JOURNAL OF VIROLOGY, Apr. 2004, p. 3244–3251 Vol. 78, No. 7
0022-538X/04/$08.00+0 DOI: 10.1128/JVI.78.7.3244–3251.2004

Ancient Coevolution of Baculoviruses and Their Insect Hosts

Elisabeth A. Herniou,[1,2]* Julie A. Olszewski,[1] David R. O'Reilly,[1]† and Jenny S. Cory[2]‡

Department of Biological Sciences, Imperial College London, Sir Alexander Fleming Building, London SW7 2AZ,[1] and Molecular Ecology and Biocontrol Group, NERC Centre for Ecology and Hydrology, Oxford OX1 3SR,[2] United Kingdom

Received 23 July 2003/Accepted 10 December 2003

If the relationships between baculoviruses and their insect hosts are subject to coevolution, this should lead to long-term evolutionary effects such as the specialization of these pathogens for their hosts. To test this hypothesis, a phylogeny of the *Baculoviridae*, including 39 viruses from hosts of the orders Lepidoptera, Diptera, and Hymenoptera, was reconstructed based on sequences from the genes *lef-8* and *ac22*. The tree showed a clear division of the baculoviruses according to the order of their hosts. This division highlighted the need to reconsider the classification of the baculoviruses to include one or possibly two new genera. Furthermore, the specialization of distinct virus lineages to particular insect orders suggests ancient coevolutionary interactions between baculoviruses and their hosts.

Coevolution, reciprocal evolution in interacting species driven by natural selection (54), is a major driving factor in the historical associations between pathogens and their hosts (13, 25, 59). Studies on the evolution of pathogen virulence and host resistance have shown that within populations both pathogens and hosts are able to adapt in response to the interactions (51, 59). However, there is much debate on how these microevolutionary scale changes can influence the patterns of speciation of the interacting species at macroevolutionary levels. Coevolution need not necessarily lead to the cospeciation of the interacting species. However, coevolutionary theories (54) all support the hypothesis that the processes of coadaptations would lead to a general trend of parasite specialization for their hosts (53), regardless of the age of the association. Retroviruses and herpesviruses, with their vertebrate hosts, are good examples of specialist pathogens for which coevolution leading to a certain level of cocladogenesis within subfamilies has been demonstrated (38–40).

The family *Baculoviridae* comprises a diverse group of arthropod-specific DNA viruses. They have been reported worldwide from over 600 host species (37), mostly from insects of the order Lepidoptera but also from the orders Diptera, Hymenoptera, and the crustacean order Decapoda (6, 16). The family *Baculoviridae* is currently subdivided into two genera based on several criteria, including the morphology of the occlusion bodies (OBs) and on mechanisms of nucleocapsid envelopment in infected cells (6). The genus *Nucleopolyhedrovirus* (NPV) is characterized by viruses forming polyhedral OBs, each containing many virions formed within the nucleus (49), whereas viruses of the genus *Granulovirus* (GV) typically produce ovoid OBs, with a single virion formed in the nucleocytoplasmic milieu (58). GVs have been described solely from lepidopteran hosts, whereas NPVs have been isolated from a wider range of arthropods. However, the taxonomic status of nonlepidopteran baculoviruses is still uncertain (6). Baculovirus phylogenies have usually been based on individual gene sequences. The polyhedrin/granulin (*polh*) gene, encoding the major matrix protein of the OBs, has been the most widely used (5, 7, 14, 34, 60), but other genes, such as *DNA polymerase*, *egt*, *gp41*, *chitinase*, *cathepsin*, and *lef2*, have also been utilized (7, 8, 10, 12, 29, 30, 34, 42). In general, these studies agree that the lepidopteran NPVs and GVs constitute distinct, well-defined groups (7, 14, 23, 24, 60).

Almost all phylogenetic studies have been based on sequences from lepidopteran baculoviruses. Mostly because of the rarity of the samples, little work has been done to try to investigate the position of nonlepidopteran baculoviruses. Resolving the relationships between viruses isolated from Hymenoptera, Diptera, and Lepidoptera would greatly enhance our understanding of the evolution of the virus family *Baculoviridae*. Early amino acid sequencing of the polyhedrin protein of *Neodiprion sertifer* NPV (NeseNPV) showed that the *polh* sequence of this hymenopteran virus is quite divergent from that of the lepidopteran viruses, including NPVs and GVs (50). This result has been confirmed by determination of the complete DNA sequence of the gene (60). These phylogenies based on the OB protein imply that the hymenopteran virus is from an ancient lineage. More recently, phylogenetic analyses based on the *p74* and *DNA polymerase* genes, including sequences from the dipteran virus *Culex nigripalpus* NPV (CuniNPV), also showed that this virus is very divergent from the lepidopteran viruses and that it is more ancestral (41). Similar results were obtained based on complete genome phylogenetic analyses (24). There are currently no baculovirus phylogenies including viruses from the three insect orders.

A comparison of nine complete genome sequences and their study in an evolutionary framework highlighted the genes that were most suitable for phylogenetic studies (23). Among them, two genes conserved in all the baculovirus genomes were chosen for the present study to address the phylogenetic relationships within the *Baculoviridae*. The gene *lef-8* encodes a sub-

* Corresponding author. Mailing address: Department of Biological Sciences, Imperial College London, Silwood Park, Ascot Berkshire SL5 7PY, United Kingdom. Phone: 44 2075942304. Fax: 44 2075942339. E-mail: e.herniou@imperial.ac.uk.

† Present address: Syngenta, Jealotts Hill International Research Station, Bracknell, Berkshire RG42 6EY, United Kingdom.

‡ Present address: Laboratory of Virology, Wageningen University, 6909 PD Wageningen, The Netherlands.

