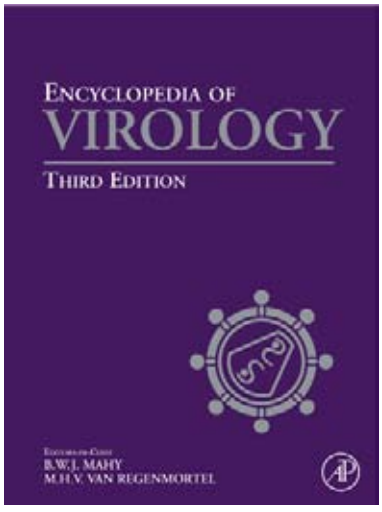

Van Etten J L and Graves M V. Phycodnaviruses. Encyclopedia of Virology, 5 vols. (B.W.J. Mahy and M.H.V. Van Regenmortel, Editors), pp. 116-125 Oxford: Elsevier.

Acknowledgments

The authors acknowledge support from the Cancer Research Institute Investigator Award (A. Grakoui), Woodruff Health Sciences Fund (A. Grakoui and E. S. Mocarski), the Georgia Cancer Coalition (E. S. Mocarski), and the USPHS (A. Grakoui and E. S. Mocarski).

See also: Apoptosis and Virus Infection.

Further Reading

Biron C and Sen G (2006) Innate immune responses to viral infection. In: Knipe DM, Howley PM Griffin DE, *et al.* (eds.) *Fields Virology,* 5th edn., pp. 249–278. Philadelphia: Lippincott Williams and Wilkins.

Braciale T, Hahn YS, and Burton D (2006) Adaptive immune responses to viral infection. In: Knipe DM, Howley PM, Griffin DE, *et al.* (eds.) *Fields Virology,* 5th edn., 279–325. Philadelphia: Lippincott Williams and Wilkins.

Virgin S (2006) Pathogenesis of viral infection. In: Knipe DM, Howley PM, Griffin DE, *et al.* (eds.) *Fields Virology,* 5th edn., pp. 2701–2772. Philadelphia: Lippincott Williams and Wilkins.

Phycodnaviruses

J L Van Etten, University of Nebraska–Lincoln, Lincoln, NE, USA
M V Graves, University of Massachusetts–Lowell, Lowell, MA, USA

History

Since the early 1970s, viruses or virus-like particles (VLPs) have been been reported in at least 44 taxa of eukaryotic algae, which include members in 10 of the 14 classes of algae. However, most of the early reports described single accounts of microscopic observations. The VLPs were not characterized because they were difficult to obtain in reasonable quantities. Several factors contributed to the low virus concentrations: (1) often only a few algal cells contained particles; (2) usually the cells only contained particles at one stage of the algal life cycle; (3) cells containing particles tended not to lyse; (4) in most cases the particles were not infectious; and (5) some hosts could not be cultured easily.

However, this situation began to change with the discovery of a family of large double-stranded DNA (dsDNA)-containing viruses that infect and replicate in certain strains of unicellular, eukaryotic, exsymbiotic, chlorella-like green algae. The first such 'chlorella viruses' were discovered in 1978 in chlorella symbiotic with *Paramecium bursaria* and in 1981 in chlorella symbiotic with the green coelenterate *Hydra viridis*. The algae from *P. bursaria* can be grown free of the paramecium in culture, and these cultured, naturally endosymbiotic *Chlorella* strains (NC64A and Pbi or their equivalents) serve as hosts for many similar viruses. The lytic chlorella viruses can be produced in large quantities and assayed by plaque formation using standard bacteriophage techniques. Recently, a plaque-forming virus that infects chlorella symbiotic with the heliozoon *Acanthocystis turfacea* was described. This virus does not infect *Chlorella* NC64A or *Chlorella* Pbi. The prototype chlorella virus is PBCV-1, which stands for *Paramecium bursaria* chlorella virus. The genomes (313–370 kbp) of several of the chloroviruses have either been sequenced or are in the process of being sequenced.

Large polyhedral, dsDNA-containing viruses that infect certain marine algae are also under active investigation. These include viruses that infect filamentous brown algae, *Ectocarpus* sp. (EsV viruses) and *Feldmannia* sp. (FsV viruses) (**Table 1**), and viruses that infect *Emiliania huxleyi* (EhV viruses). The genomes of some of these viruses have also been sequenced recently. Although all of these algal viruses arose from a common ancestor, they can have different lifestyles. For example, EsV and FsV viruses have a lysogenic phase in their life cycle and are only expressed as virus particles in sporangial cells of their host. In contrast, the chlorella viruses and EhV viruses are lytic.

The first algal viruses to be discovered were large dsDNA viruses; consequently, it was assumed for several years that algae were only infected by large dsDNA viruses. However, this scenario is changing rapidly. A positive-sense 9.1 kbp single-stranded RNA (ssRNA) virus has been discovered that infects a toxic bloom-forming alga, *Heterosigma akashiwo* (called HaRNAV) that is related to the picorna-like virus superfamily. A dsRNA, reo-like virus that infects a microalga, *Micromonas pusilla*, has been reported and finally a virus (CsNIV) with an unusual genome structure that infects diatoms in the genus *Chaetoceros* has been described. The CsNIV genome consists of a single molecule of covalently closed circular single-stranded DNA (ssDNA) (6005 nucleotides) as well as a segment of linear ssDNA (997 nucleotides). These recently discovered algal viruses are described in other articles in this encyclopedia.

