Elsevier has created a [Monkeypox Information Center](#) in response to the declared public health emergency of international concern, with free information in English on the monkeypox virus. The Monkeypox Information Center is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its monkeypox related research that is available on the Monkeypox Information Center - including this research content - immediately available in publicly funded repositories, with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the Monkeypox Information Center remains active.

Sequence and Evolutionary Relationships of African Swine Fever Virus Thymidine Kinase[1]

RAFAEL BLASCO,[2] CARLOS LÓPEZ-OTÍN,* MARIBEL MUÑÓZ,* ERNST-OTTO BOCKAMP,†
CARMEN SIMÓN-MATEO,† AND ELADIO VIÑUELA†,[3]

†Centro de Biología Molecular (CSIC-UAM), Facultad de Ciencias, Universidad Autónoma, Canto Blanco, 28049 Madrid, Spain; and
*Departamento de Biología Funcional, Facultad de Medicina, Universidad de Oviedo, 33006 Oviedo, Spain

The thymidine kinase gene of African swine fever virus was mapped in a 1.4-kb EcoRI–PstI fragment located in the left half of the EcoRI K fragment of African swine fever virus DNA by using degenerate oligonucleotide probes derived from regions of the thymidine kinase sequence conserved in several poxviruses, man, mouse, and chicken. The nucleotide sequence of this region revealed an open reading frame of 196 codons, whose translated amino acid sequence showed significant similarity to the thymidine kinases of vaccinia virus, variola virus, monkeypox virus, shope fibroma virus, fowlpox virus, capripox virus, man, mouse, and chicken. The similarity scores obtained after comparison of known thymidine kinase sequences indicated that the African swine fever virus thymidine kinase is more distantly related than the poxvirus thymidine kinases to their cellular homologs. The evolutionary implications of these findings are discussed. © 1990 Academic Press, Inc.

Animal viruses with large DNA genomes include families of icosahedral viruses (Herpesviruses, Iridoviruses) and brick-shaped viruses (Poxviruses) (1). African swine fever (ASF) virus does not fit well into any of these groups since, although its genomic structure and replication strategy are similar to those of poxviruses, its capsid symmetry is icosahedral (for reviews, see Refs. (2 and 3)).

ASF virus infects domestic and wild pigs, as well as a soft tick (Ornithodoros sp.) which acts as a vector for the virus. ASF virus can actively replicate in the tick population, in which it is transmitted both horizontally and vertically. The suggestion has been made that ASF virus was originally a virus of the ticks which eventually adapted to infect wild African pigs (4).

Various thymidine kinase (TK) sequences of different organisms are known, including man (5, 6), mouse (7), chicken (8), several poxviruses (9–14), and a number of herpesviruses (15–22). A comparison of viral and cellular TK sequences has shown the existence of two different groups of viral enzymes, one of which (Poxviruses) is closely related to cellular TKs, whereas the other (Herpesviruses) shows some resemblance to cellular thymidilate kinases (20, 23).

ASF virus induces a TK activity in infected cells (24). Given the unusual relationship of ASF virus with re-

spect to the well-established virus families, we undertook the mapping and sequencing of the TK gene of ASF virus in an attempt to define its evolutionary origins.

According to previous reports (12, 13), the amino acid sequence of the TKs of poxvirus and vertebrates share several conserved sequence motifs, of which the most conserved ones correspond to the sequences GPMFSGK and IDEGQFF. Two degenerate oligonucleotide probes, (GGNCCNATGTT(T/C)(T/A)(C/G)NGGNAA) and (AT(A/C/T)GA(T/C)GA(A/G)GGNCA(A/G)TT(T/C)TT), derived from these amino acid sequences, were synthesized and used to probe a set of ASF virus fragments covering the whole viral genome (25, 26). Both probes hybridized specifically with fragment SalI B (Fig. 1). Further mapping experiments showed that the hybridization site lies within a 1.4-kb EcoRI/PstI fragment located to the left of fragment EcoRI K. The DNA sequence around the hybridization site revealed the presence of an open reading frame (ORF), which could code for a 196 amino acid polypeptide (Fig. 2). Preliminary comparisons of this amino acid sequence with those of known TKs indicated that the ASF virus sequence showed significant similarity to both poxvirus and vertebrate TKs. However, no significant similarities were found with the sequences of herpesvirus TKs, other than the presence of the NTP-binding domain (23).

