Evidence for the evolution of ascoviruses from iridoviruses

Karine Stasiak,1 Sylvaine Renault,1 Marie-Véronique Demattei,1 Yves Bigot1,2 and Brian A. Federici2

1Université François Rabelais, UFR des Sciences et Techniques, Laboratoire d’Etude des Parasites Génétiques, F3-CNRS 2535, Parc Grandmont, 37200 Tours, France
2Department of Entomology and Interdepartmental Graduate Programs in Genetics and Microbiology, University of California, Riverside, CA 92507, USA

Ascoviruses (family Ascoviridae) are large, enveloped, double-stranded (ds)DNA viruses that attack lepidopteran larvae and pupae, and are unusual in that they are transmitted by parasitic wasps during oviposition. Previous comparisons of DNA polymerase sequences from vertebrate and invertebrate viruses suggested that ascoviruses are closely related to iridoviruses. This relationship was unexpected because these viruses differ markedly in virion symmetry, genome configuration and cellular pathology. Here we present evidence based on sequence comparisons and phylogenetic analyses of a greater range of ascovirus proteins and their homologues in other large dsDNA viruses that ascoviruses evolved from iridoviruses. Consensus trees for the major capsid protein, DNA polymerase, thymidine kinase and ATPase III from representative ascoviruses, algal viruses (family Phycodnaviridae), vertebrate and invertebrate iridoviruses (family Iridoviridae) and African swine fever virus (ASFV; family Asfarviridae) showed that ascovirus proteins clustered most closely with those of the lepidopteran iridovirus Chilo iridescent virus (CIV) (Invertebrate iridescent virus 6). Moreover, analysis of the presence or absence of homologues of an additional 50 proteins encoded in the genome of Spodoptera frugiperda ascovirus (SfAV-1a) showed that about 40% occurred in CIV, with lower percentages encoded by the genomes of, respectively, vertebrate iridoviruses, phycodnaviruses and ASFV. The occurrence of three of these genes in SfAV-1a but not CIV was indicative of the evolutionary differentiation of ascoviruses from invertebrate iridoviruses.

INTRODUCTION

Different families of viruses are distinguished from one another based primarily on the structural and biochemical characteristics of their virions (van Regenmortel, 2000). The most important characteristics are the size, shape and complexity of the virion, nucleic acid type and physical configuration of the genome, and presence or absence of an envelope. Using these criteria, more than 50 virus families are recognized, ranging from small single-stranded (ss)RNA viruses of the family Reoviridae to large, enveloped, double-stranded (ds)DNA viruses of the families Poxviridae (Moyer et al., 2000) and Herpesviridae (Minson et al., 2000).

In contrast to organisms, all of which are thought to have evolved from a common ancestral prokaryote, many types of viruses apparently originated and evolved independently of one another. This is implicit in the lack of a higher taxonomic classification than family for most types of viruses (van Regenmortel, 2000). Phenotypically, vertebrate viruses of the families Enteroviridae, Poxviridae and Flaviviridae, plant viruses of the families Tobamoviridae and Geminiviridae, and insect viruses of the families Iridoviridae, Baculoviridae and Ascaroviridae, each produce characteristic different virions and vary significantly in biology. Major differences among the characteristics of these viruses are suggestive of separate origins and evolutionary lineages. Nevertheless, there is evidence that some viruses of different families are related and thus may have evolved from a common ancestral virus. For example, algal viruses (family Phycodnaviridae), African swine fever virus (family Asfarviridae) and iridoviruses (family Iridoviridae) all produce large icosahedral virions containing an internal lipid membrane between the core and the capsid and a linear dsDNA genome (Dixon et al., 2000; Goorha & Murti, 1982; Heppel & Bethiaume, 1992; van Etten, 2000; Ward & Kalmakoff, 1991; Williams et al., 2000). Phylogenetic analyses of their DNA polymerase, major capsid protein and several other virally encoded structural proteins and enzymes suggest these viruses originated from a common ancestral icosahedral nucleocytoplasmic DNA virus (He et al., 2002; Iyer et al., 2001; Knopf, 1998; Stasiak et al., 2000), which also may be the evolutionary source of poxviruses (Iyer et al., 2001; Salas et al., 1999). There is little
evidence, however, that two virus families that produce virions with markedly different structural properties are closely related and evolved from a common ancestor, or that one evolved from another. Recent studies of DNA polymerases of large dsDNA viruses indicate that ascoviruses and iridoviruses are closely related (He et al., 2002; Knopf, 1998; Stasiak et al., 2000). However, phylogenetic analysis of this protein alone is insufficient to infer the evolutionary relationship of these viruses, and specifically whether they descended from a common ancestor or one evolved from the other. Thus, the purpose of this study was to clone a greater range of ascovirus genes and use these to examine the phylogenetic relationship of these viruses to other large dsDNA viruses, especially the iridoviruses, from which they differ so significantly in biology and virion structure.

