Journal of Invertebrate Pathology 137 (2016) 10–22

Contents lists available at ScienceDirect

Journal of Invertebrate Pathology

journal homepage: www.elsevier.com/locate/jip

Geographic isolates of *Lymantria dispar multiple nucleopolyhedrovirus*: Genome sequence analysis and pathogenicity against European and Asian gypsy moth strains

Robert L. Harrison [a,*], Daniel L. Rowley [a], Melody A. Keena [b]

[a] Invasive Insect Biocontrol and Behavior Laboratory, Beltsville Agricultural Research Center, USDA Agricultural Research Service, 10300 Baltimore Avenue, Beltsville, MD 20705, USA
[b] Northern Research Station, USDA Forest Service, 51 Mill Pond Road, Hamden, CT 06514, USA

ARTICLE INFO

Article history:
Received 25 September 2015
Revised 7 March 2016
Accepted 29 March 2016

Keywords:
Baculovirus
Nucleopolyhedrovirus
Gypsy moth
Asian gypsy moth
Lymantria dispar
LdMNPV

ABSTRACT

Isolates of the baculovirus species *Lymantria dispar multiple nucleopolyhedrovirus* have been formulated and applied to suppress outbreaks of the gypsy moth, *L. dispar*. To evaluate the genetic diversity in this species at the genomic level, the genomes of three isolates from Massachusetts, USA (LdMNPV-Aba624), Spain (LdMNPV-3054), and Japan (LdMNPV-3041) were sequenced and compared with four previously determined LdMNPV genome sequences. The LdMNPV genome sequences were collinear and contained the same *homologous repeats* (hrs) and clusters of *baculovirus repeat orf* (*bro*) gene family members in the same relative positions in their genomes, although sequence identities in these regions were low. Of 146 non-*bro* ORFs annotated in the genome of the representative isolate LdMNPV 5-6, 135 ORFs were found in every other LdMNPV genome, including the 37 core genes of *Baculoviridae* and other genes conserved in genus *Alphabaculovirus*. Phylogenetic inference with an alignment of the core gene nucleotide sequences grouped isolates 3041 (Japan) and 2161 (Korea) separately from a cluster containing isolates from Europe, North America, and Russia. To examine phenotypic diversity, bioassays were carried out with a selection of isolates against neonate larvae from three European gypsy moth (*Lymantria dispar dispar*) and three Asian gypsy moth (*Lymantria dispar asiatica* and *Lymantria dispar japonica*) colonies. LdMNPV isolates 2161 (Korea), 3029 (Russia), and 3041 (Japan) exhibited a greater degree of pathogenicity against all *L. dispar* strains than LdMNPV from a sample of Gypchek. This study provides additional information on the genetic diversity of LdMNPV isolates and their activity against the Asian gypsy moth, a potential invasive pest of North American trees and forests.

Published by Elsevier Inc.

1. Introduction

Lymantria dispar multiple nucleopolyhedrovirus is a species in the genus *Alphabaculovirus* of the insect virus family *Baculoviridae* (Herniou et al., 2012). Viruses of this family form virions consisting of a single double-stranded circular DNA genome contained in an enveloped, rod-shaped capsid (Harrison and Hoover, 2012). Baculovirus DNA replication and progeny virus assembly occur in the nucleus of the host. Initially, a type of progeny virion, referred to as budded virus (BV), is produced when nucleocapsids exit the nucleus and bud through the host plasma membrane, acquiring an envelope in the process. Later during the replication cycle, nucleocapsids are enveloped, either singly or in bunches, within the nucleus. The resulting virions, known as occlusion-derived virus (ODV), are assembled into a paracrystalline matrix composed of a single viral protein, polyhedrin, which is synthesized at very high levels in infected cells. This occlusion process results in the formation of occlusion bodies (OBs; also known as polyhedra) which contain multiple virions (Slack and Arif, 2007).

OBs serve to transmit viral infection horizontally within a host population (Fuxa, 2004). Larval stages of insects from orders Lepidoptera, Diptera, and Hymenoptera become infected when they ingest OBs. The OB protein matrix dissolves in the host gut, and the liberated ODV enter the host midgut epithelial cells. From this point, progeny BV secreted from infected cells serve to disseminate infection within the host. Baculovirus infections generally result in the death of the host. Cadavers of baculovirus-killed hosts break down due to the action of viral-encoded degradative enzymes, and progeny OBs are released into the environment. The OB matrix

* Corresponding author at: Invasive Insect Biocontrol and Behavior Laboratory, Beltsville Agricultural Research Center, USDA Agricultural Research Service, Building 007, Room 301, BARC-W, 10300 Baltimore Ave., Beltsville, MD 20705, USA.
E-mail addresses: Robert.L.Harrison@ars.usda.gov (R.L. Harrison), Daniel.Rowley@ars.usda.gov (D.L. Rowley), mkeena@fs.fed.us (M.A. Keena).

http://dx.doi.org/10.1016/j.jip.2016.03.014
0022-2011/Published by Elsevier Inc.

confers a degree of environmental persistence to the virions occluded within, which allows for the retention of infectivity until another larva ingests the OBs and renews the cycle.

Baculovirus OBs have been produced and formulated for use as safe, ecologically and environmentally friendly biopesticides (Moscardi, 1999). Isolates of *Lymantria dispar multiple nucleopolyhedrovirus* have been used in this way to control outbreaks of its host, *L. dispar*, the gypsy moth, a pest of trees and forests (Solter and Hajek, 2009). Populations of the subspecies *Lymantria dispar dispar* (common name: European gypsy moth) are found in Europe and in North America, where it is an invasive pest. *L. dispar dispar* has been spreading throughout the Northeast corner of the USA and adjacent areas in Canada since its introduction in Massachusetts, USA in the late 1860s (Pogue and Schaefer, 2007). Its dispersal has been slow, likely due to efforts to hinder its spread and to the inability of adult females to fly. There are also two subspecies, *Lymantria dispar asiatica* and *Lymantria dispar japonica*, which are part of the Asian gypsy moth complex (Pogue and Schaefer, 2007). Populations of *L. dispar asiatica* are found in Russia east of the Ural Mountains, northern China, and the Korean peninsula, while *L. dispar japonica* populations are found in Japan. The Asian gypsy moth subspecies are currently not established in North America, but Asian gypsy moths have been detected and eradicated in the United States on at least 23 occasions between 1991 and 2014 (USDA/APHIS/PPQ, 2015). The Asian gypsy moth poses a serious invasive threat to North American trees and forests due to the broader plant host range of the larvae and the ability of adult females to fly.

The isolate Lymantria dispar multiple nucleopolyhedrovirus LDP-67 (LdMNPV LDP-67) has formed the basis for Gypchek, a biocontrol product currently produced by the USDA Forest Service and Sylvar Technologies Inc. (Canada) for use against outbreaks of gypsy moth in North America (Reardon et al., 2012). Other LdMNPV isolates have been tested for activity against gypsy moth populations in North America and Asia (Duan et al., 2012; Lewis et al., 1984; Narang et al., 2001; Shapiro et al., 1984). An isolate has also been used to develop the product Virin-ENSh for use against gypsy moth in the former Soviet Union (Alyoshina, 1980). Results from comparative bioassays with LdMNPV isolates have raised the possibility that there may be differences in the susceptibilities of European and Asian gypsy moth larvae to LdMNPV infection (Ebling et al., 2004). Differences in pathogenicity may affect the capacity of current formulations of Gypchek to control invading populations of Asian gypsy moth.

Basic research on the molecular biology and genetics of LdMNPV has been carried out primarily with strain LDP-67 or clonal isolates derived from it (McClintock et al., 1986; Slavicek and Podgwaite, 1992; Slavicek et al., 1995, 1992). The entire genome sequence of the plaque isolate LdMNPV 5-6 was determined by Sanger dideoxy sequencing in 1999 (Kuzio et al., 1999). The advent of next-generation sequencing technologies have facilitated the sequencing of baculovirus genomes, and genome sequences have now been determined for isolates LdMNPV-2161 from South Korea (Harrison et al., 2014); LdMNPV-27 from Western Siberia, Russia (Kabilov et al., 2015); and LdMNPV-3029, a sample from the biopesticide Virin-ENSh (Harrison and Rowley, 2015). The data from these sequences, along with data from partial sequencing of the *lef-8* gene from several additional LdMNPV isolates in a USDA insect virus collection, suggest that viruses from the *L. dispar* populations in Europe and North America have diverged from viruses found in Asian *L. dispar* populations (Harrison et al., 2014).

In this study, three additional LdMNPV genomes – one from a plaque isolate derived from a Massachusetts (USA) population, an isolate from Spain, and an isolate from Japan – were completely sequenced in order to amass more information on the genetic diversity of this group of viruses at the genomic level and confirm the grouping of LdMNPV isolates into European/North American and Asian assemblages. The pathogenicities of the isolates in Gypchek and Virin-ENSh and Asian LdMNPV isolates were compared in bioassays with European and Asian gypsy moth colonies to obtain information on LdMNPV phenotypic diversity and to evaluate the control potential of different LdMNPV isolates against European and Asian gypsy moth populations.

2. Materials and methods

2.1. Virus isolates and insects

LdMNPV-Ab-a624 is a plaque isolate obtained by plating hemolymph of larvae infected with an Abington, MA LdMNPV sample on the cell line IPLB-LdEIta (Lynn et al., 1993). Other LdMNPV isolates featured in this study are from a USDA Agricultural Research Service insect virus collection maintained in Beltsville, MD, and include LdMNPV-3049, a sample of Gypchek deposited in September 1997; LdMNPV-2161, an isolate collected in South Korea by D. K. Reed (Pemberton et al., 1993) and deposited September 28, 1993; LdMNPV-3029, a sample of Virin-ENSh; LdMNPV-3041, collected in Japan; and LdMNPV-3054, an isolate from Spain deposited November 26, 1980 (Harrison et al., 2014). Virus isolates were grown in 3rd and 4th instar larvae of the New Jersey Standard Strain of *L. dispar*, reared from eggs obtained from the USDA APHIS rearing facility in Otis AFB, MA on *L. dispar*-specific diet from Southland Products (Lake Village, AR) at 28 °C on a 14:10 light:dark cycle.

