J. Hyg., Camb. (1985), **94**, 1–7
Printed in Great Britain

SPECIAL ARTICLE

Biochemical techniques for the genetic and phenotypic analysis of viruses: 'Molecular Epidemiology'

New technologies now enable virologists to study small genetic and antigenic differences between field isolates of animal viruses at a higher level of discrimination than has been possible using conventional serological techniques. The most important of the laboratory techniques revolve around the use of monoclonal antibodies, peptide mapping, fingerprinting of whole RNA virus genomes, RNA:RNA hybridization, 'electropherotyping' of virus RNA or polypeptides, restriction enzyme analysis of virus DNA genomes, cloning of genes and rapid sequencing of viral DNAs and RNAs, in the latter case using primer extension techniques (reviewed by Palese & Roizman, 1980). From a practical point of view, genetic and phenotypic heterogeneity among viruses may be of considerable importance in attempts to control certain virus diseases by chemo- or immuno-prophylaxis.

These techniques can be used to determine how biologically and genetically different various viruses are as they circulate in the community. For example, how much genetic variation exists in rotavirus, parainfluenza, influenza, picorna or herpes viruses and is the variation also present in the 'working' protein products of the genetic information, i.e. phenotypic variation? Clinical observation tells us already that some viruses are apparently more variable than others. For example influenza can alter its antigenic phenotype with relative ease and thus escape neutralizing antibodies. Many serotypes of common cold viruses exist, whereas, as far as is known, many fewer serotypes of herpes virus exist. But is the 'field' epidemiological situation even more complicated than this? Even within a number of influenza viruses of the same antigenic subtype genetic variants may exist and flourish with other attributes such as faster replication rate, easier spreadability, low immunogenicity and so on. Antigenic variants of measles virus exist but why are they not detected easily in the community? Is there a range of herpes, influenza or picornaviruses with different virulence characteristics?

We should note that random genetic mutation can occur at a very high frequency with a number of RNA viruses (Holland *et al.* 1982), and with influenza virus, for example, antigenic mutants occur in virus pools at a frequency as high as 1 in 10^5 (Yewdell, Webster & Gerhard, 1979). These high frequencies of genetic mutation are not found in DNA viruses, presumably because of the fidelity of the DNA transcription and replication process and the presence in cells of DNA repair enzymes. This high mutation rate of RNA viruses means that mutants could possibly emerge during a single infection in an individual and, given a selective advantage, such mutants could predominate in the progeny virus. Two important selective pressures which are already known are specific antibody (Yewdell, Webster & Gerhard, 1979) and host cell receptors (Schild *et al.* 1983). Indeed, recent studies using monoclonal antibodies have highlighted the extent of antigenic

variation among a number of viruses other than influenza, including rabies (Wiktor & Koprowski, 1980), measles (Birrer *et al.* 1981) and picornaviruses (Minor *et al.* 1982).

In a recent issue of the *Journal* Follett *et al.* (1984) have used biochemical techniques to analyse the electrophoretic properties of RNA of 146 rotaviruses isolated in Scotland, and this paper gives an excellent idea of the range of data generated by application of these biochemical methods to field isolates. RNA was extracted from semi-purified virus isolated from faecal samples and separated on 2·8% polyacrylamide slab gels containing 6 M urea. The 11 RNA bands were visualized by a sensitive silver staining method (Merril, Switzer & Van Keuren, 1979). Several interesting and novel features were apparent from this biochemical analysis. The authors established clearly that differences in the prevalence of certain electropherotypes existed between different regions of one country during the seasonal high incidence of rotavirus infections. 'Long' or 'short' electropherotypes of rotavirus could be distinguished with ease (the terminology depending upon the relative migration of RNA segments 10 and 11). To make matters more complicated, migration pattern variants (e.g. Lv) were detected which differed in migration of other RNA bands. The L and Lv types started in the east and west of Greater Glasgow in October/November 1981 and spread to the south in December, the peak being in January 1982. In contrast, the S electropherotype disappeared almost completely. The authors were also able to establish that in no single case where RNA migration patterns were obtained from more than one stool specimen was a double infection with rotaviruses with differing RNA patterns found. In other studies, rotaviruses of different RNA migration pattern have been isolated from infected children in different areas of the world either sequentially (Rodger *et al.* 1981; Schnagl, Rodger & Holmes, 1981) or at the same time (Espejo *et al.* 1980; Flores *et al.* 1982; Follett & Desselberger, 1983; Dimitrov *et al.* 1984). Electropherotyping of virus RNA using single-dimension gels is easy to perform and sensitive, but of course only gives unambiguous data under closely controlled conditions. Similarly, Squire *et al.* (1983) reported an analysis of 200 field isolate of bluetongue virus, a reovirus. A wide genetic diversity was apparent but the majority of isolates, especially those isolated in the same year and from the same region, had similar electropherotypes.

