can be accommodated by the second-dimension gel mold, resolution of the diagnostic RNase T1-resistant oligonucleotides is improved in the second dimension by continued electrophoresis. Migration of the bromphenol blue dye marker to 30 cm or more is frequently advantageous.

*e. Comments.* Several conditions must be carefully controlled to assure production of high-quality fingerprints. It is particularly important to avoid excessive heating of the gels during electrophoresis. Overheating in the first-dimension separation may cause extensive streaking of the oligonucleotides. Poor contact between the first- and second-dimension gels, caused, for example, by incomplete removal of urea, may also result in localized heating and thus produce distorted fingerprints. Bubbles in the gels or nonuniform polymerization of the second-dimension gel can also cause distorted patterns.

The formulations described here incorporate several important modifications of the original De Wachter and Fiers (1972) fingerprinting system. The Bis cross-linker content is lower in the first-dimension and higher in the second-dimension gels. Urea is omitted from the first-dimension electrode buffer. First-dimension electrophoresis is run at pH 3.3 because it was found empirically that better separation of the oligonucleotides was obtained (Lee and Wimmer, 1976). The original Tris–citrate second-dimen-

sion buffer system is replaced by a Tris-borate system, of lower conductivity (Frisby *et al.*, 1976), thus allowing for higher voltage gradients. Additional changes have been introduced by other investigators. Polymerization of the first-dimension gel may be catalyzed with TEMED and ammonium persulfate (Holland *et al.*, 1979). Pedersen and Haseltine (1980b) omitted urea in the first-dimension gels and included 7 M urea in the second-dimension gels. Shorter oligonucleotides are better resolved, and the urea-containing gels may be frozen without breaking at  $-70^{\circ}\text{C}$ , allowing for higher fluorescent yields with intensifying screens (Laskey, 1980). Second-dimension gels may also be frozen if they contain 0.5% linear polyacrylamide (Freeman *et al.*, 1979; Stern and Kennedy, 1980).

## 2. Homochromatography

The original systems for fingerprinting complete viral genomes were based on the homochromatographic technique of Brownlee and Sanger (1969). Homochromatography was developed to overcome problems encountered earlier in resolving long-chain oligonucleotides (Sanger *et al.*, 1965), and its availability opened the way for a series of important structural and comparative studies of RNA viral genomes (Horst *et al.*, 1972; Duesberg and Vogt, 1973; Billeter *et al.*, 1974; Wang *et al.*, 1975). Homochromatography is also used in sequence analysis of end-labeled RNA (Rensing and Schoenmakers, 1973). A brief description of the elements of this method is given here, and further details are given by Brownlee and Sanger (1969) and Barrell (1971). A mini-fingerprinting technique, in which second-dimension homochromatographic separation is performed on small polyethyleneimine (PEI)-cellulose thin-layer plates, has been described by Volckaert *et al.* (1976).

Separation in the first dimension is by high-voltage electrophoresis on cellulose acetate strips in 5% acetic acid-0.5% pyridine (pH 3.5), 7 M urea. Oligonucleotides are transferred by blotting from the cellulose acetate strip to a thin-layer plate of DEAE-cellulose. After washing the plate with ethanol to remove urea, second-dimension fractionation is performed by ascending chromatography at  $60^{\circ}\text{C}$ . The chromatographic solvent, called a homomixture, contains 7 M urea and 5% yeast tRNA (previously partially base hydrolyzed, then dialyzed against water). Other homomixtures, differing in the pretreatment of the tRNA, are used to resolve shorter chain oligonucleotides. Separation in the second dimension is based upon the relative efficiencies at which the labeled oligonucleotides are displaced by the polynucleotides of the homomixture. Short-chain low-valence oligonucleotides are more readily displaced along the DEAE stationary phase by the polynucleotides of the mobile phase, and thus migrate more rapidly than oligonucleotides with longer chains and higher valences. As a result, oligonucleotides are resolved in the second dimension primarily according to

their chain lengths. However, the contributions of base composition are important, and pronounced deviations from an inverse relationship between mobility and molecular weight are frequently observed.

### 3. *One-Dimensional Fingerprints*

A rapid one-dimensional fingerprinting procedure has been used to distinguish isolates of foot-and-mouth disease virus (La Torre *et al.*, 1982). RNase T1 digests were resolved according to chain length on sequencing gels of 20% polyacrylamide, 8.3 M urea, 1 × 90 mM Tris-90 mM boric acid (pH 8.3)-2 mM EDTA (1 × TBE) (Sanger and Coulson, 1978). Patterns were distinguished by the migrations of the larger oligonucleotides as well as by the relative migrations of the tracts of polycytidylic acid [poly(C)], a genomic feature peculiar to the aphthovirus and cardiovirus groups. Band patterns of closely related viruses were very similar, but isolates with distinct two-dimensional fingerprints also had different one-dimensional band patterns. Although much less sensitive to small genetic differences than two-dimensional procedures, this approach may be generally useful as a preliminary screen for genetic relatedness.

## D. AUTORADIOGRAPHY

After electrophoresis, one plate is removed and the gel is covered with polyethylene wrap. A screen-type X-ray film, 35 × 43 cm (e.g., DuPont Cronex 4, Kodak XRP, Kodak XAR, or Fuji RX), is placed over the gel. The film may be preflashed to improve sensitivity (Laskey, 1980). A calcium tungstate-type intensifying screen (e.g., DuPont Lightning Plus or Fuji Mach 2) is placed over the film, and a Plexiglas plate (36 cm × 43 cm × 1/4 in.) is set on top to assure good contact between the gel and the film. The assembly is securely taped to prevent accidental displacement of the film along the gel, and sealed in a black plastic bag (Picker International) for exposure at 4°C. Good fingerprints are obtained with picornavirus RNA (approximately 7500 bases) after 16 hr of exposure when 5 × 10<sup>6</sup> dpm are applied to the gel. Because diffusion of the larger oligonucleotides in 22% polyacrylamide is very slow at 4°C, gels can be exposed for up to 2 weeks without detectable loss of resolution. Well-defined fingerprints can thus be obtained with as little as 10<sup>5</sup> dpm of RNA.

## E. ANALYSIS OF ISOLATED OLIGONUCLEOTIDES

### 1. *Extraction from Polyacrylamide Gels*

If the oligonucleotides are to be analyzed further, reference spots of <sup>32</sup>P-labeled radioactive ink are made on the polyethylene sheet at each corner of the gel and covered with clear tape. Oligonucleotides, located by align-



ment of reference spots on the developed film and the gel, are excised with a sterile scalpel or cork borer and transferred to a 1.5-ml Eppendorf tube. Oligonucleotides can be recovered with good yields (>80%) by mechanical extraction. The brittle second-dimension gel piece is crushed with a flame-sealed Eppendorf pipet tip (bulb diameter 2–3 mm) attached to a glass rod (Lee and Wimmer, 1976). Excessive grinding should be avoided to prevent substantial release of linear polyacrylamide, which copurifies with nucleic acids. The crushed gel is soaked for 2 hr at 37° C in 750  $\mu$ l of elution buffer [0.5 M NH<sub>4</sub> OAc, 10 mM Mg(OAc)<sub>2</sub>, 0.1% SDS, 2 mM EDTA] (Maxam and Gilbert, 1980) containing 50  $\mu$ g of carrier tRNA. The gel particles are removed by centrifugation (500 g, 2 min) through a spin filter (see below), washed with 200  $\mu$ l of elution buffer, and centrifuged again. To the filtrate is added 2.5 ml of cold 95% ethanol. After 20 min in a dry ice–ethanol bath, the precipitate is pelleted in a swinging-bucket rotor (10,000 g, 4° C, 30 min), washed with 95% ethanol, centrifuged again (10,000 g, 4° C, 5 min), and lyophilized.

The spin filter set consists of a 1-ml Eppendorf pipet tip plugged with silanized glass wool [Alltech Associates, or prepared as described by Maxam and Gilbert (1980)], mounted through an adapter to the top of a 12  $\times$  75 mm polypropylene tube. The reusable adapters are made from 1.5-ml Eppendorf tubes from which the bottoms and the snap caps have been removed.

A different approach is described by Stewart and Crouch (1981). The entire fingerprint is transferred to DEAE paper by uniform suction of elution buffer through the gel. After drying, the oligonucleotides are stably associated with the DEAE paper, and the transferred fingerprints are visualized by autoradiography with intensifying screens at –70° C. Individual oligonucleotides are excised from the DEAE paper and efficiently recovered by elution with triethylamine carbonate, which is removed by lyophilization.