Table 1 Taxonomy and general characteristics of some phycodnaviruses

Genus[a]	*Type species*[a]	*Known host range*[a]	*Source*	*Particle diameter (nm)*	*Genome size (kbp) and conformation*	*Latent period (h)*	*Burst size*
Chlorovirus	*Paramecium bursaria* chlorella virus 1 (PBCV-1)	*Chlorella* NC64A *Chlorella* Pbi *Chlorella* SAG 3.83	FW	190	313–370 Closed linear dsDNA, hairpin termini	6–8	200–350[b]
Coccolithovirus	*Emiliania huxleyi* virus 86 (EhV-86)	*Emiliania huxleyi*	MW	160–200	407–415 Circular	4–6	400–1000
Phaeovirus	*Ectocarpus siliculousus* virus 1 (EsV-1)	*Phaeophycea* *Ectocarpus siliculousus* *Ectocarpus fasciculatus* *Feldmannia simplex* *Feldmannia irregularis* *Feldmannia species* *Hincksia hincksiae* *Myriotrichia clavaeformis* *Pilayella littoralis*	MW	150–170	170–340 Open linear, single stranded regions	ND	$>1\times10^6$
Prasinovirus	*Micromonas pusilla* virus SP1 (MpV-SP1)	*Micromonas pusilla*	MW	115	200	7–14	1800–4100
Prymnesiovirus	*Chrysochomulina brevifilum* virus PW1 (CbV-PW1)	*Haptophyceae (aka Prymnesiophyceae)* *Chrysochomulina brevifilum* *Chrysochomulina strobilus* *Chrysochomulina globosa*	MW	145–170	510	ND	800–1000
Raphidovirus	*Heterosigma akashiwo* virus 01 (HaV01)	*Heterosigma akashiwo*	MW	202	294	30–33	770

[a]Data abstracted from http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/index.htm.
[b]PFU/cell.
FW, freshwater; MW, marine/coastal water; ND, not determined.
Reproduced from Dunigan DD, Fitzgerald LA, and Van Etten JL (2006) Phycodnaviruses: A peek at genetic diversity. *Virus Research* 117: 119–132, with permission from Elsevier.

Taxonomy and Classification

Members and prospective members of the family *Phycodnaviridae* constitute a genetically diverse, but morphologically similar group of viruses with eukaryotic algal hosts from both fresh and marine waters. Accumulating genetic evidence indicates that the phycodnaviruses together with the poxviruses, iridoviruses, asfarviruses, and the 1.2 Mbp mimivirus have a common evolutionary ancestor, perhaps, arising at the point of eukaryogenesis, variously reported to be 2.0–2.7 billion years ago. All of these viruses share nine gene products and 33 more gene products are present in members of at least two of these five viral families. Collectively, these viruses are referred to as nucleocytoplasmic large DNA viruses (NCLDV).

Phycodnaviruses are large (mean diameter 160 ± 60 nm) icosahedrons, which encapsidate 160–560 kbp dsDNA genomes. Where known, the viruses have an internal-membrane that is required for infection. Phylogenetic analyses of their δ-DNA polymerases indicate that they are more closely related to each other than to other dsDNA viruses and that they form a monophyletic group, consistent with a common ancestor. However, the phycodnaviruses fall into six clades which correlate with their hosts and each has been given genus status. Often the genera can be distinguished by additional properties, for example, lytic versus lysogenic lifestyles or linear versus circular genomes. Members of the genus *Chlorovirus* infect freshwater algae, whereas, members of the other five genera (*Coccolithovirus*, *Phaeovirus*, *Prasinovirus*, *Prymnesiovirus*, and *Raphidovirus*) infect marine algae.

Structure and Composition

Chlorella virus particles are large (molecular weight $\sim 1 \times 10^9$ Da) and complex. The PBCV-1 virion contains more than 100 different virus-encoded proteins. The PBCV-1 54 kDa major capsid protein is a glycoprotein and comprises ~40% of the total protein. The major capsid protein consists of two eight-stranded, antiparallel β-barrel, jelly-roll domains related by pseudo sixfold rotation. This structure resembles the major coat proteins from some other dsDNA viruses that infect all three domains of life including bacteriophage PRD1, human adenoviruses, and a virus STIV infecting the Archaea, *Sulfolobus solfataricus*. This finding led to the suggestion that these three viruses may also have a common evolutionary ancestor with the NCLDVs, even though there is no significant amino acid sequence similarity among their major capsid proteins.

Cryoelectron microscopy and three-dimensional image reconstruction of the PBCV-1 virion (**Figure 1**) indicate that the outer capsid is icosahedral and covers a lipid bilayered membrane. The membrane is required for infection because the virus loses infectivity after exposure to organic solvents. The outer diameter of the virus capsid ranges from 1650 Å along the two- and threefold axes to 1900 Å along the fivefold axis. The capsid shell consists of 1680 doughnut-shaped trimeric capsomers plus 12 pentameric capsomers at each icosahedral vertex. The trimeric capsomers are arranged into 20 triangular facets (trisymmetrons, each containing 66 trimers) and 12 pentagonal facets (pentasymmetrons, each containing 30 trimers and one pentamer at the icosahedral vertices). Assuming all the trimeric capsomers are identical, the outer capsid of the virus contains 5040 copies of the major capsid protein. The virus has a triangulation number of 169. However, PBCV-1 is not the largest phycodnavirus; phaeocystis pouchetti virus (PpV01) has an icosahedral capsid with a triangulation number of 219.

Structural proteins of many viruses, such as herpesviruses, poxviruses, and paramyxoviruses, as well as the chlorella viruses, are glycosylated. Typically, viral proteins are glycosylated by host-encoded glycosyltransferases located in the endoplasmic reticulum (ER) and Golgi and then transported to a host membrane. Nascent viruses acquire the glycoprotein(s) and only become infectious by budding through the membrane, usually as

