View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Virology 346 (2006) 180 – 193
VIROLOGY
www.elsevier.com/locate/yviro

Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera

Johannes A. Jehle [a,*], Martin Lange [a], Hualin Wang [a,b], Zhihong Hu [b], Yongjie Wang [a], Rüdiger Hauschild [a]

[a] *Laboratory for Biotechnological Crop Protection, Department of Phytopathology, Agricultural Service Center Palatinate (DLR Rheinpfalz), Breitenweg 71, 67435 Neustadt an der Weinstraße, Germany*

[b] *State Key Laboratory of Virology and Joint Laboratory of Invertebrate Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, People's Republic of China*

Received 30 August 2005; returned to author for revision 27 September 2005; accepted 25 October 2005
Available online 28 November 2005

Abstract

PCR amplification of the highly conserved baculovirus genes *late expression factor 8* (*lef-8*), *late expression factor 9* (*lef-9*) and *polyhedrin/granulin* (*polh/gran*) combined with molecular phylogenetic analyses provide a powerful tool to identify lepidopteran-specific baculoviruses and to study their diversity. In the present investigation, we have improved the degenerate oligonucleotides and corroborated the approach that was recently described by Lange et al. (Lange, M., Wang, H., Zhihong, H., Jehle, J.A., 2004. Towards a molecular identification and classification system of lepidopteran-specific baculoviruses. Virology 325, 36–47.). Baculovirus DNA was isolated from 71 uncharacterized historic baculovirus samples, and partial gene sequences were amplified by using gene-specific degenerate PCR primers. The obtained PCR products were directly sequenced, and the deduced amino acid sequences were compiled and aligned with published sequences of these target genes. A phylogenetic tree of 117 baculoviruses was inferred using maximum parsimony and distance methods. Based on the comprehensive phylogenetic analysis of the partial *lef-8*, *lef-9* and *polh/gran* genes, we propose a phylogenetic species criterion for lepidopteran-specific baculoviruses that uses the genetic distances of these genes for species demarcation.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Baculoviridae; Nucleopolyhedrovirus; Granulovirus; Virus identification; Molecular phylogeny; Polymerase chain reaction

Introduction

Baculoviruses are arthropod-specific, enveloped, rod-shaped viruses with a circular double-stranded DNA genome that replicate in the nucleus of infected host cells. To date, more than 600 baculoviruses have been described that infect species from the insect orders Lepidoptera, Diptera and Hymenoptera, and it is likely that baculoviruses represent the largest and most diverse family of DNA viruses (Martignoni and Iwai, 1981). More than 90% of the presently known baculoviruses were isolated from lepidopteran species (Martignoni and Iwai, 1981). The present classification of the family Baculoviridae into two genera, the Nucleopolyhedroviruses (NPVs) and Granuloviruses (GVs), is mainly based on the morphology of their occlusion bodies (OBs). NPVs have OBs with many virions and have been isolated from lepidopteran and non-lepidopteran hosts. In contrast, granuloviruses (GVs) OBs contain a single virion and were isolated only from lepidopteran insects (Theilmann et al., 2005). This classification into NPVs and GVs has been recently challenged by genome sequence analyses demonstrating that the NPVs from Diptera and Hymenoptera represent ancient groups that are phylogenetically separated from the monophyletic lepidopteran-specific NPVs and GVs (Afonso et al., 2001; Garcia-Maruniak et al., 2004; Lauzon et al., 2004; Herniou et al., 2004).

Despite the large number of baculoviruses in nature, only a small fraction of these viruses has been studied in any detail. Fully sequenced genomes have been deposited in Genbank for only 24 lepidopteran-specific NPVs and GVs as well as the two hymenopteran-specific *Neodiprion sertifer* (Nese) NPV and *N. lecontei* (Nele) NPV and one dipteran-specific *Culex nigripal-*

* Corresponding author. Fax: +49 6321 671222.
E-mail address: Johannes.Jehle@dlr.rlp.de (J.A. Jehle).

0042-6822/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2005.10.032

pus (Cuni) NPV. Knowledge on the genetic relationship of the majority of baculoviruses is still fragmentary. Recently, PCR-based methods were established, which allow fast and reliable identification of baculovirus isolates. By comparing homologues of the *late expression factor 8* (*lef-8*), *lef-9* and *polyhedrin* (*polh*) of *Autographa californica* (Ac) MNPV, highly conserved DNA regions were identified and were used as targets for degenerate PCR primers (Lange et al., 2004). These oligonucleotides allowed amplification of the target regions from a diverse range of baculoviruses. Recently, degenerate primers also were developed for *ac22* homologues, which encode *per os infectivity factor 2* (*pif-2*) (Pijlman et al., 2003; Herniou et al., 2004).

The *lef-8* and *lef-9* genes encode for subunits of the baculovirus RNA polymerase, which initiates transcription from late and very late promoters (Guarino et al., 1998). *Lef-8* and *lef-9* were identified in all completely sequenced baculovirus genomes and were previously shown to be suitable for studying baculovirus phylogeny (Herniou et al., 2001, 2003; Lange et al., 2004). The *polh* gene encodes polyhedrin, the major matrix protein of the occlusion bodies (OBs) in NPVs. Its homologue in GVs is *gran* that encodes granulin. The *polh/gran* gene is one of the most conserved genes in lepidopteran-specific baculoviruses and has been used in several phylogenetic studies (Rohrmann, 1986; Zanotto et al., 1993; de Moraes and Maruniak, 1997). Based on the *polh* phylogeny, the separation of lepidopteran-specific NPVs into groups I and II was proposed (Zanotto et al., 1993). This grouping correlates also with the utilization of two different envelope fusion proteins by groups I and II NPVs (Pearson and Rohrmann, 2002). Recently, it was suggested that the *polh* from AcMNPV (group I) is a chimeric gene that resulted from recombination with a group II NPV (Jehle, 2004). This exemplified the limited utility of single genes for phylogenetic studies. On the other hand, it was shown that phylogenetic analyses based on a combined data set of concatenated partial *polh/gran*, *lef-8* and *lef-9* sequences were congruent with a tree derived from 30 concatenated baculovirus core genes of 22 completely sequenced baculovirus genomes. This demonstrated the usefulness of these gene fragments in phylogenetic analysis and species identification (Lange et al., 2004). In the present study, we describe the partial *lef-8*, *lef-9* and *polh/gran* sequences of 71 hitherto uncharacterized NPV and GV isolates from historic baculovirus collections. The phylogenetic analyses based on the concatenated sequences as well as their potential for identification and species classification are discussed.

Results and discussion

The PCR approach for baculovirus identification reported by Lange et al. (2004) was validated by using a wide range of different historic virus samples from a broad host range and various geographic origins. The degenerate oligonucleotides used in the present study remained unchanged in their gene-specific binding region from those described previously. However, the sequencing and analyses of the *lef-8* PCR products were simplified by introducing a new extension (T7 universal sequencing primer sequence) for direct sequencing (Table 1). In some cases, when the *polh/gran* fragments generated by prPH-1 and prPH-2 could not be sequenced, prPH-1 was replaced with prPH-1B, and the PCR was repeated. The prPH-1B contained a modified sequencing primer extension with the universal BGHrev sequencing primer tag. The *lef-9*-specific primers prL9-1 and prL9-2 remained unchanged (Table 1).

The partial sequences of *polh/gran*, *lef-8* and *lef-9* genes were determined for the first time for 48 NPV and 23 GV historic samples. The virus names, abbreviations and accession numbers for the partial *polh/gran*, *lef-8* and *lef-9* sequences are listed in Table 2. Sequences obtained for the *polh/gran* gene were 510 bp long, representing 2/3 of the complete *polh* open reading frame (ORF). Amplified sequences for *lef-8* differed in length between 670 and 800 bp, whereas the full-length ORF of completely sequenced *lef-8* genes ranges from 2484 bp (*Phthorimea operculella* (Phop) GV, NC004032) to 2757 bp (*Spodoptera litura* (Splt) NPV) (Pang et al., 2001). The sequences obtained for the *lef-9* gene were about 260 bp long, while the length of the complete *lef-9* ORF is 1470 bp to 1560 bp. The virus samples were obtained from 55 different host species belonging to 16 lepidopteran families. Two species, *Agrotis segetum* and *Hyphantria cunea*, were hosts to both NPV and GV. The wide range of baculoviruses that originated from a large number of different host species and lepidopteran host families underline the universal nature of the degenerate primers used for gene fragment amplification. Thus, a reliable identification and phylogenetic analysis of these viruses

Table 1
Degenerate oligonucleotide primers used for PCR

Target gene	Primer name	AcMNPV genome position	Sequencing extension	Primer Sequence[a,b]
Polh/gran	prPH-1	4684–4698	M13 forward	TGTAAAACGACGGCCAGTNRCNGARGAYCCNTT
	prPH-2	5210–5224	M13 reverse	CAGGAAACAGCTATGACCDGGNGCRAAYTCYTT
	prPH-1B	4684–4698	BGHrev	TAGAAGGCACAGTCGAGGNRCNGARGAYCCNTT
lef-8	prL8-1	42,075–42,088	M13 reverse	CAGGAAACAGCTATGACCCAYGGHGARATGAC
	prL8-2	41,373–41,389	M13 reverse	CAGGAAACAGCTATGACCAYRTASGGRTCYTCSGC
	prL8-1B	42,075–42,088	T7	TAATACGACTCACTATAGGGCAYGGHGARATGAC
lef-9	prL9-1	49,748–49,763	M13 reverse	CAGGAAACAGCTATGACCAARAAYGGITAYGCBG
	prL9-2	50,027–50,043	M13 forward	TGTAAAACGACGGCCAGTTTGTCDCCRTCRCARTC

[a] B = C, G or T; D = A, G or T; H = A, C or T; I = Inosin; N = C, A, T or G; R = A or G; S = C or G; Y = C or T.
[b] Underlined nucleotides indicate standard sequencing primers M13 forward, M13 reverse, T7 and BGHrev.