TABLE 1. Baculoviruses included in this study

Virus name[a]	Host order	Abbreviation	Collection source	GenBank accession no.[b]	Reference
***Culex nigripalpus* NPV**	Diptera	CuniNPV		AF4033738	1
Gilpinia hercyniae NPV239	Hymenoptera	GiheNPV239	NERC, CEH		This study
Gilpinia hercyniae NPVi7	Hymenoptera	GiheNPVi7	J. Cory	AY449800-779	This study
Gilpinia hercyniae NPVk14	Hymenoptera	GiheNPVK14	J. Cory		This study
***Neodiprion lecontei* NPV**	Hymenoptera	NeleNPV	B. Arif		
Neodiprion lecontei NPV726	Hymenoptera	NeleNPV726	NERC, CEH	AY449791-770	This study
Neodiprion sertifer NPV345	Hymenoptera	NeseNPV345	NERC, CEH	AY449785-765	This study
Neodiprion sertifer NPV411	Hymenoptera	NeseNPV411	NERC, CEH		This study
Neodiprion sertifer NPV413	Hymenoptera	NeseNPV413	NERC, CEH	AY449786	This study
Neodiprion sertifer NPVk5	Hymenoptera	NeseNPVk5	J. Cory		This study
Neodiprion sertifer NPV-Virox	Hymenoptera	NeseNPVv	NERC, CEH		This study
Abraxas grossulariata NPV112	Lepidoptera	AbgrNPV112	NERC, CEH	AY449781-761	This study
Achaea janata GV835	Lepidoptera	AcjaGV835	NERC, CEH	AY449793-772	This study
Actias selene NPV47	Lepidoptera	AcseNPV47	NERC, CEH	AY449789-768	This study
Antheraea polyphemus NPV30	Lepidoptera	AnpoNPV30	NERC, CEH	AY449783-763	This study
***Autographa californica* MNPV**	Lepidoptera	AcMNPV		L22858	4
***Bombyx mori* NPV**	Lepidoptera	BmNPV		L33180	18
Bombyx mori NPV460	Lepidoptera	BmNPV460	NERC, CEH	AY449788-767	This study
***Choristoneura fumiferana* MNPV**	Lepidoptera	ChMNPV	P. Krell	AF512031	
***Cydia pomonella* GV**	Lepidoptera	CpGV		U53466	35
***Epiphyas postvittana* MNPV**	Lepidoptera	EppoMNPV		AY043265	27
Harrisina brillians GVm2	Lepidoptera	HabrGVm2	B. Federici	AY449801-780	This study
***Helicoverpa armigera* SNPV**	Lepidoptera	HaSNPV		AF271059	9
***Helicoverpa zea* SNPV**	Lepidoptera	HzSNPV		AF334030	11
Helioconius erato NPV789	Lepidoptera	HeerNPV789	NERC, CEH	AY449792-771	This study
***Lymantria dispar* MNPV**	Lepidoptera	LdMNPV		AF081810	31
Lymantria monacha NPVb4	Lepidoptera	LymoNPVb4	J. Cory	AY449796-775	This study
***Mamestra configurata* NPVA**	Lepidoptera	MacoNPVA		AF467808	33
***Mamestra configurata* NPVB**	Lepidoptera	MacoNPVB		AY126275	32
Natada nararia GV254	Lepidoptera	NanaGV254	NERC, CEH	AY449782-762	This study
***Orgyia pseudotsugata* MNPV**	Lepidoptera	OpMNPV		U75930	2
Orgyia pseudotsugata NPVb5	Lepidoptera	OpNPVb5	J. Cory	AY449797-776	This study
***Phthorimaea operculella* GV**	Lepidoptera	PhopGV		AF499596	
Phthorimaea operculella GVc4	Lepidoptera	PhopGVc4	J. Cory	AY449799-778	This study
Pieris brassicae GV384	Lepidoptera	PbGV384	NERC, CEH	AY449787-766	This study
Pieris rapae GV95	Lepidoptera	PiraGV95	NERC, CEH	AY449794-773	This study
Plodia interpunctella GVb3	Lepidoptera	PiGVb3	J. Cory	AY449795-774	This study
***Plutella xylostella* GV**	Lepidoptera	PlxyGV		AF270937	20
***Rachiplusia ou* MNPV**	Lepidoptera	RoMNPV		AY145471	19
***Spodoptera exigua* MNPV**	Lepidoptera	SeMNPV		AF169823	28
Spodoptera littoralis GV66	Lepidoptera	SpliGV66	NERC, CEH	AY449790-769	This study
***Spodoptera litura* NPV**	Lepidoptera	SpltMNPV		AF325155	43
Trichoplusia ni NPVc2	Lepidoptera	TnNPVc2	NERC, CEH	AY449798-777	This study
Wiseana cervinata NPV344	Lepidoptera	WiceNPV344	NERC, CEH	AY449784-764	This study
***Xestia c-nigrum* GV**	Lepidoptera	XecnGV		AF162221	21

[a] Completely sequenced baculoviruses are shown in boldface type.
[b] GenBank accession numbers of the sequences obtained in this study are indicated as *lef-8-ac22*; only the last three digits of the *ac22* accessions are reported, as they have the same starting code as the *lef-8* accession numbers.

unit of the baculovirus RNA polymerase, and *ac22* encodes a per os infectivity factor (*pif-2*) (44). The *polh* gene was not considered for this study primarily because CuniNPV does not harbor a homologue of this gene (1). This suggests that other divergent baculoviruses might not possess a homologue of *polh* to encode their major OB protein. At the time of carrying out the analysis, 18 sequences, including that of CuniNPV, were available from the database for *lef-8* and *ac22*. We supplemented this information with 22 novel sequences for these two genes from lepidopteran and hymenopteran baculoviruses. This allowed the reconstruction of phylogenetic trees including baculoviruses isolated from hosts of the arthropod orders Lepidoptera, Hymenoptera, and Diptera to improve our understanding of the early evolution of the virus family *Baculoviridae*.

Traditionally, two competing evolutionary hypotheses have been put forward to explain the current host distribution of the baculoviruses (16). The first hypothesis states that baculoviruses could have evolved within one group of arthropods, such as the Lepidoptera, and switched to other insect groups (48). The second proposes that the association between baculoviruses and their hosts dates back to the origin of insects or even arthropods and that they coevolved during evolutionary time with the viruses colonizing the insect orders as they arose (by cocladogenesis) (16). We propose to examine these two hypotheses in this study with the reconstruction of a baculovirus

TABLE 2. Degenerate oligonucleotide primers used for amplification of a diverse range of baculoviruses

Gene	Oligonucleotide	Sequence[a]	Amino acid motif[b]	Size (bp)[c]
lef-8	L8F2	gtaaaacgacggccagtNNNACNRCNGARGAYCC	XTAEDP	450
	L8R2	aacagctatgaccatgMMNCCYTTYTGNCCRTG	HGQKGV	
ac22	Ac22F	gtaaaacgacggccagtGGWNNTGYATNSGNGARGAYCC	W(TSN)CI(AP)EDP	400
	Ac22R	aacagctatgaccatgRTYNCCRCANTCRCANRMNCC	G(EVF)C(ED)CG(DN)	

[a] M13 primer sequences are in lowercase type (this part of the primer allows for the direct sequencing of PCR products), and degenerate baculovirus primers are in uppercase type. R, A or G; Y, T or C; M, A or C; W, A or T; N, A, C, G, or T.
[b] Amino acid sequence corresponding to the primer site (single letter code, X = any).
[c] Expected size of the amplification product.

phylogeny including, for the first time, viruses from three distinct insect orders.

Furthermore, this study might shed new lights on the interrelationships between baculoviruses and question the phylogenetic validity of the present classification of the *Baculoviridae*, which divides the family into two genera.