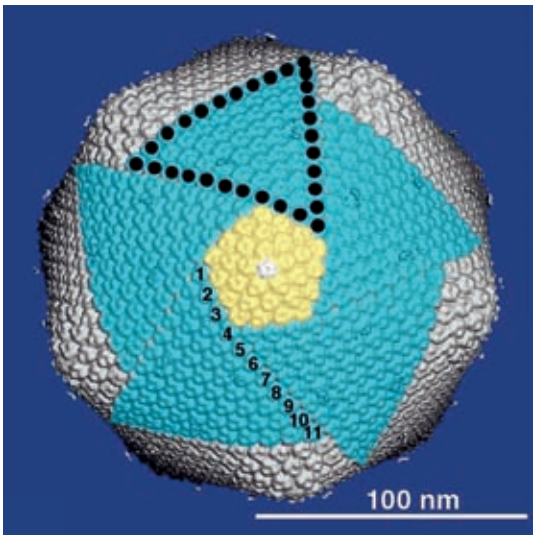

Figure 1 Three-dimensional image reconstruction of chlorella virus PBCV-1 from cryoelectron micrographs. The virion capsid consists of 12 pentasymmetrons and 20 trisymmetrons. Five trisymmetrons are highlighted in the reconstruction (blue) and a single pentasymmetron is colored yellow. A pentavalent capsomer (white) lies at the center of each pentasymmetron. Each pentasymmetron consists of one pentamer plus 30 trimers. Eleven capsomers form the edge of each trisymmetron (black dots) and therefore each trisymmetron has 66 trimers. Reprinted, with permission, from Van Etten JL (2003) Unusual life style of giant chlorella viruses. *Annual Review of Genetics* 37: 153–195, ©2003 by Annual Reviews.

they are released from the cell. Consequently, the glycan portion of virus glycoproteins is host specific.

However, glycosylation of PBCV-1 major capsid protein differs from this paradigm. Accumulating evidence indicates that PBCV-1 encodes most, if not all, of the enzymes involved in constructing the complex oligosaccharides attached to its major capsid protein and that the process occurs independently of the ER and Golgi. Furthermore, five of six putative PBCV-1-encoded glycosyltransferases are predicted to be located in the cytoplasm. PBCV-1 also encodes several additional proteins involved in post-translational modification that may alter virus structural proteins. These include a prolyl-4-hydroxylase and several protein kinases and a phosphatase.

Genomes

The 331 kbp PBCV-1 genome is linear and nonpermuted. The genome termini consist of 35 nucleotide long, incompletely base-paired, covalently closed hairpin loops that exist in one of two forms (flip and flop). Each hairpin loop is followed by an identical 2.2 kbp inverted repeat sequence; the remainder of the genome consists primarily of single-copy DNA. The PBCV-1 genome has ~695 open reading frames (ORFs) that have 65 or more codons, of which ~366 are probably protein encoding. The putative protein-encoding genes are evenly distributed on both strands and intergenic space is minimal, 275 ORFs are separated by less than 100 nucleotides (**Figure 2**). One exception is a 1788 bp sequence near the middle of the genome. This sequence has a polycistronic gene containing 11 tRNAs (**Figure 2**).

Approximately 40% of the 366 PBCV-1 gene products resemble proteins in the databases, many of which have not previously been associated with viruses (some are listed in **Figure 3**). Eighty-four ORFs have paralogs within PBCV-1, forming 26 groups. The size of these groups ranges from two to six members.

Some PBCV-1 genes are interrupted by introns; a gene encoding a transcription factor-like protein contains a self-splicing type I intron, whereas the δ-DNA polymerase gene contains a spliceosomal-processed type of intron. In addition, one of the PBCV-1 tRNA genes is predicted to have an intron.

One unusual feature of PBCV-1 DNA, as well as the other chlorella virus DNAs, is that they contain methylated bases. Chlorella virus genomes contain 5-methylcytosine in amounts varying from 0.1% to 48% of the total cytosines. Many viral DNAs also contain N^6-methyladenine with concentrations up to 37% of the total adenines. This led to the discovery that many chlorella viruses encode multiple DNA methyltransferases, as well as site-specific (restriction) endonucleases.

The 407 kbp circular dsDNA genome of one of the EhV viruses is predicted to have 472 protein-encoding genes. Only 66 (14%) of these 472 gene products match a sequence in GenBank. The EhV virus encodes several unexpected genes never found in a virus before, including four gene products involved in sphingolipid biosynthesis and two additional gene products which encode desaturases.

The structure of the 335 kbp EsV genome is unknown. Although several experiments suggest that the genome is circular, DNA sequencing indicates that it has defined ends with inverted repeats. The virus is predicted to contain 231 protein-encoding genes, 48% of which resemble various proteins in the public databases. About 12% of the EsV genome consists of tandem repeats and portions of the genome have ssDNA regions. Collectively, PBCV-1, EsV, and EhV viruses have in excess of 1000 unique ORFs. A total of 123 putative ORFs from these three viruses is organized into metabolic domains (**Figure 3**). Interestingly, these three viruses only have 14 gene products in common. Not surprisingly, several of these common ORFs are involved in DNA replication, such as the δ-DNA polymerase, large and small subunits of ribonucleotide reductase, proliferating cell nuclear antigen (PCNA), superfamily II and III helicases, and the newly recognized archaeo-eukaryotic primases. Another common ORF among the three phycodnaviruses is the major capsid protein.

Virus Replication

PBCV-1 attaches rapidly, specifically, and irreversibly to the external surface of cell walls, but not to protoplasts, of host *Chlorella* NC64A. Attachment always occurs at a virus vertex, followed by degradation of the wall at the attachment point. Following wall degradation, the internal membrane of the virus probably fuses with the host membrane resulting in entry of the virus DNA and virion-associated proteins into the cell, leaving an empty capsid on the surface. Several observations suggest that the infecting DNA, plus associated proteins, is rapidly transported to the host cell nucleus and the host transcription machinery is reprogrammed to transcribe virus RNAs. This process occurs rapidly because early PBCV-1 transcripts can be detected within 5–10 min post infection (p.i.). PBCV-1 translation occurs on cytoplasmic ribosomes and early PBCV-1-encoded proteins can be detected within 10 min p.i.