## 2. Composition Analysis

The base compositions of uniformly labeled oligonucleotides can be determined after complete secondary digestion with RNase A. The double digestion products have the general structure (Ap)<sub>0–n</sub>Xp, where X = C, G, or U. Because Gp is present only at the 3' terminus of an RNase T1-generated oligonucleotide, only one G-containing product should be observed upon secondary digestion, with all other base residues present in integral molar ratios relative to Gp. Two basic approaches are most commonly used to identify the products of secondary digestion. The first, described in detail by Barrell (1971), utilizes high-voltage electrophoresis on DEAE paper at pH 3.5. The products are identified by their mobilities relative to markers

of known composition. Identification of longer products (three or more A residues) can be confirmed by elution from the DEAE, digestion with RNase T2 (non-base specific), and electrophoresis as before at pH 3.5. In the second method, developed by Volckaert and Fiers (1977), two-dimensional chromatographic separation occurs on small ( $6.7 \times 10$  cm) PEI-cellulose thin-layer plates (CEL 300, Machery-Nagel, Brinkmann). Samples are spotted 1 cm from each edge and resolved in two steps for each dimension. Elution in the first dimension (along the short axis) is first with water until the front moves 1.5 cm; then the plate is resolved with 20–22% formic acid until the solvent reaches the top. After the plate is air dried, second-dimension separation is at pH 4.3, first in 0.1 M formic acid-pyridine (to 2 cm), then in 1.0 M formic acid-pyridine until the front reaches the top. In the acidic conditions of the first dimension (pH < 2), Ap and Cp have no net negative charges and thus have minimal ionic interaction with the PEI support. Because migration rates increase with decreasing negative charge, the digestion products are resolved primarily according to the identities of their 3'-terminal bases, in the order Cp > Gp > Up. At pH 4.3, migration of the products is reduced with increasing chain length. Thus the base composition of a digestion product can be deduced from its position on the chromatogram.

### 3. Oligonucleotide Sequence Determination

For determination of sequences, either 5'- or 3'-<sup>32</sup>P end-labeled oligonucleotides can be utilized in the technique of partial digestion with base-specific RNases (Simoncsits *et al.*, 1977; Donis-Keller *et al.*, 1977; Donis-Keller, 1980). A commercial kit for enzymatic sequencing of RNA is available from P/L Biochemicals.

Sequences can also be determined from oligonucleotides labeled at either the 5'- or 3'-ends by fingerprinting the products of partial alkaline hydrolysis (wandering-spot analysis). The techniques of separation in wandering-spot analysis are essentially the same as described for total genome fingerprinting and can utilize either electrophoresis-homochromatography (Rensing and Schoenmakers, 1973) or two-dimensional polyacrylamide gel electrophoresis (Nomoto *et al.*, 1981). By examining partial hydrolyzates, wandering-spot analysis can reconstruct the effects on the mobility of an oligonucleotide upon stepwise removal of each base residue. Because, for first-dimension mobilities, removal of a cytosine residue causes the greatest shift toward the anode and removal of a uridine residue causes the greatest shift toward the cathode, wandering-spot analysis is most efficient at locating pyrimidines. Enzymatic sequencing, in contrast, most readily distinguishes purine residues, because of the base specificities of the RNases T1 (G) and U2 (A + G). The complementary strengths of the two sequencing

methods may be combined to sequence oligonucleotides (Darlix *et al.*, 1979; Rommelaere *et al.*, 1979; Nomoto *et al.*, 1981). Adenosine residues are located by partial cleavage with RNase U2 and separation on sequencing gels. The same partial hydrolyzate used to form the "ladder" in the sequencing gel is also fingerprinted, and the positions of C and U residues are determined.

The third direct RNA sequencing technique, involving base-specific chemical cleavages, may also be used, but determinations can be performed only with oligonucleotides labeled at their 3'-ends (Peattie, 1979).

## F. ISOLATION OF INDIVIDUAL RNA MOLECULES FOR FINGERPRINTING

### 1. *Gel Electrophoresis*

Resolution of individual genome segments, subgenomic transcripts, or full-length viral genomes from defective-interfering RNA species for separate fingerprint analysis is usually achieved by gel electrophoresis. The choice of gel system depends upon the size range of the RNA molecules and whether the chains are single or double stranded.

Large, single-stranded RNA molecules over the size range of  $0.5-4 \times 10^6$  can be separated on preparative gels of the high-resolution acid agarose-urea system (Lehrach *et al.*, 1977). The gel contains 1.5% low-gelling-temperature agarose (Marine Colloids, distributed by Miles Laboratories), 6 M urea, and 25 mM sodium citrate (pH 3.5). The electrode buffer is 25 mM sodium citrate (pH 3.5). If the gels are run horizontally, 6 M urea is included in the electrode buffer. Better resolution is generally obtained with vertical gels, cast between frosted glass plates and supported by a 5% polyacrylamide plug. Electrophoresis is at 4°C at a potential of 2-5 V/cm. RNA is eluted from gel slices by dilution in 5-10 volumes of 20 mM Tris-HCl (pH 7.8), 1 mM EDTA, and melting the mixture at 65°C for 5 min. RNA is recovered from the mixture by phenol extraction, reextraction of the interphase, and ethanol precipitation of the combined aqueous phases (Weislander, 1979). If the RNA is to be subsequently labeled *in vitro*, oyster glycogen (Calbiochem) can be substituted for tRNA as a carrier. Two micrograms of glycogen is used for each microgram of tRNA.

Single-stranded RNA molecules over the size range  $0.3-1.5 \times 10^6$  (e.g., influenza virus genome segments) can be resolved on slab gels containing 2-3% polyacrylamide (acrylamide:Bis 20:1), 6 M urea, with both the gel and electrode buffers consisting of Loening's E buffer [36 mM Tris-30 mM NaH<sub>2</sub>PO<sub>4</sub> (pH 7.8)-1 mM EDTA] (Scheurch *et al.*, 1975; Loening, 1969). RNA can be recovered from the gels by mechanical extraction essentially as described for isolation of oligonucleotides.

Good resolution of double-stranded RNA molecules is obtained on a

preparative scale by electrophoresis on gels containing 7.5% acrylamide, 0.2% Bis, 6.3 M urea, 0.1% SDS, in Loening's E buffer. The electrode buffer is Loening's E buffer containing 0.1% SDS (Ito and Joklik, 1972). The discontinuous SDS-polyacrylamide gel system of Laemmli (1970) may also be used for preparative isolation of double-stranded RNA molecules (Clewley and Bishop, 1979a). Double-stranded RNA molecules may be recovered from polyacrylamide gels by mechanical extraction or by electroelution into dialysis membranes (Scheurch *et al.*, 1975; McDonell *et al.*, 1977). Gels are soaked for several hours in deionized water to remove SDS and stained for 30 min with 10  $\mu\text{g}/\text{ml}$  ethidium bromide in  $0.2\times$  TBE. The RNA bands, visualized under long-wavelength ultraviolet light, are excised from the gel. Each gel slice is placed in a dialysis bag (previously boiled in 5 mM EDTA, then autoclaved for 10 min in deionized water) containing a small volume of  $0.2\times$  TBE, and the bag is sealed at both ends without trapping air bubbles. Bags are placed longitudinally between the electrodes of an electroelution tank containing just enough  $0.2\times$  TBE to submerge the bags. Electrophoresis is at 100 V (0–4°C) until the ethidium bromide-stained RNA migrates from the gel (usually 2–3 hr) and forms a fluorescent streak along the inside of the bag. The polarity of electrophoresis is reversed for 2 min, the buffer is removed from the bag, and the bag is washed once with  $0.2\times$  TBE. The RNA-containing buffer is twice extracted with phenol, and the RNA is recovered by ethanol precipitation.

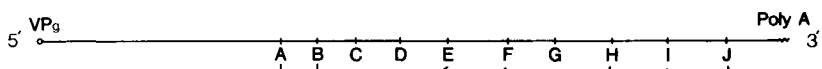
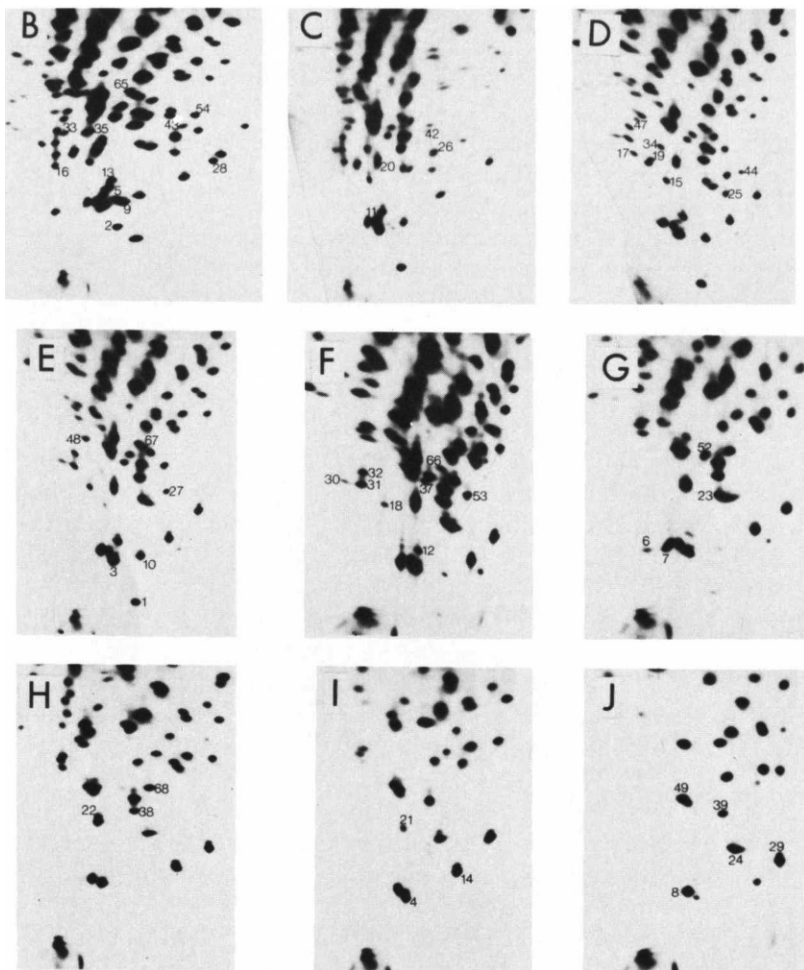
## 2. Hybridization Selection