MATERIALS AND METHODS

Molecular sequences. The samples examined for this study belonged to the historical insect virus collection held at the Natural Environment Research Council (NERC), Centre for Ecology and Hydrology (CEH), Oxford, England. They include nine isolates of hymenopteran baculoviruses from three sawfly host species (*Gilpinia hercyniae*, *Neodiprion lecontei*, and *N. sertifer*) and 17 lepidopteran baculoviruses, including 9 NPVs and 8 GVs (Table 1). OBs were dissolved in 10 mM NaOH (pH 12.5) for 15 min (modified from the method of Moser et al. [41]). The DNA was then purified with the DNAeasy kit (Qiagen).

Degenerate primers (Table 2) were designed to amplify fragments of the genes *lef-8* and *ac22* (*pif-2*). These primers included universal primer [M13(−20) and M13R] tails to allow for direct sequencing of the PCR products. PCR amplifications were performed with Ready-to-Go PCR beads (Amersham Pharmacia) under touchdown amplification cycles (95°C for 5 min; 94°C for 30 s, 55 to 43°C for 20 s, and 72°C for 30 s [3 times], down 3°C after each third cycle, for 15 cycles; 94°C for 30 s, 60°C for 20 s, and 72°C for 45 s [20 times]; and 72°C for 5 min). Successful amplifications were purified by using the Qiaquick PCR purification kit (Qiagen). Direct cycle sequencing of the entire PCR fragments was performed in both directions by using M13(−20) and M13R universal primers with the Big Dye terminator reaction mix (Applied Biosystems). The sequences were run on a 3700 ABI automated sequencer. Chromatograms were checked, and contiguous sequences were assembled in Sequencher 4.1 (Gene Codes Corporation). BLAST searches (3) were performed on all the new sequences to verify authenticity before any phylogenetic analyses were undertaken.

Phylogenetic analyses. The DNA sequences obtained from the PCR fragments were aligned in MacClade 4, based on amino acid coding (36), against sequences of the same genes from 18 completely sequenced baculoviruses available from the databases (Table 1). The alignments were trimmed to the size of the PCR fragments. They have been deposited in TreeBASE (http://www.treebase.org) under the accession numbers S1005 and M1697.

Maximum-likelihood (ML) analyses were performed in PAUP*, version 4.0b10 (52). Each alignment was analyzed by using a statistical model-fitting approach implemented in MODELTEST, version 3.06, to choose between substitution models (45, 46). The selected models were used to calculate a tree by using the neighbor-joining method under ML distances. This tree was then used to start an ML heuristic search including branch swapping by nearest-neighbor interchange to find shorter trees.

Bayesian phylogenetic analyses of the combined data set were conducted with MrBayes, version 3.0b4 (26). Five Markov chains were run for 1 million generations, and the ML parameters were estimated for each gene partition in every analysis. Trees were sampled every 100th generation; 1,000 trees obtained in the early phase of the analysis were discarded before computing the consensus of the remaining 9,001 trees to assess the posterior probability of each node.

The robustness of the tree topologies was also evaluated by bootstrap analysis under the following conditions: ML heuristic searches with 100 replicates and maximum-parsimony (MP) methods with 1,000 replicates. For the MP reconstructions, uninformative characters were excluded from the data matrices, the trees were built by stepwise addition, and tree bisection reconnection branch swapping was performed to find the best MP tree at each replication step. Differences in tree topologies were assessed by using one-tailed Kishino-Hasegawa (KH) and Shimodaira-Hasegawa (SH) tests implemented for ML tree scores in PAUP* (52).

RESULTS

Phylogenetic analyses. Phylogenetic analyses were performed to elucidate the relationships of the hymenopteran and dipteran baculoviruses within the *Baculoviridae*. Two genes were used in these analyses, *ac22* (*pif-2*) and *lef-8*. In total, sequences from 39 virus isolates were used (Table 1). Only three new sawfly sequences were used in the phylogenetic analyses, as sequences obtained from additional samples were very similar (345 and 413) or identical (i7 and K14) to each other between samples from the same host species. The *lef-8* PCR fragments produced an alignment of 513 nucleotides, which resulted in 74.8% parsimony informative and 18.5% constant sites. For *ac22*, the 357-nucleotide-long alignment resulted in 70% informative and 19% constant sites. A partition homogeneity test was performed in PAUP*. The resulting *P* value ($P = 0.27$) showed that both data sets were congruent and could be combined into one data set. Furthermore, assessment of the tree topologies obtain for both genes, with KH and SH tests, showed that they were not significantly incongruent (data not shown).

For the Lef-8 & Ac22 data set, the best-fit model of evolution selected by MODELTEST (45) was characterized by 9.4% of invariable sites (I) and a gamma shape parameter (G = 1.317%), which reflects the heterogeneity of variation rates across sites, and the substitution model had variable transition rates (A<>G = 2.08; T<>C = 2.82). The tree obtained from the combined alignment (Fig. 1, T1) showed the lepidopteran NPVs divided into two groups and the sawfly viruses with CuniNPV clearly separated from the GVs. However, it also showed that the GVs might be paraphyletic (i.e., split within the tree). To test whether this was strongly supported by the data, ML heuristic searches were performed again to find the most likely tree (T2) under the constraint that the GVs should be monophyletic (enforcing that the GVs should all derive from a single common ancestor). To measure the significance of the differences between T1 and T2, KH and SH tests were performed. The constrained tree (T2) had a likelihood value (−lnL = 19,508.2) only marginally lower than that of the best tree (T1) (−lnL = 19,507.9; delta = 0.32). The KH and SH tests, used to assess the differences between the topologies of T1 and T2, showed that they were not significantly different

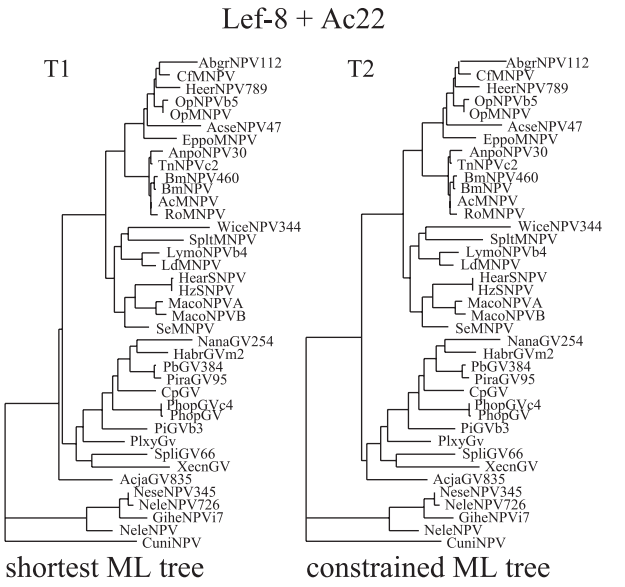

FIG. 1. ML trees obtained from the Lef-8 & Ac22 data set. T1, unconstrained tree; T2, GV monophyletic constraint tree. The trees were found by a heuristic search starting with neighbor joining and nearest-neighbor interchange branch swapping. Virus abbreviations are shown in Table 1.