The putative ASF virus TK gene can code for a polypeptide with a molecular weight of about 22 kDa. This value is slightly lower than those predicted for the eukaryotic TKs (about 25 kDa) and slightly higher than those predicted for poxvirus TKs (about 20 kDa).
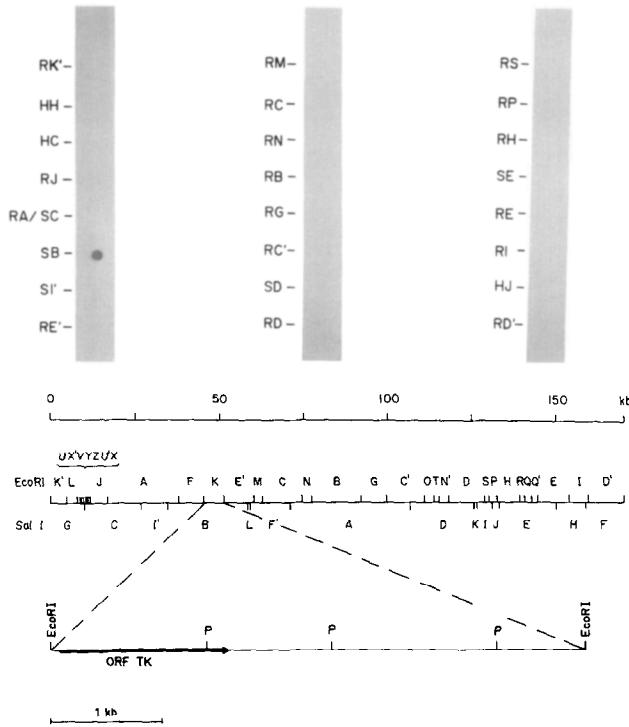
FIG. 1. Mapping of the ASF virus TK gene. The hybridization of the degenerate oligonucleotide probe to restriction fragments of ASF virus DNA is shown at the top. The nomenclature of ASF virus restriction fragments has been described (26). Restriction maps of the viral genome were taken from Almendral et al. (25). ORF TK is indicated with an arrow.

A progressive alignment method (27) was used to generate a multiple alignment of the ASF virus, poxvirus, and vertebrate TK sequences. The TKs of monkeypox virus and variola virus were not included in the alignment, since they show only a few amino acid replacements with respect to the vaccinia virus sequence (11). From the alignment of the different TK sequences (Fig. 3) it is noticeable that, despite the great divergence of the ASF virus TK, there are several regions with a high degree of similarity to the other known TK sequences. Conservation of amino acid sequences is particularly striking at the regions containing the nucleotide binding motif, located around positions 30 and 105 in the multiple alignment, suggesting that those regions are maintained by selective pressure. ASF virus TK shows enlarged N-terminal and C-terminal sequences when compared to the sequences of poxviruses. This fact suggests that the evolution of the terminal portions of the protein is the result of small deletions or additions rather than the result of the insertion of a poxvirus-like TK gene into another cellular gene, as has been previously hypothesized (13).

Pairwise comparison of the ASF virus TK sequence with those of poxviruses and vertebrates showed a de-

gree of similarity ranging from 25.1 to 31% (Table 1). In contrast, TKs from poxviruses were 51.1 to 67.6% identical to those of vertebrates. Therefore, the enzymes of vertebrates are more closely related to poxvirus TKs than to ASF virus TK.

