JOURNAL OF VIROLOGY, June 1992, p. 3557–3565
0022-538X/92/063557-09$02.00/0

Vol. 66, No. 6

Evolution of the Capsid Protein Genes of Foot-and-Mouth Disease Virus: Antigenic Variation without Accumulation of Amino Acid Substitutions over Six Decades

M. A. MARTÍNEZ,[1,2] J. DOPAZO,[3] J. HERNÁNDEZ,[1] M. G. MATEU,[1] F. SOBRINO,[3] E. DOMINGO,[1]* AND N. J. KNOWLES[2]

Centro de Biología Molecular, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid,[1] and Instituto Nacional de Investigaciones Agrarias-Sanidad Animal, 28012 Madrid,[3] Spain, and Agriculture and Food Research Council Institute for Animal Health, Pirbright Laboratory, Woking, Surrey GU24 0NF, United Kingdom[2]

Received 23 December 1991/Accepted 19 March 1992

The genetic diversification of foot-and-mouth disease virus (FMDV) of serotype C over a 6-decade period was studied by comparing nucleotide sequences of the capsid protein-coding regions of viruses isolated in Europe, South America, and The Philippines. Phylogenetic trees were derived for VP1 and P1 (VP1, VP2, VP3, and VP4) RNAs by using the least-squares method. Confidence intervals of the derived phylogeny (significance levels of nodes and standard deviations of branch lengths) were placed by application of the bootstrap resampling method. These procedures defined six highly significant major evolutionary lineages and a complex network of sublines for the isolates from South America. In contrast, European isolates are considerably more homogeneous, probably because of the vaccine origin of several of them. The phylogenetic analysis suggests that FMDV CGC Ger/26 (one of the earliest FMDV isolates available) belonged to an evolutionary line which is now apparently extinct. Attempts to date the origin (ancestor) of the FMDVs analyzed met with considerable uncertainty, mainly owing to the stasis noted in European viruses. Remarkably, the evolution of the capsid genes of FMDV was essentially associated with linear accumulation of silent mutations but continuous accumulation of amino acid substitutions was not observed. Thus, the antigenic variation attained by FMDV type C over 6 decades was due to fluctuations among limited combinations of amino acid residues without net accumulation of amino acid replacements over time.

Foot-and-mouth disease virus (FMDV) is an aphthovirus of the family *Picornaviridae* which causes an economically important disease of cattle and other cloven-hoofed animals (reviews in references 2, 6, 20, and 57). FMDV populations are genetically and antigenically heterogeneous (9, 12–17, 27, 47, 51, 58, 60, 63), and variants showing different competitive abilities relative to their parental populations arise at high rates (28, 46). Thus, the virus conforms to a quasispecies structure, a feature shared with most other RNA viruses and retroviruses (7, 11, 18, 19, 31–36, 40, 43, 53, 59, 67, 68).

Long-term evolution of FMDV, which often involves relevant antigenic changes, must be the result of events that occur during at least three stages of the epizootiological cycle of the virus: (i) genetic variation within an infected host (9, 17, 27), (ii) genetic variation and random sampling events during a disease outbreak (17, 47, 52, 65, 70, 71), and (iii) survival of virus outside infected hosts and existence of transport routes that may lead to initiation of outbreaks at distant geographical locations (see reviews on the epizootiology of FMDV in references 2, 57, and 62). The available information does not permit an assessment of which of these three stages contributes most to FMDV diversification. During persistent infections of FMDV (27) or influenza virus type A (59) or in the course of acute poliomyelitis (41), the replicating viruses may evolve to a greater extent in a single individual than after multiple rounds of multiplication and transmission in an epidemic outbreak. It has been suggested that the extent of virus diversification in an infected individual at the time of transmission to a susceptible host, as well as the number of infectious particles involved in the transmission, may influence the rate of fixation of mutations in the field (41, 70).

During an extended FMD episode (Spain, 1970 to 1982), periods of rapid evolution alternated with others of relative genetic stability (stasis), for reasons that are not well understood (65, 70, 71). In the periods of rapid evolution, mutations accumulated with time in an approximately linear fashion. This analysis involved the computation of mutations between any two viral RNAs. Such a pairwise comparison necessitates that all viruses share a common recent ancestor to ensure computation of replacements among directly related genomes, ideally from a single lineage (70, 71). Obviously, this approach is not suitable for derivation of phylogenetic relationships among viruses isolated over several decades on different continents.

Phylogenetic methods have been applied to several viruses. Evolutionary relationships among different picornaviruses were first established by Palmenberg (56). For FMDV, analysis of the VP1 gene—which encodes important antigenic determinants of the virus (69)—of 15 isolates of serotype A, O, or C produced an extremely ramified tree that reflected the groupings of classical serology (21). However, the number of sequences available at that time did not allow a study of the genetic diversification of an FMDV serotype. The object of the present analysis was to define the main evolutionary lineages for FMDV of serotype C from 1926 (the year of isolation of the first FMDV available) until the present. We determined 12 new VP1 RNA sequences of viruses isolated in Europe, South America, and The Philippines. Together with previously reported sequences, they

* Corresponding author.

TABLE 1. FMDV isolates used in the phylogenetic analysis

FMDV designation[a]	Isolation date (mo/yr)	Place of isolation	Reference(s) for VP1, P1 RNA nucleotide sequences
CGC Ger/26[b]	Circa 1926	Germany	This report
C_2 997 UK/53	1953	United Kingdom	This report
C_1 Loupoigne Bel/53	5/1953	Loupoigne, Belgium	This report
C_1 Vosges Fr/60	1960	Vosges, France	This report
C_1 Oberbayern Ger/60	6/1960	Niederroth, Germany	3
C_1 Turup Den/61	1961	Turup, Denmark	This report
C_1 Noville Sw/65	1965	Noville, Switzerland	This report
C_1 Haute Loire Fr/69	1969	Haute Loire, France	This report
C_1 Sta Pau Sp/70C (C-S8c1)	2/1970	Santa Pau, Girona, Spain	64, 65
C_1 Serra de Daró Sp/81 (C-S15)	1/1981	Serra de Daró, Girona, Spain	64, 65
C_1 Barcelona Sp/82 (C-S30)	1/1982	Barcelona, Spain	47
C Monte Pulciano It/88	6/1988	Monte Pulciano, Tuscany, Italy	This report
C Brescia It/88	3/1989	Brescia, Lombardy, Italy	This report
C_2 Pando Ur/44	1944–1945	Pando, Uruguay	This report
C_3 Resende Br/55	5/1955	Resende, Rio de Janeiro, Brazil	58, this report
C_4 Tierra del Fuego Arg/66	12/1966	Rio Grande, Tierra del Fuego, Argentina	This report
C_5 Argentina/69	6/1969	Pehuajó, Buenos Aires, Argentina	This report
C_3 Indaial Br/71	10/1971	Indaial, Santa Catarina, Brazil	10
C_3 Indaial Br/71-78	10/1971	Indaial, Santa Catarina, Brazil	45, this report
C_3 Santa Fe Arg/75c	1/1975	San Carlos, Santa Fe, Argentina	This report
C_3 Argentina/83c	3/1983	Daireaux, Buenos Aires, Argentina	This report
C_3 Argentina/84	9/1984	General Roca, Córdoba, Argentina	58
C_3 Argentina/85	12/1984	Marcos Juárez, Córdoba, Argentina	58
C Philippines/3/87	10/1987	Santa Cruz, Lubao, Pampanga, The Philippines	This report
C Philippines/1/88	2/1988	Cotabato, Mindanao, The Philippines	This report

[a] The three groups given correspond to isolates of serotype C from Europe, South America, and The Philippines, in chronological order; c indicates that the virus was plaque purified.