Iridescence viruses cause acute fatal diseases among a diverse group of invertebrates and ectothermic vertebrates including nematodes, crustaceans, numerous insect species, amphibians and fishes (Ward & Kalmakoff, 1991; Williams et al., 2000). Their virions are icosahedral, large (> 125 nm in diameter), contain an internal lipid membrane located between the core and the capsid, and have a linear circularly permuted dsDNA genome which, depending on the species, ranges from 160 to greater than 200 kbp. Virions assemble in the cytoplasm where, in many invertebrates, they accumulate in paracrystalline arrays that impart an iridescent hue to infected hosts, from which the name of this group is derived (Williams et al., 2000). Mechanisms of transmission are not well understood, though transmission horizontally through wounds and vertically through eggs is known. Alternatively, ascoviruses cause a chronic fatal disease restricted to larvae and pupae of lepidopteran insects (Bigot et al., 1997a, b; Federici, 1983; Federici & Govindarajan, 1990; Federici et al., 1990, 2000). The disease is unique in that ascoviruses induce host cells to cleave into clusters of virion-containing vesicles (Federici, 1983). Virions are large, 300–400 nm long by 100–150 nm in diameter, enveloped, and vary in shape from ovoidal to bacilliform among different species. The genome is circular and, depending on the species, is from 115 to 180 kbp (Bigot et al., 1997a; Cheng et al., 1999; Federici et al., 2000). Another unusual and important feature of ascoviruses is that their principal mechanism of transmission from one lepidopteran host to another is by endoparasitic hymenopteran females during oviposition (Bigot et al., 1997b, Govindarajan & Federici, 1990; Hamm et al., 1985, 1986). Most invertebrate iridoviruses and ascoviruses do share one unusual property in that neither group replicates in midgut epithelial tissue (Federici, 1993).

In the present study, we provide phylogenetic evidence that ascoviruses originated from an iridovirus ancestor, probably an invertebrate iridovirus. This evidence is derived from three different lines of investigation. The first is a more extensive analysis of DNA polymerase genes from a wider range of dsDNA viruses including several ascoviruses. The second is a comparative analysis of three other viral proteins, the major capsid protein (MCP) and two enzymes, thymidine kinase (TK) and ATPase III (ATP) from a range of ascoviruses, iridoviruses and other large icosahedral dsDNA viruses including phycodnaviruses and ASFV. In addition, using more than 50 genes of Spodoptera frugiperda ascovirus (SFav-1a) sequenced recently, we show that the greatest number and the closest homologues to these occur in Chilo iridescent virus (CIV) (Invertebrate iridescent virus 6), a lepidopteran iridovirus, with decreasingly fewer homologues detected, respectively, in the vertebrate iridoviruses, phycodnaviruses and ASFV.

**METHODS**

**Viruses.** Ascoviruses studied were Spodoptera frugiperda ascovirus (SFav-1a), Trichoplusia ni ascovirus (TnAv-2a), Heliothis virescens ascovirus (HvAv-3b and c) and Diadromus pulchellus ascovirus (DpAv-4a), all described previously (Bigot et al., 1997a, b; Federici, 1983; Federici et al., 1990, 1997a, b; Pelleck et al., 1996; Stasiak et al., 2000). Regular mosquito iridovirus (RMIV; Invertebrate iridescent virus 3) was provided by J. Becnel (USDA, Gainesville, FL, USA). Virions and viral DNA were produced and purified as described (Bigot et al., 1997b; Federici et al., 1990).

**Random libraries of ascovirus DNA.** Approximately 20 μg of genomic DNA of SFav-1a, HvAv-3c or DpAv-4a was sonicated in 2–5 M NaCl, 10 mM Tris/HCl, pH 7–2, 1 mM EDTA for 2–5 min at 20 and power 10 (Vibra Cell 72446). The sonicated DNA was blunt-ended with S1 nuclease (5 min at 37 °C), and ligated with EcoRI adaptors (Promega). Fragments of approximately 1 kbp were amplified by PCR with a primer similar in sequence to the MCP gene was amplified from its genome by PCR using the following primers derived from the MCP sequence of HvAv-3c: 5′-GGGGTACCTATAGATACACTTATTTCTCAAAAG-3′ and 5′-GGGATCCCTTAGCTGCGGAG-3′. Amplification products were separated on an agarose gel, eluted with a Qiagen kit, and cloned using the pGEM-T Easy kit (Promega).