PBCV-1 DNA synthesis and late virus transcription begins 60–90 min p.i. Ultrastructural studies of PBCV-1-infected chlorella suggest that the nuclear membrane remains intact, at least during early stages of virus replication. At approximately 2–3 h p.i., assembly of virus

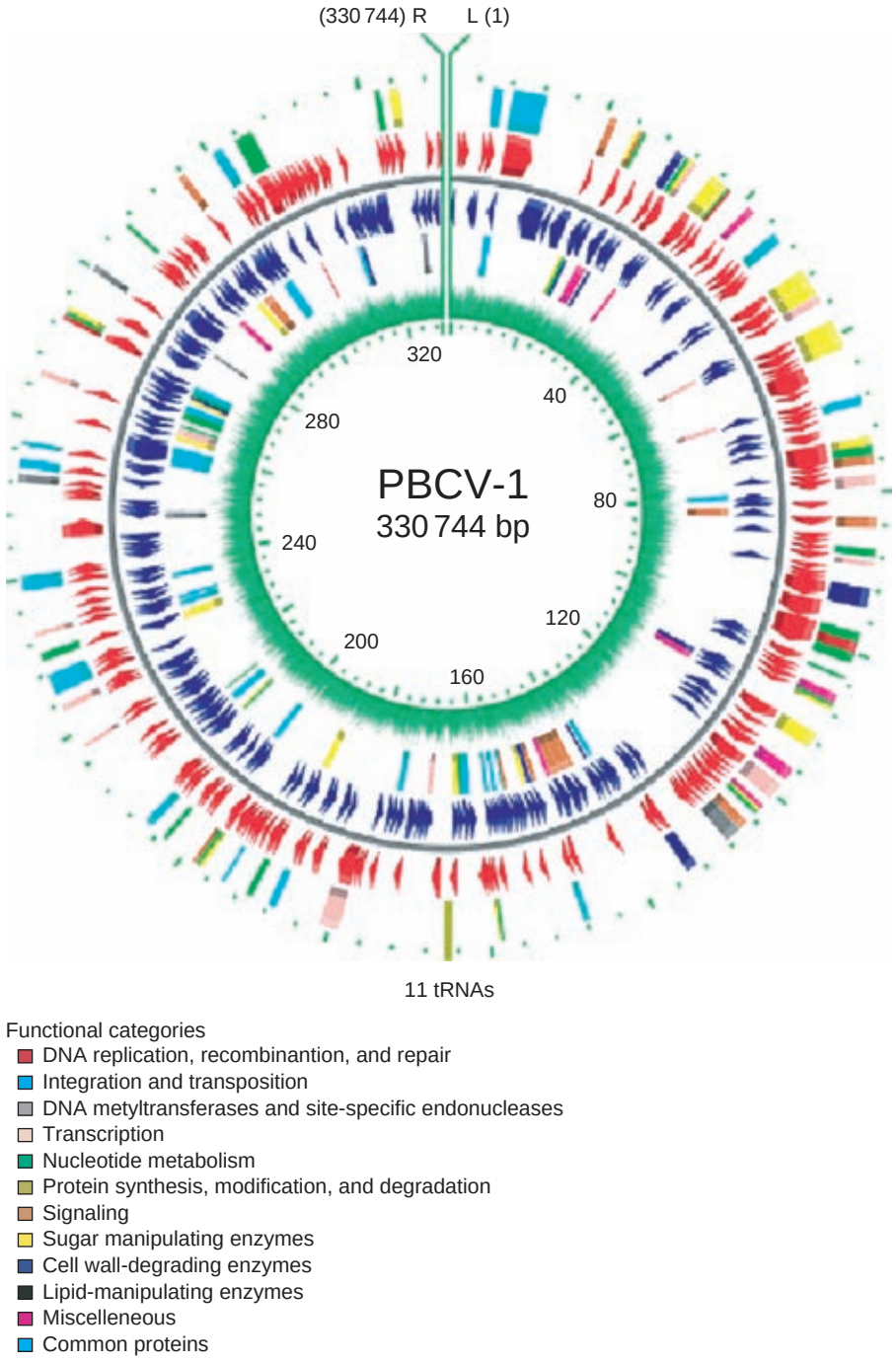

Figure 2 Map of the chlorella virus PBCV-1 genome visualized as a circle using Circular Genome Viewer (Paul Stothard, Genome Canada). However, the genome is a linear molecule and the ends are depicted at the top of the figure as green lines (L and R represent the left and right ends of the genome, respectively). The red and blue arrows represent the 366 protein-encoding genes; red arrows depict genes transcribed in the rightwards direction and blue arrows genes transcribed in the leftwards direction. The two rings that flank the protein-encoding genes show the predicted functions of the proteins, color-coded by function (see insert in the figure). The location of the polycistronic gene encoding the 11 tRNAs is indicated in the outermost ring (i.e., this gene is transcribed in the rightwards direction). The innermost ring (in green) represents the A + T content determined using a 25 bp window. Note that the A + T content is fairly constant over the genome (60% A + T).

capsids begins in localized regions in the cytoplasm, called virus assembly centers, which become prominent at 3–4 h p.i.. By 5 h p.i., the cytoplasm is filled with infectious progeny virus particles (~1000 particles/cell) and by 6–8 h p.i. localized lysis of the host cell releases progeny. Of the progeny released, 25–50% of the particles are infectious. Mechanical disruption of cells releases infectious virus 30–50 min prior to cell lysis, indicating the

DNA replication, recombination, and repair			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
Archaeo-eukaryotic primase			
ATPase[a]	4	2	
DNA ligase	*		
DNA topoisomerase II	*		
DNA-binding protein	#		
Exonuclease[b]	*	2	
Helicase			2
Helicase-Rec BCD-like			
Helicase-Superfamily III			
Nucleic acid-binding protein			
PCNA[c]	2		2
Pyrimidine dimer-specific glycosylase	*		
Replication factor C (lg subunit)			
Replication factor C (sm subunit)		4	
RNaseH			
δ-DNA polymerase			

