Epidemiol. Infect. (2002), **129**, 107–112. © 2002 Cambridge University Press
DOI: 10.1017/S0950268802006982 Printed in the United Kingdom

Partial VP1/2A gene sequence based molecular epidemiology of wild type 1 poliovirus isolates from some parts of India

R. ANAND[1], D. GHOSH[2], A. V. BHUPATIRAJU[2], S. BROOR[2], S. T. PASHA[1], S. KHARE[1], M. KUMAR[1], K. K. DUTTA[1] AND A. RAI[1]*

[1] *Division of Biochemistry and Biotechnology and Microbiology, National Institute of Communicable Diseases, Delhi, India*
[2] *Department of Microbiology, All India Institute of Medical Sciences, New Delhi, India*

(*Accepted 22 January 2002*)

SUMMARY

Genomic variability within the sequences of VP1/2A junction among polioviruses from across the globe has revealed the existence of several endemic genotypes and their epidemiological inter-relationships; but such data on Indian isolates are scanty. The present work was intended to ascertain the persistence and transmission pattern of different genotypes of wild type 1 polioviruses circulating in India. Forty-eight wild type 1 poliovirus isolates obtained from different parts of India during 1996–8 were subjected to RT–PCR and nucleotide sequencing using M13 tailed primers. A 293 base pair region was amplified and sequenced for genetic variation study. Considering the 15% divergence of the sequences from Sabin 1, the isolates from six different states of India confirmed a single dominant genotype 4. Phylogenetic analysis revealed the circulation and active inter-state transmission of many genetically distinct strains of wild poliovirus type 1 belonging to genotype 4. This warrants the need for insisting on more efficient surveillance mechanisms so as to assess the impact of an extensive pulse polio immunization programme in India.

INTRODUCTION

There is a global effort to eradicate poliomyelitis through extensive immunization and other integrated strategies. These initiatives are frequently hampered due to the endemicity of several circulating genetic subtypes, particularly in India and other developing countries. Active surveillance of circulating wild polioviruses in communities is considered crucial for the success of the polio eradication initiatives of the WHO.

Despite extensive vaccination coverage and a pulse polio immunization programme in India, the circulation of wild poliovirus strains remains unabated in many parts of the world [1–3]. Wild strains can also be transmitted rapidly to distant communities causing outbreaks far from the originating foci [4, 5]. Since a high proportion of poliovirus infection are frequently sub-clinical, evidence for epidemiological links must be obtained from the laboratory data. Although, conventional serologic methods are able to distinguish vaccine strains from those of wild ones [6], the potential of polioviruses to undergo rapid genetic variation during replication in humans [7–9], about 1–2 nucleotide substitution per week [10] resulting in antigenic variability raises doubts about the specificity of methods based on antigenic characterization. Other methods such as nucleic acid hybridization and RT–PCR also have some limitations. In recent years, poliovirus gene sequence analysis has become a strong molecular epidemiology tool to study the predominant lineages and ensure the elimination of wild poliovirus strains from the community. Nucleotide

* Author for correspondence: Division of Biochemistry and Biotechnology, National Institute of Communicable Diseases, 22, Shamnath Marg, Delhi: 110054, India.

sequence analysis can also help to recognize various strains, their epidemiological links, pathways of virus transmission, identify local reservoirs and unusual vaccine-derived wild polioviruses [4, 10, 15]. In fact, gene sequencing is now recommended as the ultimate method of certification of polio eradication [11].

The present study was undertaken to unveil the circulating genotypes in six different states of India. Forty-eight such isolates of wild poliovirus type 1 collected over a period of 3 years were subjected to nucleotide sequence relatedness study to ascertain the circulating strains and their inter-state transmission pattern.