Ortin *et al.* (1980) isolated influenza virus segments for fingerprinting by hybridization of total viral RNA to plasmid DNA, immobilized on nitrocellulose filters, containing the sequences of specific segments. After hybridization at 37°C for 24–36 hr in 50% formamide, 0.5% SDS,  $5\times$  SSC ( $1\times$  SSC = 150 mM NaCl, 15 mM sodium citrate, pH 7.0), the filters were washed with  $2\times$  SSC, and the hybrids were treated with 10 units/ml RNase T1 in 10 mM Tris-HCl (pH 7.5), 0.5 M NaCl, in the presence of 20  $\mu\text{g}/\text{ml}$  carrier tRNA. The filters were again washed with  $2\times$  SSC, and the resistant RNA was eluted at 50°C in formamide and recovered by ethanol precipitation. Hybridization selection should find wider application as a preparative method with the increasing availability of cloned copies of viral genes.

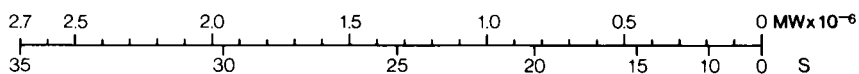
## G. OLIGONUCLEOTIDE MAPPING

### 1. Genomes with Terminal Poly(A) Sequences

Physical genomic maps of RNA viruses can be constructed by determination of the order of the unique RNase T1 oligonucleotides along the RNA molecule. Ordering of oligonucleotides is most easily accomplished with



40,43,55, 87,28,36, 73,78,62, 2,33,35, 65, 9,60.	5, 16,20, 54,13,42, 72,77,81, 82,84,26,	46,11,51, 44,47,45, 69,41,17, 25,56,57.	59,91,15, 19,34, 1, 83,27,74,	48, 3,10, 67,18,37, 12,66,53, 61,71,86,	80,30,31, 32,63,64, 79, 6, 7, 23,	52,38,68, 22,	89,21,58, 85,76, 4, 14,50,75, 88,92.	8,24,29, 39,49,70, 90
--	--	--	-------------------------------------	--	--	------------------	---	-----------------------------



genomes having terminal sequences of polyadenylic acid [poly(A)]. For these, the basic analytical strategy is partial fragmentation of the viral RNA, selection of the poly(A)-containing fraction, size fractionation of the co-terminating fragments, and fingerprinting to determine which oligonucleotides disappear with decreasing chain length (Wang *et al.*, 1975). Several conditions have been used to produce partial fragmentation of the viral RNA, including limited base hydrolysis (Wang *et al.*, 1975), sonication (Kennedy, 1976), boiling in Tris-EDTA buffer (Merregaert *et al.*, 1981), partial digestion with RNase III (Stewart *et al.*, 1980), or radioactive decay of  $^{32}\text{P}$ -labeled phosphodiester bonds (Faller and Hopkins, 1978). Nested sets of polynucleotides are generally obtained with common 3'-ends by selection of the poly(A)-containing fraction on oligo(dT)-cellulose (Aviv and Leder, 1972) or poly(U)-Sepharose (Lindberg and Persson, 1972). Harris *et al.* (1980) obtained fragments of varying lengths extending from the poly(C) tract near the 5'-end of foot-and-mouth disease virus RNA by selection on oligo(dG)-cellulose. Separation according to chain length is generally performed in denaturing sucrose gradients (Wang *et al.*, 1975; Darlix *et al.*, 1979). Better separation of individual size classes of fragments is achieved by electrophoresis on SDS-polyacrylamide-agarose gels (Stewart *et al.*, 1980), but sedimentation gradients are usually preferred because the fractions are more easily recovered.

Results of a typical mapping experiment, following the procedures of Wang *et al.* (1975), are shown in Fig. 5. The basic steps are given below. Partial base hydrolysis of *in vivo*  $^{32}\text{P}$ -labeled virion RNA, incubated at 50°C in low-salt buffer [LSB; 10 mM Tris-HCl (pH 7.8), 10 mM NaCl, 1 mM EDTA], is brought about by addition of 0.05 volume of 1.0 M  $\text{Na}_2\text{CO}_3$  to increase the pH to 10.8. At various intervals (e.g., 0, 15 sec, 30 sec, 1 min, 2 min, . . .) small aliquots are neutralized in 0.1 volume of 0.5 M acetic

---

FIG. 5. Mapping of the oligonucleotides of the Leon strain of type 3 poliovirus using the methods of Wang *et al.* (1975) and Nomoto *et al.* (1979b). Virion RNA labeled *in vivo* was partially base hydrolyzed, and the poly(A)-containing RNA fragments were selected by two cycles of affinity chromatography on oligo(dT)-cellulose. The poly(A)-containing fragments were separated according to size by sucrose gradient sedimentation, followed by resedimentation of each of the 10 pooled fractions. Peak fractions from each gradient were fingerprinted (B-J). The fraction A fingerprint is not shown because it was identical to that of untreated virion RNA (Fig. 1). Spots are numbered as in Fig. 1. (Lower panel) A map of oligonucleotide order determined by visual comparisons of relative spot intensities in the fingerprints above. Numbers given in brackets are arranged in sets of ascending values according to the intensities of the spots in the next (lower S value) gradient pool. Spots that are missing in the next fingerprints are listed first; those that are present but visibly lighter than the others are given in succeeding sets depending upon intensity. The overall map is compressed relative to the physical genome, probably because the relationship between genome size and experimentally determined S value deviates from the equation of Spirin (1963) used to calibrate the map.

acid. Optimal incubation and sampling conditions vary with different RNA preparations. The neutralized samples are then pooled, and ethanol precipitated. To select for the poly(A)-containing fraction, the RNA is resuspended in 500  $\mu$ l of oligo(dT) binding buffer [10 mM Tris-HCl (pH 7.8), 0.5 M NaCl, 1mM EDTA, 0.5% SDS], applied to a column containing 500 mg of oligo(dT)-cellulose, and the column is washed with 20 volumes of binding buffer. Poly(A)-containing RNA is eluted from the oligo(dT) by addition of elution buffer [10 mM Tris-HCl (pH 7.8), 1 mM EDTA, 0.05% SDS], and the peak fractions are pooled. Two cycles of oligo(dT) selection are generally required to eliminate non-poly(A)-containing RNA contaminants. After ethanol precipitation, the poly(A)-containing RNA fragments are resuspended in 0.5 ml of LSB/0.1% SDS, heated at 70°C for 2 min, and resolved by sedimentation at 20°C in a linear 15–30% sucrose gradient in LSB/0.1% SDS. Pools of gradient fractions may be fingerprinted directly. However, resedimentation of each gradient fraction pool may be necessary to obtain greater size homogeneity within each sedimentation fraction (Nomoto *et al.*, 1979b). In some procedures, oligo(dT) selection is the final step before fingerprinting, with each size class pool selected in small Pasteur pipet columns containing approximately 100 mg of oligo(dT).

Oligonucleotide maps may be constructed by visual comparison of the fingerprints of each size class and recording the order in which oligonucleotide spots become weak and then absent with decreasing fragment size. More quantitative comparisons are made by fingerprinting digests of each fragment size class mixed with differentially labeled full-length RNA. The RNA fragments are labeled with [<sup>3</sup>H]uridine and the full-length reference RNA with <sup>32</sup>P. The unique oligonucleotides are excised from the fingerprints, and their radioactivity is determined by scintillation counting. The order of the oligonucleotides is deduced by plotting the <sup>3</sup>H/<sup>32</sup>P ratios for each oligonucleotide as a function of the size fractionation range of the gradients (Coffin *et al.*, 1978).