Table 2
Isolates of nucleopolyhedroviruses (NPVs) and granuloviruses (GVs) analyzed in this study

Virus name	Abbreviation	Host family	Accession no. *polh/gran*	Accession no. *lef-8*	Accession no. *lef-9*	AT[a] (%)	Isolate[b]	Origin[c]
Actias selene NPV	AcseNPV	*Saturniidae*	AY706680	AY706527	AY706592	48.4	S1	IVCAS 1.105
Actias selene NPV	AcseNPV	*Saturniidae*	AY706678	AY706525	AY706590	48.3	S2	IVCAS 1.239
Agraulis sp. NPV	Ag. sp. NPV	*Nymphalidae*	AY706682	AY706532	AY706597	49.8	M34-3	Steinhaus collection
Agrotis segetum NPV	AgseNPV	*Noctuidae*	AY706683	AY706535	AY706600	51.9	A12-3	BBA-165: from Burgerjon (ca. 1978)
Amorbia cuneacapsa NPV	AmcuNPV	*Tortricidae*	AY706685	AY706537	AY706602	54.7	A8-3	BBA
Anagrapha falcifera NPV	AnfaNPV	*Noctuidae*	AY706686	AY706539	AY706604	61.3	A5-3	BBA-235
Antheraea pernyi NPV	AnpeNPV	*Saturniidae*	DQ231347	DQ231348	DQ231349	48.4	S4	IVCAS 1.5
Antheraea pernyi NPV	AnpeNPV	*Saturniidae*	AY706687	AY706540	AY706605	48.4	S5	IVCAS 1.29
Apocheima cinerarium NPV	ApciNPV	*Geometridae*	AY706688	AY706541	AY706606	65.3	S7	IVCAS 1.194
Archips rosanus NPV	ArroNPV	*Tortricidae*	n.d.	DQ231335	DQ231334	–	A8-2	BBA-207, n.d.a.
Autographa biloba NPV	AubiNPV	*Noctuidae*	AY737724	AY737722	AY737723	59.8	A23-1	Steinhaus collection
Autographa californica MNPV	AcMNPV	*Noctuidae*	AY706681	AY706528	AY706593	59.8	A12-2	BBA-139: from P.V. Vail, Phoenix, in 1974
Autographa californica MNPV	AcMNPV	*Noctuidae*	AY706679	AY706526	AY706591	59.4	S43	Steinhaus collection
Bombyx mori NPV	BmNPV	*Bombycidae*	DQ231336	DQ231337	DQ231338	60.5	S9	IVCAS 1.9
Bombyx mori NPV	BmNPV	*Bombycidae*	DQ231339	DQ231340	DQ231341	60.3	S12	IVCAS 1.131
Buzura suppressaria NPV	BusuNPV	*Geometridae*	DQ231342	DQ231343	DQ231344	62.9	S13	IVCAS 1.10
Catopsilia pomona NPV	CapoNPV	*Pieridae*	n.d.	DQ231345	DQ231346	62.3	S16	IVCAS 1.228
Dasychira plagiata NPV	DaplNPV	*Lymantriidae*	AY706690	AY706545	AY706610	58.4	M36-8	Steinhaus collection
Dirphia peruvianus NPV	DipeNPV	*Saturniidae*	AY706691	AY706546	AY706611	70.8	A3-1	Steinhaus collection
Ectropis grisescens NPV	EcgrNPV	*Geometridae*	AY706692	AY706547	AY706612	60.3	S22	IVCAS 1.0310
Euproctis digramma NPV	EudiNPV	*Lymantriidae*	AY706693	AY706549	AY706614	58.6	S24	IVCAS 1.189
Euproctis pseudoconspersa NPV	EupsNPV	*Lymantriidae*	AY706694	AY706551	AY706616	60.8	A13-1, A4-5	BBA-181: Tea Res. Inst. Cin. Acad. Agri. Sci. Hangzou via J.R. Adams 1979
Galleria mellonella MNPV	GmMNPV	*Pyralidae*	AY706696	AY706553	AY706618	60.1	A11-3	BBA-26: from J. Veber in 1960, CSSR
Galleria mellonella MNPV	GmMNPV	*Pyralidae*	AY706698	AY706555	AY706620	60.3	A3-6	BBA-192: from G. Croizier in 1981, St. Christol-les-Ales (France)
Galleria mellonella MNPV	GmMNPV	*Pyralidae*	AY706697	AY706554	AY706619	59.3	A16-3	BBA164: 1978 from Germany
Hemerocampa vetusta NPV	HeveNPV	*Lymantriidae*	AY706699	AY706558	AY706623	57.6	A24-5	Steinhaus collection
Hyphantria cunea NPV	HycuNPV	*Arctiidae*	AY706700	AY706560	AY706625	52.8	S27	IVCAS 1.53
Lymantria dispar MNPV	LdMNPV	*Lymantriidae*	n.d.	DQ235246	DQ235249	–	A24-6	Steinhaus collection
Lymantria monacha NPV	LymoNPV	*Lymantriidae*	AY706702	AY706563	AY706628	46.6	A19-3	BBA-183: from Forstl. Bundes-Vers. Anst. Vienna (Austria)
Lymantria monacha NPV	LymoNPV	*Lymantriidae*	AY706701	AY706562	AY706627	46.6	A14-3	BBA-51: 1970 South Sweden
Lymantria xylina NPV	LyxyNPV	*Lymantriidae*	AY706703	AY706564	AY706629	47.0	S31	IVCAS 1.29
Malacosoma americanum NPV	MaamNPV	*Lasiocampidae*	AY706704	AY706565	AY706630	56.8	M39-4	BBA-176: from J.R. Adams, Beltsville (USA) in 1979
Malacosoma neustria NPV	ManeNPV	*Lasiocamp.*	AY706708	AY706569	AY706634	58.5	S32	IVCAS 1.30
Mamestra brassicae MNPV	MbMNPV	*Noctuidae*	AY706706	AY706567	AY706632	56.4	A3-5	n.d.a.
Mamestra brassicae MNPV	MbMNPV	*Noctuidae*	AY706705	AY706566	AY706631	57.8	A10-1	BBA-218: from Aschersleben (Germany) in 1991
Mamestra brassicae MNPV	MbMNPV	*Noctuidae*	AY706707	AY706568	AY706633	57.5	S33	IVCAS 1.50
Nepytia phantasmaria NPV	NephNPV	*Geometridae*	AY706709	AY706571	AY706636	51.7	A25-5	Steinhaus collection
Peridroma margaritosa NPV	PemaNPV	*Noctuidae*	AY706710	AY706572	AY706637	45.3	A25-4	Steinhaus collection

Plusia acuta NPV	PlacNPV	*Noctuidae*	AY706712	AY706577	AY706642	59.7	A14-5	BBA-115: Aventis (E548/68)
Plutella maculipennis NPV	PlmaNPV	*Plutellidae*	AY706713	AY706578	AY706643	59.5	A15-2	BBA-142: 1974 n.d.a.
Samia cynthia NPV	SacyNPV	*Saturniidae*	AY706711	AY706574	AY706639	48.2	S36	IVCAS 1.0034
Spilosoma phasma NPV	SpphNPV	*Arctiidae*	AY706684	AY706536	AY706601	54.7	S3	IVCAS 1.120
Spodoptera litura NPV	SpltNPV	*Noctuidae*	AY706714	AY706580	AY706645	56.1	A17-3	BBA-180: Inst. For Zool., Chin. Acad. Sci. Beijing via J.R. Adams in 1979
Spodoptera litura NPV	SpltNPV	*Noctuidae*	AY706715	AY706581	AY706646	55.6	S37	IVCAS 1.234
Spodoptera terricola NPV	SpteNPV	*Noctuidae*	AY706716	AY706582	AY706647	56.0	A26-1	Steinhaus collection
Spodoptera littoralis NPV	SpliNPV	*Noctuidae*	AY706717	AY706585	AY706650	53.8	A26-5	Steinhaus collection
Spodoptera littoralis NPV	SpliNPV	*Noctuidae*	AY706718	AY706586	AY706651	53.9	A9-1	from Dr. B. Inceoglu
Thysanoplusia orichalcea NPV	ThorNPV	*Noctuidae*	AY706719	AY706587	AY706652	61.1	A28-1	Cheng et al., 2005
Tineola bisselliella NPV	TibiNPV	*Tineidae*	AY706720	AY706588	AY706653	64.2	M50-4	n.d.a
Adoxophyes orana GV	AdorGV	*Tortricidae*	AY706658	AY706530	AY706595	64.9	S45	IVCAS 1.156
Adoxophyes orana GV	AdorGV	*Tortricidae*	AY706657	AY706529	AY706594	65.0	A6-5	BBA-190: Swiss (Wallis) Isolate via Flückinger in 1980
Agrotis exclamationis GV	AgexGV	*Noctuidae*	AY706659	AY706531	AY706596	65.5	S46	IVCAS 1.0001
Agrotis segetum GV	AgseGV	*Noctuidae*	AY706661	AY706534	AY706599	65.9	S47	IVCAS 1.0002
Agrotis segetum GV	AgseGV	*Noctuidae*	AY706660	AY706533	AY706598	65.4	A17-5	BBA-196: from B. Bolet, Zool. Inst. University Kopenhagen (Denmark)
Andraca bipunctata GV	AnbiGV	*Bombycidae*	AY706662	AY706538	AY706603	65.2	S48	IVCAS 1.0060
Choristoneura murinana GV	ChmuGV	*Tortricidae*	AY706663	AY706543	AY706608	68.2	A11-1 M50-3	BBA-2: n.d.a.
Clostera anachoreta GV	ClanGV	*Notodontidae*	AY706664	AY706544	AY706609	58.0	S49	IVCAS 1.56
Cydia pomonella GV	CpGV	*Tortricidae*	AY706667	AY706556	AY706621	55.3	A6-4	BBA-201: from infected *Grapholita funebrana* (1980)
Cydia pomonella GV	CpGV	*Tortricidae*	AY706668	AY706557	AY706621	55.6	M39-1	BBA-143: from infected *Grapholita molesta* (1974)
Cydia pomonella GV	CpGV	*Tortricidae*	AY706670	AY706561	AY706626	55.5	A11-2	BBA-198: Russian Isolate from Falcon Berkeley (USA) in 1981
Erinnyis ello GV	ErelGV	*Sphingidae*	AY706665	AY706548	AY706313	64.4	M34-4	Steinhaus collection
Estigmene acrea GV	EsacGV	*Arctiidae*	DQ235250	n.d.	DQ235251	–	M30-3	Steinhaus collection
Euxoa ochrogaster GV	EuocGV	*Noctuidae*	AY706666	AY706550	AY706615	60.6	A24-1	Steinhaus collection
Hyphantria cunea GV	HycuGV	*Arctiidae*	AY706669	AY706559	AY706624	66.0	A5-1/ A18-3	BBA-226: Russian isolate
Peridroma morpontora GV	PemoGV	*Noctuidae*	AY706672	AY706573	AY706638	59.3	A25-3	Steinhaus collection
Pieris brassicae GV	PbGV	*Pieridae*	DQ235253	DQ235252	n.d.	–	S54	IVCAS 1.35.
Pieris rapae GV	PiraGV	*Pieride*	AY706673	AY706575	AY706640	68.7	S55	IVCAS 1.37
Plathypena scabra GV	PlscGV	*Noctuidae*	AY706675	AY706579	AY706644	64.4	A25-6	Steinhaus collection
Scotogramma trifolii GV	SctrGV	*Noctuidae*	AY706676	AY706583	AY706648	60.6	A26-3	Steinhaus collection
Spodoptera androgea GV	SpanGV	*Noctuidae*	n.d.	DQ235248	DQ235247	–	A25-7	Steinhaus collection
Spodoptera frugiperda GV	SpfrGV	*Noctuidae*	AY706677	AY706584	AY706649	54.1	A12-4	BBA-208: from G. Biane, La Minére (France) in 1981