MATERIALS AND METHODS

Isolation of virus

Poliovirus strains were isolated in HEp-2 cell line from stool specimens of patients of acute flaccid paralysis at the regional (WHO) reference laboratory of Microbiology division NICD. The isolates were typed as poliovirus type 1 by standard procedures [12, 13]. Intratypic differentiation was done by ELISA using cross-absorbed rabbit mono specific anti-sera as well as by probe hybridization on all the strains.

Primers for RT–PCR

A 293 bp was used as a target region for molecular epidemiological studies.

Upstream primer VP1 (nt 3235–3251)

5′ TGT AAA ACG ACG GCC AGT GTC AAT GAT CAC AAC CC 3′

Downstream primer 2A (nt 3508–3527)

5′ CAG GAA ACA GCT ATG ACC AAG ACG TCT CTA TTC CAC AT 3′ (M13 universal primer sequences underlined)

RT–PCR

The RNA was extracted from HEp-2 infected cultures by Trizol method (Gibco–BRL). cDNA was synthesized with 3 μl of the extracted RNA and 0·75 μM of downstream primer using the RT–PCR kit. (Applied Biosystems).

cDNA was further used for amplification with the same kit using the forward primer (0·15 μM) and the PCR profile was as follows: Initial denaturation: 94°/2′; 40 cycles of 94°/30 sec, 42°/45 sec, 60°/1′; final extension: 60°/7′. The PCR products were analysed on a 2% agarose gel stained with ethidium-bromide to quantitate and check for its integrity.

Sequencing of PCR products

The amplicons were purified by PCR purification kit (Qiagen) and sequencing reactions were performed using the di-deoxy chain terminator protocol as well as primer chemistry using either fluroscein dye labelled M13 primer or di-deoxy nucleotides (Applied Biosystems). The sequencing reaction products were then purified and electrophoresed in an automated DNA sequencer (ABI 373). Each product was sequenced in both the directions and appropriate kit controls were used with each reaction.

Phylogenetic analysis

Nucleotide sequences of 293 bp fragment from the VP1-2A region (nt 3235–3527), of the 48 isolates were changed to appropriate format using the SeqED facility on DNASTAR. The 150 nt stretch out of 293 bp sequence in the VP1-2A junction region in case of each of the 48 strains obtained from 6 different states of India were aligned together and compared with polio type 1 Sabin reference strain. Published genotypes 1, 2, 3 and 4 sequences from different parts of the world and India were also used in this alignment.

Sequences were then aligned using Clustal W. Distance matrix was calculated using Kimura 2 parameter and the tree was constructed by Neighbour Joining method using MEGA Version 2.1 (Molecular Evolutionary Genetic Analysis). Published sequences from Romania (Genotype 2), Egypt (Genotype 3), Bulgaria (Genotype 1) have also been considered as out-groups. Genotype 4 sequences reported earlier were also used in this alignment (kindly provided by Dr Olen Kew, CDC, Atlanta).

RESULTS

The micro-neutralization typing revealed that the present isolates under study were all poliovirus wild type 1 isolates. This was also confirmed by the method of probe hybridization [12].

The sequence analysis revealed that all the Indian isolates included in this study had a divergence of more than 15% from the Sabin strain. However, they showed a 12–14% divergence from the isolates (1600IND82, 955SRL87) of reported genotype 4

https://doi.org/10.1017/S0950268802006982 Published online by Cambridge University Press

