

Multigene Families in African Swine Fever Virus: Family 110

JOSÉ M. ALMENDRAL, FERNANDO ALMAZÁN, RAFAEL BLASCO,† AND ELADIO VIÑUELA*

Centro de Biología Molecular (Consejo Superior de Investigaciones Científicas-UAM), Facultad de Ciencias, Universidad Autónoma, Canto Blanco, 28049 Madrid, Spain

Received 15 September 1989/Accepted 2 January 1990

The genome of African swine fever virus was screened for the existence of repetitive sequences by hybridization between different cloned restriction fragments covering the viral DNA. Several sets of repeated sequences were detected in fragments located close to the DNA ends. One of these groups of repetitions involved fragments located at both ends of the genome. The remaining groups involved fragments that were located exclusively at the left end. The sequence of a 3.2-kilobase segment spanning from 7.5 to 11 kilobases from the left DNA end, which showed a complex pattern of cross-hybridizations, was determined. Two short and three long blocks of direct repeated sequences were found in this DNA region, which accounted for the hybridization results. The repeated sequences formed a family of five homologous genes with an average length of 116 codons (multigene family 110), one of which had a dimeric structure. Transcripts of the five members of the family were detected both in RNA synthesized *in vitro* by purified African swine fever virions and in RNA isolated at early times after infection. Comparison of the predicted protein sequences revealed a striking conservation of a cysteine-rich domain in the central part of the proteins. In addition, a highly hydrophobic NH₂-terminal sequence present in all the proteins suggests that these proteins are processed through the endoplasmic reticulum.

African swine fever (ASF) virus causes an important disease of domestic pigs and related species of the *Suidae* family (for reviews, see references 37 and 38). No neutralizing antibodies are produced during infection, although virus-specific antibodies can easily be detected in the sera of infected animals (17). The basic mechanisms underlying this immune response are not understood. Nevertheless, the lack of efficient neutralization seems to be due to characteristic features of ASF virus, since chronically infected or recovered animals can still respond efficiently to foot-and-mouth disease virus (7). This special nature of the interaction of ASF virus with its host has led us to carry out more-detailed studies on the genetic and biochemical properties of the virus. It is particularly important to characterize the distinctive features of the ASF virus genome, which could shed light on the special properties of the virus.

The ASF virus genome is a double-stranded DNA molecule of about 170 kilobases (kb) which contains hairpin loop structures linking both strands at the ends (12). Most of the viral DNA molecule has been cloned (14), and restriction maps for endonucleases *SalI*, *KpnI*, *SmaI*, and *PvuI* have been completed. *EcoRI* restriction fragments, except for several small fragments located between 5 and 10 kb from the left DNA end (1), were also mapped.

ASF virus induces the synthesis of about 100 proteins in infected cells (31). Genes transcribed late in infection map in the central part of the genome (29). Genetic variation among different virus isolates occurs mainly in two regions located close to the DNA ends, in which large deletions take place (3, 4).

Sogo et al. (34) described the presence of terminal and internal inverted repetitions at the ends of ASF virus DNA. Terminal inverted repetitions are composed of direct repeats

arranged in tandem (A. González, V. Calvo, and E. Viñuela, unpublished data).

As a further step to understanding the ASF virus genome organization, we have studied the distribution of repeated sequences along the viral DNA, since they are frequently associated with unstable elements in large DNA genomes and with mechanisms of immunological escape in microorganisms (for example, see the work of Murphy et al. [24]). We describe here the presence of internal repetitions in the ASF virus genome which are related to two multigene families. One of them (multigene family 110) showed a hydrophobic NH₂-terminal sequence and a conserved cysteine-rich domain.

MATERIALS AND METHODS

Cells and viruses. The BA71V strain of ASF virus, a cell-culture-adapted virus strain (9), was used to infect Vero cells (ATCC CCL 81).

Recombinant plasmids. Most of the recombinant plasmids harboring *EcoRI* restriction fragments of ASF virus DNA have been described elsewhere (14). Plasmids containing fragments HC/SG and SC/HC were selected from a library of cloned *SalI-HindIII* restriction fragments from the whole ASF virus genome by hybridization with *EcoRI* fragments L and J, respectively.

Dot blot hybridization. Individual restriction fragments were excised from the recombinant plasmids and purified from the vector by gel electrophoresis in low-melting-point agarose (13). Samples (20 ng) of each restriction fragment were treated and applied on nitrocellulose as described previously (15). Each restriction fragment was uniformly labeled with ³²P by nick translation (28) to a specific activity of 10⁸ cpm/μg and used as a probe. Hybridizations were done under either stringent (68°C) or relaxed (37°C and 20% formamide) conditions in 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 2× Denhardt solution (1× Denhardt solution is 0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone, and 0.02% Ficoll), and 50 μg of

* Corresponding author.

† Present address: Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892.