Because of the limited resolving power of sedimentation gradients and the nonuniform distribution of unique oligonucleotides along the genome, the maps obtained by the above procedures are only approximate and are not a precise measure of the physical distances between oligonucleotides. The largest fragments are least efficiently resolved because they have the most similar sedimentation values; therefore the largest errors in assigning the oligonucleotide order are expected to occur for the 5'-terminal region. Oligonucleotide maps have been useful in identifying and characterizing recombinant genomes produced by intramolecular exchange (Beemon *et al.*, 1974), in correlating biological properties with the presence of certain mapped oligonucleotides (Wang, 1978), in mapping deletion mutants of viral genomes (Kennedy, 1976), and in localizing some mutational differences between closely related strains (Nomoto *et al.*, 1981).

## 2. Negative-Stranded Viruses

Oligonucleotides of the negative-stranded viruses vesicular stomatitis virus (Freeman *et al.*, 1979; Clewley and Bishop, 1979a) and Sendai virus (Amesse and Kingsbury, 1982) have been assigned to specific structural genes by identifying the set of oligonucleotides protected from nuclease digestion in mRNA-virion RNA duplexes. Freeman *et al.* (1979) separately hybridized  $^{32}\text{P}$ -labeled virion RNA to saturating amounts of each viral mRNA. The partial duplexes were digested with RNase T1 and fingerprinted, and the oligonucleotides appearing in the fingerprints were assigned to the gene encoding the missing mRNA. Clewley and Bishop (1979a) applied the reciprocal approach and produced separate duplexes with each mRNA, then digested the single-stranded regions with a nuclease mixture containing S1 nuclease and RNases A, T1, and T2. The protected duplexes were purified on polyacrylamide gels and fingerprinted. The observed oligonucleotides were assigned to the gene encoding the protecting mRNA. Because the structural genes had been ordered by independent means, a partial oligonucleotide map could be constructed. Further ordering of oligonucleotides within each gene would require partial hydrolysis and fractionation of the poly(A)-containing mRNAs. This approach was useful in mapping deletions in defective-interfering mutants (Freeman *et al.*, 1979; Amesse *et al.*, 1982) and in comparing the genes of closely related rhabdoviruses (Clewley and Bishop, 1979a).

## H. GENOME COMPLEXITY ANALYSIS

Fingerprinting has been used to provide a measure of the number of nucleotide residues per unique genome. The analysis is based on the principle that, for uniformly labeled RNA, the specific activity of any unique oligonucleotide is equivalent to that of the complete RNA molecule (Sinha *et al.*, 1965; Fiers *et al.*, 1965). Thus the chain length per unique genome can be calculated from the relationship

$$\text{chain length (RNA)} = \frac{\text{chain length (oligonucleotide)} \times \text{cpm (RNA)}}{\text{cpm (oligonucleotide)}}$$

In practice, uniformly labeled viral RNA is fingerprinted and several unique oligonucleotides are analyzed for radioactivity and chain length, determined from either composition data (Billeter *et al.*, 1974) or sequence data (Lee *et al.*, 1979). Several oligonucleotides are examined because of variable recoveries of oligonucleotides from the fingerprints. Some losses of the input RNA counts occur during fingerprinting, and correction for losses has been approached by the addition of oligonucleotides of known specific activities (internal standards), by parallel analyses of RNAs of



known size and complexities (external standards) (Billeter *et al.*, 1974), or by counting the entire fingerprint gel (Lee *et al.*, 1979).

The complexity, or level of ploidy, of an RNA viral genome can be established by comparing the minimum genome size with other physical measures of the size of the viral RNA. Such analyses have been widely used to determine the complexities of the polyploid retroviral genomes (Billeter *et al.*, 1974; Beemon *et al.*, 1976; Friedrich *et al.*, 1976; Vigne *et al.*, 1977). Complexity analysis has also been used to determine the chain lengths of the genomic RNAs of Semliki Forest virus (Lomniczi and Kennedy, 1977), avian infectious bronchitis virus (Lomniczi and Kennedy, 1977), poliovirus types 1 and 2 (Lee *et al.*, 1979), and the genome segments of the bunyaviruses Uukuniemi virus (Pettersen *et al.*, 1977) and snowshoe hare virus (Clewley *et al.*, 1977b).

#### I. OLIGONUCLEOTIDE MAPPING IN GENOME SEQUENCE ANALYSIS

The ultimate identification of a virus strain is the determination of its complete genomic sequence. In studies to determine the total genome sequence of the Mahoney strain of type 1 poliovirus (Kitamura *et al.*, 1981), RNase T1- and RNase A-resistant oligonucleotides were used to prime the synthesis of DNA chains in the DNA polymerase I-mediated dideoxy sequencing system (Kitamura and Wimmer, 1980). Because the viral genome and the oligonucleotide primers are of the same polarity, the template was cDNA synthesized from virion RNA. The resultant sequence, unlike those determined from cloned DNA copies, represents the average genome sequence of the viral population (see the next section). Oligonucleotides have also been used in sequence studies with cloned DNA, serving as hybridization probes to characterize clones (Boothroyd *et al.*, 1981; van der Werf *et al.*, 1981).

### *IV. Applications*

#### A. FINGERPRINTING AS A MEANS OF STUDYING LARGE VIRUS POPULATIONS

The fingerprint obtained experimentally represents the average genotype of a large virus population. Because of the extremely high inherent mutabilities of RNA genomes (Holland *et al.*, 1982), substantial genetic polymorphism may exist in the virus population. Model fingerprinting studies with phage Q $\beta$  revealed that the genome of most individual members of the population differed from the average genome at one or more sites (Dom-

ingo *et al.*, 1978). Further, more than one average genome may exist for the population, particularly if the virus population is in a state of disequilibrium (Holland *et al.*, 1982). If the alternative average genomes have different fingerprints, then the variant oligonucleotides will be detected as lighter, submolar spots. It is the experience of many laboratories that virus stocks maintained in the laboratory by low-multiplicity passage in cell culture appear to maintain a stable genetic equilibrium, such that their fingerprints remain unaltered with increased passage.

#### B. ANALYSIS OF NATURAL VIRUS ISOLATES

During natural transmission a virus population may encounter numerous complex conditions that favor disequilibrium and the selection of new variants. Given a succession of selective conditions, genome evolution is continual and is detected as cumulative changes in the fingerprints of successive isolates. The nature of the selective pressures is poorly defined, but appears to involve factors in addition to host antibodies, as natural genomic evolution can involve changes in sequences encoding nonstructural proteins (Young *et al.*, 1979; Domingo *et al.*, 1980; Nottay *et al.*, 1981) and can occur during persistent infection of cultured cells (Holland *et al.*, 1979; Meinkoth and Kennedy, 1980) and during infection of persons with severe immunodeficiencies (Yoneyama *et al.*, 1982).

In fact, the observed rates of genome evolution in nature vary widely among RNA viruses. Presumably, this is due to varying degrees of selective pressure encountered during the natural life cycles, as well as to possible differences in the basic mutation rate (Pringle *et al.*, 1981).

For some viruses, e.g., influenza A virus (Young *et al.*, 1979; Ortin *et al.*, 1980), foot-and-mouth disease virus (Domingo *et al.*, 1980), and poliovirus (Nottay *et al.*, 1981), the rate of natural evolution is so rapid that it is possible to reconstruct the general pathways of epidemiological transmission based upon the pattern of fixation of changes into the fingerprints. Such applications of fingerprint analysis have sometimes been described as "molecular epidemiology."

Many other viruses, including vesicular stomatitis virus (Clewley *et al.*, 1977a) and Western equine encephalitis virus (Trent and Grant, 1980), evolve so slowly that viruses with similar fingerprints can be isolated over a period of many years. In such cases, transmission pathways cannot be determined by fingerprint analysis, although the distribution of a genotype over space and time may be followed. Many virus strains appear to be restricted geographically, possibly because of the limited range of their reservoirs and vectors, and disappear from the environment as a genotype identifiable by fingerprinting, either by continued evolution or by displace-

ment by strains of the same group but with distinct fingerprints. Virus genotypes, recognized by fingerprinting and geographic distribution, are called "topotypes."

Representatives of nearly every RNA virus group have been fingerprinted. The available data constitute a growing fingerprint catalog of RNA virus genotypes. For some viruses, the catalog may require frequent updating, because of the rapid evolution, whereas for others updating may not be necessary for decades.

### 1. *Molecular Epidemiology*

*a. Influenza Viruses.* Fingerprint analysis of influenza A (H1N1) isolates from the 1977-1979 pandemic revealed their close genetic relationship to H1N1 strains isolated in 1950 (Nakajima *et al.*, 1978). The 1977-1979 isolates showed a pattern of continual evolution by fixation of mutations, allowing construction of an evolutionary tree based upon fingerprint changes (Young *et al.*, 1979). In view of the rapid evolution of the pandemic virus, the great similarities in the fingerprints of isolates obtained 27 years apart suggested that the 1950 virus had reemerged after being maintained in a state of genetic dormancy, the exact nature of which is unknown. The evolution of H3N2 viruses has also been studied by fingerprinting (Ortin *et al.*, 1980; Nakajima *et al.*, 1982).

