CHAPTER EIGHT

Current capsid assembly models of icosahedral nucleocytoviricota viruses

Yuejiao Xian and Chuan Xiao*

Department of Chemistry and Biochemistry, The University of Texas at El Paso, El Paso, Texas, United States
*Corresponding author: e-mail address: cxiao@utep.edu

Contents

Abstract

Nucleocytoviricota viruses (NCVs) belong to a newly established phylum originally grouped as Nucleocytoplasmic large DNA viruses. NCVs are unique because of their large and complicated genomes that contain cellular genes with homologs from all kingdoms of life, raising intensive debates on their evolutional origins. Many NCVs pack their genomes inside massive icosahedral capsids assembled from thousands of proteins. Studying the assembly mechanism of such capsids has been challenging until breakthroughs from structural studies. Subsequently, several models of the capsid assembly were proposed, which provided some interesting insights on this elaborate process. In this review, we discuss three of the most recent assembly models as well as supporting experimental observations. Furthermore, we propose a new model that combines research developments from multiple sources. Investigation of the assembly process of these vast NCV capsids will facilitate future deciphering of the molecular mechanisms driving the formation of similar supramolecular complexes.

Advances in Virus Research, Volume 108
ISSN 0065-3527
https://doi.org/10.1016/bs.aivir.2020.09.006

1. The nucleocytoplasmic large DNA viruses (NCLDVs)

Originally, viruses were thought to be small infectious agents that could pass through ultrafilters that were used for removing bacteria. They were later defined as infectious nucleoproteinic entities that possess DNA or RNA as their genomic materials and can't grow or undergo binary fission on their own due to the lack of biological synthesis (Lwoff, 1957). In the 1990s, Carl Woese et al. revealed that there are three different types of ribosomes in the living world, which led to the current universal tree of life that includes Archaea, Bacteria, and Eukarya (Woese et al., 1990). Viruses, however, due to the lack of ribosomes, are still missing from this classification despite being the most abundant biological entities on earth (Bergh et al., 1989; Raoult and Forterre, 2008; Reche et al., 2018; Suttle, 2005). Viruses infect cells from all three domains of life while they share several common features, suggesting that viruses may have originated very early in the evolution of life before the division of the three domains of life (Forterre, 2006a,b). Furthermore, viruses play important roles in various evolutionary events including the origins of DNA (Forterre, 2002, 2006a,b; Haig, 2012), the nucleus (Bell, 2001; Takemura, 2001), and the mammalian placenta (Chuong, 2013; Dunlap et al., 2006; Dupressoir et al., 2009). Therefore, understanding the origin and phylogenetic classification of viruses is of great interest in deciphering the evolutionary transitions as well as the formation of different domains of life.

The phylogenetic classification of viruses became more interesting after the study of comparative protein sequence analyses of four families of large DNA viruses (Poxviridae, Asfarviridae, Iridoviridae, and Phycodnaviridae) (Iyer et al., 2001). This study showed that there are 9 genes shared by all the viruses from these four viral families, and 22 genes that are present in at least three of the families. Given the fact that all these viruses harbor a large double-stranded DNA genome and they either completely or partly replicate in the cytoplasm of eukaryotic cells, these viruses are referred as nucleocytoplasmic large DNA viruses (NCLDVs) (Iyer et al., 2001). Since then, more members of NCLDVs were identified. The phylogenetic reconstruction of NCLDVs, including the newly discovered ones, showed that these viruses share a set of approximately 40 genes that are essential for virion morphogenesis and replication, suggesting that they may have evolved from a common ancestor (Iyer et al., 2006).

One major characteristic of NCLDVs is the large size of their virions. The existence of large virus particles has been appreciated since the poxvirus

in 1960s (Dales and Siminovitch, 1961), jumbo bacteriophages in 1971 (Donelli et al., 1972, 1975), and the phycodnaviruses in the early 1980s (Van Etten et al., 1982). However, it was the discovery of the Mimivirus (La Scola et al., 2003) with its extremely high genome complexity, gigantic physical size, and unique structural characteristics that drew massive attention to giant viruses from scientists (Claverie et al., 2006). Mimivirus was serendipitously discovered during an investigation of a hospital outbreak of pneumonia and was originally assumed to be a bacterium due to its large size and gram-stainable fibers on the surface. The virus was named Mimivirus for "mimicking microbe" (La Scola et al., 2003). Further characterization of this virus showed that it contains a 1.2-Mbp genome packed in a protein shell with a diameter of 500 nm (Raoult et al., 2004; Xiao et al., 2005). Furthermore, this protein shell is decorated with a dense layer of 125 nm-long fibers, making the total diameter of the particle about 750 nm. The Mimivirus was officially classified by the International Committee on Taxonomy of Viruses (ICTV) (Lefkowitz et al., 2018) as the founder of a new family, Mimiviridae (La Scola et al., 2003; Raoult et al., 2004), belonging to the clade of NCLDVs.

Subsequently, dozens of Mimivirus relatives and other NCLDVs were discovered, leading to the expansion of Mimiviridae family (Yutin et al., 2013) as well as the establishment of new viral families (Arslan et al., 2011; Yutin and Koonin, 2012; Yutin et al., 2009), including Marseilleviridae (Boyer et al., 2009), Ascoviridae (Asgari et al., 2017), Pandoraviridae (Philippe et al., 2013), Pithoviridae (Legendre et al., 2014), and Mininucleoviridae (Subramaniam et al., 2020). These viruses carry by far the most complicated known viral genomes ranging from 100 kb to 2.5 mb. Most interestingly, cellular genes that were never seen in viruses before were found in NCLDVs (Yutin and Koonin, 2012). These genes carry unprecedented functions including components of the translation machinery, DNA repair, and metabolic pathways. In addition, many of these cellular genes have homologs in all three kingdoms of life (Boyer et al., 2009; Fischer et al., 2010; Moniruzzaman et al., 2014; Philippe et al., 2013; Raoult et al., 2004). The complexity of their genomes has led to intensive and controversial debates on whether they should be considered as the fourth domain of life. It has also been proposed that NCLDVs descended from a common cellular ancestor through reductive evolution (Claverie and Abergel, 2009; Claverie et al., 2006, 2009; Raoult et al., 2004). This hypothesis has been challenged by other researchers who argued that NCLDVs were evolved from a small DNA viral ancestor by horizontal gene transfer and gene duplications (Filee, 2009; Filee and Chandler, 2010; Koonin et al., 2015; Yutin and Koonin, 2012;

Yutin et al., 2009, 2014). The increasing number of available NCLDV genomes led to more comparative evolutionary genomics analyses. Using several most strongly conserved representatives from the nine core genes that are shared by all the NCLDVs, the phylogenetic analyses done by Yutin et al. suggested that NCLDVs evolved from multiple lineages of smaller viruses on a minimum of three different occasions, rather than being a fourth domain of life (Yutin et al., 2014). These findings led to the proposal of multiple convergent origins of NCLDVs (Koonin and Yutin, 2018, 2019).

2. The double jelly roll and the nucleocytoviricota viruses (NCVs)

Viruses are often grouped together based on their genetic sequence similarities, which has provided us with an enormous amount of insight in the classification of the virosphere. However, this approach becomes difficult in the absence of observable sequence similarities, especially in the attempt to identify long-range evolutionary relationships. Therefore, structure-based phylogenetic analysis (SBPA) has become a valuable method to understand the evolutionary relationship of genetically distinct viruses using distances calculated from pairwise comparison of protein structures (Ng et al., 2020). One example of SBPA is the proposed "PRD1-adenovirus" lineage based on the structural similarities of the double jelly roll (DJR) fold of capsid proteins from genetically distinct viruses infecting hosts from all three kingdoms of life (Bahar et al., 2011; Bamford et al., 2005; Benson et al., 2004; Khayat et al., 2005; Krupovic and Bamford, 2008; Yutin et al., 2018). PRD1 is a bacteriophage that was first isolated and studied in 1970s (Olsen et al., 1974). The DJR fold is the main structural element of the major capsid proteins (MCPs) of many double-stranded DNA (dsDNA) viruses such as bacteriophage PRD1 (Benson et al., 1999) and adenoviruses (Roberts et al., 1986; Stewart et al., 1991) as well as many NCLDVs (Bahar et al., 2011; Bamford et al., 2002; Fang et al., 2019; Liu et al., 2019a,b; Nandhagopal et al., 2002; Wang et al., 2019; Yan et al., 2009; Zhang et al., 2011). The structural characteristics of DJR in NCLDVs will be further discussed in the following section. These MCP DJR folds share very low sequence homology but strikingly high structural similarities among many dsDNA viruses. In addition to this, recently, FLiP (Flavobacterium-infecting, lipid-containing phage), a single stranded DNA virus, was also found to process the DJR capsid protein, expanding the PRD1-adenovirus lineage to ssDNA viruses (Laanto et al., 2017). These results collectively

suggest that viruses within this DJR PRD1-adenovirus lineage could be very ancient and share a common ancestor (Benson et al., 2004; Krupovic and Bamford, 2008; Ng et al., 2020; San Martin and van Raaij, 2018), although convergent evolution cannot be completely ruled out.

In agreement with the PRD1-adenovirus lineage that is largely based on SBPA, new comparative genomics and metagenomics analyses showed the presence of evolutionary conservation in these jelly roll genes, leading to the recent ICTV-approved megataxonomic framework (Koonin et al., 2019; Walker et al., 2019). Within this framework, a new realm Varidnaviria was created to include DNA viruses that encode an MCP containing the vertical (single or double) jelly roll fold and forming pseudo-hexameric capsomers. Within the Varidnaviria realm is the kingdom Bamfordvirae for those viruses using the DJR to form their MCPs. The kingdom of Bamfordvirae was named after Dr. Dennis H. Bamford for his contributions in the PRD1-adenovirus lineage (Bamford, 2003; Bamford et al., 2005; Benson et al., 2004; Krupovic and Bamford, 2008). Furthermore, NCLDVs are officially classified as a new phylum of Nucleocytoviricota under the kingdom of Bamfordvirae. Currently, Nucleocytoviricota contains two different classes, Megaviricetes and Pokkesviricetes. The name of Megaviricetes is modified from a previously proposed name of Megavirales (Colson et al., 2012) with the "-viricetes" suffix for class taxa, and Pokkesviricetes is named after Middle English "pokkes," meaning pox. Megaviricetes includes three orders: (1) Pimascovirales, that includes the families of Ascoviridae, Iridoviridae, and Marseilleviridae; (2) Imitervirales, that includes the family of Mimiviridae; and (3) Algavirales, that includes the family of Phycodnaviridae. Within the Pokkesviricetes class are the two orders: (1) Asfuvirales that includes the family of Asfarviridae and (2) Chitovirales that includes the family of Poxviridae. Among the 10 families from the original NCLDV clades, Pithoviridae, Mininucleoviridae and Pandoraviridae are yet to be classified. The stories of NCLDVs being discovered and studied in the last two decades are a perfect demonstration that these highly complicated viruses are interesting and fascinating targets for evolutionary and structural studies. This review will use the newly ICTV-approved name Nucleocytoviricota viruses (NCVs) as a replacement of NCLDVs in the following sections.

3. The structure of icosahedral NCVs

Based on their virion morphology, NCVs from the 10 families can be roughly classified as icosahedral NCVs, including members from

Asfarviridae, Iridoviridae, Marseilleviridae, Mimiviridae, Phycodnaviridae, and Mininucleoviridae; and non-icosahedral NCVs from the families of Poxviridae, Pandoraviridae, Ascoviridae, and Pithoviridae. In this paper, we will review the structures as well as the assembly models of icosahedral NCVs' capsids. As some of NCVs are comparable to small bacteria and archaea in both particle and genome size (Benamar et al., 2016; Boyer et al., 2009; Legendre et al., 2014; Philippe et al., 2013; Raoult et al., 2004), their structural studies have been a great challenge and were mainly studied through combining multiple structural determination techniques (Xiao and Rossmann, 2011).

Electron microscopy (EM) has been commonly utilized to obtain the first glimpse of viral morphology, which provides some basic structural information including the shape and the surface features of viruses. For example, it was EM that revealed the icosahedral capsid of the Mimivirus and differentiated it from bacteria (La Scola et al., 2003). Since the 1990s, cryogenic electron microscopy (cryo-EM) with its three-dimensional reconstruction techniques has become an increasingly powerful tool for studying the structures of NCVs. Besides Mimivirus (Xiao et al., 2005, 2009), other NCVs that have been studied by cryo-EM include but are not limited to: *Paramecium bursaria* Chlorella virus 1 (PBCV-1) (Fang et al., 2019; Nandhagopal et al., 2002; Yan et al., 2000; Zhang et al., 2011), Chilo iridescent virus (CIV) (Yan et al., 2000, 2009), *Phaeocystis pouchetii* virus (PpV01) (Yan et al., 2005), Cafeteria roenbergensis virus (CroV) (Xiao et al., 2017), Faustovirus (FAUV) (Klose et al., 2016), Melbournevirus (MelV) (Okamoto et al., 2018), Pacmanvirus (Andreani et al., 2017), and African swine fever virus (ASFV) (Andres et al., 2020; Liu et al., 2019a; Wang et al., 2019), whose diameters and other characteristics are summarized in Table 1.

Early low resolution cryo-EM structural studies on these NCVs revealed that they share some common features in their morphology. The most commonly seen architecture of icosahedral NCVs consists of three major components: an icosahedral protein capsid, a lipid membrane, and a genomic core (Fig. 1A). While most of these NCVs have an inner lipid membrane located directly underneath their protein capsids, an interesting exception is FAUV that replaces the lipid membrane with an inner protein layer (Fig. 1C). Another exception is ASFV that possesses a multilayer structure with an additional outer lipid envelope as well as an icosahedral protein shell between the inner membrane and the genomic core (Fig. 1B). The surfaces of the NCVs have various features. For example, the surfaces of

Table 1 The Mathematical characteristics of large icosahedral Nucleocytoviricota virus capsids.