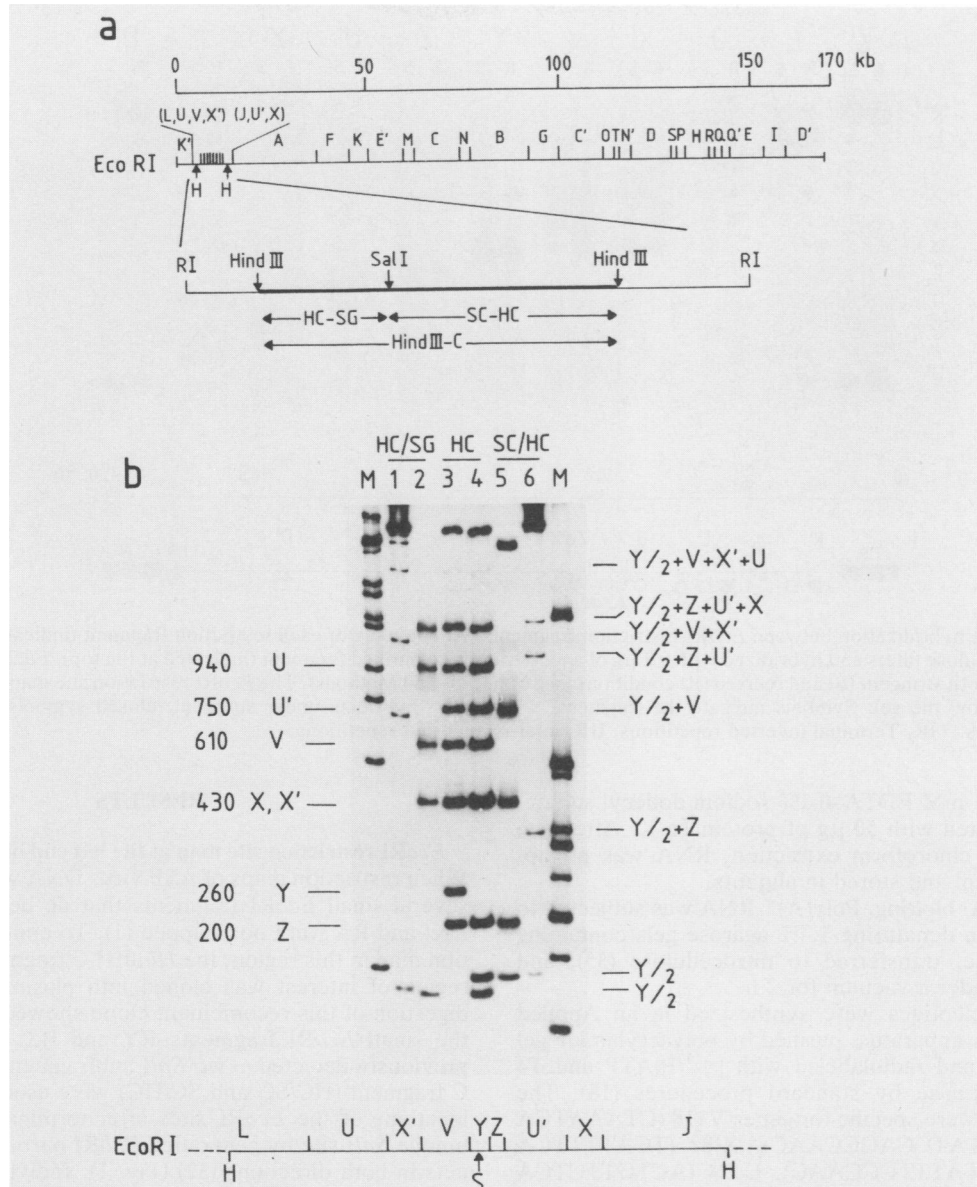


FIG. 1. *Eco*RI restriction site map of the left region of ASF virus DNA. (a) *Eco*RI restriction site map of ASF virus DNA. The expanded region shows the locations of the fragments used in the mapping experiments. (b) Mapping of the low-molecular-weight *Eco*RI fragments. ³²P-labeled total or partial digestion products of the different clones were run in a polyacrylamide gel, dried, and exposed for autoradiography. Lanes: 2, 3, and 5, purified fragments HC/SG, HC, and SC/HC, respectively, digested with *Eco*RI and labeled by end filling with Klenow enzyme; 4, same as lane 3, digested with *Eco*RI and *Sal*I; 1 and 6, recombinant plasmids harboring fragments HC/SG and SC/HC, respectively, labeled at the *Sal*I site and partially digested with *Eco*RI; M, terminally labeled lambda *Eco*RI-*Hind*III fragments (left) and pBR322 *Hin*fI fragments (right) as markers. Y₂ refers to any of the *Sal*I subfragments of *Eco*RI fragment Y. Products of the partial *Eco*RI digestion are indicated on the right. Numbers on the left indicate sizes, in nucleotides, of the *Eco*RI fragments. The deduced order of *Eco*RI fragments is shown at bottom.

salmon sperm DNA per ml for 40 h. After hybridization, the membranes were washed in 2× SSC at either 68°C (stringent hybridizations) or 50°C (relaxed hybridizations).

DNA sequencing. DNA sequencing was carried out by either the chemical degradation procedure (19) or the chain termination procedure (30). The nucleotide sequence was obtained from both strands of ASF virus DNA restriction fragments cloned in pUC plasmids or M13mp phages (21, 22).

RNA isolation. Viral RNA synthesized in vitro was obtained from purified ASF virus as reported elsewhere (29) and enriched in poly(A)⁺ mRNA by oligo(dT) chromatogra-

phy (18). Early RNA was purified from cultures infected in the presence of 250 µg of cycloheximide per ml at 4 h postinfection. Late RNA was purified from infected cultures at 14 h postinfection. Total RNA was isolated by the guanidinium hydrochloride procedure (5). Briefly, cells from three roller bottle cultures were harvested and sedimented in a Sorvall GS3 rotor at 5,000 × g for 10 min at 4°C. The pellet was suspended in 15 ml of 7 M guanidine hydrochloride–2% 2-mercaptoethanol–0.1 M sodium acetate and disrupted. RNA was then purified by centrifugation at 35,000 rpm for 20 h at 18°C through a cushion of 5.7 M CsCl–25 mM sodium acetate, pH 5.2. The pellet was suspended in 1 ml of 0.3 M

