

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Papers in Plant Pathology

Plant Pathology Department

---

1999

## Giant Viruses Infecting Algae

James L. Van Etten

*University of Nebraska-Lincoln*, [jvanetten1@unl.edu](mailto:jvanetten1@unl.edu)

Russel Meints

*Oregon State University*

Follow this and additional works at: <https://digitalcommons.unl.edu/plantpathpapers>



Part of the [Plant Pathology Commons](#)

---

Van Etten, James L. and Meints, Russel, "Giant Viruses Infecting Algae" (1999). *Papers in Plant Pathology*. 129.

<https://digitalcommons.unl.edu/plantpathpapers/129>

This Article is brought to you for free and open access by the Plant Pathology Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Plant Pathology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Giant Viruses Infecting Algae

James L. Van Etten<sup>1</sup> and Russel H. Meints<sup>2</sup>

<sup>1</sup>Department of Plant Pathology, University of Nebraska–Lincoln,  
Lincoln, Nebraska 68583-0722; email: [jvanetten@unlnotes.unl.edu](mailto:jvanetten@unlnotes.unl.edu)

<sup>2</sup>Department of Botany and Plant Pathology, Center for Gene Research and  
Biotechnology, Oregon State University, Corvallis, Oregon 97331-2906

## Abstract

*Paramecium bursaria* chlorella virus (PBCV-1) is the prototype of a family of large, icosahedral, plaque-forming, double-stranded-DNA-containing viruses that replicate in certain unicellular, eukaryotic chlorella-like green algae. DNA sequence analysis of its 330,742-bp genome leads to the prediction that this phycodnavirus has 376 protein-encoding genes and 10 transfer RNA genes. The predicted gene products of ~40% of these genes resemble proteins of known function. The chlorella viruses have other features that distinguish them from most viruses, in addition to their large genome size. These features include the following: (a) The viruses encode multiple DNA methyltransferases and DNA site-specific endonucleases; (b) PBCV-1 encodes at least part, if not the entire machinery to glycosylate its proteins; (c) PBCV-1 has at least two types of introns—a self-splicing intron in a transcription factor-like gene and a spliceosomal processed type of intron in its DNA polymerase gene. Unlike the chlorella viruses, large double-stranded-DNA-containing viruses that infect marine, filamentous brown algae have a circular genome and a lysogenic phase in their life cycle.

**Keywords:** chlorella viruses, brown algal viruses, DNA restriction/modification enzymes, virus-encoded glycosylation, lysogeny

## Contents

|  |     |
|--|-----|
| Introduction .....   | 448 |
| Chlorella Viruses .....  | 450 |
| PBCV-1 Morphology .....  | 451 |
| Physical and Chemical Properties of PBCV-1 .....                                   | 451 |
| PBCV-1 Genome .....  | 453 |
| PBCV-1 Life Cycle .....  | 455 |
| PBCV-1 Transcription .....   | 457 |
| Similarities of PBCV-1 Open Reading Frames to Proteins in the Databases .....      | 458 |
| PBCV-1 Deletion Mutants .....  | 466 |
| Viruses Encode DNA Methyltransferases and<br>DNA Site-Specific Endonucleases ..... | 467 |
| Glycosylation of PBCV-1 Structural Proteins Is Unusual .....                       | 470 |
| Natural History of the Chlorella Viruses .....                                     | 471 |

|  |     |
|--|-----|
| <i>Characteristics of Marine Algal Viruses</i> .....         | 475 |
| <i>Host Algae</i> .....                                      | 475 |
| Evolutionary Relationship of the Algal Viruses to Each Other |     |
| and to Other Viruses .....                                   | 480 |
| Concluding Remarks .....                                     | 483 |

## Introduction

Since the early 1970s, viruses or virus-like particles have been reported in at least 44 taxa of eukaryotic algae, which include members in 10 of the 14 classes of alga (171). However, most of these reports described isolated accounts of microscopic observations, and the virus particles were not characterized further for many reasons. This situation changed in the early 1980s with the discovery of large double-stranded-DNA (dsDNA)-containing viruses that infect and replicate in certain strains of unicellular, eukaryotic, ex-symbiotic chlorella-like green algae (also referred to as zoochlorellae) and more recently with the finding of tractable virus systems in brown algae.

In 1978, Kawakami & Kawakami (60) described the appearance of large (180-nm diameter), lytic viruses in zoochlorellae after these algae were released from the protozoan *Paramecium bursaria*. No virus particles were detected in zoochlorellae growing symbiotically inside the paramecium. Independently, lytic viruses were described in zoochlorellae isolated from the green coelenterate *Hydra viridis* (93, 173) and also from *P. bursaria* (174). As in the 1978 report, viruses appeared only after the zoochlorellae were separated from their hosts. Fortunately, the zoochlorellae from *P. bursaria* can be grown free of the paramecium in culture, and ensuing experiments have revealed that these cultured algae serve as hosts for many closely related viruses. These lytic chlorella viruses can be produced in large quantities and assayed by plaque formation with standard bacteriophage techniques (159, 167). The chlorella viruses have several unique properties including the following: (a) They have large, 330- to 380-kb genomes (128, 195), and analyses of the recently sequenced 330,742-bp genome of the prototype chlorella virus, *Paramecium bursaria* chlorella virus 1 (PBCV-1), have revealed many interesting and unexpected putative genes. (b) The viruses encode multiple DNA methyltransferases and DNA site-specific (restriction) endonucleases. (c) PBCV-1 encodes at least part, if not all, of the intracellular machinery used to glycosylate its proteins. (d) PBCV-1 has at least two different types of introns, a self-splicing intron in a transcription factor (TF)-like gene and a spliceosomally processed intron in its DNA polymerase gene. Of the 10 PBCV-1-encoded transfer RNA (tRNA) genes, 1 may also contain a small intron.

In addition to the chlorella viruses, a few large icosahedral, dsDNA-containing viruses that infect marine algae are under active investigation. These include viruses that infect the unicellular alga, *Micromonas pusilla* (MpV viruses) (19, 87), and viruses that infect the filamentous brown algae, *Ectocarpus*

**Table 1.** Representative algal viruses

| Host                          | Virus  | Genome size (kb) | %5mC <sup>a</sup> | %6mA <sup>b</sup> |
|-------------------------------|--------|------------------|-------------------|-------------------|
| <i>Chlorella</i> NC64A        | CA-4B  | ~330             | 0.12              | ND <sup>c</sup>   |
|                               | AL-1A  | ~330             | 0.45              | ND                |
|                               | PBCV-1 | 330              | 1.9               | 1.5               |
|                               | IL-3A  | ~335             | 9.7               | ND                |
|                               | SC-1A  | ~330             | 1.9               | 7.3               |
|                               | NC-1A  | ~335             | 7.1               | 7.3               |
|                               | XZ-3A  | ~330             | 12.8              | 2.2               |
|                               | BJ-2C  | ~330             | 12.8              | 11.5              |
|                               | NE-8A  | ~330             | 14.3              | 8.1               |
|                               | NYs-1  | ~345             | 47.5              | 11.3              |
|                               | AL-2A  | ~330             | 35.8              | 14.6              |
|                               | CA-4A  | ~335             | 39.8              | 19.6              |
|                               | NY-2A  | ~380             | 44.9              | 37.0              |
|                               | XZ-6E  | ~330             | 21.2              | 15.2              |
|                               | XZ-4C  | ~330             | 46.7              | 20.8              |
|                               | XZ-4A  | ~330             | 44.1              | 28.3              |
| <i>Chlorella</i> Pbi          | CVA-1  | >300             | 43.1              | ND                |
|                               | CVB-1  | >300             | 42.7              | 17.7              |
|                               | CVG-1  | >300             | 19.2              | ND                |
|                               | CVM-1  | >300             | 41.9              | 10.1              |
|                               | CVR-1  | >300             | 14.2              | ND                |
| <i>Ectocarpus siliculosus</i> | EsV1   | 320              | 1                 | 3                 |
| <i>Feldmannia</i> sp.         | FsV-1  | 158              | ?                 | ? <sup>d</sup>    |
|                               | FsV-2  | 178              | ?                 | ?                 |
| <i>Micromonas pusilla</i>     | MpV-1  | ?                | ?                 | ?                 |

<sup>a</sup> Percentage of 5mC per C plus 5mC plus deoxyuridine.

<sup>b</sup> Percentage of 6mA per A plus 6mA plus deoxyinosine.

<sup>c</sup> ND, is none detected.

<sup>d</sup> Question marks are unknown.

sp. (EsV viruses) (104) and *Feldmannia* sp. (FsV viruses) (50) (Table 1). EsV viruses, which are transmitted from cell to cell in a Mendelian fashion, and FsV viruses have a lysogenic phase in their life cycles and produce virus particles only in differentiated sporangial cells of their hosts.

The chlorella viruses and the marine algal viruses are widely distributed in nature, and they have been isolated from fresh water or seawater collected throughout the world. Phylogenetic trees constructed from virus-encoded

DNA polymerases indicate that all of these algal viruses probably have a common evolutionary heritage. This review focuses primarily on the chlorella and brown algal viruses. However, because the chlorella viruses and brown algal viruses have different life styles, they are described separately. Information on the early history of the algal viruses has been reviewed by Van Etten *et al.* (171) and Fuhrman & Suttle (29). Muller *et al.* (103) have written a recent review on certain aspects of the viruses in marine brown algae, and Suttle (153) has written a review on ecological aspects of these fascinating viruses. Finally, the first meeting ever devoted exclusively to algal viruses was held in Bergen, Norway, in June 1998.

## Chlorella Viruses

Algae included in the genus *Chlorella* are among the most widely distributed and frequently encountered algae on Earth. *Chlorella* algae include small, spherical, unicellular, nonmotile, asexually reproducing green algae (164). *Chlorella* species have a rigid cell wall and typically have a single chloroplast, which sometimes contains a pyrenoid body. They have a simple developmental cycle and reproduce by mitotic division. Vegetative cells increase in size and, depending on the species and environmental conditions, divide into two, four, eight, or more progeny, which are released by rupture or enzymatic digestion of the parental walls. However, algae assigned to the genus *Chlorella* are more heterogeneous than their simple morphology suggests. This degree of heterogeneity is illustrated with two examples: (a) The nuclear DNAs of algae assigned to the genus have G+C contents ranging from 43% to 79% (49). However, most isolates that have been assigned to the same species contain similar G+C contents. (b) The cell wall polysaccharides of *Chlorella* species also vary widely (80, 158, 197), even among isolates assigned to the same species, e.g. *Chlorella vulgaris* and *Chlorella ellipsoidea* (197). Because of the heterogeneity in this genus, properties discovered in one *Chlorella* species or isolate may not apply to another species.

Most *Chlorella* species are free-living in nature. However, some forms, called zoochlorellae or *Chlorella*-like algae, live as hereditary endosymbionts within fresh water and, to less extent, marine animals (e.g. 122, 162). To our knowledge, the only hosts for the chlorella viruses are these symbiotic zoochlorellae, some of which can be cultured under laboratory conditions. These zoochlorellae have not been assigned species status.

Chlorella viruses infect either *Chlorella* strain NC64A and its equivalent strains (NC64A viruses) or *Chlorella* strain Pbi (Pbi viruses). About 50 of the NC64A viruses have been partially characterized. These viruses can be grouped into at least 16 classes identified by plaque size, antiserum sensitivity, DNA restriction patterns, sensitivity of the DNAs to restriction endonucleases, and, most importantly, the nature and abundance of methylated bases in their DNAs. One member of each of the 16 virus classes is listed in Table 1. Five Pbi viruses are also listed in Table 1.

### *PBCV-1 Morphology*

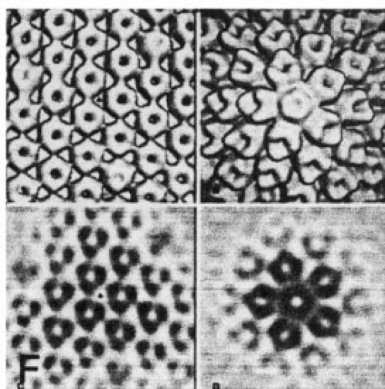
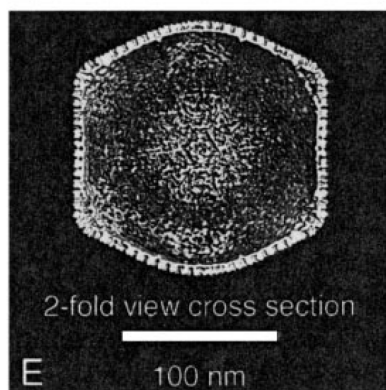
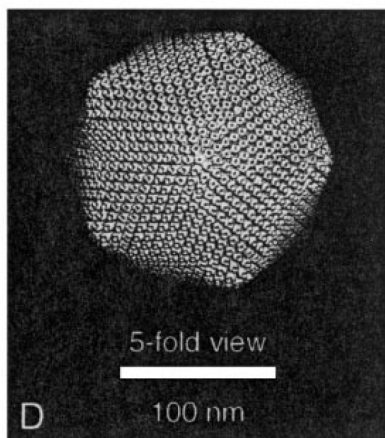
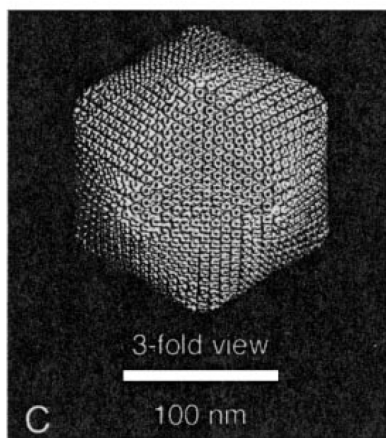
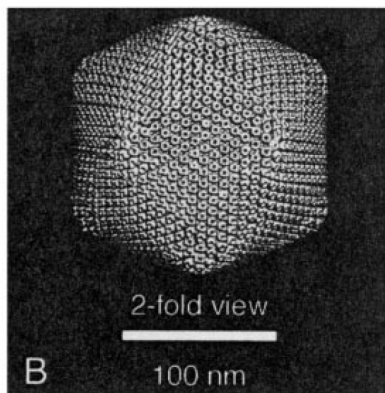
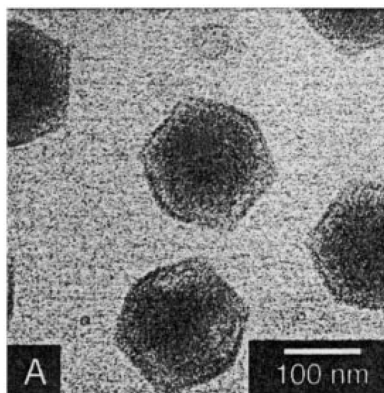
The chlorella viruses are large icosahedra with multilaminar shells surrounding electron-dense cores. Depending on the microscopic techniques, diameters of 140–190 nm have been reported for the viruses (171). A three-dimensional reconstruction of PBCV-1 was recently created from 356 cryo-electron micrographs of PBCV-1 (Figure 1) (199). The image, which has a resolution limit of 26 angstroms, reveals several features about PBCV-1 morphology. The capsid has a distinct icosahedral shape, with a maximum diameter of 190 nm along the fivefold axes. The outer dimensions of the virion at the two- and threefold axes are both ~165 nm. The capsid consists of 1692 doughnut-like capsomers arranged in a  $T = 169$ , skew icosahedral lattice ( $h = 8$ ,  $k = 7$ ) (14). Twelve capsomers, each ~6.4 nm in diameter, occur as pentamers at the fivefold vertices. The remaining 1680 capsomers are trimeric structures, each about 6.7 nm in diameter and 7.5 nm high. The prominent, cylinder portion of each trimeric capsomer extends 5 nm above the surface of the capsid shell, and most appear to have axial channels. The capsomers interconnect at their bases in a contiguous shell with a thickness of 2–2.5 nm. The stoichiometry of the trimers indicates that the PBCV-1 capsid contains 5040 copies of the major capsid protein, Vp54. The difference in oligomer state and size of the two types of capsomers suggests that the pentamers are composed of one of the minor capsid proteins. In addition, unlike the trimeric capsomers, each pentamer has a cone-shaped, axial cavity at its base.

Other types of microscopy suggest that the surface of PBCV-1 has additional features beyond that of an icosahedron. For example, PBCV-1 particles prepared by the quick-freeze, deep-etch procedure have flexible, hairlike fibers that extend from at least some of the virus vertices. The tips of these appendages are swollen and may aid in attachment of the virus to the host (see Figure 3B and Figure 8 in reference 171). Negative staining suggests that the virions contain a distinctive 20- to 25-nm spike or tail extending from one vertex, and that the DNA within the particle is retracted from this unique vertex (see Figure 3C of reference 171). The vertex containing this spike may be the one involved in digesting the host cell wall at the point of attachment (see below). Ultrastructural studies on a Pbi virus also suggested that one vertex is unique (7).

### *Physical and Chemical Properties of PBCV-1*

PBCV-1 has a sedimentation coefficient of about 2300 S in sucrose density gradients (170) and an estimated molecular mass of  $1 \times 10^9$  Daltons (201). From 25% to 50% of the viral particles are infectious and form plaques (170). The PBCV-1 virion contains 64% protein, 21% to 25% DNA, and 5% to 10% lipid (143). The lipid component is located inside the outer glycoprotein capsid shell and is required for virus infectivity. PBCV-1 particles contain more than 50 polypeptides, which range in size from 10 to 280 kDa (117, 143). The PBCV-1 major capsid protein weighs 53,790 Daltons (Vp54) (42; R. Cerny, J.





L. Van Etten, unpublished results) and composes about 40% of the total viral protein. Protein Vp54, along with two other PBCV-1 proteins Vp51 and Vp27.5, is myristoylated, and at least six of the viral proteins are phosphorylated (117, 118). The myristic acid is attached by amide linkages to the three PBCV-1 proteins. Surprisingly, the myristic acid of Vp54 is not attached to the N terminus but probably to an internal lysine  $\epsilon$ -amino group (117).

At least three PBCV-1 proteins, the major capsid protein Vp54 as well as Vp280 and Vp260, are glycosylated (42, 117, 179). The glycan portion of Vp54 is on the external surface of PBCV-1 and undoubtedly contributes to the protease resistance of the virus and to its antigenicity. The glycan is probably linked to the protein by a Ser/Thr O-linkage (117, 179). The amino acid sequence of the major capsid protein predicts a molecular size of 48,116 Daltons (the N-terminal Met is removed from the protein). Therefore, the additional weight of the protein (5674 Daltons) results from glycosylation, myristoylation, and possibly other unknown post-translational modifications.