Virus	Diameter (nm)	Trisymmetron size[a]	Pentasymmetron size[b]	T number	h	k	Hexameric capsomers[c]	Jelly rolls[d]	MCP copies
CIV (Yan et al., 2000, 2009)	185	55	30 + 1	147	7	7	1460	8820	4380
PBCV-1 (Fang et al., 2019; Zhang et al., 2011)	190	66	30 + 1	169	7	8	1680	10,140	5040
PpV-01 (Yan et al., 2005)	220	91	30 + 1	219	7	10	2180	13,140	6540
FAUV (Klose et al., 2016)	260	120	30 + 1	277	7	12	2760	16,620	8280
ASFV (Liu et al., 2019a; Wang et al., 2019)	260	120	30 + 1	277	7	12	2760	16,620	8280
MelV (Okamoto et al., 2018)	232	136	30 + 1	309	7	13	3080	18,540	9240
Pacmanvirus (Andreani et al., 2017)	250	136	30 + 1	309	7	13	3080	18,540	9240
CroV (Xiao et al., 2017)	300	231	30 + 1	499	7	18	4980	29,940	14,940

[a]The sizes of trisymmetron (triangular patch of capsomers with the same orientation) are presented by the number of pseudo-hexameric capsomers in each symmetron.
[b]The sizes of pentasymmetron (the pentagonal patch of capsomers around the icosahedral fivefold vertices of the capsid) are presented by the number of pseudo-hexameric capsomers plus one pentameric capsomer (penton) in each symmetron.
[c]The number of pentons for all icosahedral NCVs are 12, hence, only the number of pseudo-hexameric capsomers are listed here.
[d]The total number of jelly rolls including the single jelly rolls from pentons and the double jelly rolls from pseudo-hexameric capsomers.

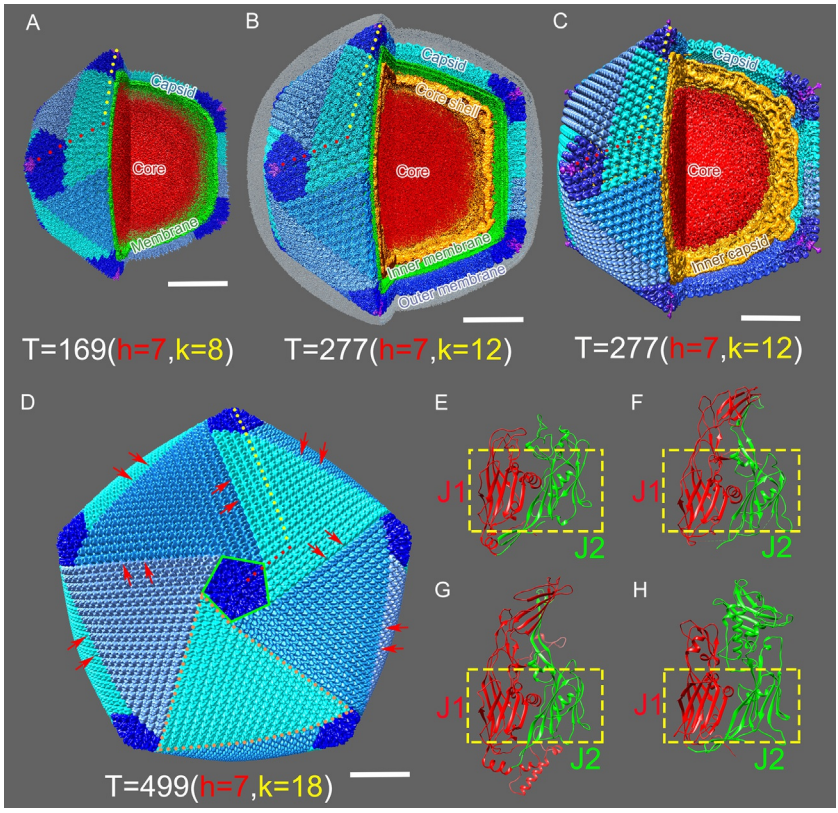

Fig. 1 The cryo-EM structures of icosahedral Nucleocytoviricota viruses (NCVs) and the high-resolution structures of double jelly roll (DJR) major capsid proteins (MCPs). (A–D) are the cryo-EM reconstructions of *Paramecium bursaria* Chlorella virus 1(PBCV-1) (Electron Microscopy Data Bank, EMDB, accession code EMD-0436, and, EMD-5378), African swine fever virus (ASFV) (EMDB accession code EMD-0815), Faustovirus (FAUV) (EMDB accession code EMD-8144), and Cafeteria roenbergensis virus (CroV) (EMDB access code EMD-8748), respectively. The capsomers located along the h and k axes for Triangulation number (T number) calculations are marked by red and yellow dots, respectively. The T number and its (h, k) values are labeled beneath each virus. A quarter of the virus in panels (A–C) is removed to show their layered architecture: core (red), capsid (various shade of blues), the inner membrane of PBCV-1 and ASFV (green), the outer membrane of ASFV (gray), the core shell of the ASFV and the inner capsid of FAUV (orange). The viral capsids are colored in four different shapes of blue to shown the capsomer arrangement into pentasymmetrons (deep blue) and trisymmetrons (cornflower blue, dodger blue, and cyan). The outer membrane of ASFV is shown with 50% transparency to allow better visualization of the capsomer arrangement in the capsid. One representative pentasymmetron and trisymmetron in panel (D) are highlighted with a green pentamer (solid line) and orange triangle (dashed line), respectively; the discontinuous lines on the CroV surface at the boundaries of symmetrons are indicated with red arrows. (E–H) Ribbon diagrams of the MCPs of PBCV-1, ASFV, and FAUV as well as the scaffold protein D13 of Vaccinia virus (VACV). The conserved DJR folds are highlighted by yellow dashed line boxes. These structures are acquired from the Protein Data Bank (PDB) with IDs 1M4X, 6KU9, 5J7U, and 2YGB, respectively.

Mimivirus and its close relatives such as Samba virus (SMBV) (Campos et al., 2014) and Tupanvirus (Abrahao et al., 2018) are covered by a dense layer of long fibers. CIV has some shorter fibers decorating most of its capsomers (Yan et al., 2000, 2009) while PBCV-1 and PpV01 have fibers attached on a few of their capsomers (Yan et al., 2005). These fibers were hypothesized to play roles in host recognition and attachment (Claverie et al., 2006; Klose et al., 2010; Xiao et al., 2009; Zauberman et al., 2008). Beside the fibers, Tupanvirus also has a unique large cylindrical structure attached to the capsid (Abrahao et al., 2018). In addition, a unique portal on one of the icosahedral fivefold vertices has been observed in several NCVs, such as the stargate of Mimivirus (Klose et al., 2010; Kuznetsov et al., 2010; Xiao et al., 2009; Zauberman et al., 2008) and the spike portal of PBCV-1 (Cherrier et al., 2009; Zhang et al., 2011).

Despite the variations on their surfaces and portals, early cryo-EM studies demonstrated that these capsids of NCVs are assembled in a highly conserved manner following the icosahedral symmetry. Like most small viruses using repeating protein units for building their viral capsids (Rossmann and Johnson, 1989), NCVs use building blocks called capsomers for assembling their large protein shells (Lwoff et al., 1959). These capsomers are arranged into triangular and pentagonal arrays (Fig. 1D) that were defined as trisymmetrons and pentasymmetrons by Wrigley et al. in 1969 based on his EM study on broken Sericesthis iridescent virus (SIV) (Wrigley, 1969). In this study, the decomposition of SIV resulted in individual triangular and pentagonal patches. The fracture lines between neighboring symmetrons were clearly visualized at the early stage of viral decay. These symmetron patches were also seen in broken FAUV samples (Klose et al., 2016). Furthermore, discontinuous lines were observed on the viral surfaces in low resolution cryo-EM reconstructions of PBCV-1 (Yan et al., 2000), CIV (Yan et al., 2000, 2009), PpV01 (Yan et al., 2005), FAUV (Klose et al., 2016), CroV (Fig. 1D) (Xiao et al., 2017), and MelV (Okamoto et al., 2018). These discontinuous lines divide the viral capsids into triangular and pentagonal areas that are equivalent to the symmetrons.

In addition to cryo-EM, X-ray crystallography was also utilized to determine the structures of individual capsid proteins, such as the MCP of PBCV-1 (Nandhagopal et al., 2002). The atomic structure of PBCV-1 MCP has the aforementioned DJR fold. A single "jelly roll" (SJR) fold is a wedge-shaped β-barrel structure built form ∼eight β-sheets, which

were first found in small single stranded RNA viruses (Rossmann and Johnson, 1989). The SJR fold has become the most common fold in viruses with various sizes. A recent comparative study showed that the SJR fold occupies approximately 28% of the capsid protein in the entire classified virosphere (Krupovic and Koonin, 2017). A DJR fold contains two tandemly connected SJRs and is believed to have evolved from the SJR via gene duplication and combination (Gil-Carton et al., 2015; Krupovic and Koonin, 2017). Unlike in most SJR capsid proteins, the β-barrels in DJR MCPs are orientated vertically (perpendicularly to the viral surfaces) (Benson et al., 2004). In addition, these DJR MCPs trimerize to form the capsomers; therefore, a single capsomer consists of six SJRs, contributing to its pseudo-sixfold symmetry. However, in the same DJR, the two jelly rolls are positioned slightly differently, where one jelly roll (J2 in Fig. 1E–H) is placed higher (toward the exterior of the virus) than the other one (J1 in Fig. 1E–H). The lengths of the surface loops on top of the two jelly rolls also vary (Fig. 1E–H). These two structural features contribute to the true trimeric symmetry of the capsomers. Structural studies demonstrated that the DJR MCP is highly conserved among icosahedral NCVs (Fig. 1E–H) (Andres et al., 2020; Klose et al., 2016; Liu et al., 2019a,b; Wang et al., 2019; Xiao et al., 2017; Yan et al., 2009) and the scaffold protein D13 of the immature poxvirus particles (Fig. 1H) (Bahar et al., 2011). It is noteworthy that while the mature virions of poxviruses are in brick shape, their immature virions are spherical with hexagonal arrays of D13 on their surface.

Cryo-EM and X-ray crystallography have been combined to study the structures of NCVs. By fitting the atomic structures of the MCPs determined by X-ray crystallography into low resolution cryo-EM density maps, the pseudo-atomic structures of the whole icosahedral NCV capsids were obtained (Klose et al., 2016; Nandhagopal et al., 2002). The pseudo-atomic structures allow the studies of the capsomer arrangements within the viral capsids (Simpson et al., 2003). Due to the trimeric nature of the capsomers, they can be packed into two different orientations where one capsomer is (1) sharing the same orientation with the neighboring capsomers; and (2) 60° rotated to neighboring capsomers. The pseudo-atomic structures of NCVs revealed that while the capsomers within the same trisymmetron are packed in the same orientation, they are 60° rotated when compared to those in the neighboring trisymmetrons. This orientation difference results in the discontinuous lines between different symmetrons mentioned

above (Klose et al., 2016; Nandhagopal et al., 2002; Simpson et al., 2003; Xiao et al., 2017; Yan et al., 2005, 2009). The discontinuous lines were also observed in between the pentasymmetrons and trisymetrons. It is worthy to mention that underneath the discontinuous lines, extra protein densities were seen in cryo-EM reconstructions with improved resolution (Klose et al., 2016; Yan et al., 2009; Zhang et al., 2011). These densities were suggested to belong to the "cement" or "zipper" proteins that function in stabilizing the neighboring symmetrons. In addition to the zipper proteins, densities of other minor capsid proteins (mCPs) such as finger proteins or anchor proteins were also found underneath the MCPs. These mCPs were suggested to facilitate the assembly of capsomers by functioning as inter-capsomer crosslinks or capsid membrane anchors (Yan et al., 2009). However, due to their small molecular weights and weak densities in cryo-EM reconstructions, the structural nature of these mCPs remained unknown until recent resolution breakthroughs in cryo-EM reconstruction of NCVs.

Due to the revolutionary improvements in both cryo-EM hardware and software, the resolution of small particle reconstructions has significantly improved to the atomic level (Henderson, 2018). However, due to the much larger sizes of NCVs, it remained challenging to determine their atomic resolution structures, until the recent development of the "block-based" reconstruction technique (Xian and Xiao, 2020; Zhu et al., 2018). This method resulted in a 3.5 Å reconstruction of PBCV-1, the first near-atomic structure of an icosahedral NCV (Fang et al., 2019). A similar method was used in ASFV studies, leading to its near-atomic structure with 4.5 Å resolution (Liu et al., 2019a; Wang et al., 2019). With these breakthroughs, 14 and four different types of mCPs were identified in PBCV-1 and ASFV, respectively. One of the mCPs is the SJR protein that pentamerizes to form the pentameric capsomer (penton) located at the fivefold vertices. The other mCPs form an intensive network beneath the MCP layer, interacting with both MCPs and the inner membranes (Fig. 2). Therefore, these mCPs were suggested to play significant roles in stabilizing the capsomer assembly and in anchoring the capsomers onto the membrane. The characterization of these mCPs deepened our understanding of the architecture of icosahedral NCV capsids, leading to the proposals of two models on the viral capsid assembly, which will be further discussed in the later sections.

