

Evolution of Circulating Wild Poliovirus and of Vaccine-Derived Poliovirus in an Immunodeficient Patient: a Unifying Model

GENE V. GAVRILIN,¹ ELENA A. CHERKASOVA,¹ GALINA Y. LIPSKAYA,¹
OLEN M. KEW,² AND VADIM I. AGOL^{1,3*}

*A. N. Belozersky Institute of Physical-Chemical Biology, Moscow State University, Moscow 119899,¹ and
M. P. Chumakov Institute of Poliomyelitis and Viral Encephalitides, Russian Academy of
Medical Sciences, Moscow Region 142782,³ Russia, and National Center for
Infectious Diseases, Centers for Disease Control and
Prevention, Atlanta, Georgia 30333²*

Received 15 March 2000/Accepted 18 May 2000

We determined nucleotide sequences of the VP1 and 2AB genes and portions of the 2C and 3D genes of two evolving poliovirus lineages: circulating wild viruses of T geotype and Sabin vaccine-derived isolates from an immunodeficient patient. Different regions of the viral RNA were found to evolve nonsynchronously, and the rate of evolution of the 2AB region in the vaccine-derived population was not constant throughout its history. Synonymous replacements occurred not completely randomly, suggesting the need for conservation of certain rare codons (possibly to control translation elongation) and the existence of unidentified constraints in the viral RNA structure. Nevertheless the major contribution to the evolution of the two lineages came from linear accumulation of synonymous substitutions. Therefore, in agreement with current theories of viral evolution, we suggest that the majority of the mutations in both lineages were fixed as a result of successive sampling, from the heterogeneous populations, of random portions containing predominantly neutral and possibly adverse mutations. As a result of such a mode of evolution, the virus fitness may be maintained at a more or less constant level or may decrease unless more-fit variants are stochastically generated. The proposed unifying model of natural poliovirus evolution has important implications for the epidemiology of poliomyelitis.

Analysis of polioviruses isolated either during a given outbreak of poliomyelitis (30, 35, 51) or from sequential fecal samples from infected individuals (29, 31, 34) has revealed rapid changes in the nucleotide sequence of the viral 7.5-kb RNA. Oligonucleotide fingerprinting (51) and genome sequencing (30, 34) suggested that, during the epidemic, the nucleotide substitutions ranged from 1 to 2% per year. The molecular basis of such genetic instability, common to all RNA viruses, resides in a high error rate of the viral RNA-dependent RNA polymerases, estimated to be, for poliovirus, on the order of 10^{-4} to 10^{-5} substitutions per base per replication (12, 70, 71), and the absence of proofreading mechanisms for the correction of the errors made. The combination of these properties results in a high heterogeneity (the “quasispecies” nature) of all populations of the virus (8, 10, 11, 63).

Less understood, however, are the rules governing the fixation of specific mutations upon passages of a viral population. Generally speaking, a mutation may decrease or increase the level of virus fitness for a particular ecological niche or leave it unchanged. Nucleotide substitutions associated with changes in the “sense” of codons (nonsynonymous mutations) are less likely to be neutral than substitutions resulting in no change of the codon meaning (synonymous mutations). Fixation of mutations conferring a selective advantage is readily understood in the framework of positive Darwinian selection. Adverse mutations may be eliminated by negative selection.

The likelihood of fixation of a mutation depends not only on the associated changes in virus fitness but, to a large extent, on the sizes of propagated populations (48, 52–54). Consecutive

passages involving small viral populations, the situation very typical of the natural poliovirus infection, may well result in the accumulation of neutral or even debilitating mutations (5, 7, 13, 18), the phenomenon known as the Muller’s ratchet (5, 6, 8). Thus, the consecutive changes in viral lineages are not necessarily adaptive but may be due to random sampling. There are reasons to believe that the majority of generated mutations may in fact be neutral or nearly so (33). Since the same appears to be true of the fixed mutations (60), the rate of nucleotide changes should be approximately constant in time (33). This concept of a “molecular clock” was suggested to be valid for the evolution of RNA viruses (17, 24, 69) and is not infrequently used for estimations of the viral “age” (31, 65, 73).

Here, we undertook an analysis of genomic changes in two evolving poliovirus populations. One was represented by a lineage of wild-type (wt) viruses of T geotype (40) in circulation in the territory of the former Soviet Union (FSU) during 1991 to 1995. The other consisted of the strains successively isolated from an immunodeficient patient (IDP) with paralytic poliomyelitis (31). The obtained sequences revealed basic similarities in the character of accumulation of mutations in the two lineages, even though some quantitative differences were also evident. Different regions of the viral genome have evolved at nonidentical rates, and these rates might have changed during the history of a population. Nevertheless, the major contribution to the evolving populations came from the linear accumulation of predominantly synonymous substitutions. From the perspective of current theories of viral evolution (8, 24), the results obtained suggest a unifying model for the evolution of circulating poliovirus on the one hand and for virus growing in the gut of an individual on the other. This model has important implications for the epidemiology of poliomyelitis.

* Corresponding author. Mailing address: Institute of Poliomyelitis, Moscow Region 142782, Russia. Phone: 7 (095) 439 9026. Fax: 7 (095) 439 9321. E-mail: viago@ipive.genebee.msu.su.

TABLE 1. Poliovirus isolates studied

Strain ^a	Geographic location	Geotype ^b	Date of sampling (da/mo/yr, mo/yr or yr)
PV1/2677USA81, ^c day 23 major	Missouri, USA ^e		23/08/81
PV1/2677USA81, ^c day 48 major	Missouri, USA		17/09/81
PV1/3038USA81, ^c day 126 major	Missouri, USA		04/12/81
PV1/3038USA82, ^c day 158 major	Missouri, USA		05/01/82
PV1/3038USA82, ^c day 200 major	Missouri, USA		16/02/82
PV1/169AZB59	Baku, Azerbaijan		11/59
PV1/7TAJ91 ^d	Kurgan-Tyube, Tajikistan	T	02/08/91
PV1/4TAJ91 ^d	Kurgan-Tyube, Tajikistan	T	03/08/91
PV1/5341UKR92 ^d	Ukraine	T	11/92
PV1/5794UZB94 ^d	Kashkadaryinsk region, Uzbekistan	T	11/07/94
PV1/6011TAJ94 ^d	Tajikistan	T	17/11/94
PV1/6064CHE95 ^d	Chechnya	T	14/06/95
PV1/6484CHE95 ^d	Grozny region, Chechnya	T	17/08/95
PV1/6486CHE95 ^d	Grozny region, Chechnya	T	22/08/95
PV1/6490CHE95 ^d	Grozny region, Chechnya	T	08/95
PV1/6405ING95 ^d	Ingush Republic	T	28/09/95
PV1/6427ING95 ^d	Nasyr-Kort, Ingush Republic	T	14/10/95
PV1/827GEO85	Tbilisi, Georgia	T	05/02/85
PV1/832GEO85	Tbilisi, Georgia	T	12/11/85
PV1/434MOL91	Kishinev, Moldova	T	29/08/91
PV1/6070CHN94	China	T	1994
PV1/5990TAJ94	Shaartuz, Tajikistan	T	24/09/94
PV1/6013TAJ94	Kolkhozabad, Tajikistan	T	17/11/94
PV1/6430PAK95	Pakistan	T	1995
PV1/6433PAK95	Pakistan	T	1995
PV1/422RUS91	Inza, Russia	A	17/06/91
PV1/919GEO90	Zugdidi, Georgia	G	03/12/90
PV1/5937RUS94	Moscow, Russia	G	03/06/94
PV1/717RUS94	Moscow region, Russia	G	12/08/94

^a Each strain designation consists of the serotype, number given at isolation, and place and year of the corresponding polio case.

^b As shown in reference 40.

^c IDP lineage.

^d T geotype lineage.

^e USA, United States.

MATERIALS AND METHODS

Virus isolation and typing. Two sets of polioviruses of serotype 1 were analyzed (Table 1; Fig. 1). One contained 24 wt strains isolated in the territory of the FSU, China, and Pakistan from 1991 to 1995. Of these, 19 strains belonged to the T geotype (40). The other set included five vaccine-related strains (GenBank accession numbers for VP1 sequences, AF083933 to AF083937) obtained from an IDP with paralytic poliomyelitis from Missouri in 1981 (31).

Virus isolation from stool samples was done by standard methods (72). wt isolates were passaged once on monolayers of RD or HEP-2 cell lines and were typed in microneutralization tests with type-specific sera. The isolation and typing of vaccine-related viruses were described earlier (31).