In another recent issue of the *Journal* (volume **93**, pp. 237–250) Moosai, Alcock & Madeley (1984) attempt to exploit the occurrence of these rotavirus electropherotypes in order to group rotaviruses for epidemiological analysis. The 'Rotacode' is an interesting proposal, but is likely to cause controversy amongst rotavirologists! However, the authors emphasize that the Rotacode is purely descriptive, since the nine RNA segments are not numbered but rather the resulting cryptogram is based on the relative distance of migration of adjacent bands in each RNA size class – it is a shorthand system and not a classification scheme. New strains can be coded instantly.

In an earlier study more complex biochemical techniques were applied to the genetic analysis of re-emerging influenza A (H1N1) viruses, in what must now be regarded as a classical study (Nakajima, Desselberger & Palese, 1978). It was rapidly established by oligonucleotide mapping of the RNA that the 'new' H1N1 virus was, in reality, not new at all, but almost identical genetically to viruses circulating in the 1950s. Similar methodology was soon applied to an analysis of

field strains and, in an extensive study of influenza A H1N1 viruses circulating in Japan in 1978–9, Nakajima *et al.* (1980) concluded that at least four types of H1N1 viruses could be distinguished by the oligonucleotide pattern. The first type was prevalent throughout Japan from January to March 1978, whereas the second and third types caused outbreaks in the next winter of 1978–9. The fourth type was a recombinant and caused influenza outbreaks in the winter of 1979. Thus, even within a single antigenic subtype (H1N1) a genetically heterogeneous group of viruses was circulating and causing influenza outbreaks. At present it is not known if the genetically different viruses differed in their virulence and 'spreadability', and these are important questions to be answered by further studies. The oligonucleotide mapping technique is particularly useful for analysing genetically related groups of viruses. With viruses where there is more than 10% divergence of nucleic acid sequence the technique will give completely different fingerprints (Aaronson, Young & Palese, 1982). An alternative technique which is under-used at present, but which gives quite extensive semi-quantitative data, is RNA:RNA hybridization (Ghendon *et al.* 1981; Oxford *et al.* 1984).

As an example of the use of monoclonal antibodies in the analysis of field isolates, we have used monoclonal antibodies prepared against influenza B virus HA to examine, in detail, the antigenic composition of a series of virus isolates from a single school outbreak of influenza (Oxford *et al.* 1983). The school represents a semi-closed community, with the boys boarding at the school. Studies of virus evolution during a single virus outbreak can therefore be carried out with less likelihood of new introductions of viruses from the local community as the epidemic proceeds, although the major problem of artifacts induced during laboratory passage still exists. A major wave of influenza B virus occurred approximately two weeks after the first cases were noted. Approximately 161 out of 786 boys fell ill and 81 influenza B virus isolates were obtained by the Public Health Service Laboratory at Guildford. By analysis with a panel of monoclonal antibodies to the HA, the viruses were established as an antigenically heterogeneous group and 13 distinguishable antigenic patterns could be detected, although only two major antigenic groupings were present at all times throughout the outbreak and constituted over half the viruses analysed. In parallel studies the polypeptide and RNA profiles of a number of viruses isolated from this and other schools were examined. Influenza B virus isolates from simultaneous outbreaks in three different schools could be distinguished school by school on the basis of electrophoretic migration rate differences in NS1 polypeptides or by their RNA pattern, although all the viruses from a particular school were very closely related biochemically.