(KH: $P = 0.416$; SH: $P = 0.975$; P value showing a significant difference with best tree, <0.05). This indicates that T2 represents a hypothesis for the evolution of these viruses that is equivalent to that presented by T1. Therefore, in view of the present classification and phylogeny of the baculoviruses, we believe that T2 represents a satisfactory hypothesis for the phylogeny of the baculoviruses. Bayesian phylogenetic analysis further confirmed this evolutionary hypothesis, as the majority rule consensus tree derived from this analysis showed the monophyly of the GVs. The topology of the Bayesian consensus tree was also found to be not significantly different from T1 or T2 by KH and SH tests (data not shown).

Robustness. The robustness of the phylogenies was evaluated by bootstrap analysis and analysis of Bayesian posterior probability. The backbone of the T2 tree is fairly well supported (Fig. 2). It shows that the lepidopteran NPVs are well separated from other baculoviruses, with separate group I NPVs being strongly supported; although T2 shows a monophyletic group II, there is little support for this group from any analysis. The separation of the nonlepidopteran baculoviruses at the base of the tree is well supported by high Bayesian posterior probability. The GVs as a whole are supported by 62% of the trees obtained in the Bayesian analysis. Within the GVs, a group including *Natada nararia* GV, *Harrisina brillians* GV, *Pieris brassicae* GV, *Pieris rapae* GV, *Cydia pomonella* GV, *Phthorimaea operculella* GVs, and *Plodia interpunctella* GV seems to detach within the group (Fig. 2). Terminal nodes leading to closely related viruses generally have high bootstrap and probability values with all methodologies of phylogenetic reconstruction. This includes the hymenopteran viruses, which strongly cluster together to the exclusion of other viruses. Within this group, the relationships are also well defined (Fig. 2).

DISCUSSION

Should the *Baculoviridae* comprise more genera? The family *Baculoviridae* is currently split into two genera, NPV and GV (6). So far, viruses from nonlepidopteran hosts have been classified within the NPV genus because their morphology and cytopathology fulfilled the criteria of this genus. However there is no evidence of the monophyly of this genus. Prior to this study, it was clear that the mosquito baculovirus CuniNPV is a distant relative to the lepidopteran baculoviruses and could represent a new genus (24, 41). There was also some indication that the sawfly virus NeseNPV is distantly related to the lepidopteran baculoviruses (60).

The phylogenetic analyses described here include baculoviruses from hosts of the arthropod orders Lepidoptera, Diptera, and Hymenoptera. They indicate that there are at least three, and possibly four, distinct groups of baculoviruses (Fig. 3A). The lepidopteran NPVs clearly form a discrete group, which is distinct from the rest of the baculoviruses. The branch leading to this group is quite long ($l = 0.26$) (Fig. 3A) and well supported by high bootstrap values and Bayesian posterior probabilities. The GVs, however, appear to be more genetically diverse than the NPVs, and the branch leading to the group is comparatively short ($l = 0.03$) and poorly supported (Fig. 2 and 3A). This suggests that the GVs are a much older group than the lepidopteran NPVs, that the sampling of the GVs was wider, or that both groups speciate at different speeds. Furthermore, the phylogeny shows that neither the dipteran virus nor the hymenopteran viruses belong to either the GVs or the lepidopteran NPVs. The branch separating the lepidopteran from the nonlepidopteran baculoviruses is long ($l = 0.36$) (Fig. 3A) and well supported. The branch lengths between the mosquito virus and the sawfly viruses are also large ($l = 0.88$ and 0.54) (Fig. 3A). This suggests that hymenopteran and dipteran baculoviruses probably belong to distinct and separate groups. The members of the *Baculoviridae* appear to be clearly divided according to the classification of their hosts.

In the past, baculoviruses (NPVs) have been reported from a wide variety of nonlepidopteran insects including three families of Coleoptera, six families of Diptera, four families of Hymenoptera, two families of Neuroptera, and one family of Trichoptera (37). The taxonomic status of most of these viruses remains uncertain. Most of them are rare, poorly characterized, correspond to isolates identified by light microscopy only, and have been removed from the International Committee on Taxonomy of Viruses (ICTV) baculovirus list for lack of molecular data.

Unsuccessful attempts were made to obtain sequences from virus isolates from lacewings (Neuroptera, Chrysopa PV-330, and Hemerobius NPV-318 and -440; NERC, CEH) and craneflies (Diptera, Tipula oleracea NPV-35 and −281; NERC, CEH) that had been classified as NPVs (22). These viruses might be extremely divergent baculoviruses beyond the range of our degenerate primers, but these results may also indicate that they belong to other virus families yet to be identified. No samples of crustacean baculoviruses were available for this study. A baculovirus of the shrimp *Peneaus monodon* is still included in the ICTV tentative baculovirus list (6). A recent morphological description of *Monodon baculovirus* might correspond to this virus (47). However, the lack of molecular