Reverse transcription, PCR, and sequencing of poliovirus RNAs. RNA was extracted from cell lysates with Ultraspec 3 (Biotecx) and was reverse transcribed using random hexamer primers (Boehringer Ingelheim) with avian myeloblastosis virus reverse transcriptase (Promega). DNA copies of selected genomic portions were amplified by PCR with appropriate primers (the primer sequences are available upon request), and the PCR products were purified with the QIAquick system (Qiagen) and were directly sequenced either automatically (Applied Biosystems) or manually by using the Sequenase kit, version 2.0 (U.S. Biochemicals). The following regions were analyzed in all the 29 strains: the complete VP1 (906 nucleotides [nt]; nt 2500 to 3408), 2A (447 nt; nt 3409 to 3855), and 2B (291 nt; nt 3856 to 4146) genes as well as a portion of the 2C gene (51 nt; nt 4147 to 4197; called 2C'). A portion of the 3D gene (792 nt; nt 6601 to 7392; called 3D') in PV1/169AZB59, -/7TAJ91, -/5794UZB94, -/6484CHE95, -/6490CHE95, -/6405ING95, -/422RUS91, and -/919GEO90 was analyzed. The coordinates are given according to the consensus numbering system of Toyoda et al. (68).

The sequences of the following regions have been previously reported: the VP1/2A junction (150 nt; nt 3319 to 3368) of PV1/827GEO85, -/832GEO85, -/919GEO90, -/434MOL91, -/422RUS91, -/4TAJ91, and -/7TAJ91 (40) as well as the VP1 gene of the IDP strains (31).

Comparative analysis of nucleotide sequences. Multiple alignment of the determined sequences was performed by the program CLUSTAL W, version 1.74 (67). Phylogenetic analysis was carried out by the PHYLIP, version 3.5c, program package (20), PUZZLE, version 4.0 (64), and METREE (58). Analysis under maximum-likelihood optimality criteria was done with empirical base

frequencies and the HKY85 model of substitution (25). The following algorithms have also been used for testing the consistency of branching: maximum parsimony (16), neighbor joining (59), minimum evolution (57), and the unweighted average-linkage method of clustering. Evolutionary distances were estimated using the two-parameter method of Kimura (32) and the Tamura-Nei model (66). The reliability of clustering was tested by 10,000 iterations in the quartet-puzzling method (PUZZLE, version 4.0) and by 1,000 bootstrapping replicates in other methods mentioned above.

To assess the degrees of nonsynonymous and synonymous nucleotide divergence separately, the estimation of pairwise genetic distances was performed according to Li's algorithm (38). A site is considered fourfold degenerate if all possible changes at the site are synonymous; a site is twofold degenerate if one of the three possible changes is synonymous; at a nondegenerate site, all possible changes are nonsynonymous. The following parameters were calculated: the percentage of mutated synonymous sites among all synonymous sites (K_s), the percentage of mutated nonsynonymous sites among all nonsynonymous sites (K_a), the percentage of mutated sites among all sites (K_t), the overall number of mutations (N_t), the number of synonymous differences (N_s), and the number of nonsynonymous differences (N_a) (50). The rate of mutation fixation was calculated by regression analysis.

Statistical analysis of rare-codon conservation. A codon was defined as rare if its abundance was no more than 30% of that of its most abundant synonym in the poliovirus codon usage table. Five families of codons (encoding Ala, Arg, Pro, Ser, and Thr) contain codons meeting this criterion. The same codons were also scarce in humans. To eliminate the bias due to the close relatedness of some of the strains in our data set, the 6 most-diverged VP1 and 2AB sequences were selected out of 29 sequences available. Upon pairwise comparisons, these sequences exhibited divergences higher than 14 and 18%, respectively. The following strains were used in the analysis of both regions: PV1/169AZB59, -/7TAJ91, -/832GEO85, -/5937RUS94, and -/Day 200; in addition, the strains PV1/6070CHN94 and -/6013TAJ94 were used for the VP1 and 2AB sets, respectively. It was assumed that all of the sites mutated independently and that the mutation rate was uniform across the region in question. The ratios of the respective rare codons to all their synonyms in the whole poliovirus (Sabin 1) RNA were calculated. The Bernoulli formula was used to determine the probability of the occurrence of a certain number of rare codons (1 to 6) in the

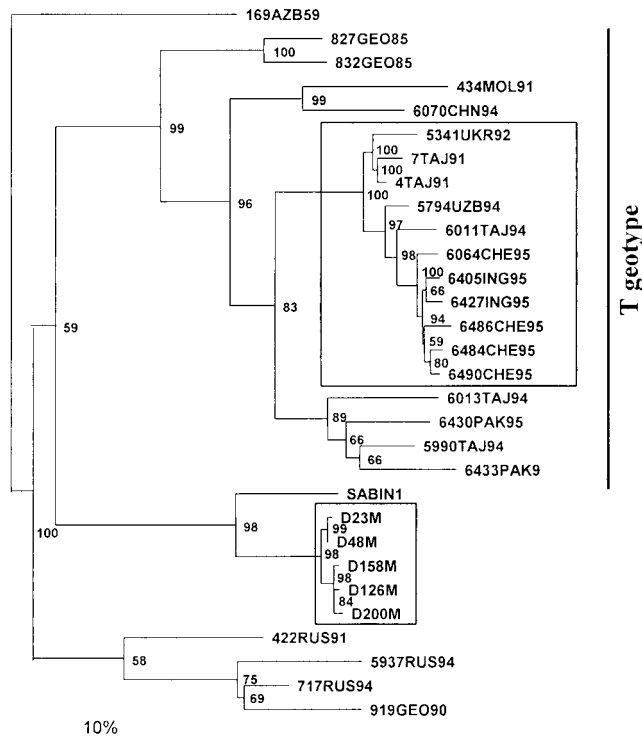


FIG. 1. The VP1/2ABC' maximum-likelihood bootstrap consensus tree (PUZZLE, version 4.0) rooted on PV1/169AZB59 showing the relationships of poliovirus type 1 wild and vaccine strains. The scale corresponds to 10% nucleotide divergence. The isolates belonging to the T geotype and IDP lineages are boxed.

positions of the six aligned sequences that corresponded to each of the above amino acid families. The positions having significant deviations ($P < 0.05$) from random codon usage were considered anomalously conserved.

Nucleotide sequence accession numbers. The nucleotide sequence data reported in this paper are available from the GenBank nucleotide sequence database under accession no. AF233098 to AF233222.

RESULTS

General characterization of the two poliovirus populations.

Two evolving poliovirus populations of serotype 1 were studied (Table 1). One corresponded to 11 wt strains isolated from poliomyelitis patients in the territory of the FSU in 1991 to 1995 and constituted a distinct cluster of the T geotype (Fig. 1). The other set of viruses was represented by five vaccine-related strains successively isolated from an IDP with paralytic poliomyelitis (31). The primary structures of the 150-nt VP1/2A region of the RNA of most of these T geotype isolates and of the VP1 genes of the vaccine-derived viruses have been determined previously (31, 40). Here, the sequenced portions were

extended to include the entire 1,694-nt VP1/2ABC' region (nt 2500 to 4197) of all of the viruses and a 792-nt portion of the 3D gene (nt 6601 to 7392) of all IDP isolates and five T geotype strains.

The selected cluster of T geotype strains was considered to represent a single lineage since they appeared to have a recent common ancestor, a strain imported to the FSU from Pakistan in 1991 (49). Indeed, the earliest T geotype strain of this lineage (PV1/7TAJ91) shared 148 out of 150 nt of the VP1/2A junction region with Pakistani strain PV1/18643PAK91 isolated the same year. The strains of this group were closely related to each other, having 92.3 to 98.8% nucleotide identity in the entire VP1/2AB region. All of them but one demonstrated a close relatedness in the 3D' region as well (97.9 to 99% identity). The 3D' region of the deviating strain, PV1/7TAJ91, differed from those of the other T geotype strains by 12.8 to 13.4%. It seems that a recombination event involving the 3'-terminal part of the genome took place in the history of the T geotype lineage. The 3D' sequence of PV1/7TAJ91 was excluded from the subsequent comparative analysis.