[b] CGC Ger/26 was received in 1933 at the World Reference Laboratory (Pirbright, United Kingdom) from Insel Riems, Germany. Another preparation of the same virus was independently maintained at the Pan-American Foot-and-Mouth Disease Center, Rio de Janeiro, Brazil (termed CGC Ger/26c in Fig. 1 to 3). The two preparations were analyzed by nucleotide sequencing, and they differ in 24 of 2,193 P1 positions. FMDV CGC Ger/26 is the first serotype C isolate of FMDV (72).

were used to derive a phylogenetic tree. The reliability of node positions and branch lengths was assessed by the statistical bootstrap procedure (22, 23, 44), here applied for the first time to a highly variable RNA virus. New sequences for the entire capsid-coding (P1) region and the corresponding tree are also reported, and they confirm the relationships established with VP1 RNA. The results define six main evolutionary lineages for FMDV of serotype C and provide the first evidence of the probable extinction of an FMDV lineage in the field.

MATERIALS AND METHODS

Viruses. The FMDV isolates of serotype C analyzed are listed in Table 1. The procedures used for virus isolation and purification and extraction of genomic RNA have already been described (17, 47, 52, 63).

Nucleotide sequencing. FMDV genomic RNA was sequenced by primer extension and dideoxy-chain termination using reverse transcriptase (65, 75). VP1 RNAs from FMDV CGC Ger/26 and C_1 Haute Loire Fr/69 were also amplified by using the polymerase chain reaction, and the DNA products were sequenced as previously described (38). The oligodeoxynucleotide primers used for RNA or cDNA sequencing were complementary to positions 185 to 203, 326 to 345, 481 to 499, 712 to 729, 856 to 874, 943 to 961, 1184 to 1209, 1417 to 1436, 1677 to 1692, 1834 to 1850, and 2101 to 2118 of P1 RNA or to positions 34 to 54 of 2AB RNA. In all cases, the average sequence present in the populations of RNA or cDNA molecules was determined.

Phylogenetic trees. Nucleotide sequences of VP1 or P1 RNA were aligned as previously described (21). The genetic distance (d) between any two nucleotide sequences was calculated by using the formula $d = -0.75 \cdot \ln[1 - (4/3 \cdot p]$, where p is the proportion of different nucleotides between the two sequences (39). The phylogenetic tree was derived from the distance matrix by using the least-squares method (25) as implemented in the FITCH program from the PHYLIP package, version 3.3 (24). To estimate the accuracy of the tree topology and of the branch lengths, the bootstrap method was used (22, 23). First, 10^3 replicates of the original set of sequences were obtained and for each replica a phylogenetic tree was derived (24, 25). Then a consensus topology for the tree and confidence intervals for each branching point were obtained by using the CONSENSE program from the PHYLIP package (24). Once a consensus topology was established, 10^3 additional bootstrap replicates of the original set of sequences were obtained. Again, for each replicate a phylogenetic tree was derived but in this case its topology was forced to be the consensus topology by using the "user tree" option of the FITCH program. This gave 10^3 replicates for the branch lengths, and from them estimates of means, standard deviations, and confidence intervals for branch lengths were calculated (22).

The origin (ancestor) of the FMDVs analyzed was dated by following previously described procedures (29) using the least-squares fit relating genetic distances (from each isolate to the origin of the tree) to isolation dates. Errors and confidence intervals were calculated for the rate of fixation of mutations and the time of origin of the ancestor by taking

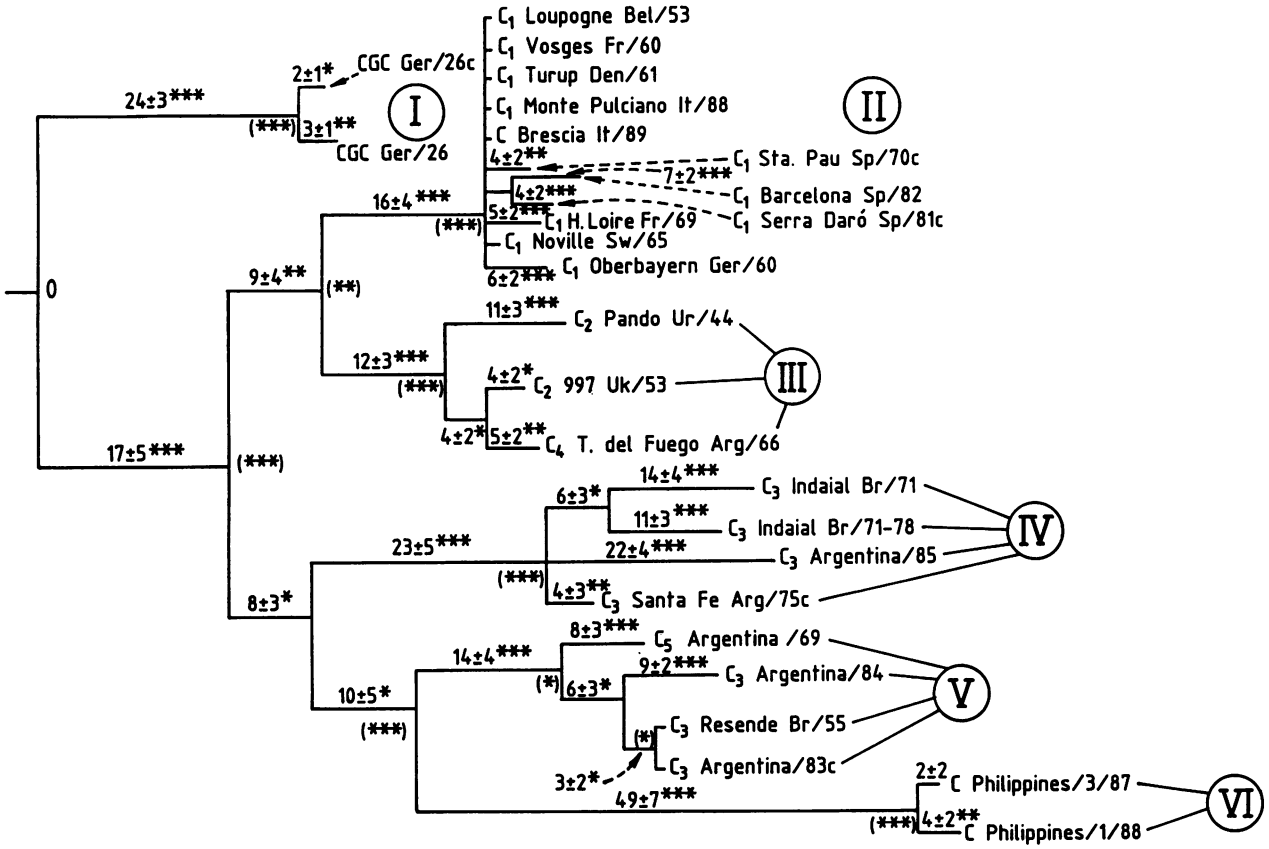

FIG. 1. Phylogenetic tree derived with VP1 RNA sequences of FMDV isolates of serotype C. The origins of the viruses are listed in Table 1. The procedures used for tree construction and determination of significance levels and standard deviations of branch lengths are described in Materials and Methods. Significance levels for nodes are given in parentheses; symbols: *, $0.05 > P > 0.01$; **, $0.01 > P > 0.001$; ***, $0.001 > P$; no symbols on a branch length means no significant difference from zero; no symbols on a node (except the origin, O) means that the branching order at this point is not defined. The main lineages with a high degree of significance are labeled I to VI; I includes only CGC Ger/26, and II includes a total of 11 European isolates. The same tree was obtained by using the neighbor-joining method (61, 66).

into account the individual error of each point estimated by the bootstrap procedure (22). Synonymous versus nonsynonymous mutations were computed by using the algorithm of Nei and Gojobori (55).