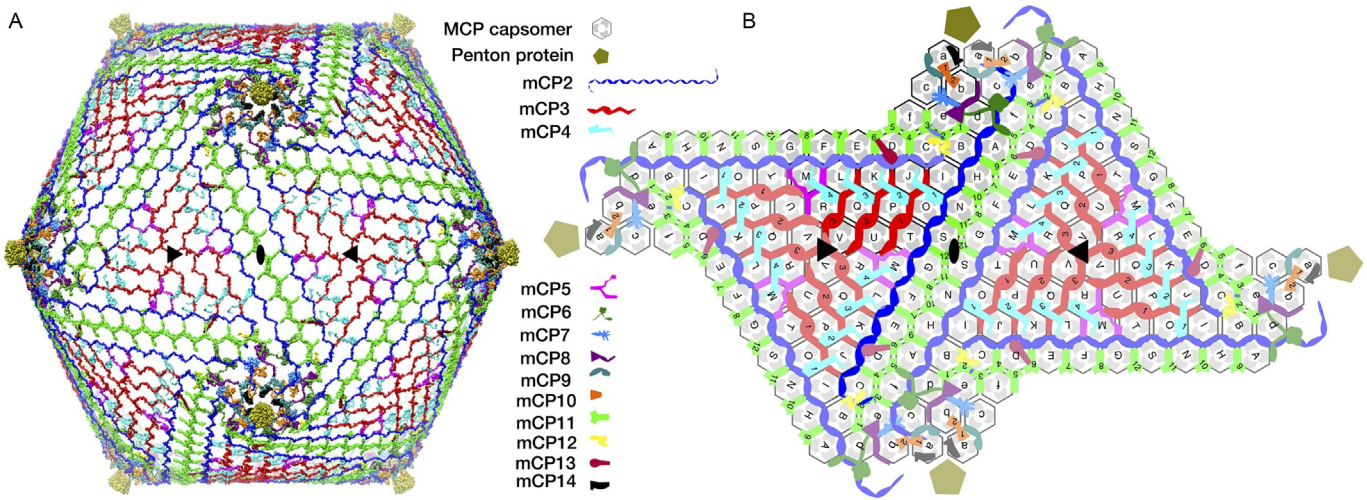

Fig. 2 The minor capsid proteins (mCPs) of PBCV-1. (A) The hexagonal network beneath the MCPs formed by the 14 mCPs; (B) A schematic diagram of the capsomers and the mCPs viewed from inside of PBCV-1. Capsid proteins in both panels are colored as indicated by the legend in the middle. *Images from both panels are adapted from Fang, Q., Zhu, D., Agarkova, I., Adhikari, J., Klose, T., Liu, Y., Chen, Z., Sun, Y., Gross, M.L., Van Etten, J.L., Zhang, X., Rossmann, M.G., 2019. Near-atomic structure of a giant virus. Nat. Commun. 10 (1), 388 with the permission.*

4. The triangulation numbers of the icosahedral NCV capsids

Like many small spherical viruses, the capsids of icosahedral NCVs are also assembled from identical protein units in a particular manner following symmetries. It is noteworthy that this assembly mechanism was suggested as early as 1956 by Crick and Watson (Crick and Watson, 1956). Before any atomic structures of viruses were determined, Casper and Klug proposed a quasi-equivalent theory for constructing an icosahedral virus (Caspar and Klug, 1962). Their theory involved folding up a plane of equilateral-triangular lattice, where six triangular units were joined to form a hexamer at each lattice point. By removing one triangular unit at the particular lattice points to introduce the pentamers, the original plane can be folded into an icosahedron with various sizes (Fig. 3A). Consequently, this folding was formulated by Caspar and Klug into the triangulation number (T number) with equation:

$$T = \mathrm{h}^2 + \mathrm{hk} + \mathrm{k}^2 \tag{1}$$

when one particular pentamer is chosen as the origin (0,0) and the two axes along the lattice points are defined as H and K with angle of 60°, the (h,k) are then the coordinates of a neighboring pentamer on the triangulated plane (Fig. 3A and B). In this coordination system, each lattice point can represent either a pentamer or a hexamer. If we define each subunit (e.g., the jelly roll fold) that forms the pentamer or hexamer at each lattice point as a structural unit, of a given icosahedral virus capsid, the T number is thus the number of the structural units in one asymmetric unit (ASU) (Fig. 3A). Therefore, the number of structural units in the whole viral capsid can be calculated through the following equation with the 60-fold icosahedral symmetry:

$$N_{strutural\ units} = 60T \tag{2}$$

Their theory presented possibly the most efficient approach to assemble a spherical capsid with icosahedral symmetry where the assembly of the structural units are quasi-equivalent and require only minimal distortion in the subunit bondings. This theory was later confirmed in many small viruses whose high-resolution structures became avaliable (Harrison et al., 1978; Hogle et al., 1985; Rossmann et al., 1985). Based on the geometry of icosahedron, the number of pentamers is always 12, as they are located at the

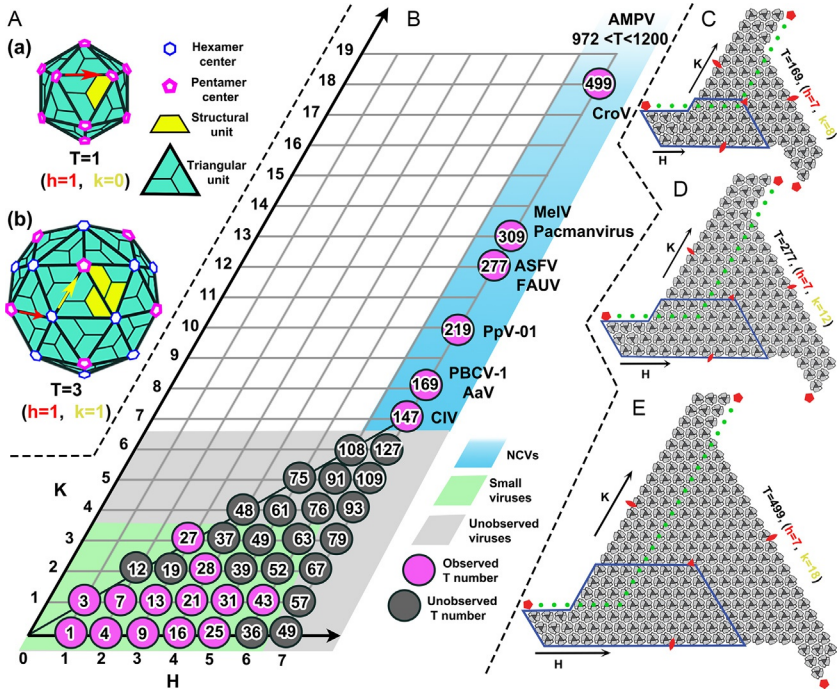

Fig. 3 The Triangulation number (T number) and examples in viruses with different sizes. (A) Two examples of $T=1$ and $T=3$ icosahedral capsids are shown in subpanels (a) and (b); In panel (a and b), the triangular unit (cyan) is a trimeric form of the structural unit (yellow); The centers of the pentamers and hexamers are indicated with pink pentagons and blue hexagons, respectively. The paths for calculating their T numbers from the original pentamer center to its neighboring pentamer center are indicated with the red and yellow arrows. (B) A matrix of T numbers with different h and k values. The corresponding T numbers of each (h, k) are shown on top of colored circles (pink, observed T number; gray, not yet observed T number); the T number area associated with NCVs, small viruses, and gap area are shaded with light blue, light green, and light gray, respectively; (C–E) the schematic diagrams of one icosahedral face of PBCV-1, ASFV, and CroV, respectively. Icosahedral five-, three-, and twofold symbols are indicated in red and one asymmetric unit (ASU) is outlined in blue. Directions for axes K and H are indicated with black arrows. The capsomers for counting h and k are labeled with green dots. *Panel (C and E) are adapted from Xiao, C., Fischer, M.G., Bolotaulo, D.M., Ulloa-Rondeau, N., Avila, G.A., Suttle, C.A., 2017. Cryo-EM reconstruction of the Cafeteria roenbergensis virus capsid suggests novel assembly pathway for giant viruses. Sci. Rep. 7 (1), 5484 with permission.*

icosahedral vertices, and the number of hexamers can be calculated based on the given T number using following equation:

$$N_{hexamer} = 10(T - 1) \quad (3)$$

As mentioned above, the capsids of icosahedral NCVs are also assembled from identical protein units, the trimeric DJR MCPs. If we consider an SJR as the structural unit, the T number, as well as the number of hexameric capsomers, jelly rolls, and MCPs can be calculated using the methods and equations developed by Caspar and Klug (Table 1). In NCVs, the pentameric capsomer (penton) or pseudo-hexameric capsomer occupies one lattice point. Therefore, the T number of NCVs can be calculated from h and k that are the number of capsomers and pentons along the shortest path on the axes of H and K between two neighboring fivefold pentons (Figs. 1A–D and 3C–E). Other numbers can then be calculated from the T number. For example, PBCV-1 has a triangulation number of 169, therefore, its capsid contains 10,140 = (169 × 60) jelly rolls, 1680 = (169–1) × 10 pseudo-hexameric capsomers, and 5070 = 1680 × 3 MCPs.

The assembly of icosahedral NCV capsids is unique in that the capsomers are arranged into pentasymmetrons and trisymmetrons. N. G. Wrigley, who defined the symmetrons in 1969, presented a modified Goldberg's diagram that described all the possible sizes of symmetrons.

Besides penta- and trisymmetrons, N. G. Wrigley also proposed the presence of disymmetrons. However, such a disymmetron has never been observed in any icosahedral NCVs. Later, the possible combinations of trisymmetron and pentasymmetron sizes were thoroughly studied by Sinkovits and Baker (2010). As indicated by both studies, various values of h are allowed to form an icosahedron using symmetrons. However, it is interesting that the h values of all the structurally characterized icosahedral NCVs are always seven, while the k values vary (Figs. 1A–D and 3B–E; Table 1). This is because their pentasymmetrons are always the same size, including one penton located at the fivefold axis and 30 pseudo-hexameric capsomers arranged in three layers (Xiao et al., 2017). Therefore, when counting capsomers along the h axis, from a penton toward a neighboring penton, the h number always includes three counts of hexamers in the pentasymmetron and four counts in the trisymmetron. The count in the trisymmetron remains four because it is parallel to the neighboring pentasymmetron (which includes three counts of the hexamers and one count of the penton) (Figs. 1A–D and 3C–E). It is intriguing that the size

of the pentasymmetrons among all known icosahedral NCVs remains the same (Table 1). One can speculate that a pentasymmetron with 30 hexameric capsomers and a penton is the optimal size, however, the regulation factors on the size of the pentasymmetron remain unknown.

While the structures of many viruses with various sizes were solved and shown to follow the proposed quasi-equivalent theory, there is a large gap between $T=43$ and $T=147$ (Fig. 3B T numbers on gray circles). Is such a gap because our current data is limited and the viruses with an intermediate T numbers are yet to be discovered? Or perhaps, the missing viruses implies that the assembly of such viruses requires a more difficult mechanism? Are there any underlining evolutional factors that lead to the two groups of small and giant viruses? Answering these questions requires more viruses with various sizes to be identified and structurally characterized to comprehend their assembly mechanisms.

5. Viral factory and NCV assembly

The assembly of NCV capsids occurs in viral factories, which are membrane-separated compartments formed in the infected cellular host during the viral infections (Novoa et al., 2005). Within the viral factories, massive reorganization of the host cell cytoskeletons and membranes occurs for viral genome replication and virion assembly. Such viral factories, commonly seen during the infection of many viruses (Ahlquist, 2006; Fischer, 2011; Fontana et al., 2010; Milrot et al., 2016; Salas and Andres, 2013; Sodeik et al., 1993; Suarez et al., 2013), allow the viral components to escape from the host defense mechanisms, and to be concentrated and act together to significantly increase the efficiency of viral replication (Novoa et al., 2005; Zhang et al., 2015). Imaging technologies such as the three-dimensional (3D) electron tomography have significantly deepened our understanding of viral factories. For instance, in 2016, Milrot et al. studied the PBCV-1 viral factories (Fig. 4A–B). Based on their 3D structural study, the authors proposed a well-supported model of viral factory generation and virion assembly (Milrot et al., 2016). In their model, the viral factories start as viral products accumulating in the cytoplasm and excluding the organelles from the host cell. Among these viral products are the membrane cisternae that are ruptured from rough endoplasmic reticulum membranes and serve as scaffolds for viral capsid protein assembly. The association of viral capsid proteins with membrane cisternae leads to the formation of a procapsid initiating at the icosahedral fivefold vertex. The assembly of the procapsid continuously

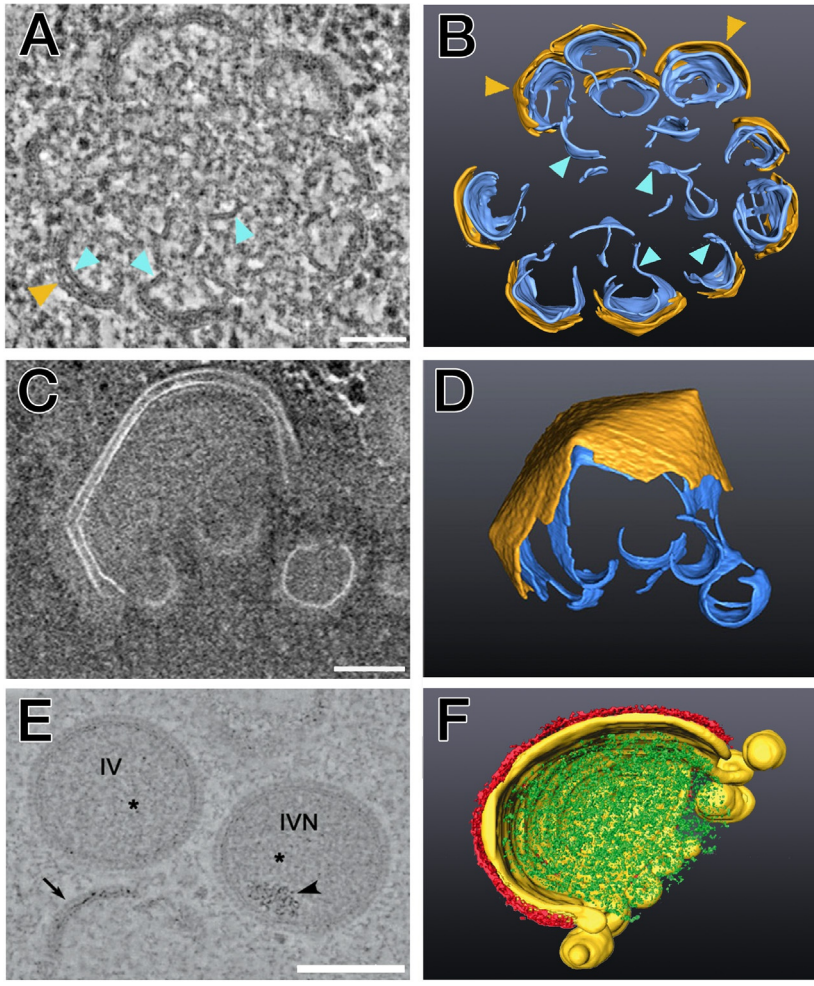

Fig. 4 The EM studies of NCV capsid assembly. (A), (C), and (E) are slices from EM tomograms of the viral factories of PBCV-1, Mimivirus, and VACV, respectively; Scale bar,100 nm; (B) and (D) are the 3D reconstructions of the viral factories showed in panel (A and C), respectively. The capsid and open membranes are colored in yellow and blue, respectively. In panel (A and B), yellow and blue triangles are pointing to the capsid and membrane, respectively. In panel (E), viral particles were immunolabeled with anti-A14, a membrane protein of VACV. The long and short arrows point to the VACV assembly precursor and genomic materials, respectively. IV and IVN with the asterisks stand for immature virus and immature virus with nucleoid. (F) The 3D reconstruction of a VACV crescent. The crescent membrane, scaffold protein, and viral core proteins are