The isolates from the IDP exhibited a serotype 1 specificity and were closely related to the Sabin 1 virus in the primary structure of the VP1 gene (31). However, more-extensive sequencing demonstrated that at least a portion of their 2B and 2C genes was derived from the Sabin 2 genome. The crossover region appeared to map to a short sequence between positions 3973 and 3996, exhibiting similar degrees of relatedness to Sabin 1 and Sabin 2 genomes (Fig. 2). The 583-nt 2AB sequence upstream of the arbitrarily assumed crossover point (in the middle of the crossover region, that is, between nt 3983 and 3984) showed 7.0 and 17.9% divergences from the Sabin 1 and Sabin 2 sequences, respectively, and for the VP1 gene, the corresponding values were 9.7 and 34.1%. On the other hand, corresponding values were 20.5 and 10.4% for the remaining 155-nt portion of the 2B gene and the 51-nt 5'-terminal portion of the 2C gene, respectively. The 3D' region showed again a closer similarity to the Sabin 1 sequence (4.7%) than to the Sabin 2 sequence (11.7%). We concluded that the IDP isolates were represented by double (type 1/type 2/type 1) recombinants. The portions of the 2B gene upstream and downstream of the assumed crossover point were denoted 2B₁ and 2B₂, respectively.

Time course of genomic evolution of the poliovirus populations. The total number of accumulated mutations in the sequenced portion of the genomes was plotted against time. The time course of mutation accumulation could be well approximated by straight lines (the regression coefficient, r , was ≥ 0.95) corresponding to the rates of 1.19×10^{-2} and 2.62×10^{-2} substitutions/site/year for the T geotype and IDP lineages, respectively (Fig. 3A and D).

The majority of the detected mutations ($\sim 75\%$ in both populations) were synonymous substitutions. The linear ($r \geq 0.95$) rates of accumulation of synonymous substitutions per synon-

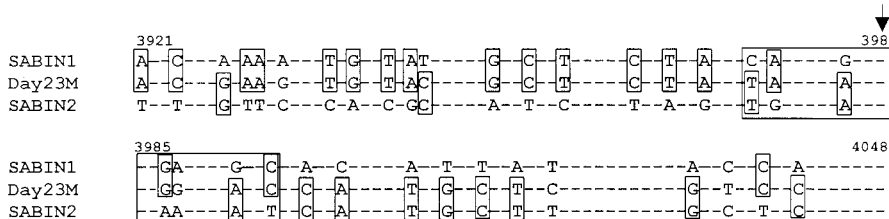


FIG. 2. The crossover region within the 2B gene of the strain Day23M. Dashes, identical bases in all three sequences; framed segment (nt 3973 to 3996), hypothetical recombination region; arrow, beginning of the 2B₂ region. Identical nucleotides are boxed.

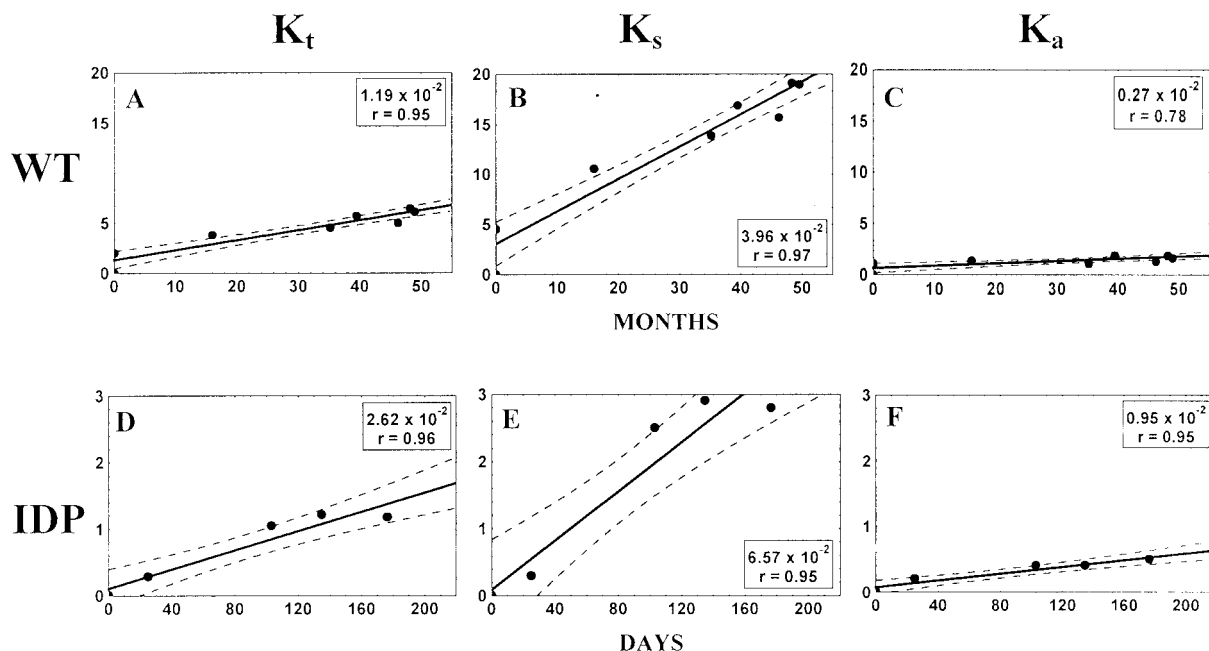


FIG. 3. Mutation fixation rates in the sequenced regions of 11 wt (VP1/2ABC') and 5 IDP (VP1/2ABC'/3D') isolates in substitutions per site per year (r , correlation coefficient) for all (K_t), synonymous (K_s), and nonsynonymous (K_a) sites. y axis, percentage of mutated sites. The position of the first strain in each of the lineages corresponds to $x = 0$ and $y = 0$. Dashed lines, 95% confidence intervals.

ymous site per year were 3.96×10^{-2} and 6.57×10^{-2} for the T geotype and IDP lineages, respectively (Fig. 3B and E). The corresponding parameters for nonsynonymous substitutions (per nonsynonymous site per year) were significantly lower, 0.27×10^{-2} and 0.95×10^{-2} , respectively (Fig. 3C and F).

The preponderance of synonymous substitutions among the fixed mutations and the linear character of mutation fixation were consistent with the notion that random sampling was a major factor contributing to the evolution of both the viral populations studied.

Nonuniform rates of evolution of different parts of the viral genome. The character of mutations accumulated was analyzed for the three genomic regions (VP1, 2AB, and 3D') separately (Fig. 4 and 5). The total and synonymous mutations accumulated linearly with time in all the studied regions of both lineages (with r varying between 0.89 and 0.97). However, in both sets of viruses, different RNA regions varied significantly in their relative mutabilities. The 2AB region evolved 2.5 to 5.5 times faster than the VP1 gene, and again a major contribution

to this difference came from synonymous substitutions, with the relevant rate for the 2AB region in the IDP strains being exceptionally high, 18.25×10^{-2} (Fig. 5E). No significant difference between the rates of evolution of the 2A and 2B genes studied separately was noted (not shown).

The rate of evolution of the portion of 3D' gene in the IDP lineage was found to be the lowest among the regions investigated, being more than two- and threefold lower than the rates of accumulation of total and synonymous substitutions, respectively, in the VP1 gene (Fig. 5G and H). The lowest mutability of the 3D' region was also noted for the T lineage sequences (data not shown), but due to the insufficient number of 3D' sequences available, a quantitative comparison with the other regions in this case seemed unwarranted.

The rates of fixation of nonsynonymous substitutions were in general significantly lower and less regular. For the T geotype, the relevant values for 2AB appeared to be higher than those for VP1 (Fig. 4C), but the reverse was true of the IDP population (Fig. 5C and F). The accumulation of nonsynonymous

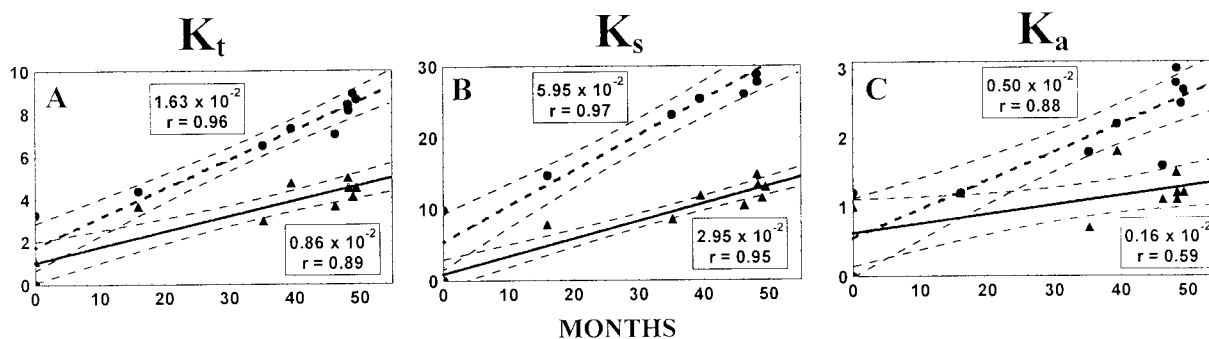


FIG. 4. Mutation fixation rates in the T geotype (11 isolates) lineage in all (K_t), synonymous (K_s), and nonsynonymous (K_a) sites. Solid and boldface dashed lines, data for VP1 and 2AB regions, respectively. For other details, see the legend to Fig. 3.