**Cloning genes.** The genes encoding the MCP of HvAv-3c and DpAv-4a were initially obtained from the random sequencing of 50 library fragments that were identified by BLAST searches of iridovirus MCP sequences in databases (Jakob et al., 2001; Tidona et al., 1998). Full genes were then obtained by inverse PCR (IPCR). The TnAv-2a MCP gene was amplified from its genome by PCR using the following primers derived from the MCP sequence of HvAv-3c: 5′-GGGATCCCTAGATACACTTATTTCTCAAAAG-3′ and 5′-GGGATCCCTTAGCTGCGGAG-3′. The genes encoding the DNA polymerase of HvAv-3c and RMIV were obtained by IPCR using primers designed from internal fragments previously described (Stasiak et al., 2000). Similar procedures were used to obtain genes encoding protein homologues of TK, RNase III and ATPase III in the SFav-1a, HvAv-3c and DpAv-4a genomes, as well as homologues of CIV open reading frames (ORF) 50L (CIV-O50L), 118L (CIV-O118L), 98R (CIV-O32R), 307L (CIV-O307L), 347L (CIV-O347L), 359L (CIV-O359L), 393L (CIV-O393L) in the SFav-1a and DpAv-4a genomes. SFav-1a, DpAv-4a and SeAv-5 homologues of ORFs 10R, 22R and 23R of SFav-1a were identified in our previous publication (Stasiak et al., 2000).

**IPCR.** Experiments were performed by standard procedures and conditions as previously described (Ausubel et al., 1994; Stasiak et al., 2000), employing specific nucleotide primers (25-mers) and restriction sites first defined from partial sequences obtained by sequencing clones from our random libraries. Briefly, viral DNA was digested with a specific restriction enzyme, DNA was diluted to 1 ng μl⁻¹, and fragments were circularized by incubation for 12 h with T4 DNA ligase.
DNA ligase at 16°C. Genes or gene fragments were amplified by PCR from the circularized DNA using specific primers and a Long Expand System kit (Boehringer Mannheim) following the manufacturer’s recommendations.

**Cloning of PCR or IPCR fragments.** Amplified fragments were purified after separation by agarose gel electrophoresis and eluted using the GeneClean II kit according to the manufacturer’s instructions (Bio 101). PCR fragments were cloned using a pGEM-T Easy kit (Promega).

**Sequencing.** Plasmids were isolated and purified by alkaline lysis followed by precipitation with LiCl. Double-stranded templates were used for dideoxynucleotide sequencing reactions with the SequiTherm Long-Read cycle sequencing kit (Epicentre Technologies) and labelled using universal and reverse IRD800-primers. A LI-COR 4000 unit was used for sequencing. Sequences reported here have been deposited in the DDBJ/EMBL/GenBank database. Accession numbers are provided in Table 1.

**Source of other gene sequences.** Sequences of other genes used in our analyses were obtained from the complete genomes of *Lymphophyctis disease virus* (LCDV), ASFV, Chlorella virus (CHV) and *Paramecium bursaria Chlorella virus 1* (PBCV-1), and partial sequences of Red sea bream virus (RSBV) and *Frog virus 3* (FV3) located in databases. Accession numbers of all these genes are provided in Table 1.

**Computer-assisted analyses.** The Infobiogen facilities were used for database searches, extractions, sequence alignments and calculations (http://www.lovelace.infobiogen.fr). Three programs were used to perform the protein sequence designations and alignments: tBlastN, BlastP and CLUSTALW. All sequence similarities identified and sequence alignments we developed were verified by comparison with their structural profiles determined by hydrophobic cluster analysis (HCA) as described (http://smi.snv.jussieu.fr/hca/hca-seq.html). This method enabled detection of local misalignments that most methods for evaluating optimal alignment quality using a Z-score test failed to validate (Callebaut et al., 1997). The phylogenetic analyses were performed using parsimony and neighbour-joining programs of the PHYLIP package, version 3.5c (Felsenstein, 1993). The default setting was used for all parsimony calculations. The neighbour-joining trees were derived from matrices of distances based on the category distance model (George et al., 1988). For each of the four data sets analysed, the consensus trees obtained with parsimony and neighbour-joining programs were similar. As in the case of previous phylogenetic analyses of amino acid sequences from dsDNA virus polymerases, all the consensus trees for all proteins analysed here were rooted with the genes of the most distantly related viruses (Stasiak et al., 2000). These were either the asfarivirus, ASFV, if the gene sequence was available or, if not, those of the phycodnavirus PBCV-1. In addition to computer-generated alignments, alignments were improved by direct examination. All alignments presented here have been deposited in the EMBL database under accession numbers ALIGN_#000102–000109. All trees were based on alignments of full-length sequences. However, aligned regions corresponding to long insertions or deletions in one or a few sequences were deleted. This was done to avoid over weighting with the PHYLIP package. When reporting results, the percent similarity refers to identity plus silent substitutions.