Bioassays were carried out with a selection of Asian and European gypsy moth strains maintained at the USDA Forest Service Northern Research Station Quarantine Facility in Ansonia, CT. These strains included *L. dispar japonica* strain JN from Nagoya, Japan; two *L. dispar asiatica* strains, including strain RM from Mineralni, Primorski in Far East Russia; strain RB from Bellyk, Krasnoyarsk in Siberia, Russia; and three *L. dispar dispar* strains, including strain LJ from Juodkrante, Kuzsin Nezijos in Lithuania; strain KG from Kavála, Macedonia in Greece; and strain UC from Bethany, New Haven County in Connecticut, USA. These strains and their maintenance are described in Keena et al. (2008). Each generation is produced from 100 randomly-selected egg masses to maintain genetic diversity. The identities of these colonies have been confirmed in a recent barcoding study (Chen et al., 2016).

2.2. Genomic DNA preparation and 454 sequencing

For each virus isolate to be sequenced, genomic DNA was isolated and sequenced as previously described (Harrison and Lynn, 2007; Harrison et al., 2014). Sequencing reads from a Roche 454 GS Junior instrument were sorted and assembled using the SeqMan NGEN V3.0 assembler program (Lasergene; DNASTAR, Inc., Madison, WI) with default parameters. Gaps were closed and regions with ambiguous sequences or unusual features were resolved or confirmed by PCR amplification and Sanger dideoxy sequencing. The Lasergene SeqManPro (version 9) sequence editor was used to prepare the final contigs of the consensus genome sequences. Sequence coverage and GenBank accession numbers for each isolate are listed in Table 1.

Open reading frames (ORFs) were manually annotated for each genome by selecting ORFs of at least 50 codons that did not overlap adjacent ORFs by >75 bp. ORFs were also selected for which annotated homologues existed in other baculovirus genomes, including other genomes of LdMNPV. BLASTp queries were carried out to determine the relatedness of predicted amino acid sequences to those of LdMNPV 5-6 and other baculoviruses. Intergenic *homologous repeat* (*hr*) sequences were identified by searching the

Table 1
Isolates of *Lymantria dispar* nucleopolyhedrovirus with completely sequenced genomes.

LdMNPV isolate	Source	Reference	Genome size, bp (coverage)	Annotated ORFs[a]	*hrs*	Annotated *bro* genes	GenBank ID
5–6 (representative isolate)	Connecticut, USA	Kuzio et al. (1999)	161,046	163 (176)	13	16	AF081810
2161	South Korea	Harrison et al. (2014)	163,138 (19.34×)	174 (179)	13	20	KF695050
3029 (Virin EnSH)	Russia	Harrison and Rowley (2015)	161,712 (132.7×)	168 (176)	12	15	KM386655
27	Western Siberia, Russia	Kabilov et al. (2015)	164,108	162 (172)	13	18	KP027546
Ab-a624	Massachusetts, USA	This study	161,321 (127.3×)	176	13	15	KT626572
3041	Japan	This study	162,658 (221.5×)	178	12	19	KT626571
3054	Spain	This study	164,478 (219.6×)	175	13	17	KT626570

[a] ORF numbers in parentheses include ORFs that were not annotated in the original publication of the genome sequence.

genome for repeated sequences matching the *hr* consensus sequence described by Kuzio et al. (1999). Supplementary Tables 1 through 4 list all of the ORFs and *hrs* for isolates 3029, Ab-a624, 3041, and 3054, respectively, and includes information on nucleotide locations and amino acid sequence similarities with LdMNPV 5-6 orthologues.

Genome sequences of the LdMNPV isolates were aligned with the genome sequence of LdMNPV 5-6 using the Martinez-Needleman-Wunsch method of MegAlign (Lasergene, v. 12) with default parameters. The percent nucleotide sequence identity between isolate genome sequences in pairwise alignments was determined by dividing the product of the Similarity Index and the Consensus Length by the difference between the Consensus Length and the Gap Length [(Similarity Index × Consensus Length)/(Consensus Length − Gap Length)]. A global alignment of the LdMNPV genome sequences, either by themselves or with the genome of Lymantria xylina multiple nucleopolyhedrovirus-5 (LyxyMNPV-5; Nai et al., 2010), was carried out with Mauve 2.4.0 (Darling et al., 2004) using the mauveAligner algorithm with default parameters.

2.3. Phylogenetic inference

For the seven LdMNPV isolates that have been completely sequenced (Table 1), the nucleotide sequences for the 37 currently recognized core genes proposed to be present in all viruses of family *Baculoviridae* (Garavaglia et al., 2012) were aligned by CLUSTAL W (Thompson et al., 1994) using MegAlign with default parameters. Alignments also included sequences from LyxyNPV-5, which was used as an outgroup for phylogenetic inference.

Nucleotide sequence alignments were concatenated with BioEdit (Hall, 1999) and phylogenetic trees inferred from the alignments in MEGA 6.06 (Tamura et al., 2013). Maximum likelihood (ML), minimum evolution (ME), and maximum parsimony (MP) tree construction methods were used with bootstrap re-sampling. The Tamura 3-parameter substitution model was used for ML and ME analysis, and value of the shape parameter for the discrete gamma distribution used for modeling rate differences among sites was estimated from the alignment for the ME analysis.

2.4. Bioassays

Bioassays were carried out with neonate *L. dispar* larvae by the droplet feeding method (Hughes and Wood, 1981) as previously described (Harrison et al., 2014). For the second and third iterations of the bioassay, the concentration range for LdMNPV-3049/Gypchek was modified to 1×10^4, 3×10^4, 1×10^5, 3×10^5, and 9×10^5 OB/mL. Larvae were allowed to drink the OB dilutions and were then transferred to 32-cell trays (1 tray/concentration; Frontier Agricultural Sciences, Newark, DE) containing 7 mL of Southland Products gypsy moth diet/cell. To provide levels of dietary iron optimal for larval growth and development (Odell et al., 1997), the diet was supplemented with ferric citrate (Sigma–Aldrich, St. Louis, MO; catalog #F3388) at a rate of 0.03 g/L for strains UC and KG, 0.07 g/L for strains LJ and RB, and 0.11 g/L for strains RM and JN. Three replicate bioassays were carried out at 8-month intervals over the course of 17 months when neonate larvae were available. The LC_{50} values and the slopes and intercepts of probit concentration–response lines were calculated using PoloPlus 2.0 (Robertson et al., 2007). LC_{50}s were compared using the lethal dose ratio test described in Robertson et al. (2007).

3. Results

3.1. Features of LdMNPV genome sequences

Three novel LdMNPV genome sequences of isolates LdMNPV-Ab-a624, LdMNPV-3041, and LdMNPV-3054 were determined. Table 1 lists the general features of these genome sequences and the sequences of four other LdMNPV isolates, including the representative isolate for the species, LdMNPV 5-6. The genome sizes among the isolates exhibited a 0.3% difference between the largest and smallest genomes, ranging from 161,046 bp (isolate 5-6) to 164,478 bp (3054). Comparable numbers of ORFs, *homologous repeat* (*hr*) regions, and *baculovirus repeat ORF* (*bro*) family members were identified in each isolate.

Pairwise alignments of genome sequences with the reference isolate LdMNPV 5-6 (Table 2) revealed overall nucleotide sequence identities ranging from 96.8% (5-6 × 3041) to 99.2% (5-6 × Ab-a624). However, hundreds of gaps were inserted to optimize the alignments. The total lengths of these gaps ranged from >4 to >17 kbp. A portion of the gap length could be accounted for by large deletions unique to individual isolates, such as the 697-bp deletion in isolate 5-6 that removes half of an ORF encoding the P24 capsid protein and all of an upstream ORF (Slavicek and Hayes-Plazolles, 2003). A global alignment of all the LdMNPV genome sequences with Mauve confirmed that the LdMNPV genomes were collinear with each other. The Mauve alignment also revealed a previously described inversion in the genome of the closely related LyxyNPV-5 (Nai et al., 2010) (Supplementary Figure 1).

Gaps in the pairwise alignments were clustered in the genome regions containing *homologous repeat* (*hr*) sequences and members

Table 2
Pairwise alignment and comparison of sequenced LdMNPV genomes with LdMNPV 5–6.

Alignment	% Sequence identity	Number of gaps	Total gap length, bp
5-6 × Ab-a624	99.2	203	4467
5-6 × 3029	98.2	410	8607
5-6 × 3054	97.8	494	12,636
5-6 × 27	97.8	498	12,508
5-6 × 2161	97.5	566	12,101
5-6 × 3041	96.8	743	17,658

of the *baculovirus repeated orf* (*bro*) gene family (van Oers and Vlak, 2007). Nevertheless, all seven isolates contained at least 12 of the 13 h reported for isolate 5-6 (Fig. 1). Isolates 3029 and 3041 did not contain a single-repeat *hr* that is part of a cluster of four *hr*s (*hr7a* –*hr7d*) in isolate 5-6.

In addition, all of the LdMNPV genomes contained four clusters of *bro*s, located in the same relative positions on the genomes: between LdMNPV 5-6 ORFs 30 and 34, *chitinase* and ORF 81 (*ac111*), ORF111 and *dutpase*, and *sod* and *pif-1*, as well as single *bro* adjacent to *hr8* (Fig. 1). The numbers of *bro*s in the four clusters differed among the genomes, and BLAST queries with *bro* amino acid sequences indicated a low degree of conservation among *bro* sequences in the same genomic locations of the different isolates. This low degree of conservation is also evident in a Mauve alignment of the LdMNPV nucleotide genome sequences (Fig. 2). Regions of the Mauve alignment containing the *bro* genes either exhibited a low similarity profile or were excluded from the alignment's locally collinear blocks, indicating a lack of significant sequence similarity among the isolates in these regions.