Thus, even in an influenza B virus outbreak in a single school, antigenic variants can co-exist and spread. It is not known at present if the outbreak was initiated by several distinguishable viruses or, for example, whether mutation occurred in the gene coding for HA during the outbreak itself. A possible explanation is that the initially infected boy excreted a genetically heterogeneous collection of virions which were then cloned *in nature* among the remaining boys in the school. In certain boys with a high level of antibody and/or cell-mediated immunity, no infection would result. In others, partial immunity due to antibody to cross-reactive antigenic determinants would result in strong selective pressures. This would give rise to the emergence of virions with antigenically different HAs, in the same way

as detected *in vitro* when cloned influenza virions are suppressed with monoclonal antibodies and antigenic mutants emerge at a rate of approximately 1 in 10^5 parental virions (Yewdell, Webster & Gerhard, 1979). But caution has to be shown in deductions made from epidemiological studies and especially with a virus such as influenza where, for example, it is known that pressures can be exerted from the time of first isolation of the virus in cells or eggs (Schild *et al.* 1983).

As a final example of application of biochemical analysis to RNA viruses, a number of field isolates of polioviruses type I and type III have been analysed by SDS PAGE and oligonucleotide mapping. The latter technique is of particular value because isolates of attenuated virus vaccine origin can be distinguished from 'wild' strains. The genomes of both vaccine strains and epidemic strains evolve at measurable rates during human intestinal passage (Nottay *et al.* 1981) and oligonucleotide mapping has been used by the above group of workers to estimate molecular evolution of type I poliovirus during a single epidemic. In April 1978 paralytic polio occurred in the USA in a non-immunized religious community, apparently transmitted from the Netherlands to Alberta and Ontario and thence to Pennsylvania. The last case was in June 1978, and the total number of reported cases was 133, with 101 paralytic cases and 32 cases of aseptic meningitis. Comparison of the maps of the earliest and last virus gave an estimate of the minimum rate of genome evolution, which was estimated to be about 1–2 % of the genome.

In contrast to the marked biochemical similarities of the above closely related epidemic strains of polio, oligonucleotide maps of epidemiologically unrelated strains were distinct. A large number of distinguishable polio type 1 strains exist worldwide, as shown by oligonucleotide fingerprinting. Isolates from different epidemics within the same country showed no genetic relationship to each other. Also contemporary isolates from neighbouring countries with endemic polio generally appeared different. During a short period several genetically unrelated type I strains were found in the same country. Even in geographically small countries such as Nicaragua more than one epidemic strain could be isolated in one year (Nottay *et al.* 1981).

Similar studies applying the oligonucleotide fingerprinting method have been carried out in our institute with type III polio viruses, since the latter are often the cause of the few cases of polio in populations with a high degree of immunity (Minor, 1980; Minor *et al.* 1982). The maps of three wild viruses isolated before the use of oral polio vaccine showed the viruses to be genetically distinct and completely different from the vaccine pattern. A virus strain (119) was characterized biochemically in some detail, since this was isolated from the central nervous system of a fatal case of poliomyelitis and has been shown to be neurovirulent for monkeys. The fingerprints showed the virus to be vaccine-like. These data suggest a vaccine origin of the infection, but also indicate that changes in virulence are not necessarily detectable in the oligonucleotide map. Presumably mutations in this case could have occurred in the lower-molecular weight oligonucleotides.

A versatile analytical technique for investigations of molecular epidemiology of DNA viruses such as herpes virus has been the use of restriction endonucleases for DNA fingerprinting. With such genome complexity even a minimal degree of mutation would be enough to create a complicated genetic heterogeneity in these

herpes viruses (Chaney *et al.* 1983). Because of the association of HSV-1 or 2 with a variety of clinical disorders ranging from facial and genital lesions to encephalitis it is of considerable interest to determine if specialized substrains of virus exist. From a technical point of view several restriction endonuclease enzymes must be used in conjunction to obtain a series of digests for comparison. In an early study more than 150 HSV-1 and HSV-2 isolates from different clinical and geographic origins were analysed by Buchman *et al.* (1980). Electrophoretic separation of fragments of DNA (cleaved by using restriction endonucleases) of related and unrelated isolates revealed that repeated isolates from a single site in a single individual (isolates more than 12 years apart) or from sexual partners were indistinguishable. In addition, epidemiologically unrelated isolates were easily distinguishable from each other. These authors concluded that mutations occurred during the evolution of HSV and that virus variants accumulated and persisted in the population, presumably without displacing each other. Since HSV viruses are latent and only multiply periodically it can be assumed that intimate contact is required for cross-infection. Viral variants would tend to be perpetuated within clusters of individuals defined by the existence of close personal contact. In a similar type of study, Lonsdale *et al.* (1980) analysed the DNA of 44 ganglion isolates of herpes virus from 21 individuals. Multiple isolates from a single ganglion (6 cases), isolates from different ganglia from a single individual (7 cases) and isolates from left and right ganglia from a single individual (7 cases) were the same virus strain. It would appear that the latent virus in ganglia obtained from the same individual comes from the same initial infection. A study of the restriction enzyme profiles of lip lesion obtained from four individuals on two occasions showed that the virus isolate from each individual could be distinguished. Variations in the restriction endonuclease profiles were not confined to the loss or gain of restriction sites, since variability was also observed in the mobility of certain restriction fragments which could have arisen due to sequence translocations or to local sequence inversions around an asymmetrically located restriction site.

