ASFV Genome Replication and Packaging

Subjects: Virology

Contributor: Ana Catarina Urbano

Genome condensation and packaging are essential processes in the life cycle of viruses. Mimivirus and many other nucleocytoplasmic large DNA virus (NCLDV) subfamilies have evolved a unique genome packaging mechanism that is comparable to chromosome segregation in bacteria and archaea and requires a number of specific enzymes, such as packaging ATPases, recombinases, DNA polymerases, helicases, and topoisomerases, as well as histones or histone-like proteins. Although the mechanisms of assembly and genome encapsidation in African Swine Fever Virus (ASFV) have not been fully characterized, the similarities in genome structure with Poxviruses and data from electron microscopy suggest that the ASFV packaging machinery is similar to Mimivirus and other NCLDVs. ASFV encodes up to 70 structural proteins, 16 of which, at least, are thought to be involved in assembly of the virus particle and include a predicted packaging A32L ATPase (B354L), a lambda-like recombinase (D345L), a type II topoisomerase (P1192R), and the histone-like DNA-binding protein pA104R.

Keywords: African swine fever ; ASFV ; genome packaging ; viral DNA-packaging proteins ; NCLDV

1. Viral DNA-Packaging Proteins

In cellular organisms, DNA-packaging proteins bind DNA and promote its bending, organizing it into highly compacted structures (called chromatin) which have a central role in the regulation of gene expression. Evolutionary analysis has shown that the primary DNA-packaging proteins involved in the organization of chromatin are different across the three domains of life. In bacteria, the primary DNA-packaging proteins are members of the HU/IHF superfamily (also called Type II DNA-binding proteins—DNABII—or bacterial histone-like proteins, pfam PF00216)[1]. Conversely, most eukaryotes and several archaea contain histones, highly basic proteins that form the characteristic octameric DNA structural (and functional) unit termed the nucleosome[2].

Double-stranded DNA viruses display a large variety of proteins that interact with host chromatin whose distribution seems to be influenced mainly by viral genome size and the domain to which the host of the virus belongs[3]. Smaller viruses (e.g., Papoviruses in eukaryotes, Salterproviruses in archaea, and Tectiviruses in bacteria) usually possess a minimal DNA replication apparatus consisting of a few multifunctional proteins that mediate several distinct interactions with host chromatin proteins and viral or host DNA[3]. Larger viruses such as the animal Adenoviruses and Herpesviruses, archaeal Lipothrixviruses and Baculoviruses, and several lineages of caudate bacteriophages possess distinctive virus-specific DNA-binding and/or adaptor domains (e.g., RRM-like domain in E2 from Papillomaviruses and LANA/EBNA1 from vertebrate Herpesviruses), and additionally encode several enzymes which catalyze chromatin status and chromosomal architecture (e.g., SWI2/SNF2 P-loop ATPases from archaeal Lipothrixviruses and several bacteriophages), and covalently modify chromatin components (e.g., SET domain histone methyltransferase from *Paramecium bursaria Chlorella virus*)[3].

The largest DNA-viruses (over 150 kb and extending up to 2.5 Mbp (*Pandoravirus salinus*)[4] typically encode hundreds to thousands of proteins. Eukaryotic viruses in this range include Polydnaviruses such as the wasp *Cotesia congregata Bracovirus*[5], and the large nucleocytoplasmic DNA virus (NCLDV) clade which consists of seven distinct families of eukaryotic dsDNA viruses, namely *Phycodnaviridae*, *Poxviridae*, *Asfarviridae* (ASFV), *Asco-* and *Iridoviridae*, *Mimiviridae*, *Marseilleviridae*[6], and the proposed novel *Medusaviridae*[7]. Mimivirus and many other NCLDV subfamilies have evolved a unique genome packaging mechanism that is comparable to chromosome segregation in bacteria and archaea [8][9] and requires a number of specific enzymes, such as packaging ATPases, recombinases, DNA polymerases, helicases, and topoisomerases, as well as histones or histone-like proteins [10]. Mimiviruses package their genome into preformed procapsids through a nonvertex portal driven by the vaccinia virus A32-type virion packaging ATPase [11], homologous to the bacterial FtsK/HerA family prokaryotic chromosome segregation and packaging motors [8][9], and use similar revolving mechanism for genome packaging [12]. The ATPase interacts with other genome packaging components such as recombinase/s, a type II topoisomerase, and possibly several as yet unidentified components to form a complex that is competent for both resolving the genomes into unit lengths and translocating them