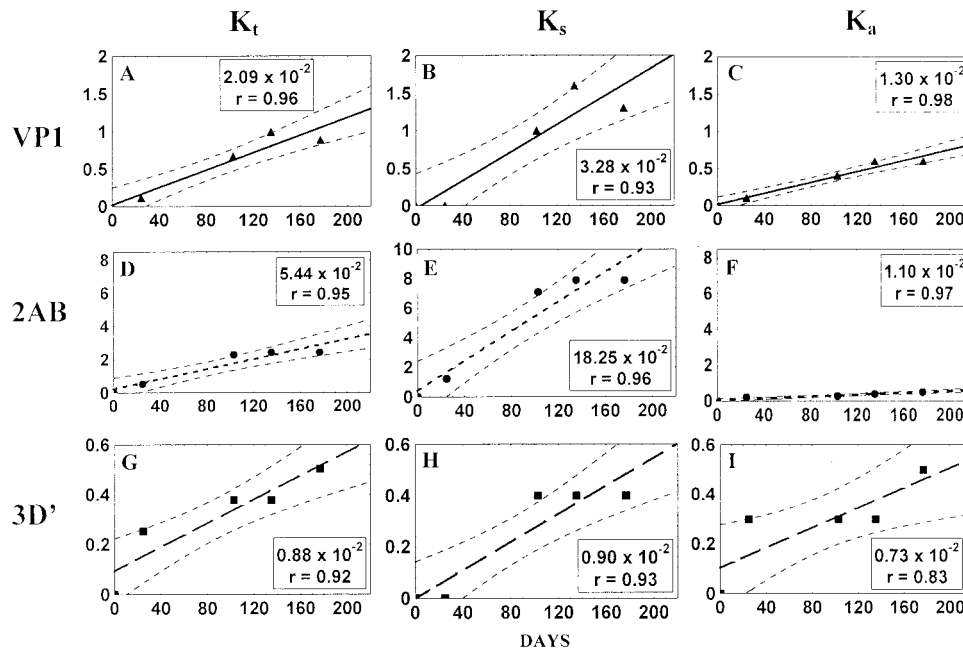


FIG. 5. Mutation fixation rates in the IDP lineage (five isolates) in all (K_t), synonymous (K_s), and nonsynonymous (K_a) sites in the VP1, 2AB, and 3D' regions. For other details, see the legend to Fig. 3.

mutations in the VP1 genes of the former population deviated from linearity ($r = 0.59$) (Fig. 4C).

Character of mutation fixation. In the vaccine-derived viruses, a significant proportion of mutations, once fixed, continued to be present in the genomes of all of the subsequent isolates. This was true of the mutations accumulated both before and after the isolation of the first strain (Fig. 6). Such a character of mutation accumulation was consistent with the assumption that the successive isolates of the IDP lineage were related to each other in the parent-progeny fashion. Such a direct relationship was less evident for the T genotype populations, where a significant proportion of mutations present in a given isolate was absent from some or all subsequent isolates (Fig. 6).

Remarkably, about half of the detected replacements in the VP1 and 2AB₁ regions of the IDP viruses resulted in the return to the Sabin 1 sequence, and such reversions were more or less evenly distributed over these regions (no reversions were observed in 3D' out of a total of three replacements). A smaller but significant (22%) proportion of reversions to Sabin 2 in the 2B₂ region was noted.

Codon usage and the nonuniform evolution of the viral genome. A possible explanation of different rates of fixation of synonymous mutations along the viral genome may be related to the codon usage bias. Different synonymous codons are known to be used with different frequencies in different organisms, different tissues, and even different genes, contributing thereby to the translational control of gene expression (61, 74). It can be argued that highly preferred codons are favorable for efficient translation and may therefore be less prone to be replaced during evolution. Consequently, the genomic regions with an abundance of such codons might be expected to undergo nucleotide replacements relatively less frequently. No significant heterogeneity in codon usage among the genomic regions studied could be found (not shown). From this analysis, however, a statistically significant conservation of rarely used codons at certain positions of the VP1, but not the 2AB,

region was noticed (Fig. 7A). In this analysis, not only the true conservation of a rare codon but also the cases in which a rare codon was replaced by another synonymous rare codon were taken into account. At certain positions, the consecutive conserved rare codons were clustered (Fig. 7A). In the vicinity of 9 out of the 15 conserved rare codons (at a distance of 1 to 3 codons) additional rare codons were also present in at least 75% of the 29 aligned VP1 sequences (not shown).

It should be noted, however, that the number of conserved rare codons in the VP1 gene was too small to account for the observed difference in the rates of synonymous replacements in the VP1 and 2AB regions.

Does the RNA secondary structure contribute to the non-uniform evolution of the viral genome? It could be hypothesized that local minima in the accumulation of synonymous substitutions were due to constraints imposed by the necessity to maintain specific high-order RNA structures. An analysis of this problem was hampered by the dynamic rather than static nature of the RNA structural organization. RNA regions with relatively stable secondary structure elements are expected to possess a low value of the so-called pairing number (P-num), a quantitative measure of the propensity of a certain base to pair with alternative partners in a collection of suboptimal folds (28, 56). We calculated that the proportions of bases with low P-num values (<100) for the VP1 and 2AB regions of the Mahoney genome were quite similar, being 91 out of 906 nt (10%) and 62 out of 738 nt (8.4%), respectively. This fact failed to support the notion that the two RNA regions differed from one another with regard to possessing strictly conserved secondary structures. No such support was obtained from attempts to derive a consensus secondary structure by comparative sequence analysis of these regions either. However, a higher percentage (17.7% or 140 out of 792 nt) of nucleotides with low P-num values was found in the 3D' region. This observation was indicative of substantial structural constraints, which may be partially responsible for lowering the mutation fixation rate in the sequenced portion of the polymerase gene.