Influenza C viruses appear to be much more stable in nature, as isolates from four continents obtained over a 32-year period had generally similar fingerprints (Meier-Ewert *et al.*, 1981).

*b. Enteroviruses.* We use fingerprinting in our laboratory as a routine method for identifying the poliovirus genotype associated with an epidemic. The results of an epidemiological study are shown in Fig. 6 (Nottay *et al.*, 1981). It has been possible, on the basis of fingerprint data alone, to reveal unsuspected epidemiological links between cases or outbreaks that occur at different locations and at different times (Nottay *et al.*, 1981). An important fraction of isolates from the few remaining cases of paralytic poliomyelitis in developed countries has been shown, by fingerprinting, to be genetically related to the live poliovaccines (see, e.g., Fig. 2) (Minor, 1980; Nottay *et al.*, 1981; Kew *et al.*, 1981). Case isolates of enterovirus 70 from the 1980-1981 pandemic of acute hemorrhagic conjunctivitis were shown by fingerprinting to be closely related to each other and to some pre-pandemic isolates (Kew *et al.*, 1983).

*c. Foot-and-Mouth Disease Viruses (FMDV).* Isolates of type C FMDV obtained from diseased swine and cattle in Spain in 1979 were shown to be genetic variants of a 1970 strain from the same region (Domingo *et al.*, 1980). Similar findings were presented by this group for a type O strain. FMDV isolates from a 1981 outbreak in England and France were shown

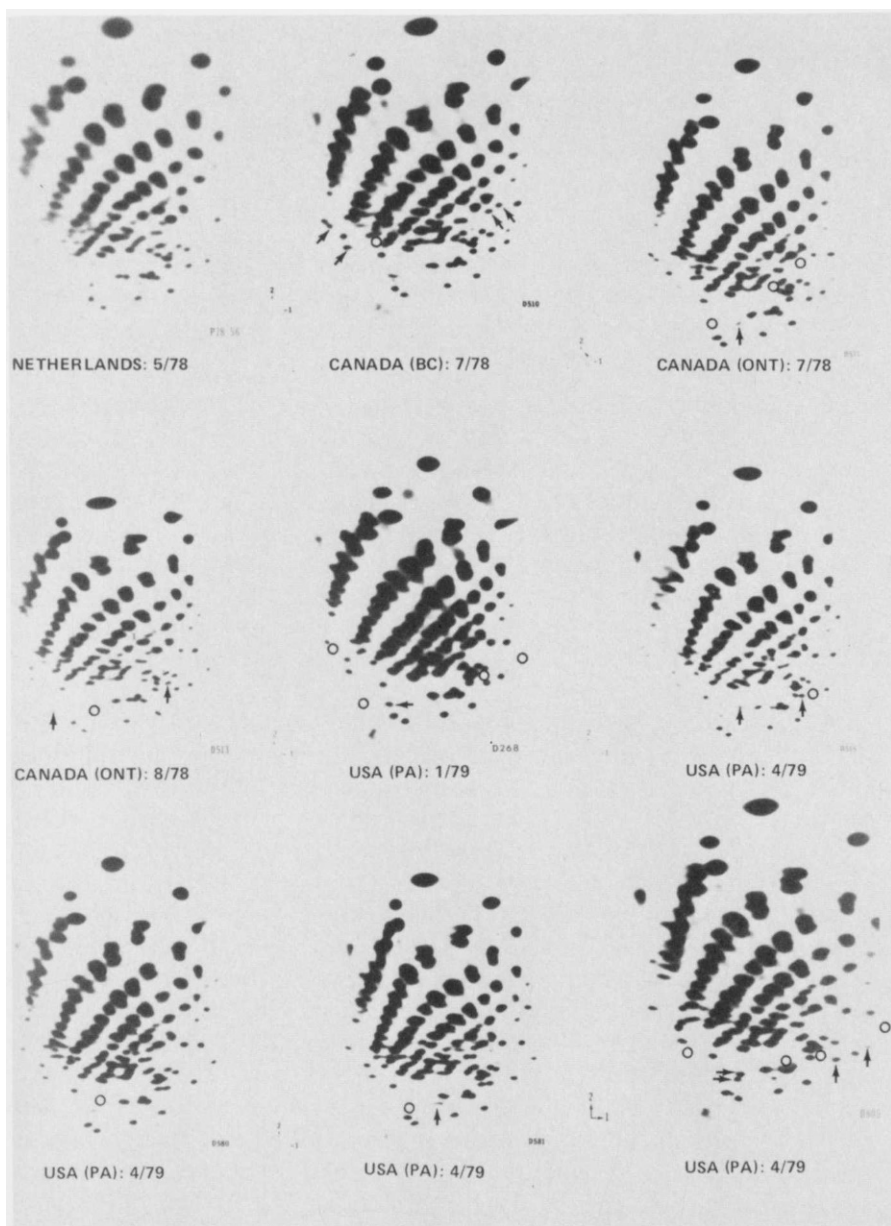


FIG. 6. Fingerprints of type 1 poliovirus isolates from the 1978-1979 Netherlands-Canada-United States epidemic. Fingerprints are arranged according to date of isolation. Note the general similarities among the fingerprints of isolates from this epidemic, and the obvious differences between this group and the fingerprint of a wild-type 1 poliovirus isolate from a different epidemic (Fig. 2). Note also evidence of evolution of the virus during epidemic transmission. Arrows (new spots) and open circles (missing spots) indicate differences in each pattern from the next earlier isolate's fingerprint. From Fig. 5 of Nottay *et al.* (1981).

to have fingerprints nearly identical to that of a vaccine strain isolated in 1965. Because of the high mutability of FMDV during epizootic transmission, unaltered survival of the 1965 strain in the field is unlikely, strongly suggesting that the outbreak was caused by accidental reintroduction of the virus into the field (King *et al.*, 1981).

## 2. *Comparison of Natural Isolates Obtained at Different Times and Locations*

*a. Picornaviruses.* Frisby *et al.* (1976) fingerprinted representatives of three of the four major picornavirus subgroups. Except for the cardioviruses encephalomyocarditis virus and Mengo virus, which had similar fingerprints, different viruses within and across subgroups had dissimilar fingerprints. The genetic relationships between isolates of swine vesicular disease virus and coxsackievirus B5 were studied by fingerprinting and competition hybridization (Harris *et al.*, 1977). Fingerprint analysis showed that the eight independent isolates of Theiler's murine encephalomyelitis virus belong to two distinct genetic subgroups (Lorch *et al.*, 1981). Attenuated strains of FMDV (Harris and Brown, 1977) and poliovirus (Nomoto *et al.*, 1981; Nottay *et al.*, 1981) have been compared by fingerprinting with their virulent parents.

*b. Alphaviruses.* The alphaviruses have undergone extensive evolutionary divergence, as little sequence homology is detected by hybridization across the three serocomplexes, Eastern equine encephalitis (EEE), Western equine encephalitis (WEE), and Venezuelan equine encephalitis (VEE) (Wengler *et al.*, 1977). For both the VEE (Wengler *et al.*, 1977; Trent *et al.*, 1979) and WEE (Trent and Grant, 1980) groups, fingerprint analysis generally corroborated serological classification. Isolates that were very close serologically had similar fingerprints, clearly distinguishable from those of close, but distinct, serological relatives. Individual alphavirus strains appear to evolve slowly, as very similar fingerprints were characteristic of WEE isolates widely separated temporally and geographically (Trent and Grant, 1980).

*c. Flaviviruses.* The flaviviruses are divided into seven subgroups, with very limited genetic homology across subgroups. Fingerprints of viruses of different subgroups (Wengler *et al.*, 1978) or different viruses within a subgroup (the prototypes of dengue virus types 1, 2, 3, and 4) (Veza *et al.*, 1980a) showed no similarities. However, when comparisons were restricted to independent isolates of the same flavivirus, substantial genetic homology among representatives of a strain was evident. Trent *et al.* (1981) surveyed numerous North and Central American isolates of St. Louis encephalitis (SLE) virus from humans, birds, rodents, and mosquitoes obtained over a 47-year period. Fingerprint similarities were seen among all

isolates, with differences among genotypic variants primarily correlating with differences in geographic region of isolation.

*d. Rhabdoviruses.* Fingerprints of the genetically divergent rhabdoviruses vesicular stomatitis virus (VSV), rabies, Cocal, Chandipura, and spring viremia of carp virus were, as expected, quite distinct (Clewley and Bishop, 1979b). However, independent isolates of the VSV Indiana group, obtained from diverse biological sources at different locations over a 50-year period, shared many similarities in their fingerprints (Clewley *et al.*, 1977a). The existence of two genetic subtypes of VSV New Jersey was demonstrated by fingerprint analysis. Three different laboratory strains (Clewley and Bishop, 1979b) of rabies have been fingerprinted.