An evolutionary analysis was carried out with the sequences aligned in Fig. 3, using both distance matrix and parsimony methods (28–32). Figure 4 shows the genealogic tree obtained using the distance matrix method. Optimal trees from protein parsimony methods showed minor differences with respect to this tree, mainly concerning the branching of fowlpox and Shope fibroma viruses within the poxvirus group. This, along with the fact that some branches leading to these poxviruses are very short, makes it possible that the tree does not reflect the exact phylogeny of the poxvirus sequences. However, the different methods gave consistent results regarding the branching of the groups of organisms (vertebrates, poxvirus, ASF virus) being considered.

The genealogic tree suggests the existence of a common ancestral TK sequence leading to all present-day poxvirus sequences. Similarly, all vertebrate sequences are clustered together, in a way consistent with the phylogeny of those organisms. The branching order in the tree supports the hypothesis that a poxvirus ancestor acquired the TK gene from its host. According to the tree, this would have happened before the divergence of birds and mammals.

From the unrooted tree in Fig. 4 two different rooted trees depending on the positioning of the ancestral sequence could be suggested. One of them implies either that the ancestral sequence was viral or that ASF virus acquired the TK gene long before poxviruses,

```
TTACCAGAATGGATCGCGTCTTACTAAAAGTGAATTGGAGAAAAAAATTAAACGGTCAAA    -1

        M  N  I  I  R  K  L  K  P  G  T  I  S  L  V  L  G  P  M  F    20
ATGAATATAATTAGGAAGCTTAAGCCTGGAACAATTAGCCTTGTGCTGGGACCCATGTTT    60

        A  G  K  T  T  F  L  I  H  C  I  Y  M  L  E  R  L  E  K  K    40
GCCGGCAAAACTACGTTTCTTATTCATTGCATTTACATGCTCGAACGTTTGGAAAAAAAA    120

        V  V  F  I  K  S  T  K  N  T  R  D  K  T  I  K  T  H  S  G    60
GTAGTCTTCATAAAATCTACCAAAAACACCCGAGACAAAACTATTAAAACACACTCCGGT    180

        I  Q  L  R  P  K  Q  C  K  I  I  E  S  T  Q  L  S  D  V  G    80
ATACAGCTACGACCCAAGCAATGTAAAATCATAGAAAGCACACAGTTATCTGACGTGGGA    240

        S  L  T  D  I  H  A  V  V  V  D  E  A  H  F  F  D  D  L  I    100
TCTCTCACCGATATCCATGCAGTTGTCGTAGATGAAGCGCATTTTTTTGACGATTTAATA    300

        K  C  R  T  W  A  E  E  E  K  I  I  I  L  A  G  L  N  A  S    120
AAATGCCGCACCTGGGCAGAGGAAGAAAAAATTATTATTCTTGCGGGACTCAATGCTTCC    360

        F  E  Q  K  M  F  P  P  I  V  R  I  F  P  Y  C  S  W  V  K    140
TTCGAGCAGAAAATGTTTCCGCCCATCGTTCGTATTTTTCCTTACTGCAGCTGGGTTAAG    420

        Y  I  G  R  T  C  M  K  C  N  Q  H  N  A  C  F  N  V  R  K    160
TATATTGGCCGCACCTGTATGAAATGTAACCAACATAATGCATGCTTTAATGTGCGTAAG    480

        N  A  D  K  T  L  I  L  A  G  G  S  E  L  Y  V  T  C  C  N    180
AACGCAGACAAGACGCTTATCCTTGCGGGAGGAAGTGAACTGTACGTAACATGTTGTAAC    540

        N  C  L  K  N  T  F  I  K  Q  L  Q  P  I  K  Y                 196
AACTGTCTAAAAAATACATTTATTAAGCAGTTGCAACCTATTAAATATTAAAAATCTTAT    600

ACAATAATGGATCATTATCTTAAAAAATTACAAGATGTCGAAAACCAATGAAAAAGAGTT    660
```