3.2. Conservation of ORFs among LdMNPV isolates

Table 3 lists the 146 non-*bro* ORFs originally annotated for LdMNPV 5-6 and their distribution among the other six completely sequenced LdMNPV genomes, along with the sizes of the encoded proteins and amino acid sequence identities with the 5-6 orthologues. This list includes the *p24* (*ac129*) ORF, but excludes 5-6 ORFs 133 and 134, whose annotation was an artifact of the deletion in isolate 5-6 that removed *p24* (Slavicek and Hayes-Plazolles, 2003). Of these 146 ORFs, 135 are conserved in all of the other LdMNPV genomes. Among these 135 conserved ORFs are the 37 core genes of *Baculoviridae* listed in Table 1 of Garavaglia et al., 2012. In addition, all LdMNPV genomes contain the 9 genes found in all sequenced alpha-, beta-, and gammabaculoviruses, the 16 genes found in all sequenced alpha- and betabaculoviruses, and *ac23*, which occurs in the single deltabaculovirus genome that has been sequenced in addition to all alpha- and betabaculovirus genomes (Garavaglia et al., 2012). Also present in every LdMNPV genome are an additional 33 orthologues of ORFs found in

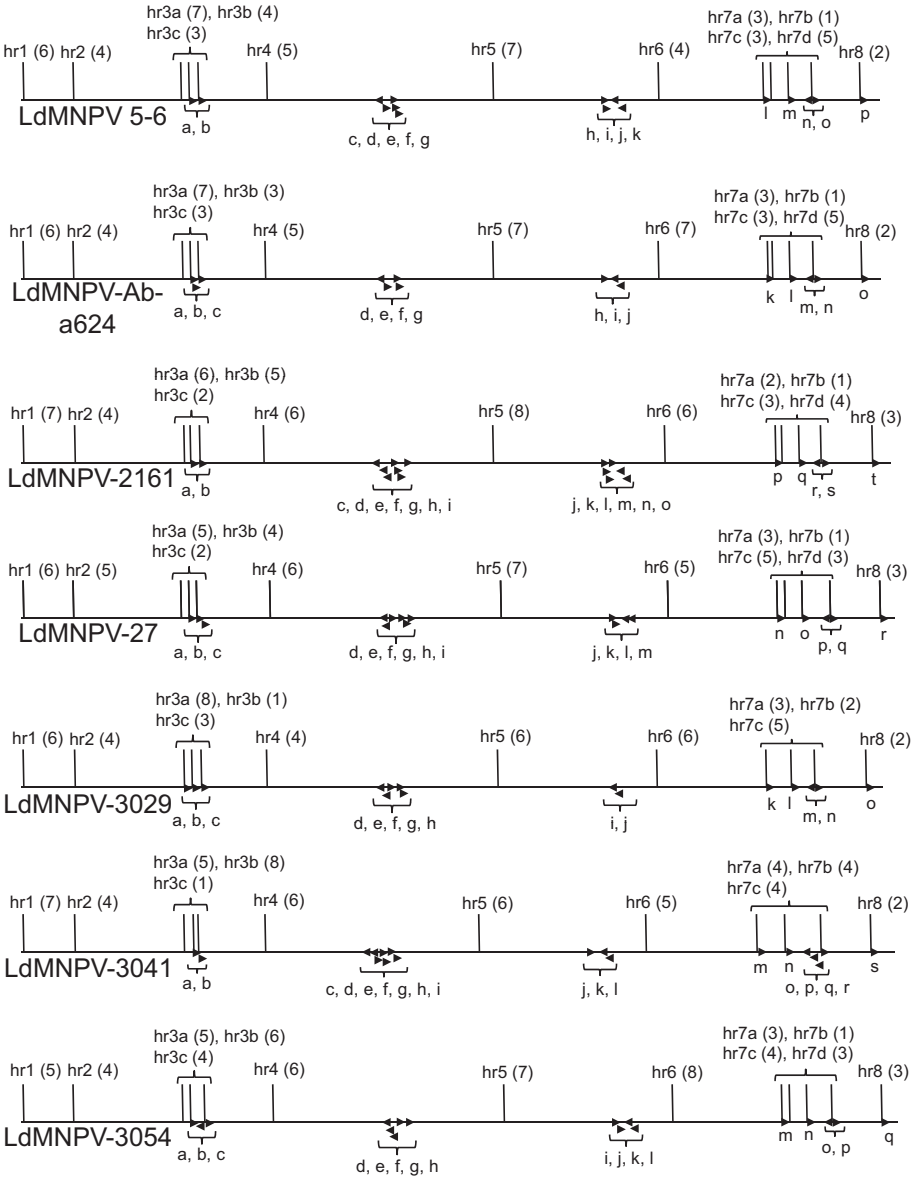

Fig. 1. Distribution and positions of *homologous repeat* (*hr*) regions and copies of *baculovirus repeated orf* (*bro*) genes in the genomes of seven LdMNPV isolates. The *hr*s are indicated by vertical lines, with the number of unit palindromes for each *hr* given in parentheses. The *bro*s are indicated by lettered arrowheads, and the direction of the arrow indicates orientation on the genome. Brackets denote clusters of *hr*s or *bro*s.

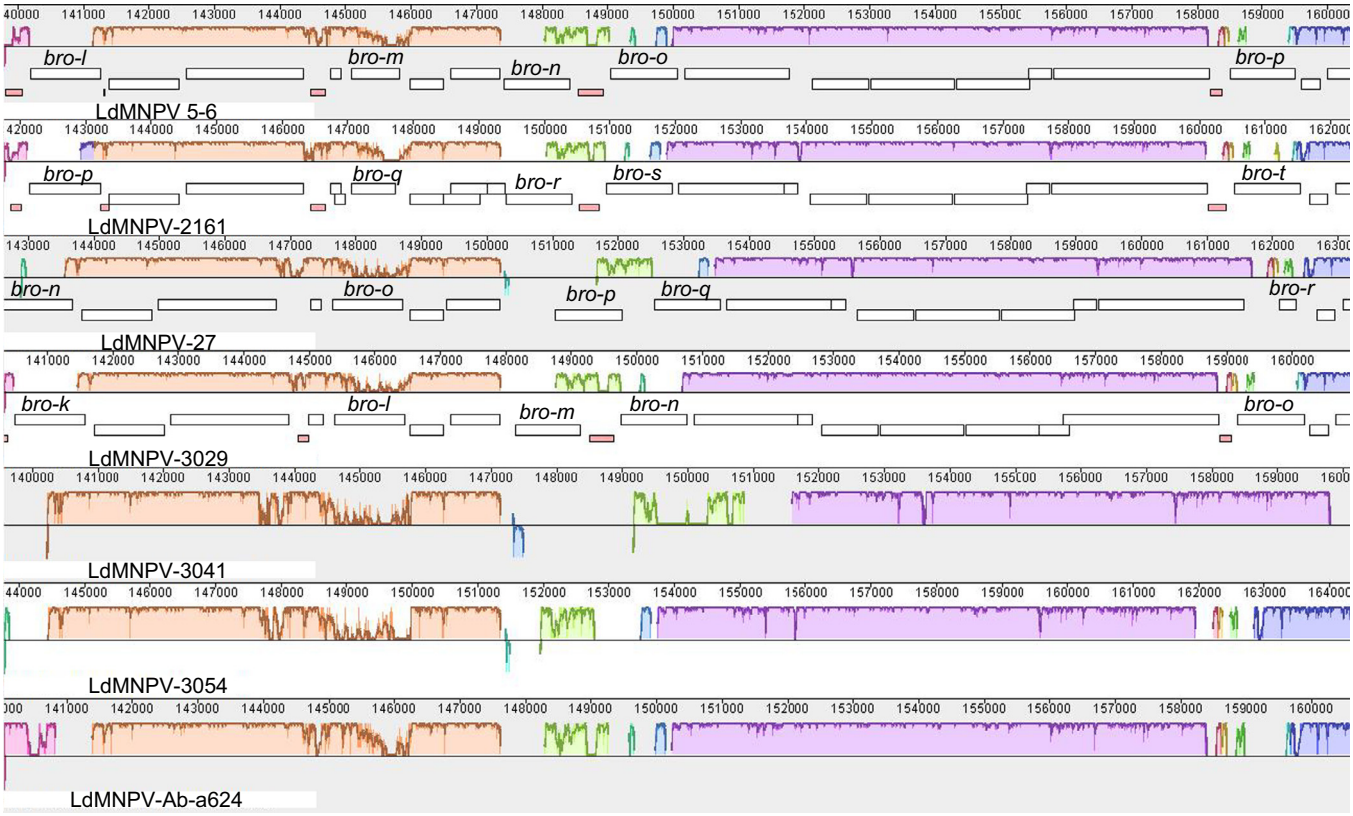

Fig. 2. Mauve alignment of the last approximately 25 kbp of seven LdMNPV genome sequences, showing the low degree of conservation of *bro* gene nucleotide sequences relative to other regions of the genomes. Block outlines of the same color correspond to Locally Collinear Blocks (LCBs), which are segments of the sequence that are conserved among the isolates and free of internal rearrangements. Nucleotide sequence positions for each genome are indicated on a line above the LCBs. The height of the profile within each LCB corresponds to the average level of sequence conservation among the isolates in that region of the genome sequence. Regions outside the LCBs lack a significant degree of sequence identity among the isolate genome sequences. White boxes below the genomes for LdMNPV isolates 5-6, 2161, 27, and 3029 correspond to annotate ORFs, and red boxes correspond to *hr*s (which are not annotated in the LdMNPV-27 GenBank record). The *bro* genes for these four isolates are indicated.

Autographa californica multiple nucleopolyhedrovirus C6 (AcMNPV-C6), the representative isolate for *Alphabaculovirus* type species *Autographa californica multiple nucleopolyhedrovirus* (Ayres et al., 1994).