Given are the virus names, virus sigla, insect host family and accession numbers of partial *polh/gran, lef-8* and *lef-9* sequences, isolate identifiers and origin of the virus isolate.

[a] Predicted AT content (%) of the genome.

[b] Isolate identifier.

[c] Origin of virus sample: BBA = Federal Biological Research Center for Agriculture and Forestry, Institute for Biological Control, Darmstadt, Germany. Steinhaus collection = historic collection of Edward Steinhaus deposited at the University of Berkeley, US. IVCAS = China Center for General Viruses Culture Collection/Wuhan Institute of Virology, Chinese Academy of Virology, Wuhan, China. n.d.a. = no description available, n.d. = not determined.

became possible, even when only small amounts of material or badly preserved samples were available.

Phylogenetic analysis

In the case of fully sequenced baculoviruses, it was shown that the topology of the phylogenetic tree derived from concatenated partial amino acid sequences of *polh/gran*, *lef-8* and *lef-9* is fully compatible with the tree produced by the concatenated 30 baculovirus core genes (Lange et al., 2004). Consequently, these partial sequences allowed the identification and classification of the newly analyzed isolates and the comparison with viruses whose genomes have been fully sequenced. A phylogenetic tree of the deduced amino acid sequences of the concatenated *lef-8*, *lef-9* and *polh/gran* fragments comprising 71 newly analyzed virus samples and 46 previously published viruses was established using maximum parsimony (MP) and UPGMA distance methods. The division of the individual virus samples into the major groups NPV and GV was consistent with descriptions of the historic virus samples used in this study. This tree included 38 group I NPVs (Fig. 1A), 39 group II NPVs (Fig. 1B) and 37 GVs (Fig. 1C). The dipteran-specific (CuniNPV) and the hymenopteran-specific NeleNPV and NeseNPV were included to root the lepidopteran NPV and GV tree.

Within the group I NPVs, two monophyletic clades were identified (Fig. 1A). The first clade (Ia) included AcMNPV, *Rachiplusia ou* (Ro) MNPV, *Bombyx mori* (Bm) NPV and *Thysanoplusia orichalcea* (Thor) NPV as the best-characterized species. The second clade (Ib) included the completely sequenced *Epiphyas postvittana* (Eppo) NPV, *Choristoneura fumiferana* (Cf) MNPV, *C. fumiferana* DEF (CfDEF) MNPV, and *Orgyia pseudotsugata* (Op) MNPV as the most prominent species. In clade Ia, most host species belonged to the Noctuidae. *B. mori* (Bombycidae) and *Galleria mellonella* (Pyralidae) were represented with several virus samples, and Plutellidae and Nymphalidae were only represented with one species each. Interestingly, group I NPVs specific for Noctuidae, Bombycidae, Pyralidae and Plutellidae were exclusively found in clade Ia while NPVs from Nymphalidae fell in both clades Ia and Ib. Clade Ib comprises of OpMNPV, EppoNPV, CfMNPV and CfDEFNPV and a species complex consisting of *Actias selene* (Acse) NPV, *Antheraea pernyi* (Anpe) NPV and *Phryganidia californica* (Phca) NPV. The insect host family distribution for clade Ib viruses included the Arctiidae, Geometridae, Lymantriidae, Notodontidae, Nymphalidae, Saturniidae and Tortricidae. This was a considerably broader host distribution than for clade Ia viruses. Interestingly, an NPV that is highly similar to CfDEFNPV was isolated from *Amorbia cuneacapsa*. It remains to be analyzed whether this virus can replicate productively in *A. cuneacapsa*.

Six NPVs (*Coloradia pandora* (Copa) NPV, *Dirphia peruvianus* (Dipe) NPV, *Tineola bisselliella* (Tibi) NPV, *Aporia crataegi* (Apcr) NPV, *Pterolocera amplicornis* (Ptam) NPV and *Catopsilia pomona* (Capo) NPV) were placed among the group I NPVs, but they did not belong to either clade Ia or Ib. Their positions were not well resolved by MP and UPGMA analyses, and the bootstrap support was low. The two NPVs infecting Saturniidae family members in this set (CopaNPV and DipeNPV) differed from the other NPVs we analyzed from the same family of insects in that the latter belonged to clade Ib. NPVs derived from the other three host families (Anthelidae, Pieridae, Tineidae) are not represented within clade Ia or Ib.

Two of them, CopaNPV and DipeNPV, were isolated from Saturniidae, whereas all other viruses of Saturniidae belonged to clade Ib. It is striking that the other three host families (Anthelidae, Pieridae, Tineidae) of these viruses are not represented within clade Ia or Ib. In total, 13 lepidopteran families could be identified for the 38 analyzed group I NPVs. It is likely that further clades of closely related viruses will be distinguished when more viruses from this subgroup are analyzed.

The tree of group II NPVs contains 39 viruses from only five insect families, i.e. Noctuidae, Lymantridae, Lasiocampidae, Geometridae and Tortricidae (Fig. 1B). The bootstrap values of the basis nodes of this tree are low, and the basis branching was not well resolved. Even so, the MP and UPGMA bootstrap values of the root node of group II NPVs were 80% and 65%, respectively. Thus, a monophyletic origin of these viruses is suggested. Clusters of closely related viruses were observed for viruses from *Spodoptera*, *Lymantria*, *Mamestra*, *Malacosoma* and *Helicoverpa*, and many viruses from different hosts (but within the same genus) appeared to be the same species (see below). The associations between different virus isolates and host families were not as close as in group I NPVs. For example, viruses from Noctuidae and Lymantridae were found in several distinct branches of the group II NPV tree, suggesting that group II NPV phylogeny is linked more to the host genus than to the host families. The long basal branches and the lack of major clades indicate that the group II NPVs are much more diverged than group I NPVs. However, all analyzed group II NPVs infect members of the macrolepidoptera with the only exception being the unusual *Adoxophyes honmai* (Adho) NPV, which is pathogenic to a tortricid. Despite the apparent divergence of group II NPVs, their host range is restricted to a few lepidopteran families.

NPVs from the three Spodoptera species, *S. litura* (SpltNPV), *S. littoralis* (SpliNPV) and *S. terricola* (SpteNPV), showed some interesting aspects. The two SpliNPV (A26-5, A9-1) samples clustered together, and their closest relative is SpltNPV (NC003102). A neighbor group consisted of two other SpltNPV (A17-3, S37) and SpteNPV (A26-1). Provided that the hosts of these virus samples had been correctly determined, this suggests that there are two different NPVs from *S. litura*. A similar situation is noticed for the cluster of *Mamestra brassicae* (Mb) and *M. configurata* (Maco) NPVs that are also distributed in two terminal branches. The MbMNPV (A10-1 and S33) shares nearly sequence identity with MacoNPV-B, whereas MbMNPV (A3-5) is more closely related to MacoNPV-A. From the biological and genomic differences between MacoNPV-A and MacoNPV-B, it was concluded that both represent separately evolving viruses (Li et al., 2002a, 2002b, 2005). The branching pattern of NPVs that