FIG. 2. Nucleotide sequence and predicted polypeptide sequence of the ASF virus TK gene.

```
                                                                    70
HUM    MSCINLPTVLPGSPSKTRGQIQVILGDMFSGKSTELMRRVRRFQIAQYKCLVIKYAKDTRY SSSF  C
MOU    MSYINLPTVLPSSPSKTRGQIQVILGFMISGKSTELMRRVRRFQIAQYKCLVIKYAKDTRY SNSF  S
CHI    MNCLTVPGVHPGSPGRPRKQIQVIFLFMISLKSTELMRRVRRFQLAQYRCLLVKYAKDTRYCTTGV  S
VV            MNGCHIQLIIGFMISLKSTELIRRVRRYQIAQYKCVTIKYSNDNRY GTGL  W
SFV           MYGCHIHLIIGFMIALKSTEIRLVRRYQIAKHKCLVVKYEKDIRY GNGV  C
CPOX          MDYGYIHLIIGFMISLKSTELIRIVKRYQIAQYKCCVVKYLKDIRY GNSV  Y
FPOX          MSSGSIHVITGFMISLKTSELVRRIKRFMLSNFKCIIIKHCGDNRYNEDDINKVY
ASFV   MNIIRKLKPGTISLVLGFMFAGKTTFLIHCIYMLERLEKKVVFIKSTKNTRDKTIKTHSGI

                                                                    140
HUM    THDRNTMEALPACLLRDVAQEAL   GVAVIGIDEGQFFPDIMEFCEAMANAGKTVIVAALDGTFQRKPFG
MOU    THDRNTMDALPACMLRDVTQELL   GVAVIGIDEGQFFPDIVDFCEMMANEGKTVIVAALDGTFQRKAFG
CHI    THDRNTMEARPACALGDVYQEAL   GSAVIGIDEGQFFPDIVEFCEKMANTGKTVIVMALDGTFQRKAFG
VV     THDKNNFEALEATKLCDVLESIT   DFSVIGIDEGQFFPDIVEFCERMANEGKIVIVAALDGTFQRKPFN
SFV    THDNMSITAVCTPSLDKIDSVAE   NAEVIGIDEGQFFPNIATFCERMANRGKVLIVAALDGTFQRKPFS
CPOX   THDNNHVSAISTTNNYDVVDKIM   NFDIIGIDEGQFFKDIVSFSENMANMGKIIIIAADSTFQRKEFN