**RESULTS**

**DNA polymerase phylogeny**

In an earlier study (Stasiak et al., 2000), based on an analysis of DNA polymerase genes, we provided evidence that ascoviruses and iridoviruses are closely related, and more distantly related to the asfaravirus, ASFV, and phycodnaviruses. This study was carried out prior to the availability of any sequences from invertebrate iridoviruses and was based on only two ascovirus DNA polymerase sequences, those of SfAV-1a and DpAV-4a, and two vertebrate iridoviruses, RSBV and LCDV. To improve the rigour of our previous analysis, in the present study we added new DNA polymerase sequences from two invertebrate iridoviruses, one that we cloned (RMIV), and another obtained from a database (CIV; Jakob et al., 2001), and one from an ascovirus (HvAV-3c) that we sequenced. Thus, our present analysis was based on the DNA polymerases of three ascoviruses, SfAV-1a, HvAV-3c and DpAV-4a; four iridoviruses, CIV, LCDV, RMIV and RSBV; two phycodnaviruses, CHV and PBCV-1; and the asfarivirus, ASFV. In the resulting tree, ascovirus and iridovirus DNA polymerases clustered together on a branch clearly separated from those of ASFV and phycodnaviruses (Fig. 1a). Moreover, ascovirus DNA polymerases segregated along two branches, one bearing SfAV-1a and HvAV-3a, the other DpAV-4a.

This analysis confirmed that the families *Ascoviridae* and *Iridoviridae* are closely related, and more distantly related to the *Asfarviridae* and *Phycodnaviridae* (Stasiak et al., 2000). To further evaluate this relationship, we then analysed sequence data from three other sets of genes from a much greater range of viruses, specifically those encoding the MCP, TK and ATPase III, as reported below.

**Major capsid protein phylogeny**

Four ascovirus MCP genes were cloned and sequenced, those of SfAV-1a, TnAV-2a, HvAV-3c and DpAV-4a (Table 1). In all four, the MCP had a mass of approximately 50 kDa. The number of amino acids ranged from 434 in DpAV-4a to 461 in SfAV-1a, with intermediate lengths of 456 amino acids for both TnAV-2a and HvAV-3c. Comparison of the amino acid sequences showed that the MCP was well conserved among these ascoviruses. The MCKs of TnAV-2a and HvAV-3c MCP were highly similar, with a value of 98–7% at the amino acid level. The SfAV-1a MCP had similarity values of 66% when compared with, respectively, TnAV-2a and HvAV-3a. The DpAV-4a MCP was more distantly related to the latter ascoviruses, with an average similarity value of only 21%. Comparison showed that LCDV and CIV iridovirus MCPs were 22–24% similar to TnAV-2a and HvAV-3a, 23–24% similar to SfAV-1a and 30–32% similar to DpAV-4a. The sequence variability of the MCP within and between virus families was not randomly distributed. Seven conserved domains were identified in all 13 full MCP sequences analysed (four ascoviruses, seven iridoviruses, one phycodnavirus and one asfarivirus), and used for alignment (Fig. 2).

Phylogenetic analyses yielded trees with significant bootstrap values (>75%), indicating that ascovirus and iridovirus MCPs are closely related (Fig. 1b). When rooted using the ASFV MCP, the tree had the same branching as that
observed in the DNA polymerase analysis (Fig. 1a), showing ascovirus and iridovirus MCPs clustered on two separate but close branches.

**Thymidine kinase phylogeny**

TK genes were cloned and sequenced for three ascoviruses, SfAV-1a, HvAV-3c and DpAV-4a (Table 1). TK proteins ranged from 186 amino acids in DpAV-4a, to 210 for SfAV-1a, to 216 for HvAV-3c. Amino acid sequence comparisons showed that SfAV-1a, HvAV-3c and DpAV-4a TKs were, respectively, 55, 50 and 53 % similar to that of CIV.