into empty capsids [8][9]. In Vaccinia virus (*Poxviridae*) genome packaging is slightly different. The packaging ATPase complex collaborates with host type II topoisomerase for decatenation and genome replication after which the packaging ATPase assembly docks a copy of the genome onto the capsid vertex for packaging and leaves [9]. Although the mechanisms of assembly and genome encapsidation in ASFV have not been fully characterized, the similarities in genome structure with Poxviruses and the presence of replication intermediates consisting of head to head genome concatemers suggests they may share a similar replication model [13]. Further, data from electron microscopy indicates that in ASFV the viral DNA begins to condense into a pronucleoid and is then inserted, at a single vertex, into an empty particle which then goes through an intermediate phase of consolidation to produce the characteristic mature virions [13]. Thus, the overall composition of the ASFV packaging machinery is probably similar to Mimivirus and other NCLDVs.

2. ASFV Genome Structure

ASFV genomic organization also resembles that of other NCLDVs. The ASFV genome is a single molecule of linear double stranded DNA organized in a central relatively conserved and evolutionary stable region of about 125 kbp capped by two variable regions with a length of 38–47 kbp for the left, and 13–16 kbp for the right DNA ends [14]. Each DNA strand is covalently closed at both ends by a 37 nt hairpin loop followed by terminal inverted repeats of 2.1 kbp, which are characterized by numerous tandem repeat arrays. When examined in opposite polarities the AT-rich hairpin loops are inverted and complementary [15]. The genome varies in length between 170 and 190 kbp and encodes between 151 and 167 open reading frames (ORF) spaced closely along both chains of the viral DNA and separated by short intergenic regions. About half of them lack any known or probable function [16][17][18]. Differences in genome length mainly result from deletions or additions of up to 8.6 kbp in the left- and right-hand variable regions, with gain or loss of members of different multigene families (MGFs) [14][16]. Interestingly, MGFs do not share similarity to other known NCLDV genes. Transcription of viral genes is tightly regulated and acts as the main switch on ASFV gene expression in coordination with DNA replication. In total, 20 genes are currently considered to be involved in the transcription and modification of mRNAs, comprising approximately 20% of the ASFV genome. Four classes of mRNAs have been identified by their distinctive accumulation kinetics: immediate early and early genes, expressed before the onset of DNA replication, and intermediate and late genes, expressed after [18]. This transcriptional machinery gives ASFV precise configurational and temporal control of gene expression and considerable independence from its host.

Despite having a relatively low overall genomic mutation rate, the evolutionary rate of the ASFV variable regions seems to approach those of RNA viruses and can greatly affect ASFV genome structure [19][20]. ASFV strains showing MGFs gene duplication are often associated with a more violent phenotype while attenuating loss of MGFs occurs after viral passage in in vitro cell culture [16][20][21]. Comparative genomic analyses have also identified a range of genes in the constant region undergoing positive selection (e.g., CD2v/EP402R and C-type lectin/EP153R) that represent another source of genetic diversity among ASFV isolates [5][20][22]. These evolutionary processes are of considerable interest as they are key drivers of change in specific genes and encoded antigens, arguably influencing vaccination strategies and/or the stability of live attenuated vaccines, as well as diversity in host response.

3. ASFV Structural Proteins and Proteins Involved in Assembly

Up to 70 structural proteins have been identified from the ASFV virion, 16 of which, at least, are thought to be involved in assembly of the virus particle [23] (Table 1). These include the major capsid protein p72 (ORF B646L) which is essential for assembly of the icosahedral capsid on the inner envelope [24]; the mature proteins derived from polyprotein pp220 (CP2474L) and polyprotein pp62 (CP530R) who assemble to form the core shell that surrounds the DNA-containing nucleoid [25]; the membrane protein p17 (D117L), required for assembly of the capsid layer on the inner envelope [26]; the phosphoprotein p14.5 (E120R) a capsid component which mediates intracellular virus transport [27]; the enzyme responsible for polyprotein processing, encoded by ORF S273R (cysteine proteinase); and the major DNA-binding proteins p10 (K78R) and pA104R which are located in the nucleoid of mature ASFV particles [25], consistent with a role in this viral domain. pA104R, specifically, has, on the basis of knockdown experiments with small interfering RNAs [28], been shown to be involved in viral transcription, DNA replication, and genome packaging.