(Continued)

progresses from the initial vertex to its neighboring ones on the convex side of the growing membrane. The procapsid remains incomplete until the genome is packaged through a large portal. Similar viral assembly processes were also seen in Mimivirus (Fig. 4C and D) (Mutsafi et al., 2013, 2014; Suarez et al., 2013), Vaccinia virus (VACV) (Fig. 4E and F) (Chlanda et al., 2009; Liu et al., 2014; Risco et al., 2002), and ASFV (Andres et al., 1998; Suarez et al., 2015), suggesting a general mechanism of viral membrane formation and capsid assembly of NCVs.

Although the assemblies of NCVs share common features such as the viral factories and membrane reorganization, there are some significant variations. One unique example is ASFV, that has an outer lipid membrane and a protein core shell (Salas and Andres, 2013). The core shell formation and DNA packaging progress simultaneously underneath the concave face of the membrane as the capsomers progressively assemble on the convex face of the viral membrane forming a polyhedral capsid (Andres et al., 1997). The process of assembling the inner membrane and the capsid is largely independent of the assembly of the core shell and the nucleoid, as the "empty" icosahedral particles can form in the absence of core shell polyproteins (Andres et al., 2002). After the completion of capsid assembly, ASFV acquires its outer membrane while budding through the plasma membrane (Salas and Andres, 2013). Another unique example is the assembly of VACV, in which the ruptured endoplasmic reticulum membranes associates with the preassembled scaffold protein D13 forming the so-called crescent precursors (Chlanda et al., 2009; Liu et al., 2014; Risco et al., 2002; Sodeik et al., 1993). These crescents then continuously grow into spherical immature virus particles while the genomic DNA and core proteins are packed inside (Fig. 3E and F). The immature virus particles then transform into mature virions through multiple events including the cleavage of D13 protein (Chichon et al., 2009; Liu et al., 2014).

Fig. 4—Cont'd colored in yellow, red, and green, respectively. *With permission, panel (A and B) are adapted from Milrot, E., Mutsafi, Y., Fridmann-Sirkis, Y., Shimoni, E., Rechav, K., Gurnon, J.R., Van Etten, J.L., Minsky, A., 2016. Virus-host interactions: insights from the replication cycle of the large Paramecium bursaria chlorella virus. Cell Microbiol. 18 (1), 3–16; panel (C and D) are adapted from Mutsafi, Y., Shimoni, E., Shimon, A., Minsky, A., 2013. Membrane assembly during the infection cycle of the giant Mimivirus. PLoS Pathog. 9 (5), e1003367; panel (E) is adapted from Chlanda, P., Carbajal, M.A., Cyrklaff, M., Griffiths, G. Krijnse-Locker, J., 2009. Membrane rupture generates single open membrane sheets during vaccinia virus assembly. Cell Host Microbe 6 (1), 81–90; and panel (F) is adapted from Krijnse Locker, J., Chlanda, P., Sachsenheimer, T., Brugger, B., 2013. Poxvirus membrane biogenesis: rupture not disruption. Cell Microbiol. 15 (2), 190–199.*

While EM studies on viral factories of NCVs have provided large amount of details in their morphogenesis, the molecular mechanisms that modulate the assembly of the icosahedral NCV capsids are yet to be demystified. It's clear from these EM studies that the assembly process of capsids is closely associated with the remodeling of viral membranes, and that the assembly initiates at one of the icosahedral fivefold vertices and proceeds continuously (Milrot et al., 2016; Mutsafi et al., 2014; Suarez et al., 2013). However, many questions remain to be answered. Thorough investigations on the architecture of the viral capsids are the first essential step that will provide a wealth of information about their assembly process and guide further assembly studies.

6. The current models in the capsid assembly of NCVs

6.1 The symmetron assembly model

The observations of the individual symmetrons in SIV as well as other NCVs have led to the assumption that the viral capsids were assembled from prepacked symmetrons (Wrigley, 1969, 1970). This model became even more convincing as low resolution cryo-EM reconstructions of NCVs show discontinuous lines at the boundaries of symmetrons (Klose et al., 2016; Xiao et al., 2017; Yan et al., 2000; Zhang et al., 2011). It was then assumed that capsomers were preassembled into the symmetrons that will then assemble into the icosahedron with the help of the mCPs such as the zipper proteins (Yan et al., 2009). The near-atomic structure of PBCV-1 determined by Fang et al. showed some unique structural properties of mCPs. Five different mCPs were found to form a hexagonal network underneath the trisymmetrons, among which mCP2, mCP3 and mCP11 contribute the most. The mCP2 is the so-called tape measure protein (TmP), which is located near the boundaries of the trisymmetrons. The TmP is an elongated protein with a highly extended fiber-like conformation. There are three copies of the TmPs in each trisymmetron (totally 60 copies in the whole virion), each of which spans from the center of one pentasymmetron to the edge of a neighboring pentasymmetron. The first TmP homolog was seen in PRD1 bacteriophage and was suggested to play roles in regulating the size of the icosahedral viral capsid (Abrescia et al., 2004; Bamford et al., 2005). The mCP11 is the so-called zip protein found at the boundaries of symmetrons; mCP11 also has the highest copy number among all 14 types of mCPs identified in PBCV-1. The mCP3, having 9 copies in each trisymmetron (totally 180 copies in the whole virion), is also in extended

fiber-like conformation. Interestingly, the three mCP3s located near the threefold center possess an additional N-terminal helix bundling at the center of the trisymmetron. In addition, the mCP3s span from this helix bundle to near the edge of the trisymmetron connecting with the mCP2 TmPs. Based on these observations, the authors proposed an assembly model in which the capsid assembly starts at the center of the trisymmetron, where the mCP3 helix bundle serves as the initiation site for capsomer gathering (Fig. 5A and B). While more capsomers are assembled, the mCP4 and mCP5 are added to secure the capsomer assembly (Fig. 5C and D). The assembly proceeds with more capsomers added until mCP3 proteins contact with mCP2 TmPs, establishing the edge of the trisymmetron (Fig. 5E and F).

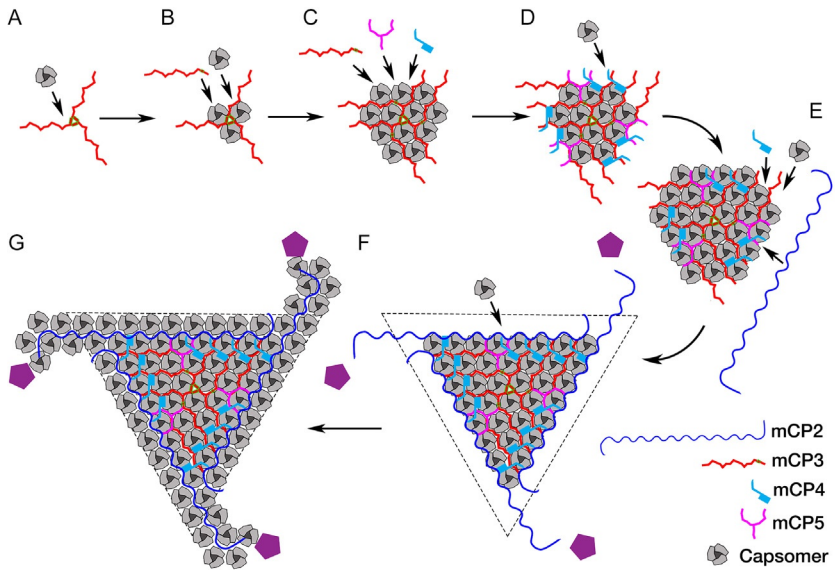

Fig. 5 A schematic diagram of the proposed symmetron model. (A) The helix bundle formed by three mCP3s located at the threefold center serving as an initiation site for capsomer assembly; (B) The assembly of the first layer of capsomers and more mCP3s; (C) The addition of a second layer of capsomers and more mCPs including mCP3s, mCP4 and mCP5; (D) The assembly of the third layer of the capsomers with the help of all three mCPs; (E) The binding of mCP2 tape measure proteins (TmPs) to the third layer of capsomers and the C-terminals of the mCP3s while more capsomers are assembled with the help of mCPs; (F) The TmPs establish the boundaries of trisymmetrons and link the nearby fivefold vertices (purple pentagons). (G) The completion of a trisymmetron as more capsomers are assembled onto the TmPs. Furthermore, with the help of TmPs and possibly other mCPs, the capsomers continue to assemble in the pentasymmetron regions.

The trisymmetron then connects to its neighboring pentasymmetrons with TmPs (Fig. 5G). Meanwhile, the zipper proteins glue the neighboring symmetrons together by binding to the capsomers located on the edges of the symmetrons.

This model emphasizes that the mCP3 plays essential roles in the capsid assembly by not only providing the initiation site for the assembly, but also regulating the size of the trisymmetrons. Both mCP2 TmP and mCP3 are involved in differentiating the size of NCVs. However, this assembly model suggests the assembly starts from the flat center of the trisymmetron, which has never been observed in any EM studies. On the contrary, the assembly was shown to be initiated from one of the vertices by many studies mentioned above (Fig. 4). In addition, no preassembled trisymmetrons or large arrays of capsomers have been observed inside the infected cells. Assembling from prepacked arrays would show a step-wise process where only certain sizes of procapsids would be observed, which is inconsistent with the aforementioned EM observations (Chlanda et al., 2009; Krijnse Locker et al., 2013; Milrot et al., 2016; Mutsafi et al., 2013).

6.2 The spiral assembly model

As mentioned, the capsomers within the same trisymmetron are packed in the same orientation, while they are 60° rotated when being compared to those in the neighboring trisymmetrons (Fig. 6A). Therefore, if the 30 hexagonal capsomers in the same pentasymmetron were to be divided into five small triangular arrays, one would assume, based on the packing of the trisymmetrons, that the six capsomers in the same triangular array would share the same orientation. However, the orientations of these capsomers within the pentasymmetrons had not been thoroughly examined until 2017 when Xiao et al. elucidated that one capsomer in the third layer (capsomer P1d in Fig. 6Aa) of the triangular array is rotated by 60° compared to the other five capsomers (Fig. 6Aa). Meanwhile, this uniquely orientated capsomer shares the same orientation as the capsomers from its counterclockwise neighboring triangular array (Fig. 6A). As a result, the orientation of capsomers in the pentasymmetron shows an interesting pattern that resembles five interlocked heads of golf clubs (Fig. 6A). More interestingly, those uniquely orientated capsomers (capsomer P1d in Fig. 6Aa and its fivefold related equivalents) share the same orientations with those in the nearby trisymmetrons forming a spiral pattern (Fig. 6A). This spiral pattern is conserved in all the icosahedral NCVs. Xiao et al. then proposed an assembly

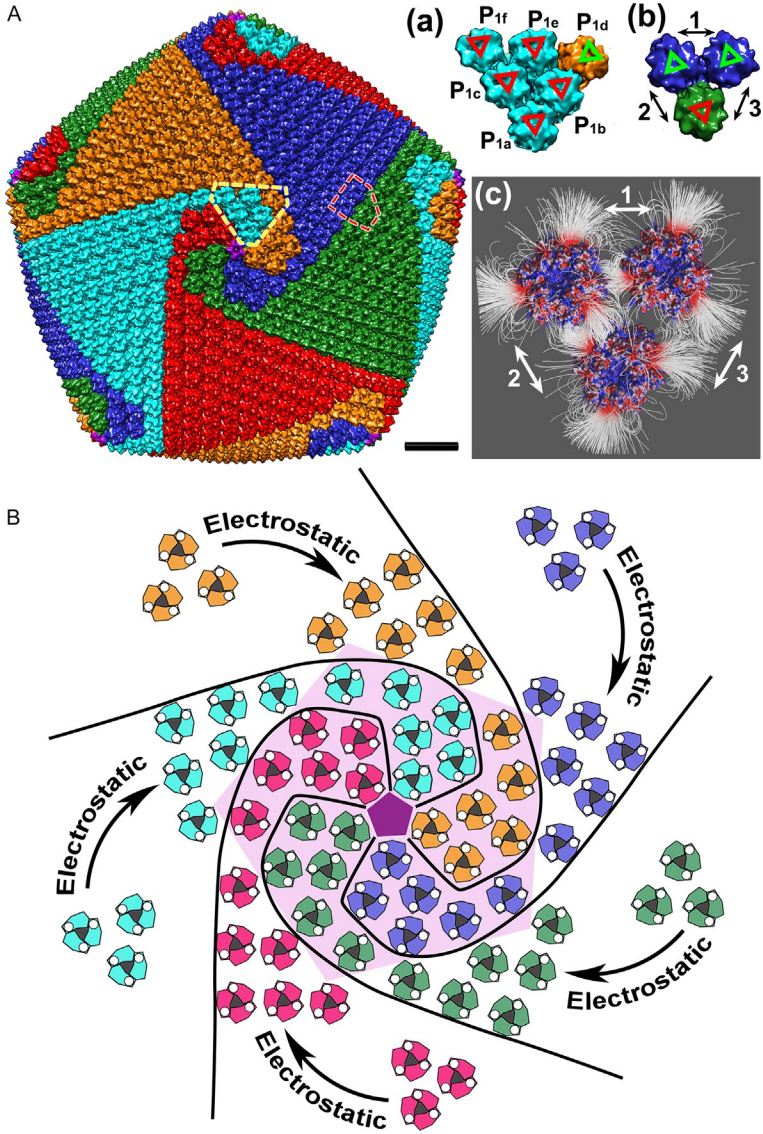

Fig. 6 The capsid structural features and a schematic diagram of the proposed spiral assembly model. (A) The capsid of PBCV-1, where the hexameric capsomers are colored based on their packing orientations in red, green, blue, orange, and cyan, respectively. The pentons located at the fivefold vertices are colored in purple. One ASU of the pentasymmetron and the three capsomers located at the twofold center are outlined by yellow and red dashed lines, respectively; Subpanel (a) An enlarged view of the ASU of the pentasymmetron outlined by yellow dashed lines in panel (A). The orientation of the capsomers is highlighted with red and green triangles, respectively. The vertices of these triangles correspond to the locations of jelly roll 2 (J2 in Fig. 1E);

model where the assembly of the icosahedral viral capsid initiates from the fivefold vertex, as the capsomers continuously assemble in a spiral fashion (Fig. 6B) (Xiao et al., 2017).