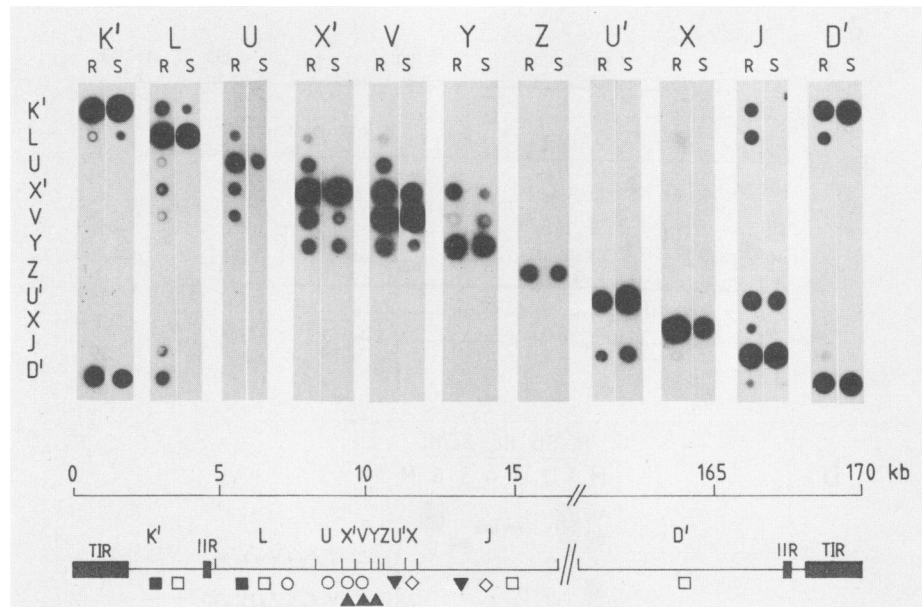


FIG. 2. Dot blot hybridization between *EcoRI* restriction fragments. About 20 ng of each restriction fragment (indicated on the left) was applied on nitrocellulose filters and hybridized with 0.1 μ g of a nick-translated purified fragment (indicated at the top). Each hybridization was carried out under both stringent (S) and relaxed (R) conditions (see Materials and Methods). The *EcoRI* restriction site map of ASF virus DNA ends is shown below the gel. Symbols indicate the fragments which cross-hybridize under stringent (closed symbols) or relaxed (open symbols) conditions. TIR, Terminal inverted repetitions; IIR, internal inverted repetitions.

sodium acetate–1 mM EDTA–0.1% sodium dodecyl sulfate, pH 5.2, and treated with 50 μ g of proteinase K. After two phenol and one chloroform extraction, RNA was precipitated with ethanol and stored in aliquots.

Northern (RNA) blotting. Poly(A)⁺ RNA was subjected to electrophoresis in denaturing 1.4% agarose gels containing 6% formaldehyde, transferred to nitrocellulose (35), and baked at 80°C under a vacuum for 2 h.

Oligodeoxynucleotides were synthesized in an Applied Biosystems 380A apparatus, purified by polyacrylamide gel electrophoresis, and radiolabeled with [γ -³²P]ATP and T4 polynucleotide kinase by standard procedures (18). The oligonucleotides were specific for genes V118 (CTAAATCA TACATACATTTATCCAGCCAAC), X'82 (TCATTTCA AGAATTGTTGTATTTCCAAC), U124 (ACTGTTCTCA ATAATCAATGGCATGCTCTC), U104 (GGTGAATCATA CAGTGTTCATGGGATAGC), L270A (GTAGTAGATAC AGAATCATTGCGACGATA), and L270B (TATCATACC ATTTAAAGTGTGTAGGTTGGG). Hybridization was done in 5 \times SSC–5 \times Denhardt solution–0.5% sodium dodecyl sulfate–10 mM Tris hydrochloride (pH 7.5) at 45°C. The equivalent hybridization temperature was adjusted for each oligonucleotide probe to 20°C below the predicted T_m value by the addition of formamide. Washes were done at 10°C below the predicted T_m for each oligonucleotide.

Computer analysis. Analysis of DNA sequences was performed with the software package of the University of Wisconsin genetics computer group (8) in a DEC VAX 730 computer running under the VMS operating system. The programs COMPARE and BESTFIT were used to carry out dot matrix comparisons (16) and alignment of DNA sequences (33). Searches of the National Biomedical Research Foundation were done with the WORDSEARCH program of the University of Wisconsin software package (40) and the FASTA program (25).

Secondary structure predictions were done by following the Chou and Fasman procedure (6).

RESULTS

***EcoRI* restriction site map at the left end of ASF virus DNA.** When restriction maps of ASF virus DNA were constructed, several small *EcoRI* fragments that lie between fragments RK' and RA were not mapped (1). To complete the restriction map in this region, the *HindIII* C fragment spanning the region of interest was cloned into plasmid pUC8. *EcoRI* digestion of this recombinant clone showed the presence of the small *EcoRI* fragments RY and RZ, which were not previously detected. Two *SalI* subfragments of the *HindIII* C fragment (HC/SG and SC/HC) were used to establish the locations of the *EcoRI* sites after terminal labeling at the unique *SalI* site by generating *EcoRI* partial digestion products in both directions (32) (Fig. 1). Individual *EcoRI* fragments U, U', V, X, X', Y, and Z were cloned in pUC plasmids. These fragments, along with cloned *EcoRI* fragments from the rest of the viral DNA, include the complete ASF virus genome, except for small fragments deleted by S1 nuclease treatment in the cloning process of the terminal fragments (14).

Repeated sequences in ASF virus DNA. The presence of repeated sequences in ASF virus DNA was investigated by a dot hybridization assay. *EcoRI* restriction fragments purified from recombinant clones which covered about 99% of the ASF virus genome were blotted onto nitrocellulose. Each *EcoRI* fragment was then used to probe these blots under either relaxed or moderately high stringency conditions. Figure 2 shows the results obtained with fragments which gave a positive cross-hybridization. The hybridization between the terminal *EcoRI* fragments K' and D' was expected, because these fragments are known to contain terminal and internal inverted repetitions (34). Several additional groups of cross-hybridizing fragments were identified. One of the groups included *EcoRI* fragments K', L, J, and D' and is described elsewhere (11). The other groups included restriction fragments located only at the left end of the ASF