Nucleotide sequence accession numbers. The nucleotide sequences reported here have been deposited in the GenBank data base under accession no. M84360, M90055, and M90367 to M90380.

RESULTS

Multiple evolutionary lineages in FMDV of serotype C. On the basis of the genetic distance matrix obtained as described in Materials and Methods, a phylogenetic tree relating the VP1 genes of 26 FMDV type C isolates was derived by using the least-squares method (Fig. 1). Confidence levels of node positions, as well as the standard deviations of branch lengths, were determined by application of the bootstrap method (22). Six significantly distinct clusters of isolates, designated I to VI in Fig. 1, were distinguished. It is interesting that most European isolates cluster in group II whereas the South American isolates are widely distributed among groups III, IV, and V. The position of C_2 997 UK/53 among South American viruses in group III was unexpected, even though this group is the closest to European isolates. This observation agrees with the serologic and antigenic behavior of C_2 997 UK/53 in being closer to South American than to other European isolates (48, 49). Genetic distances between any two European viruses of group II are, on average, sevenfold lower than among South American isolates, and this difference is maintained even if the comparison is restricted to isolates from the last 2 decades. Thus, there must be a fundamental difference in FMDV epidemiology between the two continents (see below and Discussion).

To determine whether the lineages defined by the VP1 gene reflected the evolution of the entire capsid, a tree was derived with the P1-coding regions of eight viruses belonging to subtype C_1, C_2, C_3, or C_4 (Fig. 2). Both trees define similar relationships among different C subtypes.

Evidence of extinction of an early lineage. FMDV CGC Ger/26 was isolated during an epizootic outbreak in Germany in 1926. This virus differs considerably from any other type C FMDV examined, and it defines a separate lineage (I in Fig. 1). Preliminary partial sequencing of VP1 genes of type C viruses from Angola, Bhutan, Ethiopia, India, Kuwait, and Sri Lanka show that these viruses are different from other type C viruses, including CGC Ger/26 (42a). Even though CGC Ger/26 is the isolate closest to the tree origin, the results render very unlikely the possibility that it was the direct ancestor of subsequent type C viruses. The isolation date of CGC Ger/26 is not within the confidence interval

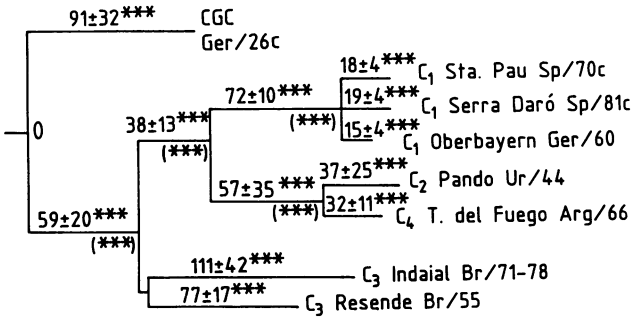

FIG. 2. Phylogenetic tree derived with P1 (the four capsid protein genes) RNA sequences of isolates of FMDV subtypes C_1, C_2, C_3, C_4, and CGC Ger/26c. Symbols are as in Fig. 1. P1 RNA sequences for C_1 viruses were reported previously (3, 64) (see also Table 1); other sequences are from this report.

calculated for the branching point that leads to all other viruses (O in Fig. 1 and 2; see also below). Rather, the analysis suggests that an ancestor for which no representative isolate is available gave rise to at least two sublines—one including FMDV CGC Ger/26 and the other constituting the origin of all of the other type C isolates examined (Fig. 1 and 2). Since no virus similar to CGC Ger/26 has been found among a substantial number of isolates from Europe and South America, CGC Ger/26 appears to be either extinct or circulating to a very limited extent, at least on these two continents.

Amino acid sequences of capsid proteins of representative FMDV C subtypes. The alignment of amino acid sequences of the capsid proteins of eight FMDV isolates of serotype C (Fig. 3) shows considerable diversity, mainly (but not exclusively) at loops exposed on the virion surface. A total of 52 amino acids are present only once in a given position (Fig. 3). FMDV CGC Ger/26 includes 19 such unique residues, which also occur in the other viruses, albeit to a lesser extent—ranging from 12 in C_3 Indaial Br/71-78 to only 2 in FMDV C_1 Sta Pau Sp/70c. The capsid of FMDV CGC Ger/26 shows somewhat closer amino acid identity to South American than to European isolates. However, this trend is lost at main antigenic site A of VP1 (loop GH in Fig. 3), in which CGC Ger/26 shares a similar number of residues with the two groups of viruses. This is reflected in a spectrum of reactivities of CGC Ger/26 with monoclonal antibodies directed to site A which is intermediate between those shown by European and South American viruses (48).

Rates of long-term evolution. Attempts to date the origin of FMDV type C. A plot of genetic distances versus isolation dates for all of the viruses under analysis showed a poor correlation coefficient (Fig. 4A). As a consequence, it was not possible either to derive reliable values for the rates of fixation of mutations or to date the tree origin (O in Fig. 1). In particular, the European viruses isolated after 1964 deviated most from a line defined by all of the other isolates (Fig. 4B). Thus, either there was no accumulation of mutations in VP1 RNA with time or the recent European isolates do not reflect the natural evolution of FMDV. We favor the second possibility, since there is considerable evidence that after implementation of compulsory vaccination in Europe (around 1965) the number of FMD outbreaks was greatly reduced (20), but several of them were probably caused by reintroductions of vaccine strains in the field (4, 8). When European viruses isolated after 1964 were omitted from the plot, a reasonable correlation was established between the number of mutations and isolation dates, and the extrapolated date for the origin of the FMDVs analyzed was 1897 (with a confidence interval of 1876 to 1912 at 95%) (Fig. 4B). From this plot, the average rate of evolution was $(1.43 \pm 0.16) \times 10^{-3}$ substitution per nucleotide per year. It could be argued that these values could be affected by a possible vaccine origin of the viruses of group V, which are related to C_3 Resende Br/55, a vaccine strain used in South America. Among other possibilities, Piccone et al. (58) considered that the similarity between C_3 Resende Br/55 and C_3 Argentina/84 could be explained by a vaccine origin of the latter, although the evidence is not conclusive. However, exclusion of C_3 Argentina/83c or C_3 Argentina/84 (or both) from our calculation did not significantly affect the extrapolated date of origin or the rate of accumulation of mutations (data not shown).

Antigenic variation without net accumulation of amino acid substitutions over time. Interestingly, this evolution of FMDV involves essentially synonymous (silent) replacements (Fig. 4C and Table 2), which define a rate of accumulation of mutations [$(2.1 \pm 0.7) \times 10^{-3}$ substitution per nucleotide per year, calculated relative to the maximum possible number of silent replacements according to reference 55] and an extrapolated date of origin (year, 1859; confidence interval of 1834 to 1877 at 95%) within the range of values calculated from the total number of mutations. The slope of the line defined by mutations leading to amino acid substitutions (Fig. 4C) is not significantly different from zero. Thus, the remarkable antigenic variation during 60 years of FMDV type C diversification (48, 49) was achieved without net accumulation of amino acid replacements in VP1.