For the phylogenetic analyses, the three ascovirus TK sequences were aligned with those of three iridoviruses, FV3, LCDV and CIV; one phycodnavirus, CHV; and the asfarvirus, ASFV. The tree topology was similar to the MCP tree; the ascoviruses clustered along two branches, SfAV-1a and HvAV-3c on one, and DpAV-4a on another (Fig. 1c). With this enzyme, however, the branching was somewhat different than that obtained for the MCP and the DNA polymerase in that the lepidopteran iridovirus, CIV, clustered on a major branch of the tree along with the ascoviruses (Fig. 1c). This relationship was supported by a bootstrap value of 82 % (Fig. 1c).

---

**Fig. 1.** Consensus trees resulting from phylogenetic analyses of four ascovirus and iridovirus proteins. (a) δ DNA polymerase; (b) major capsid protein; (c) thymidine kinase; and (d) ATPase III. The sequences are from one asfarvirus, *African swine fever virus* (ASFV), two phycodnaviruses, *Chlorella virus* (CHV) and *Paramecium bursaria Chlorella virus* (PBCV-1), nine iridoviruses, IV1, IV9, IV16, IV22, *Chilo iridescent virus* (CIV) (*Invertebrate iridescent virus 6*), *Frog virus 3* (FV3), *Lymphocystis disease virus* (LCDV), *Regular mosquito iridovirus* (RMIV) and *Red sea bream iridovirus* (RSBIV), and four ascoviruses, *Spodoptera frugiperda ascovirus 1a* (SfAV-1a), *Trichoplusia ni ascovirus 2a* (TnAV-2a), *Heliothis virescens ascovirus 3b* and 3c (HvAV-3b and 3c), and *Diadromus pulchellus ascovirus 4a* (DpAV-4a). Consensus trees were rooted with the corresponding proteins of ASFV in (a), (b) and (c), or CHV in (d). For each of the analysed proteins, the consensus trees obtained with parsimony and neighbour-joining programs were similar. The trees presented are those developed using the parsimony procedure. Numbers given at each node correspond to the percent bootstrap values for 1000 repetitions.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SfAV-1a</td>
<td>δ DNA polymerase MCP TK homologues</td>
<td>ATPase III homologues RNase III homologues</td>
<td>CIV-O50L homologues CIV-O118L homologues CIV-O307L homologues CIV-O393L homologues</td>
<td>CIV-O98R homologues CIV-O359L homologues</td>
</tr>
<tr>
<td>TnAV-2a</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>HvAV-3b</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>HvAV-3c</td>
<td>AJ312696*</td>
<td>AJ312704*</td>
<td>AJ312697*</td>
<td>AJ312692*</td>
</tr>
<tr>
<td>DpAV-4a</td>
<td>AJ279812</td>
<td>AJ312705*</td>
<td>AJ279812</td>
<td>AJ279815*</td>
</tr>
<tr>
<td>IV1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>IV9</td>
<td>M33542</td>
<td>AF025774</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>IV16</td>
<td>AF025775</td>
<td>AF025775</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>IV22</td>
<td>AF042341</td>
<td>AF042341</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>CIV (IV6)</td>
<td>AF303741</td>
<td>AF303741</td>
<td>AF303741</td>
<td>AF303741</td>
</tr>
<tr>
<td>FV3</td>
<td>U36913</td>
<td>U15575</td>
<td>M80551</td>
<td>P14358</td>
</tr>
<tr>
<td>LCDV</td>
<td>L63545</td>
<td>L63545</td>
<td>Abs. in L63545</td>
<td>Abs. in L63545</td>
</tr>
<tr>
<td>RMIV</td>
<td>AJ312708*</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>RSBIV</td>
<td>AB007366</td>
<td>NA</td>
<td>NA or Abs</td>
<td>NA or Abs</td>
</tr>
<tr>
<td>PBCV-1</td>
<td>AB011500</td>
<td>U42580</td>
<td>Abs. in U42580</td>
<td>Abs. in U42580</td>
</tr>
<tr>
<td></td>
<td>1–2843</td>
<td>202750–203313</td>
<td>Abs. in U42580</td>
<td>Abs. in U42580</td>
</tr>
</tbody>
</table>

NA, Data not available. Abs, gene absent from databases.
ATPase III phylogeny

ATPase III genes were cloned and sequenced from three ascoviruses, SfAV-1a, HvAV-3b and HvAV-3c (Table 1). All three proteins contained 233 amino acids. BLAST analysis showed that the SfAV-1a, and HvAV-3b and -3c proteins were, respectively, 64 and 65% similar, and all these were about 59% similar to RSBIV ATPase III.