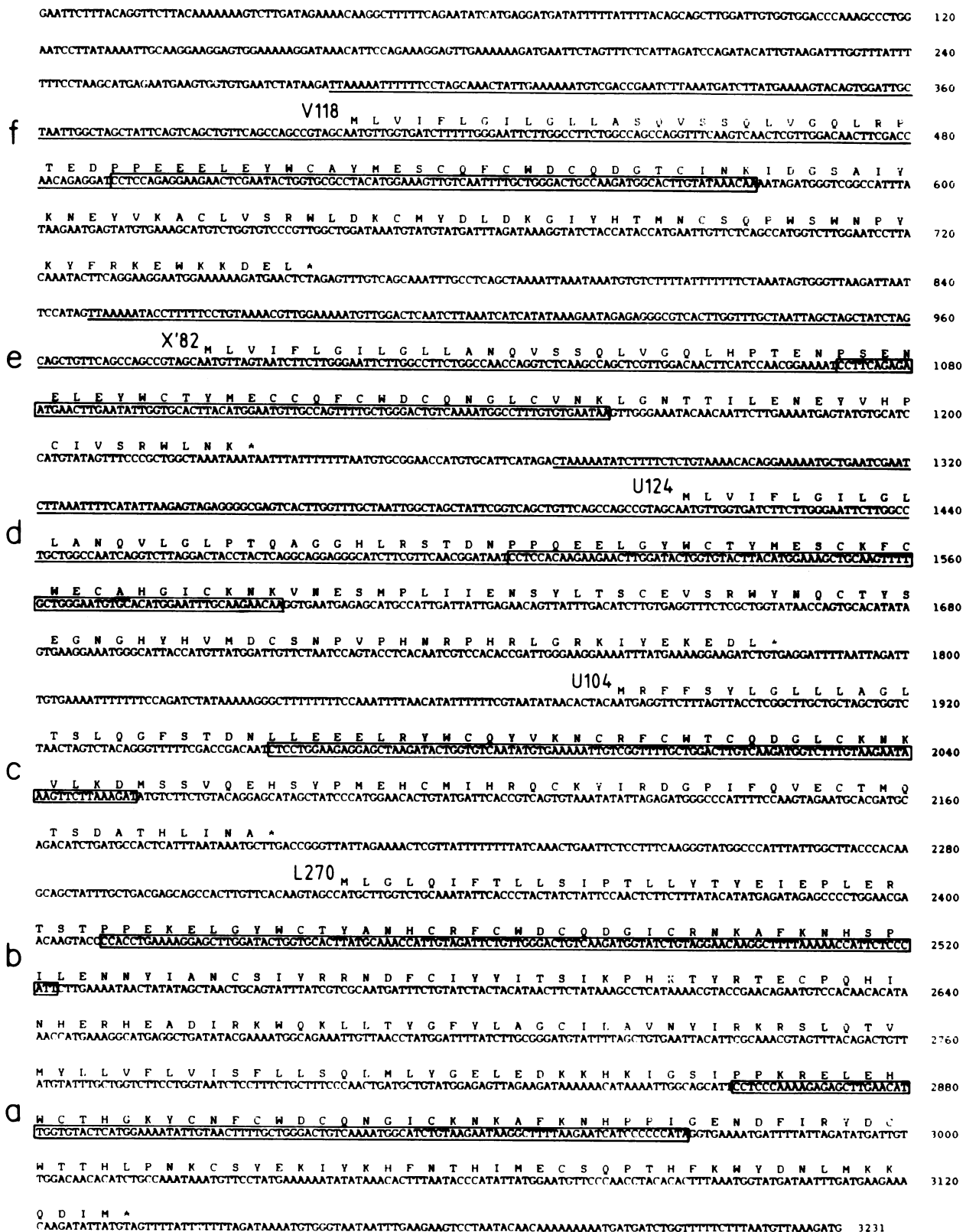


FIG. 3. DNA sequence of the region located between 7.5 and 11 kb from the left end of ASF virus DNA. The sequence is shown in the 5'-to-3' direction and from right to left according to the restriction map. Underlined sequences correspond to the block 2 repeated sequences, and boxed sequences correspond to the block 1 repeated sequences. The derived protein sequence from each reading frame is shown in the single-letter amino acid code.