Protein synthesis, modification, and degradation			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
ATPase (AAA+ class)			
ATP-dependent protease proteolytic subunit			
Esterase			
Hydrolase			
Prolyl 4-hydroxylase	*#		
Protease-Cysteine			
Protease-OTU like cysteine			2
Protease-Serine			5
Protein disulfide isomerase	#		
SKP1 protein			
Thiol oxidoreductase	#		
Translation elongation factor-3			
Ubiquitin C-terminal hydrolase			
Zn metallopeptidase			
tRNAs[d]	11		5

Integration and transposition			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
Antirepressor of lysogeny			
Endonuclease			3
Homing endonuclease GIY-YIG	7		
Homing endonuclease HNH[e]	6	2	
Integrase			
Protelomerase			
Repressor of lysogeny			
Transposase		2	

Sugar and lipid manipulation			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
Alginate mannuronan epimerase			
D-lactate dehydrogenase			
Fatty acid desaturase			2
Fructose-2, 6 bisphosphatase			
Fucose synthetase	*		
GDP-D-mannose dehydratase	*		
Glucosamine synthetase	*		
Glycerophosphoryl diesterase			
Glycosyltransferase	6		
Hyaluronan synthase	*		
Lipase			
Lipid phosphate phosphatase			
Lysophospholipase			
N-acetyltransferase			
Patatin phospholipase			
Salidase			
Serine palmitoyltransferase			
Sterol desaturase			
Sugar lyase			
Transmembrane fatty acid elongation protein			
UDP-glucose dehydrogenase	*		

Signaling			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
Dual specificity phosphatase	#		
Glutamate receptor			
Hybrid histidine kinase		6	
Ligand-gated channel protein			
Phosphoshuttle			
Potassium channel protein	*#		
Ser/Thr protein kinase[f]	8*#	4	2

DNA restriction/modification			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
Adenine DNA methylase	2*		
Cytosine DNA methylase	3*		
DNA restriction endonuclease	2*#		

Transcription			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
DNA-dependent RNA polymerase II largest			
DNA-directed RNA polymerase II subunit			
DNA-directed RNA polymerase subunit			3
Histone H3, Lys 27 methylase	*#		
mRNA guanylyltransferase	*		
Oligoribonuclease			
RNA triphosphatase	*		
RNase III	*		
Superfamily II helicase[g]	3#		
SWI/SNF helicase	#		
Transcription factor TFIIB			
Transcription factor TFIID			
Transcription factor TFIIS			
VLTF2-type transcription factor			

Cell wall degradation			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
Chitinase	2*#		
Chitosanase	*#		
β and α 1, 4-linked glucuronic lyase	*		
β-1, 3-glucanase	*		

Nucleotide metabolism			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
Aspartate transcarbamylase	*		
Cytosine deaminase			
dCMP deaminase	*		
Deoxynucleoside kinase			
dUTP pyrophosphatase	*		
Glutaredoxin	*		
NTP pyrophosphohydrolase			
Nucleic acid-independent nucleoside			
Ribo. reductase (large subunit)			
Ribo. reductase (small subunit)			
Ribonuclease			
Thermonuclease			
Thioredoxin	#		2
Thymidylate kinase			
Thymidylate synthase X	*		
Thymidylate synthase-bifunct. dihydrofolate			

Miscellaneous			
Description	**PBCV-1**	**EsV-1**	**EhV-86**
ABC transporter protein	#		
Agmatine iminohydrolase	*		
Amidase	*#		
Cu/Zn-superoxi dedismutase	*#		
Fibronectin binding protein			
Histidine decarboxylase			
Homospermidine synthase	*#		
Monoamine oxidase	#		
N-carbamoylput amidohydrolase	*		
O-methyltransferase			
Ornithine decarboxylase	*		
Pathogenesis-related protein			
Calcium-binding protein			
Collagen-like protein			
Thaumatin-like protein			
Lectin protein			2
Longevity-assurance family protein			
Major facilitator			
Phosphate permease			
Phosphoglycerate mutase			

Present in all 3 viruses	
PBCV-1 + EsV-1	
PBCV-1 + EhV-86	
EsV-1 + EhV-86	
Unique to one virus	
Functional Enzyme	*
Virion Associated	#

Figure 3 Continued

virus does not acquire its glycoprotein capsid by budding through the host plasma membrane as it is released from the cell. Other chlorella viruses have longer replication cycles than PBCV-1. For example, virus NY-2A requires approximately 18 h for replication and consequently forms smaller plaques.

Virus EsV initiates its life cycle by infecting free-swimming, wallless gametes of its host. Virus particles enter the cell by fusion with the host plasma membrane and release a nucleoprotein core particle into the cytoplasm, leaving remnants of the capsid on the surface. The viral core moves to the nucleus within 5 min p.i. One important feature that distinguishes the EsV life cycle from the other phycodnaviruses is that the viral DNA is integrated into the host genome and is transmitted mitotically to all cells of the developing alga. The viral genome remains latent in vegetative cells until it is expressed in the algal reproductive cells, the sporangia or gametangia. Massive viral DNA replication occurs in the nuclei of these reproductive cells, followed by nuclear breakdown and viral assembly that continues until the cell becomes densely packed with virus particles. Virus release is stimulated by the same factors that induce discharge of gametes from the host, that is, changes in temperature, light, and water composition. This synchronization facilitates interaction of viruses with their susceptible host cells.

In contrast to the chlorella and EsV viruses, very little is known about the EhV life cycle, including how it infects its host. One property that clearly distinguishes EhV from the other phycodnaviruses is that it encodes six RNA polymerase subunits; in contrast, neither PBCV-1 nor EsV encodes a recognizable RNA polymerase component. Thus like the poxviruses, EhV may carry out its entire life cycle in the cytoplasm of its host. The latent period for EhV is 4–6 h with a burst size of 400–1000 particles per cell.