Virion proteins from NC64A virus CVK2, which is closely related to PBCV-1, were separated into 10 capsid proteins and at least 40 core proteins (196). Comparison of the N-terminal amino acids of seven of the capsid proteins with the corresponding open reading frames (ORFs) deduced from the PBCV-1 genomic sequence indicated that the N termini of six of the CVK2 proteins were proteolytically processed (146). Three of the proteins were missing the N-terminal methionine, presumably removed by methionine aminopeptidase. Two of the apparent PBCV-1 homologs were missing a hydrophobic 25- to 30-amino-acid signal peptide-like region at their amino ends. Each of the predicted proteolytic cleavage sites for both of these proteins contains a lysine with an acidic amino acid on its N-terminal side. Seven of the insoluble core proteins had strong DNA-binding activity (193).

### *PBCV-1 Genome*

The PBCV-1 genome is a linear, 330-kb, nonpermuted dsDNA with covalently closed hairpin termini (35, 128). Some of the chlorella viruses have genomes as large as 380 kb (128, 195). The termini of the PBCV-1 genome

---

**Figure 1.** (facing page) **A.** Micrograph of vitrified chlorella virus PBCV-1. This image shows a well-defined outer capsid and a nonuniformly distributed interior mass. **B-D.** Shaded-surface view of a three-dimensional reconstruction viewed along **(B)** a twofold axis, **(C)** a threefold axis, and **(D)** a fivefold axis. **(E)** Reconstruction of a cross section of the virus from a twofold view. **(F)** Close-up views along threefold **(left)** and fivefold **(right)** axes of the PBCV-1 density map, represented as shaded-surfaces **(top)** or as density projections **(bottom)**. The trimeric nature of the hexavalent capsomers is clearly evident. Projection images include planar slabs of density, 3.3 nm **(left)** and 4.4 nm **(right)** thick, and include the most radially extended features in each view.



consist of 35-nucleotide-long, incompletely base-paired, covalently closed hairpin loops that exist in one of two forms; the two forms are complementary when the 35 nucleotide sequences are inverted (flip-flop) (208). An identical 2221-bp inverted repeat is adjacent to each hairpin end (152). The remainder of the PBCV-1 genome contains primarily single-copy DNA (35).

The sequences of the inverted-repeat regions differ among NC64A viruses. For example, 8 of 37 viruses did not hybridize with the PBCV-1 repeat element, and the remaining 29 viruses hybridized to various degrees (152). Furthermore, the genome of NC64A virus CVK1 has inverted terminal repeats of about 1 kb; the sequence of this repeat has no obvious resemblance to the PBCV-1 repeat region (194).

The PBCV-1 genome (not counting the 35-nucleotide-long hairpin ends) was recently sequenced (330,742 bp) and analyzed (68, 77, 78, 82, 83). The PBCV-1 genome contains 701 potential protein-coding regions or ORFs, defined as continuous stretches of DNA that translate into a polypeptide initiated by an ATG translation start codon and extend for 65 or more codons. The ORFs were numbered consecutively starting with ORF a1L at the extreme left terminus of the genome and ending with ORF a692R. The 701 ORFs are higher than the last numbered ORF a692R because 10 small ORFs were inadvertently overlooked during the initial analysis. In addition, correction of a DNA-sequencing error revealed that two adjacent ORFs are fused (L. Sun & J. L. Van Etten, unpublished results). The letters R and L following the ORF number indicate that the transcript is either in a left-to-right or right-to-left orientation, respectively.

The 701 ORFs were divided into 376, mostly nonoverlapping ORFs (major ORFs, labeled "A"), which are believed to encode proteins, and 325 minor ORFs (labeled "a") which may or may not encode proteins. When ORFs overlapped extensively, the larger ORF was classified as a major ORF and the smaller ORF(s) was classified as a minor ORF. The 50 nucleotides preceding the start codon of most (293 of 376) of the major ORFs have A+T concentrations >70%. In contrast, the 50 bases preceding the start codon of most (289 of 325) of the minor ORFs have A+T concentrations <70%. Previous studies indicate that promoter regions of protein encoding ORFs are usually A+T rich (135).

The 376 major ORFs are evenly distributed along the genome and, with one exception, there is little intergenic space between them. The exception is a 1788-bp sequence near the middle of the genome. This region, which has numerous stop codons in all reading frames, encodes 10 tRNA genes. The middle 900 bp of this intergenic region also resemble a "CpG island" (2).

The PBCV-1 genome shows interspersed leftward and rightward transcription throughout the genome. Overall, 183 major PBCV-1 ORFs are transcribed in the left-to-right orientation and 193 in the right-to-left orientation. The first four PBCV-1 ORFs, a1L, A2L, A3R, and a4L, reside in the inverted terminal repeat region of the PBCV-1 genome and hence are duplicated at the right end of the genome (83, 152).

One unusual feature of chlorella virus DNAs is that they contain relatively high levels of methylated bases (175). In fact, chlorella viruses can be

distinguished from one another by the site specificity and abundance of DNA methylation (Table 1) (see reviews in 108, 109, 171). DNA from each *Chlorella* virus contains 5-methylcytosine (5mC) in amounts ranging from 0.1% to 47.5% of the total cytosine. Many viruses also contain  $N^6$ -methyladenine (6mA) in amounts ranging from 1.45% to 37% of the total adenine (Table 1). Regardless of the level of methylation, all of the NC64A viruses have a G+C content of ~40%, and the Pbi viruses have a G+C content of ~46% (125).

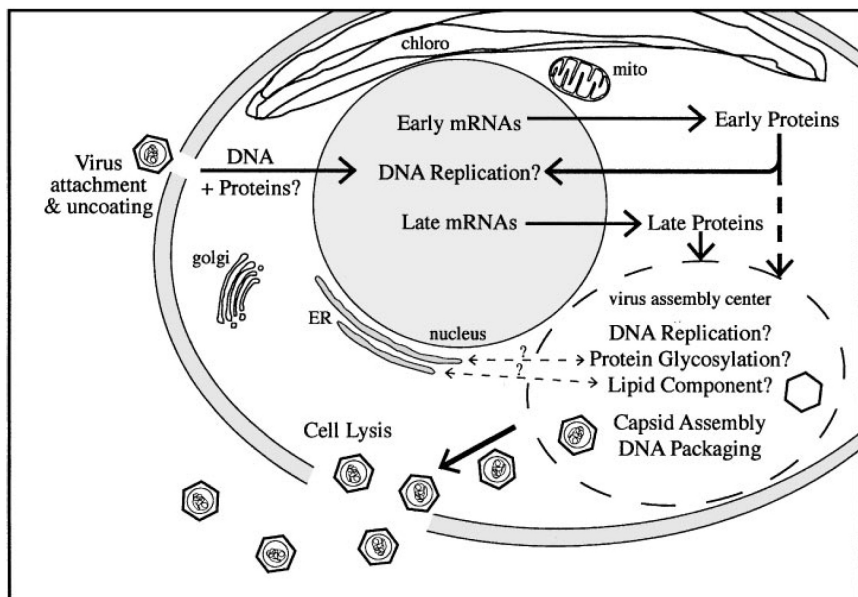
### *PBCV-1 Life Cycle*

PBCV-1 attaches rapidly and irreversibly to the external surface of *Chlorella* NC64A cell walls but not to other *Chlorella* strains (91). The virus attaches to the cell wall via one of its hexagonal vertices, digests the wall at the attachment point, and releases viral DNA into the host leaving an empty capsid on the cell surface. Because the virus also attaches to and digests cell wall fragments, which have been boiled or extracted by harsh procedures, the wall-degrading enzyme(s) must be packaged in the virus particles (90, 126). Release of the virus DNA must require a host function because attachment to wall fragments and digestion of the wall at the point of attachment does not release viral DNA.

Stereo views reveal that PBCV-1 attaches to the wall by the hairlike fibers that originate from at least some virus vertices. These micrographs suggest that the tips of the hairlike fibers are responsible for the initial recognition and attachment of the virus to the host receptor (see Figure 8 of reference 171). Preliminary results indicate that the host receptor for PBCV-1 is the sugar portion of a lipopolysaccharide-like component (T. S. Hagge & R. L. Parady, unpublished results). The discovery that *Chlorella* NC64A cell walls have a lipopolysaccharide component (129) was unexpected because lipopolysaccharides had previously been found only in the walls of gram-negative bacteria.

Two observations indicate that the infecting PBCV-1 DNA and probably DNA-associated proteins are targeted to the nucleus, where early virus transcription occurs. First, PBCV-1 does not encode a recognizable RNA polymerase gene(s), nor were we able to detect RNA polymerase activity in isolated virions (J. Rohozinski & J. L. Van Etten, unpublished results). Second, a small intron with the splice site sequences characteristic of a nuclear-spliced mRNA is present in the PBCV-1 DNA polymerase gene (41). Presumably this intron would need to be excised in the infected cell nucleus. The signal(s) involved in targeting the infecting virus DNA to the nucleus is unknown. However, the process is rapid because early virus mRNAs are synthesized within 5–10 min postinfection (p.i.). Virus replication does not require a functional host nucleus because PBCV-1 replicates, albeit poorly and with a small burst size, in UV-irradiated cells (168). PBCV-1 proteins are synthesized on cytoplasmic ribosomes and not on organelle ribosomes because cycloheximide, but not chloramphenicol, inhibits virus replication (170).

Virus DNA replication begins ~1 h p.i. and is followed by transcription of late virus genes (134, 166). Virus DNA synthesis presumably takes place in



**Figure 2.** Proposed life cycle for chlorella virus PBCV-1. The virus uncoats at the surface of the alga, and the DNA, possibly with associated proteins, goes to the nucleus where early gene transcription begins within 5–10 min after infection. The early mRNAs are transported to the cytoplasm for translation, and the early proteins presumably return to the nucleus to initiate DNA replication beginning 60–90 min after infection, followed by late gene transcription. Late mRNAs are transported to the cytoplasm for translation, and many of these late proteins are targeted to the cytoplasmically located virus assembly centers, where virus capsids are formed. However, some of the virus capsid proteins are glycosylated, myristoylated, and phosphorylated at some intracellular location before assembly into capsids. Virus DNA also has to move from the nucleus to the virus assembly centers for additional replication and packaging into virions. Infectious PBCV-1 particles appear in the cytoplasm 50–60 min before release from the cell, which occurs by localized lysis of the cell wall.

the nucleus along with late transcription. An intracellular traffic pattern for PBCV-1 replication is proposed in Figure 2. Host RNA and protein syntheses are inhibited at least 75% after PBCV-1 infection. Chloroplast rRNAs, but not cytoplasmic rRNAs, are degraded beginning ~30 min p.i. Host nuclear and chloroplast DNAs are degraded beginning 1 h p.i., even though the total DNA in the cell increases 4- to 10-fold by 4 h p.i. (88, 166). Thus PBCV-1 DNA synthesis requires a large increase in deoxyribonucleoside triphosphates (dNTP) after infection. Degraded host DNA may supply some of these intermediates; however, virus replication requires 4- to 10-fold the dNTPs that recycling can provide. As noted below, PBCV-1 encodes many putative enzymes involved in dNTP metabolism.

PBCV-1 infection rapidly inhibits host CO<sub>2</sub> fixation (170) and photosynthesis (136). PBCV-1 does not require host photosynthesis for replication because the virus replicates in dark-grown algae as well as light-grown cells treated with the photosynthetic inhibitor DCMU (170). However, the virus burst size is reduced ~50% under these conditions. Like most bacterial viruses (48), PBCV-1 replicates most efficiently in actively growing host cells and poorly in stationary-phase cells (170).

The host cells change cytologically beginning ~1 to 2 h p.i.; the nucleus, mitochondria, and Golgi apparatus become appressed to the chloroplast, leaving one or more finely granulated electron-translucent areas in the cytoplasm (92). Because PBCV-1 virions ultimately assemble in these translucent areas, the areas have been designated virus assembly centers. Capsids usually form at the periphery of the virus assembly centers, occasionally producing a rosette pattern. Complete virus capsids appear to assemble before DNA packaging. Presumably DNA is packaged through an opening in the capsid. By 4 to 5 h p.i., the cells contain many filled virus particles, which are distributed throughout the cytoplasm. Progeny PBCV-1 are first released ~4 to 5 h p.i., and the majority of infectious virus particles are released by 8 h p.i. PBCV-1 release involves localized lysis of the cell wall, presumably by a late viral gene product(s). Mechanical disruption of the cells releases infectious virions 30–50 min before spontaneous lysis (170). Consequently, infectious PBCV-1 is assembled inside the host, and the virus does not acquire its lipid components and glycoproteins by budding through an exterior host membrane. The typical PBCV-1 burst size is 200–350 plaque-forming units (PFU) (170).

Because PBCV-1 is assembled in specific regions in the cytoplasm, the cytoskeleton may participate in either targeting the virus proteins to the virus assembly centers or assembling virus proteins into capsids. However, several cytoskeleton-disrupting agents, which inhibit either tubulin or actin functions, had no effect on the formation of PBCV-1 virus assembly centers or PBCV-1 replication even at concentrations several fold higher than those that inhibited host *Chlorella* growth (110). As noted above, some PBCV-1-encoded proteins are extensively posttranslationally modified, for example, the N-terminal methionine or signal peptides are removed, and proteins are phosphorylated, myristoylated, and/or glycosylated. The intracellular site(s) where these events take place is unknown. One difference between PBCV-1 and other glycoprotein-containing viruses is that PBCV-1 encodes at least part, if not all, of the machinery required to glycosylate its three structural glycoproteins (see below).

### *PBCV-1 Transcription*

PBCV-1 transcription can be divided into early and late stages. The junction between these stages is 60 to 90 min p.i., which coincides with the initiation of virus DNA synthesis (134). It is interesting that the incorporation of [<sup>3</sup>H]adenine into polyadenylate-containing RNA also changes abruptly at about 1 h p.i. Early virus transcripts are probably polyadenylated, but late transcripts are virtually devoid of polyadenylate segments (176).

Only two studies (134, 135) have examined PBCV-1 RNA synthesis. Consequently only general statements can be made about virus transcription. (a) Viral infection rapidly inhibits host RNA synthesis. (b) PBCV-1 transcription is programmed, and early transcripts appear within 5–10 min p.i. (c) Late viral transcription begins ~60 to 90 min p.i., after viral DNA synthesis begins. (d) Some early transcripts continue to be synthesized after virus DNA synthesis begins. (e) Early and late virus genes are interspersed throughout the PBCV-1 genome. (f) The sum of the sizes of the mRNAs hybridizing to viral DNA probes is often 40% to 60% larger than the probe, suggesting that PBCV-1 either contains overlapping genes, transcribes both strands of DNA, or extensively processes RNA after transcription or that some transcripts are polycistronic. Transcription studies with individual PBCV-1 genes indicate that some genes produce the expected size mRNA as either an early or late gene, whereas other genes produce complex RNA patterns, even with single-stranded-DNA (ssDNA) probes. (g) Transcriptional mapping of seven PBCV-1 genes indicates that the start sites are within 150 nucleotides upstream from the translational start codon, producing mRNAs with 50 untranslated regions of as little as 14 nucleotides up to 149 nucleotides; transcripts also extend beyond the translational stop codon (42, 43, 135). To date, no early or late PBCV-1 promoters have been identified, although Schuster et al (135) noted that the regions upstream from the translation start codons for five major ORFs are A+T rich. As noted above, the content of the 50 bases preceding most of the major PBCV-1 ORFs is at least 70% ACT. (h) Codon usage by PBCV-1 is biased to codons ending in XXA/U (63%) over those ending in XXC/G (37%) (135). This bias is expected, because PBCV-1 DNA is 40% G+C. In contrast, the codon usage of the host tubulin gene has a strong bias to codons ending in XXC/G (67%) (RH Meints, MV Graves, unpublished results). This result is also consistent with the 67% GCC content of the host nuclear DNA (175). Therefore, the fact that PBCV-1 encodes 10 tRNA genes is not surprising.

### *Similarities of PBCV-1 Open Reading Frames to Proteins in the Databases*

Computer analyses of the 376 PBCV-1 major ORFs indicate that ~40% resemble proteins in databases, including many interesting and unexpected enzymes. Table 2 lists 77 genes that match database ORFs of known function and, in a few cases, indicates whether the PBCV-1 ORF is transcribed early (E) or late (L). The predicted enzyme activities of some PBCV-1-encoded proteins have been confirmed by complementation of *Escherichia coli* mutants and/or bioassay of recombinant proteins (indicated with an asterisk in Table 2). Additional PBCV-1 ORFs have significant homology to ORFs of unknown function.

Some PBCV-1-encoded proteins resemble those of bacteria and phages, such as DNA restriction endonucleases and methyltransferases. However, other virus-encoded proteins resemble those of eukaryotic organisms and

their viruses, such as translation elongation factor-3 (EF-3), RNA guanylyl-transferase, and proliferating cell nuclear antigen (PCNA). The PBCV-1 genome thus is a mosaic of prokaryotic and eukaryotic genes, suggesting considerable gene exchange in nature during the evolution of these viruses. A brief description of some of the PBCV-1 encoded proteins follows.

**DNA Replication- and Repair-Associated Proteins** PBCV-1 encodes at least 11 ORFs that are involved in either replicating or interacting with DNA, including a B-family ( $\alpha$ -like) DNA polymerase (41), an ATP DNA ligase (53), a type-II DNA topoisomerase, three helicases, two ORFs that resemble PCNA proteins, a cyclin A/cdk-associated protein, and a pyrimidine dimer-specific glycosylase endonuclease (PDG) (30).

The gene encoding the PBCV-1 DNA polymerase contains a 101-nucleotide intron with 5'-AG/GUGAG and 3'-GCAG/U splice site sequences, as well as a possible branch point UCAC sequence (41). These sequences are hallmarks of spliceosomal spliced mRNA introns (44). A similar 86-bp intron is also present in the DNA polymerase gene of another NC64A virus NY-2A (41). Experiments in progress indicate that the DNA polymerase genes from 38 other NC64A viruses also have an intron at the same position (Y. Zhang, D. B. Burbank, J. L. Van Etten, unpublished results). Interestingly, the DNA polymerase gene from five viruses that infect *Chlorella* Pbi lack the intron, as do the DNA polymerase genes from the brown algal virus FsV (75a) and several micromonas viruses and chrysochromulina viruses (16).

ATP-dependent DNA ligases, which are encoded by several viruses—including vaccinia virus, African swine fever virus (ASFV), and bacteriophages T3, T4, and T7—join two DNA fragments by catalyzing the formation of internucleotide phosphodiester bonds. They range in size from 268 amino acids in *Haemophilus influenzae* (17) to 1070 amino acids in *Xenopus laevis* (76). The 298-amino-acid-residue PBCV-1 enzyme, which is the second smallest DNA ligase described to date, has been expressed in *E. coli*, and the recombinant enzyme has been extensively characterized (53, 147–149).

The 1061-amino-acid-residue PBCV-1 type-II DNA topoisomerase has >45% amino acid identity to type-II topoisomerases from several eukaryotic organisms. Type-II topoisomerases are ATP dependent and act by passing a dsDNA segment through a transient double-strand break (127). ASFV is the only other virus infecting a eukaryotic cell that encodes a type-II DNA topoisomerase (31). However, the PBCV-1 and ASFV enzymes have only 23% amino acid identity. The PBCV-1-encoded enzyme is essential for virus replication because type-II topoisomerase inhibitors, such as amsacrine and etoposide, but not type-I topoisomerase inhibitors, inhibit plaque formation (M. Nelson, Y. Zhang, T. Morehead, J. L. Van Etten, unpublished results). The PBCV-1 DNA topoisomerase II is  $\geq 130$  amino acids smaller than the other type-II topoisomerases currently in the databases.