Eleven ORFs annotated in isolate 5-6 are not conserved in every LdMNPV genome, including ORFs 5, 6, 8, 10, 13, 31, 49, 65, 66, 69, and 121 (Table 3). ORF31 is the only 5-6 ORF that is not present in any of the other isolate genomes. The encoded 59-amino acid sequence of this ORF does not share significant identity with any other sequence in GenBank. The ORF itself is located upstream of *hr3b*, and the sequence containing the start codon for the ORF does not occur in other LdMNPV sequences. ORF66 is annotated as *ctl-2* in isolate 5-6, and encodes a conotoxin-like peptide homologue. Orthologues of this ORF are also found in the genomes of LyxyMNPV-5 (Nai et al., 2010), Orygia pseudotsugata multiple nucleopolyhedrovirus (Ahrens et al., 1997), and at least five other related alphabaculoviruses. A truncated *ctl-2* ORF occurs in LdMNPV-3041, but not in any other LdMNPV isolate. An upstream ORF, ORF65, is annotated as *vef-1* and encodes one of two enhancins expressed by LdMNPV 5-6 (Kuzio et al., 1999). A 2.3 kbp deletion has removed this ORF entirely from the genome of LdMNPV-3041.

Frameshift mutations and the occurrence of in-frame stop and start codons have significantly altered some of the ORFs conserved among LdMNPV isolates. Perhaps the most striking example of this phenomenon is ORF4, *mucin-like*. The orthologues of this ORF in the other isolates are 149–282 codons longer than isolate 5-6 *mucin-like* due to the occurrence of upstream in-frame start codons (Table 3). In absence of empirical data, it is unclear if the additional N-terminal codons in these orthologues are actually transcribed and translated. Other examples of differences in ORF size of >20% include ORF29 (*ac4*), ORF 48 (*ac40*; *p47*), ORF49, ORF58 (*ac55*), ORF145 (*ac31*; *sod*), and ORF156 (*ac32*; *fgf*). A deletion occurring in the LdMNPV-Ab-a624 orthologue of enhancin-encoding ORF160 (*vef-2*) has truncated the ORF, such that the coding sequence for this gene is split between two ORFs. It is unclear if either of these ORFs encodes a functional enhancin.

Amino acid sequence identities of the conserved ORFs with their orthologues in LdMNPV 5-6 tended to be >90%, with 9 of the ORFs sharing 100% sequence identity between LdMNPV 5-6 and every other LdMNPV isolate. ORFs with <90% amino acid sequence identity between one or more pairs of isolates included LdMNPV 5-6 ORF12, 28, 29 (*ac4*), 34, 49, 55, 61 (*ac59*), 66 (*ctl-2*), 91 (*ac83*; *vp91*), 101 (*ac100*; *p6.9*), 121, and 132.

Twenty-three ORFs not originally annotated in the LdMNPV 5-6 genome were identified and annotated in genomes of the other isolates (Table 4). Thirteen of these ORFs can also be found in the 5-6 genome sequence, and eleven of these thirteen are present in every LdMNPV genome. Two ORFs have annotated orthologues in multiple NPVs, and six have been annotated only in the LyxyNPV-5 genome.

3.3. Phylogenetic relationships among LdMNPV isolates

Phylogenetic inference with concatenated nucleotide sequence alignments of the 37 baculovirus core genes yielded clades with generally strong bootstrap support (Fig. 3). Two clades consisting of isolates from Asia (LdMNPV-2161 and LdMNPV-3041) and

Table 3
Distribution of conserved LdMNPV 5-6 ORFs among other isolates of LdMNPV.

5-6		Ab-a624		27		2161		3029		3041		3054	
Number (name)	Size (aa)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)
1 (ac8, polh)	245	245	100% (245/245)	245	100% (245/245)	245	100% (245/245)	245	100.0% (245/245)	245	100% (245/245)	245	100.0% (245/245)
2 (ac9, pp78)	555	555	100% (555/555)	527	90.3% (504/558)	552	93.4% (524/561)	547	96.4% (535/555)	552	93.6% (528/564)	528	90.7% (506/558)
3 (ac10, pk-1)	274	274	100% (274/274)	274	98.9% (271/274)	274	98.9% (271/274)	274	99.3% (272/274)	274	98.9% (271/274)	274	99.3% (272/274)
4 (mucin-like)	1029	1278	89.1% (959/1076)	1178	83.3% (870/1045)	1311	85.2% (946/1110)	1282	91.9% (974/1060)	1286	86.7% (944/1089)	1262	91.2% (954/1046)
5	189	189	99.5% (188/189)			189	94.7% (179/189)	189	98.4% (186/189)	189	94.2% (178/189)	189	98.4% (186/189)
6	80	80	97.5% (78/80)					71	97.4% (38/39)	63	100% (43/43)		
7	191	180	98.9% (177/179)	180	98.9% (177/179)	180	98.9% (177/179)	180	97.8% (175/179)	182	98.4% (179/182)	180	99.4% (178/179)
8	119	119	100% (119/119)							137	90.6% (29/32)	97	94.6% (88/93)
9	172	172	99.4% (171/172)	171	93.0% (160/172)	187	93% (160/172)	171	94.8% (163/172)	167	87.8% (151/172)	171	93.0% (160/172)
10	73	73	100% (73/73)					69	89.9% (62/69)				
11	244	244	99.6% (243/244)	252	93.3% (235/252)	257	83.4% (216/259)	252	94.0% (237/252)	254	84.8% (217/256)	252	94.4% (238/252)
12	172	190	85.9% (164/191)	178	92.7% (165/178)	182	89.1% (164/184)	172	95.3% (162/170)	178	86.5% (154/178)	178	92.7% (165/178)
13	66	56	94.7% (18/19)										
14 (ac148, odv-e56)	356	356	99.7% (355/356)	356	98.9% (352/356)	356	98.9% (352/356)	356	98.3% (350/356)	356	98.9% (352/356)	356	98.9% (352/356)
15 (ac147, ie-1)	566	564	96.6% (547/566)	561	97.0% (549/566)	563	95.9% (543/566)	563	97.7% (553/566)	563	97.0% (549/566)	561	97.0% (549/566)
16 (ac146, ep23)	208	208	99.0% (206/208)	206	96.6% (199/206)	207	94.7% (197/208)	208	98.1% (204/208)	213	93.7% (193/206)	206	96.6% (199/206)
17 (ac145)	92	92	100% (92/92)	92	98.9% (91/92)	92	100% (92/92)	92	98.9% (91/92)	92	98.9% (91/92)	92	98.9% (91/92)
18 (ac144, odv-ec27)	283	283	99.6% (282/283)	283	99.3% (281/283)	283	99.3% (281/283)	283	99.3% (281/283)	283	99.3% (281/283)	283	99.3% (281/283)
19 (ac143, odv-e18)	88	88	100% (88/88)	88	100% (88/88)	88	100% (88/88)	88	100% (88/88)	88	100.0% (88/88)	88	100.0% (88/88)
20 (ac142, p49)	483	483	99.6% (481/483)	483	99.7% (482/483)	483	99.6% (481/483)	483	99.8% (482/483)	483	99.8% (482/483)	483	99.8% (482/483)
21 (ac141, exon-0)	258	258	99.2% (256/258)	258	98.8% (255/258)	258	98.8% (255/258)	258	98.4% (254/258)	269	98.4% (254/258)	258	98.8% (255/258)
22 (dna ligase)	548	548	100% (548/548)	548	97.4% (535/549)	549	97.4% (535/549)	548	97.4% (535/549)	549	97.3% (534/549)	548	97.3% (534/549)
23 (ac139, me53)	342	342	100% (342/342)	342	99.1% (339/342)	341	98.8% (337/341)	341	99.7% (340/341)	341	99.4% (339/341)	342	99.1% (339/342)
24	208	208	99.5% (207/208)	208	97.1% (202/208)	208	97.1% (202/208)	208	98.1% (204/208)	208	97.1% (202/208)	208	97.6% (203/208)
25	154	154	100% (154/154)	154	96.8% (149/154)	154	97.4% (150/154)	154	99.4% (153/154)	154	96.8% (149/154)	154	96.1% (148/154)
26	72	72	100% (72/72)	72	98.6% (71/72)	72	98.6% (71/72)	72	100% (55/55)	72	98.6% (71/72)	72	100.0% (72/72)
27 (ac138, p74)	672	672	99.3% (667/672)	672	99.3% (667/672)	672	99.0% (665/672)	672	99.3% (667/672)	672	98.8% (664/672)	672	99.1% (666/672)
28	379	379	98.4% (373/379)	360	78.5% (296/377)	360	79% (298/377)	379	98.7% (374/379)	374	94.7% (357/377)	360	78.5% (296/377)
29 (ac4)	146	146	96.6% (141/146)	88; 84	76.8% (63/82); 97.6% (82/84)	148	91.5% (129/1)	90	75.0% (63/84)	148	85.8% (127/148)	88	76.8% (63/82)
30 (ac150)	94	94	100% (94/94)	94	96.8% (91/94)	94	92.6% (87/94)	84	100% (84/84)	94	91.5% (86/94)	94	96.8% (91/94)
31	59												
34	253	253	100% (253/253)	261	82.4% (206/250)	266	94.7% (234/247)	253	100% (253/253)	253	96.8% (245/253)	261	82.8% (207/250)
35 (ac11)	359	359	99.7% (358/359)	359	95.5% (343/359)	361	95.8% (344/359)	361	98.6% (354/359)	361	95.5% (343/359)	359	95.5% (343/359)
36 (ac26)	123	123	100% (123/123)	123	100% (123/123)	123	99.2% (122/123)	123	99.2% (122/123)	123	97.6% (120/123)	123	100.0% (123/123)
37 (ac25, dbp)	239	239	100% (239/239	239	97.9% (234/239)	239	99.2% (237/239)	239	99.2% (237/239)	239	98.7% (236/239)	239	98.7% (236/239)
38 (ac28, lef-6)	159	157	98.1% (156/159)	159	100% (159/159)	159	98.7% (157/159)	161	98.8% (159/161)	160	98.8% (158/160)	159	100.0% (159/159)
39 (ac29)	68	68	100% (68/68)	68	100% (68/68)	68	98.5% (67/68)	68	100% (68/68)	68	98.5% (67/68)	68	100.0% (68/68)
40 (ac136, p26)	253	253	100% (253/253)	252	98.8% (251/254)	252	97.2% (247/254)	252	99.6% (252/253)	252	97.6% (248/254)	252	98.8% (251/254)
41 (ac137, p10)	77	77	100% (77/77)	77	100% (77/77)	77	100% (77/77)	77	100% (77/77)	77	100.0% (77/77)	77	100.0% (77/77)
42 (ac34)	188	188	99.5% (187/188)	188	99.5% (187/188)	188	98.9% (186/188)	188	99.5% (187/188)	188	98.9% (186/188)	188	(187/188) 99.5%
43 (ac35, v-ubi)	150	150	98.7% (148/150)	151	96.0% (145/151)	150	96.7% (145/150)	150	98.7% (148/150)	150	96.0% (144/150)	151	96.0% (145/151)
44 (ac36, 39k/pp31)	264	264	100% (264/264)	264	100% (264/264)	263	98.5% (260/264)	263	98.5% (260/264)	263	99.6% (263/264)	264	100.0% (264/264)
45 (ac37, lef-11)	187	187	99.5% (186/187)	187	99.5% (186/187)	187	98.9% (185/187)	187	98.4% (184/187)	187	100.0% (187/187)	187	99.5% (186/187)
46 (ac38, bv-e31)	247	247	100% (247/247)	247	100% (247/247)	250	98.4% (246/250)	247	100.0% (247/247)	250	98.4% (246/250)	247	100.0% (247/247)
47 (ac25, dbp)	257	308	99.2% (255/257)	308	97.3% (250/257)	308	96.9% (249/257)	308	97.7% (251/257)	308	97.3% (250/257)	308	97.3% (250/257)
48 (ac40, p47)	390	561	100% (390/390)	390	99.7% (389/390)	390	99.7% (389/390)	390	99.7% (389/390)	390	99.5% (388/390)	390	99.7% (389/390)
49	106	189	85.8% (97/113)			60	83.9% (52/62)	138	73.9% (102/138)	84	69.1% (67/97)	154	94.1% (96/102)
50 (helicase-2)	460	461	99.3% (458/461)	460	99.3% (457/460)	458	97.8% (451/461)	458	97.6% (449/460)	458	97.8% (451/461)	460	99.3% (457/460)
51 (ac50, lef-8)	874	875	99.0% (866/875)	873	99.0% (864/873)	871	98.4% (858/872)	872	99.2% (865/872)	872	98.7% (861/872)	873	98.9% (863/873)
52 (bjdp)	299	299	100% (299/299)	298	96.7% (289/299)	291	93.3% (279/299)	299	97.0% (291/300)	288	92.6% (277/299)	298	94.3% (281/298)
53 (ac52)	300	300	100% (300/300)	300	98.7% (296/300)	229	97.8% (224/229)	300	99.3% (298/300)	229	98.3% (225/229)	300	99.0% (297/300)