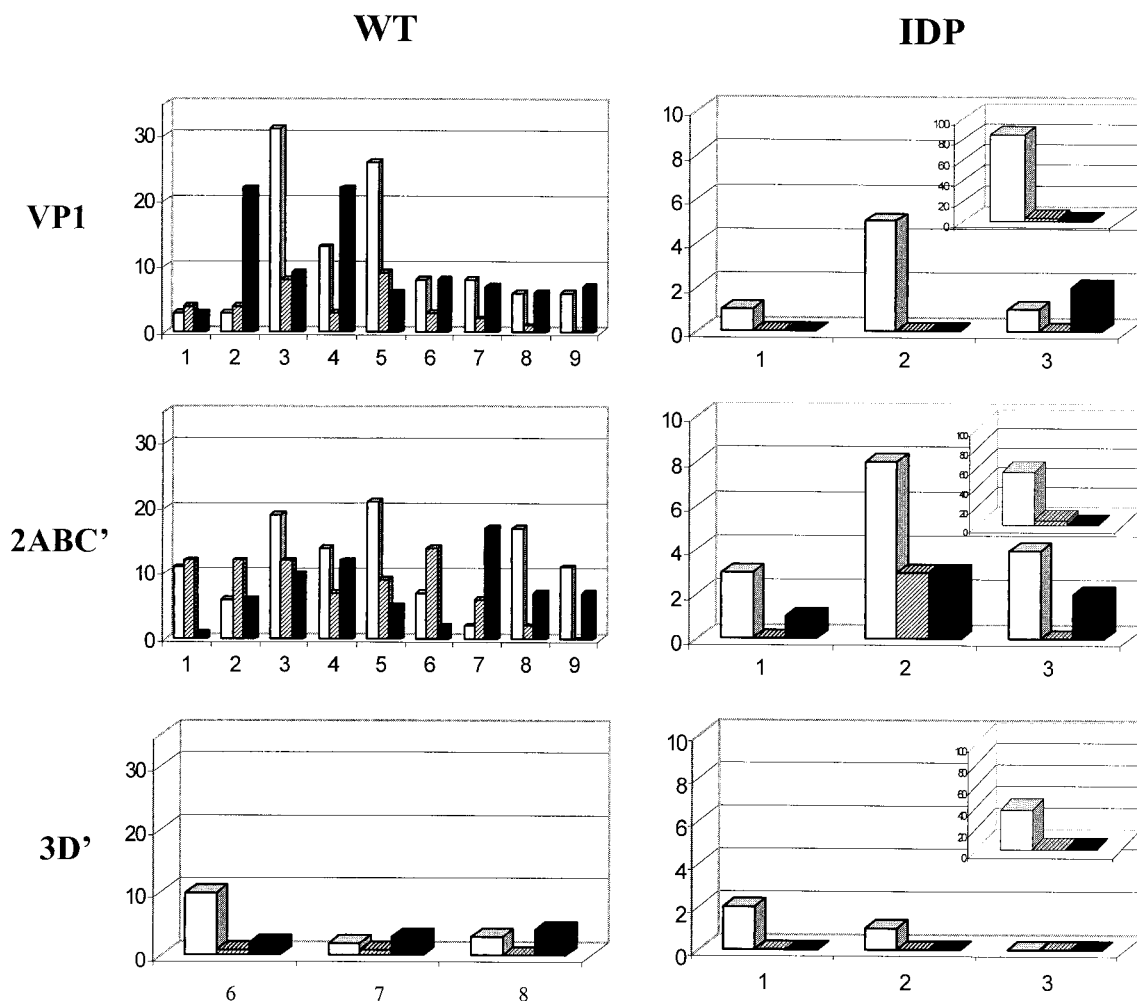


FIG. 6. Characterization of mutation fixation in the wt and IDP lineages. y axis, numbers of mutations that once appeared, became fixed in all (open bars) or at least some (hatched bars) of the subsequent isolates, as well as numbers of solitary mutations (black bars). Insets in the IDP panels, comparison of the first isolate (day 23) to the vaccine sequences. For the wt lineage, the numbers on the x axis correspond to the following isolates: 1, PV1/4TAJ91 (compared to PV1/7TAJ91); 2, PV1/5341UKR92; 3, PV1/5794UZB94; 4, PV1/6011TAJ94; 5, PV1/6064CHE95; 6, PV1/6484CHE95; 7, PV1/6486CHE95; 8, PV1/6490CHE95; and 9, PV1/6405ING95; for the IDP lineage, the numbers correspond to the following isolates: 1, day 48 (compared to day 23); 2, day 126; 3, day 158. The last isolates in the lineages are omitted.

The close relatedness of the 13 sequences available did not enable us to search for a consensus secondary structure of the 3D' region.

Deviation from the molecular clock mode of evolution. As shown above, the accumulation of total, synonymous, and non-synonymous mutations appeared to linearly depend on time in all the RNA regions studied with only one exception (nonlinearity of nonsynonymous substitutions in the T genotype VP1 gene). However, a more serious deviation from linearity was revealed by an analysis of mutation rates throughout the whole history of the Sabin vaccine-derived population. Indeed, the rate of synonymous substitutions in 2AB, during the time interval studied here, was approximately fivefold higher than that in VP1. On the other hand, the proportions of synonymous substitutions (K_s values) in the VP1 and 2AB₁ regions of the earliest IDP isolate, compared to the corresponding values for the Sabin progenitor virus, were similar, 30.4 and 27.9, respectively, suggesting that these regions evolved at comparable rates. This discrepancy became clearly evident when the apparent times of divergence of the relevant IDP lineage genes from the Sabin vaccine were calculated by extrapolations using

the estimated evolution rates taken from Fig. 5B and E. For the synonymous nucleotide replacements in the VP1 region (Fig. 8A), this time (9.6 years) corresponded reasonably well to the time (6.9 years) elapsed between the last oral polio vaccine vaccination and the onset of paralysis (31). However, the 2AB₁ region (Fig. 8B) appeared to require only one-fourth of the time (~2.5 years) to achieve the observed divergence from the progenitor. Possible reasons for this paradox are discussed below.

DISCUSSION

Basic mechanisms of poliovirus evolution. The rate and character of fixation of mutations generated by an error-prone viral RNA polymerase are controlled by positive selection of more-fit variants, negative selection of variants with decreased fitness, and accumulation of neutral or nearly neutral (33) replacements through random sampling of viral exemplars from the "swarm" of genomic variants (7, 9, 11, 13, 14, 18). Admittedly, the discussion of the nature of evolutionary changes would have been more factual if accompanying alterations in

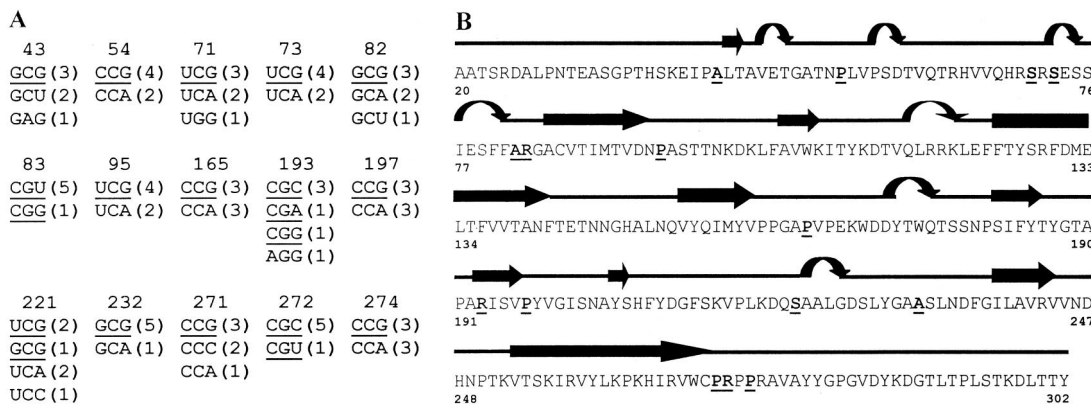


FIG. 7. Conserved rare codons of the VP1 gene. (A) Sites with anomalous conservation of the rare codons in the six most divergent sequences of the type 1 poliovirus. The codons are numbered from the start of the VP1 gene. The rare codons are underlined. The number of each of the individual codons in a given position of the six sequences is shown in parentheses. (B) VP1 secondary structure wire plot (Protein Data Bank; www.biochem.ucl.ac.uk/bsm/pdbsum/2plv/main.html). The amino acids corresponding to the conserved rare codons in panel A are underlined and in boldface. Straight arrows, beta structural elements; curved arrows, alpha helices. The amino acids are numbered from the start of the VP1 gene.

the level of viral fitness were known (8). Unfortunately, no adequate animal model to assess the relative fitness of polioviruses in their natural niche, the human gut, is available. It seems likely, however, that all three of the above-mentioned mechanisms contributed, to a certain extent, to the changes registered in this study. Although the exact contribution of each of them cannot as yet be defined, the preponderance of synonymous replacements and the constant rates of mutation fixation suggest a major role for random sampling as the mechanism of neutral evolution. This appeared to be true of both poliovirus populations studied here, even though they differed from one another in several respects: (i) the starting populations were represented by wt (presumably relatively well adapted) and attenuated (less-fit) viruses; (ii) the wt lineage was interrupted by human-to-human transfers, whereas the other one corresponded to a continuously growing population in a single host; and (iii) accordingly, the wt isolates were hardly related to each other in the direct ancestor-progeny fashion, whereas this was almost certainly true of the IDP isolates.

The accumulation of a significant proportion of synonymous substitutions in the IDP population requires some specification. Too little is known about the mode of poliovirus reproduction in the human gut to explain how consecutive random sampling from a heterogeneous viral population can be achieved in this case. A relatively low content of infectious virus in the feces (on the order 3.0 to 6.5 log₁₀ 50% tissue culture infective doses/g [45]) makes it likely that either relatively few susceptible cells are available at each given moment or only a tiny minority of susceptible cells are effectively infected. If so, virus reproduction in the gut may be to some extent likened to the propagation of a virus by consecutive small-population passages, resulting in the accumulation of neutral and adverse mutations (7, 18). In line with this reasoning, polioviruses inducing neurological disease are believed to represent just stochastic variants of the heterogeneous population present in the gut (23). It is worthwhile to note that Muller's ratchet can be enabled only if the initial poliovirus variants are sufficiently optimized for their particular ecological niche, the gut, through adaptive selection. Previous studies indicated that selection against the attenuating mutations in the Sabin viruses did indeed occur at the initial steps of infection (15).

Note that the rates of evolution of the populations studied here fit well values, on the order of 10⁻² to 10⁻³ substitution/

site/year, observed for other picornaviruses (65, 69, 73) and other RNA-containing viruses (12, 24, 26).

Nonuniform rates of evolution of different genomic regions and different genomes. A nonuniform rate of mutation accumulation along the viral RNA is not unique for poliovirus. The bias in nonsynonymous mutation fixation in different genomic regions of different RNA viruses has been observed (24, 27) and can readily be explained by differences in constraints imposed on the structures of different viral proteins. More interesting, in the context of the present study, were unequal rates of accumulation of synonymous substitutions.

