The identification of an African swine fever gene with conserved helicase motifs and a striking homology to herpes virus origin binding protein, UL9

M.D.Sussman, Z.Lu, G.F.Kutish, C.A.Afonso and D.L.Rock*
Plum Island Animal Disease Center, Agricultural Research Service, USDA, PO Box 848, Greenport, NY 11944, USA

Received March 12, 1993; Accepted April 1, 1993 GenBank accession no. L12174

African swine fever virus (ASFV), the causative agent of African swine fever in domestic pigs (3, 8, 12), is a large icosahedral virus with a double stranded DNA genome of 170–190 Kb that shares many characteristics with poxviruses. Although ASFV and poxviruses are distinct morphologically, they both replicate in the cytoplasm, exhibit temporal regulation of gene expression, and have similar genome structures, which include terminal inverted repeats, terminal crosslinks, a central conserved region, and variable regions at each end of the genome (1, 6). In contrast to the poxviruses, ASFV replication includes a distinct nuclear stage where early viral DNA replication occurs (5). ASFV is the sole member of an unnamed family of viruses.

The open reading frame (ORF), LMW6DL (1102 amino acids), located on the SalI 'C' fragment 0.28 m.u. from the left terminus of the genome of ASFV isolate Malawi-Lil 20/1 (4) contains a large central to carboxy terminal region which bears significant similarity to UL9 of herpes simplex 1-HSV1 (11), ORF 51 of *Varicella Zoster* virus-VZV (2), ORF 53 of equine herpes virus-EHSV (14) and gene 19R of human herpes virus 6-HH6 (9). Furthermore, each of six highly conserved motifs identified for the helicase superfamily (7) and required for function of the origin binding protein (UL9) of HSV1 are conserved in LMW6DL (10). Sequence analysis with the FASTA algorithm (13) shows LMW6DL most closely related to HH6 (28.5% amino acid identity/341 residues; 51.3% similarity, FASTA = 289, z = 25.2), followed by VZV (25.9% amino acid identity/312 residues; 51.7% similarity, FASTA = 177, z = 15.4), HSV1 (27.5% amino acid identity/325 residues; 50.7% similarity, FASTA = 126, z = 10.6) and EHSV1 (26.8% amino acid identity/325 residues; 51.2% similarity, FASTA = 139, z = 11.5). The alignment of each of the six highly conserved motifs is shown in Figure 1. UL9 of Herpes simplex virus is essential for viral DNA replication, codes for a protein that specifically recognizes sequences within the viral origin of replication, and possesses independent helicase and DNA-dependent ATPase activities (10). The striking relationship between LMW6DL and the Herpes virus origin binding protein indicates a possible role for this gene product in ASFV DNA replication.

REFERENCES

1. Costa,J.V. (1990) In Darai,G. (ed.), *Molecular Biology of Iridoviruses*. pp. 247–270.
2. Davison,A.J. and Scott,J.E. (1986) *J. Gen. Virol.* **67**, 1759–1816.
3. DeKock,G., Robinson,E.M. and Keppel,J.J.G. (1940) *Ondertepoort, J. Vet. Sci. Animal Ind.* **14**, 31–93.
4. Dixon,L. (1988) *J. Gen. Virol.* **69**, 1683–1694.
5. Garcia-Beato,R., Salas,M.L., Vinuela,E. and Salas,J. (1992) *Virology* **188**, 637–649.
6. Gonzalez,A., Talavera,A., Almendral,J.M. and Vinuela,E. (1986) *Nucleic Acids Res.* **14**, 6835–6844.
7. Gorbalenya,A.E., Koonin,E.V., Donchenko,A.P. and Binov,V.M. (1989) *Nucleic Acids Res.* **17**, 4713–4730.
8. Hess,W.R. (1971) *Virol. Monographs* **9**, 1–33.
9. Lawrence,G.L., Chee,M., Craxton,M.A., Gompels,U.A., Honess,R.W. and Barrell,B.G. (1990) *J. Virol.*, **64**, 287–299.
10. Martinez,R., Shao,L. and Weller,S. (1992) *J. Virol.* **66**, 6735–6746.
11. McGeoch,D.J., Dalrymple,M.A., Dolan,A., McNab,D., Perry,L.J., Taylor,P. and Challberg,M.D. (1988) *J. Virol.* **62**, 444–453.
12. Mebus,C.A. (1988) *Adv. Virus Res.* **35**, 251–269.
13. Pearson,W.R. (1990) *Methods Enzymol.* **183**, 63–98.
14. Telford,E.A., Watson,M.S., McBride,K. and Davison,A.J. (1992) *Virology* **189**, 304–316.

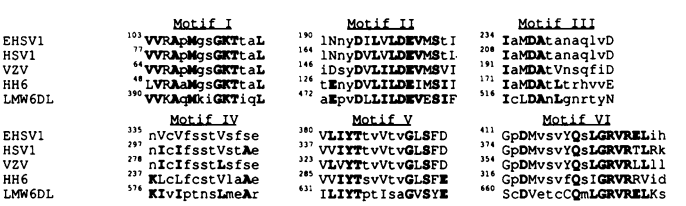

Figure 1. Six highly conserved motifs of DNA and RNA helicases of LMW6DL. Residues which are identical to LMW6DL are shown as bold upper case letters, while conservative substitutions are depicted by upper case letters only. Numbers to the left indicate the starting amino acid of each motif in the protein. Protein names are given in the text.

* To whom correspondence should be addressed