(*continued on next page*)

Table 3 (*continued*)

5-6		Ab-a624		27		2161		3029		3041		3054	
Number (name)	Size (aa)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)
54 (ac53)	142	142	100% (142/142)	142	99.3% (141/142)	142	97.9% (139/142)	142	100% (142/142)	142	97.2% (138/142)	142	100.0% (142/142)
55	361	363	97.2% (353/363)	353	92.8% (335/361)	362	82.6% (303/367)	353	93.1% (336/361)	357	81.7% (300/367)	351	92.5% (334/361)
56 (ac53a, lef-10)	76	76	100% (76/76)	76	100% (76/76)	84	97.4% (74/76)	76	100% (76/76)	84	97.4% (74/76)	76	100.0% (76/76)
57 (ac54, vp1054)	332	332	100% (332/332)	332	99.7% (331/332)	332	98.2% (326/332)	332	99.7% (331/332)	332	98.2% (326/332)	332	99.7% (331/332)
58 (ac55)	64	64	100% (64/64)	64	100% (64/64)	89	96.9% (62/64)	64	98.4% (63/64)	89	98.4% (63/64)	64	100.0% (64/64)
59	53	53	100% (53/53)	53	100% (53/53)	53	98.1% (52/53)	53	100% (53/53)	53	98.1% (52/53)	53	100.0% (53/53)
60 (ac57)	164	164	99.4% (163/164)	160	94.5% (155/164)	160	93.9% (154/164)	163	98.2% (161/164)	160	93.3% (153/164)	160	94.5% (155/164)
61 (ac59, chaB1)	189	189	99.5% (188/189)	194	94.8% (184/194)	191	62.1% (128/206)	190	90.0% (171/190)	182	65.3% (128/196)	196	89.8% (176/196)
62 (ac60)	95	96	100% (69/69)	100	90.0% (90/100)	98	94.2% (65/69)	99	100% (63/63)	95	89.5% (85/95)	102	100.0% (69/69)
63 (ac61, fp25k)	217	217	100% (217/217)	220	97.7% (215/220)	220	96.4% (212/220)	256	98.2% (213/217)	220	92.7% (204/220)	220	97.7% (215/220)
64 (ac62, lef-9)	496	496	100% (496/496)	496	99.4% (493/496)	496	99.6% (494/496)	528	99.8% (495/496)	495	98.8% (488/494)	496	99.2% (492/496)
65 (vef-1)	783	783	99.7% (781/783)	784	97.7% (766/784)	783	97.2% (761/783)	783	97.5% (744/763)			783	97.7% (765/783)
66 (ctl-2)	92									53	86.7% (39/45)		
67 (hrf-1)	218	218	100% (218/218)	218	97.2% (212/218)	218	95.0% (207/218)	218	98.2% (214/218)	218	95.9% (209/218)	218	97.2% (212/218)
68 (ac64, gp37)	269	274	98.5% (263/267)	269	97.8% (263/269)	269	98.1% (264/269)	269	99.6% (268/269)	269	97.4% (262/269)	269	98.1% (264/269)
69	50			50	100% (50/50)	50	96.0% (48/50)						
70 (ac126, chitinase)	558	558	100% (558/558)	558	98.9% (552/558)	558	99.6% (556/558)	558	99.6% (556/558)	558	99.5% (555/558)	558	98.9% (552/558)
76 (ac111)	87	89	97.8% (87/89)	90	95.6% (86/90)	91	93.4% (85/91)	92	100% (45/45)	88	96.6% (85/88)	90	95.6% (86/90)
77	214	216	96.2% (202/210)	214	98.6% (211/214)	207	92.3% (191/207)	214	98.6% (211/214)	207	92.8% (192/207)	214	98.1% (210/214)
78 (ac127, v-cath)	356	356	100% (356/356)	356	99.7% (355/356)	360	97.2% (350/360)	356	99.7% (355/356)	360	98.1% (353/360)	356	99.4% (354/356)
79 (ac71, iap-2)	234	234	99.6% (233/234)	234	98.3% (230/234)	235	93.6% (220/235)	234	98.3% (230/234)	233	97.0% (227/234)	233	97.4% (228/234)
80 (ac68)	128	133	100% (126/126)	133	99.2% (125/126)	129	97.5% (117/120)	135	100% (113/113)	129	97.5% (117/120)	133	99.2% (125/126)
81 (ac67, lef-3)	374	374	99.2% (371/374)	373	96.8% (362/374)	373	96.8% (362/374)	373	98.6% (356/361)	373	97.3% (364/374)	373	96.8% (362/374)
82 (ac66, desmoplakin)	778	777	98.1% (763/778)	791	93.7% (747/797)	773	95.6% (745/779)	773	95.4% (742/778)	770	95.1% (744/782)	774	96.7% (753/779)
83 (ac65, dnapol)	1014	1014	98.4% (999/1015)	1014	98.5% (1000/1015)	1010	98.7% (1001/1014)	1012	98.0% (995/1015)	1011	98.3% (998/1015)	1014	98.6% (1001/1015)
84 (ac75)	128	128	100% (128/128)	128	100% (128/128)	128	100% (128/128)	128	100% (128/128)	128	100.0% (128/128)	128	100.0% (128/128)
85 (ac76)	86	86	100% (86/86)	86	100% (86/86)	86	100% (86/86)	86	100% (86/86)	86	100.0% (86/86)	86	100% (86/86)
86 (ac77, vlf-1)	378	378	99.7% (377/378)	378	99.7% (377/378)	378	99.2% (375/378)	378	99.7% (352/353)	378	99.2% (375/378)	378	99.7% (377/378)
87 (ac78)	113	112	98.2% (111/113)	111	96.5% (109/113)	112	95.6% (108/113)	111	87.6% (99/113)	111	94.7% (107/113)	111	94.7% (107/113)
88 (ac80, gp41)	323	323	100% (323/323)	323	100% (323/323)	323	100% (323/323)	323	100% (323/323)	323	100.0% (323/323)	323	100.0% (323/323)
89 (ac81)	219	219	100% (219/219)	219	98.2% (215/219)	222	97.3% (216/222)	216	99.5% (203/204)	222	98.2% (214/218)	226	98.2% (215/219)
90 (ac82, tlp-20)	223	223	100% (223/223)	223	97.3% (217/223)	226	94.7% (214/226)	220	95.1% (212/223)	226	92% (208/226)	229	96.1% (220/229)
91 (ac83, vp91)	864	860	96.9% (837/864)	853	96.0% (831/866)	850	91% (789/867)	854	95.5% (828/867)	839	89.3% (773/866)	844	95.9% (829/864)
92 (ac89, vp39)	352	352	100% (352/352)	354	98.0% (348/355)	346	94.6% (332/351)	345	100% (318/318)	348	98.8% (324/328)	351	99.1% (349/352)
93 (ac90, lef-4)	485	485	99.8% (484/485)	485	98.8% (479/485)	485	98.1% (476/485)	485	99.0% (480/485)	485	98.1% (476/485)	485	99.4% (482/485)
94 (ac92, p33)	251	251	100% (251/251)	251	98.4% (247/251)	251	100% (251/251)	251	98.8% (248/251)	251	100.0% (251/251)	251	99.6% (250/251)
95 (ac93)	159	159	100% (159/159)	159	100% (159/159)	159	100% (159/159)	159	100% (159/159)	159	100.0% (159/159)	159	100.0% (159/159)
96 (ac94, odv-e25)	217	217	99.1% (215/217)	217	98.6% (214/217)	217	98.6% (214/217)	217	98.3% (170/173)	217	98.6% (214/217)	217	99.5% (216/217)
97 (ac95, dnahel)	1218	1218	99.9% (1217/1218)	1218	99.8% (1216/1218)	1222	98.6% (1206/1223)	1218	99.8% (1216/1218)	1223	99.1% (1212/1223)	1218	99.8% (1216/1218)
98 (ac96, odv-e28)	173	173	100% (173/173)	173	100% (173/173)	173	100% (173/173)	173	99.4% (172/173)	173	100.0% (173/173)	173	99.4% (172/173)
99 (ac98, 38k)	322	322	99.7% (321/322)	322	97.8% (315/322)	322	98.1% (316/322)	322	98.0% (295/301)	322	97.8% (315/322)	322	98.8% (318/322)
100 (ac99, lef-5)	278	278	98.9% (275/278)	278	98.6% (274/278)	278	98.2% (273/278)	278	98.6% (274/278)	278	98.2% (273/278)	278	98.9% (275/278)
101 (ac100, p6.9)	99	102	97.1% (99/102)	104	95.2% (99/104)	103	89.7% (70/78)	102	97.1% (99/102)	104	95.2% (99/104)	102	96.1% (74/77)
102 (ac101, p40)	381	381	100% (381/381)	381	100% (381/381)	381	100% (381/381)	381	100% (381/381)	381	100.0% (381/381)	381	99.7% (380/381)
103 (ac102, p12)	121	121	100% (121/121)	121	100% (121/121)	121	99.2% (120/121)	121	100% (103/103)	121	99.2% (120/121)	121	100.0% (121/121)
104 (ac103, p45)	389	389	100% (389/389)	389	99.2% (386/389)	389	99.0% (385/389)	434	99.7% (347/348)	389	98.7% (384/389)	389	99.5% (387/389)
105 (ac104, vp80)	964	973	98.9% (962/973)	938	95.5% (923/966)	957	95.8% (914/954)	958	98.0% (948/967)	962	96.4% (938/973)	957	98.0% (945/964)
106 (ac110)	56	56	100% (56/56)	56	100% (56/56)	56	100% (56/56)	56	100.0% (56/56)	56	98.2% (55/56)	56	100.0% (56/56)
107 (ac109, odv-ec43)	366	366	100% (366/366)	366	99.7% (365/366)	366	99.7% (365/366)	366	99.7% (365/366)	366	99.7% (365/366)	366	99.7% (365/366)
108 (ac108)	97	97	100% (97/97)	97	97.9% (95/97)	96	95.9% (93/97)	96	97.9% (95/97)	96	92.8% (90/97)	97	100.0% (97/97)