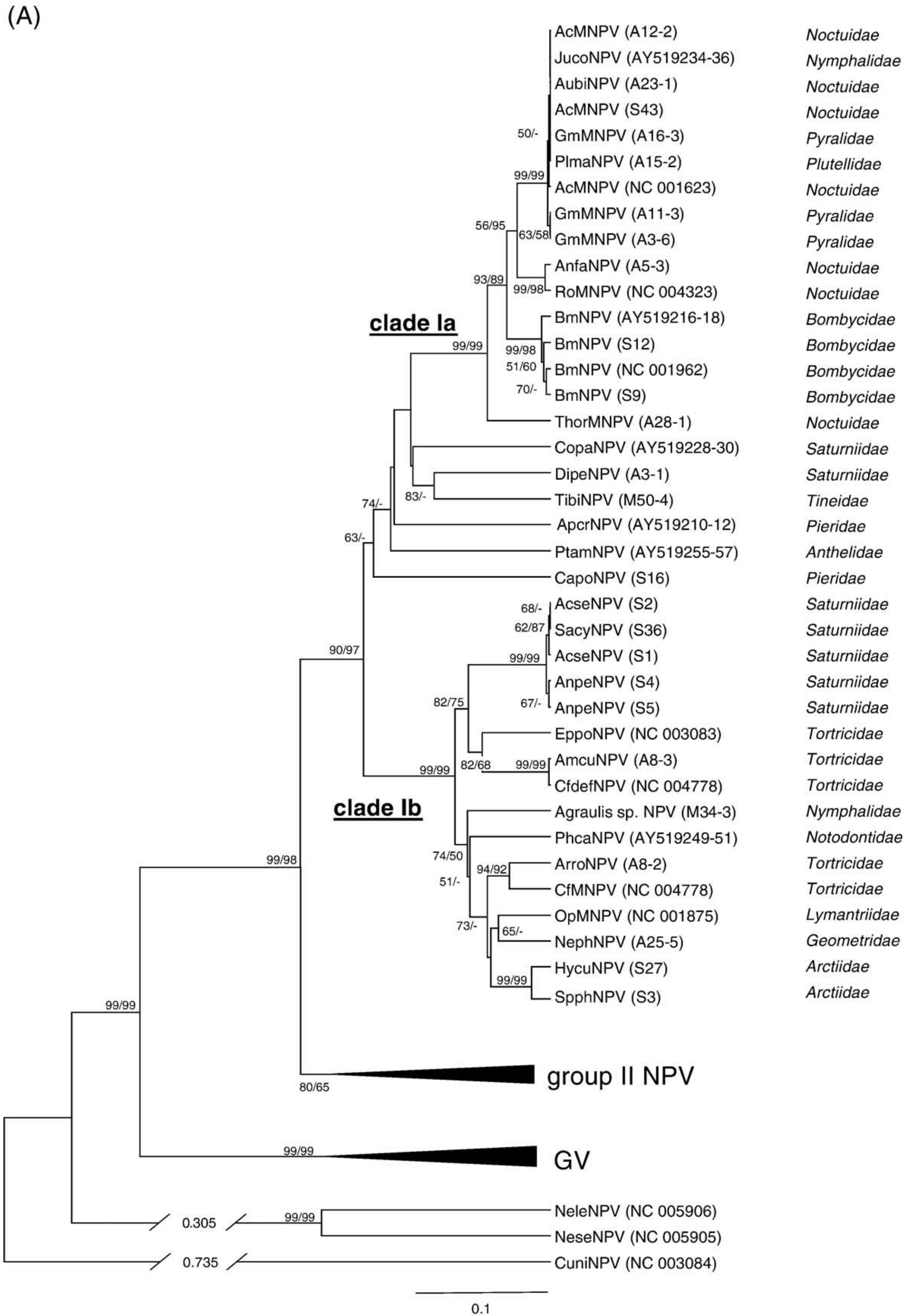

Fig. 1. UPGMA distance tree based on concatenated amino acid sequences of the partial *polh/gran*, *lef-8* and *lef-9* genes of 117 baculoviruses including viruses of Table 2 (indicated by isolate number) and those published previously (indicated by accession numbers) (Lange et al., 2004). Numbers at the nodes indicate the bootstrap values (1000 replicates) from a maximum parsimony (MP) and UPGMA, respectively (MP/UPGMA). CuniNPV was used as an outgroup. The tree is split into three subtrees depicting (A) group I NPVs, (B) group II NPVs and (C) Granuloviruses (GVs). The families of the host species are given to the right.

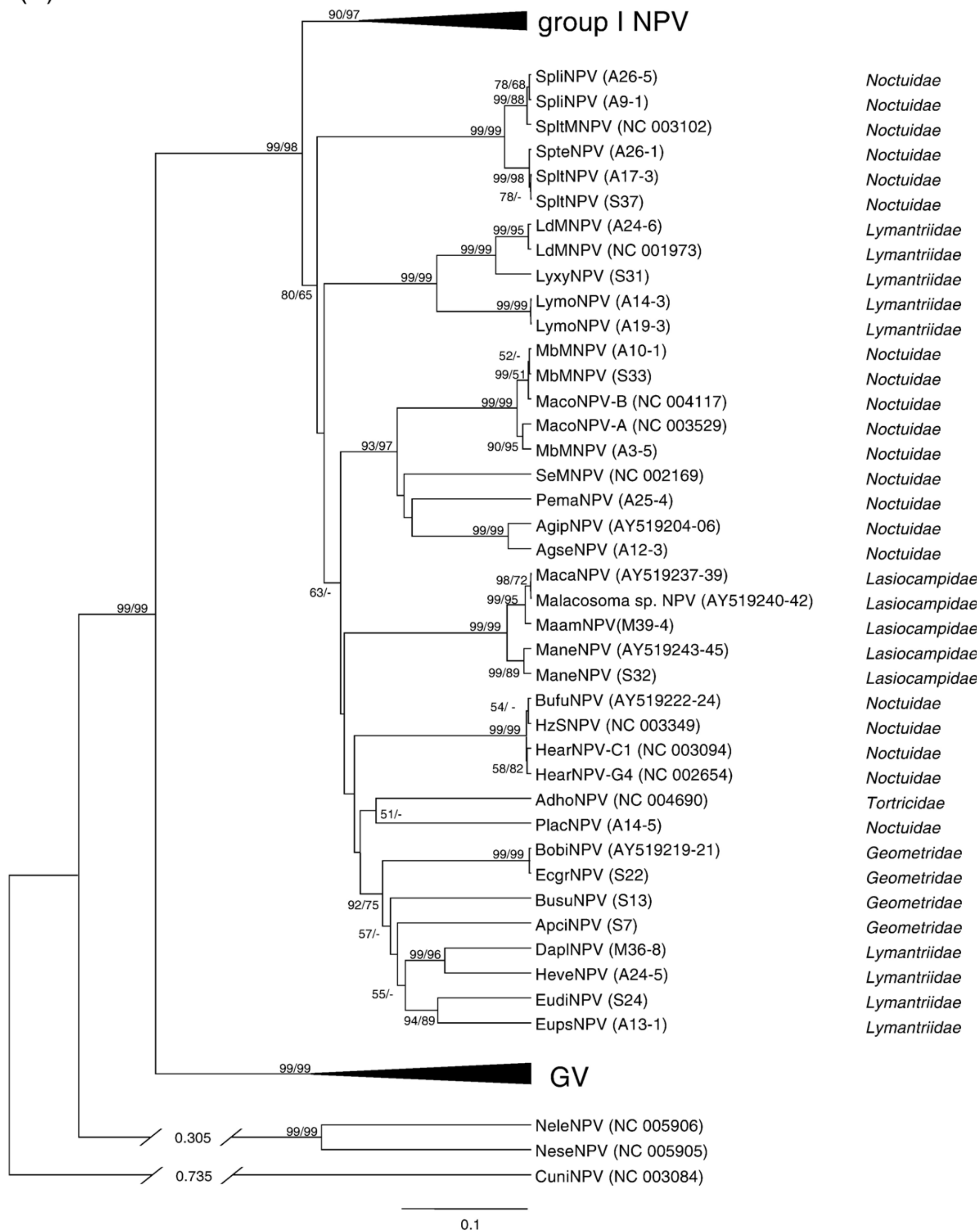

Fig. 1 (*continued*).

infected *L. dispar* (LdMNPV), *L. xylina* (LyxyNPV) and *L. monacha* (LymoNPV) corresponds closely to the host species separation within the host genus. This relationship apparently reflects a tight co-evolution between the viruses and their hosts.

Gene fragments of 37 GVs from nine different host families were analyzed (Fig. 1C). Among the GVs, only one clearly distinct clade comprising a group of Noctuidae-specific viruses could be defined. This clade branched off from other GVs at the basis of the GV tree and contained GVs with a slow pathology such as *Xestia c-nigrum* (Xecn) GV and *Trichoplusia ni* (Tn) GV. In contrast to fast killing GVs that show similar pathology as NPVs, these slow acting GVs cause protracted infection only in the fat body for 2–3 weeks, and the larvae die in the last instar (Federici, 1997). Another slow acting GV, *Adoxophyes orana* (Ador) GV, clusters with Tortricid infecting GVs, signifying that this type of pathology is not a