The nature of the selective pressure on synonymous sites is not clear. A higher rate of mutation fixation in 2AB than in VP1 and 3D' may be caused by the additional contribution of either positive selection directed to 2AB or negative selection directed to VP1 and 3D'. The slowing down of the evolution of a genomic region could be due to the need for conservation of certain sequences. In our case, these sequences were represented predominantly by synonymous codons. Hypothetically, conservation of certain rare codons could be important to preserve an optimal translation elongation rate and hence the

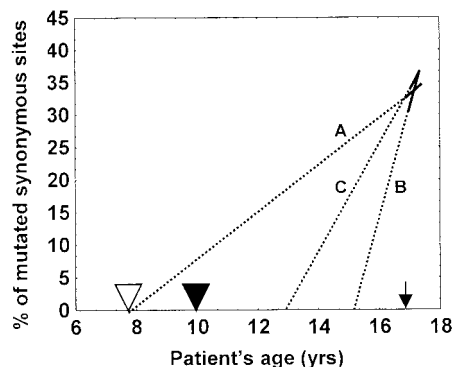


FIG. 8. Extrapolation of evolution rates of synonymous sites calculated for the IDP lineage for VP1 (A), 2AB₁ (B), and combined VP1/2AB₁/3D' (C) regions to zero divergence with Sabin 1. The solid upper portions of the lines correspond to the data taken from Fig. 5B and E and 3E. Open arrowhead, extrapolated age of initial infection as calculated by Kew et al. (31); solid arrowhead, time of the last vaccination with oral polio vaccine; arrow, time of the paralysis onset.

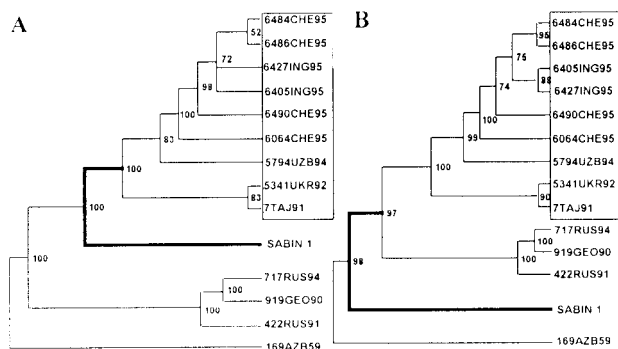


FIG. 9. Different topologies of the maximum-likelihood bootstrap consensus trees (PUZZLE, version 4.0) of the poliovirus type 1 wild and vaccine (Sabin 1) strains (rooted on PV1/169AZB59) when based on the VP1 (A) and 2AB (B) sequences. For the sake of brevity, a selected set of strains was used; a comparison of the complete trees also demonstrated the same difference in topology (not shown). Strains of the T geotype lineage are boxed. PV1/422RUS91 belongs to geotype A; PV1/919GEO90 and PV1/717RUS94 correspond to geotype G (40).

proper folding of nascent proteins (1, 36, 37, 55, 74). Statistically significant conservation of rare codons at certain positions of the VP1 gene (but not in the 2AB region) may be interpreted as indirect evidence for this hypothesis. Remarkably, the conserved loci with one or more rare codons tend to be positioned close to the borders of structured elements of VP1 (Fig. 7B), again consistent with their hypothetical role in proper folding. However, the number of such sites is too small to account for the difference between the evolutionary rates of VP1 and 2AB nucleotide sequences.

In addition to other mechanisms, conservation of some synonymous codons could be caused by the necessity to maintain the RNA structure as such. A relatively low variability in the subterminal coding regions of the hepatitis G virus RNA was explained by postulating the participation of these regions in the formation of structured replicative *cis* elements (62). Similar reasons may explain the extraordinary evolutionary stability of the 3D' region. The internal coding regions of picornavirus RNAs are also known to harbor *cis*-acting replicative signals, usually in the form of hairpin elements (41, 43). Such signals may well be present in the sequenced region of the poliovirus genome. However, they hardly occupy extended RNA regions and therefore could not make a significant contribution to the observed difference in the rates of evolution of VP1 and 2AB. On the other hand, the importance of RNA secondary structure may be related not only to certain conserved domains but also to the rule "fold as you please, but fold you must" (2). Implementation of this rule may result in negative selection of certain synonymous substitutions and hence in retardation of the evolutionary rate.

On the other hand, a relatively high rate of mutation fixation in the 2AB region could be caused by positive selection of certain substitutions (see below).

Regardless of the mechanism, local differences in the mutation fixation rates may result in the ambiguity of dendrograms illustrating the relatedness of the viral strains. Indeed, the dendrograms of type 1 viruses based on the VP1 and 2AB sequences are topologically different with regard to the position of the Sabin 1 strain (Fig. 9).

Disturbances in the operation of molecular clocks. Although the accumulation of mutations is generally linear, deviation from the molecular clock is possible. Mutations within the 2AB region of the IDP lineage appeared to accumulate nonlinearly

with time, the rate of evolution being apparently higher during the last period of the population history.

Although the nonlinearity of the viral evolutionary rates has been reported previously (17, 69), the underlying mechanism(s) remains obscure. The nonlinearity of 2AB evolution could hardly be explained by changes in the fidelity of the viral RNA polymerase or in the number of viral generations per time unit because of the local nature of the anomaly. One may speculate that the apparent increase in the rate of 2AB evolution was caused by the prior acquisition of an adverse mutation(s) within or outside this region. If so, the enhanced mutability of 2AB might reflect the necessity to mutually readjust different portions of the genome and thus was largely adaptive, even though it involved predominantly synonymous substitutions. For example, one may imagine that a recombination event destroyed (or destabilized) an extended RNA secondary structure element in the 2AB region, causing a concomitant drop of the fitness level. The selection pressure tended to restore the function, which could be achieved by stepwise accumulation of numerous slightly advantageous (nearly neutral) mutations.

On the other hand, the nonlinearity of 2AB evolution could be only apparent. For example, recombination with a bona fide Sabin strain could have occurred ~2.5 years before the isolation of the first IDP strain. As a result, the IDP lineage could have acquired an unaltered segment of the 2AB region which evolved rapidly but with a constant rate afterwards. An apparent deviation from linearity might be encountered if the time interval of observations is too long. Indeed, due to a high mutation rate, the probability of back mutations also becomes high, resulting in illusory "slowing down" of the accumulation of detectable mutations.

Given the real or apparent deviations from the molecular clock, one should be cautious in the determination of the viral age by back-extrapolating the rates of mutation fixation. The longer is the sequence used for the estimation of the mutation fixation rate, the more reliable should be the results of such extrapolations. However, even combining the sequences of the VP1, 2AB₁, and 3D' regions for such calculations appeared to result in an underestimated time of divergence of the IDP lineage from the Sabin vaccine (3.8 years) (Fig. 8C).

A hypothetical scenario of viral evolution in the immunodeficient host. The events that occurred in the IDP host between the vaccination and the isolation of the first virus are not documented. Indirect evidence suggests that the lineage studied might have become predominant not long before the first isolate (PV1/2677USA81) was obtained. First, the relevant fecal sample contained two variants, one of which, the minor variant (m), subsequently disappeared (31), suggesting that the major variant (M) exhibited better fitness. This selective advantage of M should have likely precluded the long-term coexistence of the two variants in a single host. Second, the paralytic disorder that developed in the IDP after several years of asymptomatic virus carriage was most likely due to a newly arisen more pathogenic variant.

Thus, the evolution of the Sabin virus in the immunodeficient host might occur according to the following hypothetical scenario. The vaccine virus evolved as a quasispecies, and a predecessor of the M variant diverged from the predominantly m population relatively shortly before the onset of paralysis (31) and existed as a minor component. Then a stochastic increase in fitness (9, 14, 19, 52) and hence in neurovirulence occurred in the M lineage due to either a mutation(s) (not necessarily in the sequenced region) or recombination (21, 22, 39). The newly emerged M variant eventually outcompeted the

then-predominant m variant and caused paralytic poliomyelitis.