The three ascovirus sequences were aligned with those of one invertebrate iridovirus, CIV, three vertebrate iridoviruses, RSBIV, LCDV and FV3, and one phycodnavirus, CHV. No ATPase III sequence is present in the ASFV genome and no ATPase III has been identified yet for DpAV-4a (as only 30% of the genome has been sequenced). The consensus tree yielded a topology very similar to that obtained for TK proteins; the ascoviruses clustered together, and close to CIV, but with the latter on a different branch of a major tree trunk (Fig. 1d).

An assessment of the phylogenic trees obtained for the four proteins analysed – DNA polymerase, MCP, TK and ATPase III – confirmed that ascoviruses and iridoviruses are closely related, and that they are more closely related to each other than to the asfarvirus and phycodnaviruses. In addition, the trees obtained with all four proteins showed that iridoviruses and ascoviruses shared a common ancestor, with the DNA polymerase and MCP trees suggesting that these viruses were two independent virus lineages that evolved from this ancestor, whereas the TK and ATPase III trees indicated that ascoviruses were more closely related to the lepidopteran iridovirus CIV than to the vertebrate iridoviruses. The latter finding suggested that ascoviruses originated from the invertebrate iridovirus lineage. Phylogenic analyses (unpublished data) done with homologues of the RNA polymerase subunit 2 (CIV-O428L) and CIV-O347L yielded similar results, although these analyses contained fewer representatives of ascoviruses (SfAV-1a; sequence AJ437060, positions 3624–4265 and 11223–14708) and iridoviruses (CIV and LCDV).

Considered together, the implication of these results was that analysis of the homologues of a few specific proteins could not resolve the phylogeny of ascoviruses and iridoviruses, as the analysis of one protein only reflects the history of this protein. Thus, the phylogenetic history of two different proteins in a cluster of related viruses can exhibit alternative phylogenies as a result of amino acid changes that occurred at different rates during evolution due to differences in selection pressures on these proteins in different lineages. However, aside from the evolution of viral genes via point mutations, viruses with a large dsDNA genome evolve by acquiring and losing genes (Bugert & Darai, 2000; Iyer et al., 2001). In essence, the presence or absence of these genes/proteins can be used as evolutionary markers of a lineage if several constraints are followed for their definition. First, the only proteins that can be used as markers are those present or absent when two lineages are compared, indicating the acquisition or loss of a protein with respect of one lineage to another. Second, genes
acquired must be present in all representative viruses of a lineage, and the proteins they encode must be more similar to each other than to their eukaryotic homologues. This criterion avoids use of marker proteins with features that result from interactions with their eukaryotic hosts (Bugert & Darai, 2000; Dall et al., 2001; Tidona & Darai, 2000).

Third, phylogenetic analyses of marker proteins in a sublineage must agree with those of genes used for the study of the complete lineage. These conditions have been applied to studying the evolution of herpesviruses (Alba et al., 2001), in which the need for rigorous conditions was demonstrated for correlating data on marker genes with phylogenies obtained for other genes used in the phylogenetic analyses.

To further clarify the evolutionary relationship of ascoviruses and iridoviruses, using the concept of marker proteins, we investigated a much greater range of proteins, determining the presence or absence of these in representatives of the large dsDNA viruses we studied. The presence or absence of specific proteins in different lineages indicates events of gene acquisition and deletion, and thus is useful in evaluating the evolutionary history of virus families. Here we used these marker proteins to further evaluate the hypothesis that ascoviruses evolved from iridoviruses.

**Proteins that mark the differentiation of ascoviruses from iridoviruses**

ATPase is an example of a marker protein for the families we studied. This protein occurs in phycodnaviruses, iridoviruses and ascoviruses, but not in asfarvirus. Moreover, the phylogenetic relationships of the former three virus types that it yields are similar to those obtained with DNA polymerase, MCP and TK. This protein can thus be used as a marker for differentiating the ASFV lineage from the others. We therefore searched for other marker proteins that could be used to differentiate the lineages of these virus families as they diverged from one another during evolution. To test this possibility, we used several known ascovirus genes to search databases for homologues among dsDNA viruses. This was possible owing to the availability of...
of the complete sequences for genomes of PBCV-1 (a phycodnavirus), ASFV (an asfarvirus), LCDV (a vertebrate iridovirus) and CIV (an invertebrate iridovirus). In addition, we sequenced a large number of genes from SfAV-1a and DpAV-4a. TnAV-2a and HvAV-3c could not be used in these searches because few of their genes have been sequenced.