The usefulness and validity of the restriction endonuclease technique for analysis of the DNA of varicella zoster isolates is described by Straus *et al.* (1983). All epidemiologically distinct isolates were shown to be unique by DNA analysis whilst two isolates recovered from members of a family infected in a common source outbreak were identical to each other but distinct from other strains. Furthermore, three isolates recovered at different times during the course of a single episode of zoster in an individual were identical but distinct from other isolates.

The new applied scientific discipline of 'molecular epidemiology' will undoubtedly contribute to our understanding of how viruses spread and change in their attempts to avoid barriers to transmission. In turn this may expose weak links which can be exploited for the development of novel strategies of virus control.

J. S. Oxford
National Institute for Biological Standards and Control,
Holly Hill,
Hampstead.
London NW3 6RB

REFERENCES

AARONSON, R. P. YOUNG, J. F. & PALESE, P. (1982). Oligonucleotide mapping: evaluation of its sensitivity by computer-simulation. *Nucleic Acids Research* **10**, 237–246.

BIRRER, M. J., UDEM, S., NATHENSON, S. & BLOOM, B. R. (1981). Antigenic variants of measles virus. *Nature* **293**, 67–69.

BUCHMAN, T. G., SIMPSON, T., NOSAL, C. & ROIZMAN, B. (1980). The structure of herpes simplex virus DNA and its application to molecular epidemiology. *Annals of the New York Academy of Science* **354**, 279–290.

CHANEY, S. M. J., WARREN, K. G., KETTYLS, J., ZBITNUE, A. & SUBAK-SHARPE, J. H. (1983). A comparative analysis of restriction enzyme digests of the DNA of herpes simplex virus isolated from genital and facial lesions. *Journal of General Virology* **64**, 357–371.

DIMITROV, D. H., GRAHAM, D. Y., LOPEZ, J., MUCHINIK, G., VELASCO, G., STENBACK, W. A. & ESTES, M. K. (1984). RNA electropherotypes of human rotaviruses from North and South America *Bulletin of the World Health Organization* **62**, 321–329.

ESPEJO, R. T., MUNOZ, O., SERAFIN, F. & ROMERO, P. (1980). Shift in the prevalent human rotavirus detected by ribonucleic acid segment differences. *Infection and Immunity* **27**, 351–354.

FLORES, J., PEREŻ, I., WHITE, L., PEREZ, M., KALICA, A. R., MARQUINA, R., WYATT, R. G., KAPIKIAN, A. Z. & CHANOCK, R. M. (1982). Genetic relatedness among human rotaviruses as determined by RNA hybridization. *Infection and Immunity* **37**, 648–655.

FOLLETT, E. A. C. & DESSELBERGER, U. (1983). Co-circulation of different rotavirus strains during a local outbreak of infantile gastroenteritis. Monitoring by rapid and sensitive nucleic acid analysis. *Journal of Medical Virology* **11**, 49–52.

FOLLETT, E. A. C., SANDERS, R. C., BEARDS, G. M., HUNDLEY, F. & DESSELBERGER, U. (1984). Molecular epidemiology of human rotaviruses. Analysis of outbreaks of acute gastroenteritis in Glasgow and the west of Scotland 1981/82 and 1982/83. *Journal of Hygiene* **92**, 209–222.

GHENDON, Y., KLIMOV, A., GORODKOVA, N. & DOHNER, L. (1981). Genome analysis of influenza A virus strains isolated during an epidemic of 1979–1980. *Journal of General Virology* **56**, 303–313.