Following the proposal of the spiral model, to understand the reason behind the spiral pattern and the selective orientation of capsomers, a computational study was performed in the same lab (Xian et al., 2019). Using PBCV-1, which has the highest resolution cryo-EM map as well as the atomic structure of its MCPs, this study calculated and compared the free binding energies between capsomers that share the same orientation and those with different orientations at the boundary of the trisymmetron (Fig. 6Ab). Their results revealed three different binding modes (Fig. 6Ab–c) of capsomers with different binding strengths. Binding mode 1, with intermediate strength, is found among all capsomers located inside the same trisymmetron (sharing the same orientations). Binding mode 1 is also the dominate one throughout the capsid since most capsomers are in trisymmetrons sharing the same orientation. The other two binding modes (2 and 3) were found between capsomers located on the boundaries of trisymmetrons. Binding mode 2 has the strongest binding energy while mode 3 has the weakest. Detailed energy calculations showed that electrostatic interaction is the main factor that distinguishes these three binding modes (Fig. 6Ac). Therefore, the electrostatic potential on the capsomer surface was calculated and showed an interesting distribution pattern: while the six side vertices of the hexagonal capsomer are always dominated by positive charged residues, the grooves between vertices have alternating negative and positive charges. Furthermore, simulation of interactions between two capsomers positioned at different distances and orientations showed that the electrostatic forces can guide capsomers toward

Subpanel (b) An enlarged view of capsomers located at the twofold center outlined by red dashed lines in panel (A). The orientation of the capsomers are highlighted with red and green triangles, respectively. The three binding modes are indicated by double-headed arrows; Subpanel (c) The electrostatic interactions between the capsomers at the twofold center showing the electrostatic field lines. The three capsomers have the same orientations as in subpanel (b) but are separated 20 Å away from each other in order to show the electrostatic field lines. Negatively and positively charged capsomer surface areas are colored with a scale from −1.0 to 1.0 kT/Å. (B) A simplified diagram of the spiral assembly model where the assembly initiates from the penton (the purple pentagon) at the fivefold vertex; the capsomers assemble around the penton layer by layer in a spiral pattern, which is indicated by the five curved black lines. The pentasymmetron region is shaded with a light pink pentagon in the background. The electrostatic interaction may facilitate the assembly of capsomers by favoring their packing orientations into the spiral pattern. *Panel (Ac): The image is adapted from reference Xian, Y., Karki, C.B., Silva, S.M., Li, L., Xiao, C., 2019. The roles of electrostatic interactions in capsid assembly mechanisms of giant viruses. Int. J. Mol. Sci. 20 (8) with permission*

the final positions and stabilize their assembly. A similar charge distribution was also observed in the capsomer of ASFV (Liu et al., 2019a), suggesting that the capsomers from different icosahedral NCVs may share a similar mechanism where electrostatic interactions are involved in favoring their assembly in the described orientation that follows a spiral pattern (Fig. 6E).

To the best of our knowledge, the spiral assembly model is the first one proposed that is in agreement with current EM studies showing that the assembly is a continuous process initiating from the fivefold vertex (Chlanda et al., 2009; Krijnse Locker et al., 2013; Milrot et al., 2016; Mutsafi et al., 2013). This spiral model emphasizes the role of the MCP capsomers in forming into a high-ordered structure. The crucial role of MCP in the icosahedral capsid assembly is supported by experimental observations. Infecting host cells with the recombinant ASFV whose MCP expression was repressed resulted in the accumulation of extended membrane-like structures instead of the icosahedral procapsids. By introducing functional MCP back into the infected cell with various methods, the formation of icosahedral capsid was rescued (Epifano et al., 2006; Garcia-Escudero et al., 1998). These results collectively suggested that the MCPs play a significant role in assembling the icosahedral capsids.

The spiral model provided some clear descriptions of the initialization of the continuous assembly process where the electrostatic properties of the MCPs might guide the assembly of capsomers. However, the spiral model left two unanswered questions about how the spiral pattern progresses from the initial fivefold vertex to the neighboring ones and how the size of the icosahedral NCVs is controlled.

6.3 The mCP scaffold model

The recent epidemic of ASFV from 2018 to 2019 in China led to a devastating impact to the food industry and economy in general (Zhou et al., 2019). In the efforts to develop vaccines or drugs against the disease, two near-atomic cryo-EM structural studies of ASFV were published in 2019 (Liu et al., 2019a; Wang et al., 2019). In these two structural studies, besides the MCP (p72 of ASFV), four mCPs were characterized (Table 2) with different nomenclatures. In this review, we follow the names used by Wang et al. Compared to the 14 types of mCPs identified in PBCV-1, the lower number of mCPs in ASFV might be due to the lower resolution of the map. Alternatively, this might be because ASFV has the icosahedral core shell to stabilize the virion, thus requiring less mCPs.

Among the four ASFV mCPs, the penton protein and TmP are structural homologs to those in PBCV-1, respectively. Another mCP p17 was found

Table 2 Names of mCPs in ASFV from the two structural studies.

mCPs	(Wang et al., 2019)	(Liu et al., 2019a)	Gene ID
Penton protein	Penton protein	Penton protein	H240R
TmP	pM1249L	P2	M1249L
p17	p17	P1	D117L
p49	p49	Not observed	B438L
Zipper protein	Not observed	P3	To be identified

on the inner viral membrane of ASFV forming a similar hexagonal network to that of PBCV-1 without any helical bundles located at the threefold center. It was suggested that p17 adapts a different conformation at the boundaries of the symmetrons (Wang et al., 2019) while the other reconstruction study suggested the presence of a different mCP (P3) that needs to be identified (Liu et al., 2019a). This p17, with comparable abundance to MCP, was also found on the viral membrane precursor and other capsid assembly intermediates of ASFV (Rouiller et al., 1998; Suarez et al., 2010). The absence of p17 blocked the ASFV morphogenesis, leading to the accumulation of viral membrane precursors as well as the delocalization of the MCPs. These studies suggested a strong relationship between the p17 and MCPs, which is crucial to drive the assembly process of intact icosahedral particles. The last mCP, p49, was proposed to be located below the penton based on the weak lantern-like densities connecting the penton and inner membrane while associating the five neighboring capsomers (Wang et al., 2019), consistent with immunoelectron microscopic studies on intact virions (Alejo et al., 2018). However, these densities were too weak for determining the atomic structure, and were not observed in the other reconstruction of ASFV (Liu et al., 2019a). In the viral factory of the host cell infected by ASFV recombinants with repressed expression of p49, large tubular structures accumulated instead of normal icosahedral virus particles. These tubular structures were covered by a capsid-like layer that consists of MCPs, p17, and core shell components (Epifano et al., 2006). These results suggested that the p49 is likely to be located at the fivefold vertex and is essential in initializing the capsid assembly.

In addition to the published mCPs (Table 2), Liu et al. pointed out that there are additional densities on the ASFV capsid that may represent more mCPs to be identified in future higher resolution structures. Nevertheless, the structural and molecular virological studies of mCPs have led to another interesting model on the assembly of NCV capsids (Fig. 7)

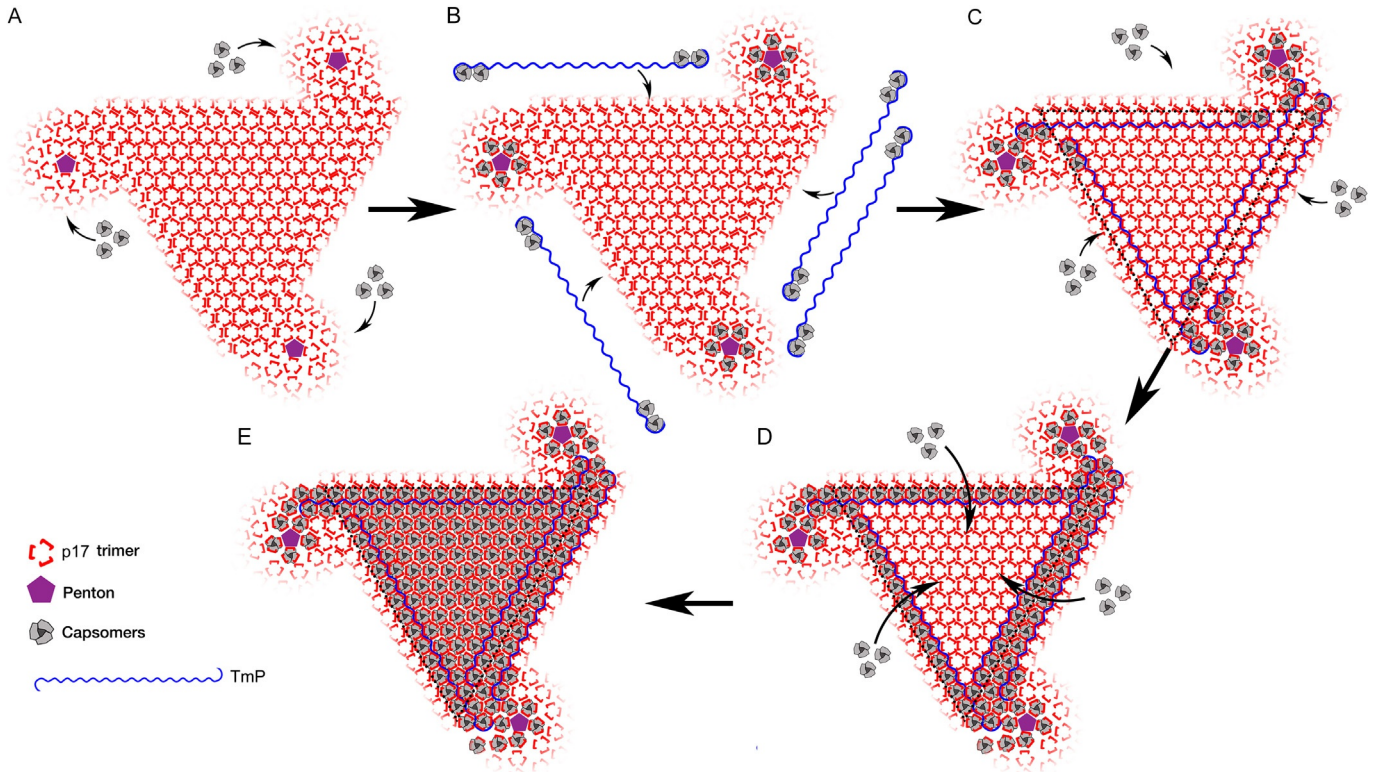

Fig. 7 A schematic diagram of the proposed mCP scaffold model. (A) The initiation of the capsid assembly at the fivefold vertices; (B) The attachment of TmPs and the carried capsomers; (C) The docking of capsomers near the TmPs; (D) The continuous docking of capsomers in the framework; (E) The completed docking of capsomers in the frameworks. The p17, penton, pseudo-hexameric capsomer, and TmP are represented as indicated by the legend on the lower left.

(Wang et al., 2019): prior to capsid assembly, p17 randomly floats on the inner membrane. Initiated by p49, the penton docks on the inner membrane and recruits the MCPs to the vertex (Fig. 7A). The TmPs then function as the skeletons and carry more capsomers to expand the penton core to form the pentasymmetron (Fig. 7B). The TmPs will also establish the boundaries of the trisymmetron (Fig. 7C), forming the polyhedral framework that allows MCPs to dock on the membrane surface with the help of p17 (Fig. 7D). As the TmPs extend the framework to form an icosahedral cage with the concaved membrane inside, the MCPs continue to dock on the membrane surface until the capsid assembly is complete (Fig. 7E).