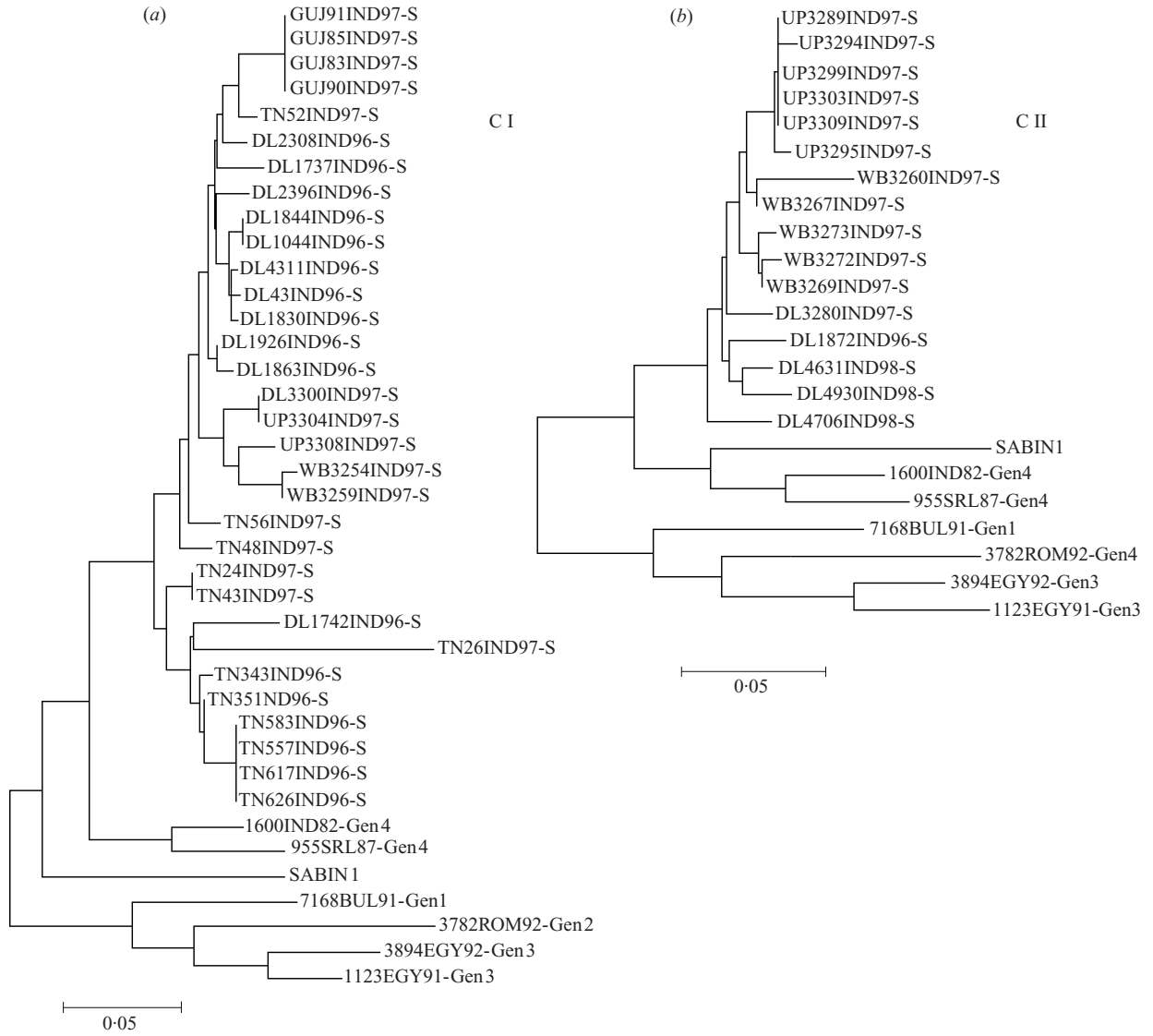

Fig. 1. Dendrogram of 48 wild type 1 polioviruses from some parts of India, isolates from other regions of the world and the Sabin 1 vaccine strain. Trees were constructed according to their sequence relatedness across the interval of nt 3235–3527 (VP1/2A region). Isolates are identified by two-letter state code, laboratory number, three-letter country code, and year of isolation. The Indian isolates included in the present study have been designated with a code – 'S' and the isolates representing different genotypes are designated as 'Gen1', 'Gen2', 'Gen3', 'Gen4'. Sequences were aligned using Clustal W. Distance matrix was calculated using Kimura 2 parameter and the trees were constructed by Neighbour Joining method using MEGA version 2.1 (Molecular Evolutionary Genetic Analysis). For ease of comparison, the dendrogram is split into two clusters referred as cluster I and cluster II (CI and CII as shown in Fig. 1*a*, *b*, respectively). Country abbreviations: BUL, Bulgaria; ROM, Romania; EGY, Egypt; SRL, Srilanka; IND, India. Abbreviations of Indian states: UP, Uttar Pradesh; WB, West Bengal; TN, Tamil Nadu; GJ, Gujarat, DL, Delhi. Sabin strain is designated as Sabin 1.