Table 3 (*continued*)

5-6		Ab-a624		27		2161		3029		3041		3054	
Number (name)	Size (aa)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)	Size (aa)	% ID (range)
109 (ac112/113; apsup)	336	333	98.5% (328/333)	333	98.5% (326/331)	445	97.6% (325/333)	334	98.8% (326/330)	333	97.9% (324/331)	334	98.5% (325/330)
110 (ac24, pkip-1)	179	179	100% (179/179)	179	99.4% (178/179)	179	99.4% (178/179)	179	99.4% (178/179)	179	99.4% (178/179)	179	99.4% (178/179)
111	100	100	99.0% (99/100)	100	99.0% (99/100)	100	98.0% (98/100)	100	99.0% (99/100)	100	98.0% (98/100)	100	100.0% (100/100)
116 (dutpase)	149	149	99.3% (148/149)	149	99.3% (148/149)	149	100% (149/149)	149	100% (149/149)	149	99.3% (148/149)	149	99.3% (148/149)
117 (ac63)	154	154	100% (154/154)	154	98.1% (151/154)	154	96.1% (148/154)	154	97.4% (150/154)	154	95.5% (147/154)	154	98.7% (152/154)
118 (ac20/21, arif-1)	269	269	100% (269/269)	265	97.4% (262/269)	266	94.4% (255/270)	309	97.1% (237/244)	275	91.7% (221/241)	269	99.6% (268/269)
119 (ac22, pif-2)	407	407	99.3% (404/407)	407	99.5% (405/407)	407	99.5% (405/407)	407	99.8% (406/407)	407	99.3% (404/407)	407	100.0% (407/407)
120 (rr2b)	348	341	98.5% (331/336)	348	99.4% (346/348)	359	99.4% (335/337)	352	98.8% (334/338)	359	96.1% (345/359)	348	99.7% (347/348)
121	78	78	98.7% (77/78)					89	89.8% (44/49)			78	88.5% (69/78)
122 (ac13, 38.7k)	200	199	97.5% (196/201)	206	96.6% (200/207)	202	97.0% (196/202)	195	94.5% (190/201)	203	98.0% (199/203)	203	98.5% (200/203)
123 (ac14, lef-1)	234	234	100% (234/234)	234	100% (234/234)	234	99.6% (233/234)	234	100% (234/234)	234	99.6% (233/234)	234	100.0% (234/234)
124	133	133	100% (133/133)	132	97.7% (130/133)	132	98.5% (131/133)	132	98.5% (131/133)	134	97.8% (131/134)	133	99.2% (132/133)
125 (ac15, egt)	560	560	100% (560/560)	560	99.8% (559/560)	561	98.9% (555/561)	560	99.8% (559/560)	561	98.4% (552/561)	560	99.8% (559/560)
126	55	55	98.2% (54/55)	55	100% (55/55)	55	98.2% (54/55)	55	100% (55/55)	55	96.4% (53/55)	55	100.0% (55/55)
127	192	192	100% (192/192)	194	97.4% (189/194)	190	98.4% (182/185)	199	92.0% (184/200)	193	98.4% (190/193)	192	98.4% (188/191)
128	226	226	100% (226/226)	226	99.6% (225/226)	226	99.1% (224/226)	226	99.1% (224/226)	226	97.8% (221/226)	226	99.6% (225/226)
129	884	876	98.1% (867/884)	863	96.4% (855/887)	888	95.4% (857/898)	876	97.3% (860/884)	866	94.8% (840/886)	859	96.7% (855/884)
130 (ac23, F protein)	676	676	99.9% (675/676)	676	99.1% (670/676)	676	99.0% (669/676)	675	99.3% (671/676)	675	99.0% (669/676)	676	99.4% (672/676)
131 (ac46, odv-e66)	665	654	92.5% (605/654)	654	100% (566/566)	654	92.4% (604/654)	654	99.8% (565/566)	654	99.8% (565/566)	654	100.0% (566/566)
132	81	81	100% (81/81)	88	90.9% (80/88)	87	69% (60/87)	80	96.3% (78/81)	79	88.9% (72/81)	94	77.0% (67/87)
(ac129, p24)	223	222	99.1% (221/223)	223	99.1% (221/223)	223	98.7% (220/223)	223	100% (198/198)	223	98.7% (220/223)	223	99.1% (221/223)
135	123	123	100% (123/123)	124	99.2% (123/124)	123	100% (123/123)	123	100.0% (123/123)	123	99.2% (122/123)	123	100.0% (123/123)
136 (ac131, pep)	313	313	99.7% (312/313)	314	99.0% (311/314)	313	99.7% (312/313)	314	99.7% (313/314)	313	99.7% (312/313)	314	99.4% (312/314)
137 (ac6, lef-2)	216	216	100% (216/216)	218	95.4% (208/218)	219	95.9% (210/219)	215	96.8% (209/216)	216	96.8% (209/216)	218	94.0% (205/218)
138	291	291	99.7% (290/291)	291	98.3% (286/291)	291	97.6% (284/291)	291	98.3% (286/291)	291	97.6% (284/291)	291	98.3% (286/291)
139 (iap-3)	155	156	98.1% (153/156)	156	99.4% (155/156)	156	96.8% (151/156)	155	99.4% (154/155)	155	99.4% (154/155)	156	90.4% (141/156)
140 (ac106-107)	246	246	95.6% (239/250)	251	96.8% (244/252)	245	97.2% (239/246)	245	99.2% (244/246)	243	98.4% (242/246)	247	89.9% (223/248)
141	542	542	99.8% (541/542)	542	99.3% (538/542)	542	98.7% (535/542)	542	98.9% (536/542)	542	98.0% (531/542)	542	99.1% (537/542)
142	118	129	100% (116/116)	127	98.3% (114/116)	127	98.1% (106/108)	130	99.0% (100/101)	127	98.3% (114/116)	127	98.0% (99/101)
143 (ac115, pif-3)	203	203	100% (203/203)	203	99.0% (201/203)	203	99.0% (201/203)	203	100.0% (203/203)	202	98.0% (196/200)	209	98.5% (200/203)
144	113	113	100% (113/113)	113	99.1% (112/113)	113	98.2% (111/113)	116	100% (86/86)	113	97.3% (110/113)	143	99.1% (112/113)
145 (ac31, sod)	154	154	100% (154/154)	154	99.4% (153/154)	154	98.1% (151/154)	195	98.1% (151/154)	154	98.7% (152/154)	208	99.3% (144/145)
147 (rr2a)	359	359	100% (359/359)	359	98.3% (354/360)	359	96.7% (348/360)	358	98.3% (353/359)	359	97.5% (351/360)	359	97.5% (351/360)
148 (rr1)	596	596	99.7% (594/596)	596	99.0% (590/596)	595	97.3% (580/596)	596	99.0% (590/596)	595	97.3% (580/596)	596	99.2% (591/596)
149 (ac3, ctl-1)	53	53	100% (53/53)	53	100% (53/53)	53	100% (53/53)	75	100% (53/53)	53	(53/53) 100.0%	53	100% (53/53)
151 (ac12)	168	168	100% (168/168)	168	98.8% (166/168)	168	97.6% (164/168)	168	95.2% (160/168)	352	97.6% (160/164)	169	98.2% (162/165)
152	250	250	99.6% (249/250)	269	98.4% (246/250)	187	98.4% (182/185)	250	99.2% (248/250)	250	96.0% (240/250)	269	98.4% (246/250)
155 (ac119, pif-1)	530	530	98.9% (524/530)	533	96.8% (516/533)	534	94.8% (506/534)	530	98.5% (522/530)	533	95.1% (507/533)	533	96.6% (515/533)
156 (ac32, fgf)	285	285	99.6% (284/285)	285	99.3% (283/285)	285	98.2% (280/285)	343	99.0% (202/204)	284	97.9% (279/285)	285	99.6% (270/271)
157 (ac133, alk-exo)	420	419	98.8% (415/420)	420	100% (420/420)	419	99.3% (417/420)	419	99.0% (416/420)	420	99.5% (418/420)	419	99.8% (419/420)
158 (ac18)	373	373	100% (373/373)	373	98.7% (368/373)	373	98.7% (368/373)	373	98.9% (369/373)	373	98.7% (368/373)	373	98.7% (368/373)
159 (ac19)	118	118	100% (118/118)	118	100% (118/118)	118	100% (118/118)	160	99.2% (117/118)	118	100.0% (118/118)	163	118/118(100%)
160 (vef-2)	788	514/450	99.7% (308/309); 98.0% (441/450)	733	92.9% (456/491)	787	95.6% (753/788)	786	97.6% (730/748)	788	94.7% (747/789)	788	96.7% (723/748)
162 (ac12)	91	134	100% (91/91)	90	93.4% (85/91)	89	100% (68/68)	91	100% (62/62)	87	100.0% (63/63)	90	93.4% (85/91)
163	329	327	95.7% (315/329)	327	93.9% (309/329)	327	92.4% (304/329)	327	96.0% (316/329)	327	92.1% (303/329)	325	90.6% (298/329)