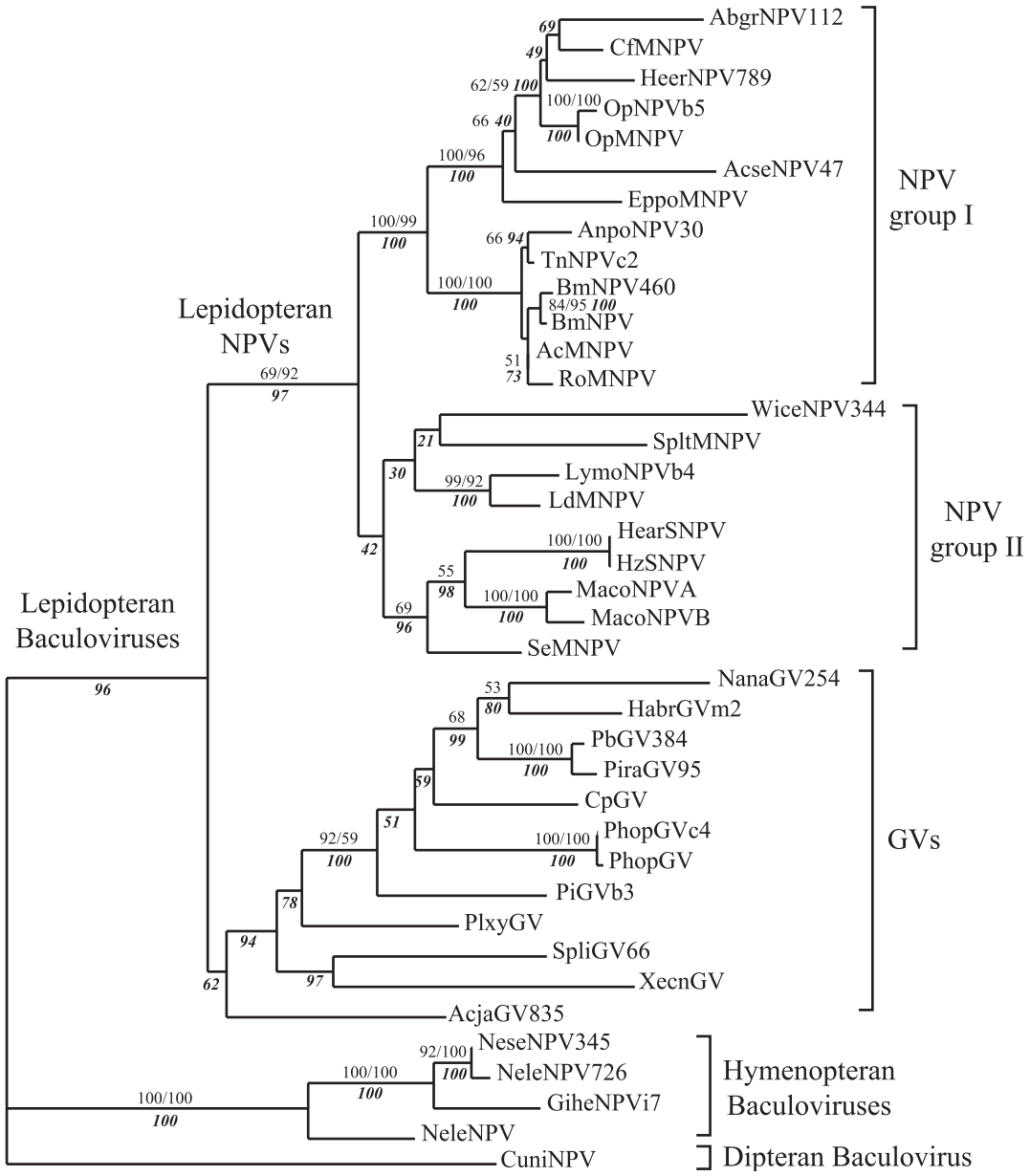

FIG. 2. Robustness of the Lef-8 & Ac22 ML tree (T2). Numbers in roman type (ML/MP ratio of >50) indicate bootstrap scores obtained by the ML method with 100 replicates. The second number, when present, indicates the score obtained by the MP method with 1,000 replicates. Numbers in bold italic type indicate the Bayesian posterior probabilities of the nodes.

sequences for this virus still casts doubt on its affiliation with the *Baculoviridae*, especially as another shrimp virus, the white spot syndrome virus, formerly classified as a baculovirus, is in the process of being classified to its own virus family, *Nimaviridae*, since sequences have become available (55, 56). Thus, it was not possible to confirm the presence of baculoviruses in the Crustacea, nor in insect orders other than Lepidoptera, Diptera, and Hymenoptera.

From the unrooted tree (Fig. 3A), it is possible to envisage a scenario where the GVs could have given rise to all the NPVs (lepidopteran and nonlepidopteran). This would reconcile the phylogeny with the present classification of the baculoviruses into two genera. However, evidence from DNA polymerase phylogenies (41) and comparative genomics studies including GVs, NPVs, and CuniNPV genomes (24) showed that the GVs and lepidopteran NPVs are more closely related to each other than they are to the mosquito virus. This therefore reinforces the idea that in a phylogenetic context the NPV genus might not include viruses from a nonlepidopteran background.

Taxonomic proposals. If the ICTV was to consider using phylogenetic concepts for the classification of baculoviruses, this would require the genera to be monophyletic. This study shows that under the present ICTV classification, the NPV genus is polyphyletic. So from a phylogenetic perspective, CuniNPV and the sawfly NPVs should be removed from the NPV genus and classified under the unclassified baculovirus

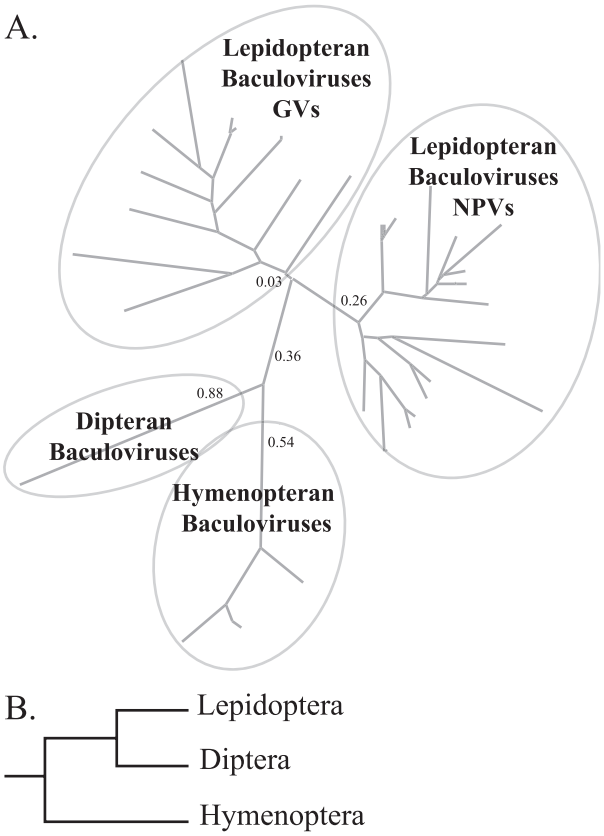

FIG. 3. Evolution of the *Baculoviridae*. (A) Phylogeny of the baculoviruses, highlighting four main groups of the unrooted tree (T2); numbers indicate branch lengths in substitutions per site. (B) Relationships of the three arthropod orders infected by the baculoviruses shown in panel A (57).

section. Our evidence would also support a taxonomic proposal to create one or two new genera of baculoviruses. The number of new genera would depend on further evidence to cluster together or keep apart the dipteran and hymenopteran baculoviruses.

BLAST search results showed that NeseNPV345, NeseNPV413, and NeseNPV726 are isolates of the species *N. sertifer* NPV (taxonomic code 00.006.0.01.017.). The other virus isolated from *N. lecontei* should still be classified as a separate species, *N. lecontei* NPV (taxonomic code 00.006.0.01.323.). Comparisons of phylogenetic distances and distinct host ranges (15) suggest that NPVs of *G. hercyniae* are part of a third distinct species.