sequences [15]. The sequences of the present isolates as apparent from Figure 1*a* and 1*b* are grouped into two distinct clusters (CI and CII) within the genotype 4. The isolates in CI are 7–12% divergent from (1600IND82, 955SRL87) the known genotype 4 sequences. Similarly, the isolates in CII cluster are also 8–13% divergent from the above known sequences of genotype 4. The maximum divergence observed between CI and CII was found to be 12%.

From the dendrogram of cluster CI (Fig. 1*a*) it is evident that there were two distinct foci in 1996; one of which originates from Delhi (DL) and the other from Tamil Nadu (TN). The divergence observed among the isolates from two different foci within this cluster was 4%. The isolates of Gujarat (GJ), Uttar Pradesh (UP), and West Bengal (WB) which were observed in 1997 were found close to the 1996 Delhi isolates.

https://doi.org/10.1017/S0950268802006982 Published online by Cambridge University Press

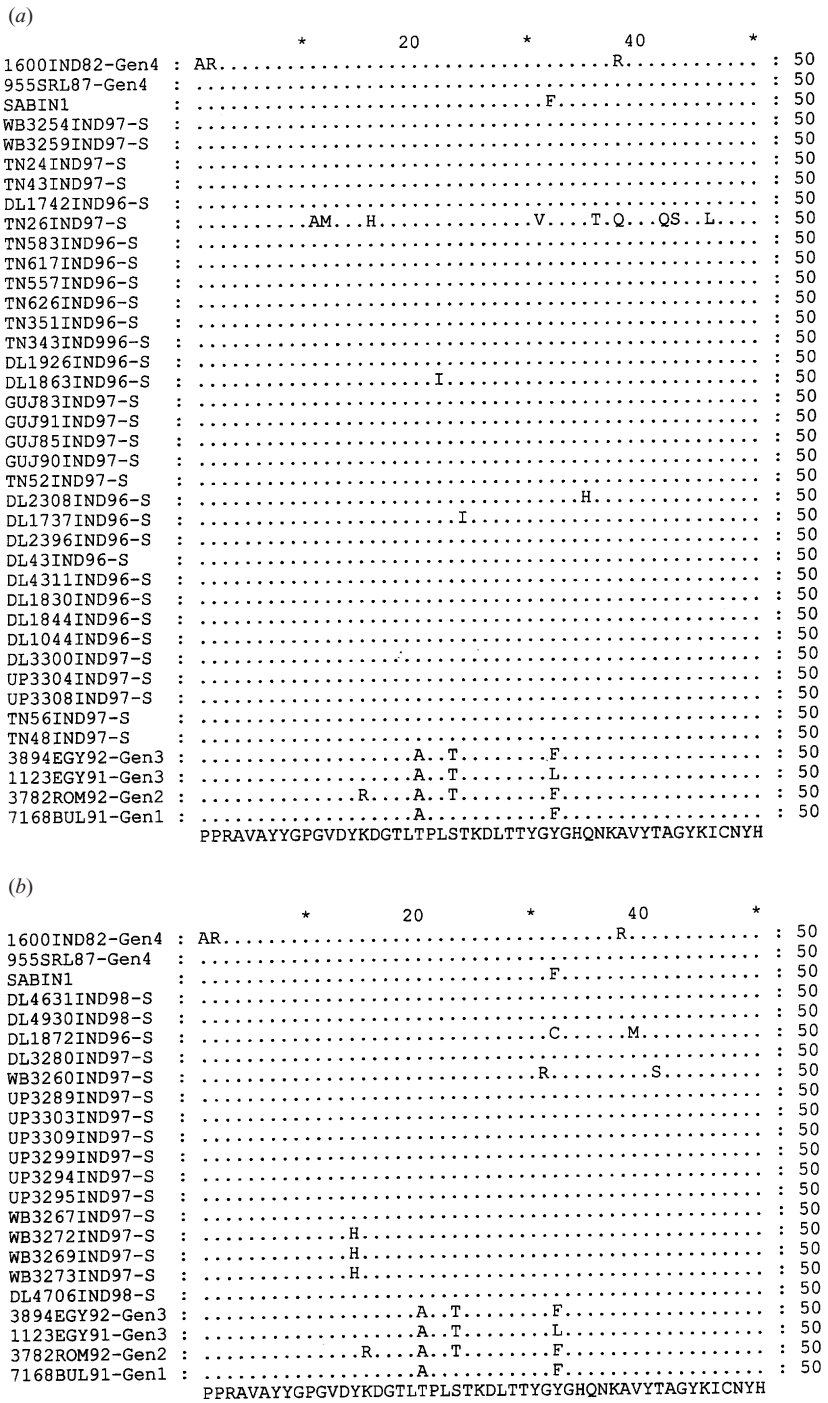

Fig. 2. Amino acid sequences obtained from the 150 bases of nucleotide sequence used for determination of genetic relatedness amongst the poliovirus type 1 isolates. The code IND represents India. The consensus sequence of amino acids is represented in the bottom line. The consensus sequence implies the sequence obtained after aligning the sequences of different isolates followed by averaging of sequences and is hypothetical. (*a*) amino acid sequences of the isolates within cluster I (CI), (*b*) amino acid sequences of the isolates within cluster II (CII).

The isolate TN52IND97 was found to be similar to the Delhi isolates of 1996 and the Gujarat isolates of 1997. Similarly, isolates TN56IND97 and TN48-IND97 were also similar to the Delhi isolate of 1996 and the UP isolate of 1997. On the other hand the Delhi isolate DL1742IND96 was similar to the Tamil Nadu isolates. This may represent inter-state transmission. The Cluster II (Fig. 1*b*) includes isolates mainly obtained in 1997, 1998 from Delhi, West Bengal and Uttar Pradesh. These isolates are similar