Virus Transcription

Detailed studies on transcription are lacking for all of the algal viruses and only a few general statements can be made about PBCV-1 transcription:

1. PBCV-1 infection rapidly inhibits host RNA synthesis.
2. Viral transcription is programmed and early transcripts appear within 5–10 min p.i. Late viral transcription first begins about 60–90 min p.i.
3. Some early viral transcripts are synthesized in the absence of *de novo* protein synthesis. As expected, the synthesis of later transcripts requires translation of early virus genes.
4. Early and late genes are dispersed in the PBCV-1 genome.
5. PBCV-1 ORFs are tightly packed on both DNA strands and the coding regions of some of the genes overlap. The largest distance between PBCV-1 ORFs is a 1788-nucleotide stretch in the middle of the genome. This sequence has a polycistronic gene containing 11 tRNAs (**Figure 2**).
6. Consensus promoter regions for early and late genes have not been identified, although the 50 nucleotides

Figure 3 Selected ORFs in the PBCV-1, EsV, and EhV genomes are arranged by their metabolic domains. If a genome encodes a putative protein more than once, a number in the box indicates the number of genes of this type per genome. Color-coding is indicated on the figure and is used to depict the relationship between viruses. Red indicates proteins that are encoded by all three viruses; yellow indicates proteins that are encoded by PBCV-1 and EsV, but not EhV; green indicates proteins that are encoded by PBCV-1 and EhV, but not EsV; orange indicates proteins that are encoded by EsV and EhV, but not PBCV-1; and blue indicates there are no shared homologs. Solid colored boxes indicate that the putative proteins are homologs. A diagonally stripped box indicates that the putative proteins are nonhomologous, and a checkered box indicates that the putative proteins are a mix of homologous, nonhomologous, or unique ORFs. In this case, a footnote has been added to clarify the specific differences; in parentheses the ORF has been defined by the gene number and any ORF beginning with an 'A' is from the PBCV-1 genome, an 'EsV' is from the EsV-1 genome, and an 'EhV' is from the EhV-86 genome. Proteins known to be functional are indicated with a star (*) and proteins known to be associated with the virion are indicated with a pound sign (#). Superscript 'a' indicates ATPase – One homolog between all three viruses (A392R, EsV-26, and EhV072), one homolog between PBCV-1 and EsV (A565R and EsV-171), and two PBCV-1 ATPases which have no homologs in EsV or EhV (A561L and A554/556/557L). Superscript 'b' indicates Exonuclease – One homolog between PBCV-1 and EsV (A166R and EsV-64) and one unique to EsV (EsV-126). Supercript 'c' indicates PCNA – One homolog between all three viruses (A193L, EsV-132, and EhV020), one homolog between PBCV-1 and EhV (A574L and EhV020). However, EhV-440, another EhV-encoded PCNA, has no homologs in PBCV-1 or EsV. Superscript 'd' indicates tRNAs – PBCV-1 has 11 tRNA genes, encoding AA Leu (2), Ile, Asn (2), Lys (3), Tyr, Arg, and Val. EhV has five tRNA genes, encoding AA Leu, Ile, Gln, Asn, and Arg. Supercript 'e' indicates Homing endonuclease HNH – One homolog between two viruses (A422R and EhV087). PBCV-1 (A87R) and EsV (EsV-119) are homologous. Four other HNH endonucleases are unique to PBCV-1 (A267L, A354R, A478R, and A490L) and one is unique to EsV (EsV-1–16). Superscript 'f' indicates Ser/Thr protein kinase – Four PBCV-1, two EsV-1, and one EhV-86 S/T kinases ORFs are homologous (A248R, A277L, A282L, A289L, EsV-82, EsV-111, and EhV451). The remaining S/T kinase ORFs from the three genomes (A34R, A278L, A614L, A617R, EsV-104, EV-156, and EhV-402) are unique. Suerscript 'g' indicates that superfamily II helicase, PBCV-1 (A153R), and EsV (EsV-66) are homologous. The two additional PBCV-1-encoded helicases (A241R and A363R) and EhV104 are unique. Reproduced from Dunigan DD, Fitzgerald LA, and Van Etten JL (2006) Phycodnaviruses: A peek at genetic diversity. *Virus Research* 117: 119–132, with permission from Elsevier.

preceding the ATG start codon of most functional PBCV-1 genes are at least 70% A+T.

7. Transcription of some PBCV-1 genes appears to be complex. For example, some gene transcripts exist as multiple bands and these patterns change between early and late times in the virus life cycle.

Additional *Chlorella* Viruses

Several hundred plaque-forming chloroviruses have been characterized to various degrees. They infect either *Chlorella* NC64A cells (NC64A viruses), an endosymbiont of *P. bursaria* isolated from North America, or *Chlorella* Pbi cells (Pbi viruses) that are endosymbiotic with a paramecium isolated in Europe. Like PBCV-1, each of these viruses contain many structural proteins, a large (>300 kbp) dsDNA genome, and they are chloroform sensitive. The DNAs of some of these viruses hybridize strongly with PBCV-l DNA, while others hybridize poorly.

Three additional chlorella virus genomes have been sequenced recently and others are nearing completion. The largest, the 370 kbp genome of virus NY-2A, contains ~400 protein-encoding genes. Most common genes are colinear in viruses PBCV-1 and NY-2A, which infect the same host chlorella. However, almost no colinearity exists between common genes in Pbi virus MT325 and those in PBCV-1, NY-2A, and AR158 suggesting plasticity in the chlorella virus genomes. Additionally, the G + C contents of the three NC64A viruses range from 40% to 41% whereas the G + C contents of the Pbi viruses are approximately 45%. These last two observations suggest that these two virus groups have been separated for considerable evolutionary time. Viruses morphologically similar to the NC64A and Pbi viruses have also been isolated from chlorella symbiotic in the coelenterate *Hydra viridis* and very recently from the heliozoon *Acanthocystis turfacea*. The symbiotic hydra chlorella have not been cultured and so very little is known about these viruses. However, the *A. turfacea* viruses can be isolated by plaque formation.