PBCV-1 encodes two ORFs that resemble PCNA proteins. PCNA is one of several proteins involved in forming a processive DNA replication complex (64, 150), and it is also involved in DNA repair (4). The two PBCV-1



**Table 2.** Putative open reading frames encoded by chlorella virus PBCV-1<sup>a</sup>

|   |  |  |                                     |
|---|--|--|-------------------------------------|
| DNA replication & repair                      |  | Phosphorylation/dephosphorylation      |                                     |
| (E) A185R                                     | DNA polymerase                         | (L) A248R*                             | Phosphorylase B kinase              |
| (E) A544R*                                    | DNA ligase                             | A277L                                  | Ser/Thr protein kinase              |
| (E) A583L*                                    | DNA topoisomerase II                   | A278L                                  | Ser/Thr protein kinase              |
| A193L   | PCNA                                   | A282L                                  | Ser/Thr protein kinase              |
| A574L   | PCNA                                   | A289L                                  | Ser/Thr protein kinase              |
| A153R   | Helicase                               | A305L                                  | Tyr phosphatase                     |
| A241R   | Helicase                               | A614L                                  | Protein kinase                      |
| A548L   | Helicase                               | A617R                                  | Tyr-protein kinase                  |
| (E) A50L*                                     | T4 endonuclease V                      | Cell Wall Degrading                    |                                     |
| A39L  | CyclinA/cdk-associated protein         | (E) A181/182R*                         | Chitinase                           |
| A638R   | Endonuclease                           | (L) A260R*                             | Endochitinase                       |
| Nucleotide metabolism                         |  | (L) A292L*                             | Chitosanase                         |
| (E) A169R*                                    | Aspartate transcarbamylase             | A94L                                   | $\beta$ -1,3 Glucanase              |
| A476R   | Ribo. reductase (small subunit)        | DNA restriction/modification $\square$ |                                     |
| A629R   | Ribo. reductase (large subunit)        | A251R*                                 | Adenine DNA methylase (M.CviAII)    |
| A427L   | Thioredoxin                            | A252R*                                 | Restriction endonuclease (R.CviAII) |
| A438L   | Glutaredoxin                           | (E) A517L*                             | Cytosine DNA methylase (M.CviAIII)  |
| A551L   | dUTP pyrophosphatase                   | (L) A530R*                             | Cytosine DNA methylase (M.CviAIV)   |
| A596R   | dCMP deaminase                         | (E) A581R*                             | Adenine DNA methylase (M.CviAI)     |
| A200R   | Cytosine deaminase                     | (E) A579L*                             | Restriction endonuclease (R.CviAI)  |
| A416R   | dG/dA kinase                           | A683L                                  | Cytosine DNA methylase (M.CviAV)    |
| A363R   | Phosphohydrolase                       | Sugar and lipid manipulation           |                                     |
| A392R   | ATPase                                 | (L) A64R                               | Galactosyl transferase              |
| A554L   | ATPase                                 | (E) A98R*                              | Hyaluronan synthase                 |
| A674R   | <i>Dictyostelium</i> Thy protein       | (E) A100R*                             | Glucosamine synthase                |
| Transcription                                 |  | A114R                                  | Fucosyltransferase                  |
| A107L   | RNA transcription factor TFIIB         | (E) A118R                              | GDP-D-mannose dehydratase           |
| A125L   | RNA transcription factor TFIIS         | A222R                                  | Cellulose synthase                  |
| A166R   | Exonuclease                            | A295L                                  | Fucose synthase                     |
| A422R   | Endonuclease                           | (E) A473L                              | Cellulose synthase                  |
| (E) A103R*                                    | RNA guanylttransferase                 | (E) A609L*                             | UDP-glucose dehydrogenase           |
| A464R   | RNase III                              | A49L                                   | Glycerophosphoryl diesterase        |
| Protein synthesis, modification & degradation |  | A53R                                   | 2-Hydroxyacid dehydrogenase         |
| (E,L) A666L                                   | Translation elongation factor-3        | A271L                                  | Lysophospholipase                   |
| A44L  | 26S protease subunit                   |  |                                     |
| A85R  | Prolyl 4-hydroxylase alpha-subunit     |  |                                     |
| A105L   | Ubiquitin C-terminal hydrolase         |  |                                     |
| A448L   | Protein disulfide isomerase            |  |                                     |
| A623L   | 10 tRNAs Ubiquitin-like fusion protein |  |                                     |

(Continued)

**Table 2** (*Continued*)

| Miscellaneous                |                                    | Pro-rich proteins       |
|------------------------------|------------------------------------|-------------------------|
| (E,L) A207R*                 | Ornithine decarboxylase            | A35L                    |
| (L) A237R                    | Homospermidine synthase            | A41R                    |
| A78R                         | $\beta$ -Alanine synthase          | A67R                    |
| A245R                        | Cu/Zn-superoxide<br>dismutase      | A145R<br>A189R          |
| A284L*                       | Amidase                            | A246R                   |
| A445L                        | Ubiquitous biosynthesis<br>protein | A316R<br>A333L          |
| A465R                        | Yeast ERV1 protein                 | A384L                   |
| A598L                        | Histidine decarboxylase            | A405R                   |
| A250R                        | K <sup>+</sup> Ion channel protein | A428L                   |
| A625R                        | Transposase                        | A488R<br>A500L<br>A678R |
| Contain aukyrin-like repeats | A5R                                |                         |
| A7L                          |                                    |                         |
| A8L                          |                                    | Gene families           |
| A682R                        |                                    | 16 with 2 members       |
| A330R                        |                                    | 8 with 3 members        |
| A607R                        |                                    | 3 with 6 members        |
| A672R                        |                                    | 2 with 8 members        |
| Virion structural proteins   |                                    |                         |
| A34R                         |                                    |                         |
| A122R                        |                                    |                         |
| A168R                        |                                    |                         |
| A203R                        |                                    |                         |
| (L) A430L                    |                                    |                         |
| A523R                        |                                    |                         |
| (L) A622L                    |                                    |                         |

<sup>a</sup> E and L refer to early and late genes, respectively. An asterisk means that the gene encodes a functional enzyme as determined either by complementation or by enzyme activity of a recombinant protein.

ORFs resemble PCNA proteins from other organisms more than each other (26% amino acid identity), indicating that PBCV-1 probably acquired the two genes independently.

Perhaps one of the more interesting PBCV-1 genes is one that resembles the bacteriophage T4 *denV* gene (30). The *denV* gene encodes a well-characterized, pyrimidine dimer-specific glycosylase, endonuclease enzyme called endonuclease V, that initiates repair of UV-induced thymidine dimers in DNA (79, 200). Although T4 endonuclease V-like enzyme activity has been discovered in several microorganisms, these enzymes differ significantly from T4 endonuclease V in size and amino acid sequence. The discovery of the PBCV-1-encoded T4 endonuclease V homolog (41% amino acid identity) led to structural and functional comparisons between the two enzymes. The PBCV-1 enzyme cleaves both *cis-syn* and *trans-syn*-II cyclobutane pyrimidine dimers (89), whereas the T4 enzyme cleaves only the *cis-syn* cyclobutane pyrimidine isomer. The PBCV-1 enzyme is also more processive than the T4 enzyme (89).

**Transcription-Associated Proteins** PBCV-1 does not encode a recognizable RNA polymerase or an RNA polymerase component, but it does encode TFIIB- and TFIIS-like elements, together with three ORFs that resemble endo- or exonucleases, as well as RNase III. The lack of a virus-encoded RNA polymerase supports the notion that the infecting virus DNA is targeted to the nucleus and that transcription is initiated by a host RNA polymerase, possibly in conjunction with one or more virus-packaged transcription factors.

PBCV-1 also encodes an mRNA-capping enzyme that catalyzes the transfer of GMP from GTP to the 5' diphosphate end of RNA (52). In its size, amino acid sequence, and biochemical properties, the PBCV-1 enzyme is more closely related to the yeast RNA guanylyltransferases than to the multifunctional capping enzymes encoded by other large DNA viruses such as the poxviruses and ASFV. To cap its transcripts, PBCV-1 must either encode additional 5'-processing activities or else rely on the host alga to provide these functions. However, the 5'-end structure of PBCV-1 mRNAs is unknown. The PBCV-1-capping enzyme has been crystallized (46, 47).

**Protein Synthesis, Modification, and Degradation** The chlorella viruses are the first viruses to encode a translation elongation factor protein. An 896-amino-acid-residue ORF from PBCV-1 has 95% amino acid identity with an EF-3-like protein from NC64A virus CVK2 (192) and 39% to 47% amino acid identity to EF-3 from fungi (8). Fungi require EF-3 for growth, and the protein stimulates EF-1 $\alpha$ -dependent binding of aminoacyl-tRNA to the ribosome. Like fungal EF-3 proteins, the CVK2 and PBCV-1 enzymes have an ABC transporter family signature and two ATP/GTP binding-site motifs. The CVK2-encoded EF-3 gene is expressed at both early and late stages of viral infection but not when viral DNA replication begins (192). Possibly the virus-encoded EF-3 alters the host protein synthetic machinery so that viral mRNAs are preferentially translated after virus infection.

PBCV-1 also encodes five other ORFs that resemble proteins involved in either post-translational modification [excluding phosphorylation and glycosylation (see below)], prolyl 4-hydroxylase  $\alpha$ -subunit and protein disulfide isomerase or protein-degrading events, a ubiquitin carboxy-terminal hydrolase, a ubiquitin-like fusion protein, and a 26S protease subunit.

Finally, PBCV-1 encodes 10 tRNAs: 3 for Lys, 2 for Asn, and 1 each for Leu, Ile, Tyr, Arg, and Val. The Tyr-tRNA gene is predicted to have a 13-nucleotide intron. None of these tRNAs have a CCA sequence encoded at the 3' end of the acceptor stem of the tRNA. Typically these 3 nucleotides are added separately to tRNAs. Some NC64A viruses encode up to 14 tRNAs (T. Yamada, K. Nishida, M. Fujie, S. Usami, unpublished results). The tRNA genes are cotranscribed at both early and late stages of virus replication as a large precursor RNA and processed via various intermediates. Codon usage analyses of several virus-encoded proteins indicate a strong correlation between the abundance of virus-encoded tRNAs and the virus protein codons.

**Nucleotide Metabolism-Associated Proteins** PBCV-1 encodes at least 13 enzymes involved in nucleotide metabolism. These enzymes are important because the DNA concentration in a virus-infected cell increases rapidly after infection (166). Consequently, large quantities of dNTPs need to be synthesized *de novo* during this time period. The PBCV-1-encoded enzymes include aspartate transcarbamylase, both subunits of ribonucleotide reductase, thioredoxin, glutaredoxin, dUTP pyrophosphatase (dUTPase), deoxycytidylate (dCMP) deaminase, cytosine deaminase, deoxyguanosine/adenosine kinase, nucleoside triphosphate phosphohydrolase, and two ATPases.

PBCV-1 is the first virus known to encode aspartate transcarbamylase, the key regulatory enzyme in the *de novo* biosynthetic pathway of pyrimidines. The PBCV-1 gene has been expressed in *E. coli*, and the recombinant protein has the expected enzyme activity (71). The gene is expressed early in PBCV-1 replication.

PBCV-1, like many large DNA-containing viruses such as the poxviruses, ASFV, herpes virus, and lymphocystis disease virus, encodes both ribonucleotide reductase subunits. Both subunits are also encoded by the brown algal virus FsV (A. M. Lee & R. H. Meints, unpublished results). Because inhibitors of ribonucleotide reductase activity, such as hydroxyurea and *cis*-platinum (M. Nelson, N. Padhye, J. L. Van Etten, unpublished results), reduce PBCV-1 plaque formation, the enzyme is important for virus replication.

PBCV-1 encodes two putative enzymes, dUTPase and dCMP deaminase, that produce dUMP, the substrate for thymidylate synthetase. Interestingly, PBCV-1 does not encode a thymidylate synthetase. However, PBCV-1 ORF A674R is about the same size and has 29% amino acid identity to the Thy1 protein from *Dictyostelium discoideum*. Thy1 complements the thymidine growth requirement of a *Dictyostelium* mutant (26). A674R also has 51% amino acid identity to an ORF from a *Synechocystis* sp. (57) and about 25% amino acid identity to ORFs from *Brevibacterium lactofermentum* and *Corynebacterium glutamicum* (115). It seems likely that all of these putative proteins participate in a thymidine synthesis pathway that lacks a traditional thymidylate synthetase.

**Phosphorylation/Dephosphorylation-Associated Proteins** Seven PBCV-1 ORFs have similarities to protein kinases, and one ORF resembles a phosphatase. At least five of these ORFs are predicted to be Ser/Thr protein kinases. Interestingly, five of these ORFs are tightly clustered on the PBCV-1 genome (82). The gene for a 35-kDa Ser/Thr protein kinase (ORF A248R) was expressed in *E. coli* and shown to have both autophosphorylation and exogenous phosphorylation activities (118). This gene was expressed late in the PBCV-1 life cycle. This protein kinase was not packaged in nascent virions, although disrupted virions had protein kinase activity. Three protein kinase activities have been detected in virions of the NC64A virus CVK2 (193).

As compared with other viruses, this large number of virus-encoded proteins involved in phosphorylation/dephosphorylation is unusual. Because

protein phosphorylation is a common mechanism for regulating cellular processes, we assume that the PBCV-1-encoded proteins are involved in one or more signal transduction pathways important for virus replication.

**Sugar-and Lipid-Manipulating Proteins** A major surprise was the discovery that 12 PBCV-1 ORFs have high identity to enzymes involved in manipulating sugars, synthesizing polysaccharides, or metabolizing lipids. Three enzymes, glutamine:fructose-6-phosphate amidotransferase (GFAT), UDP-glucose dehydrogenase (UDP-GlcDH), and hyaluronan synthase, are involved in the synthesis of hyaluronan, a linear polysaccharide composed of alternating  $\beta$ 1,4-glucuronic acid and  $\beta$ 1,3-*N*-acetylglucosamine groups. The three PBCV-1 genes have been expressed in *E. coli*, and each recombinant protein has the expected enzyme activity (23, 70). The three genes are transcribed early in PBCV-1 infection, and hyaluronan accumulates on the external surface of the host *Chlorella* cells (41a). This was unexpected because, heretofore, hyaluronan had been found only in the extracellular matrix of vertebrates (73) and the extracellular capsules of a few bacterial pathogens (54, 131). It is interesting that the predicted amino acid sequences of the PBCV-1-encoded GFAT and UDP-GlcDH enzymes resemble bacterial enzymes, whereas the amino acid sequence of the PBCV-1-encoded hyaluronan synthase most resembles the vertebrate enzyme (70). These observations suggest that the GFAT and UDP-GlcDH genes were acquired by the viruses separately from the hyaluronan synthase gene.

Two of the putative PBCV-1-encoded enzymes, GDP-D-mannose dehydratase and fucose synthase, compose a pathway that converts GDP-D-mannose to GDP-L-fucose (1). These genes have been expressed in *E. coli*, and recombinant proteins are currently being assayed for enzyme activity.

Three of the PBCV-1 ORFs resemble lipid-metabolizing enzymes, a glycerophosphoryl diesterase, a 2-hydroxyacid dehydrogenase, and a lysophospholipase. PBCV-1 deletion mutants, which lack the neighboring glycerophosphoryl diesterase and the 2-hydroxyacid dehydrogenase genes, grow in the laboratory (69).

**Cell Wall-Degrading Enzymes** PBCV-1 encodes five ORFs potentially involved in degrading *Chlorella* cell walls, including two chitinases, a chitosanase, a  $\beta$ -1,3-glucanase, and an amidase. We initially reported that PBCV-1 encoded three chitinase genes (82). However, re-examination of the PBCV-1 sequence revealed that two adjacent chitinase ORFs are components of a single ORF (L. Sun, J. L. Van Etten, unpublished results). All five genes have been expressed in *E. coli*, and the recombinant proteins have the expected enzyme activities (L. Sun, Y. Ye, J. L. Van Etten, unpublished results).

The chitosanase gene, a late gene, has been characterized from another NC64A virus, CVK2 (196). The CVK2 gene encodes two proteins, a 37-kDa protein of the expected size from the DNA gene sequence and a 65-kDa protein. The larger protein is packaged in nascent virions, whereas the smaller

protein is found only in the infected cells. Interestingly, PBCV-1 encodes the same chitosanase gene flanked by two ORFs that also exists in virus CVK2. However, the CVK2 genome has an extra ORF inserted immediately downstream of its chitosanase gene, which contributes to the bigger 65-kDa protein, presumably as a read-through product. It will be interesting to determine the size of the PBCV-1 chitosanase protein.

The presence of chitinase and chitosanase genes in PBCV-1 and CVK2 was unexpected because chitin, a linear homopolymer of  $\beta$ -1,4-linked *N*-acetylglucosamine residues that is a normal component of fungal cell walls and the exoskeleton of insects and crustaceans (37), rarely occurs in algal walls (51). However, Kapaun & Reisser (58) reported a chitin-like glycan in the cell wall of *Chlorella* Pbi. We assume that these enzymes degrade the host cell wall during virus infection and/or virus release.

**Restriction-Modification (R-M) Enzymes** As discussed below, the *Chlorella* viruses encode multiple R-M genes. PBCV-1 encodes two 6mA DNA methyltransferases (MTases), each with a corresponding restriction endonuclease and three 5mC DNA MTases.

**Miscellaneous Proteins** PBCV-1 is predicted to encode several additional enzymes, including the polyamine-metabolizing enzymes ornithine decarboxylase (ODC) and homospermidine synthase. ODC catalyzes decarboxylation of ornithine to putrescine, which is a key regulatory step in polyamine biosynthesis (22). ODC, which turns over rapidly, is regulated by an antizyme-protein:protein interaction in mammalian cells (18). The predicted size of the PBCV-1-encoded ODC is smaller (372 codons) than other ODCs (ca. 460 codons), and the smaller size can be attributed to four deletions. The PBCV-1 *odc* gene has been expressed in *E. coli*, and the recombinant protein is enzymatically active (T. Morehead and J. L. Van Etten, unpublished results).

The PBCV-1-encoded homospermidine synthase (*hss*) gene has also been expressed in *E. coli*, and the recombinant protein synthesizes homospermidine from two molecules of putrescine (A. Kaiser, M. Vollmert, D. Tholl, K. W. Nickerson, M. V. Graves, J. Gurnon, J. L. Van Etten, unpublished results). The *hss* mRNA is expressed late during PBCV-1 infection, and PBCV-1 virions accumulate homospermidine. However, homospermidine, a rare polyamine, is not the major polyamine in the virions. The significance of a PBCV-1-encoded pathway that synthesizes homospermidine is unknown.