We conclude that genetic and antigenic diversification of FMDV has been extensive in the course of 6 decades. Genetic diversification was mainly the result of accumulation of silent mutations. Antigenic diversification did not entail a net accumulation of amino acid substitutions. Six main evolutionary lineages of FMDV type C have been defined. The analysis suggests that a virus which circulated early in this century is now extinct.

DISCUSSION

Significance of phylogenetic trees for highly variable RNA viruses. Few studies on the establishment of phylogenetic relationships among virus isolates have included a statistical evaluation of the significance levels of node positions and branch lengths. In a recent derivation by cluster analysis of a tree relating the *env* genes of immunodeficiency viruses, it was not possible to specify the branching order of several human and simian viruses, owing to noise accumulation (42). A more common procedure to probe the reliability of a tree is provided by the bootstrap method, a statistical procedure based on reiteration of calculations using random samples of the original sets of sequences (22, 23, 44). This procedure has indicated as highly significant the six main FMDV lineages (I to VI) depicted in Fig. 1. However, sublines which seem obvious from an epidemiological standpoint (for example, the branchings leading to C_2 997 UK/53 and C_4 Tierra del Fuego Arg/66 or to C_3 Resende Br/55 and C_3 Argentina/83c; compare Fig. 1) are close to the limit of significance. Even though group I is represented only by the early CGC Ger/26 isolate, RNA preparations of this virus maintained in different laboratories for many years have yielded similar sequences (Tables 1 and 2 and Fig. 1). This rules out the possibility that its deviant nucleotide se-

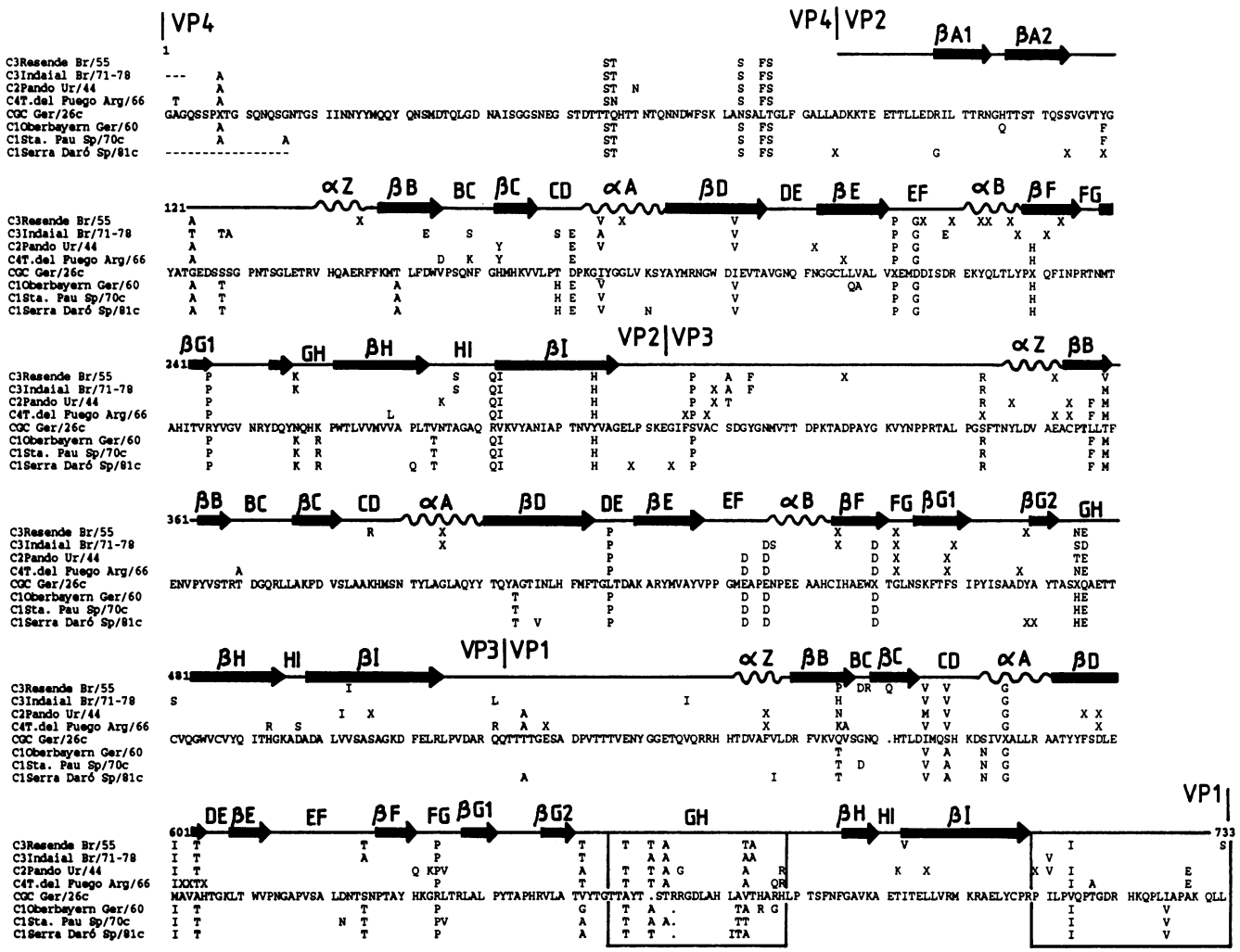

FIG. 3. Alignment of amino acid sequences of the four capsid proteins of FMDV subtypes C_1, C_2, C_3, C_4, and CGC Ger/26, as deduced from the corresponding nucleotide sequences. The amino acid sequence of capsid protein VP1 of these viruses has been previously reported (48). The single-letter amino acid code is used. Only amino acid changes relative to the FMDV CGC Ger/26 sequence are indicated. Symbols: dot, deletion; X, ambiguity in the sequencing gel; dash, sequence not determined. On top of the first sequence, secondary-structure motifs are indicated assuming that, upon alignment, they correspond to the equivalent residues of FMDV O_1 BFS 1860 (1). Symbols for secondary structure: thin line, N and C termini and loops; arrow, β-sheet; wavy line, α-helix. Boxed residues indicate the central region of the FMDV loop (amino acids 656 to 678), which includes antigenic site A, and the C terminus, which includes antigenic site C (amino acids 710 to 733) in VP1 of FMDV type C (48, 50).

quence—no longer represented among later isolates analyzed—was the result of an artifactual divergence in the laboratory.

The bootstrap method takes into account uncertainties derived from the limited number of replacements between sequences. Obviously, it does not consider an additional biological uncertainty derived from the quasispecies structure of RNA virus populations (17–21, 27, 28, 32, 36, 46, 47, 51, 59, 63, 65, 68). Any infected animal includes multiple variants (9, 17), and viruses may change upon adaptation to cell culture (16, 63), which is classically used for FMDV isolation. For example, C_3 Indaial Br/71 and C_3 Indaial Br/71-78 are from the same isolate but differ in passage history in the laboratory (quoted in reference 20), and yet the two viruses define two sublines (Fig. 1). This probably represents an extreme case of VP1 diversification in the laboratory; other adaptations to cell culture which have been analyzed have resulted in more modest genetic and antigenic changes of the average FMDV populations (48, 48a, 51, 63).