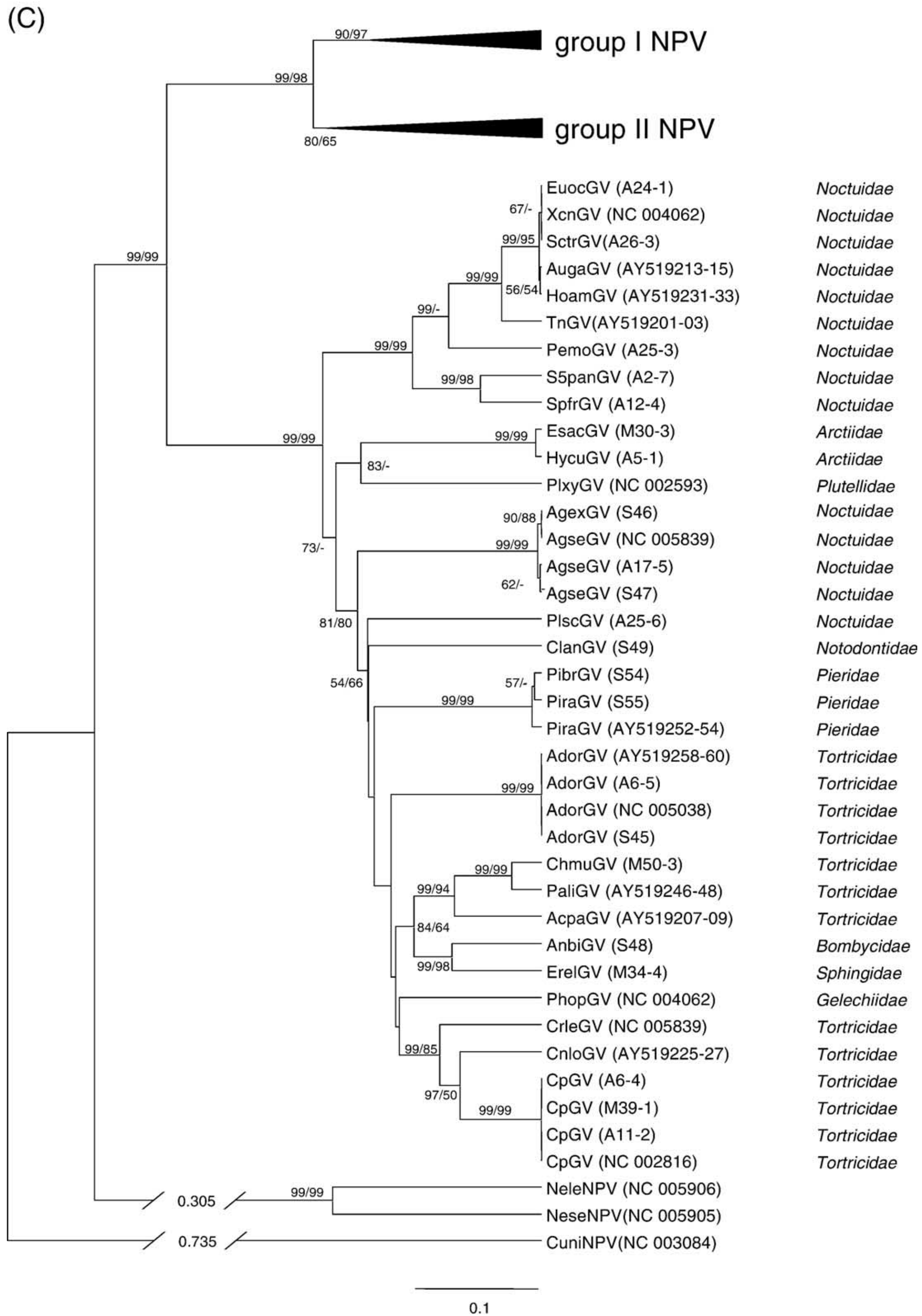

Fig. 1 (*continued*).

monophyletic trait among GVs and thus not a suitable marker for natural classification (Wormleaton and Winstanley, 2001). Apparently, this type of pathology had developed more than once, likely due to an adaptation to the biology of the host.

As shown in Fig. 1C, it is most likely that all Tortricidae infecting GVs have a monophyletic origin irrespective of their speed of replication and pathogenesis. Although the bootstrap support for this conclusion is below 50% in the present study, it had been strictly ascertained by the 30-core gene trees including the fully sequenced tortricid-specific GVs (Lange et al., 2004). GVs from Gelechiidae (PhopGV), Bombicidae (AnbiGV) and Sphingidae (ErleGV) clustered within the

tortricid infecting clade, which suggests that these host families were invaded by GVs that were originally established in tortricid hosts.

The majority of the GVs analyzed here belong to only two host families, Noctuidae and Tortricidae. From the topology of the GV tree, it follows that tortricid infecting GVs most likely originated from taxa infecting noctuid species. This is in contrast to group I and group II NPVs where Torticidae infecting viruses can be found on basal branches that did not derive from Noctuidae-specific viruses (compare Figs. 1A–C).

Evolutionary differences among group I NPVs, group II NPVs and GVs

By comparing the base composition of the completely sequenced baculovirus genomes with those of the PCR amplified fragments of *lef-8*, *lef-9* and *polh/gran*, a regression formula was established that allowed the prediction of the AT content of the baculovirus genomes (Lange et al., 2004). This formula was based on 21 baculovirus genomes. We attempted to validate this formula by analyzing recently published baculovirus genomes. The formula predicts AT contents of 54.4% for CfDEFNPV (actual 54.2%), 59.3% for *Chrysodeixis chalcites* NPV (61.0%), 60. 5% for *T. ni* SNPV (61.0%), 55.3% for MacoNPV-A (v90/4) (58.3%) and 51.9% for CfMNPV (49.1%) (Lauzon et al., 2005; van Oers et al., 2005; Willis et al., 2005; Li et al., 2005; de Jong et al., 2005). The predictions diverged from the real AT content of the respective genomes generally by less than 2%. The exceptions were MacoNPV-A (v90/4) and CfMNPV, for which the predictions were 2.95% below and 2.8% above the real AT percentage, respectively. Thus, it can be concluded that the prediction of AT contents on the basis of the short sequence fragments of *lef-8*, *lef-9* and *polh/gran* is highly reliable and can be used as a powerful estimator for the nucleotide composition of unknown genomes. The estimated AT contents of the newly characterized viruses are given in Table 2. These calculations suggest that group II NPVs have AT contents with an average of 56.4% followed by group I NPVs (average 57.5%) and GVs having a considerably higher AT content with 63.5% (Fig. 2). The various AT contents between lepidopteran-specific NPVs and GVs are striking, the reason for which has not yet been elucidated.

A further remarkable observation was that the forces driving the evolution of the selected marker genes *lef-8*, *lef-9* and *polh/gran* seems to follow different patterns for group I NPVs as well as for group II NPVs and GVs (Tables 3–6). The *lef-9* sequence distances of those viruses belonging to the *Helicoverpa* sp. NPV group (Table 4) and for those belonging to the XecnGV group (Table 5) were always close to zero. This indicated that the *lef-9* fragments of these viruses have a very similar nucleotide sequence and are extremely conserved. The small distances between these viruses were mainly constituted by mutations of a few nucleotides in the *lef-8* and *polh/gran* genes. In contrast, the closely related group I NPVs within the AcMNPV/RoMNPV/BmNPV cluster showed considerable distances in their *lef-9* genes (Table 3), suggesting that *lef-9*

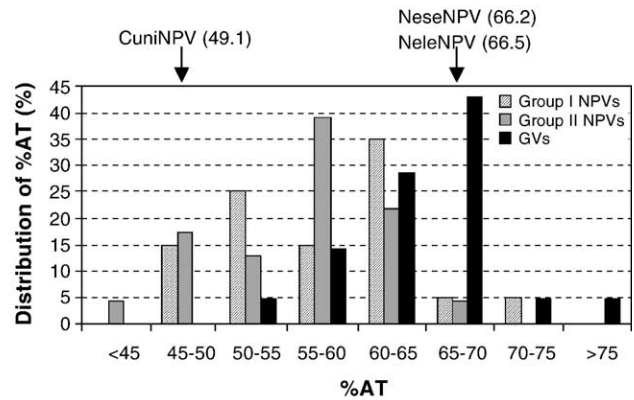

Fig. 2. Distribution of predicted AT contents of group I NPVs, group II NPVs and GVs of the lepidopteran-specific baculovirus isolates that were used to infer the phylogenetic tree in Fig. 1. Arrows indicate the AT contents of non-lepidopteran baculoviruses CuniNPV, NeleNPV and NeseNPV.

genes of these viruses were less conserved than those within group II and GV species. On the other hand, there was a strong sequence identity of the *polh/gran* sequences within the AcMNPV cluster indicating that in these isolates there was a stronger pressure to conserve the *polh/gran* gene than the *lef-9* and *lef-8* genes. These examples make it obvious that even highly conserved genes like *lef-8*, *lef-9* and *polh/gran* were constrained by different evolutionary speeds and forces. These differences apparently correlate with the phylogenetic affiliation to different clades within the lepidopteran-specific baculoviruses and might be the notable consequence of different molecular constraints that are exploited by the viruses during the infection process. Because different genes evolve at different speeds in different baculovirus clades, the utility of single gene trees is limited. Herniou et al. (2001) already pointed out the restricted significance of single gene trees for inferring Baculovirus phylogeny. Because single genes evolve at different rates within different baculovirus clades, single gene trees of even highly conserved genes, such as *lef-8*, *lef-9* and *polh/gran*, are inherently flawed. The combined usage of a set of highly conserved genes like *lef-8*, *lef-9* and *polh/gran* might equilibrate the different evolutionary signals originating from the single genes and are therefore a reasonable and straightforward strategy for baculovirus species identification and a classification based on natural phylogeny.

Baculovirus species demarcation

The phylogenetic analyses demonstrated that the distances between AcMNPV, *Junonia coenia* (Juco) NPV, *Autographa biloba* (Aubi) NPV, *G. mellonella* (Gm) MNPV and *Plutella maculipennis* (Plma) NPV were close to zero. It is well documented that AcMNPV has a broad host range (Gröner, 1986). In many cases, the same virus was isolated from different insect hosts and was consequently given different names. This was also encountered with *Anagrapha falcifera* (Anfa) NPV and RoMNPV (Harrison and Bonning, 1999, 2003) and HearNPV and HzSNPV (Chen et al., 2002). Lange et al. (2004) have already suggested that XecnGV, *Autographa gamma* (Auga) GV and *Hoplodrina ambigua* (Hoam) GV

Table 3
Pairwise distances of the nucleotide sequences of (A) *lef-9* and *lef-8* fragments and of (B) *polh* and concatenated *polh/lef-8/lef-9* fragments of the AcMNPV/BmNPV/RoMNPV cluster