PBCV-1 also encodes an ORF with 50% amino acid identity to Cu/Zn-superoxide dismutases from a variety of aerobic organisms. Superoxide dismutases convert superoxide radical anions into molecular oxygen and hydrogen peroxide (6). Presumably, the PBCV-1-encoded enzyme reduces superoxide accumulation in sunlight. Two other large dsDNA-containing viruses, *Autographa californica* nuclear polyhedrosis virus (161) and vaccinia virus (10), also encode Cu/Zn superoxide dismutases.



PBCV-1 ORF A250R encodes a 94-amino-acid-residue protein that resembles a small  $K^+$  ion channel protein. This protein is apparently important for PBCV-1 growth, because the two classical  $K^+$  ion channel protein inhibitors, amantadine and  $Ba^{2+}$  ions, inhibit PBCV-1 plaque formation (M. Nelson, Y. Zhang, unpublished results).

**Gene Families or Gene Duplications** A total of 29 of the PBCV-1 ORFs resemble 1 or more other PBCV-1 ORFs, suggesting that they might be either gene families or gene duplications. A total of 16 families have two members, 8 families have three members, 3 families have six members, and 2 families have eight members. One family with six members contains ORFs with multiple ankyrin-like repeats, that is, G-TPLH-AA-GH-(V/A)-LL-GA-(N/A)-like sequences (114). Five members in another family resemble the major capsid protein of PBCV-1. One intriguing possibility is that, in its natural environments, PBCV-1 has an additional host(s) and that these capsid-like genes are important for infection of this other host(s).

### *PBCV-1 Deletion Mutants*

Not all of the PBCV-1 genes are required for virus replication in the laboratory. For example, four spontaneously derived, antigenic variants of PBCV-1 that contain 27- to 37-kb deletions at the left end of the 330-kb genome were isolated (69). Two of these mutants had deletions that began at map positions 4.9 kb or 16 kb and ended at position 42 kb. In total, the two deleted regions, which probably resulted from recombination, encoded 28 putative protein-encoding ORFs. The other two mutants, which probably arose from nonhomologous recombination, lacked the entire left-terminal 37-kb of the PBCV-1 genome, including the 2.2-kb inverted-repeat region. The deleted left terminus was replaced by the transposition of an inverted 7.7- or 18.5-kb copy from the right end of the PBCV-1 genome. The deleted regions encoded 26 single-copy ORFs, of which 23 were common to those deleted in the first two mutant viruses. Taken together, the results indicate that 40 kb of single-copy DNA encoding 31 ORFs at the left end of the genome, or 12% of the PBCV-1 genome, are unnecessary for PBCV-1 replication. However, replication of these mutants is attenuated because their burst sizes are about half those of the parent virus. Apparently the inverted terminal repeats in PBCV-1 can vary in size, and virus DNA packaging tolerates large changes in genome size. Similar results were obtained with NC64A virus CVK1 (145, 194). CVK1-infected cells exposed to UV radiation produced 30- to 45-kb deletion mutants. Like PBCV-1, these deletions occurred in the left terminal portion of the virus genome.

The sizes and locations of the deletions and deletion/transpositions found in the chlorella viruses resemble those of deletion mutants in the poxviruses (for a review, see 163) and ASFV (e.g. 9, 11). Like PBCV-1, poxviruses and ASFV have inverted terminal repeats and covalently closed hairpin ends. Mod-

els proposed to explain the generation of deletions and deletion/transpositions in the poxvirus genomes (140, 163) may be relevant to the chlorella viruses.

### *Viruses Encode DNA Methyltransferases and DNA Site-Specific Endonucleases*

The chlorella virus genomes contain different levels of 5mC and 6mA (Table 1). Therefore, it is not surprising that the viruses encode multiple 5mC and 6mA DNA methyltransferases (MTases). However, it was unexpected that the viruses also encode DNA site-specific (restriction) endonucleases. Thus the virus-infected chlorella are a source of DNA site-specific endonucleases and the first source from a nonprokaryotic system (for reviews, see 108, 109).

Some of the chlorella virus-encoded endonucleases have recognition and cleavage specificities identical to bacterial type II restriction endonucleases [e.g. *R.CviAI* (/GATC) (185) and *R.CviBI* (G/ANTC) (187)], others are heteroschizomers of bacterial endonucleases [e.g. *R.CviAII* (C/ATG) (205) and *R.CviQI* (G/ TAC) (189)], and still others have novel recognition sites [e.g. *R.CviJI* (RG/CY) (186) and *R.CviRI* (TG/CA) (56)]. Statistically, *R.CviJI* is the first site-specific endonuclease to recognize a three-base sequence, and, under certain conditions, called *R.CviJI*\*, the enzyme also cleaves RGCR and YGCY sequences (156, 186). Because *R.CviJI* and *R.CviJI*\* cleave DNA so frequently, the enzyme has novel uses. For example, partial digestion of DNA with *R.CviJI*\* has been used to produce DNA "shotgun" libraries (27, 34).

Two of the site-specific endonucleases [*NYs1*-nickase (/CC) (188) and *NY2A*-nickase (R/AG) (206)] cleave only one strand of dsDNA; that is, they are base-specific nicking enzymes. Like bacterial restriction endonucleases, all of the viral endonucleases are inhibited by either 5mC or 6mA in their recognition site.

The characteristics of the chlorella virus R-M systems can be summarized as follows: (a) Some viruses, such as NY-2A, encode as many as 10 different DNA MTases and at least two site-specific DNA endonucleases (206). The 380-kb NY-2A genome is predicted to contain 434 protein-encoding sequences. At a minimum, the NY-2A genome contains 12 R-M genes, or 1 of every 36 NY-2A genes is an R-M gene, making the NY-2A genome one of the most concentrated sources of R-M genes known. Even PBCV-1, with a low level of methylated bases (1.9% 5mC and 1.5% 6mA), has seven genes involved in methylation and cleavage: two 6mA MTases with corresponding site-specific endonucleases, two functional 5mC MTases, and one nonfunctional 5mC MTase. (b) Viruses with high levels of methylated bases encode some MTases that recognize short (2-to 3-bp) target sites, including *M.CviPI* from virus NYs-1, which recognizes GC sequences (191). (c) Not all MTases have a companion site-specific endonuclease. We estimate that 25–30% of the MTases have a cognate endonuclease. (d) Under laboratory conditions, at least some R-M genes can be deleted without affecting viral replication (13, 107). (e) Viral DNAs also contain nonfunctional 5mC MTase genes. For ex-

ample, PBCV-1 encodes a nonfunctional 5mC MTase that differs by six amino acids from a functional 5mC MTase *M.CviJI* present in virus IL-3A. Changing one of the six amino acids restored enzyme activity (207).

About 15 chlorella virus MTases, as well as three companion site-specific endonucleases, have been cloned and sequenced. Bacterial type II R-M genes (182) have many similarities to the chlorella virus R-M genes, including the following: (a) Bacterial R-M genes are always located near one another, although the spacing and relative orientation of the two genes can vary. The genes for the three sequenced virus R-M systems are also adjacent to one another. In the PBCV-1 *CviAI* system, the divergent *M.CviAI* and *R.CviAI* genes are separated by 82 bases (78), whereas, in the *CviAII* system, the TAA termination codon of the *M.CviAII* gene overlaps the ATG translation start site of the *R.CviAII* endonuclease gene (205). The *R.CviJI* endonuclease gene from virus IL-3A begins 18 bases after the *M.CviJI* gene stop codon (142, 157). (b) The virus 6mA- and 5mC-MTases contain the same highly conserved amino acid motifs found in bacterial 6mA (63, 84) and 5mC (67, 116) MTases. (c) Some of the virus 6mA MTases have significant similarity to bacterial DNA MTases. For example, *M.CviAI* has up to 39% amino acid identity to several bacterial 6mA MTases that also methylate GATC sequences, including *M.PgiI*, *M.DpnII*, and *M.LlaII*. *M.CviAII* (C<sup>m</sup>ATG) has 36% amino acid identity to *M.NlaIII* (C<sup>m</sup>ATG) and type-IIS *M.FokI* (GG<sup>m</sup>ATG and C<sup>m</sup>ATCC) (205). These similarities are significant because some bacterial isomethylator pairs that share the same recognition sequence have <20% amino acid identity (182).

The similarity between bacterial and viral genes suggests that bacteria and chlorella viruses have exchanged genes in the past. Consistent with this, most of the virus MTases cloned to date use their own promoters when expressed in *E. coli*. In addition, the five cloned NY-2A virus MTases have a sequence that resembles the bacterial -35 (TTGACA) promoter sequence (only one base different in each case). However, the similarity and the 16–19 nucleotide spacing to a bacterial consensus -10-like (TATAAT) sequence are poor for these five genes (206).

Although the evidence suggests that bacterial and chlorella virus MTases may have common evolutionary origins, the amino acid similarity data and the promoter sequence data suggest that the MTase genes have a long association with the chlorella viruses. Comparison of the G+C content of the five NY-2A MTase genes with total NY-2A DNA also supports this hypothesis. If the G+C contents of these five genes differed significantly from total NY-2A DNA, then the MTase genes were probably acquired relatively recently (3). In contrast, if the G+C contents are similar to the total genomic DNA, the genes have probably existed in the genome for a long time. The *m.Cviqi*, *m.Cviqiii*, *m.Cviqv*, *m.Cviqvi*, and *m.Cviqvii* genes have 38%, 37%, 42%, 38%, and 41% G+C contents, respectively, which are close to the 41.5% G+C content of the entire NY-2A genome (206).

Most, but not all, of the chlorella virus MTases that methylate the same sequence have ≥ 75% amino acid identity. For example, the 6mA C<sup>m</sup>ATG MTases, *M.CviAII*, *M.CviSII*, and *M.CviQVII*, from viruses PBCV-1, SC-1A, and NY-2A,

respectively, have 80–94% amino acid identity in pairwise comparisons (119, 206). However, exceptions occur. For example, *M.CviSI* and *M.CviQV*, both of which methylate TGC<sup>m</sup>A sequences, have only 32% amino acid identity (119, 206). Assuming that chlorella virus MTase genes were originally acquired from bacteria, it seems likely that these latter two genes were acquired independently.

The biological function(s) of the virus-encoded DNA site-specific endonucleases and MTases is unknown. Bacterial R-M systems confer resistance to foreign DNAs and DNA viruses. In fact, the term “restriction” refers to the role of these endonucleases and MTases in excluding foreign DNA. Bacterial MTases prevent self-digestion of host DNA. Two functions have been considered for the chlorella virus R-M enzymes. (a) Chlorella virus endonucleases help degrade host DNA, thus providing deoxynucleotides for recycling into virus DNA. Methylation of nascent virus DNAs by the cognate MTases protects the DNA from self-digestion. (b) Chlorella virus endonucleases prevent infection of a cell by a second virus.

Three observations are consistent with the first hypothesis. (a) Chlorella nuclear and chloroplast DNAs, but not virus DNA, are digested by the virus-encoded site-specific endonuclease(s) *in vitro*. (b) *In vivo* degradation of host nuclear and chloroplast DNAs coincides with the appearance of DNA site-specific endonuclease activity (185). (c) Initiation of virus DNA synthesis *in vivo* coincides with the appearance of DNA MTase activity (190).

The isolation of three independently derived deletion mutants of virus IL-3A, which had lost their MTase (*M.CviJI*) and site-specific endonuclease (*R.CviJI*) activities, allowed us to test the host DNA degradation hypothesis directly (13). If *R.CviJI* activity was essential for host DNA degradation, nuclear and/or chloroplast DNA should be preserved or at least degraded more slowly in cells infected with the three mutants than in cells infected with wild-type IL-3A. However, both nuclear and chloroplast DNA levels decreased at nearly identical rates after infection with each of the viruses (13). Therefore, we concluded that *R.CviJI* activity was not essential for host DNA degradation. However, this finding does not exclude participation of the enzyme in the degradation process.

To determine whether the endonuclease(s) are involved in excluding infection of a cell by a second virus, *Chlorella* cells were inoculated with pairs of viruses and plaques arising from infective centers were distinguished by immunoblotting (15). These experiments revealed that chlorella viruses, like certain bacteriophage (e.g. 24), exclude one another. However, this exclusion was independent of the known site-specific endonuclease activities (15).

Therefore, the biological functions of the DNA MTases and DNA site-specific endonucleases are unknown. However, if a virus contains a functional 5mC MTase gene, then its expression is apparently required for virus growth, because there is a direct correlation between increasing 5mC concentrations in the virus genome and sensitivity of virus replication to 5-azacytidine (13). It should also be noted that many viral DNAs are more heavily methylated than is necessary for protection from their site-specific endonucleases. The reason for this apparent excess methylation is unknown.

## *Glycosylation of PBCV-1 Structural Proteins Is Unusual*

Structural proteins of many viruses, such as herpesviruses, poxviruses, paramyxoviruses, and orthomyxoviruses, as well as PBCV-1, are glycosylated. Viral glycoproteins are involved in various functions, including virus assembly, release of newly formed virus particles, and attachment to susceptible hosts. Viral proteins are generally glycosylated by host-encoded glycosyltransferases located in the rough endoplasmic reticulum and Golgi apparatus (25, 130). The glycoproteins are then transported to a host membrane and the nascent viruses acquire these glycoprotein(s) by budding through the membrane, usually as they are released from the cell. Consequently, the glycan portion of virus glycoproteins is generally specific to the host (e.g. 25, 120, 130).

PBCV-1, however, encodes at least part, if not all, of the machinery required to glycosylate its structural proteins. The initial observation that led to this conclusion arose from antibody studies. Polyclonal antiserum prepared against intact PBCV-1 virions completely inhibits PBCV-1 plaque formation by agglutinating the particles. However, spontaneously derived antiserum-resistant, plaque-forming mutants of PBCV-1 occur at a frequency of about  $10^{-6}$  (179). These antiserum-resistant mutants fall into four serologically distinct groups (117, 179). Polyclonal antisera prepared against each of these four antigenic variants reacted exclusively with the immunizing virus. The major capsid protein Vp54 as well as the two other viral glycoproteins Vp280 and Vp260 or their equivalents from the mutants migrated in a distinctive, but predictable fashion on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Western blot analyses of the mutant capsid proteins, before and after deglycosylation, revealed that the antigenic differences reflected differences in the carbohydrate moiety(ies) of the major capsid protein and the two minor glycoproteins. The ratio of 6 and/or 7 sugars (fucose, galactose, glucose, xylose, mannose, and arabinose and/or rhamnose) associated with the major capsid protein of PBCV-1 and the mutants also varied in a predictable manner related to their serology and the migration of the major capsid protein (179).

Two other observations are relevant. First, unlike many viruses that acquire their glycoprotein(s) by budding through a membrane, intact infectious PBCV-1 particles accumulate inside the cell 30–40 min before virus release (170). Second, even though the PBCV-1 major capsid protein contains ~10% carbohydrate, the protein appears homogeneous on sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and it has a single mass peak at 53,790 (R. Cerny, J. L. Van Etten, unpublished results). This is unusual because glycoproteins are typically micro-heterogeneous.

Because all of the PBCV-1 antigenic variants were grown in the same host and the predicted amino acid sequences of the major capsid protein from PBCV-1 and two of the antigenic variants were identical (179), glycosylation of chlorella virus proteins differs from other viruses. A simple explanation is that the antigenic variants arose from mutations in a common PBCV-1-encoded pathway involved in glycosylation, presumably a family of glycosyltransferases. Obviously, identification of the PBCV-1 glycosylation genes is

a high priority. Unfortunately, methods for tagging specific PBCV-1 genes are currently unavailable. As an alternative strategy, we isolated additional PBCV-1 antigenic variants, anticipating that some of these mutants might contain easily detected deletions in common DNA restriction fragments, thus locating some of the glycosylation genes.

This strategy resulted in the isolation of four antigenic variants of chlorella virus PBCV-1, which contained 27- to 37-kb deletions in the left end of the 330-kb genome (these deletion mutants were described above). Analysis of these mutants suggested that one of the five putative genes located between map positions 37 and 42.2 is involved in glycosylation and that another gene should be in the first 37 kb of DNA. It is interesting that the 638-codon ORF A64R at about map position 35 has some similarity with portions of a galactosyltransferase from yeasts (M. V. Graves, C. Berndt, J. L. Van Etten, unpublished results). Furthermore, ORF A64R contains short motifs present in glycosyltransferases from other organisms (202), indicating that the *a64r* gene product might be responsible for converting one virus phenotype to another. We will test this prediction once techniques for doing gene disruption and gene replacement experiments are available.

## Natural History of the Chlorella Viruses

Several unexpected discoveries about the diversity of the chlorella viruses and their genes are mentioned in this section. The natural history of these viruses is poorly understood, and major advances are needed in our understanding of the biology of these viruses before we can appreciate the ecology and explain the evolutionary origin of these viruses.

The two known hosts for the chlorella viruses, *Chlorella* NC64A and *Chlorella* Pbi, normally exist as hereditary endosymbionts in green isolates of the protozoan *P. bursaria*. However, the zoochlorella are resistant to virus infection when they are in the symbiotic relationship and are infected only when they are separated from the ciliate (124). In the symbiotic unit, algae are enclosed individually in perialgal vacuoles surrounded by a host-derived membrane (122). The initial establishment of a successful symbiotic relationship and the long-term maintenance of symbiosis require that the algae resist digestion by the paramecium. Reassociation studies with different *Chlorella* species and alga-free *P. bursaria* indicate that only the original symbiotic algae readily reestablish symbiosis with the ciliate (122). Other chlorellae are digested. Although there have been numerous studies and considerable speculation on the factor(s) that allow the *P. bursaria* to distinguish chlorellae that are suitable for symbiosis formation from those that are not, the specific recognition factor(s) is unknown.

*Chlorella* NC64A and *Chlorella* Pbi were originally isolated from American and European strains, respectively, of *P. bursaria*. Viruses that infect *Chlorella* NC64A neither infect nor attach to *Chlorella* Pbi, and Pbi viruses neither infect nor attach to *Chlorella* NC64A (124). Because the viruses can distinguish



the two *Chlorella* isolates, we hypothesized that the receptor for the viruses might also serve as the recognition factor for the paramecium. This hypothesis was incorrect, however, because both *Chlorella* NC64A and *Chlorella* Pbi reestablish stable symbiotic relationships with either American or European isolates of *P. bursaria* (124).