Evolution of the *Baculoviridae*. Among baculovirologists, two views are commonly held for the evolutionary origins of the *Baculoviridae* (16). The first hypothesis proposes that the baculoviruses originated within the Lepidoptera, with subsequent horizontal transmissions to other insect orders from lepidopteran virus clades (48). The second postulates that the origin of baculoviruses dates back to the origin of arthropods, with the cocladogenesis of the viruses and their hosts (16).

If the first hypothesis was true, a phylogeny including baculoviruses from different orders of hosts would not show any clear clustering of baculoviruses according to host order. It would show nonlepidopteran viruses, either in clusters or singlets, arising within the NPVs or GVs, making lepidopteran NPVs or GVs paraphyletic. The phylogeny of baculoviruses including hosts from three different insect orders (Fig. 3A) seems to reject this hypothesis, as viruses strongly cluster according to the order of insects from which they have been isolated. We believe that our lepidopteran virus sampling is diverse enough to address this question. However, this does not exclude the discovery of a nonlepidopteran baculovirus belonging to the lepidopteran NPVs or GVs.

The second hypothesis would lead to a phylogeny where the relationships between groups of baculoviruses mirror the evolutionary relationships of insect orders, with the ages of the different baculovirus lineages reflecting those of their hosts. The orders Diptera and Lepidoptera are more closely related to each other than to the Hymenoptera (Fig. 3B) (57). The phylogeny obtained in this study could be consistent with the phylogenetic host tracking of insect orders by baculoviruses, as it is possible to root the tree on the hymenopteran baculovirus branch (Fig. 3A). However, further evidence needs to be gathered before accepting the hypothesis, particularly the comparison of evolutionary rates between baculoviruses and their hosts. From our data set, it is not possible to infer directly a reliable rate of sequence evolution for the baculoviruses, particularly because the genes used in this study do not have any homologues outside of the *Baculoviridae*.

A third scenario can be suggested for the origin of the baculoviruses. We propose that ancestral baculoviruses were probably able to horizontally infect hosts of different orders, with ancient coevolution between the hosts and pathogens then leading to the progressive specialization of different baculovirus lineages to hosts of different orders. According to this hypothesis, a phylogeny of the baculoviruses would show a clear separation of the viruses infecting different kinds of hosts without necessarily reflecting the evolution of insect orders. The phylogeny obtained in this study supports this hypothesis (Fig. 3A).

The uncertainty of the position of the root in the baculovirus phylogeny does not allow us to completely discard the second hypothesis in favor of the third. Some elements of baculovirus biology suggest that the dipteran baculovirus might belong to the more ancestral lineage, thus favoring the third scenario. The complete genome sequence of the mosquito virus CuniNPV showed that this virus does not possess a polyhedrin gene and that another protein is the major constituent of the OBs (1, 41), whereas polyhedrin sequences have been obtained from sawfly viruses (50, 60). Furthermore, lepidopteran and hymenopteran sawfly larvae share similar feeding ecologies and are often found in the same environment in the wild, i.e., terrestrial plants. They would thus be exposed to each others' viruses, whereas mosquito larvae are aquatic. This suggests that the lepidopteran and hymenopteran viruses could be more closely related to each other than to the mosquito virus. However, the dipteran virus lineage might have undergone a nonorthologous gene displacement for its OB protein.

In terms of pathogenesis and tissue tropism, mosquito and sawfly viruses are more similar to each other than to the lepidopteran GVs or NPVs. These viruses only infect midgut epithelial cells (15, 16, 41), whereas GVs and lepidopteran NPVs generally cause systemic infections, often infecting a wide range of tissues. The restriction of infection to midgut

epithelial cells has been proposed as an ancestral characteristic of baculoviruses (16). Only one lepidopteran baculovirus has been found to have a similar pathology, *H. brillians* GV (17). However, phylogenies including this virus showed that this virus is not basal to the GV group (Fig. 2) and, therefore, that the restriction of infection to the midgut epithelial cells is not an ancestral trait in lepidopteran baculoviruses (5). Thus, it is not possible to conclude whether the mosquito and sawfly viruses are more primitive or more derived than the lepidopteran baculoviruses based on the cell specificity of their infections. Although they were based on smaller taxon sets, previous phylogenetic studies suggested that they were more ancestral than the lepidopteran NPVs or GVs (41, 50, 60).

The relationships of the deeper branches of the baculovirus phylogenies might benefit in the near future from comparative genomic analyses. If sawfly virus genomes were found to be more similar to the lepidopteran baculoviruses, then the mosquito virus could remain at the base of the tree. If they share more genomic features with CuniNPV, then the hymenopteran and dipteran viruses could be grouped together to the exclusion of the lepidopteran baculoviruses. However, if CuniNPV is more similar to the lepidopteran baculoviruses, then the hymenopteran baculoviruses could be the more ancestral lineage. This last option would favor the second theory of early cospeciation between the *Baculoviridae* and the Arthropoda, as the baculovirus phylogeny would then reflect that of the order of their hosts, although this would need to be correlated with a comparative study of evolutionary rates between hosts and pathogens.

Regardless of the position of the root of the baculovirus tree, the phylogenetic separation of the viruses into groups according to the hosts' classification indicates that baculoviruses have been specialist pathogens of insects since the diversification of the family *Baculoviridae*. The selection pressure exerted on baculoviruses by their different hosts has promoted their specialization to the point where specific baculovirus lineages are specific to particular kinds of host insects, such as larvae of Lepidoptera, Hymenoptera, and Diptera. These coevolutionary adaptations have constrained the range of possible hosts available to each virus lineage. Over time, this has translated into the phylogenetic pattern that we observe today, where the *Baculoviridae* from different insect orders belong to different evolutionary lineages.

ACKNOWLEDGMENTS

We thank Basil Arif, Just Vlak, and Andy Purvis for discussion and comments on the manuscript and Hilary Lauzon and Peter Krell for sharing the sequences of *N. lecontei* NPV and *Choristoneura fumiferana* MNPV.

The Natural Environment Research Council CASE studentship award GT04/99/TS/142 supported E.A.H.