Rates of evolution of FMDV. The estimates of the rate of evolution of FMDV over a 6-decade period may be compared with the rates calculated for FMDV during persistent infections of cattle (27) or during a defined epizootic outbreak (70, 71) (Table 3). It is not unexpected that rates of long-term evolution might be substantially lower than those measured over shorter time periods in which the virus replicates actively and continuously. This is because in the latter situation, periods of stasis are minimal or nonexistent and selective pressures may be quite different. Stasis has been noted for some gene segments of influenza A virus, particularly in avian hosts (29, 30, 37) and for North American populations of eastern equine encephalomyelitis virus (73). To estimate a rate of evolution and to extrapolate a time of origin for FMDV type C, we had to exclude seven

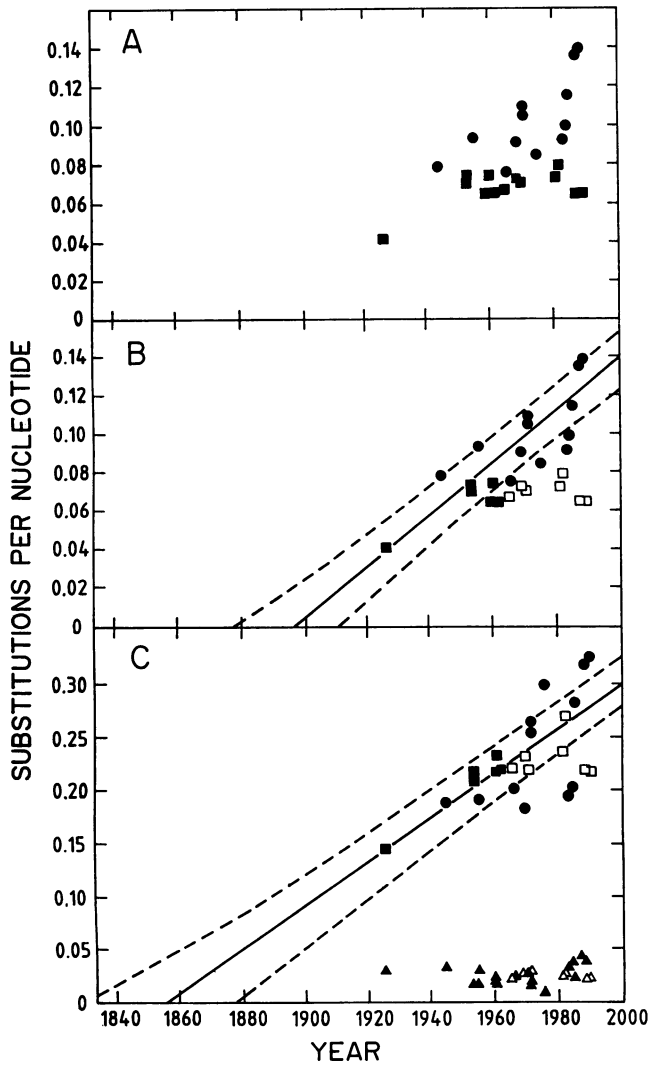

FIG. 4. Representation of genetic distances (substitutions per nucleotide in VP1 RNA of each FMDV relative to the tree origin) as a function of isolation dates. (A) All of the viruses included in Table 1 and Fig. 1. Viruses from South America and The Philippines are represented by circles, and European viruses are represented by squares. (B) Thick line: least-squares plot obtained with a subset of viruses (filled symbols) after exclusion of the European viruses isolated after 1964 (empty symbols) (see text); correlation coefficient, $r^2 = 0.79$. The broken lines delimit the 95% confidence interval derived from the standard deviation of each point. (C) Same as B but distinguishing synonymous mutations (same symbols as in B; correlation coefficient, $r^2 = 0.38$) from nonsynonymous mutations (triangles; empty symbols correspond to European viruses isolated after 1964). In this panel, substitutions per nucleotide refers to the number found relative to the maximum possible number of the corresponding type of replacement (silent or leading to an amino acid substitution), according to the algorithm of Nei and Gojobori (55). The statistical procedures used are described in Materials and Methods.

European viruses (compare Fig. 4A and B). On the basis of previous evidence, we have assumed that the excluded viruses were reintroductions of vaccine strains (4, 8) but that they do not reflect the natural evolution of FMDV. In support of our assumption that the evolution of FMDV is more accurately represented by the plot in Fig. 4B than that in Fig. 4A is the overall accumulation of mutations during an epizootic outbreak in Europe (Table 3) (52, 65, 70, 71). It is unlikely that in Europe (but not in South America) genetic variations fixed in successive outbreaks would be primarily reversions of previous mutations, even though this possibility cannot be completely excluded. When reservoirs in which the virus does not replicate are more significant (in terms of virus particles which they reintroduce in the field) than the viruses actively replicating in infected hosts, distortions of the type observed for FMDV in Europe are expected. Reintroductions (from vaccine pools or any other reservoir) will be epidemiologically more significant the lower the disease incidence. This reinforces the need to develop a synthetic vaccine not involving live virus for FMD control (5).

TABLE 2. Synonymous versus nonsynonymous mutations in the capsid protein genes of FMDV

Genomic segment	No. of mutations/ nucleotide site[a]		Synonymous/ nonsynonymous ratio
	Synonymous	Nonsynonymous	
Entire VP1 RNA	0.29	0.03	9.7
RNA encoding the FMDV loop and the C terminus of VP1[b]	0.40	0.07	5.7
Entire capsid protein-coding region (P1)	0.28	0.02	14.0

[a] Calculated as described by Nei and Gojobori (55). The pairwise comparison involved all of the viruses listed in Table 1 and was based upon alignment of nucleotide sequences (21) corresponding to the amino acid alignment of Fig. 3.
[b] The residues included in this calculation are boxed in Fig. 3.

Antigenic variation of FMDV does not require linear accumulation of amino acid substitutions over time. Both for eukaryotic cellular genes and for viral genes, the rate of fixation of synonymous nucleotide mutations is generally three- to sixfold higher than the rate for nonsynonymous replacements (review in reference 54). This bias is most dramatic in the case of FMDV type C, since no significant trend toward accumulation of amino acid substitutions was seen, even after 60 years of evolution in the field (Fig. 4C). We have recently suggested that FMDV has exploited a minimum of its potential for antigenic variation, since the variability at each amino acid residue within antigenic sites A and C was at least 40-fold lower than that possible on the basis of chemical considerations (48). Figure 5 illustrates several positions of the FMDV loop and of the carboxy terminus of VP1, in which there is a tendency to maintain alternacy between a very limited number of amino acids. The same observation applies to European viruses and to other VP1 regions, as well as to other capsid proteins (Fig. 3) (48). For the latter, however, the number of sequences

TABLE 3. Rate of evolution of VP1 RNA

Type of infection	Rate (substitutions/ nucleotide/yr)	Reference
Persistent infections of cattle[a]	0.9×10^{-2}–7.4×10^{-2}	27
Epizootic outbreak (Spain, 1979–1982)	6.5×10^{-3}	70
Long-term evolution (isolates collected over 6 decades)	1.4×10^{-3}	This report

[a] Persistent infections were established with plaque-purified FMDV.