HOLLAND, J., SPINDLER, K., HORODYSIKI, F., GRABAU, E., NICHOL, S. & VAN DE POL, S. (1982). Rapid evolution of RNA genomes. *Science* **215**, 1577–1585.

LONSDALE, D. M., BROWN, S. M., LANG, J., SUBAK-SHARPE, J. H., KOPROWSKI, H., WARREN, K. G. (1980). Variations in herpes simplex virus isolated from human ganglia and a study of clonal variation in HSV-1. *Annals of the New York Academy of Science* **354**, 291–308.

MERRIL, C. R., SWITZER, R. C. & VAN KEUREN, M. L. (1979). Trace polypeptides in cellular extracts and human body fluids detected by two-dimensional electrophoresis and a highly sensitive silver stain. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 4335–4339.

MINOR, P. D. (1980). Characterisation of strains of type 3 poliovirus by oligonucleotide mapping. *Journal of General Virology* **59**, 307–317.

MINOR, P. D., SCHILD, G. C., FERGUSON, M., MACKAY, A., MAGRATH, D. I., JOHN, A., YATES, P. J. & SPITZ, M. (1982). Genetic and antigenic variation in type 3 poliovirus: characterization of strains by monoclonal antibodies and TI oligonucleotide mapping. *Journal of General Virology* **61**, 167–176.

NAKAJIMA, K., DESSELBERGER, U. & PALESE, P. (1978). Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. *Nature* **274**, 334–339.

NAKAJIMA, S., NAKAJIMA, K., TAKEUCHI, Y. & SUGIURA, A. (1980). Influenza surveillance based on oligonucleotide mapping of RNA of H1N1 viruses prevalent in Japan 1978–1979. *Journal of Infectious Diseases* **142**, 492–502.

NOTTAY, B. K., KEW, O. M., HATCH, M. H., HEYWARD, J. T. & OBIJESKI, J. F. (1981). Molecular variation of type I vaccine related and wild polioviruses during replication in humans. *Virology* **108**, 405–423.

OXFORD, J. S., ABBO, H., CORCORAN, T., WEBSTER, R. G., SMITH, A. J., GRILLI, E. A. & SCHILD, G. C. (1983). Antigenic and biochemical analysis of field isolates of influenza B virus: evidence for intra- and inter-epidemic variation. *Journal of General Virology.*, **64**, 2367–2377.

OXFORD, J. S., KLIMOV, A. I., CORCORAN, T., GHENDON, Y. Z. & SCHILD, G. C. (1984). Biochemical

and serological studies of influenza B viruses: comparisons of historical and recent isolates. *Virus Research*, **1**, 241–258.

PALESE, P. & ROIZMAN, B. (eds). (1980). *Genetic Variation of Viruses. Annals of the New York Academy of Science* **354**, 1–507.

RODGER, S. M., BISHOP, R. F., BIRCH, L., MCLEAN, B & HOLMES, I. H. (1981). Molecular epidemiology of human rotaviruses in Melbourne, Australia, from 1973 to 1979 as determined by electrophoresis of genome ribonucleic acid. *Journal of Clinical Microbiology* **13**, 272–278.

SCHILD, G. C., OXFORD, J. S., DE JONG, J. C. & WEBSTER, R. G. (1983). Evidence for host-cell selection of influenza virus antigenic variants. *Nature* (*Lond*). **303**, 706–709.

SCHNAGL, R. D., RODGER, S. M. & HOLMES, I. H. (1981). Variation in human rotavirus electropherotypes occurring between rotavirus gastroenteritis epidemics in Central Australia. *Infection and Immunity* **33**, 17–21.

SQUIRE, K. R. E., OSBURN, B. I. CHUANG, R. Y. & DOI, R. H. (1983). A survey of electropherotype relationships of Bluetongue virus isolates from the Western United States. *Journal of General Virology* **64**, 2103–2115.

STRAUS, S. E., HAY, J., SMITH, H. & OWENS, J. (1983). Genome differences among varicella-zoster virus isolates. *Journal of General Virology* **64**, 1031–1041.

WIKTOR, T. J. & KOPROWSKI, H. (1980). Antigenic variants of rabies virus. *Journal of Experimental Medicine* **152**, 99–112.

YEWDELL, J. W., WEBSTER, R. G. & GERHARD, W. U. (1979). Antigenic variation in three distinct determinants of an influenza A haemagglutinin molecule. *Nature* **279**, 246–248.