NC64A viruses have been isolated from fresh water collected in the United States (133, 169, 177), China (204), Japan, Brazil (195), Australia (J. L. Van Etten, J. Rohozinski, unpublished results), Argentina, and Israel (Y. Zhang, M. Nelson, J. L. Van Etten, unpublished results). Pbi viruses were initially found in fresh water collected in Europe (123, 125) and more recently in water collected in Australia, Canada, and either the northern United States (Minnesota, Wisconsin, and Montana) or in higher altitudes in the western United States (M. Nelson, Y. Zhang, J. L. Van Etten, unpublished results). Therefore, the initial assumption that Pbi viruses were limited to Europe and the NC64A viruses were present only in the Americas and Eastern Asia is incorrect. In fact, two water samples collected in Australia contained both NC64A and Pbi viruses (J. L. Van Etten, J. Rohozinski, unpublished results). The most important factors influencing the distribution of NC64A and Pbi viruses are probably latitude and altitude.

Typically, the *Chlorella* virus titer in nature is 1–100 PFU/ml, but titers as high as 40,000 PFU/ml have been obtained. The natural concentration of *Chlorella* viruses is not static, but fluctuates with the seasons; the highest titers have typically been found in late spring (177, 195). However, a water sample was recently collected through the ice on a pond in Ohio, with a titer of >1000 PFU/ml (M. Nelson, unpublished results).

It is not known whether the zoochlorellae can exist outside the paramecium in natural environments. Consequently, we do not even know whether these zoochlorellae are the normal virus hosts in native waters; it is possible that NC64A and Pbi viruses replicate in another host(s). We anticipated that some of the PBCV-1 encoded proteins might help identify another host(s), if it exists. As noted above, PBCV-1 encodes three enzymes that are involved in the synthesis of hyaluronan, a glycosaminoglycan. Furthermore, hyaluronan appears as hairlike material on the external surface of the infected algae (M. V. Graves, J. Heuser, D. E. Burbank, P. DeAngelis, J. L. Van Etten, published results, 41a), suggesting that the polysaccharide plays an important role in virus survival. We have considered two biological explanations for the extracellular hyaluronan: (a) The polysaccharide prevents uptake of virus-infected zoochlorellae by *P. bursaria*. Presumably, such infected algae would lyse inside the paramecium, and the released virions would be digested by the protozoan, which would be detrimental to virus survival. (b) The viruses have another host that acquires the virus because it is attracted to or binds the hyaluronan on the virus-infected algae. A complicating factor in understanding the biological importance of the hyaluronan is the recent discovery that not all infectious NC64A viruses encode hyaluronan synthetase (M. V. Graves, D. E. Burbank, P. DeAngelis, J. L. Van Etten, published results, 41a). However, all *Chlorella* viruses encode the GFAT and UDP-GlcDH enzymes that synthesize hyaluronan precursors.

The high titers of the chlorella viruses in some indigenous waters are surprising given that the viruses are constantly exposed to solar radiation. This radiation should damage viral DNAs, inactivating the virus. For example, inactivation of bacteriophages and cyanophages occurs at rates of  $0.4\text{--}0.8\% \text{ h}^{-1}$  in full sunlight (154, 155). The chlorella viruses have apparently adapted to solar radiation by having access to two independent DNA repair systems (30). (a) PBCV-1 encodes a pyrimidine dimer-specific glycosylase (PDG) that initiates UV-induced thymidine dimer repair. This DNA repair system functions in both the light and the dark. (b) PBCV-1 also uses the host photolyase to repair UV-induced thymidine dimers. Thus, PBCV-1 can replicate whenever suitable hosts are encountered, both in the day and at night. We expected the PDG enzyme to be packaged in the virion and accompany the virus DNA into the host, where it could initiate DNA repair. However, attempts to detect the PDG protein in PBCV-1 virions were unsuccessful (30). The *pdg* gene is expressed early after virus infection.

PBCV-1 also encodes two other enzymes that may aid survival in direct sunlight. (a) The virus encodes a putative Cu/Zn superoxide dismutase that should protect DNA from reactive oxygen species. (b) A recent report (45) indicates that the polyamine spermine functions as a free-radical scavenger, thereby reducing DNA damage. Possibly this is why PBCV-1 encodes two polyamine biosynthetic enzymes, as well as packages the polyamines in the virions.

The PBCV-1-encoded PDG repair enzyme is also interesting for another reason. As mentioned above, the enzyme cleaves both *cis-syn* and *trans-syn*-II cyclobutane pyrimidine dimers, whereas its bacteriophage T4 endonuclease V homolog cleaves only the *cis-syn* cyclobutane pyrimidine isomer (89). These differences prompted the search for *pdg* genes from 41 other NC64A viruses, which were isolated from diverse geographic regions. All the chlorella viruses contained the *pdg* gene (30; L. Sun, S. R. Lloyd, J. L. Van Etten, unpublished results). Unexpectedly, however, *pdg* genes from 15 of the 41 viruses contained a 98-nucleotide intron that is 100% conserved among the viruses, and another 4 viruses contained an 81-nucleotide intron in the same position that is nearly 100% identical (one intron differed by one nucleotide). Surprisingly, the *pdg* gene protein-coding regions (exons) were less conserved than their corresponding introns. The introns in the *pdg* gene have 5'AG/GTATGT and 3'-TTGCAG/AA splice site sequences, which are characteristic of nucleus-located, spliceosomally processed pre-mRNA introns.

There is no obvious geographic correlation between the *pdg* intron-containing and intron-lacking viruses. Of the 98-nucleotide intron-containing viruses, 13 were collected throughout the United States in 1983–1985. Viruses isolated from Australia in 1995 and Argentina in 1996 contained an identical intron. Fortunately, some water samples had both 98-nucleotide intron-containing and intron-lacking viruses. For example, water samples collected in Massachusetts, North Carolina, Alabama, Illinois, and California had both intron-containing and intron-lacking viruses. The 100% identity of the 98-nucleotide intron sequence in 15 viruses and the near 100% identity of an 81-nucleotide intron sequence in another 4 viruses imply that either the intron was

acquired recently or there is strong selective pressure to maintain the DNA sequence of the intron once it is in the *pdg* gene. However, the abilities of intron-containing and intron-lacking viruses to repair UV-damaged DNA in the dark are indistinguishable (L. Sun, J. Gurnon, S. R. Lloyd, J. L. Van Etten, unpublished results). These findings contradict the widely accepted dogma that intron sequences are more variable than exon sequences.

Something similar to the sporadic occurrence of a highly conserved intron in the *pdg* gene is also observed with a self-splicing intron. The PBCV-1 TFIIS-like gene contains a 400-nucleotide self-splicing group IB intron (77). Group IB introns were originally discovered in two related NC64A viruses, CVU1 and CVB11 (198). The intron in CVU1 has 98% nucleotide identity and is in the exact same position in a TFIIS-like gene as the PBCV-1 intron, whereas the CVB11 intron has 80% nucleotide identity and is in an unidentified open reading frame. Three other NC64A viruses isolated in Japan at the same time as CVU1 and CVB11 lacked the intron. It is fascinating that two NC64A viruses, PBCV-1 isolated in the United States in 1981 and CVU1 isolated in Japan in about 1990, have a nearly identical intron located at the same place in the same gene, whereas this same intron is either located in a different gene (virus CVB11) or absent in three other Japanese chlorella viruses (198).

Recently, several hundred additional NC64A viruses, isolated in Japan, have been screened for the self-splicing intron; the intron is present in ~8% of these isolates (111). The intron was inserted in the TFIIS-like gene in ~60% of these viruses, and ~40% had the intron in the same unidentified reading frame gene as CVB11. In a few of the viruses, the major capsid protein gene contained the intron, and a couple of viruses had two copies of the intron. Nucleotide sequence analysis of the introns and the flanking regions indicated that the intron sequences are apparently under strong constraint by the exons, i.e. introns in the same gene had >99% sequence identity, whereas introns in different genes were only 72–78% identical.

The various levels of methylated bases observed in the chlorella virus genomes are another example of the natural diversity of these viruses that is difficult to explain with our current understanding of their natural history. As noted above the methylated base levels of virus genomes range from 0.1% 5mC and undetected levels of 6mA to 45% 5mC and 37% 6mA. The concentrations of 5mC and 6mA are a criterion for grouping the viruses into different classes (Table 1). Like the situation with the introns, there is no obvious geographic correlation for methylation levels in virus DNAs. Typically viruses isolated from one water sample have similar levels of methylated bases. However, exceptions are common. For example, four of six plaques originally picked from a water sample collected in New York fell into different classes. The level of methylation ranged from virus NY-2C, which contained 0.4% 5mC and no detectable 6mA, to NY-2A, which contained 45% 5mC and 37% 6mA.

Presumably the virus-encoded DNA methylation and site-specific endonuclease enzymes confer an advantage to the viruses in their native environment. We have conducted one simple experiment to determine whether the virus-encoded enzymes confer an advantage in the laboratory. Cultures of

*Chlorella* NC64A were inoculated with various ratios [multiplicity of infection (MOI) of 0.1–1, 1–1, and 1–0.1] of virus IL-3A and an IL-3A mutant deficient in its R-M system, R.CviJI and M.CviJI (one-step growth curves for the two viruses are identical). Fresh cultures of the algae were inoculated with portions of the lysates and, after five serial transfers, the lysates were plaqued. In most of these experiments 80–95% of the plaques were wild-type virus (C. Weinfeldt, J. L. Van Etten, unpublished results). Thus M.CviJI and R.CviJI appear to confer an advantage to the virus under these conditions.

Although the PBCV-1 genome is the only chlorella virus to be completely sequenced, a detailed physical map of virus CVK2 is available (K. Nishida, Y. Kimura, T. Kawasaki, M. Fujie, T. Yamada, published results, 110a), and portions of several other NC64A viruses have been sequenced. Comparisons among these viruses have led to the unexpected finding that “anonymous” genes are often either inserted or deleted between otherwise colinear genes. This phenomenon, which has been observed by several laboratories, is illustrated with two examples. (a) As described above the chitosanase gene, which is a late gene, has been characterized from NC64A virus CVK2 (196). PBCV-1 encodes the same chitosanase gene, which is flanked by two ORFs that also exist in CVK2. However, the CVK2 genome has an extra ORF inserted immediately downstream of its chitosanase gene, which is absent in PBCV-1. (b) Likewise, an extra ORF is inserted between two otherwise colinear genes in NC64A virus SC-1A, which is absent in PBCV-1 (118). We do not have an explanation for the apparent appearance and/or disappearance of these viral genes, but this finding indicates that the total number of genes encoded by the chlorella virus group exceeds that of any one isolate. In many ways, this phenomenon resembles the presence or absence of the introns in certain virus genes.

## Characteristics of Marine Algal Viruses

### *Host Algae*

There are predicted to be >100,000 species of marine algae of which only 30,000 have been identified. Photosynthetic plankton alone produce 40% of the photosynthate on the planet. It is estimated that these organisms produce  $10^{12}$  tons of cell wall material yearly, of which  $10^{11}$  tons are cellulose (38).

Despite the importance of marine algae, pathogens, including viruses and virus-like particles, are only beginning to be studied. A slowly growing literature indicates that eukaryotic marine algal viruses are large dsDNA viruses, similar to the chlorella viruses. Laboratory infection of marine algae by viruses has been demonstrated only in single-celled *Micromonas pusilla* (19, 20) and in filamentous *Sorocarpus* sp. and *Ectocarpus* sp. (104, 112). Consequently, not much is known about the infection process. Infection occurs in *Ectocarpus siliculosus* only when the host cell wall is mechanically damaged or lacking (104; R. H. Meints, unpublished results). Only viruses of the brown alga genera *Ectocarpus* and *Feldmannia* have been studied experimentally. These vi-

ruses infect uniseriate, branched, filamentous forms in the order *Ectocarpales* (5, 98). The biology of these virus-algal systems is interesting because the viral genomes are inherited in a Mendelian fashion and virus particles appear only in the reproductive organs of the algae. In general, virus infections do not affect vegetative growth of the brown algae.

***Ectocarpus* Virus** Members of the genus *Ectocarpus* inhabit temperate sea-coasts throughout the world. The most studied member, *E. siliculosus*, has a complex sexual life cycle consisting of alternating generations. In the most parsimonious version of its life cycle, the haploid growth phase begins with the formation of male and female filamentous gametophytes from single-celled meiospores. The gametophytes develop distinctive, multicellular structures called gametangia, which produce motile isogametes, discharging them into the seawater. After fusion of male and female isogametes into diploid zygotes, the zygotes grow into diploid filamentous sporophytes that develop sporangia of two types, plurilocular and unilocular. Each compartment of the plurilocular sporangia can produce a diploid spore that, upon germination, produces another sporophyte. The nucleus in the unilocular sporangia undergoes meiosis, followed by multiple rounds of mitosis, yielding 100 or more haploid meiospores per unicellular sporangium. The motile meiospores are discharged into the seawater, and each one develops into either a male or female gametophyte. The textbook version of this life cycle is probably rare in nature, with more complex paths being common. As an example, unfertilized gametes can also develop into haploid parthenosporophytes that can propagate independently, and polyploid forms are common (97).

In 1990, Müller & colleagues (104) reported that large (130- to 140-nm diameter), densely packed, polyhedral virus particles (EsV) appeared in the gametangia and both unilocular and plurilocular sporangia of an isolate of *E. siliculosus* but not in vegetative cells. Because the virus particles displaced the normal reproductive organs, the plants were sterile and could propagate only vegetatively. Occasionally, the virulence of the virus infection was reduced, and these plants produced functional meiospores and gametes. Genetic evidence suggested that the EsV genome was maintained in these meiospores and gametes and was transmitted to daughter generations. Fusion of a gamete from an infected plant with a gamete from an uninfected plant resulted in a sporophyte that formed normal sporangia-containing meiospores. These individual meiospores gave rise to healthy and infected gametophytes in a 1:1 ratio, strongly suggesting that the virus symptoms and sex alleles segregated independently (99). This same experiment was repeated recently with EsV-specific DNA probes, and the same results were obtained, that is, only 50% of the progeny gametophytes contained viral DNA (12). These two experiments suggest that one copy of the EsV genome is associated with host chromosomes. Furthermore, it means that EsV and its host can coexist in a nonlethal and pandemic manner because meiosis in heterozygotes eliminates virus from 50% of the progeny.

The appearance of virus particles and their release from reproductive organs can be induced by increasing the temperature from 12°C to 18–20°C, transferring the alga to fresh culture medium, or reducing salinity of the culture medium (99, 102). The viruses do not appear to be released from reproductive cells by virus-induced lysis. However, both sporangia and gametangia are designed to open and release their contents, typically zoospores; thus these cells might be considered ideal sites for virus replication. Good quantities of EsV can be produced in culture,  $1.7 \times 10^{13}$  to  $3.2 \times 10^{13}$  virus particles g<sup>-1</sup> fresh weight of infected algae (99). The intracellular site of virus replication is unknown, but presumably replication begins in the nucleus. Capsid assembly and DNA packaging occur in the cytoplasm, but only after degeneration of the nuclear membrane (104). The viral assembly process in a related alga, *Hincksis hinckisia*, has recently been described in a series of elegant electron and fluorescence micrographs (184).

Purified EsV can infect motile gametes or meiospores of *E. siliculosus* that lack a cell wall, but not cell-wall containing vegetative cells (104). When gametes or meiospores swim into a cloud of freshly released virus particles, flagellar activity stops, and the paralyzed cells settle to the bottom. These infected cells develop into normal-appearing filamentous plants except that the reproductive organs are filled with virus particles rather than spores. The EsV host range is not limited to *E. siliculosus*, however, because the virus can infect motile gametes of at least two other filamentous brown algae. Infection of zoospores of a *Kuckuckia* sp., which is closely related to *Ectocarpus* sp., produces infected plants that resemble those of EsV-infected *Ectocarpus* sp., that is, virus particles form only in reproductive organs (100). The virus particles grown in *Kuckuckia* sp. can infect healthy *Ectocarpus* zoospores; thus EsV maintains infectivity to *Ectocarpus* sp. after replication in another host.

EsV can also infect zoids (zoospores) of the filamentous brown alga *Feldmannia simplex*. [Other unrelated *Feldmannia* species have their own viruses (see below).] EsV infection of *Feldmannia simplex* resulted in deformed plants, including cells containing abnormal mitochondria and chloroplasts, multiple nuclei, and multiple vacuoles. Intact EsV particles are absent in these infected plants (105). The symptoms produced by EsV infection of *F. simplex*, which differ from those produced by genuine *Feldmannia* viruses (see below), persisted as the filaments continued to grow. Occasionally, a few filaments produced normal-appearing multicellular zoidangia, and the resultant zoids developed into healthy-appearing plants. Surprisingly, however, these apparently healthy *F. simplex* plants contained EsV DNA, even after 2 years of growth (101). The ability of EsV to infect several algae means it could serve as a vector for nonsexual gene transfer among organisms, as suggested by Reaney (121). There are many examples of gene transduction in bacteria, but there is much less information on viruses as potential vectors for DNA transfer among eukaryotes.

Many *E. siliculosus* isolates and related *Ectocarpus* sp. collected from various ocean coasts throughout the world contain virus particles similar to EsV (106, 113, 137; R. H. Meints, unpublished results). Thus these viruses consti-



tute an inherent element in *Ectocarpus* sp. and other brown algal populations, with implications for their genetics and evolution.

The characterization of the EsV virus particles is just beginning. EsV particles have an outer shell with two layers surrounding an electron-dense core; the thin outer layer surrounds a broader, more diffuse inner layer (61). EsV particles contain at least 16 proteins, of which at least 3 are glycoproteins. The gene for one of the EsV glycoproteins has been sequenced; the gene is predicted to encode a protein of 72 kDa, which is larger than the 60-kDa capsid protein. This observation suggests proteolytic processing (61).

The EsV genome is a large, 320-kb, circular dsDNA molecule with a 50% G+C content (72). The dsDNA is interrupted by multiple single-stranded regions of various lengths, most of which are randomly distributed over the entire length of the DNA (62). The single-stranded regions, which may result from packaging incompletely replicated DNA, make the DNA very fragile. Like the chlorella virus DNAs, the EsV DNA has methylated bases, 1% of the cytosines are 5mC, and 3% of the adenines are 6mA (72).

**Feldmannia Virus** One isolate of the genus *Feldmannia* for which uninfected algae have not been found has been designated *Feldmannia* sp. In 1992 Henry & Meints (50) described a virus, FsV, in an isolate of this marine brown alga that had originally been isolated from the New Zealand coast in 1984. The alga was maintained in the laboratory for >7 years before virus studies were initiated. This 120- to 150-nm-diameter, icosahedral dsDNA virus was present exclusively in meiotic unilocular sporangium cells. Each single-celled reproductive sporangium contains  $1 \times 10^6$ – $5 \times 10^6$  virions. Virions have not been detected in vegetative cells of either the sporophyte or gametophyte and, in contrast to *Ectocarpus* sp., are not produced in the plurilocular sporangia of the gametophyte or sporophyte. FsV particles appear to be released by rupture of the sporangium wall, probably by the normal mechanism for spore release. Like *Ectocarpus* sp., one would expect diploid sporophytic plants to develop both plurilocular (mitotic) sporangium-producing diploid spores and unilocular (meiotic) sporangium-producing haploid spores, whereas haploid gametophytes produce only plurilocular gametangia whose spores are the gametes for the sexual cycle.