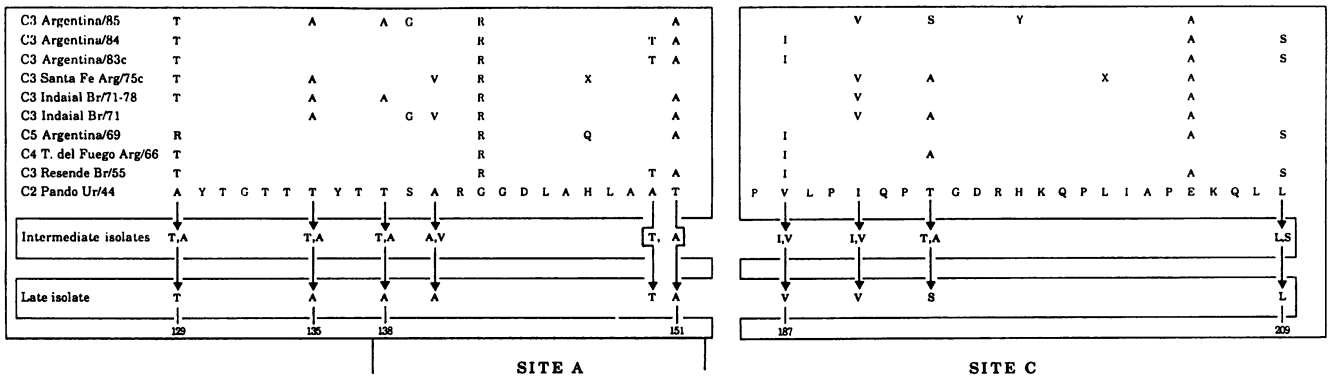

FIG. 5. Absence of accumulation of amino acid substitutions at variable sites within the FMDV loop (loop GH, which includes antigenic site A, residues 138 to 151) and at the carboxy-terminal segment (residues 187 to 209) of VP1, which includes antigenic site C. The amino acid sequences of South American isolates covering 5 decades are at the top; only amino acid substitutions relative to the early C_2 Pando Ur/44 sequence are indicated. Symbols are as in Fig. 3. We refer to a late isolate as the last one (C_3 Argentina/85) in the series compared; intermediate isolates are those isolated from 1955 until 1984. In boldface letters are those residues that show a tendency to maintain the same amino acid in intermediate and/or late isolates, as summarized at the bottom. Note the continuous alternations between Ala and Thr residues at positions 129, 135, 138, 150, and 151. The alignment of amino acid sequences is from reference 48.

available is not sufficient to allow a definitive conclusion. The lack of continuing accumulation of amino acid substitutions with time may be the result of limitations to the number and type of replacements tolerated by the FMDV capsid proteins. Such limitations may apply to antigenic sites, even though they appear to be structurally less constrained than internal capsid residues (1). Thus, considerable antigenic variation, possibly sufficient to render anti-FMD vaccines less effective (review in reference 20), is the result of fluctuations among limited combinations of amino acid residues but not of linear accumulation of amino acid changes over time.

Synonymous versus nonsynonymous mutations. The abundance of synonymous mutations during RNA virus evolution has been taken as evidence for the neutral theory of molecular evolution. Conversely, the dominance of nonsynonymous replacements at antigenic sites has been attributed to positive selection driven by the host immune response (for example, see reference 31 and references therein). However, serial passage of FMDV in the absence of anti-FMDV antibodies has repeatedly resulted in the dominance of antigenic variants in which most mutations had led to an amino acid substitution (15, 16, 46, 48a). This suggests that in addition to immune selection, other mechanisms may act to produce a bias in favor of or against nonsynonymous replacements during RNA genome evolution. As is the case for the viral proteins, the genomic RNA, owing to its need for functional foldings, must be subjected to selective forces and to limitations of the variations it may undergo (26, 74). A characteristic stem-loop structure in the RNA that encodes the carboxy-terminal half of VP1 is predicted in most of the variant type C FMDVs examined, despite considerable primary sequence variation (69a). Until more is known about the nature of selective constraints, a bias between synonymous and nonsynonymous substitutions (Table 2) will be difficult to interpret. Viral quasispecies evolve by competition among continuously arising variants subjected to continuous negative selection and also by episodes of positive selection acting on viruses in infected hosts and on free particles. It is not surprising that unpredictable genomic shifts and irregularities are observed during monitoring of long-term evolution in such systems.

ACKNOWLEDGMENTS

We thank M. Dávila for expert technical assistance and E. L. Palma and A. I. Donaldson for valuable suggestions.

Work at CBM and INIA was supported by a grant from the Comisión Interministerial de Ciencia y Tecnología (BIO88-0452-CO5) (Spain). Work at CBM was also supported by the European Community and the Fundación Ramón Areces. Work at WRL was supported by MAFF (United Kingdom). M.A.M. visited WRL supported by a fellowship from the European Community. J.H. is a predoctoral fellow from Comunidad Autónoma de Madrid.

REFERENCES

1. **Acharya, R., E. Fry, D. Stuart, G. Fox, D. Rowlands, and F. Brown.** 1989. The three-dimensional structure of foot-and-mouth disease virus at 2.9 A resolution. Nature (London) **337**:709–715.
2. **Bachrach, H. L.** 1968. Foot-and-mouth disease virus. Annu. Rev. Microbiol. **22**:201–244.
3. **Beck, E., G. Feil, and K. Strohmaier.** 1983. The molecular basis of the antigenic variation of foot-and-mouth disease virus. EMBO J. **2**:555–559.
4. **Beck, E., and K. Strohmaier.** 1987. Subtyping of European foot-and-mouth disease virus strains by nucleotide sequence determination. J. Virol. **61**:1621–1629.
5. **Brown, F.** 1989. The next generation of foot-and-mouth disease vaccines, p. 127–142. *In* J. P. Tam and E. T. Kaiser (ed.), Synthetic peptides: approaches to biological problems. Alan R. Liss, Inc., New York.
6. **Brown, F.** 1990. Picornaviruses, p. 153–169. *In* M. H. V. van Regenmortel and A. R. Neurath (ed.), Immunochemistry of viruses. II. The basis for serodiagnosis and vaccines. Elsevier, Amsterdam.
7. **Carpenter, S., L. H. Evans, M. Sevoian, and B. Chesebro.** 1990. *In vivo* and *in vitro* selection of equine infectious anemia virus variants. Appl. Virol. Res. **2**:99–115.
8. **Carrillo, C., J. Dopazo, A. Moya, M. González, M. A. Martínez, J. C. Saíz, and F. Sobrino.** 1990. Comparison of vaccine strains and the virus causing the 1986 foot-and-mouth disease outbreak in Spain: an epizootiological analysis. Virus Res. **15**:45–56.
9. **Carrillo, C., J. Plana, R. Mascarella, J. Bergada, and F. Sobrino.** 1990. Genetic and phenotypic variability during replication of foot-and-mouth disease virus in swine. Virology **179**:890–892.
10. **Cheung, A., J. DelaMarter, S. Weiss, and H. Küpper.** 1983. Comparison of the major antigenic determinants of different serotypes of foot-and-mouth disease virus. J. Virol. **48**:451–459.
11. **Coffin, J. M.** 1990. Genetic variation in retroviruses. Appl.