(A)

lef-9 / *lef-8*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 AcMNPV	**-**	**0.004**	**0.012**	**0.016**	**0.020**	**0.020**	**0.016**	**0.016**	**0.004**	0.004	0.004	0.024	0.016	0.020	0.020	0.191
2 AcMNPV A12-2	**0.000**	**-**	**0.008**	**0.012**	**0.016**	**0.016**	**0.012**	**0.012**	**0.000**	0.000	0.000	0.020	0.012	0.016	0.016	0.187
3 AcMNPV S43	**0.000**	**0.000**	**-**	**0.004**	**0.008**	**0.008**	**0.004**	**0.004**	**0.008**	0.008	0.008	0.020	0.012	0.016	0.016	0.177
4 AubiNPV A23-1	**0.003**	**0.003**	**0.003**	**-**	**0.004**	**0.004**	**0.000**	**0.000**	**0.012**	0.012	0.012	0.016	0.008	0.012	0.012	0.171
5 GmMNPV A3-6	**0.005**	**0.005**	**0.005**	**0.002**	**-**	**0.000**	**0.004**	**0.004**	**0.016**	0.016	0.016	0.020	0.012	0.016	0.016	0.176
6 GmMNPV A11-3	**0.005**	**0.005**	**0.005**	**0.002**	**0.000**	**-**	**0.004**	**0.004**	**0.016**	0.016	0.016	0.020	0.012	0.016	0.016	0.176
7 GmMNPV A16-3	**0.000**	**0.000**	**0.000**	**0.003**	**0.005**	**0.005**	**-**	**0.000**	**0.012**	0.012	0.012	0.016	0.008	0.012	0.012	0.171
8 JucoNPV M30-5	**0.000**	**0.000**	**0.000**	**0.003**	**0.005**	**0.005**	**0.000**	**-**	**0.012**	0.012	0.012	0.016	0.008	0.012	0.012	0.171
9 PlmaNPV A15-2	**0.000**	**0.000**	**0.000**	**0.003**	**0.005**	**0.005**	**0.000**	**0.000**	**-**	0.000	0.000	0.020	0.012	0.016	0.016	0.187
10 AnfaNPV A5-3	0.038	0.038	0.038	0.042	0.043	0.043	0.038	0.038	0.038	**-**	**0.000**	0.020	0.012	0.016	0.016	0.187
11 RoMNPV	0.040	0.040	0.040	0.043	0.045	0.045	0.040	0.040	0.040	**0.005**	**-**	0.020	0.012	0.016	0.016	0.187
12 BmMNPV	0.032	0.032	0.032	0.035	0.037	0.037	0.032	0.032	0.032	0.062	0.063	**-**	**0.008**	**0.012**	**0.004**	0.177
13 BmMNPV M28-4	0.033	0.033	0.033	0.037	0.038	0.038	0.033	0.033	0.033	0.063	0.065	**0.008**	**-**	**0.012**	**0.004**	0.182
14 BmMNPV S09	0.038	0.038	0.038	0.042	0.043	0.043	0.038	0.038	0.038	0.069	0.070	**0.013**	**0.008**	**-**	**0.008**	0.167
15 BmMNPV S12	0.033	0.033	0.033	0.037	0.038	0.038	0.033	0.033	0.033	0.064	0.065	**0.014**	**0.009**	**0.011**	**-**	0.177
16 ThorNPV A28-1	0.217	0.217	0.217	0.222	0.220	0.220	0.217	0.217	0.217	0.233	0.229	0.217	0.219	0.219	0.226	-

(B)

Polh / *polh/lef-8/lef-9*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 AcMNPV	**-**	**0.000**	**0.000**	**0.000**	**0.002**	**0.002**	**0.000**	**0.002**	**0.000**	0.304	0.296	0.273	0.277	0.270	0.270	0.304
2 AcMNPV A12-2	**0.001**	**-**	**0.000**	**0.000**	**0.002**	**0.002**	**0.000**	**0.002**	**0.000**	0.304	0.296	0.273	0.277	0.270	0.270	0.304
3 AcMNPV S43	**0.002**	**0.002**	**-**	**0.000**	**0.002**	**0.002**	**0.000**	**0.002**	**0.000**	0.304	0.296	0.273	0.277	0.270	0.270	0.304
4 AubiNPV A23-1	**0.005**	**0.004**	**0.002**	**-**	**0.002**	**0.002**	**0.000**	**0.002**	**0.000**	0.304	0.296	0.273	0.277	0.270	0.270	0.304
5 GmMNPV A3-6	**0.007**	**0.006**	**0.005**	**0.002**	**-**	**0.000**	**0.002**	**0.005**	**0.002**	0.308	0.300	0.277	0.281	0.273	0.273	0.308
6 GmMNPV A11-3	**0.007**	**0.006**	**0.005**	**0.002**	**0.000**	**-**	**0.002**	**0.005**	**0.002**	0.308	0.300	0.277	0.281	0.273	0.273	0.308
7 GmMNPV A16-3	**0.003**	**0.002**	**0.00**	**0.002**	**0.004**	**0.004**	**-**	**0.002**	**0.000**	0.304	0.296	0.273	0.277	0.270	0.270	0.304
8 JucoNPV M30-5	**0.004**	**0.003**	**0.002**	**0.002**	**0.005**	**0.005**	**0.001**	**-**	**0.002**	0.304	0.296	0.273	0.277	0.269	0.269	0.304
9 PlmaNPV A15-2	**0.001**	**0.000**	**0.002**	**0.004**	**0.006**	**0.006**	**0.002**	**0.003**	**-**	0.304	0.296	0.273	0.277	0.270	0.270	0.304
10 AnfaNPV A5-3	0.103	0.102	0.104	0.107	0.109	0.109	0.105	0.105	0.102	**-**	**0.005**	0.065	0.065	0.067	0.067	0.162
11 RoMNPV	0.102	0.101	0.103	0.106	0.108	0.108	0.104	0.104	0.101	**0.004**	**-**	0.059	0.059	0.062	0.062	0.155
12 BmMNPV	0.097	0.096	0.096	0.097	0.100	0.100	0.095	0.095	0.096	0.054	0.053	**-**	**0.007**	**0.002**	**0.002**	0.142
13 BmMNPV M28-4	0.097	0.096	0.096	0.097	0.100	0.100	0.095	0.095	0.096	0.053	0.052	**0.008**	**-**	**0.010**	**0.010**	0.139
14 BmMNPV S09	0.099	0.098	0.098	0.099	0.101	0.101	0.097	0.097	0.098	0.057	0.057	**0.009**	**0.009**	**-**	**0.000**	0.139
15 BmMNPV S12	0.096	0.095	0.095	0.096	0.099	0.099	0.094	0.094	0.095	0.055	0.054	**0.008**	**0.008**	**0.007**	**-**	0.139
16 ThorNPV A28-1	0.238	0.237	0.235	0.236	0.237	0.237	0.234	0.234	0.237	0.201	0.197	0.185	0.186	0.183	0.188	-

ThorNPV was included as the nearest neighbor. The distances were calculated using MEGA (Kimura 2-parameter model) (Kumar et al., 2004). Bold numbers within the same box belong to isolates that are considered as the same species.

represent the same virus species with only minor differences in the nucleotide sequences of *lef-8*, *lef-9* and *polh/gran*. Two other virus isolates, *Euxoa ochrogaster* (Euoc) GV and *Scotogramma trifolii* (Sctr) GV, were found belonging to the same group.

According to the definition set by the International Committee on Taxonomy of viruses (ICTV), a virus species is considered as "a polythetic class of viruses that constitute a replicating lineage and occupy a particular ecological niche" (van Regenmoertel, 2000). Thus, a single criterion, such as sequence divergence alone, cannot serve for species demarcation. Virus speciation has to be considered as a continuous process in time, thus a cut-off point for species differentiation does not exist. On the other hand, it is very clear that the genetic distance between two viruses correlates with the differences in other biological and physico-chemical properties, which, together, may be applied to the definition of a species. Genetic distances are thus useful diagnostic properties that should aid in the allocation of a given virus isolate to a virus species.

The Kimura 2-parameter (K-2-P) distances between the aligned *polh/gran*, *lef-8* and *lef-9* nucleotide sequences are shown for single gene sequences and for the concatenated sequences of different species groups, including the AcMNPV/RoMNPV/BmNPV groups (Table 3), the HearNPV group (Table 4), the MacoNPV cluster (Table 5) and the XecnGV group (Table 6). Within the distance matrix of the *AcMNPV/RoMNPV/BmNPV* groups, three groups of viruses corresponding to the three mentioned viruses can be clearly distinguished. The distances within each species groups ranged from 0.000 to 0.012 for *polh/gran* and *lef-8*. Only the single *lef-9* sequence distances exceeded 0.012 in some cases. For the concatenated *polh/gran*, *lef-8* and *lef-9* sequences, the intra-group distances were generally lower than 0.012. In contrast,

Table 4
Pairwise distances of the nucleotide sequences of (A) *lef-9* and *lef-8* fragments and of (B) *polh* and concatenated *polh/lef-8/lef-9* fragments of the HearNPV cluster

(A)

lef-9 / *lef-8*	1	2	3	4
1 BufuNPV A2-4	**-**	**0.004**	**0.004**	**0.000**
2 HearNPV-C1	**0.010**	**-**	**0.000**	**0.004**
3 HearNPV-G4	**0.010**	**0.000**	**-**	**0.004**
4 HzSNPV	**0.001**	**0.011**	**0.011**	**-**

(B)

polh / *polh-lef-8-lef-9*	1	2	3	4
1 BufuNPV A2-4	**-**	**0.018**	**0.012**	**0.008**
2 HearNPV-C1	**0.011**	**-**	**0.006**	**0.014**
3 HearNPV-G4	**0.009**	**0.002**	**-**	**0.008**
4 HzSNPV	**0.003**	**0.011**	**0.009**	**-**

The distances were calculated using MEGA (Kimura 2-parameter model) (Kumar et al., 2004). Bold numbers within the same box belong to isolates that are considered as the same species.

the distances between the species groups were much bigger and exceeded 0.050, although all members of the AcMNPV/RoMNPV/BmNPV complex can be considered as narrowly related. The distance between the *polh* sequence from AcMNPV on one hand and the BmNPV and RoMNPV on the other hand is far larger than the corresponding distances for *lef-8* and *lef-9*. This can be explained with the mosaic structure of the AcMNPV polyhedrin gene (Jehle, 2004). A similar situation was observed when examining the HearNPV group (Table 4) and the MacoNPV cluster (Table 5) within group II. HearNPV-G4, HearNPV-C1 and HzSNPV have been completely sequenced and are considered as variants of the same virus species (Chen et al., 2001, 2002; Zhang et al., 2005). The

Table 5
Pairwise distances of the nucleotide sequences of (A) *lef-9* and *lef-8* fragments and of (B) *polh* and concatenated *polh/lef-8/lef-9* fragments of the MacoNPV cluster

(A)

lef-9 / *lef-8*	1	2	3	4	5
1 MbMNPV A10-1	**-**	**0.000**	**0.000**	0.060	0.073
2 MbMNPB S33	**0.008**	**-**	**0.000**	0.060	0.073
3 MacoNPV-B	**0.005**	**0.003**	**-**	0.060	0.073
4 MacoNPV-A	0.099	0.101	0.105	**-**	**0.016**
5 MbMNPV A3-5	0.101	0.103	0.107	**0.020**	**-**

(B)

polh / *polh-lef-8-lef-9*	1	2	3	4	5
1 MbMNPV A10-1	**-**	**0.014**	**0.002**	0.086	0.081
2 MbMNPV S33	**0.008**	**-**	**0.012**	0.075	0.070
3 MacoNPV-B	**0.003**	**0.006**	**-**	0.088	0.084
4 MacoNPV-A	0.086	0.083	0.089	**-**	**0.018**
5 MbMNPV A3-5	0.089	0.086	0.092	**0.018**	**-**

The distances were calculated using MEGA (Kimura 2-parameter model) (Kumar et al., 2004). Bold numbers within the same box belong to isolates that are considered as the same species.