https://doi.org/10.1017/S0950268802006982 Published online by Cambridge University Press

to the isolate DL1872IND96. It thus appears that these strains originated in 1996 but became predominant in 1997 and 1998.

The nucleotide substitutions observed were mostly transitions and were silent changes except for the isolate TN26IND97. This isolate showed nine amino acid changes as depicted in Figure 2*a*.

DISCUSSION

Sequence analysis is now considered as the ultimate method to ascertain the extent of poliovirus eradication in the endemic communities because of its inherent potential to provide important clues in establishing epidemiological links between the cases of poliomyelitis. Poliovirus genome evolves rapidly ($\sim 10^{-2}$ nt substitutions/site/year) during replication in humans, therefore the resolving power of molecular studies based on nucleotide comparisons is very high. Most nucleotide substitutions are silent mutations, not causing any change in the amino acid composition [10]. The estimation of an epidemiological link is based on this rapid evolution of poliovirus genome by studying the pattern of nucleotide variation.

In the present study VP1/2A junction sequence was chosen to study the sequence diversity on Indian isolates because a large database on almost all the known genotypes of poliovirus isolates from different parts of the world is now available for comparison [14, 15]. Each serotype contains several distinct genotypes, which are defined as groups of polioviruses sharing > 85% nucleotide similarity with the VP1/2A interval. At least four major genotypes have been found circulating of which genotypes 1 and 2 are found mainly in Europe and Central Asia and the other two predominantly in the Eastern Mediterranean and Southeast Asia (genotypes 3 and 4).