REFERENCES

1. **Afonso, C. L., E. R. Tulman, Z. Lu, C. A. Balinsky, B. A. Moser, J. J. Becnel, D. L. Rock, and G. F. Kutish.** 2001. Genome sequence of a baculovirus pathogenic for *Culex nigripalpus.* J. Virol. **75:**11157–11165.
2. **Ahrens, C. H., R. L. Q. Russell, C. J. Funk, J. T. Evans, S. H. Harwood, and G. F. Rohrmann.** 1997. The sequence of the Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus genome. Virology **229:**381–399.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
4. **Ayres, M. D., S. C. Howard, J. Kuzio, M. Lopez-Ferber, and R. D. Possee.** 1994. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. Virology **202:**586–605.
5. **Bideshi, D. K., Y. Bigot, and B. A. Federici.** 2000. Molecular characterization and phylogenetic analysis of the Harrisina brillians granulovirus granulin gene. Arch. Virol. **145:**1933–1945.
6. **Blissard, G. W., B. Black, N. Crook, B. A. Keddie, R. Possee, G. F. Rohrmann, D. A. Theilmann, and L. Volkman.** 2000. Family *Baculoviridae*, p. 195–202. *In* M. H. V. Van Regenmortel et al. (ed.), Virus taxonomy: seventh report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, Calif.
7. **Bulach, D. M., C. A. Kumar, A. Zaia, B. Liang, and D. E. Tribe.** 1999. Group II nuclepolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerases gene sequences. J. Invertebr. Pathol. **73:**59–73.
8. **Chen, X., W. F. J. Ijkel, C. Dominy, P. Zanotto, Y. Hashimoto, O. Faktor, T. Hayakawa, C.-H. Wang, A. Prekumar, S. Mathavan, P. J. Krell, Z. Hu, and J. M. Vlak.** 1999. Identification, sequence analysis and phylogeny of the *lef-2* gene of *Helicoverpa armigera* single-nucleocapsid baculovirus. Virus Res. **65:**21–32.
9. **Chen, X., M. Li, X. Sun, B. M. Arif, Z. Hu, and J. M. Vlak.** 2000. Genomic organization of Helicoverpa armigera single-nucleocapsid nucleopolyhedrovirus. Arch. Virol. **145:**2539–2555.
10. **Chen, X. W., Z. H. Hu, J. A. Jehle, Y. Q. Zhang, and J. M. Vlak.** 1997. Analysis of the ecdysteroid UDP-glucosyltransferase gene of Heliothis armigera single-nucleocapsid baculovirus. Virus Genes **15:**219–225.
11. **Chen, X. W., W. J. Zhang, J. Wong, G. Chun, A. Lu, B. F. McCutchen, J. K. Presnail, R. Herrmann, M. Dolan, S. Tingey, Z. H. Hu, and J. M. Vlak.** 2002. Comparative analysis of the complete genome sequences of Helicoverpa zea and Helicoverpa armigera single-nucleocapsid nucleopolyhedroviruses. J. Gen. Virol. **83:**673–684.
12. **Clarke, E. E., M. Tristem, J. S. Cory, and D. R. O'Reilly.** 1996. Characterization of the ecdysteroid UDP-glucosyltransferase gene from *Mamestra brassicae* nucleopolyhedrosis virus. J. Gen. Virol. **77:**2865–2871.
13. **Combes, C.** 2001. Les associations du vivant—l'art d'être parasite. Flammarion, Paris, France.
14. **Cowan, P., D. Bulach, K. Goodge, A. Robertson, and D. E. Tribe.** 1994. Nucleotide sequence of the polyhedrin gene region of Helicoverpa zea single nucleocapsid nuclear polyhedrosis virus: placement of the virus in lepidopteran nuclear polyhedrosis virus group II. J. Gen. Virol. **75:**3211–3218.
15. **Cunningham, J. C., and P. F. Entwistle.** 1981. Control of sawflies by baculovirus, p. 379–407. *In* H. D. Burges (ed.), Microbial control of pests and plant diseases. Academic Press, London, United Kingdom.
16. **Federici, B. A.** 1997. Baculovirus pathogenesis, p. 33–56. *In* L. K. Miller (ed.), The baculoviruses. Plenum Press, New York, N.Y.
17. **Federici, B. A., and V. M. Stern.** 1990. Replication and occlusion of a granulosis virus in larval and adult midgut epithelium of the western grapeleaf skeletonizer, *Harrisina brillians*. J. Invertebr. Pathol. **56:**401–414.
18. **Gomi, S., K. Majima, and S. Maeda.** 1999. Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus. J. Gen. Virol. **80:**1323–1337.
19. **Harrison, R. L., and B. C. Bonning.** 2003. Comparative analysis of the *Rachiplusia ou* and *Autographa californica* multiple nucleopolyhedroviruses. J. Gen. Virol. **84:**1827–1842.
20. **Hashimoto, Y., T. Hayakawa, Y. Ueno, T. Fujita, Y. Sano, and T. Matsumoto.** 2000. Sequence analysis of the Plutella xylostella granulovirus genome. Virology **275:**358–372.
21. **Hayakawa, T., R. Ko, K. Okano, S.-I. Seong, C. Goto, and S. Maeda.** 1999. Sequence analysis of the *Xestia c-nigrum* granulovirus genome. Virology **262:**277–297.
22. **Herniou, E. A.** 2003. Use of comparative genomics and phylogenetics to study the evolution of the *Baculoviridae.* Ph.D. thesis. Imperial College, London, United Kingdom.
23. **Herniou, E. A., T. Luque, X. Chen, J. M. Vlak, D. Winstanley, J. S. Cory, and D. O'Reilly.** 2001. Use of whole genome sequence data to infer baculovirus phylogeny. J. Virol. **75:**8117–8126.
24. **Herniou, E. A., J. A. Olszewski, J. S. Cory, and D. R. O'Reilly.** 2003. The genome sequence and evolution of baculoviruses. Annu. Rev. Entomol. **48:**211–234.
25. **Herrera, C. M., and O. Pellmyr.** 2002. Plant-animal interactions an evolutionary approach. Blackwell Science Ltd., Oxford, United Kingdom.
26. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics **17:**754–755.
27. **Hyink, O., R. A. Dellow, M. J. Olsen, K. M. B. Caradoc-Davies, K. Drake, E. A. Herniou, J. S. Cory, D. R. O'Reilly, and V. K. Ward.** 2002. Whole genome analysis of the Epiphyas postvittana nucleopolyhedrovirus. J. Gen. Virol. **83:**959–973.
28. **IJkel, W. F. J., E. A. van Strien, J. G. M. Heldens, R. Broer, D. Zuidema, R. W. Goldbach, and J. M. Vlak.** 1999. Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. J. Gen. Virol. **80:**3289–3304.
29. **Jin, T., Y. Qi, D. Liu, and F. Su.** 1999. Nucleotide sequence of a 5892 base pairs fragment of the LsMNPV genome and phylogenetic analysis of Ls-MNPV. Virus Genes **18:**265–276.
30. **Kang, W., M. Tristem, S. Maeda, N. E. Crook, and D. R. O'Reilly.** 1998.