Table 6
Pairwise distances of the nucleotide sequences of (A) *lef-9* and *lef-8* fragments and of (B) *polh* and concatenated *polh/lef-8/lef-9* fragments of the XecnGV cluster

(A)

lef-9 / *lef-8*	1	2	3	4	5	6
1 AugaGV M39-3	**-**	**0.008**	**0.004**	**0.008**	**0.008**	0.190
2 EuocGV A24-1	**0.015**	**-**	**0.004**	**0.000**	**0.000**	0.190
3 HoamGV M39-2	**0.003**	**0.012**	**-**	**0.004**	**0.004**	0.195
4 SctrGV A26-3	**0.015**	**0.000**	**0.012**	**-**	**0.000**	0.190
5 XecnGV	**0.015**	**0.000**	**0.012**	**0.000**	**-**	0.190
6 TnGV	0.194	0.188	0.190	0.188	0.188	-

(B)

polh / *polh/lef-8/lef-9*	1	2	3	4	5	6
1 AugaGV M39-3	**-**	**0.002**	**0.000**	**0.004**	**0.002**	0.118
2 EuocGV A24-1	**0.009**	**-**	**0.002**	**0.002**	**0.000**	0.116
3 HoamGV M39-2	**0.002**	**0.007**	**-**	**0.004**	**0.002**	0.118
4 SctrGVA26-3	**0.010**	**0.001**	**0.008**	**-**	**0.002**	0.113
5 XecnGV	**0.009**	**0.000**	**0.007**	**0.001**	**-**	0.116
6 TnGV	0.166	0.162	0.165	0.161	0.162	-

TnGV was included as the nearest neighbor. The distances were calculated using MEGA (Kimura 2-parameter model) (Kumar et al., 2004). Bold numbers within the same box belong to isolates that are considered as the same species.

distances of the concatenated marker fragments also did not exceed 0.011, whereas single *polh* fragments showed maximum distances of up to 0.018. *Busseola fusca* (Bufu) NPV (A2-4, AY519222–AY519224) can be considered to be nearly identical to HzSNPV. MacoNPV-A and MacoNPV-B are viruses that were isolated from the same host, the bertha armyworm. Although their genomes are closely related, the two viruses are considered as two distinct species (Li et al., 2002a, 2005). The distance of the concatenated *polh*, *lef-8* and *lef-9* between MacoNPV-A and -B was 0.088 and exceeded the distance of RoMNPV and BmNPV (~0.055), which are also thought to be different species. Interestingly, two MbMNPV (A10-1, S33) clustered together with MacoNPV-B, whereas one MbMNPV (A3-5) is close to MacoNPV-A, suggesting that MbMNPV isolates are identical to the MacoNPV viruses. As has been noted by Lange et al. (2004), there are a number of noctuid GVs closely related to XecnGV and having *polh*, *lef-8* and *lef-9* genes that together are not more distant than 0.010. XecnGV is the GV with the broadest host range known to date.

Based on the observed sequence distances in Tables 3–6, we propose to define that two (or more isolates) belong to the same "baculovirus species" if the K-2-P distance between single and/or concatenated *polh*, *lef-8* and *lef-9* nucleotide sequences is smaller than 0.015. Furthermore, two viruses should be considered as different virus species if the distance between single and/or concatenated sequences is larger than 0.050. For distances between 0.015 and 0.050, complementary information is needed to determine whether two viruses are the same or different species. The K-2-P nucleotide substitution model was chosen since it is available in most phylogenetic software packages, e.g. Phylip, MEGA and PAUP. We also suggest that in cases where the partial sequences of these marker gene fragments are close to identity, comparison of only

two marker genes is sufficient since the nearest neighbor validates the phylogenetic position of this virus. For those viruses which show distances of more than 0.015 in the marker genes, sequence determination and phylogenetic analyses of all three marker genes are recommended.

Applying this suggestion, the following viruses isolated from different hosts would belong to the same baculovirus species [each cluster in brackets represent a single species]: [AcMNPV; AubiNPV A23-1; GmMNPV A3-6, A11-3, A16-3; JucoNPV M30-5; PlmaNPV A15-2], [AnfaNPV A5-3; RoMNPV], [AcseNPV S1, S2; SacyNPV S36; AnpeNPV S4, S5], [AmcuNPV A8-3, CfDEFNPV], [SpltNPV NC 003102; *Spodoptera littoralis* NPV A26-5, A9-1], [SpltNPV A17-3, S37; *Spodoptera terricola* NPV A26-1], [MacoNPV-A NC 003529; MbMNPV A3-5], [MacoNPV-B NC004107; MbMNPV A10-1, S33], [HearNPV-C1; HearNPV-G4; HzSNPV; BufuNPV AY519222–AY519224, A2-4], [*Boarmia bistortata* (Bobi) NPV; *Ectropis grisescens* (Ecgr) NPV S22], [XecnGV NC 004062; AugaGV M39-3; EuocGV A24-1; HoamGV M39-2; SctrGV A26-3], [*A. segetum* GV NC005839, A17-5, S47, *A. exclamationis* GV (S46)], [*Estigmene acrea GV M30-3; H. cunea GV A5-1].*

Our suggestions also shed new light on some of the present taxonomic propositions (Theilmann et al., 2005). RoMNPV is obviously much more diverged from AcMNPV than AubiNPV, GmMNPV, JucoNPV and PlmaNPV (compare Table 3). RoMNPV further differs from AcMNPV in host range (summarized in Harrison and Bonning, 2003), DNA restriction endonuclease profile (Harrison and Bonning, 1999) and in lacking the polyhedrin mosaic structure (Jehle, 2004). Based on the phylogenetic distances and the noticed biological differences, we strongly suggest to consider RoMNPV as an own species and not only as a variant of AcMNPV as the other viruses. On the other hand, MacoNPVs and MbMNPVs, which are presently acknowledged as different baculovirus species, seem to be the same viruses that were isolated from different hosts. These are just two examples where sequence data help to determine whether viruses belong to the same species or not. Baculoviruses form a group of immensely radiated viruses which co-evolved with their hosts. Beyond host range and genome sequence, there are often no further easily accessible and differentiating characters which would be suitable for species demarcation. For that reason, a phylogenetic species concept as proposed in this study can be a powerful means for baculovirus classification and taxonomy.

Baculovirus *lef-8*, *lef-9* and *polh/gran* appear to be suitable markers to mirror speciation. As already pointed out above, such a phylogenetic species demarcation cannot replace a qualified taxonomic description of a given virus. Proposing a genetic distance for baculovirus species demarcation can be taken as scaffold to describe the relationship of different baculovirus viruses based on phylogenetic evidence. It can help identify those viruses that are already well described, e.g. by genome sequencing. As shown for the AcMNPV/RoMNPV/BmNPV, HearNPV and MacoNPV clusters, all of which contain at least two completely sequenced genomes, this proposal is consistent with the emerging picture of baculovirus species diversity.

Materials and methods

DNA extraction from diseased insects

Purified baculovirus samples or baculovirus infected insect cadavers were obtained from the Institute for Biological Control, Federal Biological Research Center for Agriculture and Forestry (Darmstadt, Germany), the University of California (Berkeley, USA) and from the Wuhan Institute of Virology, (Wuhan, People's Republic of China). For DNA extractions and sample descriptions, see Lange et al. (2004).

PCR amplification and sequencing of partial polh/gran, lef-8 and lef-9 genes

The amplification of partial *polh*/*gran*, *lef-8* and *lef-9* genes were performed using degenerate primer pairs (prPH-1, prPH-2, prL8-1, prL8-2, prL9-1, prL9-2), which included universal primer tails to enable direct sequencing of the PCR products (Lange et al., 2004). In many reactions, the primers prPH-1 and prL8-1 were replaced with prPH-1B and prL8-1B, which contained differing universal sequencing primer tails BGHrev tails (5′-TAGAAGGCACAGTCGAGG-3′) and T7 (5′-TAATACGACTCACTATAGGG-3′), respectively (Table 1).

Amplification reactions with *lef-8* specific primers were performed with an initial denaturizing step of 95 °C for 4 min and 30 or 35 cycles of 95 °C for 2 min, 38–48 °C for 1 min, 72 °C for 1 min and a final extension step 72 °C for 5 min. *polh/gran* sequences were amplified at 95 °C for 3 min followed by 36–40 cycles of 95 °C for 30 s, 50–53 °C for 1 min, 72 °C for 1 min and a final extension step at 72 °C for 10 min. The *lef-9* protocol remained unchanged from Lange et al. (2004). The PCR products were purified by using the GFX PCR DNA and Gel Band Purification kit (Amersham, Freiburg, Germany), and both DNA strands were sequenced by automatic sequencing (MWG Ebersberg, Germany).

Sequence alignment and phylogenetic analyses

Base calling of DNA trace files was examined manually, and minor mistakes were corrected. In several cases, mispriming of primers prPH-1 and prPH-2 was detected. Blast searches (Altschul et al., 1990) were performed to verify that the partial *polh/gran*, *lef-8* and *lef-9* genes resulted in a consistent grouping within the GVs, group I NPVs and group II NPVs. Partial *polh/gran*, *lef-8* and *lef-9* sequences obtained by PCR and homologous partial sequences from 26 fully sequenced baculovirus genomes available from Genbank were imported into the BioEdit program (Hall, 1999). The deduced amino acid sequences of the three gene fragments were individually aligned using ClustalW (Thompson et al., 1994) and concatenated to a single data set using BioEdit (Hall, 1999).