In the FsV-infected plants, this life cycle is altered. The biologically important characteristic of the *Feldmannia* system is that only single-celled sporangia from virus-infected sporophyte plants fail to develop normally, and meiospores are not produced. Instead each unilocular sporangium is filled with virus particles, rendering these organisms asexual. Viruses are not produced in the gametophyte generation but appear to be latent in this generation; in contrast, virions are produced in the sporophyte by the apparent excision of the viral genome from the host genome by an as-yet-unknown mechanism, followed by viral genome replication. Once viral production begins, death of the cell is a certainty.

Analysis of marine algal viruses at the molecular level is preliminary, and a summary of the information on FsVs follows.

1. Like EsV particles, purified FsV particles have a capsid with two layers surrounding an electron-dense core; the thin outer layer surrounds a broader, more diffuse inner layer (50).
2. Clamped homogeneous electric field electrophoresis of DNA from purified FsV reveals two genome size classes (158 and 178 kb). These genomes are variably abundant in preparations from cultures grown at 18°C, whereas the smaller genome is virtually absent in cultures grown at 9°C (50).
3. Restriction maps of both FsV genome size classes and a subvariant of each map as circles (55). Mapping and hybridization studies indicate that genomes of the two size classes are similar and that the major difference in the two size classes results from the number of copies of a recurring 173-bp repeat element (74).
4. Characterization of a 4.6-kb region that was common to all genome size class variants revealed that three ORFs hybridized to Northern blots of RNA isolated from virus-producing sporophytes (55). ORF 1 encodes a polypeptide of 75.4 kDa and contains regions with significant homologies to a "RING" zinc finger motif and a nucleotide-binding site (65). RING zinc finger motifs are also found in other viruses, for example, poxviruses, where they are critical for virulence (138), and *Ectromelia* viruses, where they are involved in DNA replication (139). Because of its Ring signature, we speculate that this FsV protein participates in protein-protein interactions.
5. Transcriptional studies with four FsV genome-spanning cosmid probes on northern blots revealed 6 major and 18 minor transcripts (Y. Jia, R. H. Meints, unpublished results). One ORF has significant similarity to the major structural protein of chlorella virus PBCV-1 (42), the major structural protein of several iridescent viruses (28, 85, 132, 151), and the capsid protein of ASFV (81). This homology suggests that FsV is closely related to PBCV-1 and at least distantly related to these invertebrate and vertebrate viruses.
6. The FsV DNA polymerase gene was cloned and sequenced. The ORF encodes 986 amino acids and contains all of the conserved 3' to 5' exonuclease domains and catalytic domains found in B-family ( $\alpha$ -like) DNA polymerases. The codons for the FsV DNA polymerase have some bias towards G/C in the third position. At the 3' end of the gene, there is a TTTTNT sequence motif; this motif serves as a transcription termination signal for vaccinia virus early genes (203).
7. A lambda library prepared from a *Feldmannia* haploid gametophyte that contained no visible virus particles carries 1–5 copies of the viral genome presumably integrated into the host chromosomes. A second family of 173-bp repeats, related to but different from those mentioned above, was inserted into a viral ORF that contains all 12 catalytic motifs conserved in most Ser/Thr protein kinases and a potential autophosphorylation site

(75a). Perhaps this kinase functions in diverting spore formation to virus production.

8. Despite more than 100 sequential serial passages of our culture lines, we have not isolated a gametophyte free of the viral genome. This observation, along with inheritance evidence for EsV-1 (99), encouraged us to search for the sites of FsV integration into the host genome. We predicted that the point of diversity between the known sequence of the integrated virus and unknown host sequence would be the integration junction of host and virus DNAs. Such junctions were found, and DNA sequencing revealed that the same site is used for integration of both viral size classes, but that the host chromosomal integration sites varied (RG Ivey, RH Meints, unpublished results). Furthermore, recombination appears to occur without rearrangement, analogous to *Int-att*, *CRE-lox*, and yeast  $2\mu$ m *FLP-*frt** (21).

Together the data suggest an integration/excision mechanism that uses an integrase/recombinase and conservative site-specific recombination. Unlike in previously described systems, we expect blunt-end cutting and ligation or a single-base-pair overlap with specificity residing in flanking sequences.

## Evolutionary Relationship of the Algal Viruses to Each Other and to Other Viruses

The chlorella viruses, the brown algal viruses, and the micromonas viruses have several common properties. (a) They are large icosahedral particles (130–190 nm in diameter). (b) They have large dsDNA genomes (160–380 kb). (c) The genomes of many contain methylated bases. (d) Capsid assembly and DNA packaging occur in the cytoplasm. (e) The virus-host systems are global in nature. The chlorella viruses have been found in fresh water from at least five continents, and the viruses infecting the two groups of marine algae have been found in tropical and subtropical seawater.

However, differences exist between the chlorella viruses and the brown algal viruses as well. (a) The chlorella viruses enter their single-celled host by digesting its cell wall at the point of attachment, whereas brown algal viruses have no obvious lytic activity and infect only algal cells in a stage of host development during which these cells lack a cell wall. (b) The 320- to 380-kb chlorella virus genomes are linear, nonpermuted molecules with cross-linked hairpin ends. In contrast, brown algal viruses have circular dsDNA genomes. (c) Most importantly, the brown algal viruses have a lysogenic phase, whereas there is no evidence that the chlorella viruses have integrated genomes. We do, however, occasionally observe a carrier state or pseudolysogeny state with the chlorella viruses in the laboratory (Y. Xia, J. L. Van Etten, unpublished results).

Despite the differences between the chlorella viruses and brown algal viruses, phylogenetic trees with the DNA polymerase gene (16, 75) indicate that the algal viruses, including viruses that infect *Micromonas pusilla* and *Chrysochromulina* spp., are more closely related to each other than to other dsDNA viruses and that they form a distinct clade, suggesting a common, albeit ancient, ancestor (Figure 3). However, the algal viruses fall into several clades, which correlate with their algal hosts. The chlorella viruses have family status with the name *Phycodnaviridae* (165). It has been proposed that the brown algal viruses and the micromonas viruses be included in the *Phycodnaviridae* family and that the groups be assigned separate genera. Based on DNA polymerase gene sequences, herpesviruses are the most closely related out group (16). However, phylogenetic analyses of other genes such as the major capsid protein genes suggest that the *Phycodnaviridae* may be most closely related to the iridoviruses (R. H. Meints, unpublished results).

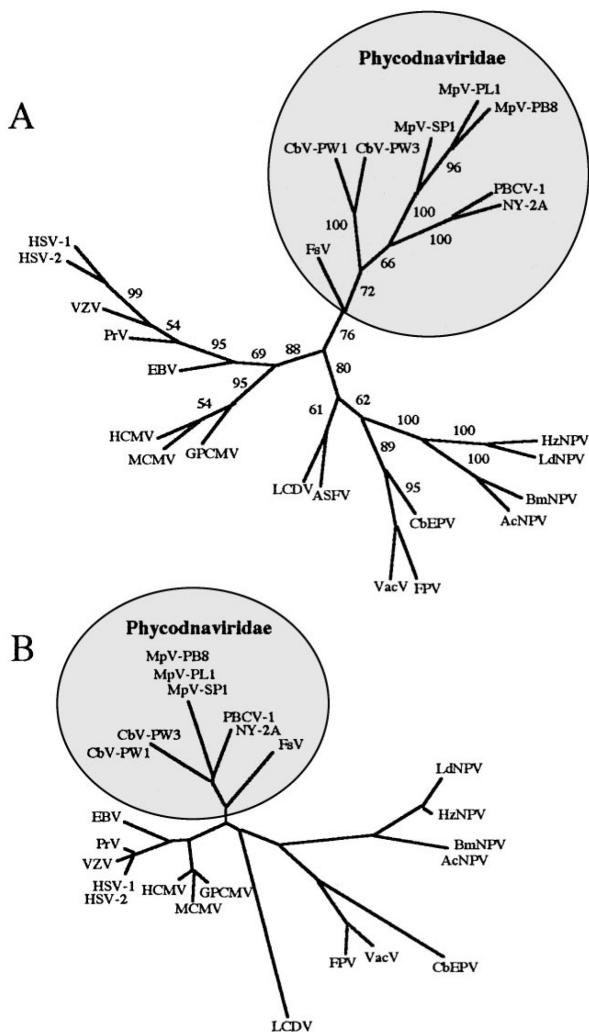
Algal viruses share several properties with certain other large dsDNA-containing viruses. Iridoviruses, which commonly infect insects and some animals, have the following properties in common with the chlorella viruses: aquatic habitats, icosahedral morphology, a large dsDNA genome, and an internal lipid component that makes up 5–10% of the viral weight (39). Frog virus 3 (FV3), the most widely studied iridovirus, is one of the few viruses that infects a eukaryotic organism and contains methylated bases; 20% of the cytosines in the FV3 genome are methylated (181). FV3 also encodes a 5mC DNA methyltransferase (59, 180). Furthermore, phylogenetic trees constructed with the major capsid proteins from several iridoviruses indicate that the iridoviruses are distantly related to the algal viruses (85, 86, 160). However, there are many differences between FV3 and PBCV-1 (172). Most importantly, the structures of their DNAs differ. FV3 DNA is linear and circularly permuted (40), whereas PBCV-1 DNA is linear and nonpermuted, with covalently closed hairpin ends, and brown algal viruses have a circular genome.

The PBCV-1 genome shares two characteristics with the vaccinia virus (a poxvirus) genome, even though the virions are morphologically distinct. Like PBCV-1, vaccinia virus DNA contains covalently closed hairpin termini and inverted terminal repetition (32, 33). However, the terminal repetitive region of vaccinia virus DNA is larger (~10 kb) than that of PBCV-1 DNA (2.2 kb), even though the PBCV-1 genome is about twice the size of the vaccinia virus genome. The vaccinia virus genome has many identical direct tandem repeats of 50–125 bases in its termini (183). Although PBCV-1 termini contain several direct repeats, with the exception of 118- and 39-base repeats, they are smaller and less extensive (152).

ASFV (presently unclassified) is the only other large polyhedral virus containing a dsDNA genome with hairpin termini and inverted terminal repetition (9, 36, 144). However, the chlorella viruses differ from ASFV in several aspects: (a) PBCV-1 DNA is about twice the size of ASFV DNA. (b) PBCV-1 DNA contains methylated bases, whereas ASFV does not. (c) ASFV lacks glycoproteins. (d) ASFV infection occurs via endocytosis, whereas the chlorella

viruses uncoat at the cell surface. Comparisons of ASFV-encoded and PBCV-1-encoded ORFs indicate low similarity between the two viruses. For example, even though these two are the only viruses infecting a eukaryotic host that encode a type-II DNA topoisomerase, there is only 23% amino acid identity between the two enzymes.

The infection process of the chlorella viruses differs from that of all DNA viruses known to infect eukaryotes, but resembles that of bacterial viruses in that uncoating occurs at the cell wall. In terms of structure, attachment, and penetration, the chlorella viruses resemble bacteriophages of the *Tectiviridae* family (94). However, tectiviruses are much smaller (~65 nm in diameter) and contain only ~40 kb of DNA.



Finally, it should be noted that a forthcoming review by Villarreal (178) suggests that DNA viruses have played a significant role in the evolution of their hosts by contributing virus genes to the host. In his review, he specifically focuses on the origin of genes encoding enzymes involved in DNA replication, and he suggests that the algal virus DNA polymerase might be related to the original progenitor of eukaryotic delta DNA polymerases. This conclusion supports the notion that the algal viruses have a long evolutionary history.

## Concluding Remarks

Studies on the algal viruses have revealed many unexpected properties, and undoubtedly many more of these viruses exist in nature and await discovery. Several practical benefits have also resulted from studies on these viruses: (a) The chlorella viruses are a new source of DNA restriction endonucleases, some of which recognize novel cleavage sites. For example, the restriction endonuclease *Cvi*JI has been used to diagnose a common mutation that causes the mitochondrial inherited genetic disease myoclonic epilepsy and ragged-red fiber disease (141). Another enzyme *Cvi*RI has been used to diagnose a common mutation producing a hereditary defect in the enzyme thiopurine S-methyltransferase (66). People lacking this enzyme are unable to methylate aromatic and heterocyclic sulfhydryl compounds, including some important thioquinane anticancer drugs. (b) Although not discussed in this review, the chlorella viruses are a new source of promoters for expressing genes in foreign hosts. For example, the upstream region of one of the viral adenine DNA methyltransferase genes functions extremely well in several higher plants (95) and in many bacteria (96; Y. Xia, J. L. Van Etten, F. Wag-

---

**Figure 3.** (facing page) **A.** Phylogenetic tree of the DNA polymerases from some large dsDNA viruses generated by using protein parsimony analysis of the 100 bootstrapped data sets. The numbers at the branches are the bootstrap values indicating the relative strengths of those branches. **B.** Phylogenetic tree generated by the neighbor-joining method. For constructing the distance tree in B, ASFV was omitted from the polymerase sequence data sets. Inclusion of ASFV created problems because it was too distantly related to other viruses. Abbreviations: FsV, *Feldmannia* sp. virus; MpV-xx, viruses that infect *Micromonas pusilla*; CbV-xx, viruses that infect *Chrysochromulina brevifilum*; PBCV-1 and NY-2A, viruses that infect *Chlorella* NC64A; HSV-1 and HSV-2, herpes simplex viruses type 1 and type 2; PrV, pseudorabies virus; VZV, varicella-zoster virus; EBV, Epstein-Barr virus; HCMV, human cytomegalovirus; GPCMV, guinea pig cytomegalovirus; MCMV, murine cytomegalovirus; VacV, vaccinia virus; CbEPV, *Choristoneura biennis* entomovirus; FPV, fowlpox virus; AcNPV, *Autographa californica* nuclear polyhedrosis virus; BmNPV, *Bombyx mori* nuclear polyhedrosis virus; HzNPV, *Helicoverpa zea* nuclear polyhedrosis virus; LdNPV, *Lymantria dispar* nuclear polyhedrosis virus; LCDV, lymphocystis disease virus; ASFV, African swine fever virus. (Modified with permission from J. Phycology—75a)



ner, unpublished results). (c) The chlorella viruses encode at least two proteins that are targets for medically important drugs. Type-II DNA topoisomerases are the object of many anticancer drugs, and a  $K^+$  ion channel defect is responsible for heart muscle arrhythmia. Because drugs that effect these two proteins inhibit chlorella virus plaque formation, the virus plaque assay has the potential to serve as a rapid method to screen for new  $K^+$  ion channel protein inhibitors and type-II DNA topoisomerase inhibitors (M. Nelson & J. L. Van Etten, unpublished results).

Many fundamental research questions remain to be answered about these viruses—ranging from their natural history to the regulation of gene transcription. Examples include the following: (a) How did these viruses acquire such a mixture of prokaryotic- and eukaryotic-like genes? (b) Do the viruses have another host or hosts? (c) How is early and late gene transcription regulated? (d) Why are some of the viral introns more conserved than the virus exons? (e) What are the mechanism and intracellular location of viral protein glycosylation? (f) Why do the viruses encode so many sugar-manipulating enzymes?

## Acknowledgments

Les Lane, Mike Nelson, Myron Brakke, and Mike Graves are acknowledged for many suggestions on early drafts of this manuscript, and Mike Graves and Jim Gurnon helped with the figures. The reconstructed micrographs used to make up Figure 1 were kindly supplied by Xiaodong Yan, Norm Olson, and Tim Baker. Takashi Yamada, Curtis Suttle, and Luis Villarreal provided copies of manuscripts before publication. Investigations were supported by NIH (JVE), NSF-EP-SCoR (JVE), the University of Nebraska Biotechnology Center (JVE), Office of Naval Research (RHM), and Seq-aGrant (RHM). We personally thank all past and present undergraduate students, graduate students, postdoctoral associates, technicians, and collaborating scientists who have worked with us on these fascinating viruses.

## Literature Cited

1. Andrianopoulos K, Wang L, Reeves PR. 1998. Identification of the fucose synthetase gene in the colanic acid gene cluster of *Escherichia coli* K-12. *J. Bacteriol.* 180:998-1001
2. Antequera F, Bird A. 1993. CpG islands. In *DNA Methylation: Molecular Biology and Biological Significance*, ed. PJ Jost, PH Saluz, pp. 169-85. Basel: Birkhauser Verlag
3. Anton BP, Heiter DF, Benner JS, Hess EJ, Greenough L, et al. 1997. Cloning and characterization of the Bg/II restriction modification system reveals a possible evolutionary footprint. *Gene* 187:19-27
4. Ayyagari R, Impellizzeri KJ, Yoder BL, Gary SL, Burgers PMJ. 1995. A mutational analysis of the yeast proliferating cell nuclear antigen indicates distinct roles in DNA replication and DNA repair. *Mol. Cell. Biol.* 15:4420-29

5. Baker JR, Evans LV. 1973. The ship fouling alga *Ectocarpus*. I. Ultrastructure and cytochemistry of plurilocular reproductive stages. *Protoplasma* 77:1-13
6. Bannister JV, Bannister WH, Rotilio G. 1987. Aspects of the structure, function, and applications of superoxide dismutase. *CRC Rev. Biochem.* 22:111-80
7. Becker B, Lesemann DE, Reisser W. 1993. Ultrastructural studies on a chlorella virus from Germany. *Arch. Virol.* 130: 145-55
8. Belfield GP, Tuite MF. 1993. Translation elongation factor 3: a fungus-specific translation factor? *Mol. Microbiol.* 9:411-18
9. Blasco R, Aguero M, Almendral JM, Vinuela E. 1989. Variable and constant regions in African swine fever virus DNA. *Virology* 168:330-38
10. Blasco R, Cole NB, Moss B. 1991. Sequence analysis, expression, and deletion of a vaccinia virus gene encoding a homolog of profilin, a eukaryotic actin-binding protein. *J. Virol.* 65:4598-608
11. Blasco R, De La Vega I, Almazan F, Aguero M, Vinuela E. 1989. Genetic variation of African swine fever virus: variable regions near the ends of the viral DNA. *Virology* 173:251-57
12. Brautigam M, Klein M, Knippers R, Muller DG. 1995. Inheritance and meiotic elimination of a virus genome in the host *Ectocarpus siliculosus* (Phaeophyceae). *J. Phycol.* 31:823-27
13. Burbank DE, Shields SL, Schuster AM, Van Etten JL. 1990. 5-Azacytidine resistant mutants of chlorella virus IL-3A. *Virology* 176:311-15
14. Caspar DLD, Klug A. 1962. Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* 27:1-24
15. Chase TE, Nelson JA, Burbank DE, Van Etten JL. 1989. Mutual exclusion occurs in a chlorella-like green alga inoculated with two viruses. *J. Gen. Virol.* 70:1829-36
16. Chen F, Suttle CA. 1996. Evolutionary relationships among large double-stranded DNA viruses that infect microalgae and other organisms as inferred from DNA polymerase genes. *Virology* 219:170-78
17. Cheng C, Shuman S. 1997. Characterization of an ATP-dependent DNA ligase encoded by *Haemophilus influenzae*. *Nucleic Acids Res.* 25:1369-13
18. Cohen SS. 1998. *A Guide to the Polyamines*. New York: Oxford Univ. Press
19. Cottrell MT, Suttle CA. 1991. Wide spread occurrence and clonal variation in viruses which cause lysis of a cosmopolitan, eukaryotic marine phytoplankter, *Micromonas pusilla*. *Mar. Ecol. Prog. Ser.* 78: 1-9
20. Cottrell MT, Suttle CA. 1993. Isolation of axenic strains of *Micromonas pusilla* (Prasinophyceae). *J. Phycol.* 29:385-87
21. Craig NL, Nash HA. 1983. The mechanism of phage lambda site specific recombination: site specific breakage of DNA by INT topoisomerase. *Cell* 35:795-803
22. Davis RH, Morris DR, Coffino P. 1992. Sequestered end products and enzyme regulation: the case of ornithine decarboxylase. *Microbiol. Rev.* 56:280-90
23. DeAngelis PL, Jing W, Graves MV, Burbank DE, Van Etten JL. 1997. Hyaluronan synthase of chlorella virus PBCV-1. *Science* 278:1800-3
24. Doermann AH. 1983. Introduction to the early years of bacteriophage T4, In *Bacteriophage T4*, ed. CK Mathews, EM Kutter, G Mosig, PB Berget, pp. 1-7. Washington, DC: Am. Soc. Microbiol.
25. Doms RW, Lamb RA, Rose JK, Helenius A. 1993. Folding and assembly of viral membrane proteins. *Virology* 193:545-62
26. Dynes JL, Firtel RA. 1989. Molecular complementation of a genetic marker in *Dictyostelium* using a genomic DNA library. *Proc. Natl. Acad. Sci. USA* 86:7966-70