This model is not only consistent with previous EM studies but also supported by many molecular virological studies that show p17 and p49 are essential for icosahedral capsid assembly (Epifano et al., 2006; Suarez et al., 2010). However, in this model, the capsid extension from the initial vertex to the symmetrons relies heavily on the TmP's ability to build the polyhedral framework. As mentioned, the TmP is an elongated protein with extended fiber-like structure. In addition, the TmP from PBCV-1 is rich in glycine residues (9.7%) indicating that it is very flexible. It is hard to imagine that such a protein can form the rigid framework. In addition, this model neglects the function of MCP, which is inconsistent with molecular and EM studies showing MCP is essential for the formation of the polyhedron. No polyhedral membrane structures were observed when MCP expression was repressed (Epifano et al., 2006, Garcia-Escudero et al., 1998).

7. The tape measure proteins (TmPs) on icosahedral NCV capsid assembly

As a non-icosahedral NCV, the assembly of VACV shares many common features with icosahedral NCVs, such as the membrane reorganization and the "crescent" precursor that contains an inner smooth lipid membrane and an external honeycomb lattice layer formed by the D13 scaffold protein (Liu et al., 2014; Risco et al., 2002; Szajner et al., 2005). The structure of the D13 scaffold protein is highly conserved with those of the MCPs from many icosahedral NCVs (Fig. 1H). However, the assembly of the scaffold protein D13 on the growing membrane results in spherical particles instead of icosahedral ones (Fig. 4E and F) (Bisht et al., 2009; Chichon et al., 2009; Chlanda et al., 2009; Liu et al., 2014). Inspired by the symmetron model and mCP scaffold model where the TmPs play crucial roles in icosahedral capsid assembly, we did a bioinformatic analysis to search for TmP homologs

in all available NCV genomes using the two TmPs identified in PBCV-1 and ASFV. Our results suggested that many icosahedral NCVs such as members from the Asfarviridae, Mimiviridae, and Phycodnaviridae posess TmPs homologs. Interestingly, no TmP homologs were found in poxviruses (Xian et al., 2020). We speculate that the lack of TmP is an essential factor that leads to the spherical instead of icosahedral symmetry of immature VACV particles (Fig. 4E and F).

The two questions left by the spiral assembly model can also be answered by the functions of TmP. The length of the TmPs are proportional to the size of the NCV capsids (Fang et al., 2019; Liu et al., 2019a; Wang et al., 2019). Since all the pentasymmetrons of icosahedral NCVs are the same, the size of the viral particles are determined by the size of the trisymmetrons (Table 1). As the TmPs connect two neighboring pentasymmetrons and "measure" the length of the edges of the trisymmetrons, it provides a mechanism for not only extending the assembly from one vertex to the next but also regulating the size of the particle.

8. A refined assembly model for giant virus capsids

These results from PBCV-1 and ASFV collectively suggest that the successful assembly of the NCVs requires the precise and efficient cooperation of both mCPs and MCPs. While the MCPs play significant roles in establishing the icosahedral geometry of the viral capsids as they progressively assemble onto the membrane, the mCPs closely interact with the MCPs as well as the viral membranes stabilizing the multilayer viral architecture. Although the three assembly models mentioned above did suggest some interesting mechanisms of icosahedral NCV capsid assembly, they all have problems. For example, the symmetron model is not consistent with the continuous process of assembly seen in many EM studies. The spiral model does not consider the roles of mCPs and only shows the initiation of the assembly. The mCP scaffold model relies on the preassembled TmP framework for MCP docking. Therefore, using the highest resolution structural information from PBCV-1, we present a refined model of icosahedral NCV capsid assembly. This model is speculative, based on the locations and the interactions among capsid proteins while considering all the valuable components in the three proposed models as well as the experimental observations discussed above. The steps of the proposed assembly model using mostly mCP names from PBCV-1 are listed below:

1. In the initial state, the membrane is ruptured from the ER by protein homologs to the p17 found in ASFV;
2. To initiate the assembly, the penton (P0) is then anchored onto the ruptured membrane with the help of mCPs: mCP14, mCP9, mCP10 and possible homologs to the p49 from ASFV (Fig. 8A);
3. The N-terminals of the TmPs are then locked onto the initial fivefold vertex through their interaction with the mCPs, especially mCP10s. This is because the individual mCP10 wraps around the N-terminus of each TmP. Meanwhile, the free ends of TmPs remain loose and flexible (Fig. 8B);

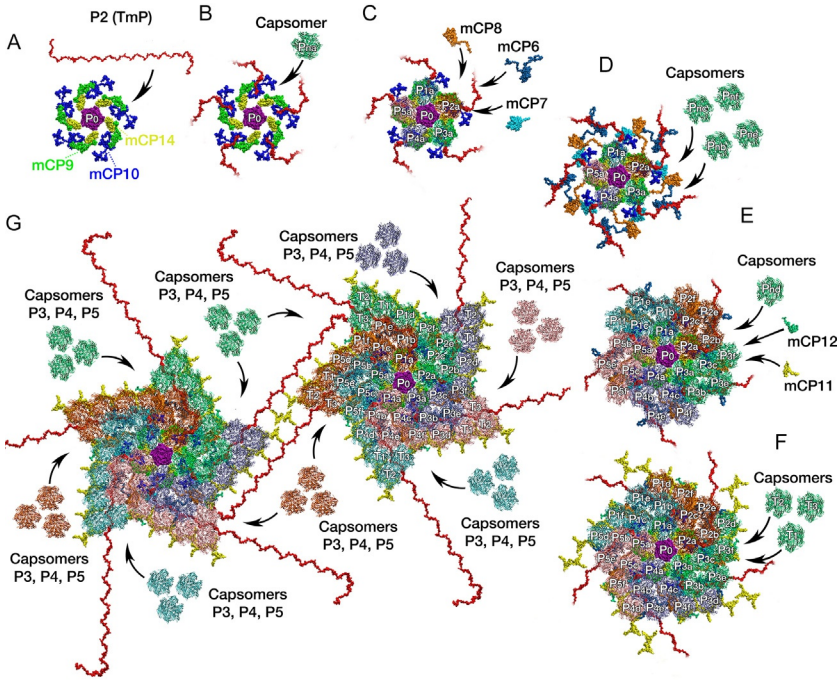

Fig. 8 The refined model of NCV assembly using mCPs identified in PBCV-1. (A) the initiation complex of the assembly at the fivefold vertex; (B) The attachment of TmPs; (C) The first layer of the pentasymmetron and addition of mCP6, mCP7, and mCP8; (D) The bending of TmP and the addition of the second and the third layer of the pentasymmetron; (E) The docking of the unique Pnd and the binding of mCP11 and mCP12; (F) Extension of pentasymmetron to trisymmetron; (G) The connection of two pentasymmetrons by TmPs and the continuation of the trisymmetron assembly with the help of mCP11 and other mCPs.

4. Five MCP capsomers then dock onto the network built by the mCPs (mCP14, mCP9, mCP10) and establish the first layer of the pentasymmetron (Fig. 8C). These capsomers are labeled as Pna where n represents ASU numbers from 1 to 5. The capsomers from the outer layer are labeled in similar ways as Pnb, Pnc, Pnd, Pne, and Pnf in the following panels. The interactions of the five Pna capsomers with the P0, the mCPs, and the membrane facilitate their precise docking (Fig. 8C);
5. More mCPs (mCP6, mCP7; mCP8) associate with the membrane as the open membrane sheet continues to grow. Alternatively, the mCPs facilitate the growth of the viral membrane by rupturing more ER membranes. These mCPs will also bind to the loose ends of the TmPs and help themselves to be associated with the ruptured membrane. It is noteworthy that a major bend of the TmP occurs right before its contact with mCP6. Thus, the mCP6 may introduce this structural bend of the TmP, resulting in the interesting spiral pattern of the five TmPs (Fig. 8D);
6. Meanwhile, more MCP capsomers are recruited to the fivefold center to assemble the second (Pnb and Pnc) and the third layers (Pne and Pnf) of the pentasymmetron; the successful assembly of these two layers allows the TmPs to wrap around the Pnb and Pne capsomers (Fig. 8D);
7. The assembly of the Pnd capsomer appears to be a rather unique scenario as the Pnd seems to be excluded from the other five capsomers (Pna, Pnb, Pnc, Pne and Pnf) by the TmP. The Pnd also interacts with the mCP6 that is speculated to bend the TmP in step 5. The mCP6 has one main body and two arms (mCP6 in Fig. 8C). The main body and one of the arms of mCP6 wrap around the Pnd while the other arm of the mCP6 interacts with the mCP8. Coincidently, the orientation of the Pnd is rotated by 60° compared to the other five capsomers located in the same ASU of pentasymmetron (for example, P1d is 60° rotated compared to P1a, P1b, P1c, P1e, P1f) (Fig. 8E); therefore, we speculate that the unique orientation of Pnd results from the interaction among the Pnd, mCP6 and TmP. While Pnd is docked, both mCP11 and mCP12 approach the capsomers at the boundaries of the pentasymmetron, securing and completing the assembly of the pentasymmetron (Fig. 8E);
8. More capsomers (Tna, Tnb, Tnc, etc., where n represents ASU numbers from 1 to 5) are recruited to assemble the trisymmetron (Fig. 8F);

9. As the assembly of trisymmetron continues, the free C-termini of TmPs are in a search mode until they find and bind to another assembled pentasymmetron. The free end of TmP from the latter pentasymmetron will bind to the initial pentasymmetron, leading to two antiparallel TmPs connecting these two pentasymmetrons. The connection established by TmPs will allow the membrane growth and capsomer assembly to proceed to the next vertex with the help of other mCPs (Fig. 8G);
10. As the assembly of viral capsid proceeds, more vertices are connected to the assembling polyhedral capsid through a similar mechanism in step 9, allowing the continuous assembly of capsomers and membrane until the assembly of viral capsid is completed with a icosahedral symmetry.

9. Conclusion

The process of giant virus assembly has always been a challenging but fascinating topic as it occurs in a very complicated biological environment and requires close and precise associations from the viral membranes and thousands of protein units. Two decades of structural studies, especially the two recent high-resolution cryo-EM reconstructions, have revealed the extremely complicated architectures of the icosahedral NCVs. In addition, the viral morphogenesis studies of NCVs focusing on their viral factories have provided fruitful information on the assembly of their capsids. Striking similarities are found in both structures and morphogenesis of NCVs, indicating these viruses are likely to be assembled in similar mechanisms and evolve from a common ancestor. Based on these analyses of the structural and molecular experimental data, scientists have proposed several models that deepened our understanding of the assembly mechanism of such large and complex viral capsids. Here, we reviewed and commented on three recent and popular models. Furthermore, we suggested a new model that takes into consideration many of the experimental evidence, as well as the recent high-resolution structures of the viral capsids. This model provides a more comprehensive description on the structures of the essential capsid proteins and their potential functions in the assembly. However, many questions remain unanswered. What are the factors that constrain the size of the pentasymmetrons? What are the molecular driving forces behind the association among the viral membrane, mCPs, and MCPs? What are the mechanisms that coordinate the viral capsid assembly and viral genome encapsidation? How are the unique portals observed in Mimivirus and PBCV-1 assembled? All proposed models remain speculations pending

further investigation of the assembly process through both structural and molecular virological studies. For instance, higher resolution structures of icosahedral NCVs will identify new viral factors that contribute to the capsid assembly. Another useful experiment that represses TmP expression using recombinant NCVs such as that of ASFV (Garcia-Escudero et al., 1998) would provide deeper understanding on critical role of the TmPs in the assembly process.

In conclusion, this review discusses the current breakthroughs as well as the existing models on the assembly process of icosahedral NCV capsids, highlights the possible roles of viral membrane and capsid proteins, and draws attention to the to-be-answered questions on the existing models that require further examination, hence, piloting the possible directions for future investigations. Deciphering the sophisticated assembly process of the large and complicated icosahedral NCV capsids will advance our knowledge in the mega-assembly of other biological supramolecular complexes.

Acknowledgments

The authors greatly appreciate Ricardo Avila and Brenda Moreno for their proofreading and constructive criticisms of the manuscript.

Funding

Research reported in this publication was supported by the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number R01GM129525. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Abrahao, J., Silva, L., Silva, L.S., Khalil, J.Y.B., Rodrigues, R., Arantes, T., Assis, F., Boratto, P., Andrade, M., Kroon, E.G., Ribeiro, B., Bergier, I., Seligmann, H., Ghigo, E., Colson, P., Levasseur, A., Kroemer, G., Raoult, D., La Scola, B., 2018. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. Nat. Commun. 9 (1), 749.

Abrescia, N.G., Cockburn, J.J., Grimes, J.M., Sutton, G.C., Diprose, J.M., Butcher, S.J., Fuller, S.D., San Martin, C., Burnett, R.M., Stuart, D.I., Bamford, D.H., Bamford, J.K., 2004. Insights into assembly from structural analysis of bacteriophage PRD1. Nature 432 (7013), 68–74.

Ahlquist, P., 2006. Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. Nat. Rev. Microbiol. 4 (5), 371–382.

Alejo, A., Matamoros, T., Guerra, M., Andres, G., 2018. A proteomic atlas of the African swine fever virus particle. J. Virol. 92 (23).

Andreani, J., Khalil, J.Y.B., Sevvana, M., Benamar, S., Di Pinto, F., Bitam, I., Colson, P., Klose, T., Rossmann, M.G., Raoult, D., La Scola, B., 2017. Pacmanvirus, a new giant icosahedral virus at the crossroads between Asfarviridae and Faustoviruses. J. Virol. 91 (14).