Identification and characterization of the *Cydia pomonella* granulovirus cathepsin and chitinase genes. J. Gen. Virol. **79:**2283–2292.
31. **Kuzio, J., M. N. Pearson, S. H. Harwood, C. J. Funk, J. T. Evans, J. M. Slavicek, and G. F. Rohrmann.** 1999. Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria dispar.* Virology **253:**17–34.
32. **Li, L. L., C. Donly, Q. J. Li, L. G. Willis, B. A. Keddie, M. A. Erlandson, and D. A. Theilmann.** 2002. Identification and genomic analysis of a second species of nucleopolyhedrovirus isolated from *Mamestra configurata.* Virology **297:**226–244.
33. **Li, Q. J., C. Donly, L. L. Li, L. G. Willis, D. A. Theilmann, and M. Erlandson.** 2002. Sequence and organization of the Mamestra configurata nucleopolyhedrovirus genome. Virology **294:**106–121.
34. **Liu, J., and J. E. Maruniak.** 1999. Molecular characterization of genes in the GP41 region of baculoviruses and phylogenetic analysis based upon GP41 and polyhedrin genes. Virus Res. **64:**187–196.
35. **Luque, T., R. Finch, N. Crook, D. R. O'Reilly, and D. Winstanley.** 2001. The complete sequence of the Cydia pomonella granulovirus genome. J. Gen. Virol. **82:**2531–2547.
36. **Maddison, D. R., and W. R. Maddison.** 2000. MacClade 4. Sinauer Associates, Sunderland, Mass.
37. **Martignoni, M. E., and P. J. Iwai.** 1981. A catalogue of viral diseases of insects, mites, and ticks, p. 897–911. *In* H. D. Burges (ed.), Microbial control of pests and plant diseases 1970–1980. Academic Press, Inc., London, United Kingdom.
38. **Martin, J., E. Herniou, J. Cook, R. W. O'Neill, and M. Tristem.** 1999. Interclass transmission and phyletic host tracking in murine leukemia virus-related retroviruses. J. Virol. **73:**2442–2449.
39. **Martin, J., P. Kabat, and M. Tristem.** 2003. Cospeciation and horizontal transmission rates in the murine leukemia-related retroviruses, p. 174–194. *In* R. D. M. Page (ed.), Tangled trees: phylogeny, cospeciation and coevolution. University of Chicago Press, Chicago, Ill.
40. **Mcgeoch, D. J., A. Dolan, and A. C. Ralph.** 2000. Toward a comprehensive phylogeny for mammalian and avian herpesviruses. J. Virol. **74:**10401–10406.
41. **Moser, B. A., J. J. Becnel, S. E. White, C. Afonso, G. Kutish, S. Shanker, and E. Almira.** 2001. Morphological and molecular evidence that Culex nigripalpus baculovirus is an unusual member of the family Baculoviridae. J. Gen. Virol. **82:**283–297.
42. **Nielsen, C. B., D. Cooper, S. M. Short, J. H. Myers, and C. A. Suttle.** 2002. DNA polymerase gene sequences indicate western and forest tent caterpillar viruses form a new taxonomic group within baculoviruses. J. Invertebr. Pathol. **81:**131–147.
43. **Pang, Y., J. X. Yu, L. H. Wang, X. H. Hu, W. D. Bao, G. Li, C. Chen, H. Han, S. N. Hu, and H. M. Yang.** 2001. Sequence analysis of the Spodoptera litura multicapsid nucleopolyhedrovirus genome. Virology **287:**391–404.
44. **Pijlman, G. P., A. J. Pruijssers, and J. M. Vlak.** 2003. Identification of *pif-2*, a third conserved baculovirus gene required for *per os* infection of insects. J. Gen. Virol. **84:**2041–2049.
45. **Posada, D., and K. A. Crandall.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817–818.
46. **Posada, D., and K. A. Crandall.** 2001. Selecting the best-fit model of nucleotide substitution. Syst. Biol. **50:**580–601.
47. **Ramasamy, P., P. R. Rajan, V. Purushothaman, and G. P. Brennan.** 2000. Ultrastructure and pathogenesis of *Monodon baculovirus* (PmSNPV) in cultured larvae and natural brooders of *Peneaus monodon.* Aquaculture **184:** 45–66.
48. **Rohrmann, G. F.** 1986. Evolution of occluded baculoviruses, p. 203–215. *In* R. Granados and B. Federici (ed.), The biology of baculoviruses, vol. 1. CRC Press, Inc., Boca Raton, Fla.
49. **Rohrmann, G. F.** 1999. Nuclear polyhedrosis viruses, p. 146–152. *In* R. G. Webster and A. Granoff (ed.), Encyclopedia of virology, 2nd ed. Academic Press, London, United Kingdom.
50. **Rohrmann, G. F., M. N. Pearson, T. J. Bailey, R. R. Becker, and G. S. Beaudreau.** 1981. N-terminal polyhedrin sequences and occluded baculovirus evolution. J. Mol. Evol. **17:**329–333.
51. **Stearns, S. C.** 1999. Evolution in health and disease. Oxford University Press, New York, N.Y.
52. **Swofford, D. L.** 2001. PAUP*. Phylogenetic analysis using parsimony (*and other methods), 4th ed. Sinauer Associates, Sunderland, Mass.
53. **Thompson, J. N.** 1994. The coevolutionary process. University of Chicago Press, Chicago, Ill.
54. **Thompson, J. N.** 1999. What we know and do not know about coevolution: insect herbivores and plant as a test case, p. 7–30. *In* H. Olff, V. K. Brown, and R. H. Drent (ed.), Herbivores: between plants and predators. Blackwell Science Ltd., Oxford, United Kingdom.
55. **van Hulten, M. C., J. Witteveldt, S. Peters, N. Kloosterboer, R. Tarchini, M. Fiers, H. Sandbrink, R. K. Lankhorst, and J. M. Vlak.** 2001. The white spot syndrome virus DNA genome sequence. Virology **286:**7–22.
56. **van Hulten, M. C. W., M.-F. Tsai, C. A. Schipper, C.-F. Lo, G.-H. Kou, and J. M. Vlak.** 2000. Analysis of a genomic segment of white spot syndrome virus of shrimp containing ribonucleotide reductase genes and repeat regions. J. Gen. Virol. **81:**307–316.
57. **Whiting, M. F.** 2002. Phylogeny of the holometabolous insect orders based on 18S ribosomal DNA: when bad things happen to good data, p. 69–84. *In* R. DeSalle, G. Giribet, and W. Wheeler (ed.), Molecular systematics and evolution: theory and practice. Birkhauser Verlag, Basel, Switzerland.
58. **Winstanley, D., and D. O'Reilly.** 1999. Granuloviruses, p. 140–146. *In* R. G. Webster and A. Granoff (ed.), Encyclopedia of virology, 2nd ed. Academic Press, London, United Kingdom.
59. **Woolhouse, M. E. J., J. P. Webster, E. Domingo, B. Charlesworth, and B. R. Levin.** 2002. Biological and biomedical implications of the coevolution of pathogens and their hosts. Nat. Genet. **32:**569–577.
60. **Zanotto, P. M. D., B. D. Kessing, and J. E. Maruniak.** 1993. Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. J. Invertebr. Pathol. **62:**147–164.