Other Algal Viruses

Field isolates representing at least six genera of filamentous brown algae contain virus particles that are morphologically similar to EsV and FsV. Virus expression is variable; virions are rarely observed in vegetative cells but often are common in unilocular sporangia (FsV) or both unilocular and plurilocular sporangia (EsV). EsV viruses only infect the free swimming, zoospore stage of *Ectacarpus* sp. All natural isolates of *Feldmania* sp. are infected with virus and so infection studies cannot be conducted.

Viruses that infect *Emiliania huxleyi*, *Micromonas pusilla*, and *Chrysochromulina brevfilum* have been isolated from many marine environments. These viruses can be distinguished by DNA restriction patterns.

Phycodnavirus Genes Encode Some Interesting and Unexpected Proteins

Many chlorella virus-encoded enzymes are either the smallest or among the smallest proteins of their class. In addition, homologous genes in the chloroviruses can differ in nucleotide sequence by as much as 50%, which translates into amino acid differences of 30–40%. Therefore, comparative protein sequence analyses can identify conserved amino acids in proteins as well as regions that tolerate amino acid changes. The small sizes and the finding that many chlorella virus-encoded proteins are 'user friendly' have resulted in the biochemical and structural characterization of several PBCV-1 enzymes. Examples include: (1) The smallest eukaryotic ATP-dependent DNA ligase, which is the subject of intensive mechanistic and structural studies. (2) The smallest type II DNA topoisomerase. The virus enzyme cleaves dsDNAs about 50 times faster than the human type II DNA topoisomerase; consequently, the virus enzyme is being used as a model enzyme to study the mechanism of topoisomerase II DNA cleavage. (3) An RNA guanylyltransferase that was the first enzyme of its type to have its crystal structure resolved. (4) A small prolyl-4-hydroxylase that converts Pro-containing peptides into hydroxyl-Pro-containing peptides in a sequence-specific fashion. (5) The smallest protein (94 amino acids) to form a functional potassium ion channel. These minimalist enzymes may represent evolutionary precursors of contemporary proteins, but it is also possible that they are products of evolutionary optimization during viral evolution.

The chloroviruses are also unusual because they encode enzymes involved in sugar metabolism. Three PBCV-1 encoded enzymes, glutamine:fructose-6-phosphate amindotransferase, UDP-glucose dehydrogenase, and hyaluronan synthase, are involved in the synthesis of hyaluronan, a linear polysaccharide composed of alternating β-1,4-glucuronic acid and β-1,3-*N*-acetylglucosamine residues. All three genes are transcribed early in PBCV-1 infection and hyaluronan accumulates on the external surface of the infected cells.

Two PBCV-1-encoded enzymes, GDP-D-mannose dehydratase and fucose synthase, comprise a three-step pathway that converts GDP-D-mannose to GDP-L-fucose. The function of this putative pathway is unknown. However, fucose, a rare sugar, is present in the glycans attached to the major capsid protein.

PBCV-1 encodes four enzymes involved in polyamine biosynthesis: ornithine decarboxylase (ODC), homospermidine synthase, agmatine iminohydrolase, and *N*-carbamoyl-putrescine amidohydrolase. ODC catalyzes the decarboxylation of ornithine to putrescine, which is the first and the rate-limiting enzymatic step in polyamine biosynthesis. Not only is the PBCV-1-encoded ODC the smallest known ODC, the PBCV-1 enzyme is also interesting because it decarboxylates arginine more efficiently than ornithine.

The genome sequences of other phycodnaviruses have revealed several interesting and unexpected gene products. However, except for an aquaglyceroporin encoded by chlorella virus MT325, none of these products have been expressed and tested for enzyme function.

Ecology

Eukaryotic algae are important components of both freshwater and marine environments; however, the significance of viruses in these systems is only beginning to be appreciated. The chlorella viruses are ubiquitous in freshwater collected throughout the world and titers as high as 100 000 infectious particles per mililiter have been reported in native waters. Typically, the titer is 1–100 infectious particles per milliliter. The titers are seasonal with the highest titers in the spring. It is not known whether chlorella viruses replicate exclusively in algae symbiotic with paramecia or if the viruses have another host(s). In fact, it is not known if paramecium chlorellae exist free of their hosts in natural environments. However, the chlorellae are protected from virus infection when they are in a symbiotic relationship with the paramecium.

The concept that viruses might have a major impact on the marine environment began about 15 years ago with the discovery that seawater contains as many as 10^7 VLPs per milliliter. This huge population consits of both bacterial and algal viruses and is important because phytoplankton, consisting of cyanobacteria and eukaryotic microalgae, fix 50–60% of the CO_2 on Earth. At any one time, 20–40% of these photosynthetic organisms are infected with a virus. Consequently, these viruses contribute to microbial composition and diversity, as well as, nutrient cycling in aqueous environments. Thus viruses, including the phycodnaviruses, have a major impact on global carbon/nitrogen cycles that is only beginning to be appreciated by scientists, including those who model such cycles.

Algal viruses are also believed to play major roles in the termination of marine algal blooms and so there are active research efforts to understand the natural history of these algal/virus systems. For example, the coccolithophorid *Emiliania huxleyi* is a unicellular alga found throughout the world; the alga can form immense coastal and mid-oceanic blooms at temperate latitudes that cover 10 000 km^2 or more. One of the primary mechanisms for terminating *E. huxleyi* blooms is lysis by the EhV viruses described above.

The filamentous brown algae, *Ectocarpus* sp. and *Feldmania* sp. isolated from around the world, are infected with lysogenic EsV and FsV viruses, respectively. Lysogeny is consistent with the observation by early investigators that VLPs appear infrequently in eukaryotic algae and only at certain stages of algal development. The apparent lack of infectivity by many of the previously observed VLPs in eukaryotic algae is also consistent with a lysogenic lifestyle. The VLPs might either infect the host and resume a lysogenic relationship or be excluded by preexisting lysogenic viruses.