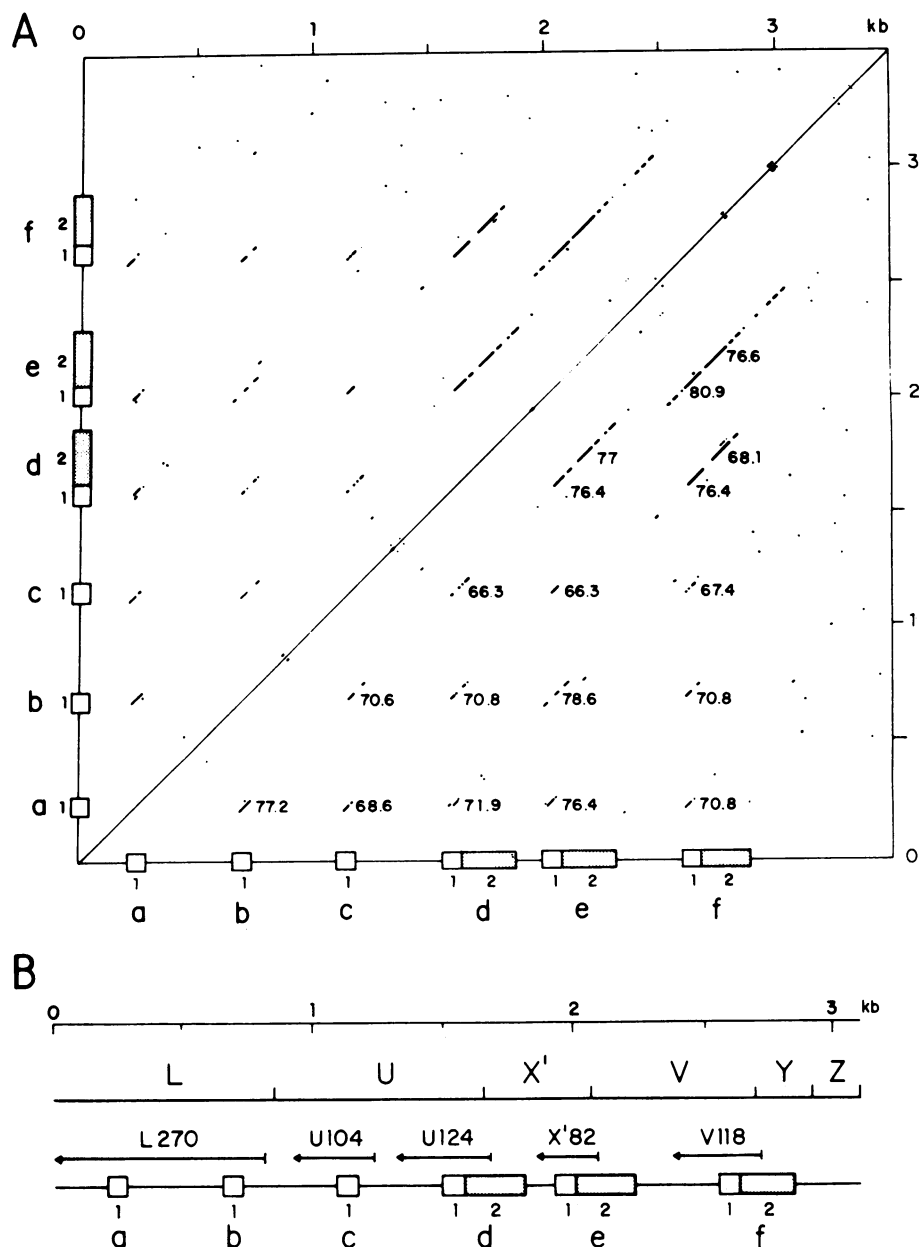


FIG. 4. Repeated sequences within the sequenced region. (A) Dot matrix analysis. The sequence is presented on both the horizontal and vertical axes. The analysis was performed with a window size of 20 nucleotides, allowing a maximum of five mismatches. The numbers near the diagonals indicate the percent identity of the different repetitions. (B) Location of the repeated sequences. The positions of the repeated sequences and ORFs are shown below the *EcoRI* restriction map.

virus genome. Two groups, one composed of *EcoRI* fragments L, U, X', and V (relaxed hybridization) and the other composed of fragments X', V, and Y (stringent hybridization), were in a contiguous DNA segment located between 5 and 11 kb from the left DNA end. In addition, fragment J hybridized with fragments U' and X. However, fragments U' and X did not cross-hybridize, indicating that their hybridization with fragment J may be due to two unrelated sequences.

Nucleotide sequence of region located 7.5 to 11 kb from the left DNA end. The hybridization results suggested that the region of ASF virus DNA encompassing *EcoRI* fragments L, U, X', V, and Y contains at least two different kinds of

repeated sequences. These were identified by sequencing a 3,230-nucleotide DNA segment. This sequence (Fig. 3) includes 993 nucleotides of fragment L and the complete sequence of fragments U, X', V, Y, and Z.

A search for repeated sequences was performed by dot matrix analysis (Fig. 4). From this analysis, six direct repetitions were found and named a through f, reading from left to right according to the restriction map. No significant inverted repeats were detected. Repeats d, e, and f were composed of two blocks of sequences: block 1, which was similar to repeats a, b, and c; and block 2, which was shared only by repeats d, e, and f.

The direct repeats described here account for the cross-

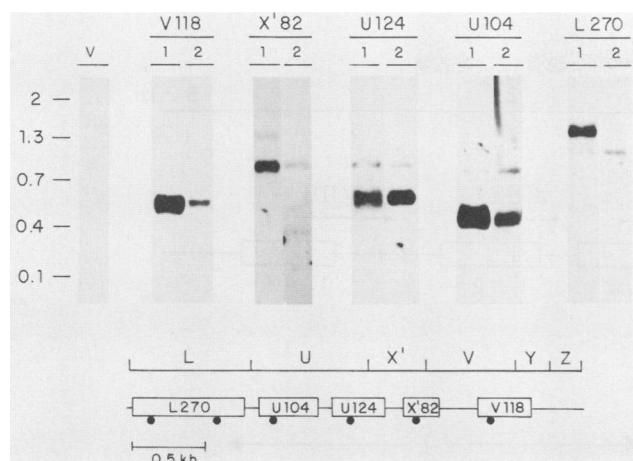


FIG. 5. Transcription of multigene family 110. Poly(A)⁺ RNA (0.15 μ g) synthesized *in vitro* by ASF virus (lanes 1) or total early RNA (20 μ g) isolated from virus-infected Vero cells (lanes 2) was hybridized as described in Materials and Methods with oligonucleotide probes specific for each gene. The dots in the lower part of the figure indicate the location of the sequence of each probe within the genes. Only one of the two hybridizations performed with oligonucleotides specific for the L270 gene, both of which gave identical results, is shown. Lane V, 30 μ g of total RNA from uninfected Vero cells hybridized with any of the probes. Sizes (in kilobases) of molecular weight markers are indicated.

hybridization between fragments in this region. Hybridizations under relaxed conditions between *Eco*RI fragments L, U, X', and V are explained by the presence of block 1 in all these fragments. Additional hybridizations under more stringent conditions between fragments X', V, and Y are consistent with the presence of block 2 in these fragments.

Multigene family 110. The coding capacity of the sequenced region in both strands showed the existence of five open reading frames (ORFs) in the leftward direction (Fig. 3). Each ORF was named according to the *Eco*RI fragment in which most of it was contained and the number of coding triplets in the frame. ORFs U124, X'82, and V118 started at equivalent positions within block 2. Interestingly, all ORFs contained a repeated sequence (block 1) between 70 and 100 bases from the ATG start codon. ORF L270 was more than twice as long as the others and included two block 1 repetitions (repetitions a and b). Since there was an ATG codon 94 nucleotides before the second block 1 (repeat a) of ORF L270, it is possible that this longer frame was generated by mutation of the stop codon separating two standard-sized members of the family. We named this group of ORFs family 110 because of the average length of the frames in coding triplets.

Transcription of family 110. To study whether the ORFs were actively transcribed genes, Northern blot hybridizations were performed with oligodeoxynucleotide probes specific for each gene. The sequence of the oligonucleotides was selected from the coding sequence of each gene and was more than 30% different from any other sequence in this region. Figure 5 shows the results of hybridizing these oligonucleotide probes to RNAs synthesized *in vitro* by ASF virions and to RNAs isolated from infected cells early during infection. These results indicated that all members of multigene family 110 were transcribed as early mRNAs. The two oligoprobes specific for L270 recognized the same RNA band. The sizes of the transcripts were consistent with the lengths of the reading frames, ranging from 0.4 kb (for the

U104 transcript) to 1.3 kb (for the L270 transcript). The major L270 transcript in the *in vitro* RNA was larger than the *in vivo* transcript. Also, differences between *in vitro* and early mRNAs were detected for low-abundance transcripts of other genes. The mRNAs for the family 110 genes were significantly less abundant late in infection (data not shown), suggesting that these early mRNAs are degraded during the course of the infection.