isolates mostly from Europe and North America (the rest of the isolates in this analysis) could be discerned, which was consistent with a previous phylogenetic analysis of relationships among LdMNPV isolates based on alignments of partial *lef-8* nucleotide sequences (Harrison et al., 2014). However, while the previously published *lef-8* phylogeny grouped isolates 3029 (Virin-ENSh) and 3054 (Spain) with the Asian isolates, the phylogeny based on all the core gene nucleotide sequences placed these isolates in a clade with the North American isolates Ab-a624 and 5-6. Isolate 27 from Western Siberia occurred in the same node as 3054.

3.4. Pathogenicity against strains of European and Asian gypsy moth

To assess the relative pathogenicities of different LdMNPV isolates against different populations of *L. dispar*, droplet-feeding bioassays were carried out against larvae from colonies derived from six different populations of *L. dispar dispar*, *L. dispar asiatica*, and *L. dispar japonica*. Overall, the LC_{50} values obtained in this study did not exhibit a large degree of variation relative to the LC_{50} values obtained in our previous study with New Jersey Standard Strain *L. dispar* larvae, even though the bioassays were carried out in a different location (Harrison et al., 2014). LdMNPV isolates 2161 (Korea), 3029 (VirinENSh/Russia), and 3041 (Japan) killed larvae with LC_{50} concentrations that ranged from 2.2- to 6-fold lower than the Gypchek stock 3049 (statistically significant at $p < 0.05$; Table 5). Significantly lower LC_{50}s for the Gypchek stock were observed with larvae from all six colonies, including the *L. dispar dispar* colonies LJ (Lithuania), UC (Connecticut), and KG (Greece). Statistically significant differences were not observed between the LC_{50} values of 2161 (Korea) and 3041 (Japan), while 3029 (VirinENSh/Russia) exhibited LC_{50}s against the LJ (Lithuania) and RB (Siberia) strains that were moderately but significantly lower ($p < 0.05$) than those of 2161 (Korea) and 3041 (Japan) isolates. Notably, isolate 3041 (Japan) did not appear to be impaired in its activity against any host strain, despite not possessing a *vef-1* gene. This result is consistent with recently published data from bioassays with an LdMNPV *vef-1* knockout mutant (Hoover et al., 2010).

4. Discussion

ORF annotation in baculovirus genomes generally involves a combination of ORF scanning and homology searches to identify ORFs that (a) are of a size longer than would be expected in a random DNA sequence and (b) are evolutionarily conserved. With the sequencing of the first baculovirus genome, it was also assumed that ORFs corresponding to real genes are distributed in a mostly non-overlapping fashion, and consequently only ORFs exhibiting a minimal degree of overlap were selected for annotation (though ORFs that are conserved in other baculoviruses are generally exempt from this criterion; Ayres et al., 1994; Possee and Rohrmann, 1997). The nucleotide distributions of LdMNPV genome sequences pose potential issues for these ORF annotation criteria. The genomes of the LdMNPV isolates analyzed in this study range from 57.25% to 57.47% G + C, and are the most GC-rich of any baculovirus. Because stop codons (TAG, TAA, and TGA) are GC-poor, they are expected to occur less frequently by chance in GC-rich sequences, which in turn lead to longer ORFs. This trend has been observed in an analysis of genome sequences that revealed that the longest ORFs are observed in the most GC-rich sequences (Oliver and Marin, 1996). During our examination of LdMNPV genome sequences, we found ORFs that were very large but did not match other annotated baculovirus ORFs. These ORFs, which also were predicted to be genes by algorithms such as fgenesV (http://linux1.softberry.com/berry.phtml) and ZCURVE_V (Guo and Zhang, 2006), exhibited significant degrees of overlap with the ORFs of other well-characterized baculovirus genes. For example, isolates Ab-a624, 3029, and 3054 include a 528-codon ORF that complete overlaps the ORF for *lef-9*, but on the opposite strand. These isolates and 3041 also contain an antisense 434-codon ORF that almost entirely overlaps both the smaller *p12* and *p45* ORFs, and a 271- to 278-codon antisense ORF that completely overlaps the *arif-1* ORF.

A comprehensive transcriptomic analysis of AcMNPV-infected *Trichoplusia ni* cells in culture revealed the presence of antisense transcripts that overlapped annotated ORFs (Chen et al., 2013). Only three of these antisense transcripts contained ORFs, and these ORFs were shorter than the sense-strand annotated ORFs. AcMNPV-C6 has a G + C content of 40.7%, which may account for the low frequency and relatively short length of antisense ORFs. We nevertheless took a conservative approach to ORF annotation and did not annotate large antisense ORFs like those described above. The question remains whether these ORFs and other ORFs in LdMNPV isolates that are not conserved among other baculoviruses are actually transcribed and translated.

LdMNPV isolates contain the largest copy numbers of *bro* gene family ORFs that have been recorded for any baculovirus. Comparisons and database searches with the LdMNPV *bro* nucleotide and protein sequences suggest that a considerable degree of intra- and intergenomic recombination has taken place within *bro* gene sequences (Harrison et al., 2014). Copies of *bro* genes appear to transverse the endpoints of the inversion of the genomic region bound by the *sod* and *dutpase* genes that distinguishes the LyxyMNPV-5 and LdMNPV genomes (Nai et al., 2010) (Supplementary Figure 1). No similar redistribution of ORFs bound by *bro* sequences appears to have occurred in any of the LdMNPV genome sequences that have been determined. While experimental data on expression and possible functions of *bro* genes have been generated for *bro* genes of Bombyx mori nucleopolyhedrovirus (BmNPV) and Spodoptera litura multiple nucleopolyhedrovirus (Gong et al., 2003; Kang et al., 2006, 1999, 2003), many baculovirus genomes contain no *bro* genes. A few of the LdMNPV *bro* genes are relatively well-conserved (for example, isolate 5-6 ORF72/*bro-d*), but most of the LdMNPV *bro* genes exhibit a low degree of sequence conservation, which suggests a low or nonexistent degree of functional conservation among the encoded BRO proteins. It is thus unclear what role, if any, the LdMNPV *bro* genes have in baculovirus replication and pathogenesis.

Phylogenetic inference with core gene nucleotide sequences supports a separation of LdMNPV isolates into two groups containing isolates from Asian host populations or isolates from North American/European host populations (Fig. 3). LdMNPV-27, from Western Siberia, and LdMNPV-3029, from an unspecified part of Russia, were grouped with the North American and European isolates in this analysis. Although *L. dispar asiatica* is described as occurring east of the Ural Mountains, recent analyses of subspecies distribution using cytochrome *c* oxidase I DNA barcodes grouped gypsy moths from Krasnoyarsk and Khakassia in Siberia not with *L. dispar asiatica*, but rather with *L. dispar dispar* samples from Europe (deWaard et al., 2010). Other genetic analyses, as well as the observation of female flight capability in the LJ strain of *L. dispar dispar*, further suggest that a significant degree of gene flow has taken place between Siberian populations of *L. dispar* and populations in adjacent regions of northeastern Europe (Chen et al., 2016; Keena et al., 2008). It is conceivable, therefore, that isolates 27 and 3029 actually originated from an *L. dispar dispar* host population in Europe and were transmitted to *L. dispar asiatica* populations in Siberia. Such an origin would be consistent with their placement among other isolates from North America and Europe.

This study was motivated in part by the desire to see if there are LdMNPV isolates that might perform better as biopesticides to control Asian gypsy moth than Gypchek. One hypothesis is that

Table 4
Additional LdMNPV ORFs not annotated in LdMNPV 5-6.