Maximum parsimony (MP) phylogenetic trees (1000 bootstrap replicates) were inferred from the amino acid sequence

alignments by using MEGA, version 3.0 (Kumar et al., 2004). Introduced gaps were treated as missing data. MP trees were analyzed by a heuristic search with the tree–bisection–reconnection (TBR) branch swapping option. Dayhoff distance corrected UPGMA distance analyses (gamma shape parameter $\alpha = 2.25$; 1000 bootstrap replicates, respectively) were performed using MEGA (Nei and Kumar, 2000).

Distance matrices from aligned nucleotide sequences were determined by using the Pairwise Distance calculation of MEGA version 3.0 applying the Kimura 2-parameter model (Kumar et al., 2004).

Acknowledgments

We are indebted to Loy Volkman and the University of California, Berkeley, for providing the historic baculovirus collection of Edward Steinhaus, which was used in part in this study. We also thank Jürg Huber, Federal Biological Research Center, Darmstadt, as well as Tianxian Li and Shenliang Chen, Wuhan Institute of Virology, for sharing baculovirus samples. The authors are indebted to Alexander Rosisko for excellent technical assistance and to Claudia Vogel (both DLR Rheinpfalz, Neustadt Wstr.) for help during preparation of the manuscript. We thank Basil Arif (Great Lakes Forestry Centre, Sault Ste. Marie, Canada) for critically reading and commenting on the manuscript. This work was supported by a grant of the Deutsche Forschungsgemeinschaft (DFG) to M.L., R.H. and Y.W. (Je245-3, Je245-7) and a grant of the Deutscher Akademischer Austauschdienst (DAAD) to H.W. It was also partially supported by grants (NSFC grant 30025003 and grant 2003CB114202) to Z.H.

References

Afonso, C.L., Tulman, E.R., Lu, Z., Balinsky, C.A., Moser, B.A., Becnel, J.J., Rock, D.L., Kutish, G.F., 2001. Genome sequence of a baculovirus pathogenic for *Culex nigripalpus*. J. Virol. 75, 11157–11165.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Chen, X., Ijkel, W.F.J, Tarchini, R., Sun, X., Sandbrink, H., Wang, H., Peters, S., Zuidema, D., Klein-Lankhorst, R., Vlak, J.M., Hu, Z.-H., 2001. The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome. J. Gen. Virol. 82, 241–257.

Chen, X., Zhang, W.-J., Wong, G., Chun, A., Lu, B.F., McCutchen, J.K., Presnail, R., Herrmann, M., Dolan, S., Tingey, S., Hu, Z.H., Vlak, J.M., 2002. Comparative analysis of the complete genome sequence of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedroviruses. J. Gen. Virol. 83, 673–684.

de Jong, J.G., Lauzon, H.A., Dominy, C., Poloumienko, A., Carstens, E.B., Arif, B.M., Krell, P.J., 2005. Analysis of the *Choristoneura fumiferana* nucleopolyhedrovirus genome. J. Gen. Virol. 86, 929–943.

de Moraes, R.R., Maruniak, J.E., 1997. Detection and identification of multiple baculoviruses using the polymerase chain reaction (PCR) and restriction endonuclease analysis. J. Virol. Methods 63, 209–217.

Federici, B.A., 1997. Baculovirus pathogenesis. In: Miller, L.K. (Ed.), The Baculoviruses. Plenum Press, New York, pp. 33–59.

Garcia-Maruniak, A., Maruniak, J.E., Zanotto, P.M., Doumbouya, A.E., Liu, J.C., Merritt, T.M., Lanoie, J.S., 2004. Sequence analysis of the genome of the *Neodiprion sertifer* nucleopolyhedrovirus. J. Virol. 78, 7036–7051.

Gröner, A., 1986. Specificity and safety of baculoviruses. In: Granados, R.R., Federici, B.A. (Eds.), The Biology of Baculoviruses, vol. 1. CRC Press, Boca Raton, pp. 177–201.

Guarino, L.A., Xu, B., Jin, J.P., Dong, W., 1998. A virus-encoded RNA polymerase purified from baculovirus-infected cells. J. Virol. 72, 7985–7991.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Harrison, R.L., Bonning, B.C., 1999. The nucleopolyhedroviruses of *Rachiplusia ou* and *Anagrapha falcifera* are isolates of the same virus. J. Gen. Virol. 80, 2793–2798.

Harrison, R.L., Bonning, B.C., 2003. Comparative analysis of the genomes of *Rachiplusia ou* and *Autographa californica* multiple nucleopolyhedroviruses. J. Gen. Virol. 84, 1827–1842.

Herniou, E.A., Luque, T., Chen, X., Vlak, J.M., Winstanley, D., Cory, J.S., O'Reilly, D.R., 2001. Use of whole genome sequence data to infer baculovirus phylogeny. J. Virol. 75, 8117–8126.

Herniou, E.A., Olezewski, J.A., Cory, J.S., O'Reilly, D.R., 2003. The genome sequence and evolution of baculoviruses. Annu. Rev. Entomol. 48, 211–234.

Herniou, E.A., Olezewski, J.A., O'Reilly, D.R., Cory, J.S., 2004. Ancient coevolution of baculoviruses and their insect hosts. J. Virol. 78, 3244–3251.

Jehle, J.A., 2004. The mosaic structure of the polyhedrin gene of the *Autographa californica* nucleopolyhedrovirus (AcMNPV). Virus Genes 29, 5–8.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. 5, 150–163.

Lange, M., Wang, H., Zhihong, H., Jehle, J.A., 2004. Towards a molecular identification and classification system of lepidopteran-specific baculoviruses. Virology 325, 36–47.

Lauzon, H.A., Lucarotti, C.J., Krell, P.J., Feng, Q., Retnakaran, A., Arif, B.M., 2004. Sequence and organization of the *Neodiprion lecontei* nucleopolyhedrovirus genome. J. Virol. 78, 7023–7035.

Lauzon, H.A., Jamieson, P.B., Krell, P.J., Arif, B.M., 2005. Gene organization and sequencing of the *Choristoneura fumiferana* defective nucleopolyhedrovirus genome. J. Gen. Virol. 86, 945–961.

Li, Q., Donly, C., Li, L., Willis, L.G., Theilmann, D.A., Erlandson, M.A., 2002a. Sequence and organization of the *Mamestra configurata* nucleopolyhedrovirus genome. Virology 294, 106–121.

Li, L., Donly, C., Li, Q., Willis, L.G., Keddie, B.A., Erlandson, M.A., Theilmann, D.A., 2002b. Identification and genomic analysis of a second species of nucleopolyhedrovirus isolated from *Mamestra configurata*. Virology 297, 226–244.

Li, L., Li, Q., Willis, L.G., Erlandson, M., Theilmann, D.A., Donly, C., 2005. Complete comparative genomic analysis of two field isolates of *Mamestra configurata* nucleopolyhedrovirus-A. J. Gen. Virol. 86, 91–105.

Martignoni, M.E., Iwai, P.J., 1981. A catalogue of viral diseases of insects, mites and ticks. In: Burges, H.D. (Ed.), Microbial Control of Pests and Plant Diseases 1970–1980. Academic Press, London, pp. 897–911.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford Univ. Press, New York.

Pang, Y., Yu, J., Wang, L., Hu, X., Bao, W., Li, G., Chen, C., Han, H., Hu, S., Yang, H., 2001. Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome. Virology 287, 391–404.

Pearson, M.N., Rohrmann, G.F., 2002. Transfer, incorporation, and substitution of envelope fusion proteins among members of the Baculoviridae, Orthomyxoviridae, and Metaviridae (insect retrovirus) families. J. Virol. 304, 81–88.

Pijlman, G.P., Pruijssers, A.J., Vlak, J.M., 2003. Identification of pif-2, a third conserved baculovirus gene required for *per os* infection of insects. J. Gen. Virol. 84, 2041–2049.

Rohrmann, G.F., 1986. Polyhedrin structure. J. Gen. Virol. 67, 1499–1513.

Theilmann, D.A., Blissard, G.W., Bonning, B., Jehle, J.A., O'Reilly, D.R., Rohrmann, G.F., Thiem, S., Vlak, J.M., 2005. Baculoviridae. In: Fauquet,

C.X., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), Virus Taxonomy—Classification and Nomenclature of Viruses. 8th Report of the International Committee on the Taxonomy of Viruses. Elsevier, Amsterdam, pp. 177–185.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

van Oers, M.M., Abma-Henkens, M.H., Herniou, E.A., de Groot, J.C., Peters, S., Vlak, J.M., 2005. Genome sequence of *Chrysodeixis chalcites* nucleopolyhedrovirus, a baculovirus with two DNA photolyase genes. J. Gen. Virol. 86, 2069–2080.

van Regenmoertel, M.H.V., 2000. Introduction to the species concept in virus taxonomy. In: van Regenmoertel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R., Wickner, R.B. (Eds.), Virus Taxonomy—Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, pp. 3–16.

Willis, L.G., Siepp, R., Stewart, T.M., Erlandson, M.A., Theilmann, D.A., 2005. Sequence analysis of the complete genome of *Trichoplusia ni* single nucleopolyhedrovirus and the identification of a baculoviral photolyase gene. Virology 338, 209–226.

Wormleaton, S.L., Winstanley, D., 2001. Phylogenetic analysis of conserved genes within the ecdysteroid UDP-glucosyltransferase gene region of the slow-killing *Adoxophyes orana* granulovirus. J. Gen. Virol. 82, 2295–2305.

Zanotto, P.M., Kessing, B.D., Maruniak, J.E., 1993. Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. J. Invertebr. Pathol. 62, 147–164.

Zhang, C.X., Ma, X.C., Guo, Z.J., 2005. Comparison of the complete genome sequence between C1 and G4 isolates of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus. Virology 333, 190–199.