FPOX   THDLLFMEATASSNLSVLVPTLINDGVQVIGIDEAQFFLDIVEFSESMANLGKTVIVAANGDFKRELFG
ASFV   QLRPKQCKIIESTQLSDVGSLT    DIHAVVVDEAHFFDDIIK CRTWAEEEKIIILAGLNASFEQKMFP

                                                                    210
HUM    AILNLVPLAESVVKLTAVCMECF REAAYTKRLGTEKEVEVIGGADKYHSVCRLCYFKKASGQPAGPDNK
MOU    SILNLVPLAESVVKLTAVCMECF REAAYTKRLGLEKEVEVIGGADKYHSVCRLCYFKKSSAQTAGSDNK
CHI    SILNLVPLAESVVKLNAVCMECY REASYTKRLGAEREVEVICGADKYHSVCRACYFQK RPQQLGSENK
VV     NILNLIPLSEMVVKLTAVCMKCF KEASFSKRLGEETEIEIIGGNDHYQSVCRKCYIDS
SFV    NISELIPLAENVTKLNAVCMYCY KNGSFSKRLGDKMEIEVIGGSDKYKSVCRKCYFF
CPOX   DILKLIPLSEKVIKLNAVCMECY KDAAFSKRITKEKEIELICGKEKYKSVCRKCYFLE
FPOX   NVYKLLSLAETVSSLTAICVKCY CDASFSKRVTENKEVMDIGGKDKYIAVCRKCFFSN
ASFV   PIVRIFPYCSWVKYIGRTCMKCNQHNACFNVRKNADKTLILAGGSELYVTCCNNC LKNTFIKQLQPIKY

                        241
HUM    ENCPVPGKPGEAVAARKLFAPQQILQCSPAN
MOU    NCLVLGQPGEALVVRKLFASQQVLQYNSAN
CHI    ENVPMGVKQLDMPASRKIFAS
```
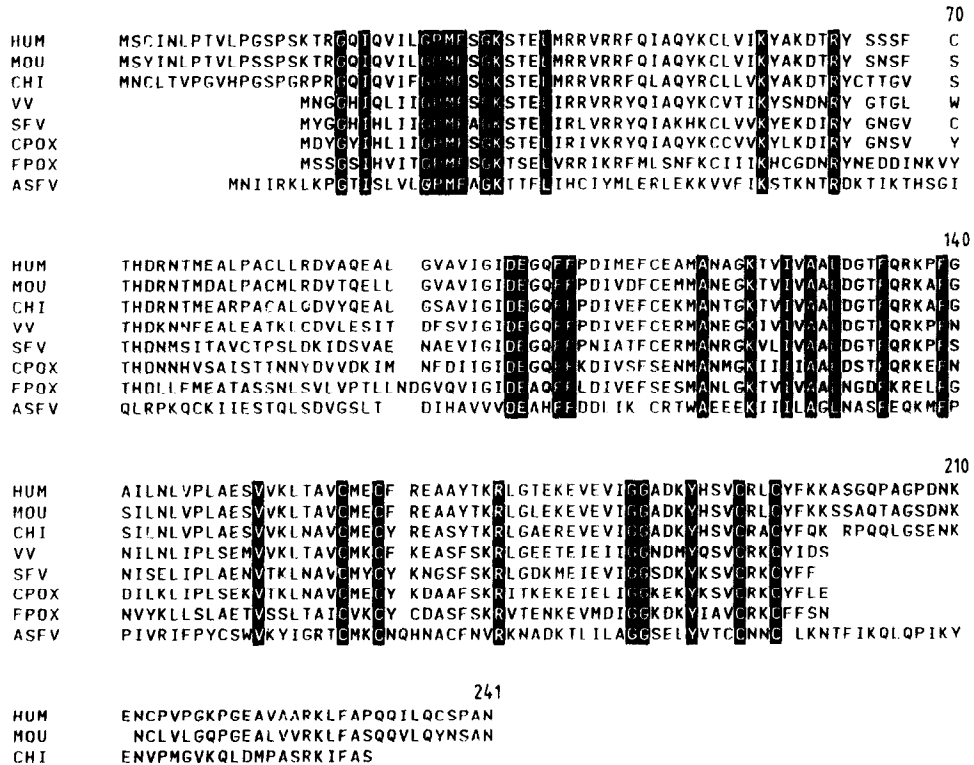
FIG. 3. Multiple alignment of the TK sequences from ASF virus, poxviruses, and vertebrates. Residues conserved in all the sequences in the alignment are shown in black boxes. Sequences are as follows: HUM, human cytoplasmic TK; MOU, mouse cytoplasmic TK; CHI, chicken cytoplasmic TK; VV, vaccinia virus; SFV, Shope fibroma virus; CPOX, capripox virus; FPOX, fowlpox virus.