Comparison of putative protein products. The amino acid sequences predicted from each ORF of multigene family 110 were aligned by the insertion of gaps to maximize similarity. The sequence of L270 was split in two parts (L270A and L270B), with the second part starting at an internal methionine. Figure 6 shows the multiple alignment of the six sequences, in which several features of the multigene family are noteworthy. The most conserved part of the sequences was between residues 34 and 64. In this segment, a group of cysteine residues arranged in the pattern C-X₅-C-X₂-C-X₂-C-X₄-C was strictly conserved among the members of the multigene family. A striking feature of the six sequences was the conservation of cysteine residues, with the exception of two cysteines in the carboxy-terminal region which were not present in X'82. Gene X'82 has a truncated structure as a consequence of a deletion affecting the 3' portion of a larger gene (R. Blasco and E. Viñuela, unpublished results).

The percentage of identity between the protein sequences of the multigene family 110 is shown in Table 1. The hydropathy profiles and secondary structure predictions for these proteins are shown in Fig. 7. All the sequences have a highly hydrophobic amino-terminal region which may represent a leader sequence (20).

Similarity searches. A search of the National Biomedical Research Foundation protein data base and the GenBank DNA data base was performed for each of the sequences of the multigene family. From this search, no significant similarity with any entry in the databases could be found. However, when a search was carried out by looking for similarities to the conserved cysteine-rich domain, a significant resemblance to the sequence of a bovine posterior pituitary peptide (26, 27) was found. Figure 8 shows the coincidence of 6 residues (including four cysteines) within a 12-residue stretch in both sequences.

DISCUSSION

A screening of the ASF virus genome in hybridization experiments has shown the existence of some regions in ASF virus DNA which contain repeated sequences. These regions mapped within about 18 kb from the left and 11 kb from the right DNA end. Within these two regions, a complex pattern of cross-hybridizations between fragments can be seen (Fig. 2). The map location of the fragments, which include repeated sequences, coincides remarkably with that of variable regions in ASF virus DNA (3). Variation in ASF virus DNA ends is mainly due to deletions, which are the results of homologous recombination between repeated sequences or of nonhomologous recombination (4).

We have determined the sequence of a cluster of direct repeats located in *Eco*RI fragments L, U, X', V, and Y (Fig. 3). Inspection of the sequence by dot matrix comparisons (Fig. 4) showed that the repeated sequences consisted of two types of blocks, which fully explains the hybridization results. This sequence contained a family of homologous genes (multigene family 110) which partially overlapped the repeated sequences. All these genes were actively transcribed at early times after infection. Early mRNAs specific

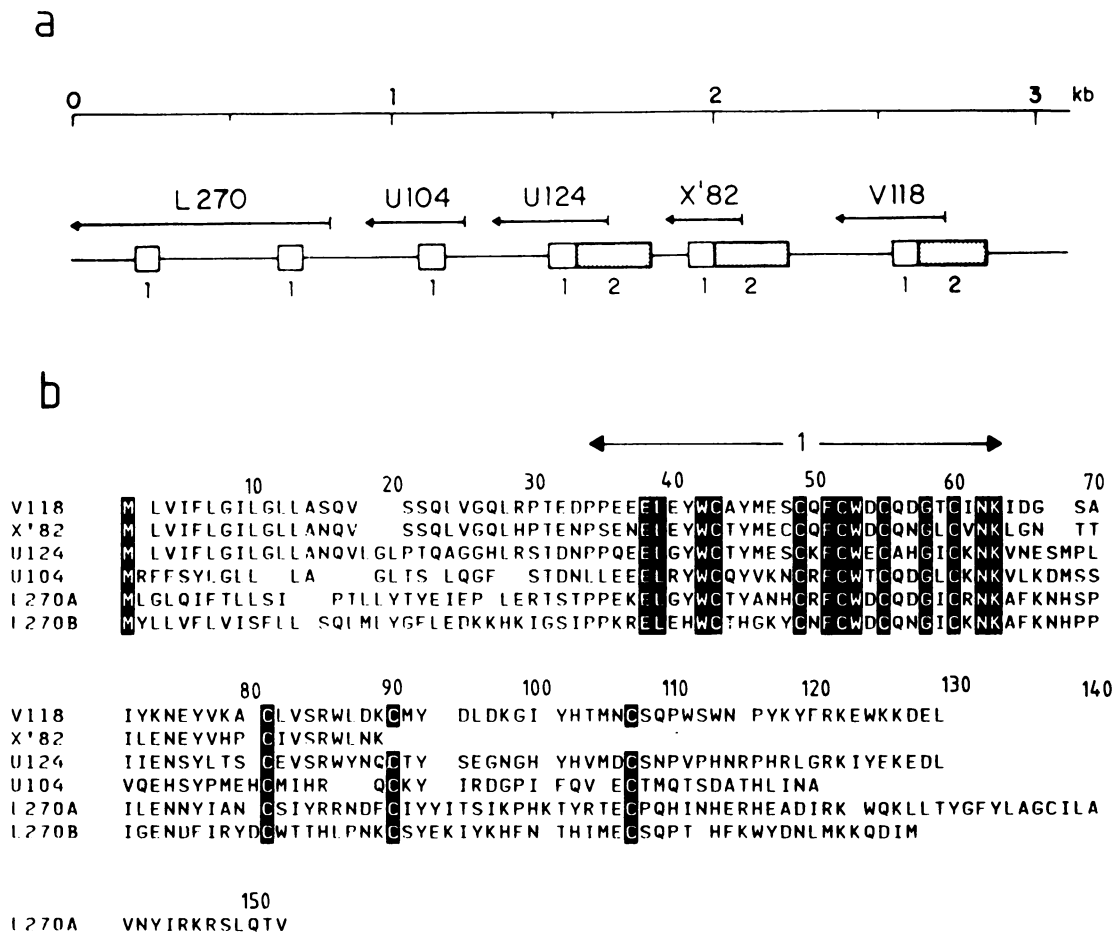


FIG. 6. Alignment of amino acid sequences of multigene family 110. (a) Relative positions of the repeated blocks and the genes. (b) Multiple alignment. The sequence of each protein of the multigene family was aligned to maximize identities. The sequence of L270 is split in two portions (L270A and L270B; see text). Boxes enclose residues conserved in all the sequences being compared. Arrows show the extent of the block 1 repeated sequences.