Although not detected in the ASFV proteome, a packaging A32L ATPase (B354L), which has orthologs in all NCLDVs has been predicted in ASFV [29], as well as a lambda-like recombinase (D345L) [10], which might be involved in processing DNA ends for strand exchange or single-strand annealing during recombination. The packaging ATPase of Mimivirus is also absent from proteomic analysis as, it has been suggested, it leaves the nonvertex packaging site after packaging and is probably reused [9]. These annotated protein sequences have yet to be functionally and biochemically characterized. The ASFV type II topoisomerase (P1192R) [30][31] is likewise absent from the ASFV proteome, but has been detected at

intermediate and late phases of infection in the cytoplasm of infected cells, accumulating in viral factories [32] and, it has been argued, may participate in genome segregation, by facilitating the separation of newly-replicated DNA molecules, as suggested for Mimivirus.

Table 1. African swine fever virus (ASFV) Structural proteins and proteins involved in assembly.

ORF	Description	Localization	Reference(s)
B646L	Major capsid protein p72	capsid	[24][33]
B438L	Protein p49	capsid	[34]
E120R	Protein p14.5	capsid	[27]
D117L	Major transmembrane protein p17	inner envelope	[26][35]
E183L	Transmembrane protein pE183L	inner envelope	[36][37]
CP2475L	Protein p5, polyprotein pp220 derived	core shell	[38]
CP2475L	Protein p14, polyprotein pp220-derived	core shell	[25][38][39][40]
CP2475L	Protein p34, polyprotein pp220-derived	core shell	[25][38][39][40]
CP2475L	Protein p37, polyprotein pp220-derived	core shell	[25][38][39][40]
CP2475L	Protein p150, polyprotein pp220-derived	core shell	[25][38][39][40]
CP530R	Protein p8, polyprotein pp62-derived	core shell [1]	[41]
CP530R	Protein p15, polyprotein pp62-derived	core shell	[25][39][42]
CP530R	Protein p35, polyprotein pp62-derived	core shell	[25][39][42]
S273R	Polyprotein processing protease	core shell	[43][44]
A104R	Histone-like DNA-binding protein	nucleoid	[28][38]
K78R	DNA-binding protein p10	nucleoid	[45]
P1192R	Topoisomerase II	ND	[30][31]
B354L	A32L ATPase [2]	ND	[29]
D345L	Lambda-like recombinase [2]	ND	[10]

[1] subviral localization inferred from their known or predicted role. [2] hypotheticals. ND—not determined

References

1. Chelikani, V.; Ranjan, T.; Zade, A.; Shukla, A.; Kondabagil, K. Genome Segregation and Packaging Machinery in Acanthamoeba polyphaga Mimivirus Is Reminiscent of Bacterial Apparatus. J. Virol. 2014, 88, 6069–6075.
2. Chelikani, V.; Ranjan, T.; Kondabagil, K. Revisiting the genome packaging in viruses with lessons from the “Giants”. Virology 2014, 466–467, 15–26.
3. Iyer, L.M.; Balaji, S.; Koonin, E.V.; Aravind, L. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. Virus Res. 2006, 117, 156–184.
4. Michel, B.; Fournier, G.; Lieffrig, F.; Costes, B.; Vanderplasschen, A. Cyprinid herpesvirus 3. Emerg. Infect. Dis. 2010, 16, 1835–1843.
5. Guo, P.; Zhao, Z. Biomotors: Linear, Rotation, and Revolution Motion Mechanisms, 1st ed.; CRC Press: Boca Raton, FL, USA, 2017; ISBN 9781351136068.
6. Dixon, L.K.; Chapman, D.A.G.; Netherton, C.L.; Upton, C. African swine fever virus replication and genomics. Virus Res. 2013, 173, 3–14.
7. Blasco, R.; Agüero, M.; Almendral, J.; Viñuela, E. Variable and constant regions in african swine fever virus DNA. Virology 1989, 168, 330–338.