Thus, the genotypes are shown to have geographic clustering. This clustering has important implications for polio eradication, indicating that local conditions (especially vaccine coverage), rather than specific properties of each virus (such as antigenicity), largely determine the extent of endemicity. While long-range importations do occur, most transmissions are between adjacent areas; and pandemic genotypes have not been reported [10].

Sequences of all the 48 isolates (1996–8) included in this study revealed a clear divergence of > 15% in relation to the reference Sabin type 1 strain but had 12–14% divergence from the known genotype 4 sequences (1600IND82, 955SRL87).

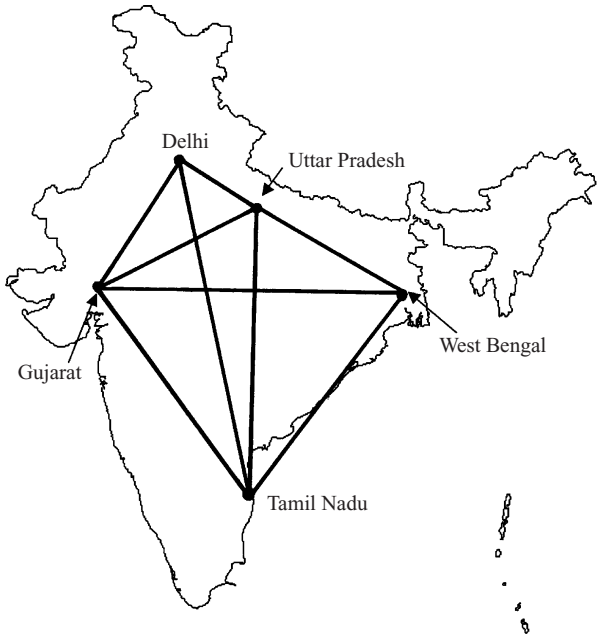

Fig. 3. Map of India showing inter-state transmission of the wild poliovirus type 1 isolates circulating in different parts of the country.

It is evident from the dendrogram generated on the basis of sequences of 48 isolates (Fig. 1) that the genotype 4 is the major genotype circulating in India. However, clustering of isolates of some states like those from West Bengal, Uttar Pradesh, Delhi suggests that circulation of the virus is through the displacement of people carrying the virus within and between the states. Air traffic between the metropolitan cities of India due to business and the corporate sector have also contributed to the displacement of virus between the states. For example, some of the Delhi isolates were genetically closer to the Tamil Nadu isolates and thus could have originated from Tamil Nadu or vice-versa.

It appears from the dendrogram that all the isolates in this study in circulation had a common pool as the isolates in both the clusters were close to either the 1996 isolates of Delhi or Tamil Nadu (Fig. 2*a*, *b*). The Gujarat, Uttar Pradesh and West Bengal isolates were close to the 1996 isolates of Delhi in CI. Also the Uttar Pradesh and West Bengal isolates in CII cluster were close to one of the 1997–8 isolates of Delhi, thereby indicating evidence of inter-state transmission. A far-flung inter-state transmission is also observed as some of the Delhi isolates of 1996 were found to have close similarity with the Tamil Nadu isolates and vice versa. This may be due to air traffic between the metropolitan cities of India due to business (Fig. 3). The 1998 isolates of Delhi were only

https://doi.org/10.1017/S0950268802006982 Published online by Cambridge University Press

2% divergent from one of the 1996 isolate showing a close epidemiological link in terms of origination of strains in 1996.

Little is known about broad patterns of poliovirus transmission in regions with high population densities. In India, the situation is further complicated by the proximity of several other large polio-endemic countries. A more comprehensive view will require the examination of many more isolates from different states of India.