for this region of the ASF virus genome have been shown to induce the synthesis of a 12.5-kilodalton protein in an in vitro translation system (29). Therefore, at least some of the transcripts (presumably those of genes U104, U124, and V118) can be translated into proteins. The presence in the multigene family of a gene (L270) showing an organization dimeric in primary structure as well as in hydropathy profile could be the result of the in-frame fusion of two contiguous genes by mutation of a stop codon. It is not clear whether this gene is functional, since although a full-size transcript is synthesized, no protein of the corresponding size can be detected after in vitro translation (29). The major transcript of the L270 gene synthesized in infected cells at early times

TABLE 1. Homology among protein products of multigene family 110

Protein product	% Identity with:				
	X'82	U124	U104	L270A	L270B
V118	73.2	55.1	35.6	36.5	32.2
X'82		61.0	41.9	41.2	37.8
U124			38.8	37.5	30.8
U104				35.9	29.8
L270A					34.4

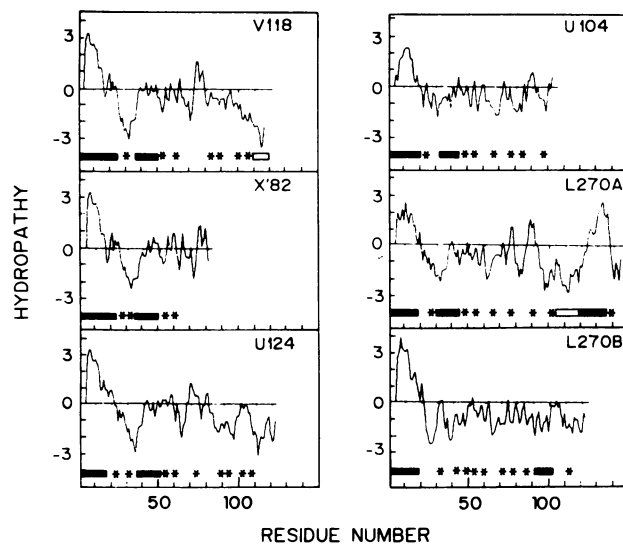


FIG. 7. Hydropathy profiles and secondary structure predictions for the members of multigene family 110. Hydropathy profiles were calculated according to the method of Kyte and Doolittle (12a), with a window size of 5 amino acids. Secondary structure predictions were carried out by the Chou and Fasman method (6), and only predictions longer than 10 residues are shown. Symbols: *, beta turn; □, alpha helix; ■, beta-pleated sheet.

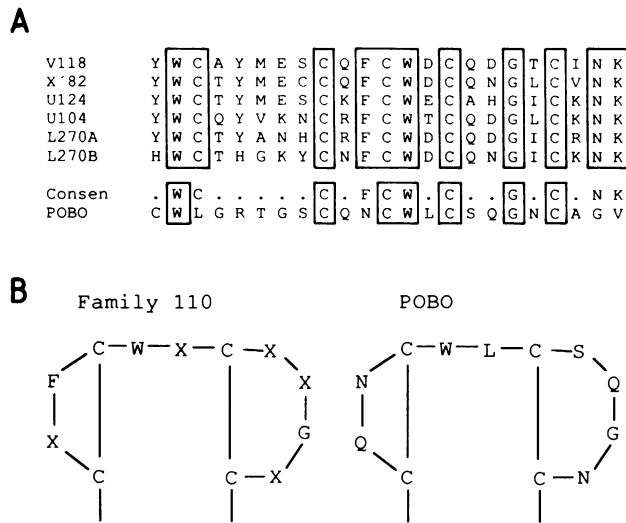


FIG. 8. Sequence similarity between multigene family 110 and a peptide from the bovine posterior pituitary (POBO). (A) Consensus sequence of multigene family 110 (consen) and sequence of the posterior pituitary peptide. Conserved residues are boxed. Dots represent amino acids that are not identical in all members of multigene family 110. (B) Predicted structures. Positions of the disulfide bridges are represented according to those present in the posterior pituitary peptide (26).

was smaller than that of the most abundant transcript synthesized *in vitro*. In addition, the relative abundance of the transcripts was different in *in vitro* and early RNAs. These two observations suggest that although the extracellular virions are able to synthesize *in vitro* mRNA from early genes, another factor(s) may influence early transcription in infected cells.

Multigene families have been described in human cytomegalovirus (39), Epstein-Barr virus (2), and Shope fibroma virus (36).

Comparison of the amino acid sequences derived from multigene family 110 (Fig. 6) shows the presence of a highly hydrophobic amino-terminal sequence, a cysteine-rich central region which is the most conserved part, and a more divergent carboxy-terminal portion. The pattern of variation among these sequences, in which several cysteines and aromatic residues are strictly conserved, suggests that selective pressure has acted on each gene during the evolution of the multigene family.

The sequence similarity between the conserved, cysteine-rich domain and the sequence of a peptide isolated from the bovine posterior pituitary (Fig. 8) is intriguing. The similarity involves cysteine, tryptophan, and glycine residues which are strictly conserved among the members of the multigene family. The cysteine residues in the bovine posterior pituitary peptide which are involved in disulfide bridge formation are present with the same spacing in the sequences of multigene family 110. Therefore, the proteins of the multigene family are likely to fold in analogous ways and produce similar structures. It seems difficult to assess the biological meaning, if any, of this similarity, since the function of the peptide is unknown.

Regions containing cysteines and aromatic residues have been shown to form metal-binding structures. In some cases, they generate a "zinc finger" domain (23), but in other cases, the metal ion serves to stabilize protein-protein interactions (10).

The arrangement of the members of multigene family 110 in the viral genome, as well as their relationship, suggests that the family originated by duplication of individual genes. The maintenance of multigene families in ASF virus could be a consequence of an evolutionary advantage conferred by the presence of homologous genes.

ACKNOWLEDGMENTS

We are grateful to R. F. Doolittle and W. M. Fitch for helpful discussions, C. López-Otín for help in secondary structure predictions, and R. J. Yáñez for help in the analysis of DNA sequences. We also thank R. Sánchez and A. Zazo for technical assistance.