therefore being two independent evolutive lineages, ASF viruses and poxviruses. An interesting possibility is that ASF virus had taken the TK gene from its invertebrate host, the soft tick, *Ornithodoros* sp. (*33, 34*). If this is the case, the divergence of the ASF virus TK would reflect, at least in part, the divergence of the invertebrate TK with respect to the vertebrate TKs.

The alternative rooted tree implies the existence of a common ancestor of the poxvirus and ASF virus sequences. Since ASF virus and the poxviruses have similar DNA structure and replication mechanism (*35–39*) it seems reasonable that they evolved from the same

ancestral virus. If this is the case, the greater divergence of the ASF virus TK should be the consequence of a greater rate of evolution of the ASF virus gene with respect to the poxvirus gene. The possibility of an evo-

TABLE 1

PERCENTAGE OF IDENTITY OF THE DIFFERENT TK SEQUENCES COMPARED

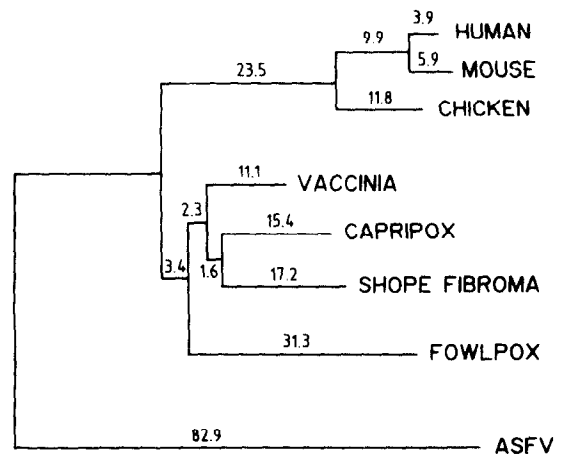|	Mouse	Chicken	VV	SFV	CPOX	FPOX	ASFV
Human	87.1	75.8	67.2	60.2	58.2	52.0	26.2
Mouse	—	74.8	66.7	59.1	59.3	52.0	26.2
Chicken		—	65.5	60.8	57.6	51.1	25.1
VV			—	65.1	64.4	50.3	31.0
SFV				—	64.6	46.3	27.3
CPOX					—	49.1	29.9
FPOX						—	25.8

FIG. 4. Genealogic tree relating different viral and cellular TK sequences. The tree was derived from a distance matrix obtained from the multiple alignment in Fig. 3. Numbers indicate the genetic distance for each branch. Horizontal lines are proportional to the genetic distance. Vertical lines are of arbitrary length, and are introduced to separate the different branches of the tree

lutive lineage common to poxviruses and ASF virus poses important questions about how those viruses evolved to have different capsid structures. On the other hand, a horizontal gene transfer between poxvirus and ASF virus ancestors cannot be ruled out until more related sequence data of the two viruses is available.