Implications for epidemiology of poliomyelitis. We propose a unifying model of evolution of both wt poliovirus upon its natural circulation and of vaccine-derived virus replicating in the human gut. According to this model, evolution proceeds essentially through consecutive sampling of small portions of a viral population, ensuring predominant fixation of neutral and possibly adverse mutations. As a result the virus fitness is expected to be maintained at a more or less constant level or to decrease, except for the early stage after the immunization, when loss of some attenuating mutations (47) may result in an enhanced fitness. The possible accumulation of adverse mutations may serve to self-limit viral reproduction in the human gut and during person-to-person circulation and may contribute to the relatively low pathogenicity of poliovirus, which is known to cause overt disease only in a tiny minority of infected nonimmune subjects (44). The fitness decrease may be particularly conspicuous for attenuated viruses, leading to the very low level of their transmissibility. Some of the deleterious mutations may be eliminated by negative selection and by intratypic or intertypic recombination (4, 21, 22, 39, 46). Concurrently, the positive selection of rarely arisen more fit and more neurovirulent variants may facilitate the development of vaccine-associated poliomyelitis in the case of vaccine-derived variants and outbreaks of poliomyelitis in the case of circulating wt viruses.

By the time this work was submitted, two papers on prolonged poliovirus replication in IDP were published (3, 42). Although the authors have not addressed the main issues studied here, a comparison of some of their and our data is of interest. The mutation fixation rates calculated in these studies for the entire poliovirus genome were comparable to that found here for all studied genomic regions combined. Also, the divergences from Sabin 1 of the VP1 and 2AB regions in the last isolates found by Bellmunt et al. (3) and those in our lineages were quite similar, (9.2 versus 9.9% and 9.3 versus 8.1%, respectively). In spite of this, the difference in nucleotide sequences of the isolates from these two IDPs amounted to 14.8 and 13.2% for the VP1 and 2AB regions, respectively. Such a striking divergence may be interpreted as additional evidence for the largely neutral (nondirected) rather than adaptive character of the evolution of vaccine viruses in IDP hosts. It may be also added that when the VP1 region sequence of the most diverged isolate from the study by Bellmunt et al. (3) was added to the analysis of the rare codons, the anomaly of the conservation of all of the rare codons determined here remained statistically significant.

ACKNOWLEDGMENTS

We thank A. V. Alekseevsky, S. A. Spirin, and E. A. Korotkova for useful suggestions on statistical analysis and Mick N. Mulders for making the VP1/2A sequence of 18643PAK91 strain available. We also express our gratitude to the virologists from NIS countries for providing us with the poliovirus strains and O. E. Ivanova from the M. P. Chumakov Institute of Poliomyelitis and Viral Encephalitis for the propagation and typing work.

Financial support for a short-term visit of E.V.G. to the Centers for Disease Control and Prevention, Atlanta, Ga., is gratefully acknowledged. This work was supported by EU grants Copernicus and INTAS as well as by the Russian Foundation for Basic Research. V.I.A. is a Soros Professor.

REFERENCES

- Adzhubei, A. A., I. A. Adzhubei, I. A. Krashennnikov, and S. Neidle. 1996. Non-random usage of 'degenerate' codons is related to protein three-dimensional structure. *FEBS Lett.* **399**:78–82.
- Arora, R., C. Priano, A. B. Jacobson, and D. R. Mills. 1996. Cis-acting elements within an RNA coliphage genome: fold as you please, but fold you must!! *J. Mol. Biol.* **258**:433–446.
- Bellmunt, A., G. May, R. Zell, P. Pring-Akerblom, W. Verhagen, and A. Heim. 1999. Evolution of poliovirus type 1 during 5.5 years of prolonged enteral replication in an immunodeficient patient. *Virology* **265**:178–184.
- Cammack, N., A. Phillips, G. Dunn, V. Patel, and P. D. Minor. 1988. Intertypic genomic rearrangements of poliovirus strains in vaccinees. *Virology* **167**:507–514.
- Chao, L. 1990. Fitness of RNA virus decreased by Muller's ratchet. *Nature* **348**:454–455.
- Chao, L. 1997. Evolution of sex and the molecular clock in RNA viruses. *Gene* **205**:301–308.
- Clarke, D. K., E. A. Duarte, A. Moya, S. F. Elena, E. Domingo, and J. Holland. 1993. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. *J. Virol.* **67**:222–228.
- Domingo, E., E. Martinez-Salas, F. Sobrino, J. C. de la Torre, A. Portela, J. Ortin, C. Lopez-Golindez, P. Perez-Brena, N. Villanueva, R. Najera, S. VandePol, D. Steinhauer, N. DePolo, and J. J. Holland. 1985. The quasi-species (extremely heterogeneous) nature of viral RNA genome populations: biological relevance—a review. *Gene* **40**:1–8.
- Domingo, E., and J. Holland. 1997. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**:151–178.
- Domingo, E., C. Escarmis, N. Sevilla, and E. Baranowski. 1998. Population dynamics in the evolution of RNA viruses. *Adv. Exp. Med. Biol.* **440**:721–727.
- Domingo, E., C. Escarmis, L. Menendez-Arias, and J. J. Holland. 1999. Viral quasispecies and fitness variations, p. 141–161. *In* E. Domingo, R. Webster, and J. J. Holland (ed.), *Origin and evolution of viruses*. Academic Press, San Diego, Calif.
- Drake, J. W., and J. J. Holland. 1999. Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. USA* **96**:13910–13913.
- Duarte, E., D. Clarke, A. Moya, E. Domingo, and J. J. Holland. 1992. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. *Proc. Natl. Acad. Sci. USA* **89**:6015–6060.
- Duarte, E. A., I. S. Novella, S. Ledesma, D. K. Clarke, A. Moya, S. F. Elena, E. Domingo, and J. J. Holland. 1994. Subclonal components of consensus fitness in an RNA virus clone. *J. Virol.* **68**:4295–4301.
- Dunn, G., N. T. Begg, N. Cammack, and P. Minor. 1990. Virus excretion and mutation by infants following primary vaccination with live oral poliovaccine from two sources. *J. Med. Virol.* **32**:92–95.
- Eck, R. V., and M. O. Dayhoff. 1966. *Atlas of protein sequence and structure*. National Biomedical Research Foundation, Silver Spring, Md.
- Elena, S. F., F. Gonzalez-Candelas, and A. Moya. 1992. Does the VP1 gene of foot-and-mouth disease virus behave as a molecular clock? *J. Mol. Evol.* **35**:223–229.
- Escarmis, C., M. Davila, N. Charpentier, A. Bracho, A. Moya, and E. Domingo. 1996. Genetic lesions associated with Muller's ratchet in an RNA virus. *J. Mol. Biol.* **264**:255–267.
- Escarmis, C., M. Davila, and E. Domingo. 1999. Multiple molecular pathways for fitness recovery of an RNA virus debilitated by operation of Muller's ratchet. *J. Mol. Biol.* **285**:495–505.
- Felsenstein, J. 1993. PHYLIP: phylogeny inference package, version 3.52c. University of Washington, Seattle, Wash.
- Furione, M., S. Guillot, D. Otelea, J. Balanant, A. Candrea, and R. Crainic. 1993. Polioviruses with natural recombinant genomes isolated from vaccine-associated paralytic poliomyelitis. *Virology* **196**:199–208.
- Georgescu, M. M., F. Delpyroux, M. Tardy-Panit, J. Balanant, M. Combiescu, A. A. Combiescu, S. Guillot, and R. Crainic. 1994. High diversity of poliovirus strains isolated from the central nervous system from patients with vaccine-associated paralytic poliomyelitis. *J. Virol.* **68**:8089–8101.
- Georgescu, M. M., J. Balanant, S. Ozden, and R. Crainic. 1997. Random selection: a model for poliovirus infection of the central nervous system. *J. Gen. Virol.* **78**:1819–1828.
- Gojobori, T., E. N. Moriyama, and M. Kimura. 1990. Molecular clock of viral evolution, and the neutral theory. *Proc. Natl. Acad. Sci. USA* **87**:10015–10018.
- Hasegawa, M., H. Kishino, and K. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**:160–174.
- Imazeki, F., M. Omata, and M. Ohto. 1990. Heterogeneity and evolution rates of delta virus RNA sequences. *J. Virol.* **64**:5594–5599.
- Ito, T., O. Gorman, Y. Kawaoka, W. Bean, and R. G. Webster. 1991. Evolutionary analysis of the influenza A virus M gene with comparison of the M1 and M2 proteins. *J. Virol.* **65**:5491–5498.
- Jaeger, J. A., D. H. Turner, and M. Zuker. 1989. Improved predictions of secondary structures for RNA. *Proc. Natl. Acad. Sci. USA* **86**:7706–7710.
- Kew, O. M., B. K. Nottay, M. H. Hatch, J. H. Nakano, and J. F. Obijeski. 1981. Multiple genetic changes can occur in the oral poliovaccines upon replication in humans. *J. Gen. Virol.* **56**:337–347.
- Kew, O. M., M. N. Mulders, G. Y. Lipskaya, E. E. da Silva, and M. A.