*e. Bunyaviruses.* The bunyavirus group contains over 200 accepted or proposed members, divided among four genera. Members of different serogroups have clearly unique fingerprints for each L, M, and S genome segment (Pettersson *et al.*, 1977; Ushijima *et al.*, 1980; Clerx and Bishop, 1981). Within the California serogroup, individual strains, distinguishable on serological grounds, were also shown to have dissimilar fingerprints (El Said *et al.*, 1979). Fingerprint similarities were evident for independent isolates of La Crosse virus (California serogroup), although differences were observed among isolates. Variants existed as topotypes, as isolates obtained over 16 years from the same region were more closely related to each other than isolates obtained at the same times from different regions. The genome segments of La Crosse appeared to evolve at different rates,  $M > L > S$  (El Said *et al.*, 1979).

*f. Arenaviruses.* Two variants of Pichinde virus (Tacaribe serocomplex) isolated 5 years apart in Colombia had clearly distinguishable fingerprints for their respective L and S segments. Vezza *et al.* (1980a) utilized these differences to provide physical evidence for independent assortment among arenaviruses.

*g. Coronaviruses.* The fingerprints of seven isolates of mouse hepatitis virus, obtained at different times and locations, indicated that all members of the group were closely related, having from 40 to 95% of their oligonucleotides in common (Lai and Stohlman, 1981). Independent isolates of infectious bronchitis viruses (IBV) had many distinct fingerprints, indicating a potential for substantial genetic variation even within a serotype. At the same time, some genotypes of IBV appeared to have wide geographic and temporal distributions (Clewley *et al.*, 1981).

*h. Measles Virus.* Fingerprints of measles virus isolated from cases of acute disease and from cases of subacute sclerosing panencephalitis revealed distinguishable patterns for each, with some large oligonucleotides shared (Stephenson and ter Meulen, 1982).

*i. Ebola Virus.* Coetaneous isolates of Ebola virus, from patients with

hemorrhagic fevers, obtained in different locations (Zaire and Sudan) had very distinct fingerprints. Subsequent isolates had fingerprints similar to the earlier strains found in each location (Cox *et al.*, 1983), suggesting a slow rate of evolution for each Ebola topotype.

*j. Orbiviruses.* Fingerprint comparisons of isolates of bluetongue virus type 11 (BTV-11), obtained over a 12-year period in the United States, indicated that evolution in the natural environment occurred by both mutational drift and reassortment, and that several BTV-11 genotypes could be in circulation in the United States at the same time (Sugiyama *et al.*, 1982). Two Australian orbiviruses have also been compared by fingerprinting (Walker *et al.*, 1980).

*k. Retroviruses.* Fingerprinting has found wide application in the study of retroviruses. The large literature on this subject was separately and comprehensively reviewed by Beemon (1978) and Wang (1978). More recently, fingerprinting was used to reveal the genetic similarity of viruses from human leukemia patients to simian and gibbon ape retroviruses (Sahagan and Haseltine, 1980), to demonstrate that retroviral sequences from lymphomas of AKR mice are produced by recombination among at least three different endogenous retroviruses (Pedersen *et al.*, 1981; Thomas and Coffin, 1982), to document the participation of retroviral mRNA molecules in the recombination process (Wang and Stacey, 1982), and to show that retroviruses isolated from wild mice in different geographic locations had divergent sequences (Lai *et al.*, 1982). Using a combination of fingerprinting and oligonucleotide sequencing, Clements *et al.* (1982) demonstrated that common nucleotide sequence changes occurred in antigenic variants of visna virus isolated from different sheep.

## C. OTHER APPLICATIONS RELATING TO VIRUS IDENTIFICATION

### 1. *Analysis of Recombinants and Reassortants*

Recombination by intramolecular exchange was demonstrated for retroviruses (Beemon *et al.*, 1974) and picornaviruses (King *et al.*, 1982) by fingerprint analysis of those progeny from mixed infections having recombinant genetic markers. In both systems, parental strains having easily distinguishable fingerprints and well-defined oligonucleotide maps were used. The approximate sites of crossover could be mapped because, for any region along the genome, only one parent contributed oligonucleotides to the recombinant RNA molecule. Natural intertypic recombinants of the live poliovaccines, isolated from humans exposed to trivalent oral vaccine, have been characterized by oligonucleotide mapping (Kew and Nottay, 1984).

Genetic exchange in viruses with segmented genomes occurs by independent

assortment. Although reassortment can be confirmed by other physical means, such as one-dimensional electrophoresis (Young and Palese, 1979; Walker *et al.*, 1980), fingerprinting of individual segments is much more sensitive and can clearly indicate whether each segment is unique. Fingerprinting has been used to characterize natural reassortants of influenza A viruses (Young and Palese, 1979; Nakajima *et al.*, 1982), bluetongue virus (Sugiyama *et al.*, 1982), and the bunyavirus Patois (Ushijima *et al.*, 1980).

### 2. *Analysis of Defective-Interfering Viral Genomes*

Defective-interfering (DI) particles containing genome deletions are produced by nearly all RNA viruses. The presence of DI RNAs may complicate comparative studies among viruses because they can contribute substantial heterogeneity to the virus RNA population. Deletions may be located by comparison of the fingerprints of DI RNAs with the mapped oligonucleotides of full-length genome. Defective RNA genomes of avian retroviruses (Wang *et al.*, 1975; Beemon, 1978), Semliki Forest virus (Kennedy, 1976; Pettersson, 1981; Soderlund *et al.*, 1981), Sindbis virus (Dohner *et al.*, 1979), poliovirus (Nomoto *et al.*, 1979b), influenza virus (Davis and Nayak, 1979; Nakajima *et al.*, 1979), vesicular stomatitis virus (Clerx-van Haaster *et al.*, 1980; Hagen and Huang, 1981), and Sendai virus (Amesse *et al.*, 1982) have been characterized by fingerprinting.

### 3. *Mapping Subgenomic Intracellular RNAs*

Subgenomic RNA molecules of genome polarity are produced intracellularly during the infectious cycles of alphaviruses, retroviruses, and coronaviruses. Fingerprinting has been used to map the subgenomic RNAs to the full-length genome for Semliki Forest virus (Kennedy, 1976; Wengler and Wengler, 1976), Rous sarcoma virus (Mellon and Duesberg, 1977), infectious bronchitis virus (Stern and Kennedy, 1980), and mouse hepatitis virus (Lai *et al.*, 1981).

## V. *Summary and Conclusions*

Fingerprinting is a definitive method of identifying RNA viruses according to their genotypes. It is not subject to the problems of antigenic drift or antigenic convergence that complicate serological identification. Furthermore, it provides a semiquantitative means of following the evolution of viral genomes in nature. Because all regions of the genome are represented by the large diagnostic oligonucleotides, a survey of the total genomic changes can be monitored.

Fingerprinting has two limitations as a diagnostic tool. First, although



highly definitive, fingerprinting is not as rapid or inexpensive as serological techniques and cannot be as easily scaled up for routine identification of a large number of samples. Second, the evolutionary range of fingerprinting is short, and relationships may not be evident for isolates of rapidly evolving viruses obtained over long intervals. However, these limitations are not large, compared to the full benefits offered the virologist by the fingerprinting method.

#### ACKNOWLEDGMENTS

We are grateful to Drs. Fred Brown, Tim Harris, Yuan Fon Lee, and Eckard Wimmer for introducing our laboratory to the fingerprinting technique. We thank Marie Knox for careful preparation of the manuscript.

#### REFERENCES

- Aaronson, R. P., Young, J. F., and Palese, P. (1982). *Nucleic Acids Res.* **10**, 237-246.
- Adesnik, M. (1971). *Methods Virol.* **5**, 126-177.
- Amesse, L. S., and Kingsbury, D. W. (1982). *Virology* **118**, 8-16.
- Amesse, L. S., Pridgen, C. L., and Kingsbury, D. W. (1982). *Virology* **118**, 17-27.
- Aviv, H., and Leder, P. (1972). *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1408-1412.
- Banerjee, A. K. (1980). *Microbiol. Rev.* **44**, 175-205.
- Barrell, B. G. (1971). In "Procedures in Nucleic Acids Research" (G. L. Cantoni and D. R. Davies, eds.), Vol. 2, pp. 751-779. Harper & Row, New York.
- Beemon, K. L. (1978). *Curr. Top. Microbiol. Immunol.* **79**, 73-110.
- Beemon, K. L., Duesberg, P. H., and Vogt, P. K. (1974). *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4254-4258.
- Beemon, K., Faras, A. J., Haase, A. T., Duesberg, P. H., and Maisel, J. E. (1976). *J. Virol.* **17**, 525-537.
- Billeter, M. A., Parsons, J. T., and Coffin, J. M. (1974). *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3560-3564.
- Boothroyd, J. C., Highfield, P. E., Cross, G. A. M., Rowlands, D. J., Lowe, P. A., Brown, F., and Harris, T. J. R. (1981). *Nature (London)* **290**, 800-802.
- Brown, F. (1982). *Ann. Neurol.* **9**, (Suppl.), 39-43.
- Brown, W. M., George, M., and Wilson, A. C. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1967-1971.
- Brownlee, G. G., and Sanger, F. (1967). *J. Mol. Biol.* **23**, 337-353.
- Brownlee, G. G., and Sanger, F. (1969). *Eur. J. Biochem.* **11**, 395-399.
- Chaconas, G., and van de Sande, J. H. (1980). *Methods Enzymol.* **65**, 75-85.
- Clements, J. E., D'Antonio, N., and Narayan, O. (1982). *J. Mol. Biol.* **158**, 415-434.
- Clerx, J. P. M., and Bishop, D. H. L. (1981). *Virology* **108**, 361-372.
- Clerx-van Haaster, C. M., Clewley, J. P., and Bishop, D. H. L. (1980). *J. Virol.* **33**, 807-817.
- Clewley, J. P., and Bishop, D. H. L. (1979a). *J. Virol.* **30**, 116-123.
- Clewley, J. P., and Bishop, D. H. L. (1979b). In "Rhabdoviruses" (D. H. L. Bishop, ed.), Vol. 1, pp. 119-168. CRC Press, Boca Raton, Florida.
- Clewley, J. P., and Bishop, D. H. L. (1982). In "New Developments in Practical Virology" (C. R. Howard, ed.), pp. 231-277. Liss, New York.
- Clewley, J. P., Bishop, D. H., Kang, C.-Y., Coffin, J., Schnitzlein, W. M., Reichmann, M. E., and Shope, R. E. (1977a). *J. Virol.* **23**, 152-166.