Andres, G., Simon-Mateo, C., Vinuela, E., 1997. Assembly of African swine fever virus: role of polyprotein pp220. J. Virol. 71 (3), 2331–2341.
Andres, G., Garcia-Escudero, R., Simon-Mateo, C., Vinuela, E., 1998. African swine fever virus is enveloped by a two-membraned collapsed cisterna derived from the endoplasmic reticulum. J. Virol. 72 (11), 8988–9001.
Andres, G., Garcia-Escudero, R., Salas, M.L., Rodriguez, J.M., 2002. Repression of African swine fever virus polyprotein pp220-encoding gene leads to the assembly of icosahedral core-less particles. J. Virol. 76 (6), 2654–2666.
Andres, G., Charro, D., Matamoros, T., Dillard, R.S., Abrescia, N.G.A., 2020. The cryo-EM structure of African swine fever virus unravels a unique architecture comprising two icosahedral protein capsids and two lipoprotein membranes. J. Biol. Chem. 295 (1), 1–12.
Arslan, D., Legendre, M., Seltzer, V., Abergel, C., Claverie, J.M., 2011. Distant mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. Proc. Natl. Acad. Sci. U. S. A. 108 (42), 17486–17491.
Asgari, S., Bideshi, D.K., Bigot, Y., Federici, B.A., Cheng, X.W., Ictv Report, C., 2017. ICTV virus taxonomy profile: Ascoviridae. J. Gen. Virol. 98 (1), 4–5.
Bahar, M.W., Graham, S.C., Stuart, D.I., Grimes, J.M., 2011. Insights into the evolution of a complex virus from the crystal structure of vaccinia virus D13. Structure 19 (7), 1011–1020.
Bamford, D.H., 2003. Do viruses form lineages across different domains of life? Res. Microbiol. 154 (4), 231–236.
Bamford, D.H., Burnett, R.M., Stuart, D.I., 2002. Evolution of viral structure. Theor. Popul. Biol. 61 (4), 461–470.
Bamford, D.H., Grimes, J.M., Stuart, D.I., 2005. What does structure tell us about virus evolution? Curr. Opin. Struct. Biol. 15 (6), 655–663.
Bell, P.J., 2001. Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus? J. Mol. Evol. 53 (3), 251–256.
Benamar, S., Reteno, D.G., Bandaly, V., Labas, N., Raoult, D., La Scola, B., 2016. Faustoviruses: comparative genomics of new Megavirales family members. Front. Microbiol. 7, 3.
Benson, S.D., Bamford, J.K., Bamford, D.H., Burnett, R.M., 1999. Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. Cell 98 (6), 825–833.
Benson, S.D., Bamford, J.K., Bamford, D.H., Burnett, R.M., 2004. Does common architecture reveal a viral lineage spanning all three domains of life? Mol. Cell 16 (5), 673–685.
Bergh, Ø., BØrsheim, K.Y., Bratbak, G., Heldal, M., 1989. High abundance of viruses found in aquatic environments. Nature 340 (6233), 467–468.
Bisht, H, Weisberg, A.S., Szajner, P, Moss, B, 2009. Assembly and disassembly of the capsid-like external scaffold of immature virions during vaccinia virus morphogenesis. J. Virol. 83 (18), 9140–9150.
Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., Robert, C., Azza, S., Sun, S., Rossmann, M.G., Suzan-Monti, M., La Scola, B., Koonin, E.V., Raoult, D., 2009. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc. Natl. Acad. Sci. U. S. A. 106 (51), 21848–21853.
Campos, R.K., Boratto, P.V., Assis, F.L., Aguiar, E.R., Silva, L.C., Albarnaz, J.D., Dornas, F.P., Trindade, G.S., Ferreira, P.P., Marques, J.T., Robert, C., Raoult, D., Kroon, E.G., La Scola, B., Abrahao, J.S., 2014. Samba virus: a novel mimivirus from a giant rain forest, the Brazilian Amazon. Virol. J. 11, 95.
Caspar, D.L., Klug, A., 1962. Physical principles in the construction of regular viruses. Cold Spring Harb. Symp. Quant. Biol. 27, 1–24.

Cherrier, M.V., Kostyuchenko, V.A., Xiao, C., Bowman, V.D., Battisti, A.J., Yan, X., Chipman, P.R., Baker, T.S., Van Etten, J.L., Rossmann, M.G., 2009. An icosahedral algal virus has a complex unique vertex decorated by a spike. Proc. Natl. Acad. Sci. U. S. A. 106 (27), 11085–11089.
Chichon, F.J., Rodriguez, M.J., Risco, C., Fraile-Ramos, A., Fernandez, J.J., Esteban, M., Carrascosa, J.L., 2009. Membrane remodelling during vaccinia virus morphogenesis. Biol. Cell 101 (7), 401–414.
Chlanda, P., Carbajal, M.A., Cyrklaff, M., Griffiths, G., Krijnse-Locker, J., 2009. Membrane rupture generates single open membrane sheets during vaccinia virus assembly. Cell Host Microbe 6 (1), 81–90.
Chuong, E.B., 2013. Retroviruses facilitate the rapid evolution of the mammalian placenta. Bioessays 35 (10), 853–861.
Claverie, J.M., Abergel, C., 2009. Mimivirus and its virophage. Annu. Rev. Genet. 43, 49–66.
Claverie, J.M., Ogata, H., Audic, S., Abergel, C., Suhre, K., Fournier, P.E., 2006. Mimivirus and the emerging concept of "giant" virus. Virus Res. 117 (1), 133–144.
Claverie, J.M., Abergel, C., Ogata, H., 2009. Mimivirus. Curr. Top. Microbiol. Immunol. 328, 89–121.
Colson, P., de Lamballerie, X., Fournous, G., Raoult, D., 2012. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. Intervirology 55 (5), 321–332.
Crick, F.H., Watson, J.D., 1956. Structure of small viruses. Nature 177 (4506), 473–475.
Dales, S., Siminovitch, L., 1961. The development of vaccinia virus in Earle's L strain cells as examined by electron microscopy. J. Biophys. Biochem. Cytol. 10, 475–503.
Donelli, G., Guglielmi, F., Paoletti, L., 1972. Structure and physico-chemical properties of bacteriophage G. I. Arrangement of protein subunits and contraction process of tail sheath. J. Mol. Biol. 71 (2), 113–125.
Donelli, G., Dore, E., Frontali, C., Grandolfo, M.E., 1975. Structure and physico-chemical properties of bacteriophage G. III. A homogeneous DNA of molecular weight 5 times 10(8). J. Mol. Biol. 94 (4), 555–565.
Dunlap, K.A., Palmarini, M., Varela, M., Burghardt, R.C., Hayashi, K., Farmer, J.L., Spencer, T.E., 2006. Endogenous retroviruses regulate periimplantation placental growth and differentiation. Proc. Natl. Acad. Sci. U. S. A. 103 (39), 14390–14395.
Dupressoir, A., Vernochet, C., Bawa, O., Harper, F., Pierron, G., Opolon, P., Heidmann, T., 2009. Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. Proc. Natl. Acad. Sci. U. S. A. 106 (29), 12127–12132.
Epifano, C., Krijnse-Locker, J., Salas, M.L., Salas, J., Rodriguez, J.M., 2006. Generation of filamentous instead of icosahedral particles by repression of African swine fever virus structural protein pB438L. J. Virol. 80 (23), 11456–11466.
Fang, Q., Zhu, D., Agarkova, I., Adhikari, J., Klose, T., Liu, Y., Chen, Z., Sun, Y., Gross, M.L., Van Etten, J.L., Zhang, X., Rossmann, M.G., 2019. Near-atomic structure of a giant virus. Nat. Commun. 10 (1), 388.
Filee, J., 2009. Lateral gene transfer, lineage-specific gene expansion and the evolution of nucleo cytoplasmic large DNA viruses. J. Invertebr. Pathol. 101 (3), 169–171.
Filee, J., Chandler, M., 2010. Gene exchange and the origin of giant viruses. Intervirology 53 (5), 354–361.
Fischer, M.G., 2011. Sputnik and Mavirus: more than just satellite viruses. Nat. Rev. Microbiol. 10 (1), 78 (author reply 78).
Fischer, M.G., Allen, M.J., Wilson, W.H., Suttle, C.A., 2010. Giant virus with a remarkable complement of genes infects marine zooplankton. Proc. Natl. Acad. Sci. U. S. A. 107 (45), 19508–19513.

Fontana, J., Lopez-Iglesias, C., Tzeng, W.P., Frey, T.K., Fernandez, J.J., Risco, C., 2010. Three-dimensional structure of rubella virus factories. Virology 405 (2), 579–591.
Forterre, P., 2002. The origin of DNA genomes and DNA replication proteins. Curr. Opin. Microbiol. 5 (5), 525–532.
Forterre, P., 2006a. The origin of viruses and their possible roles in major evolutionary transitions. Virus Res. 117 (1), 5–16.
Forterre, P., 2006b. Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. Proc. Natl. Acad. Sci. U. S. A. 103 (10), 3669–3674.
Garcia-Escudero, R., Andres, G., Almazan, F., Vinuela, E., 1998. Inducible gene expression from African swine fever virus recombinants: analysis of the major capsid protein p72. J. Virol. 72 (4), 3185–3195.
Gil-Carton, D., Jaakkola, S.T., Charro, D., Peralta, B., Castano-Diez, D., Oksanen, H.M., Bamford, D.H., Abrescia, N.G.A., 2015. Insight into the assembly of viruses with vertical single beta-barrel major capsid proteins. Structure 23 (10), 1866–1877.
Haig, D., 2012. Retroviruses and the placenta. Curr. Biol. 22 (15), R609–R613.
Harrison, S.C., Olson, A.J., Schutt, C.E., Winkler, F.K., Bricogne, G., 1978. Tomato bushy stunt virus at 2.9 A resolution. Nature 276 (5686), 368–373.
Henderson, R., 2018. From Electron crystallography to single particle CryoEM (Nobel Lecture). Angew. Chem. Int. Ed. Engl. 57 (34), 10804–10825.
Hogle, J.M., Chow, M., Filman, D.J., 1985. Three-dimensional structure of poliovirus at 2.9 A resolution. Science 229 (4720), 1358–1365.
Iyer, L.M., Aravind, L., Koonin, E.V., 2001. Common origin of four diverse families of large eukaryotic DNA viruses. J. Virol. 75 (23), 11720–11734.
Iyer, L.M., Balaji, S., Koonin, E.V., Aravind, L., 2006. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. Virus Res. 117 (1), 156–184.
Khayat, R., Tang, L., Larson, E.T., Lawrence, C.M., Young, M., Johnson, J.E., 2005. Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. Proc. Natl. Acad. Sci. U. S. A. 102 (52), 18944–18949.
Klose, T., Kuznetsov, Y.G., Xiao, C., Sun, S., McPherson, A., Rossmann, M.G., 2010. The three-dimensional structure of Mimivirus. Intervirology 53 (5), 268–273.
Klose, T., Reteno, D.G., Benamar, S., Hollerbach, A., Colson, P., La Scola, B., Rossmann, M.G., 2016. Structure of faustovirus, a large dsDNA virus. Proc. Natl. Acad. Sci. U. S. A. 113 (22), 6206–6211.
Koonin, E.V., Yutin, N., 2018. Multiple evolutionary origins of giant viruses. F1000Res 7, 1840. F1000 Faculty Rev.
Koonin, E.V., Yutin, N., 2019. Evolution of the large nucleocytoplasmic DNA viruses of eukaryotes and convergent origins of viral gigantism. Adv. Virus Res. 103, 167–202.
Koonin, E.V., Krupovic, M., Yutin, N., 2015. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. Ann. N. Y. Acad. Sci. 1341, 10–24.
Koonin, E.V., Dolja, V.V., Krupovic, M., Varsani, A., Wolf, Y.I., Yutin, N., Zerbini, M., JH, K., 2019. Create a Megataxonomic Framework, Filling all Principal Taxonomic Ranks, for DNA Viruses Encoding Vertical Jelly Roll-Type Major Capsid Proteins. ICTV TaxoProp.
Krijnse Locker, J., Chlanda, P., Sachsenheimer, T., Brugger, B., 2013. Poxvirus membrane biogenesis: rupture not disruption. Cell. Microbiol. 15 (2), 190–199.
Krupovic, M., Bamford, D.H., 2008. Virus evolution: how far does the double beta-barrel viral lineage extend? Nat. Rev. Microbiol. 6 (12), 941–948.
Krupovic, M., Koonin, E.V., 2017. Multiple origins of viral capsid proteins from cellular ancestors. Proc. Natl. Acad. Sci. U. S. A. 114 (12), E2401–E2410.

Kuznetsov, Y.G., Xiao, C., Sun, S., Raoult, D., Rossmann, M., McPherson, A., 2010. Atomic force microscopy investigation of the giant mimivirus. Virology 404 (1), 127–137.

La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., Birtles, R., Claverie, J.M., Raoult, D., 2003. A giant virus in amoebae. Science 299 (5615), 2033.

Laanto, E., Mantynen, S., De Colibus, L., Marjakangas, J., Gillum, A., Stuart, D.I., Ravantti, J.J., Huiskonen, J.T., Sundberg, L.R., 2017. Virus found in a boreal lake links ssDNA and dsDNA viruses. Proc. Natl. Acad. Sci. U. S. A. 114 (31), 8378–8383.

Lefkowitz, E.J., Dempsey, D.M., Hendrickson, R.C., Orton, R.J., Siddell, S.G., Smith, D.B., 2018. Virus taxonomy: the database of the international committee on taxonomy of viruses (ICTV). Nucleic Acids Res. 46 (D1), D708–D717.

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., Lescot, M., Poirot, O., Bertaux, L., Bruley, C., Coute, Y., Rivkina, E., Abergel, C., Claverie, J. M., 2014. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. Proc. Natl. Acad. Sci. U. S. A. 111 (11), 4274–4279.

Liu, L., Cooper, T., Howley, P.M., Hayball, J.D., 2014. From crescent to mature virion: vaccinia virus assembly and maturation. Viruses 6 (10), 3787–3808.

Liu, S., Luo, Y., Wang, Y., Li, S., Zhao, Z., Bi, Y., Sun, J., Peng, R., Song, H., Zhu, D., Sun, Y., Li, S., Zhang, L., Wang, W., Sun, Y., Qi, J., Yan, J., Shi, Y., Zhang, X., Wang, P., Qiu, H.J., Gao, G.F., 2019a. Cryo-EM structure of the African swine fever virus. Cell Host Microbe 26 (6), 836–843 (e833).