The persistence and inter-state transmission of wild type 1 polioviruses despite intensive pulse polio immunization programme in India points to the lacunae in the programme both in terms of vaccination coverage and or proper storage of the vaccine stocks during the maintenance of the cold chain. In either case, it is legitimately felt that the poliovirus eradication programme in India needs to be strengthened further and a stringent active surveillance mechanism be developed so as to achieve the target of polio eradication in a reasonable time frame.

ACKNOWLEDGEMENTS

We thank the Council of Scientific and Industrial Research (CSIR) for providing partial financial support for this study as both Dr D. Ghosh (Senior Research Associate) and Mr A. V. Bhupatiraju (Senior Research Fellow) were supported by CSIR. We also thank Dr Olen Kew (CDC, Atlanta) for providing us with the sequences of the known genotypes used in this analysis. We thank Mr M. L. Regis and Mrs Meena Datta, Microbiology Division, National Institute of Communicable Diseases, for providing technical help in this study.

REFERENCES

1. Assaad F, Lungaras-Esteves K. World overview of poliomyelitis: regional patterns and trends. Rev Infect Dis 1984; **6**: S302–S307.
2. John TJ. Poliomyelitis in India: prospects and problems of control. Rev Infect Dis 1984; **6**: S439–S441.
3. Schaap GJP, Bukerk H, Coutinho RA, Kapsenberg JG, van Wezel AL. The spread of wild poliovirus in the well vaccinated Netherlands in connection with the 1978 epidemic. Prog Med Virol 1984; **29**: 124–40.
4. Kew OM, Nottay BK. Molecular epidemiology of polioviruses. Rev Infect Dis 1984; **6**: S499–S504.
5. Hovi T, Cantell K, Houvilianen A, Kinunnen E, et al. Outbreak of paralytic poliomyelitis in Finland: widespread co-circulation of antigenically altered poliovirus type 3 in a vaccinated population. Lancet 1986; **i**: 1427–32.
6. Nakano JH, Hatch MH, Thieme ML, Nottay B. Parameters for differentiating vaccine derived and wild poliovirus strains. Prog Med Virol 1978; **24**: 178–206.
7. Nakano JH, Gelfand HM, Cole JT. The use of a modified Wecker technique for the serodifferentiation of type 1 polioviruses related and unrelated to Sabin's vaccine strain. II. Antigenic segregation of isolates from specimen's collected in field studies. Am J Hyg 1963; **78**: 214–26.
8. Crainic R, Couillin P, Blondel B, Cabau N, Boue A, Horodniceanu F. Natural variation of poliovirus neutralization epitopes. Infect Immun 1983; **41**: 1217–25.
9. Minor PD, Ferguson M, Evans DMA, Almond JW, Icenogle JP. Antigenic structure of polioviruses of serotypes 1, 2, and 3. J Gen Virol 1986; **67**: 1283–91.
10. Kew OM, Mulders MN, Lipskaya GY, da Silva EE, Pallansch MA. Molecular epidemiology of polioviruses. Sem Virol 1995; **6**: 401–14.
11. Kew OM. Genomic sequencing and epidemiologic tracking of polioviruses: Part I. Polio Lab Network 1999; **V**: 1–4.
12. World Health Organization. Manual for the virological investigation of poliomyelitis. Geneva: WHO: WHO/EPI/CDS/POLIO/90.1.
13. Grandien M, Forsgren M, Ehrnst A. Enteroviruses and reoviruses. In: Schmidt NJ, Emmons RW, eds. Diagnostic procedures for viral, rickettsial and chlamydial infections, 6th edn. Washington, DC: American Public Health Association, 1989: 513–69.
14. Mulders MN, Lipskaya GY, van der Avoort HGAM, Koopmans MPG, Kew OM, van Loon AM. Molecular epidemiology of wild poliovirus type 1 in Europe, the Middle East, and the Indian Subcontinent. J Infect Dis 1995; **171**: 1399–405.
15. Rico-Hesse R, Pallansch MA, Nottay BK, Kew OM. Geographic distribution of wild poliovirus type 1 genotypes. Virol 1987; **160**: 311–22.