- Pallansch. 1995. Molecular epidemiology of polioviruses. *Semin. Virol.* **6**:401–414.
31. Kew, O. M., R. W. Sutter, B. K. Nottay, M. J. McDonough, D. R. Prevots, L. Quick, and M. A. Pallansch. 1998. Prolonged replication of a type 1 vaccine-derived poliovirus in an immunodeficient patient. *J. Clin. Microbiol.* **36**:2893–2899.
 32. Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111–120.
 33. Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, United Kingdom.
 34. Kinnunen, L., A. Huovilainen, T. Poyry, and T. Hovi. 1990. Rapid molecular evolution of wild type 3 poliovirus during infection in individual hosts. *J. Gen. Virol.* **71**:317–324.
 35. Kinnunen, L., T. Poyry, and T. Hovi. 1992. Genetic diversity and rapid evolution of poliovirus in human hosts. *Curr. Top. Microbiol. Immunol.* **176**:49–61.
 36. Kolb, V. A., E. V. Makeyev, and A. S. Spirin. 1994. Folding of firefly luciferase during translation in a cell-free system. *EMBO J.* **13**:3631–3637.
 37. Kolb, V. A., E. V. Makeyev, A. Kommer, and A. S. Spirin. 1995. Cotranslational folding of proteins. *Biochem. Cell Biol.* **73**:1217–1220.
 38. Li, W. H., C. I. Wu, and C. C. Luo. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**:150–174.
 39. Lipskaya, G. Y., A. R. Muzychenko, O. K. Kutitova, S. V. Maslova, M. Equestre, S. G. Drozdov, R. P. Bercoff, and V. I. Agol. 1991. Frequent isolation of intertypic poliovirus recombinants with serotype 2 specificity from vaccine-associated polio cases. *J. Med. Virol.* **35**:290–296.
 40. Lipskaya, G. Y., E. A. Chervonskaya, G. I. Belova, S. V. Maslova, T. N. Kutateladze, S. G. Drozdov, M. Mulders, M. A. Pallansch, O. M. Kew, and V. I. Agol. 1995. Geographical genotypes (geotypes) of poliovirus case isolates from the former Soviet Union: relatedness to other known poliovirus genotypes. *J. Gen. Virol.* **76**:1687–1699.
 41. Lobert, P. E., N. Escrioni, J. Ruelle, and T. Michiels. 1999. A coding RNA sequence acts as a replication signal in cardiomyoviruses. *Proc. Natl. Acad. Sci. USA* **96**:11560–11565.
 42. Martin, J., G. Dunn, R. Hull, V. Patel, and P. Minor. 2000. Evolution of the Sabin strain of type 3 poliovirus in an immunodeficient patient during the entire 637-day period of virus excretion. *J. Virol.* **74**:3001–3010.
 43. McKnight, K. L., and S. M. Lemon. 1996. Capsid coding sequence is required for efficient replication of human rhinovirus 14 RNA. *J. Virol.* **70**:1941–1952.
 44. Melnick, J. L. 1996. Enteroviruses: polioviruses, coxsackieviruses, echoviruses, and newer enteroviruses. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), *Fields virology*, 3rd ed. Lippincott-Raven Publishers, Philadelphia, Pa.
 45. Melnick, J. L., and V. Rennick. 1980. Infectivity titers of enterovirus as found in human stools. *J. Med. Virol.* **5**:205–220.
 46. Minor, P. D., A. John, M. Ferguson, and J. P. Icenogle. 1986. Antigenic and molecular evolution of the vaccine strain of type 3 poliovirus during the period of excretion by a primary vaccinee. *J. Gen. Virol.* **67**:693–706.
 47. Minor, P. D., and G. Dunn. 1988. The effect of sequences in the 5' non-coding region on the replication of polioviruses in the human gut. *J. Gen. Virol.* **69**:1091–1096.
 48. Miralles, R., P. J. Gerrish, A. Moya, and S. F. Elena. 1999. Clonal interference and the evolution of RNA viruses. *Science* **285**:1745–1747.
 49. Mulders, M. N., G. Y. Lipskaya, H. G. A. M. van der Avoort, M. P. G. Koopmans, O. M. Kew, and A. M. van Loon. 1995. Molecular epidemiology of wild poliovirus type 1 in Europe, the Middle East and the Indian subcontinent. *J. Infect. Dis.* **171**:1399–1405.
 50. Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**:418–426.
 51. Nottay, B. K., O. M. Kew, M. H. Hatch, J. T. Heyward, and J. F. Obijeski. 1981. Molecular variation of type 1 vaccine-related and wild poliovirus during replication in humans. *Virology* **108**:405–423.
 52. Novella, I. S., E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. J. Holland. 1995. Exponential increases of RNA virus fitness during large population transmissions. *Proc. Natl. Acad. Sci. USA* **92**:5841–5844.
 53. Novella, I. S., S. F. Elena, A. Moya, E. Domingo, and J. J. Holland. 1995. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. *J. Virol.* **69**:2869–2872.
 54. Novella, I. S., J. Quer, E. Domingo, and J. J. Holland. 1999. Exponential fitness gains of RNA virus populations are limited by bottleneck effects. *J. Virol.* **73**:1668–1671.
 55. Oresic, M., and D. Shalloway. 1998. Specific correlations between relative synonymous codon usage and protein secondary structure. *J. Mol. Biol.* **281**:31–48.
 56. Palmenberg, A. C., and J.-Y. Sgro. 1997. Topological organization of picornaviral genomes: statistical prediction of RNA structural signals. *Semin. Virol.* **8**:231–241.
 57. Rzhetsky, A., and M. Nei. 1993. Theoretical foundation of the minimum evolution method of phylogenetic inference. *Mol. Biol. Evol.* **10**:1073–1095.
 58. Rzhetsky, A., and M. Nei. 1994. METREE: program package for inferring and testing minimum evolution trees. *Comput. Appl. Biosci.* **10**:409–412.
 59. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
 60. Sala, M., and S. Wain-Hobson. 1999. Drift and conservatism in RNA virus evolution: are they adapting or merely changing? p. 115–140. *In* E. Domingo, R. Webster, and J. J. Holland (ed.), *Origin and evolution of viruses*. Academic Press, San Diego, Calif.
 61. Sharp, P. M., and G. Matassi. 1994. Codon usage and genome evolution. *Curr. Opin. Genet. Dev.* **4**:851–860.
 62. Simmonds, P., and D. B. Smith. 1999. Structural constraints on RNA virus evolution. *J. Virol.* **73**:5787–5794.
 63. Steinhauer, D. A., and J. J. Holland. 1987. Rapid evolution of RNA viruses. *Annu. Rev. Microbiol.* **41**:409–433.
 64. Strimmer, K., and A. von Haeseler. 1996. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**:964–969.
 65. Takeda, N., M. Tanimura, and K. Miyamura. 1994. Molecular evolution of the major capsid protein VP1 of enterovirus 70. *J. Virol.* **68**:854–862.
 66. Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**:512–526.
 67. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
 68. Toyoda, H., M. Kohara, Y. Kataoka, T. Suganuma, T. Omata, N. Imura, and A. Nomoto. 1984. Complete nucleotide sequences of all three poliovirus serotype genomes. Implication for genetic relationship, gene function and antigenic determinants. *J. Mol. Biol.* **174**:561–585.
 69. Villaverde, A., M. A. Martinez, F. Sobrino, J. Dopazo, A. Moya, and E. Domingo. 1991. Fixation of mutations at the VP1 gene of foot-and-mouth disease virus. Can quasispecies define a transient molecular clock? *Gene* **103**:147–153.
 70. Ward, C. D., M. A. Stokes, and J. B. Flanagan. 1988. Direct measurement of the poliovirus RNA polymerase error frequency in vitro. *J. Virol.* **62**:558–562.
 71. Ward, C. D., and J. B. Flanagan. 1992. Determination of the poliovirus RNA polymerase error frequency at eight sites in the viral genome. *J. Virol.* **66**:3784–3793.
 72. World Health Organization. 1997. Manual for the virologic investigation of poliomyelitis. WHO/EPI/GEN/97.1. World Health Organization, Geneva, Switzerland.
 73. Zhang, G., D. T. Haydon, N. J. Knowles, and J. W. McCauley. 1999. Molecular evolution of swine vesicular disease virus. *J. Gen. Virol.* **80**:639–651.
 74. Zhou, J., W. J. Liu, S. W. Peng, X. Y. Sun, and I. Frazer. 1999. Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. *J. Virol.* **73**:4972–4982.