Liu, Q., Ma, B., Qian, N., Zhang, F., Tan, X., Lei, J., Xiang, Y., 2019b. Structure of the African swine fever virus major capsid protein p72. Cell Res. 29 (11), 953–955.

Lwoff, A., 1957. The concept of virus. J. Gen. Microbiol. 17 (2), 239–253.

Lwoff, A., Anderson, T.F., Jacob, F., 1959. Remarks on the characteristics of the infectious viral particle. Ann. Inst. Pasteur (Paris) 97, 281–289.

Milrot, E., Mutsafi, Y., Fridmann-Sirkis, Y., Shimoni, E., Rechav, K., Gurnon, J.R., Van Etten, J.L., Minsky, A., 2016. Virus-host interactions: insights from the replication cycle of the large Paramecium bursaria chlorella virus. Cell. Microbiol. 18 (1), 3–16.

Moniruzzaman, M., LeCleir, G.R., Brown, C.M., Gobler, C.J., Bidle, K.D., Wilson, W.H., Wilhelm, S.W., 2014. Genome of brown tide virus (AaV), the little giant of the Megaviridae, elucidates NCLDV genome expansion and host-virus coevolution. Virology 466–467, 60–70.

Mutsafi, Y., Shimoni, E., Shimon, A., Minsky, A., 2013. Membrane assembly during the infection cycle of the giant Mimivirus. PLoS Pathog. 9 (5), e1003367.

Mutsafi, Y., Fridmann-Sirkis, Y., Milrot, E., Hevroni, L., Minsky, A., 2014. Infection cycles of large DNA viruses: emerging themes and underlying questions. Virology 466–467, 3–14.

Nandhagopal, N., Simpson, A.A., Gurnon, J.R., Yan, X., Baker, T.S., Graves, M.V., Van Etten, J.L., Rossmann, M.G., 2002. The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. Proc. Natl. Acad. Sci. U. S. A. 99 (23), 14758–14763.

Ng, W.M., Stelfox, A.J., Bowden, T.A., 2020. Unraveling virus relationships by structure-based phylogenetic classification. Virus Evol. 6 (1), veaa003.

Novoa, R.R., Calderita, G., Arranz, R., Fontana, J., Granzow, H., Risco, C., 2005. Virus factories: associations of cell organelles for viral replication and morphogenesis. Biol. Cell 97 (2), 147–172.

Okamoto, K., Miyazaki, N., Reddy, H.K.N., Hantke, M.F., Maia, F., Larsson, D.S.D., Abergel, C., Claverie, J.M., Hajdu, J., Murata, K., Svenda, M., 2018. Cryo-EM structure of a Marseilleviridae virus particle reveals a large internal microassembly. Virology 516, 239–245.

Olsen, R.H., Siak, J.S., Gray, R.H., 1974. Characteristics of PRD1, a plasmid-dependent broad host range DNA bacteriophage. J. Virol. 14 (3), 689–699.
Philippe, N., Legendre, M., Doutre, G., Coute, Y., Poirot, O., Lescot, M., Arslan, D., Seltzer, V., Bertaux, L., Bruley, C., Garin, J., Claverie, J.M., Abergel, C., 2013. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. Science 341 (6143), 281–286.
Raoult, D., Forterre, P., 2008. Redefining viruses: lessons from Mimivirus. Nat. Rev. Microbiol. 6 (4), 315–319.
Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M., Claverie, J.M., 2004. The 1.2-megabase genome sequence of Mimivirus. Science 306 (5700), 1344–1350.
Reche, I., D'Orta, G., Mladenov, N., Winget, D.M., Suttle, C.A., 2018. Deposition rates of viruses and bacteria above the atmospheric boundary layer. ISME J. 12 (4), 1154–1162.
Risco, C., Rodriguez, J.R., Lopez-Iglesias, C., Carrascosa, J.L., Esteban, M., Rodriguez, D., 2002. Endoplasmic reticulum-Golgi intermediate compartment membranes and vimentin filaments participate in vaccinia virus assembly. J. Virol. 76 (4), 1839–1855.
Roberts, M.M., White, J.L., Grutter, M.G., Burnett, R.M., 1986. Three-dimensional structure of the adenovirus major coat protein hexon. Science 232 (4754), 1148–1151.
Rossmann, M.G., Johnson, J.E., 1989. Icosahedral RNA virus structure. Annu. Rev. Biochem. 58, 533–573.
Rossmann, M.G., Arnold, E., Erickson, J.W., Frankenberger, E.A., Griffith, J.P., Hecht, H.J., Johnson, J.E., Kamer, G., Luo, M., Mosser, A.G., et al., 1985. Structure of a human common cold virus and functional relationship to other picornaviruses. Nature 317 (6033), 145–153.
Rouiller, I., Brookes, S.M., Hyatt, A.D., Windsor, M., Wileman, T., 1998. African swine fever virus is wrapped by the endoplasmic reticulum. J. Virol. 72 (3), 2373–2387.
Salas, M.L., Andres, G., 2013. African swine fever virus morphogenesis. Virus Res. 173 (1), 29–41.
San Martin, C., van Raaij, M.J., 2018. The so far farthest reaches of the double jelly roll capsid protein fold. Virol. J. 15 (1), 181.
Simpson, A.A., Nandhagopal, N., Van Etten, J.L., Rossmann, M.G., 2003. Structural analyses of Phycodnaviridae and Iridoviridae. Acta Crystallogr. D Biol. Crystallogr. 59 (Pt. 12), 2053–2059.
Sinkovits, R.S., Baker, T.S., 2010. A tale of two symmetrons: rules for construction of icosahedral capsids from trisymmetrons and pentasymmetrons. J. Struct. Biol. 170 (1), 109–116.
Sodeik, B., Doms, R.W., Ericsson, M., Hiller, G., Machamer, C.E., Hof, W.v.'t., van Meer, G., Moss, B., Griffiths, G., 1993. Assembly of vaccinia virus: role of the intermediate compartment between the endoplasmic reticulum and the Golgi stacks. J. Cell Biol. 121 (3), 521–541.
Stewart, P.L., Burnett, R.M., Cyrklaff, M., Fuller, S.D., 1991. Image reconstruction reveals the complex molecular organization of adenovirus. Cell 67 (1), 145–154.
Suarez, C., Gutierrez-Berzal, J., Andres, G., Salas, M.L., Rodriguez, J.M., 2010. African swine fever virus protein p17 is essential for the progression of viral membrane precursors toward icosahedral intermediates. J. Virol. 84 (15), 7484–7499.
Suarez, C., Welsch, S., Chlanda, P., Hagen, W., Hoppe, S., Kolovou, A., Pagnier, I., Raoult, D., Krijnse Locker, J., 2013. Open membranes are the precursors for assembly of large DNA viruses. Cell. Microbiol. 15 (11), 1883–1895.
Suarez, C., Andres, G., Kolovou, A., Hoppe, S., Salas, M.L., Walther, P., Krijnse Locker, J., 2015. African swine fever virus assembles a single membrane derived from rupture of the endoplasmic reticulum. Cell. Microbiol. 17 (11), 1683–1698.

Subramaniam, K., Behringer, D.C., Bojko, J., Yutin, N., Clark, A.S., Bateman, K.S., van Aerle, R., Bass, D., Kerr, R.C., Koonin, E.V., Stentiford, G.D., Waltzek, T.B., 2020. A new family of DNA viruses causing disease in Crustaceans from diverse aquatic biomes. mBio 11 (1).

Suttle, C.A., 2005. Viruses in the sea. Nature 437 (7057), 356–361.

Szajner, P., Weisberg, A.S., Lebowitz, J., Heuser, J., Moss, B., 2005. External scaffold of spherical immature poxvirus particles is made of protein trimers, forming a honeycomb lattice. J. Cell Biol. 170 (6), 971–981.

Takemura, M., 2001. Poxviruses and the origin of the eukaryotic nucleus. J. Mol. Evol. 52 (5), 419–425.

Van Etten, J.L., Meints, R.H., Kuczmarski, D., Burbank, D.E., Lee, K., 1982. Viruses of symbiotic Chlorella-like algae isolated from Paramecium bursaria and Hydra viridis. Proc. Natl. Acad. Sci. U. S. A. 79 (12), 3867–3871.

Walker, P.J., Siddell, S.G., Lefkowitz, E.J., Mushegian, A.R., Dempsey, D.M., Dutilh, B.E., Harrach, B., Harrison, R.L., Hendrickson, R.C., Junglen, S., Knowles, N.J., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Nibert, M., Rubino, L., Sabanadzovic, S., Simmonds, P., Varsani, A., Zerbini, F.M., Davison, A.J., 2019. Changes to virus taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2019). Arch. Virol. 164 (9), 2417–2429.

Wang, N., Zhao, D., Wang, J., Zhang, Y., Wang, M., Gao, Y., Li, F., Wang, J., Bu, Z., Rao, Z., Wang, X., 2019. Architecture of African swine fever virus and implications for viral assembly. Science 366 (6465), 640–644.

Woese, C.R., Kandler, O., Wheelis, M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. U. S. A. 87 (12), 4576–4579.

Wrigley, N.G., 1969. An electron microscope study of the structure of Sericesthis iridescent virus. J. Gen. Virol. 5 (1), 123–134.

Wrigley, N.G., 1970. An electron microscope study of the structure of Tipula iridescent virus. J. Gen. Virol. 6 (1), 169–173.

Xian, Y., Xiao, C., 2020. The structure of ASFV advances the fight against the disease. Trends Biochem. Sci. 45 (4), 276–278.

Xian, Y., Karki, C.B., Silva, S.M., Li, L., Xiao, C., 2019. The roles of electrostatic interactions in capsid assembly mechanisms of giant viruses. Int. J. Mol. Sci. 20 (8), 1876.

Xian, Y., Avila, R., Pant, A., Yang, Z., Xiao, C., 2020. The role of tape measure protein in giant virus capsid assembly. Viral Immunol. In press.

Xiao, C., Rossmann, M.G., 2011. Structures of giant icosahedral eukaryotic dsDNA viruses. Curr. Opin. Virol. 1 (2), 101–109.

Xiao, C., Chipman, P.R., Battisti, A.J., Bowman, V.D., Renesto, P., Raoult, D., Rossmann, M.G., 2005. Cryo-electron microscopy of the giant mimivirus. J. Mol. Biol. 353 (3), 493–496.

Xiao, C., Kuznetsov, Y.G., Sun, S., Hafenstein, S.L., Kostyuchenko, V.A., Chipman, P.R., Suzan-Monti, M., Raoult, D., McPherson, A., Rossmann, M.G., 2009. Structural studies of the giant mimivirus. PLoS Biol. 7 (4), e92.

Xiao, C., Fischer, M.G., Bolotaulo, D.M., Ulloa-Rondeau, N., Avila, G.A., Suttle, C.A., 2017. Cryo-EM reconstruction of the Cafeteria roenbergensis virus capsid suggests novel assembly pathway for giant viruses. Sci. Rep. 7 (1), 5484.

Yan, X., Olson, N.H., Van Etten, J.L., Bergoin, M., Rossmann, M.G., Baker, T.S., 2000. Structure and assembly of large lipid-containing dsDNA viruses. Nat. Struct. Biol. 7 (2), 101–103.

Yan, X., Chipman, P.R., Castberg, T., Bratbak, G., Baker, T.S., 2005. The marine algal virus PpV01 has an icosahedral capsid with T=219 quasisymmetry. J. Virol. 79 (14), 9236–9243.

Yan, X., Yu, Z., Zhang, P., Battisti, A.J., Holdaway, H.A., Chipman, P.R., Bajaj, C., Bergoin, M., Rossmann, M.G., Baker, T.S., 2009. The capsid proteins of a large, icosahedral dsDNA virus. J. Mol. Biol. 385 (4), 1287–1299.

Yutin, N., Koonin, E.V., 2012. Hidden evolutionary complexity of nucleo-cytoplasmic large DNA viruses of eukaryotes. Virol. J. 9, 161.

Yutin, N., Wolf, Y.I., Raoult, D., Koonin, E.V., 2009. Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. Virol. J. 6, 223.

Yutin, N., Colson, P., Raoult, D., Koonin, E.V., 2013. Mimiviridae: clusters of orthologous genes, reconstruction of gene repertoire evolution and proposed expansion of the giant virus family. Virol. J. 10, 106.

Yutin, N., Wolf, Y.I., Koonin, E.V., 2014. Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. Virology 466–467, 38–52.

Yutin, N., Backstrom, D., Ettema, T.J.G., Krupovic, M., Koonin, E.V., 2018. Vast diversity of prokaryotic virus genomes encoding double jelly-roll major capsid proteins uncovered by genomic and metagenomic sequence analysis. Virol. J. 15 (1), 67.

Zauberman, N., Mutsafi, Y., Halevy, D.B., Shimoni, E., Klein, E., Xiao, C., Sun, S., Minsky, A., 2008. Distinct DNA exit and packaging portals in the virus Acanthamoeba polyphaga mimivirus. PLoS Biol. 6 (5), e114.

Zhang, X., Xiang, Y., Dunigan, D.D., Klose, T., Chipman, P.R., Van Etten, J.L., Rossmann, M.G., 2011. Three-dimensional structure and function of the Paramecium bursaria chlorella virus capsid. Proc. Natl. Acad. Sci. U. S. A. 108 (36), 14837–14842.

Zhang, Y., Cao, X., Li, D., 2015. Architecture of viral replication factories. Oncotarget 6 (31), 30439–30440.

Zhou, L., Yu, E.Y.W., Wang, S., Sun, C., 2019. African swine fever epidemic in China. Vet. Rec. 184 (23), 713.

Zhu, D., Wang, X., Fang, Q., Van Etten, J.L., Rossmann, M.G., Rao, Z., Zhang, X., 2018. Pushing the resolution limit by correcting the Ewald sphere effect in single-particle Cryo-EM reconstructions. Nat. Commun. 9 (1), 1552.