- Clewley, J., Gentsch, J., and Bishop, D. H. L. (1977b). *J. Virol.* **22**, 459-468.
- Clewley, J. P., Morser, J., and Lomniczi, B. (1981). In "Biochemistry and Biology of Coronaviruses" (V. ter Meulen, S. Siddell, and H. Wege, eds.), pp. 143-153. Plenum, New York.
- Coffin, J. M., Champion, M., and Chabot, F. (1978). *J. Virol.* **28**, 972-991.
- Commerford, S. L. (1971). *Biochemistry* **10**, 1993-1999.
- Commerford, S. L. (1980). *Methods Enzymol.* **70**, 247-252.
- Cox, N. J., McCormick, J. B., Johnson, K. M., and Kiley, M. P. (1983). *J. Infect. Dis.* **147**, 272-275.
- Darlix, J. L., Levray, M., Bromley, P. A., and Spahr, P. F. (1979) *Nucleic Acids Res.* **6**, 471-485.
- Davis, A. R., and Nayak, D. P. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3092-3096.
- De Wachter, R., and Fiers, W. (1972). *Anal. Biochem.* **49**, 184-197.
- De Wachter, R., and Fiers, W. (1982). In "Gel Electrophoresis of Nucleic Acids: A Practical Approach" (D. Rickwood and B. D. Hames, eds.), pp. 77-116. IRL Press, Oxford.
- Dohner, D., Monroe, S., Weiss, B., and Schlesinger, S. (1979). *J. Virol.* **29**, 794-798.
- Domingo, E., Sabo, D., Taniguchi, T., and Weissman, C. (1978). *Cell* **13**, 735-744.
- Domingo, E., Davila, M., and Ortin, J. (1980). *Gene* **11**, 333-346.
- Donis-Keller, H. (1980). *Nucleic Acids Res.* **8**, 3133-3142.
- Donis-Keller, H., Maxam, A., and Gilbert, W. (1977). *Nucleic Acids Res.* **4**, 2527-2538.
- Duesberg, P. H., and Vogt, P. K. (1973). *J. Virol.* **12**, 594-599.
- El Said, L. H., Vorndam, V., Gentsch, J. R., Clewley, J. P., Calisher, C. H., Klimas, R. D., Thompson, W. H., Grayson, M., Trent, D. W., and Bishop, D. H. L. (1979). *Am. J. Trop. Med. Hyg.* **28**, 364-386.
- England, T. E., and Uhlenbeck, O. C. (1978). *Nature (London)* **275**, 560-561.
- Faller, D. V., and Hopkins, N. (1978). *J. Virol.* **26**, 143-152.
- Fiers, W., Lepoutre, L., and Vandendriessche, L. (1965). *J. Mol. Biol.* **13**, 432-450.
- Freeman, G. J., Rao, D. D., and Huang, A. S. (1979). *Gene* **5**, 141-157.
- Friedrich, R., Morris, V. L., Goodman, H. M., Bishop, J. M., and Varmus, H. E. (1976). *Virology* **72**, 330-340.
- Frisby, D. (1977). *Nucleic Acids Res.* **4**, 2975-2996.
- Frisby, D. P., Newton, C., Casey, N. H., Fellner, P., Newman, J. F. E., Harris, T. J. R., and Brown, F. (1976). *Virology* **71**, 379-388.
- Hagen, F., and Huang, A. S. (1981). *J. Virol.* **37**, 363-371.
- Harris, T. J. R., and Brown, F. (1977). *J. Gen. Virol.* **34**, 87-105.
- Harris, T. J. R., Robson, K., and Brown, F. (1977). *J. Gen. Virol.* **35**, 299-315.
- Harris, T. J. R., Robson, K., and Brown, F. (1980). *J. Gen. Virol.* **50**, 403-418.
- Henry, C. (1967). *Methods Virol.* **2**, 427-462.
- Holland, J. J., Grabau, E., Jones, C. L., and Semler, B. (1979). *Cell* **16**, 495-504.
- Holland, J. J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and Vande Pol, S. (1982). *Science* **215**, 1577-1585.
- Horst, J., Kieth, J., and Fraenkel-Conrat, H. (1972). *Nature (London) New Biol.* **240**, 105-109.
- Huang, E. S., and Pagano, J. S. (1977). *Methods Virol.* **6**, 457-497.
- Ito, Y., and Joklik, W. K. (1972). *Virology* **50**, 189-201.
- Jordan, E. M., and Raymond, S. (1969). *Anal. Biochem.* **27**, 205-211.
- Kennedy, S. I. T. (1976). *J. Mol. Biol.* **108**, 491-511.
- Kew, O. M., and Nottay, B. K. (1984). In "Modern Approaches to Vaccines" (R. M. Chanock and R. A. Lerner, eds.), pp. 357-362. Cold Spring Harbor Laboratory, New York.
- Kew, O. M., Nottay, B. K., Hatch, M. H., Nakano, J. H., and Obijeski, J. F. (1981). *J. Gen. Virol.* **56**, 337-347.

- Kew, O. M., Nottay, B. K., Hatch, M. H., Hierholzer, J. C., and Obijeski, J. F. (1983). *Infect. Immun.* **41**, 631-635.
- King, A. M. Q., Underwood, B., McCahon, D., Newman, J. W. I., and Brown, F. (1981). *Nature (London)* **293**, 479-480.
- King, A. M. Q., McCahon, D., Slade, W. R., and Newman, J. W. I. (1982). *Cell* **29**, 921-928.
- Kitamura, N., and Wimmer, E. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3196-3200.
- Kitamura, N., Semler, B. L., Rothberg, P. G., Larsen, G. L., Adler, C. J., Dorner, A. J., Emini, E. A., Hanecak, R., Lee, J. J., van der Werf, S., Anderson, C. W., and Wimmer, E. (1981). *Nature (London)* **291**, 547-553.
- Laemmli, U. K. (1970). *Nature (London)* **227**, 680-685.
- Lai, M. M. C., and Stohlman, S. A. (1981). *J. Virol.* **38**, 661-670.
- Lai, M. M. C., Brayton, P. R., Armen, R. C., Patton, C. D., Pugh, C. H., and Stohlman, S. A. (1981). *J. Virol.* **39**, 823-834.
- Lai, M. M. C., Shimuzu, C. S., Rasheed, S., Pal, B. K., and Gardner, M. B. (1982). *J. Virol.* **41**, 605-614.
- Laskey, R. A. (1980). *Methods Enzymol.* **65**, 363-371.
- La Torre, J. L., Underwood, B. O., Lebendiker, M., Gorman, B. M., and Brown, F. (1982). *Infect. Immun.* **36**, 142-147.
- Lee, Y. F., and Fowlks, E. (1982). *Anal. Biochem.* **119**, 224-235.
- Lee, Y. F., and Wimmer, E. (1976). *Nucleic Acids Res.* **3**, 1647-1658.
- Lee, Y. F., Kitamura, N., Nomoto, A., and Wimmer, E. (1979). *J. Gen. Virol.* **44**, 311-322.
- Lehrach, H., Diamond, D., Wozney, J. M., and Boedtke, H. (1977). *Biochemistry* **16**, 4743-4751.
- Lindberg, V., and Persson, T. (1972). *Eur. J. Biochem.* **31**, 246-254.
- Loening, U. E. (1969). *Biochem. J.* **113**, 131-138.
- Lomniczi, B., and Kennedy, I. (1977). *J. Virol.* **24**, 99-107.
- Lorch, Y., Friedmann, A., Lipton, H. L., and Kotler, M. (1981). *J. Virol.* **40**, 560-567.
- McDonnell, M. W., Simon, M. N., and Studier, F. W. (1977). *J. Mol. Biol.* **110**, 119-146.
- Markham, R., and Smith, J. D. (1952). *Biochem. J.* **52**, 558-565.
- Maxam, A., and Gilbert, W. (1980). *Methods Enzymol.* **65**, 499-560.
- Meier-Ewert, H., Petri, T., and Bishop, D. H. L. (1981). *Arch. Virol.* **67**, 141-147.
- Meinkoth, J., and Kennedy, S. I. T. (1980). *Virology* **100**, 141-155.
- Mellon, P., and Duesberg, P. H. (1977). *Nature (London)* **270**, 631-634.
- Merregaert, J., Barbacid, M., and Aaronson, S. A. (1981). *J. Virol.* **39**, 219-228.
- Minor, P. D. (1980). *J. Virol.* **34**, 73-84.
- Minson, A. C., and Darby, G. (1982). In "New Developments in Practical Virology" (C. R. Howard, ed.), pp. 185-229. Liss, New York.
- Nakajima, K., Desselberger, U., and Palese, P. (1978). *Nature (London)* **274**, 334-339.
- Nakajima, K., Ueda, M., and Suguira, A. (1979). *J. Virol.* **29**, 1142-1148.
- Nakajima, K., Nakajima, S., and Suguira, A. (1982). *Virology* **120**, 504-509.
- Nakajima, S., Cox, N. J., and Kendal, A. P. (1981). *Infect. Immun.* **32**, 287-294.
- Nomoto, A., Kajigaya, S., Suzuki, K., and Imura, N. (1979a). *J. Gen. Virol.* **45**, 107-117.
- Nomoto, A., Lee, Y. F., Babich, A., Jacobson, A., Dunn, J. J., and Wimmer, E. (1979b). *J. Mol. Biol.* **128**, 165-177.
- Nomoto, A., Kitamura, N., Lee, J. J., Rothberg, P. G., Imura, N., and Wimmer, E. (1981). *Virology* **112**, 217-227.
- Nottay, B. K., Kew, O. M., Hatch, M. H., Heyward, J. T., and Obijeski, J. F. (1981). *Virology* **108**, 405-423.
- Ogra, P. L., Kargon, D. T., Righthand, F., and MacGillivray, M. (1968). *N. Engl. J. Med.* **279**, 893-900.