27. Fitzgerald MC, Skowron P, Van Etten JL, Smith LM, Mead DA. 1992. Rapid shotgun cloning utilizing the two base recognition endonuclease *Cvi*JI. *Nucleic Acids Res.* 20:3753-62
28. Flugel RM. 1985. Lymphocystis disease virus. *Curr. Top. Microbiol. Immunol.* 116: 133-50
29. Fuhrman JA, Suttle CA. 1993. Viruses in marine planktonic systems. *Oceanography* 6:51-63
30. Furuta M, Schrader JO, Schrader HS, Kokjohn TA, Nyaga S, et al. 1997. *Chlorella* virus PBCV-1 encodes a homolog of the bacteriophage T4 UV damage repair gene *DenV*. *Appl. Environ. Microbiol.* 63:1551-56
31. Garcia-Beato R, Freije JMP, Lopez-Otin C, Blasco R, Vinuela E, Salas ML. 1992. A gene homologous to topoisomerase II in African swine fever virus. *Virology* 188:938-47
32. Garon CF, Barbosa E, Moss B. 1978. Visualization of an inverted terminal repetition in vaccinia virus DNA. *Proc. Natl. Acad. Sci. USA* 75:4863-67
33. Geshelin P, Berns KI. 1974. Characterization and localization of the naturally occurring cross-links in vaccinia virus DNA. *J. Mol. Biol.* 88:785-96
34. Gingrich JC, Boehrer DM, Basu SB. 1996. Partial *Cvi*JI digestion as an alternative approach to generate cosmid sublibraries for large-scale sequencing projects. *Bio- Techniques* 21:99-104
35. Girtton LE, Van Etten JL. 1987. Restriction site map of the *Chlorella* virus PBCV-1 genome. *Plant Mol. Biol.* 9:247-57
36. Gonzalez A, Talavera A, Almendral JM, Vinuela E. 1986. Hairpin loop structure of African swine fever virus DNA. *Nucleic Acids Res.* 14:6835-44
37. Gooday GW, Humphreys AM, McIntosh WH. 1986. Role of chitinases in fungal growth. In *Chitin in Nature and Technology*, ed. R Muzzarelli, C Jeuniaux, GW Gooday, pp. 83-91. New York: Plenum
38. Goodwin TW, Mercer EI. 1983. *Introduction to Plant Biochemistry*. Oxford, UK: Pergamon. 677 pp.
39. Goohra R, Granoff A. 1979. Icosahedral cytoplasmic deoxyriboviruses. In *Comprehensive Virology*, Vol. 14, ed. H Fraenkel-Conrat pp. 347-399. New York: Plenum
40. Goorha R, Murti KG. 1982. The genome of frog virus 3, an animal DNA virus, is circularly permuted and terminally redundant. *Proc. Natl. Acad. Sci. USA* 79:248-52
41. Grabherr R, Strasser P, Van Etten JL. 1992. The DNA polymerase gene from *Chlorella* viruses PBCV-1 and NY-2A contains an intron with nuclear splicing sequences. *Virology* 188:721-31
- 41a. Graves MV, Burbank DE, Roth R, Heuser J, DeAngelis PL, Van Etten JL. 1999. Hyaluronan synthesis in virus PBCV-1-infected *Chlorella*-like green algae. *Virology* 257:15-23
42. Graves MV, Meints RH. 1992. Characterization of the major capsid protein and cloning of its gene from algal virus PBCV1. *Virology* 188:198-207
43. Graves MV, Meints RH. 1992. Characterization of the gene encoding the most abundant in vitro translation product from virus-infected *Chlorella*-like algae. *Gene* 113:149-55
44. Green MR. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu. Rev. Cell Biol.* 7: 559-99
45. Ha HC, Sirisoma NS, Kuppusamy P, Zweier JL, Woster PM, Casero RA. 1998. The natural polyamine spermine functions directly as a free radical scavenger. *Proc. Natl. Acad. Sci. USA* 95:11140-45
46. Hakansson K, Doherty AJ, Shuman S, Wigley DB. 1997. X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell* 89:543-53
47. Hakansson K, Wigley DB. 1998. Structure of a complex between a cap analogue and

- mRNA guanylyl transferase demonstrates the structural chemistry of RNA capping. *Proc. Natl. Acad. Sci. USA* 95:1505–10
48. Hayes W. 1968. *The Genetics of Bacteria and Their Viruses*, 2nd ed. Oxford, UK: Blackwell Sci. 925 pp.
49. Hellmann V, Kessler E. 1974. Physiologische und Biochemische Beiträge zur Taxonomie der gattung *Chlorella*. VIII. Die Basenzusammensetzung der DNS. *Arch. Microbiol.* 95:311–18
50. Henry EC, Meints RH. 1992. A persistent virus infection in *Feldmannia* (Phaeophyceae). *J. Phycol.* 28:517–26
51. Herth W, Mulisch M, Zugenmaier P. 1986. Comparison of chitin fibril structure and assembly in three unicellular organisms. In *Chitin in Nature and Technology*, ed. R Muzzarelli, C Jeuniaux, GW Gooday, pp. 107–20. New York: Plenum
52. Ho CK, Van Etten JL, Shuman S. 1996. Expression and characterization of an RNA capping enzyme encoded by chlorella virus PBCV-1. *J. Virol.* 70:6658–64
53. Ho CK, Van Etten JL, Shuman S. 1997. Characterization of an ATP-dependent DNA ligase encoded by chlorella virus PBCV-1. *J. Virology* 71:1931–37
54. Husmann LK, Yung DL, Hollingshead SK, Scott JR. 1997. Role of putative virulence factors of *Streptococcus pyogenes* in mouse models of long-term throat colonization and pneumonia. *Infect. Immun.* 65:1422–30
55. Ivey RG, Henry EC, Lee AM, Klepper L, Krueger SK, Meints RH. 1996. A *Feldmannia* algal virus has two genome size-classes. *Virology* 220:267–73
56. Jin A., Zhang Y, Xia Y, T aylor E, Nelson M, Van Etten JL. 1994. New restriction endonuclease CviRI cleaves DNA at TG/CA sequences. *Nucleic Acids Res.* 22:3928–29
57. Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, et al. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 3:109–36
58. Kapaun E, Reisser W. 1995. A chitin-like glycan in the cell wall of a *Chlorella* sp. (Chlorococcales, Chlorophyceae). *Planta* 197:577–82
59. Kaur K, Rohozinski J, Goorha R. 1995. Identification and characterization of the frog virus 3 DNA methyltransferase gene. *J. Gen. Virol.* 76:1937–43
60. Kawakami H, Kawakami N. 1978. Behavior of a virus in a symbiotic system, *Paramecium bursaria*-zoochlorella. *J. Protozool.* 25:217–25
61. Klein M, Lanka STJ, Knippers R, Muller D. 1995. Coat protein of the *Ectocarpus siliculosus* virus. *Virology* 206:520–26
62. Klein M, Lanka S, Muller D, Knippers R. 1994. Single-stranded regions in the genome of the *Ectocarpus siliculosus* virus. *Virology* 202:1076–78
63. Klimasauskas S, Timinska A, Menkevicius S, Butkiene D, Butkus V, Janulaitis A. 1989. Sequence motifs characteristic of DNA [cytosine-N4] methyltransferases: similarity to adenine and cytosine-C5 DNA methylases. *Nucleic Acids Res.* 17:9823–32
64. Krishna TSR, Kong XP, Gary S, Burgers PM, Kuriyan J. 1994. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* 79:1233–43
65. Krueger SK, Henry EC, Ivey RG, Meints RH. 1996. A 4.5-kbp nucleotide sequence fragment from a large dsDNA brown algal virus contains ORFs and a “RING” zinc finger motif. *Virology* 219:301–3
66. Krynetski EY, Schuetz JD, Galpin AJ, Pui CH, Relling MV, Evans WE. 1995. A single point mutation leading to loss of catalytic activity in human thiopurine S-methyltransferase. *Proc. Natl. Acad. Sci. USA* 92:949–53
67. Kumar S, Cheng X, Klimasauskas S, Mi S, Posfai J, et al. 1994. The DNA (cytosine5)

- methyltransferases. *Nucleic Acids Res.* 22:1-10
68. Kutish GF, Li Y, Lu Z, Furuta M, Rock DL, Van Etten JL. 1996. Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map positions 182 to 258. *Virology* 223:303-17
69. Landstein D, Burbank DE, Nietfeldt JW, Van Etten JL. 1995. Large deletions in antigenic variants of the chlorella virus PBCV-1. *Virology* 214:413-20
70. Landstein D, Graves MV, Burbank DE, DeAngelis P, Van Etten JL. 1998. Chlorella virus PBCV-1 encodes functional glutamine:fructose-6-phosphate amidotransferase and UDP-glucose dehydrogenase enzymes. *Virology* 250:388-96
71. Landstein D, Minberg M, Arad S, Tal J. 1996. An early gene of the chlorella virus PBCV-1 encodes a functional aspartate transcarbamylase. *Virology* 221:151-58
72. Lanka STJ, Klein M, Ramsperger U, Muller DG, Knippers R. 1993. Genome structure of a virus infecting the marine brown alga *Ectocarpus siliculosus*. *Virology* 193:802-11
73. Laurent TC, Fraser JRE. 1992. Hyaluronan. *FASEB J.* 6:2397-404
74. Lee AM, Ivey RG, Henry EC, Meints RH. 1995. Characterization of a repetitive DNA element in a brown algal virus. *Virology* 212:474-80
- 75a. Lee AM, Ivey RG, Meints RH. 1998. The DNA polymerase gene of a brown algal virus: structure and phylogeny. *J. Phycol.* 34:608-15
- 75b. Lee AM, Ivey RG, Meints RH. 1998. Repetitive DNA insertion in a protein kinase ORF of a latent FSV (*Feldmannia* sp. virus) genome. *Virology* 248:35-45
76. Lepetit D, Thiebaud P, Aoufouchi S, Prigent C, Guesne R, Theze N. 1996. The cloning and characterization of a cDNA encoding *Xenopus levis* DNA ligase I. *Gene* 172:273-77
77. Li Y, Lu Z, Burbank DE, Kutish GF, Rock DL, Van Etten JL. 1995. Analysis of 43 kb of the chlorella virus PBCV-1 330 kb genome: map position 45 to 88. *Virology* 212:134-50
78. Li Y, Lu Z, Sun L, Ropp S, Kutish GF, Rock DL, Van Etten JL. 1997. Analysis of 74 kb of DNA located at the right end of the 330-kb chlorella virus PBCV-1 genome. *Virology* 237:360-77
79. Lloyd RS, Linn S. 1993. Nucleases involved in DNA repair. In *Nucleases*, ed. S Linn, RS Lloyd, RJ Roberts, Vol. II, pp. 263-316. Cold Spring Harbor, NY: Cold Spring Harbor Lab. Press
80. Loos E, Meindl D. 1982. Composition of the cell wall of *Chlorella fusca*. *Planta* 156:270-73
81. Lopez-Otin C, Freije JM, Parra F, Mendez E, Vinuela E. 1990. Mapping and sequence of the gene coding for protein p72, the major capsid protein of African swine fever virus. *Virology* 175:477-84
82. Lu Z, Li Y, Que Q, Kutish GF, Rock DL, Van Etten JL. 1996. Analysis of 94 kb of the chlorella virus PBCV-1 330 kb genome: map position 88 to 182. *Virology* 216:102-23
83. Lu Z, Li Y, Zhang Y, Kutish GF, Rock DL, Van Etten JL. 1995. Analysis of 45 kb of DNA located at the left end of the chlorella virus PBCV-1 genome. *Virology* 206:339-52
84. Malone TE, Blumenthal RM, Cheng X. 1995. Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. *J. Mol. Biol.* 253:618-32
85. Mao J, Hedrick RP, Chinchar VG. 1997. Molecular characterization, sequence analysis, and taxonomic position of newly isolated fish iridoviruses. *Virology* 229:212-20
86. Mao J, Tham TN, Gentry GA, Aubertin A, Chinchar VG. 1996. Cloning, sequence analysis, and expression of the major capsid protein of the iridovirus frog virus 3. *Virology* 216:431-36
87. Mayer JA, Taylor FJR. 1979. A virus which lyses the marine nanoflagellate *Micromonas pusilla*. *Nature* 281:299-301
88. McCluskey K, Graves MV, Mills D, Meints RH. 1992. Replication of chlorella virus PBCV-1 and host karyotype determination studies with pulsed-field gel electrophore-

- sis. *J. Phycol.* 28:846–50
89. McCullough AK, Romberg MT, Nyaga S, Lei Y, Wood T, et al. 1998. Characterization of a novel cis-syn and trans-syn II pyrimidine dimer glycosylase/AP lyase from a eukaryotic algal virus, *Paramecium bursaria* chlorella virus-1. *J. Biol. Chem.* 273:13136–42
  90. Meints RH, Burbank DE, Van Etten JL, Lampion DTA. 1988. Properties of the chlorella receptor for the virus PBCV-1. *Virology* 164:15–21
  91. Meints RH, Lee K, Burbank DE, Van Etten JL. 1984. Infection of a chlorella-like alga with the virus, PBCV-1: ultrastructural studies. *Virology* 138:341–46
  92. Meints RH, Lee K, Van Etten JL. 1986. Assembly site of the virus PBCV-1 in a chlorella-like green alga: ultrastructural studies. *Virology* 154:240–45
  93. Meints RH, Van Etten JL, Kuczmarski D, Lee K, Ang B. 1981. Viral infection of the symbiotic chlorella-like alga present in *Hydra viridis*. *Virology* 113:698–703
  94. Mindich L, Bamford D. 1988. Lipid-containing bacteriophages. In *The Bacteriophages*, Vol. 2, ed. R Calendar, pp. 475–520. New York: Plenum
  95. Mitra A, Higgins DW. 1994. The chlorella virus adenine methyltransferase gene promoter is a strong promoter in plants. *Plant Mol. Biol.* 26:85–93
  96. Mitra A, Higgins DW, Rohe NJ. 1994. A chlorella virus gene promoter functions as a strong promoter both in plants and bacteria. *Biochem. Biophys. Res. Commun.* 204:187–94
  97. Muller DG. 1967. Generationswechsel, kernphasenwechsel und sexualität der braunalge *Ectocarpus siliculosus* in kulturversuch. *Planta* 75:39–54
  98. Müller DG. 1972. Life cycle of the brown alga *Ectocarpus fasciculatus* var. *refractus* (Kütz.) Ardis. (Phaeophyceae, Ectocarpales) in culture. *Phycologia* 11:11–13
  99. Muller DG. 1991. Mendelian segregation of a virus genome during host meiosis in the marine brown alga *Ectocarpus siliculosus*. *J. Plant Physiol.* 137:739–43
  100. Muller DG. 1992. Intergeneric transmission of a marine plant DNA virus. *Naturwissenschaften* 79:37–39
  101. Muller DG, Brautigam M, Knippers R. 1996. Virus infection and persistence of foreign DNA in the marine brown alga *Feldmannia simplex* (Ectocarpales, Phaeophyceae). *Phycologia* 35:61–63
  102. Muller DG, Frenzer K. 1993. Virus infections in three marine brown algae: *Feldmannia irregularis*, *F. simplex*, and *Ectocarpus siliculosus*. *Hydrobiologia* 260:37–44
  103. Muller DG, Kapp M, Knippers R. 1998. Viruses in marine brown algae. *Adv. Virus Res.* 50:49–67
  104. Muller DG, Kawai H, Stache B, Lanka S. 1990. A virus infection in the marine brown alga *Ectocarpus siliculosus* (Phaeophyceae). *Bot. Acta* 103:72–82
  105. Muller DG, Parodi E. 1993. Transfer of a marine DNA virus from *Ectocarpus* to *Feldmannia* (Ectocarpales, Phaeophyceae): aberrant symptoms and restitution of the host. *Protoplasma* 175:121–25
  106. Müller DG, Stache B. 1992. Worldwide occurrence of virus infections in filamentous marine brown algae. *Helgol. Meeresunters.* 46:1–8
  107. Narva KE, Wendell DL, Skrdla MP, Van Etten JL. 1987. Molecular cloning and characterization of the gene encoding the DNA methyltransferase, M.CviBIII, from chlorella virus NC-1A. *Nucleic Acids Res.* 15:9807–23
  108. Nelson M, Burbank DE, Van Etten JL. 1998. Chlorella viruses encode multiple DNA methyltransferases. *Biol. Chem.* 379:423–28
  109. Nelson M, Zhang Y, Van Etten JL. 1993. DNA methyltransferases and DNA site-specific endonucleases encoded by chlorella viruses. In *DNA Methylation: Molecular Biology and Biological Significance*, ed. PJ Jost, PH Saluz, pp. 186–211. Basel: Birkhauser Verlag
  110. Niefeldt JW, Lee K, Van Etten JL. 1992. Chlorella virus PBCV-1 replication is not af-