8. González, A.; Talavera, A.; Almendral, J.M.; Viñuela, E. Hairpin loop structure of African swine fever virus DNA. Nucleic Acids Res. 1986, 14, 6835–6844.
9. Chapman, D.A.G.; Tcherepanov, V.; Upton, C.; Dixon, L.K. Comparison of the genome sequences of non-pathogenic and pathogenic African swine fever virus isolates. J. Gen. Virol. 2008, 89, 397–408.
10. Yáñez, R.J.; Rodríguez, J.M.; Nogal, M.L.; Yuste, L.; Enríquez, C.; Rodriguez, J.F.; Viñuela, E. Analysis of the Complete Nucleotide Sequence of African Swine Fever Virus. Virology 1995, 208, 249–278.
11. Rodríguez, J.M.; Salas, M.L. African swine fever virus transcription. Virus Res. 2013, 173, 15–28.
12. Michaud, V.; Randriamparany, T.; Albina, E. Comprehensive Phylogenetic Reconstructions of African Swine Fever Virus: Proposal for a New Classification and Molecular Dating of the Virus. PLoS ONE 2013, 8, e69662.
13. Malogolovkin, A.; Kolbasov, D. Genetic and antigenic diversity of African swine fever virus. Virus Res. 2019, 271, 197673.
14. Krug, P.W.; Holinka, L.G.; O'Donnell, V.; Reese, B.; Sanford, B.; Fernandez-Sainz, I.; Gladue, D.P.; Arzt, J.; Rodriguez, L.; Risatti, G.R.; et al. The Progressive Adaptation of a Georgian Isolate of African Swine Fever Virus to Vero Cells Leads to a Gradual Attenuation of Virulence in Swine Corresponding to Major Modifications of the Viral Genome. J. Virol. 2015, 89, 2324–2332.
15. Chapman, D.A.G.; Darby, A.C.; Da Silva, M.; Upton, C.; Radford, A.D.; Dixon, L.K. Genomic analysis of highly virulent Georgia 2007/1 isolate of African swine fever virus. Emerg. Infect. Dis. 2011, 17, 599–605.
16. De Villiers, E.P.; Gallardo, C.; Arias, M.; da Silva, M.; Upton, C.; Martin, R.; Bishop, R.P. Phylogenomic analysis of 11 complete African swine fever virus genome sequences. Virology 2010, 400, 128–136.
17. Alejo, A.; Matamoros, T.; Guerra, M.; Andrés, G. A Proteomic Atlas of the African Swine Fever Virus Particle. J. Virol. 2018, 92.
18. García-Escudero, R.; Andrés, G.; Almazán, F.; Viñuela, E. Inducible gene expression from African swine fever virus recombinants: Analysis of the major capsid protein p72. J. Virol. 1998, 72, 3185–3195.
19. Andrés, G.; Alejo, A.; Salas, J.; Salas, M.L. African swine fever virus polyproteins pp220 and pp62 assemble into the core shell. J. Virol. 2002, 76, 12473–12482.
20. Suárez, C.; Gutiérrez-Berzal, J.; Andrés, G.; Salas, M.L.; Rodríguez, J.M. African swine fever virus protein p17 is essential for the progression of viral membrane precursors toward icosahedral intermediates. J. Virol. 2010, 84, 7484–7499.
21. Andrés, G.; García-Escudero, R.; Viñuela, E.; Salas, M.L.; Rodríguez, J.M. African swine fever virus structural protein pE120R is essential for virus transport from assembly sites to plasma membrane but not for infectivity. J. Virol. 2001, 75, 6758–6768.
22. Frouco, G.; Freitas, F.B.; Coelho, J.; Leitão, A.; Martins, C.; Ferreira, F. DNA-Binding Properties of African Swine Fever Virus pA104R, a Histone-Like Protein Involved in Viral Replication and Transcription. J. Virol. 2017, 91, e02498-16.
23. Iyer, L.M.; Aravind, L.; Koonin, E. V Common Origin of Four Diverse Families of Large Eukaryotic DNA Viruses. J. Virol. 2001, 75, 11720–11734.
24. Baylis, S.A.; Dixon, L.K.; Vydelingum, S.; Smith, G.L. African swine fever virus encodes a gene with extensive homology to type II DNA topoisomerases. J. Mol. Biol. 1992, 228, 1003–1010.
25. Garcia-Beato, R.; Freije, J.M.P.; López-Otín, C.; Blasco, R.; Viñuela, E.; Salas, M.L. A gene homologous to topoisomerase II in African swine fever virus. Virology 1992, 188, 938–947.
26. Coelho, J.; Martins, C.; Ferreira, F.; Leitão, A. African swine fever virus ORF P1192R codes for a functional type II DNA topoisomerase. Virology 2015, 474, 82–93.
27. Hingamp, P.M.; Leyland, M.L.; Webb, J.; Twigger, S.; Mayer, R.J.; Dixon, L.K. Characterization of a ubiquitinated protein which is externally located in African swine fever virions. J. Virol. 1995, 69, 1785–1793.
28. Epifano, C.; Krijnse-Locker, J.; Salas, M.L.; Salas, J.; Rodríguez, J.M. Generation of filamentous instead of icosahedral particles by repression of African swine fever virus structural protein pB438L. J. Virol. 2006, 80, 11456–11466.
29. Simón-Mateo, C.; Freije, J.M.; Andrés, G.; López-Otín, C.; Viñuela, E. Mapping and sequence of the gene encoding protein p17, a major African swine fever virus structural protein. Virology 1995, 206, 1140–1144.
30. Rodríguez, J.M.; García-Escudero, R.; Salas, M.L.; Andrés, G. African swine fever virus structural protein p54 is essential for the recruitment of envelope precursors to assembly sites. J. Virol. 2004, 78, 1313–4299.
31. Rodriguez, F.; Alcaraz, C.; Eiras, A.; Yáñez, R.J.; Rodriguez, J.M.; Alonso, C.; Rodriguez, J.F.; Escribano, J.M. Characterization and molecular basis of heterogeneity of the African swine fever virus envelope protein p54. J. Virol.