- Ortin, H., Najera, R., Lopez, C., Davila, M., and Domingo, E. (1980). *Gene* **11**, 319-332.
- Peattie, D. A. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1760-1764.
- Pedersen, F. S., and Haseltine, W. A. (1980a). *J. Virol.* **33**, 349-365.
- Pedersen, F. S., and Haseltine, W. A. (1980b). *Methods Enzymol.* **65**, 680-687.
- Pedersen, F. S., Crowther, R. L., Tenney, D. Y., Reimold, A. M., and Haseltine, W. A. (1981). *Nature (London)* **292**, 167-170.
- Pettersson, R. F. (1981). *Proc. Natl. Acad. Sci. U.S.A.* **78**, 115-119.
- Pettersson, R. F., Hewlett, M. J., Baltimore, D., and Coffin, J. M. (1977). *Cell* **11**, 51-63.
- Pringle, C. R., Wilkie, D. M., Preston, C. M., Dolan, A., and McGeoch, D. (1981). *J. Virol.* **39**, 377-389.
- Ralph, R. K., and Bergquist, P. L. (1967). *Methods Virol.* **2**, 463-545.
- Rensing, V. F. E., and Schoenmakers, J. G. G. (1973). *Eur. J. Biochem.* **33**, 8-18.
- Robertson, H. D., Dickson, E., Plotch, S. J., and Krug, R. M. (1980). *Nucleic Acids Res.* **8**, 925-942.
- Robson, K. J. H., Crowther, J. R., King, A. M. Q., and Brown, F. (1979). *J. Gen. Virol.* **45**, 579-590.
- Rommelaere, J., Donis-Keller, H., and Hopkins, N. (1979). *Cell* **16**, 43-50.
- Sahagan, B. G., and Haseltine, W. A. (1980). *J. Virol.* **34**, 390-401.
- Sanger, F., and Coulson, A. R. (1978). *FEBS Lett.* **87**, 107-110.
- Sanger, F., Brownlee, G. G., and Barrell, B. G. (1965). *J. Mol. Biol.* **13**, 373-398.
- Schuerch, A. R., Mitchell, A. R., and Joklik, W. K. (1975). *Anal. Biochem.* **65**, 331-345.
- Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, R. S., and Guilley, H. (1977). *Nature* **269**, 833-836.
- Sinha, N. K., Fujimura, R. K., and Kaesberg, P. (1965). *J. Mol. Biol.* **11**, 84-89.
- Soderlund, H., Keranen, S., Lehtovaara, P., Palva, I., Pettersson, R. F., and Kaarianen, L. (1981). *Nucleic Acids Res.* **9**, 3403-3417.
- Spirin, A. S. (1963). *Prog. Nucleic Acid Res. Mol. Biol.* **1**, 301-345.
- Stephenson, J. R., and ter Meulen, V. (1982). *Arch. Virol.* **71**, 279-290.
- Stern, D. F., and Kennedy, S. I. T. (1980). *J. Virol.* **36**, 440-449.
- Stewart, M. L., and Crouch, R. J. (1981). *Anal. Biochem.* **111**, 203-211.
- Stewart, M. L., Crouch, R. J., and Maizel, J. V. (1980). *Virology* **104**, 375-397.
- Sugiyama, K., Bishop, D. H. L., and Roy, P. (1982). *Am. J. Epidemiol.* **115**, 332-347.
- Szekely, N., and Sanger, F. (1969). *J. Mol. Biol.* **43**, 607-617.
- Takahashi, K., and Moore, S. (1982). In "The Enzymes" (P. D. Boyer, ed.), 3rd ed., Vol. 15, pp. 435-468. Academic Press, New York.
- Thomas, C. Y., and Coffin, J. M. (1982). *J. Virol.* **43**, 416-426.
- Trent, D. W., and Grant, J. A. (1980). *J. Gen. Virol.* **47**, 261-282.
- Trent, D. W., Clewley, J. P., France, J. K., and Bishop, D. H. L. (1979). *J. Gen. Virol.* **43**, 365-381.
- Trent, D. W., Grant, J. A., Vorndam, A. V., and Monath, T. P. (1981). *Virology* **114**, 319-332.
- Ushijima, H., Klimas, R., Kim, S., Cash, P., and Bishop, D. H. L. (1980). *Am. J. Trop. Med. Hyg.* **29**, 1441-1452.
- van der Werf, S., Bregegere, F., Kopecka, H., Kitamura, N., Rothberg, P. G., Kourilsky, P., Wimmer, E., and Girard, M. (1981). *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5983-5987.
- Veza, A. C., Cash, P., Jahrling, P., Eddy, G., and Bishop, D. H. L. (1980a). *Virology* **106**, 250-260.
- Veza, A. C., Rosen, L., Repik, P., Dalrymple, J., and Bishop, D. H. L. (1980b). *Am. J. Trop. Med. Hyg.* **29**, 643-652.
- Vigne, R., Brahic, M., Filippi, P., and Tamalet, J. (1977). *J. Virol.* **21**, 386-395.
- Volckaert, F., and Fiers, W. (1977). *Anal. Biochem.* **83**, 222-227.

- Volckaert, F., Min-Jou, W., and Fiers, W. (1976). *Anal. Biochem.* **72**, 433-446.
- Walker, P. J., Mansbridge, J. N., and Gorman, B. M. (1980). *J. Virol.* **34**, 583-591.
- Wang, L.-H. (1978). *Annu. Rev. Microbiol.* **32**, 561-592.
- Wang, L.-H., and Stacey, D. W. (1982). *J. Virol.* **41**, 919-930.
- Wang, L.-H., Duesberg, P., Beemon, K., and Vogt, P. K. (1975). *J. Virol.* **16**, 1051-1070.
- Weislander, L. (1979). *Anal. Biochem.* **98**, 305-309.
- Wengler, G., and Wengler, G. (1976). *Virology*, **73**, 190-199.
- Wengler, G., Wengler, G., and Filipe, A. R. (1977). *Virology* **78**, 124-134.
- Wengler, G., Wengler, G., and Gross, H. J. (1978). *Virology* **89**, 423-437.
- Wimmer, E. (1972). *J. Mol. Biol.* **68**, 537-540.
- Wimmer, E. (1982). *Cell* **28**, 199-201.
- Yewdell, J. W., and Gerhard, W. (1981). *Annu. Rev. Microbiol.* **35**, 185-206.
- Yoneyama, T., Hagiwara, H., Hara, M., and Shimojo, H. (1982). *Infect. Immun.* **37**, 46-53.
- Young, J. F., and Palese, P. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76**, 6547-6551.
- Young, J. F., Desselberger, U., and Palese, P. (1979). *Cell* **16**, 73-83.