The results revealed the presence of 11 marker proteins that could be used to identify four significant phases or steps of divergence in the asfarvirus–phycodnavirus–iridovirus–ascovirus lineage (Table 1 and Fig. 3). Thus, RNase III homologues were markers for differentiating the phycodnavirus–iridovirus–ascovirus lineage from asfarvirus, and homologues of CIV-O50L, O118L, O307L and O393L differentiated the iridovirus–ascovirus lineage from the phycodnaviruses. More interestingly, we found two marker genes that differentiated the invertebrate iridovirus–ascovirus lineage from the vertebrate iridoviruses (the homologues of CIV-O98R and CIV-O359L) and three specific markers that differentiated the ascovirus lineage from iridoviruses (sequence AJ437060, SfAV-1a-O10R, SfAV-1a-O22R and SfAV-1a-O23R). The results of this marker protein analysis provide additional evidence that ascoviruses evolved from invertebrate iridoviruses.

Identification of SfAV-1a homologues in other dsDNA viruses

To further test, using marker proteins, the putative origin of the ascoviruses from invertebrate iridoviruses, a search was undertaken for homologues of SfAV-1a proteins, identified as part of a SfAV-1a genome sequencing project, in the other viruses of the phycodnavirus–iridovirus–ascovirus lineage. Three contigs representing about 60% of the SfAV-1a genome (88 kbp for a genome of 145 kbp) were used. These corresponded to two fragments of 52 and 27 kbp, obtained by sequencing the SfAV-1a genome (accession no. AJ437059 and AJ437060) and one of 9 kbp containing the DNA polymerase gene (accession no. AJ279830). Computer-assisted analysis of DNA sequences revealed the presence of 321 potential ORFs with coding capacities ranging from 50 to 1162 amino acids. Analysis of these revealed that 33% (106 ORFs) were non-overlapping. Homologous proteins were identified for 50 of these 106 ORFs. The deduced proteins had the following closest homologues; 20 iridovirus proteins, 17 eukaryotic proteins, 5 br0-like proteins, and 5 adenovirus, chordopox and baculovirus proteins. Three of the proteins were unique to SfAV-1a. Recently, a very large icosahedral dsDNA virus has been reported from amoebae (La Scola et al., 2003), but the limited number of genes cloned and sequenced from this virus did not allow us to include it in our analyses.

If the hypothesis that iridoviruses and ascoviruses share a common ancestor is correct, then it would be expected that the closest ascovirus homologues would be found in invertebrate and vertebrate iridoviruses. We observed that of the 20 iridovirus homologues we detected, 17 of the SfAV-1a sequences were closest to the invertebrate iridovirus CIV. The percentage similarity of these SfAV-1a

---

**Fig. 3.** Hypothetical evolutionary pathway from phycodnaviruses through to the ascoviruses based on the genes that differentiate these virus types and were apparently acquired during evolution. Virion and genome characteristics that differentiate these viruses are on the right, with the genes acquired during evolution indicated by angled arrows on the left. Owing to the absence of homologues to many of these genes in ASFV, the relative position of Asfarvirus is not known.
sequences was lower when compared to those of the vertebrate iridovirus LCDV. In addition, the other 3 of the 20 iridovirus–ascovirus homologues were specific for the CIV and SfAV-1a genomes; i.e. no homologues to these were found in vertebrate iridoviruses, in other dsDNA viruses or in eukaryotic genomes. These sequences were SfAV-1a-O1L/CIV-O209L and SfAV-1a-O12L/CIV-O254L (in the AJ437060 sequence of SfAV-1a) and SfAV-1a-O18L/CIV-O295L (in the AJ437059 sequence). While it is realized that there is a lack of sequence data available from other iridoviruses, this nevertheless indicates that these three proteins may be markers for differentiating the invertebrate iridovirus–ascovirus lineage from the corresponding lineage of vertebrate iridoviruses (Fig. 3).