ACKNOWLEDGMENTS

REFERENCES

1. MATTHEWS, R. E. F., Intervirology 17, 1–199 (1982).
2. VIÑUELA, E., Curr. Top. Microbiol. Immunol. 116, 151–170 (1985).
3. VIÑUELA, E., In "African Swine Fever" (Y. Becker, Ed.)., pp. 31–49. Nijhoff, Boston, 1987.
4. PLOWRIGHT, W., In "Infectious Diseases of Wild Mammals" (J. W. Davis, L. H. Karstad, and D. O. Trainer, Eds.), 2nd ed., pp. 178–190. Iowa State Univ. Press, Ames, 1981.
5. BRADSHAW, H. D., and DEININGER, P. L., Mol. Cell. Biol. 4, 2316–2320 (1984).
6. FLEMINGTON, E., BRADSHAW, H. D., JR., TRAINA-DORGE, V., SLAGEL, V., and DEININGER, P. L., Gene 52, 267–277 (1987).
7. LIN, P. F., LIEBERMAN, H. B., YEH, D. B., XU, T., ZHAO, S. Y., and RUDDLE, F. H., Mol. Cell. Biol. 5, 3149–3156 (1985).
8. KWOH, T. J., and ENGLER, J. A., Nucleic Acids Res. 12, 3959–3971 (1984).
9. HURBY, D., MAKI, R. A., MILLER, D. B., and BALL, L. A., Proc. Natl. Acad. Sci. USA 80, 3411–3415 (1983).
10. WEIR, J. P., and MOSS, B., J. Virol. 46, 530–537 (1983).
11. ESPOSITO, J. J., and KNIGHT, J. C., Virology 135, 561–567 (1984).
12. UPTON, C., and MCFADDEN, G., J. Virol. 60, 920–927 (1986).
13. BOYLE, D. B., COUPAR, B. E. H., GIBBS, A. J., SEIGMAN, L. J., and BOTH, G. W., Virology 156, 355–365 (1987).
14. GERSHON, P. D., and BLACK, D. N., J. Gen. Virol. 70, 525–533 (1989).
15. MCKNIGHT, S. L., Nucleic Acids Res. 8, 5949–5964 (1980).
16. KIT, S., KIT, M., QAVI, H., TRKULA, D., and OTSUKA, H., Biochim. Biophys. Acta 741, 158–170 (1983).
17. SWAIN, M. A., and GALLOWAY, D. A., J. Virol. 46, 1045–1050, (1983).
18. OTSUKA, H., and KIT, S., Virology 135, 316–330 (1984).
19. HONESS, R. W., CRAXTON, M. A., WILLIAMS, L., and GOMPELS, U. A., J. Gen. Virol. 70, 3003–3013 (1989).
20. ROBERTSON, G. R., and WHALLEY, J. M., Nucleic Acids Res. 16, 11,303–11,317 (1988).
21. SHEPPARD, M., and MAY, J. T., J. Gen. Virol. 70, 3067–3071 (1989).
22. SCOTT, S. D., ROSS, N. L. J., and BINNS, M. M., J. Gen. Virol. 70, 3055–3065 (1989).
23. GORBALENYA, A. E., and KOONIN, E. V., Nucleic Acids Res. 17, 8413–8440 (1989).
24. POLATNICK, J., and HESS, W., Amer. J. Vet. Res. 31, 1609–1613 (1970).
25. ALMENDRAL, J. M., BLASCO, R., LEY, V., BELOSO, A., TALAVERA, A., and VIÑUELA, E., Virology 133, 258–270 (1984).
26. LEY, V., ALMENDRAL, J. M., CARBONERO, P., BELOSO, A., VIÑUELA, E., and TALAVERA, A., Virology 133, 249–257 (1984).
27. FENG, D. F., and DOOLITTLE, R. F., J. Mol. Evol. 25, 351–360 (1987).
28. FELSENSTEIN, J., In "Statistical Analysis of DNA Sequence Data" (B. S. Weir, Ed.), pp. 133–150. Dekker, New York, (1983).
29. FELSENSTEIN, J., Annu. Rev. Genet. 22, 521 (1988).
30. FITCH, W. M., and MARGOLIASH, E., Science 155, 279–284 (1967).
31. ECK, R. V., and DAYHOFF. M. O., National Biomedical Research Foundation, Silver Spring, Maryland, (1966).
32. FITCH, W. M., Syst. Zool. 20, 406–416 (1971).
33. SÁNCHEZ BOTIJA, C., Bull. Off. Int. Epizoot. 60, 895–899 (1963).
34. PLOWRIGHT, W., PARKER, J., and PIERCE, M. A., Nature (London) 221, 1071–1073 (1969).
35. ENJUANES, L., CARRASCOSA, A. L., and VIÑUELA, E., J. Gen. Virol. 32, 479–492 (1976).
36. ORTÍN, J., ENJUANES, L., and VIÑUELA, E., J. Virol. 31, 579–583 (1979).
37. SOGO, J. M., ALMENDRAL, J. M., TALAVERA, A., and VIÑUELA, E., Virology 133, 271–275 (1984).
38. GONZÁLEZ, A., TALAVERA, A., ALMENDRAL, J. M., and VIÑUELA, E., Nucleic Acid. Res. 14, 6835–6844 (1986).
39. BLASCO, R., AGÜERO, M., ALMENDRAL, J. M., and VIÑUELA, E., Virology 168, 330–338 (1989).