This work was supported by grants from the Comisión Interministerial de Ciencia y Tecnología, European Economic Community, and Fundación Ramón Areces.

LITERATURE CITED

- Almendral, J. M., R. Blasco, V. Ley, A. Beloso, A. Talavera, and E. Viñuela. 1984. Restriction site map of African swine fever virus. *Virology* 133:258-270.
- Baer, R., A. Bankier, M. Biggins, P. Deininger, P. Farrel, G. Gibson, G. Hatfull, G. Hudson, C. Satchwell, C. Sequin, P. Fuffnell, and B. Barrel. 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature (London)* 310:207-211.
- Blasco, R., M. Agüero, J. M. Almendral, and E. Viñuela. 1989. Variable and constant regions in African swine fever virus DNA. *Virology* 168:330-338.
- Blasco, R., I. de la Vega, F. Almazán, M. Agüero, and E. Viñuela. 1989. Genetic variation of African swine fever virus: variable regions near the ends of the viral DNA. *Virology* 173:251-257.
- Chirgwin, J. M., A. E. Przbyla, R. J. McDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294-5299.
- Chou, P. Y., and F. D. Fasman. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol. Relat. Areas Mol. Biol.* 47:45-148.
- De Boer, C. J. 1967. Studies to determine neutralizing antibody in sera from animals recovered from African swine fever and laboratory animals inoculated with African virus with adjuvants. *Arch. Gesamte Virusforsch.* 20:164-179.
- Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387-395.
- Enjuanes, L., A. L. Carrascosa, and E. Viñuela. 1976. Isolation and properties of the DNA of African swine fever virus. *J. Gen. Virol.* 32:479-492.
- Frankel, A. D., D. S. Brecht, and C. O. Pabo. 1988. Tat protein from human immunodeficiency virus forms a metal-linked dimer. *Science* 240:70-73.
- González, A., V. Calvo, F. Almazán, J. M. Almendral, J. C. Ramírez, I. de la Vega, R. Blasco, and E. Viñuela. 1990. Multigene families in African swine fever virus: family 360. *J. Virol.* 64:2073-2081.
- González, A., A. Talavera, J. M. Almendral, and E. Viñuela. 1986. Hairpin loop structure of African swine fever virus DNA. *Nucleic Acids Res.* 14:6835-6844.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132.
- Langridge, J., P. Langridge, and P. L. Bergquist. 1980. Extraction of nucleic acids from agarose gels. *Anal. Biochem.* 103:246-271.
- Ley, V., J. M. Almendral, P. Carbonero, A. Beloso, E. Viñuela, and A. Talavera. 1984. Molecular cloning of African swine fever virus DNA. *Virology* 133:249-257.
- López-Otín, C., C. Simón, E. Méndez, and E. Viñuela. 1988. Mapping and sequence of the gene encoding protein p37, a major structural protein of African swine fever virus. *Virus*

- Genes 1:291-303.
16. **Maizel, J. V., and R. P. Lenk.** 1981. Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proc. Natl. Acad. Sci. USA* **78**:7665-7669.
 17. **Malmquist, W. A.** 1963. Serologic and immunologic studies with African swine fever virus. *Am. J. Vet. Res.* **24**:450-459.
 18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 19. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**:560-564.
 20. **McGeoch, D. J.** 1985. On the predictive recognition of signal peptide sequences. *Virus Res.* **3**:271-286.
 21. **Messing, J.** 1983. New M13 vectors for cloning. *Methods Enzymol.* **101**:20-78.
 22. **Messing, J., and J. Vieira.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**:259-268.
 23. **Miller, J., A. McLachan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**:1609-1614.
 24. **Murphy, G. L., T. D. Connell, D. S. Barritt, M. Koomey, and J. G. Cannon.** 1989. Phase variation of gonococcal protein II: regulation of gene expression by slipped-strand mispairing of repetitive DNA sequence. *Cell* **56**:539-547.
 25. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**:2444-2448.
 26. **Preddie, E. C.** 1965. Structure of a large polypeptide of bovine posterior pituitary tissue. *J. Biol. Chem.* **240**:4194-4203.
 27. **Preddie, E. C., and M. Saffran.** 1965. Isolation of a large polypeptide from bovine posterior pituitary powder. *J. Biol. Chem.* **240**:4189-4193.
 28. **Rigby, P., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick-translation with DNA polymerase I. *J. Mol. Biol.* **113**:237-251.
 29. **Salas, M. L., J. Rey-Campos, J. M. Almendral, A. Talavera, and E. Viñuela.** 1986. Transcription and translation maps of African swine fever virus. *Virology* **152**:228-240.
 30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
 31. **Santarén, J. F., and E. Viñuela.** 1986. African swine fever virus-induced polypeptides in Vero cells. *Virus Res.* **5**:391-405.
 32. **Smith, H. O., and M. L. Birnstiel.** 1976. A simple method for DNA restriction site mapping. *Nucleic Acids Res.* **3**:2387-2398.
 33. **Smith, T. F., and M. S. Waterman.** 1981. Identification of common molecular subsequences. *J. Mol. Biol.* **147**:195-197.
 34. **Sogo, J. M., J. M. Almendral, A. Talavera, and E. Viñuela.** 1984. Terminal and internal inverted repetitions in African swine fever virus DNA. *Virology* **133**:271-275.
 35. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA* **77**:5201-5205.
 36. **Upton, C., and G. McFadden.** 1986. Tumorigenic poxviruses: analysis of viral DNA sequences implicated in the tumorigenicity of Shope fibroma virus and malignant rabbit virus. *Virology* **152**:308-321.
 37. **Viñuela, E.** 1985. African swine fever virus. *Curr. Top. Microbiol. Immunol.* **116**:151-170.
 38. **Viñuela, E.** 1987. Molecular biology of African swine fever virus, p. 31-49. *In* Y. Becker (ed.), *African swine fever.* Martinus Nijhoff, Boston.
 39. **Weston, K., and B. G. Barrel.** 1986. Sequence of the short unique region, short repeats, and part of the long repeats of human cytomegalovirus. *J. Mol. Biol.* **192**:177-208.
 40. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA* **80**:726-730.