Virol. Res. **2**:11–33.
12. **Costa Giomi, M. P., I. E. Bergmann, E. A. Scodeller, P. Augé de Mello, I. Gomes, and J. L. La Torre.** 1984. Heterogeneity of the polyribocytidylic acid tract in aphthovirus: biochemical and biological studies of viruses carrying polyribocytidylic acid tracts of different lengths. J. Virol. **51**:799–805.
13. **Costa Giomi, M. P., I. Gomes, B. Tiraboschi, P. Augé de Mello, I. E. Bergmann, E. A. Scodeller, and J. L. La Torre.** 1988. Heterogeneity of the polyribocytidylic acid tract in aphthovirus: changes in the size of the poly(C) of viruses recovered from persistently infected cattle. Virology **162**:58–64.
14. **de la Torre, J. C., E. Martínez-Salas, J. Díez, A. Villaverde, F. Gebauer, E. Rocha, M. Dávila, and E. Domingo.** 1988. Coevolution of cells and viruses in a persistent infection of foot-and-mouth disease virus in cell culture. J. Virol. **62**:2050–2058.
15. **Díez, J., M. Dávila, C. Escarmís, M. G. Mateu, J. Dominguez, J. J. Pérez, E. Giralt, J. A. Melero, and E. Domingo.** 1990. Unique amino acid substitutions in the capsid proteins of foot-and-mouth disease virus from a persistent infection in cell culture. J. Virol. **64**:5519–5528.
16. **Díez, J., M. G. Mateu, and E. Domingo.** 1989. Selection of antigenic variants of foot-and-mouth disease virus in the absence of antibodies, as revealed by an in situ assay. J. Gen. Virol. **70**:3281–3289.
17. **Domingo, E., M. Dávila, and J. Ortín.** 1980. Nucleotide sequence heterogeneity of the RNA from a natural population of foot-and-mouth disease virus. Gene **11**:333–346.
18. **Domingo, E., and J. J. Holland.** 1988. High error rates, population equilibrium and evolution of RNA replication systems, p. 3–36. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), RNA genetics, vol. 3. CRC Press Inc., Boca Raton, Fla.
19. **Domingo, E., E. Martínez-Salas, F. Sobrino, J. C. de la Torre, A. Portela, J. Ortín, C. López-Galindez, P. Pérez-Breña, N. Villanueva, R. Nájera, S. VandePol, D. Steinhauer, N. de Polo, and J. J. Holland.** 1985. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance—a review. Gene **40**:1–8.
20. **Domingo, E., M. G. Mateu, M. A. Martínez, J. Dopazo, A. Moya, and F. Sobrino.** 1990. Genetic variability and antigenic diversity of foot-and-mouth disease virus. Appl. Virol. Res. **2**:233–266.
21. **Dopazo, J., F. Sobrino, E. L. Palma, E. Domingo, and A. Moya.** 1988. VP1 protein gene of foot-and-mouth disease virus: a quasispecies model of molecular evolution. Proc. Natl. Acad. Sci. USA **85**:6811–6815.
22. **Efron, B.** 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia.
23. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.
24. **Felsenstein, J.** 1990. PHYLIP manual version 3.3. University Herbarium, University of California, Berkeley.
25. **Fitch, W. M., and E. Margoliash.** 1967. Construction of phylogenetic trees. Science **155**:279–284.
26. **Fraile, A., and F. García-Arenal.** 1991. Secondary structure as a constraint on the evolution of plant viral satellite RNA. J. Mol. Biol. **221**:1065–1069.
27. **Gebauer, F., J. C. de la Torre, I. Gomes, M. G. Mateu, H. Barahona, B. Tiraboschi, I. Bergmann, P. Augé de Mello, and E. Domingo.** 1988. Rapid selection of genetic and antigenic variants of foot-and-mouth disease virus during persistence in cattle. J. Virol. **62**:2041–2049.
28. **González, M. J., J. C. Saiz, O. Laor, and D. M. Moore.** 1991. Antigenic stability of foot-and-mouth disease virus variants on serial passage in cell culture. J. Virol. **65**:3949–3953.
29. **Gorman, O. T., W. J. Bean, Y. Kawaoka, and R. G. Webster.** 1990. Evolution of the nucleoprotein gene of influenza A virus. J. Virol. **64**:1487–1497.
30. **Gorman, O. T., R. O. Donis, Y. Kawaoka, and R. G. Webster.** 1990b. Evolution of influenza A virus PB2 genes: implications for the evolution of the ribonucleoprotein complex and origin of human influenza A virus. J. Virol. **64**:4893–4902.
31. **Goudsmit, J., N. K. T. Back, and P. L. Nara.** 1991. Genomic diversity and antigenic variation of HIV-1: links between pathogenesis, epidemiology and vaccine development. FASEB J. **5**:2427–2436.
32. **Hernández, J., M. A. Martínez, E. Rocha, E. Domingo, and M. G. Mateu.** 1992. Generation of a subtype-specific neutralization epitope in foot-and-mouth disease virus of a different subtype. J. Gen. Virol. **73**:213–216.
33. **Holland, J. J., J. C. de la Torre, and D. A. Steinhauer.** 1992. RNA virus populations as quasispecies. Curr. Top. Microbiol. Immunol. **176**:1–20.
34. **Holland, J. J., J. C. de la Torre, D. K. Clarke, and E. Duarte.** 1991. Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. J. Virol. **65**:2960–2967.
35. **Holland, J. J., E. Domingo, J. C. de la Torre, and D. Steinhauer.** 1990. Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. J. Virol. **64**:3960–3962.
36. **Holland, J. J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol, and S. VandePol.** 1982. Rapid evolution of RNA genomes. Science **215**:1577–1585.
37. **Ito, T., O. T. Gorman, Y. Kawaoka, W. J. 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.
38. **Jansen, R. W., G. Siegl, and S. Lemon.** 1990. Molecular epidemiology of human hepatitis A virus defined by an antigen-capture polymerase chain reaction method. Proc. Natl. Acad. Sci. USA **87**:2867–2871.
39. **Jukes, T. H., and C. R. Cantor.** 1969. Evolution of protein molecules, p. 71–123. *In* H. N. Munro (ed.), Mammalian protein metabolism, vol. III. Academic Press, Inc., New York.
40. **Kinnunen, L., A. Huovilainen, T. Pöyry, and T. Hovi.** 1990. Rapid molecular evolution of wild type 3 poliovirus during infection in individual hosts. J. Gen. Virol. **71**:317–324.
41. **Kinnunen, L., T. Pöyry, and T. Hovi.** 1991. Generation of virus genetic lineages during an outbreak of poliomyelitis. J. Gen. Virol. **72**:2483–2489.
42. **Kirchhoff, F., K. D. Jentsch, B. Bachmann, A. Stuke, C. Laloux, W. Lüke, C. Stahl-Hennig, J. Schneider, K. Nieselt, M. Eigen, and G. Hunsmann.** 1990. A novel proviral clone of HIV-2: biological and phylogenetic relationship to other primate immunodeficiency viruses. Virology **117**:305–311.
42a. **Knowles, N. J.** Unpublished data.
43. **Kurstak, E., and A. Hossain.** 1990. Virus variability and impact on epidemiology and control of disease. Appl. Virol. Res. **2**:1–7.
44. **Li, W.-H., and M. Gouy.** 1990. Statistical tests of molecular phylogenies. Methods Enzymol. **183**:645–659.
45. **Makoff, A. J., C. A. Paynter, D. J. Rowlands, and J. C. Boothroyd.** 1982. Comparison of the amino acid sequence of the major immunogen from three serotypes of foot-and-mouth disease virus. Nucleic Acids Res. **10**:8285–8295.
46. **Martínez, M. A., C. Carrillo, F. González-Candelas, A. Moya, E. Domingo, and F. Sobrino.** 1991. Fitness alteration of foot-and-mouth disease virus mutants: measurement of adaptability of viral quasispecies. J. Virol. **65**:3954–3957.
47. **Martínez, M. A., C. Carrillo, J. Plana, R. Mascarella, J. Bergadá, E. L. Palma, E. Domingo, and F. Sobrino.** 1988. Genetic and immunogenic variations among closely related isolates of foot-and-mouth disease virus. Gene **62**:75–84.
48. **Martínez, M. A., J. Hernández, M. E. Piccone, E. L. Palma, E. Domingo, N. Knowles, and M. G. Mateu.** 1991. Two mechanisms of antigenic diversification of foot-and-mouth disease virus. Virology **184**:695–706.
48a. **Martínez, M. A., et al.** Unpublished data.
49. **Mateu, M. G., J. L. Da Silva, E. Rocha, D. L. De Brum, A. Alonso, L. Enjuanes, E. Domingo, and H. Barahona.** 1988. Extensive antigenic heterogeneity of foot-and-mouth disease virus of serotype C. Virology **167**:113–124.
50. **Mateu, M. G., M. A. Martínez, L. Capucci, D. Andreu, E. Giralt, F. Sobrino, E. Brocchi, and E. Domingo.** 1990. A single amino acid substitution affects multiple overlapping epitopes in the major antigenic site of foot-and-mouth disease virus of serotype C. J. Gen. Virol. **71**:629–637.