DISCUSSION

Previous studies of DNA polymerases of phycodnaviruses, iridoviruses, ascoviruses, and asfarviruses suggested these viruses were related and thus that they may have been derived from a common ancestral virus (He et al., 2002; Stasiak et al., 2000). The two viral types among these that appeared most closely related based on their DNA polymerases were ascoviruses and iridoviruses (He et al., 2002; Stasiak et al., 2000). This relationship was unexpected owing to significant differences between the structure of their virosomes and genomes, and because of major differences in their cell biology. Ascoviruses destroy the cell nucleus, assemble virosomes in the anucleate cell, and subsequently induce cell cleavage resulting in virion-containing vesicles (Federici, 1983; Federici & Govindarajan, 1990). In contrast, iridoviruses leave the nucleus intact, assemble virosomes in the cytoplasm, and cause comparatively little gross cellular pathology other than hypertrophy (Federici, 1993; Williams et al., 2000). Thus, our objective here was to evaluate more critically the unexpectedly close relationship of iridoviruses and ascoviruses, and furthermore to assess whether they evolved from common ancestor along separate lineages or whether one evolved from the other, with specific emphasis on testing the hypothesis that ascoviruses evolved from iridoviruses. These studies were made possible by the much greater range of DNA polymerase, MCP, TK and ATPase III sequences that have become available recently from a wide range of dsDNA viruses. In addition, we used the presence or absence of several marker proteins, some with known functions and others whose function is unknown, to infer putative phases in the differentiation from one another of the major viral families we studied.

Although we found variations in the topologies of the phylogenetic trees we developed for DNA polymerase, MCP, TK and ATPase III, two significant patterns emerged from our analyses. First, ascoviruses and iridoviruses were more closely related to each other than to the asfarvirus or phycodnaviruses. Second, the invertebrate iridovirus CIV clustered more closely with the ascoviruses than with any of the vertebrate iridoviruses (Fig. 1a–d). The latter could be used to infer that the ascoviruses evolved from the iridoviruses, especially based on the topologies of the TK and ATPase III trees. Alternatively, however, the trees for DNA polymerase and MCP suggested that these two virus types may have evolved along separate lineages from a common ancestor. Thus, to further assess the probable evolutionary pathway that led to ascoviruses, we examined the genomes of related dsDNA viruses for the presence or absence of genes that were apparently acquired and lost as major families evolved. This analysis enabled us to differentiate the apparent shift of the iridoviruses and ascoviruses from the phycodnaviruses as well as the differentiation of the ascoviruses from the invertebrate iridoviruses (Fig. 3). In further support of these evolutionary shifts, we examined ASPV and representatives of iridoviruses and phycodnaviruses for the presence of SfAV-1a homologues, and found that the greatest number of these occurred in CIV. Moreover, comparison of these homologues showed that the highest similarities occurred between those of SfAV-1a and CIV, with the similarities decreasing successively when the SfAV-1a proteins were compared, respectively, to phycodnavirus and ASFV homologues.

Aside from molecular data that provide evidence for ascoviruses evolving from iridoviruses, other evidence supporting this hypothesis is found in the mechanisms by which these viruses are transmitted to their hosts. An unusual property not found in other insect viruses but shared by iridoviruses and ascoviruses is that both are poorly infectious per os (Federici, 1983; Govindarajan & Federici, 1990). How iridoviruses are transmitted in nature is not well understood. Transovarial transmission is known, as is infection through wounds, but more relevant to the hypothesis that ascoviruses evolved from iridoviruses is the recent observation that an ichneumonid wasp, Eiphusona viticollle, is an efficient vector of a lepidopteran iridovirus in field populations of Spodoptera frugiperda (Lopez et al., 2002). In ascoviruses, parasitic wasps are highly efficient vectors in the laboratory and field, and this appears to be the primary mode by which these viruses are transmitted (Hamm et al., 1985, 1986). Thus, transmission biology is another important link between the iridoviruses and ascoviruses of lepidopterans.

Parasitic wasps are also known to transmit ichnoviruses (family Polydnaviridae) to caterpillars and other hosts (Webb et al., 2000), and evidence is mounting that ichnovirus particle proteins are completely encoded within the genome of their wasp vectors. The virions of ichnoviruses resemble those of ascoviruses in general structure and size (Federici, 1983). This raises the possibility that the ichnoviruses evolved from the iridoviruses or ascoviruses, the latter being more probable owing to the similarities in virion structure. At present, insufficient molecular data are available on ichnovirus structural proteins to make essential phylogenetic comparisons and infer how ichnoviruses may have originated during evolution. However, the molecular data presented here provide a basis for study of ichnovirus evolution in the near future. It is certainly
possible that these interesting particles so essential to parasitic wasps’ suppression of immune systems of their hosts had their evolutionary origin in the iridoviruses or ascoviruses, and thus earlier in the phycodnaviruses.

ACKNOWLEDGEMENTS

We thank Dr F. Rouleux-Bonnin, C. Augé-Gouillou, M.-H. Hamelin, D. K. Bideshi and J. J. Johnson for their assistance throughout our investigation. This work was supported by grants from the CNRS and the NATO to Y. Bigot and by US National Science Foundation Grant INT-9726818 to B. Federici.

REFERENCES