- fect by cytoskeletal disruptors. *Intervirology* 33:116-20
- 110a. Nishida K, Kimura Y, Kawasaki T, Fujie M, Yamada T. 1999. Genetic variation of chlorella viruses: variable regions localized on the CVK2 genomic DNA. *Virology* 255:376-84
111. Nishida K, Suzuki S, Kimura Y, Nomura N, Fujie M, Yamada T. 1998. Group I introns found in chlorella viruses: biological implications. *Virology* 242:319-26
112. Oliviera L, Bisalputra T. 1978. A virus infection in the brown alga *Sorocarpus uvaeformis* (Lyngbye) Pringsheim (Phaeophyta, Ectocarpales). *Ann. Bot.* 42:439-45
113. Parodi ER, Muller DG. 1994. Field and culture studies on virus infections in *Hinckesia hincksiae* and *Ectocarpus fasciculatus* (Ectocarpales, Phaeophyceae). *Eur. J. Phycol.* 29:113-17
114. Peters LL, Lux SE. 1993. Ankyrins: structure and function in normal cells and hereditary spherocytes. *Semin. Hematol.* 30:85-118
115. Pisabarro A, Malumbres M, Mateos LM, Oguiza JA, Martin JF. 1993. A cluster of three genes (dapA, ORF2, and dapB) of *Brevibacterium lactofermentum* encodes dihydrodipicolinate synthase, dihydrodipicolinate reductase, and a third polypeptide of unknown function. *J. Bacteriol.* 175:2743-49
116. Posfai J, Bhagwat AS, Posfai G, Roberts RJ. 1989. Predictive motifs derived from cytosine methyltransferases. *Nucleic Acids Res.* 17:2421-35
117. Que Q, Li Y, Wang IN, Lane LC, Chaney WG, Van Etten JL. 1994. Protein glycosylation and myristylation in chlorella virus PBCV-1 and its antigenic variants. *Virology* 203:320-27
118. Que Q, Van Etten JL. 1995. Characterization of a protein kinase gene from two chlorella viruses. *Virus Res.* 35:291-305
119. Que Q, Zhang Y, Nelson M, Ropp S, Burbank DE, Van Etten JL. 1997. Chlorella virus SC-1A encodes at least six DNA methyltransferases. *Gene* 190:237-44
120. Rademacher TW, Parekh RB, Dwek RA. 1988. Glycobiology. *Annu. Rev. Biochem.* 57:785-838
121. Reanney DC. 1974. Viruses and evolution. *Int. Rev. Cytol.* 37:21-52
122. Reisser W, ed. 1992. *Algae and Symbioses*. Bristol, UK: Biopress. 746 pp.
123. Reisser W, Becker B, Klein T. 1986. Studies on ultrastructure and host range of a chlorella attacking virus. *Protoplasma* 135:162-65
124. Reisser W, Burbank DE, Meints RH, Becker B, Van Etten JL. 1991. Viruses distinguish symbiotic Chlorella spp. of *Paramecium bursaria*. *Endocytobiosis Cell Res.* 7:245-51
125. Reisser W, Burbank DE, Meints SM, Meints RH, Becker B, Van Etten JL. 1988. A comparison of viruses infecting two different chlorella-like green algae. *Virology* 167:143-49
126. Reisser W, Kapaun E. 1991. Entry of a chlorella-virus into its host cell. *J. Phycol.* 27:609-13
127. Roca J. 1995. The mechanisms of DNA topoisomerases. *Trends Biochem. Sci.* 20:156-60
128. Rohozinski J, Girton LE, Van Etten JL. 1989. Chlorella viruses contain linear nonpermuted double stranded DNA genomes with covalently closed hairpin ends. *Virology* 168:363-69
129. Royce CL, Pardy RL. 1996. Endotoxin-like properties of an extract from a symbiotic, eukaryotic chlorella-like green alga. *J. Endotoxin Res.* 3:437-44
130. Schlesinger MJ, Schlesinger S. 1987. Domains of virus glycoproteins. *Adv. Virus Res.* 33:1-44
131. Schmidt KH, Gunther E, Courtney HS. 1996. Expression of both M protein and hyaluronic acid capsule by group A streptococcal strains results in a high virulence for chicken embryos. *Med. Microbiol. Immunol.* 184:169-73

132. Schnitzler P, Darai G. 1993. Identification of the gene encoding the major capsid protein of fish lymphocystis disease virus. *J. Gen. Virol.* 74:2143–50
133. Schuster AM, Burbank DE, Meister B, Skrdla MP, Meints RH. 1986. et al. Characterization of viruses infecting a eukaryotic chlorella-like green alga. *Virology* 150:170–77
134. Schuster AM, Girton L, Burbank DE, Van Etten JL. 1986. Infection of a chlorella-like alga with the virus PBCV-1: transcriptional studies. *Virology* 148:181–89
135. Schuster AM, Graves M, Korth K, Ziegelbein M, Brumbaugh J, et al. 1990. Transcription and sequence studies of a 4.3kbp fragment from a dsDNA eukaryotic algal virus. *Virology* 176:515–23
136. Seaton GR, Lee K, Rohozinski J. 1995. Photosynthetic shutdown in *Chlorella* NC64A associated with the infection cycle of *Paramecium bursaria* *Chlorella* virus-1. *Plant Physiol.* 108:1431–38
137. Sengco MR, Brautigam M, Kapp M, Muller DG. 1996. Detection of virus-DNA in *Ectocarpus siliculosus* and *E. fasciculatus* (Phaeophyceae) from various geographic areas. *Eur. J. Phycol.* 31:73– 78
138. Senkevich TG, Koonin EV, Buller ML. 1994. A poxvirus protein with a RING zinc finger motif is of crucial importance for virulence. *Virology* 198:118–28
139. Senkevich TG, Wolffe EJ, Buller ML. 1995. Ectromelia virus RING finger protein is localized in virus factories and is required for virus replication in macrophages. *J. Virol.* 69:4103–11
140. Shchelkunov SN, Totmenin AV. 1995. Two types of deletions in orthopoxvirus genomes. *Virus Genes* 9:231–45
141. Shoffner JM, Lott MT, Lezza AMS, Seibel P, Ballinger SW, Wallace DC. 1990. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA<sup>Ala</sup> mutation. *Cell* 61:931–37
142. Skowron PM, Swaminathan N, McMaster K, George D, Van Etten JL, Mead DA. 1995. Cloning and application of the two/three-base restriction endonuclease R.CviJI from IL-3A virus-infected chlorella. *Gene* 157:37–41
143. Skrdla MP, Burbank DE, Xia Y, Meints RH, Van Etten JL. 1984. Structural proteins and lipids in a virus, PBCV-1, which replicates in a chlorella-like alga. *Virology* 135:308–15
144. Sogo JM, Almendral JM, Talavera A, Vinuela E. 1984. Terminal and internal inverted repetitions in African swine fever virus DNA. *Virology* 133:271–75
145. Songsri P, Hamazaki T, Ishikawa Y, Yamada T. 1995. Large deletions in the genome of chlorella virus CVK1. *Virology* 214:405–12
146. Songsri P, Hiramatsu S, Fujie M, Yamada T. 1997. Proteolytic processing of chlorella virus CVK2 capsid proteins. *Virology* 227:252–54
147. Sriskanda V, Shuman S. 1998. Chlorella virus DNA ligase: nick recognition and mutational analysis. *Nucleic Acids Res.* 26:525–31
148. Sriskanda V, Shuman S. 1998. Specificity and fidelity of strand joining by chlorella virus DNA ligase. *Nucleic Acids Res.* 26:3536–41
149. Sriskanda V, Shuman S. 1998. Mutational analysis of chlorella virus DNA ligase: catalytic roles of domain I and motif VI. *Nucleic Acids Res.* 26:4618–25
150. Stillman B. 1994. Smart machines at the DNA replication fork. *Cell* 78:725– 28
151. Stohwasser R, Raab K, Schnitzler P, Janssen W, Darai G. 1993. Identification of the gene encoding the major capsid protein of insect iridescent virus type 6 by polymerase chain reaction. *J. Gen. Virol.* 74:873–79
152. Strasser P, Zhang Y, Rohozinski J, Van Etten JL. 1991. The termini of the chlorella virus PBCV-1 genome are identical 2.2kbp inverted repeats. *Virology* 180:763– 69
153. Suttle CA. 1999. The ecological, evolutionary and geochemical consequences of viral in-

- fection of cyanobacteria and eukaryotic algae. In *Viral Ecology*, ed. CJ Hurst, pp. xx-xx. New York: Academic. In press
154. Suttle CA, Chan AM. 1994. Dynamics and distribution of cyanophages and their effect on marine *Synechococcus* spp. *Appl. Environ. Microbiol.* 60:3167-74
155. Suttle CA, Chen F. 1992. Mechanisms and rates of decay of marine viruses in seawater. *Appl. Environ. Microbiol.* 58:3721-29
156. Swaminathan ND, George D, McMaster K, Szablewski J, Van Etten JL, Mead DA. 1994. Restriction generated oligonucleotides utilizing the two base recognition endonuclease *CviJI*\*. *Nucleic Acids Res.* 22:1470-75
157. Swaminathan ND, Mead DA, McMaster K, George D, Van Etten JL, Skowron PM. 1996. Molecular cloning of the three base restriction endonuclease R.*CviJI* from eucaryotic chlorella virus IL-3A. *Nucleic Acids Res.* 24:2463-69
158. Takeda H. 1988. Classification of chlorella strains by cell wall sugar composition. *Phytochemistry* 27:3822-26
159. Tessman I. 1985. Genetic recombination of the DNA plant virus PBCV-1 in a chlorella-like alga. *Virology* 145:319-22
160. Tidona CA, Schnitzler P, Kehm R, Darai G. 1998. Is the major capsid protein of iridoviruses a suitable target for the study of viral evolution? *Virus Genes* 16:59-66
161. Tomalski MD, Eldridge R, Miller LK. 1991. A baculovirus homolog of a Cu/Zn superoxide dismutase gene. *Virology* 184:149-61
162. Trench RK. 1979. The cell biology of plant-animal symbiosis. *Annu. Rev. Plant Physiol.* 30:485-531
163. Turner PC, Moyer RW. 1990. The molecular pathogenesis of poxviruses. *Curr. Top. Microbiol. Immunol.* 163:125-151
164. van den Hoek C, Mann DG, Jahns HM. 1995. *Algae, an Introduction to Phycology*. Cambridge, UK: Cambridge Univ. Press.
165. Van Etten JL. 1995. Phycodnaviridae. In *Virus Taxonomy*, ed. FA Murphy, CM Fauguet, DHL Bishop, SA Ghabrial, AW Jarvis, et al, pp. 100-103. New York: Springer-Verlag
166. Van Etten JL, Burbank DE, Joshi J, Meints RH. 1984. DNA synthesis in a chlorella-like alga following infection with the virus PBCV-1. *Virology* 134:443-49
167. Van Etten JL, Burbank DE, Kuczmarski D, Meints RH. 1983. Virus infection of culturable chlorella-like algae and development of a plaque assay. *Science* 219:994-96
168. Van Etten JL, Burbank DE, Meints RH. 1986. Replication of the algal virus PBCV-1 in UV-irradiated chlorella. *Intervirology* 26:115-20
169. Van Etten JL, Burbank DE, Schuster AM, Meints RH. 1985. Lytic viruses infecting a chlorella like alga. *Virology* 140:135-43
170. Van Etten JL, Burbank DE, Xia Y, Meints RH. 1983. Growth cycle of a virus, PBCV1, that infects chlorella-like algae. *Virology* 126:117-25
171. Van Etten JL, Lane LC, Meints RH. 1991. Viruses and viruslike particles of eukaryotic algae. *Microbiological Rev.* 55:586- 620
172. Van Etten JL, Lane LC, Meints RH. 1991. Unicellular plants also have large dsDNA viruses. *Semin. Virol.* 2:71-77
173. Van Etten JL, Meints RH, Burbank DE, Kuczmarski D, Cuppels DA, Lane LC. 1981. Isolation and characterization of a virus from the intracellular green alga symbiotic with *Hydra viridis*. *Virology* 113:704-11
174. Van Etten JL, Meints RH, Kuczmarski D, Burbank DE, Lee K. 1982. Viruses of symbiotic chlorella-like algae isolated from *Paramecium bursaria* and *Hydra viridis*. *Proc. Natl. Acad. Sci. USA* 79:3867-71
175. Van Etten JL, Schuster AM, Girton L, Burbank D, Swinton D, Hattman S. 1985. DNA

- methylation of viruses infecting a eukaryotic chlorella-like green alga. *Nucleic Acids Res.* 13:3471-78
176. Van Etten JL, Schuster AM, Meints RH. 1988. Viruses of eukaryotic chlorella-like algae. In *Viruses of Fungi and Simple Eukaryotes*, ed. Y Koltin, MJ Leibowitz, pp. 411-28, New York: Marcel Dekker
  177. Van Etten JL, Van Etten CH, Johnson JK, Burbank DE. 1985. A survey for viruses from freshwater that infect a eukaryotic chlorella-like green alga. *Appl. Environ. Microbiol.* 49:1326-28
  178. Villarreal LP. 1999. DNA viruses: their influence on host evolution. In *Origin and Evolution of Viruses*, ed. E Domingo, R Webster, JJ Holland, T Pickett. London, Academic. In press
  179. Wang IN, Li Y, Que Q, Bhattacharya M, Lane LC, Chaney WG, Van Etten JL. 1993. Evidence for virus-encoded glycosylation specificity. *Proc. Natl. Acad. Sci. USA* 90:3840-44
  180. Willis DB, Goorha R, Granoff A. 1984. DNA methyltransferase induced by frog virus 3. *J. Virol.* 49:86-91
  181. Willis DB, Granoff A. 1980. Frog virus 3 is heavily methylated at CpG sequences. *Virology* 107:250-57
  182. Wilson GG, Murray NE. 1991. Restriction and modification systems. *Annu. Rev. Genet.* 25:585-627
  183. Wittek R, Moss B. 1980. Tandem repeats within the inverted terminal repetition of vaccinia virus DNA. *Cell* 21:277-84
  184. Wolf S, Maier I, Katsaros C, Muller DG. 1998. Virus assembly in *Hinckia hinckiae* (Eco-carpales, Phaeophyceae). An electron and fluorescence microscopic study. *Protoplasma* 203:153-67
  185. Xia Y, Burbank DE, Uher L, Rabussay D, Van Etten JL. 1986. Restriction endonuclease activity induced by PBCV-1 virus infection of a chlorella-like green alga. *Mol. Cell. Biol.* 6:1430-39
  186. Xia Y, Burbank DE, Uher L, Rabussay D, Van Etten JL. 1987. IL-3A virus infection of a chlorella-like green alga induces a DNA restriction endonuclease with novel sequence specificity. *Nucleic Acids Res.* 15:6075-90
  187. Xia Y, Burbank DE, Van Etten JL. 1986. Restriction endonuclease activity induced by NC-1A virus infection of a chlorella-like green alga. *Nucleic Acids Res.* 14:6017-30
  188. Xia Y, Morgan R, Schildkraut I, Van Etten JL. 1988. A site-specific single strand endonuclease activity induced by NYs-1 virus infection of a chlorella-like green alga. *Nucleic Acids Res.* 16:9477-87
  189. Xia Y, Narva KE, Van Etten JL. 1987. The cleavage site of the *RsaI* isoschizomer, *CviII*, is G/TAC. *Nucleic Acids Res.* 15:10063
  190. Xia Y, Van Etten JL. 1986. DNA methyltransferase induced by PBCV-1 virus infection of a chlorella-like green alga. *Mol. Cell. Biol.* 6:1440-45
  191. Xu M, Kladde MP, Van Etten JL, Simpson RT. 1998. Cloning, characterization and expression of the gene coding for a cytosine-5-DNA methyltransferase recognizing GpC sites. *Nucleic Acids Res.* 26:3961-66
  192. Yamada T, Fukuda T, Tamura K, Furukawa S, Songsri P. 1993. Expression of the gene encoding a translational elongation factor 3 homolog of chlorella virus CVK2. *Virology* 197:742-50
  193. Yamada T, Furukawa S, Hamazaki T, Songsri P. 1996. Characterization of DNA-binding proteins and protein kinase activities in chlorella virus CVK2. *Virology* 219:395-406
  194. Yamada T, Higashiyama T. 1993. Characterization of the terminal inverted repeats and their neighboring tandem repeats in the chlorella CVK1 virus genome. *Mol. Gen. Genet.* 241:554-63

195. Yamada T, Higashiyama T, Fukuda T. 1991. Screening of natural waters for viruses which infect chlorella cells. *Appl. Environ. Microbiol.* 57:3433-37
196. Yamada T, Hiramatsu S, Songsri P, Fujie M. 1997. Alternative expression of a chitosanase gene produces two different proteins in cells infected with chlorella virus CVK2. *Virology* 230:361-68
197. Yamada T, Sakaguchi K. 1982. Comparative studies on chlorella cell walls: induction of protoplast formation. *Arch. Microbiol.* 132:1013
198. Yamada T, Tamura K, Aimi T, Songsri P. 1994. Self-splicing group I introns in eukaryotic viruses. *Nucleic Acids Res.* 22:2532-37
199. Yan X, Olson NH, Van Etten JL, Baker TS. 1998. Cryoelectron microscopy and image reconstruction of PBCV-1, an algal virus with T D 169 lattice symmetry. *Electron Microsc.* 1:775-76
200. Yasuda S, Sekiguchi M. 1970. T4 endonuclease involved in repair of DNA. *Proc. Natl. Acad. Sci. USA* 67:1839-45
201. Yonker CR, Caldwell KD, Giddings JC, Van Etten JL. 1985. Physical characterization of PBCV virus by sedimentation field flow fractionation. *J. Virol. Methods* 11:145-60
202. Yuan YP, Schultz J, Mlodzik M, Bork P. 1997. Secreted fringe-like signaling molecules may be glycosyltransferases. *Cell* 88:9-11
203. Yuen L, Moss B. 1987. Oligonucleotide sequence signaling transcriptional termination of vaccinia virus early genes. *Proc. Natl. Acad. Sci. USA* 84:6417-21
204. Zhang Y, Burbank DE, Van Etten JL. 1988. Chlorella viruses isolated in China. *Appl. Environ. Microbiol.* 54:2170-73
205. Zhang Y, Nelson M, Nietfeldt JW, Burbank DE, Van Etten JL. 1992. Characterization of chlorella virus PBCV-1CviAII restriction and modification system. *Nucleic Acids Res.* 20:5351-56
206. Zhang Y, Nelson M, Nietfeldt J, Xia Y, Burbank DE, et al. 1998. Chlorella virus NY-2A encodes at least twelve DNA endonuclease/methyltransferase genes. *Virology* 240:366-75
207. Zhang Y, Nelson M, Van Etten JL. 1992. A single amino acid change restores DNA cytosine methyltransferase activity in a cloned chlorella virus pseudogene. *Nucleic Acids Res.* 20:1637-42
208. Zhang Y, Strasser P, Grabherr R, Van Etten JL. 1994. Hairpin loop structure at the termini of the chlorella virus PBCV-1 genome. *Virology* 202:1079-82