51. **Mateu, M. G., M. A. Martínez, E. Rocha, D. Andreu, J. Parejo, E. Giralt, F. Sobrino, and E. Domingo.** 1989. Implications of a quasispecies genome structure: effect of frequent, naturally occurring amino acid substitutions on the antigenicity of foot-and-mouth disease virus. Proc. Natl. Acad. Sci. USA **86:**5883–5887.
52. **Mateu, M. G., E. Rocha, O. Vicente, F. Vayreda, C. Navalpotro, D. Andreu, E. Pedroso, E. Giralt, L. Enjuanes, and E. Domingo.** 1987. Reactivity with monoclonal antibodies of viruses from an episode of foot-and-mouth disease. Virus Res. **8:**261–274.
53. **Meyerhans, A., R. Cheynier, J. Albert, M. Seth, S. Kwok, J. Sninsky, L. Morfeldt-Manson, B. Asjo, and S. Wain-Hobson.** 1989. Temporal fluctuations in HIV quasispecies "*in vivo*" are not reflected by sequential HIV isolations. Cell **58:**901–910.
54. **Nei, M.** 1987. Molecular evolutionary genetics. Columbia University Press, New York.
55. **Nei, M., and T. Gojobori.** 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3:**418–426.
56. **Palmenberg, A.** 1989. Sequence alignments of picornavirus capsid proteins, p. 211–241. *In* B. L. Semler and E. Ehrenfeld (ed.), Molecular aspects of picornavirus infection and detection. American Society for Microbiology, Washington, D.C.
57. **Pereira, H. G.** 1981. Foot-and-mouth disease, p. 333–363. *In* E. P. G. Gibbs (ed.), Virus diseases of food animals, vol. 2. Academic Press, Inc., New York.
58. **Piccone, M. E., G. Kaplan, L. Giavedoni, E. Domingo, and E. L. Palma.** 1988. VP1 of serotype C foot-and-mouth disease virus: long-term conservation of sequences. J. Virol. **62:**1469–1473.
59. **Rocha, E., N. J. Cox, R. A. Black, M. W. Harmon, C. J. Harrison, and A. P. Kendal.** 1991. Antigenic and genetic variation in influenza A (H1 N1) virus isolates recovered from a persistently infected immunodeficient child. J. Virol. **65:**2340–2350.
60. **Rowlands, D. J., B. E. Clarke, A. R. Carroll, F. Brown, B. H. Nicholson, J. L. Bittle, R. A. Houghten, and R. A. Lerner.** 1983. Chemical basis of antigenic variation in foot-and-mouth disease virus. Nature (London) **306:**694–697.
61. **Saitou, N., and M. Nei.** 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
62. **Sellers, R. F.** 1981. Factors affecting the geographical distribution and spread of virus diseases of food animals, p. 19–29. *In* E. P. G. Gibbs (ed.), Virus diseases of food animals, vol. 1. Academic Press, Inc., New York.
63. **Sobrino, F., M. Dávila, J. Ortín, and E. Domingo.** 1983. Multiple genetic variants arise in the course of replication of foot-and-mouth disease virus in cell cultures. Virology **128:**310–318.
64. **Sobrino, F., M. A. Martínez, C. Carrillo, and E. Beck.** 1989. Antigenic variation of foot-and-mouth disease virus of serotype C during propagation in the field is mainly restricted to only one structural protein (VP1). Virus Res. **14:**273–280.
65. **Sobrino, F., E. L. Palma, E. Beck, M. Dávila, J. C. de la Torre, P. Negro, N. Villanueva, J. Ortín, and E. Domingo.** 1986. Fixation of mutations in the viral genome during an outbreak of foot-and-mouth disease: heterogeneity and rate variations. Gene **50:**149–159.
66. **Sourdis, J., and M. Nei.** 1988. Relative efficiencies of the maximum parsimony and distance-matrix methods in obtaining the correct phylogenetic tree. Mol. Biol. Evol. **5:**298–311.
67. **Steinhauer, D. A., J. C. de la Torre, E. Meier, and J. J. Holland.** 1989. Extreme heterogeneity in populations of vesicular stomatitis virus. J. Virol. **63:**2072–2080.
68. **Steinhauer, D. A., and J. J. Holland.** 1987. Rapid evolution of RNA viruses. Annu. Rev. Microbiol. **41:**409–433.
69. **Strohmaier, K., R. Franze, and K. H. Adam.** 1982. Location and characterization of the antigenic portion of the FMDV immunizing protein. J. Gen. Virol. **59:**295–306.
69a. **Villaverde, A., and J. Dopazo.** Unpublished data.
70. **Villaverde, A., M. A. Martínez, F. Sobrino, J. Dopazo, A. Moya, and E. Domingo.** 1991. Fixation of mutation at the VP1 gene of foot-and-mouth disease virus. Can quasispecies define a transient molecular clock? Gene **103:**147–153.
71. **Villaverde, A., E. Martínez-Salas, and E. Domingo.** 1988. 3D gene of foot-and-mouth disease virus. Conservation by convergence of average sequences. J. Mol. Biol. **204:**771–776.
72. **Waldmann, O., and K. Trautwein.** 1926. Experimentelle Untersuchungen über die Pluralität der Maul-und-Klauenseuchevirus. Berlin. Tierärtzl. Wochenschr. **42:**569–570.
73. **Weaver, S. C., T. W. Scott, and R. Rico Hesse.** 1991. Molecular evolution of eastern equine encephalomyelitis virus in North America. Virology **182:**774–784.
74. **Yamamoto, K., and H. Yoshikura.** 1986. Relation between genomic and capsid structures in RNA viruses. Nucleic Acids Res. **14:**389–396.
75. **Zimmern, D., and P. Kaesberg.** 1978. 3′-terminal nucleotide sequence of encephalomyocarditis virus RNA determined by reverse transcriptase and chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **75:**4257–4261.