1994, 68, 7244–7252.

32. Simón-Mateo, C.; Andrés, G.; Viñuela, E. Polyprotein processing in African swine fever virus: A novel gene expression strategy for a DNA virus. EMBO J. 1993, 12, 2977–2987.
33. Andrés, G.; Simón-Mateo, C.; Viñuela, E. Assembly of African swine fever virus: Role of polyprotein pp220. J. Virol. 1997, 71, 2331–2341.
34. Andrés, G.; García-Escudero, R.; Salas, M.L.; Rodríguez, J.M. Repression of African Swine Fever Virus Polyprotein pp220-Encoding Gene Leads to the Assembly of Icosahedral Core-Less Particles. J. Virol. 2002, 76, 2654–2666.
35. Simón-Mateo, C.; Andrés, G.; Almazán, F.; Viñuela, E. Proteolytic processing in African swine fever virus: Evidence for a new structural polyprotein, pp62. J. Virol. 1997, 71, 5799–5804.
36. Suárez, C.; Salas, M.L.; Rodríguez, J.M. African swine fever virus polyprotein pp62 is essential for viral core development. J. Virol. 2010, 84, 176–187.
37. Andrés, G.; Alejo, A.; Simón-Mateo, C.; Salas, M.L. African swine fever virus protease, a new viral member of the SUMO-1-specific protease family. J. Biol. Chem. 2001, 276, 780–787.
38. Alejo, A.; Andrés, G.; Salas, M.L. African Swine Fever virus proteinase is essential for core maturation and infectivity. J. Virol. 2003, 77, 5571–5577.
39. Muñoz, M.; Freije, J.M.; Salas, M.L.; Viñuela, E.; López-Otín, C. Structure and expression in E. coli of the gene coding for protein p10 of African swine fever virus. Arch. Virol. 1993, 130, 93–107.