Perspectives

Sequence analyses of three phycodnaviruses suggest that this family may have more sequence diversity than any other virus family. These three viruses only have 14 homologous genes, which means that there are in excess of 1000 unique ORFs in just these three viruses. Despite the large genetic diversity in these three sequenced phycodnaviruses, phylogenetic analyses of δ-DNA polymerases and DNA primases indicate that the phycodnaviruses group into a monophyletic clade within the NCLDVs. A recent analysis using eight concatenated core NCLDV genes also indicates that the phycodnaviruses cluster together and are members of the NCLDV 'superfamily'. However, it is obvious that the identification and characterization of phycodnaviruses is in its infancy. Metagenomic studies, such as DNA sequences from the Sargasso Sea samples, indicate that translation products of many unknown sequences are more similar to PBCV-1 proteins than the next known phycodnavirus.

See also: African Swine Fever Virus; Algal Viruses; Iridoviruses of Invertebrates; Marnaviruses; Poxviruses.

Further Reading

Brussaard CPD (2004) Viral control of phytoplankton populations – A review. *Journal of Eukaryotic Microbiology* 51: 125–138.

Dunigan DD, Fitzgerald LA, and Van Etten JL (2006) Phycodnaviruses: A peek at genetic diversity. *Virus Research* 117: 119–132.

Iyer ML, Balaji S, Koonin EV, and Aravind L (2006) Evolutionary genomics of nucleocytoplasmic large DNA viruses. *Virus Research* 117: 156–184.

Muller DG, Kapp M, and Knippers R (1998) Viruses in marine brown algae. *Advances in Virus Research* 50: 49–67.

Van Etten JL (2003) Unusual life style of giant chlorella viruses. *Annual Review of Genetics* 37: 153–195.

Van Etten JL, Graves MV, Muller DG, Boland W, and Delaroque N (2002) Phycodnaviridae – large DNA algal viruses. *Archives of Virology* 147: 1479–1516.

Wilson WH, Schroeder DC, Allen MJ, *et al.* (2005) Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. *Science* 309: 1090–1092.
Yamada T, Onimatsu H, and Van Etten JL (2006) Chlorella viruses. *Advances in Virus Research* 66: 293–336.

Relevant Websites

http://www.giantvirus.org – Giantvirus.org.
http://greengene.uml.edu – Greengene at UML.

Phylogeny of Viruses

A E Gorbalenya, Leiden University Medical Center, Leiden, The Netherlands

Introduction: Evolution, Phylogeny, and Viruses

Biological species, including viruses, change through generations and over time in the process known as evolution. These changes are first fixed in the genome of successful individuals that give rise to genetic lineages. Due to either limited fidelity of the replication apparatus copying the genome or physico-chemical activity of the environment, nucleotides may be changed, inserted, or deleted. Genomes of other origin may also be a source of innovation for a genome through the use of specially evolved mechanisms of genetic exchange (recombination). Accepted changes, known as mutations, may be neutral, advantageous, or deleterious, and depending on the population size and environment, the mutant lineage may proliferate or go extinct. Overall, advantageous mutations and large population size increase the chances for a lineage to succeed. The lineage fit is constantly reassessed in the ever-changing environment and lineages that, due to mutation, became a success in the past could be unfit in the new environment. Due to the growing number of mutations accumulating in the genomes, lineages diverge over time, although occasionally, due to stochastic reasons or under similar selection pressure, they may converge.

The relationship between biological lineages related by common descent is called phylogeny; the same term also embodies the methodology of reconstructing these relationships. Phylogeny deals with past events and, therefore, it is reconstructed by quantification of differences accumulated between lineages. Due to the lack of fossils and (relatively) high mutation rate, viruses were not considered to provide a recoverable part of phylogeny until the advent of molecular data proved otherwise. Comparison of nucleotide and amino acid sequences, and, occasionally, other quantitative characteristics such as distances between three-dimensional structures of biopolymers, have been used to reconstruct virus phylogeny. Results of phylogenetic analysis are commonly depicted in the form of a tree that may be used as a synonym for phylogeny. For instance, all-inclusive phylogeny of cellular species is depicted as the Tree of Life (ToL).

With few exceptions, virus phylogeny follows the theory and practice developed for phylogeny of cellular life forms. For inferring phylogeny, differences between the sequences of species members, assumed to be of a discernable common origin, are analyzed. If species in all lineages evolve at a uniform constant rate, like clock ticks, their evolution conforms to a molecular clock model. The utility of this model in relation to viruses may be very limited. Rather, related virus lineages may evolve at different and fluctuating rates and some sites may mutate repeatedly with each new mutation erasing a record about the prior change. As a result, the accumulation of inter-species differences may progress nonlinearly with the time elapsed. At present, our understanding of these parameters of virus evolution is poor and this limits our ability to assess the fit between a reconstructed phylogeny and the true phylogeny, with the latter practically remaining unknown for most virus isolates. This gap in our knowledge does not eliminate the conceptual strength of phylogentic analysis for reconstructing the relationships between biological species.

The ultimate goal of virus phylogeny is reconstructing the relationships between 'all' virus isolates and species. In contrast to cellular species, which form three compact domains (kingdoms) and whose origin is traced back to a common ancestor in the ToL, major virus classes may combine species that have originated from different ancestors. Thus, reconstructing the comprehensive virus phylogeny requires comparisons that involve genomes of virus and cellular origins. This formidable task remains largely 'work in progress'. In fact, most efforts in virus phylogeny are invested in reconstructing the relationships at the micro, rather than grand, scale and they focus on well-sampled lineages that have practical (e.g., medical) relevance. Phylogeny itself or in combination with other data may provide a deep insight into virus evolution and diverse aspects of virus life cycles, including virus interactions with their hosts.

Our knowledge about contemporary virus diversity has been steadily advancing with new viruses being