ORF[a]	Distribution among LdMNPV isolates[b]							Homologues
	5-6	Ab-a624	27	2161	3029	3041	3054	
3029 ORF13	–	–	–	+ (ORF10a)	+ (ORF13)	+ (ORF13)	–	LyxyNPV-5 ORF11
3029 ORF21	–	–	–	–	+ (ORF21)	–	–	LyxyNPV-5 ORF19
Ab-a624 ORF 27	+ (ORF26a)	+ (ORF27)	+ (ORF21a)	+ (ORF22a)	+ (ORF26a)	+ (ORF26)	+ (ORF24)	
2161 (ORF34)	–	+ (ORF40)	–	+ (ORF34)	+ (ORF40)	+ (ORF38)	+ (ORF39)	
27 ORF47	+ (ORF48a)	+ (ORF51)	+ (ORF47)	+ (ORF44a)	+ (ORF49a)	+ (ORF49)	+ (ORF48)	LyxyNPV-5 ORF45
2161 ORF 52	+ (ORF55a)	+ (ORF59)	+ (ORF54)	+ (ORF52)	+ (ORF57)	+ (ORF57)	+ (ORF56)	Homologues in multiple NPVs
3029 ORF 60	+ (ORF57a)	+ (ORF62)	–	–	+ (ORF60)	+ (ORF60)	+ (ORF57)	
2161 ORF64	–	–	+ (ORF65a)	+ (ORF64)	–	+ (ORF70)	+ (ORF69)	
Ab-a624 ORF 73	+ (ORF68a)	+ (ORF73)	+ (ORF66a)	+ (ORF65a)	+ (ORF70a)	+ (ORF72)	+ (ORF71)	LyxyNPV-5 ORF65
3029 ORF 73	+ (ORF70a)	+ (ORF75)	+ (ORF68a)	+ (ORF67a)	+ (ORF73)	+ (ORF75)	+ (ORF74)	
2161 ORF74	+ (ORF74a)	+ (ORF80)	+ (ORF72a)	+ (ORF74)	+ (ORF78a)	+ (ORF82)	+ (ORF80)	
3029 ORF118	+ (ORF116a)	+ (ORF121)	+ (ORF113a)	–	+ (ORF118)	–	+ (ORF122)	
2161 ORF124	+ (ORF121a)	+ (ORF128)	+ (ORF118)	+ (ORF124)	+ (ORF124a)	+ (ORF128)	+ (ORF128)	LyxyNPV-5 ORF136
2161 ORF123	–	+ (ORF126)	–	+ (ORF123)	+ (ORF123)	+ (ORF127)	–	
2161 ORF135	–	–	+ (ORF128a)	+ (ORF135)	–	+ (ORF139)	+ (ORF139)	Cell division protein DamX domain (PRK10905)
2161 ORF137	–	+ (ORF140)	+ (ORF130)	+ (ORF137)	+ (ORF136)	+ (ORF141)	+ (ORF141)	
2161 ORF141	+ (ORF136a)	+ (ORF144)	+ (ORF134)	+ (ORF141)	+ (ORF140)	+ (ORF145)	+ (ORF145)	LyxyNPV-5 ORF119 (9.7 kDa protein)
2161 ORF144	–	+ (ORF147)	–	+ (ORF144)	–	–	–	
2161 ORF151	–	–	+ (ORF142a)	+ (ORF151)	+ (ORF148a)	+ (ORF154)	–	
2161 ORF157	–	–	–	+ (ORF157)	–	+ (ORF160)	–	
2161 ORF160	+ (ORF151a)	+ (ORF161)	+ (ORF149a)	+ (ORF160)	+ (ORF155a)	+ (ORF162)	+ (ORF161)	
2161 ORF162	+ (ORF152a)	+ (ORF163)	+ (ORF151a)	+ (ORF162)	+ (ORF156a)	+ (ORF164)	+ (ORF163)	
2161 ORF166	+ (ORF155a)	+ (ORF167)	+ (ORF155)	+ (ORF166)	+ (ORF160)	+ (ORF170)	+ (ORF167)	Homologues in multiple NPVs

[a] Isolate and ORF number are indicated.
[b] ORF numbers followed by an "a" are ORFs that were not annotated in the original publication of the genome sequences of LdMNPV isolates 5-6, 27, 2161, and 3029.

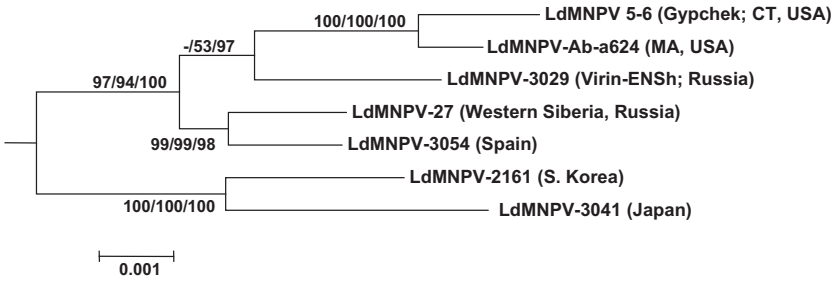

Fig. 3. Phylogenetic analysis of concatenated nucleotide sequence alignments of 37 baculovirus core genes (Garavaglia et al., 2012) showing relationships of the LdMNPV isolates listed in Table 1. LyxyNPV-5 (not shown) was used as an outgroup. Bootstrap values for each node are shown when the node occurred in trees inferred by minimum evolution (ME), maximum parsimony (MP), and maximum likelihood (ML) (ME/ML/MP).

Table 5
Concentration-mortality response ($LC_{50} \times 10^4$ in OBs/mL, with 95% confidence limits) of neonate *L. dispar* larvae infected with LdMNPV isolates.

LdMNPV isolates	Strains of *L. dispar*[a]					
	L. dispar dispar			*L. dispar asiatica*		*L. dispar japonica*
	UC (Connecticut, USA)	KG (Greece)	LJ (Lithuania)	RB (Siberia, Russia)	RM (Far East Russia)	JN (Japan)
3049 (Gypchek)	28.0 (21.54–38.4) ***a***	42.8 (20.92–177.6) ***a***	14.9 (11.55–19.71) ***a***	16.1 (9.59–30.81) ***a***	13.3 (10.15–17.83) ***a***	9.5 (5.03–20.4) ***a***
2161 (Korea)	11.1 (8.33–15.64) ***b***	7.6 (5.74–10.4) ***b***	4.8 (3.62–6.49) ***b***	4.88 (2.84–8.9) ***b***	4.3 (3.26–5.69) ***b***	3.7 (2.85–4.87) ***b***
3029 (Russia/Virin ENSh)	8.69 (6.5–12.21) ***b***	7.1 (3.98–16.0) ***b***	2.4 (1.36–4.2) ***c***	2.18 (1.74–2.7) ***c***	3.2 (2.33–4.37) ***b***	3.3 (2.51–4.42) ***b***
3041 (Japan)	7.0 (4.14–13.69) ***b***	8.5 (4.53–21.3) ***b***	3.4 (2.59–4.59) ***bc***	3.53 (2.0–6.28) ***b***	2.8 (2.04–3.86) ***b***	4.4 (3.27–5.97) ***b***

[a] For each *L. dispar* strain, LC_{50} values with different letters are significantly different as assessed by comparison of 95% confidence levels of lethal dose ratios (Robertson et al., 2007). Slopes and intercepts were not significantly different among virus treatments for any of the *L. dispar* strains.

LdMNPV isolates from Asian gypsy moth might be more pathogenic against Asian gypsy moth larvae than LdMNPV isolates from European gypsy moth (e.g. Gypchek). In a study of isolates of Spodoptera frugiperda multiple nucleopolyhedrovirus (SfMNPV), Barrera et al. (2011) proposed that baculovirus isolates are selected to maintain a high degree of infectivity towards local host populations relative to geographically distant host populations. Not all published work supports this hypothesis (e.g. Ogembo et al., 2005). Similarly, prior published bioassay data of LdMNPV isolates against strains of Asian gypsy moth have differed in the trends reported for Gypchek and Gypchek-derived virus isolates. Ebling et al. (2004) found that LdMNPV isolates from China and Japan performed better in bioassays against a Russian *L. dispar* strain than a sample of Disparvirus, a Canadian product derived from Gypchek. Duan et al. (2011) found that Disparvirus and a Chinese LdMNPV isolate, LdMNPV-H, performed comparably against a Chinese gypsy moth strain in the laboratory. The same group later found that LdMNPV-H exhibited greater pathogenicity than Disparvirus against Chinese gypsy moth populations in the field (Duan et al., 2012). In contrast, Podgwaite and coworkers (Bakhvalov et al., 2005; Podgwaite et al., 2006) found little difference between the activities of Gypchek and Virin ENSh against Western Siberian strains of gypsy moth. The same group also reported that an LdMNPV isolate from a Western Siberian host population killed Western Siberian gypsy moth larvae with LC_{50} values that were equal to or lower than LC_{50}s obtained with plaque isolates derived from Gypchek (Podgwaite et al., 2013). The results from our bioassays were consistent with those studies showing that LDP-67-derived LdMNPV virus preparations were significantly less pathogenic against Asian gypsy moth larvae compared to LdMNPV isolates from other sources. However, our Gypchek sample, LdMNPV-3049, also killed *L. dispar dispar* larvae from Connecticut, Greece, and Lithuania with significantly higher LC_{50}s compared to the other isolates from Korea (2161), Russia (3029), and Japan (3041). This observation raises the possibility that LdMNPV-3049 was inherently less pathogenic than other samples of Gypchek, perhaps due to the occurrence of mutations in genes that influence pathogenicity (Zhang et al., 2010). To address this possibility, we carried out droplet-feeding bioassays of 3049 and a sample of Gypchek from Sylvar Technologies Inc. against *L. dispar dispar* New Jersey Standard Strain larvae, but found no consistent differences in pathogenicity between these two Gypchek samples (data not shown). No significant differences in LC_{50} were observed for the LdMNPV isolates 2161 (Korea), 3029 (Russia), or 3041 (Japan) against gypsy moth colonies JN (Japan), RM (Far East Russia), KG (Greece), or UC (Connecticut), while isolate 3029 (Russia) exhibited significantly lower LC_{50} values against colonies RB (Siberia) and LJ (Lithuania). Collectively, these trends do not wholly support a correlation between pathogenicity of virus isolates and their geographic origin relative to the geographic origin or subspecies of the host population. A study of geographic isolates of the gypsy moth fungal pathogen *Entomophaga maimaiga* using some of the same gypsy moth colonies also found no correlation between pathogenicity and virulence against the colonies and the geographic origin of the isolates (Nielsen et al., 2005). However, it is conceivable that some degree of adaptation and selection has taken place within the colonies since their establishment that may have affected susceptibility to viral or fungal infection.

This study presents a more detailed view of the relationships among isolates of LdMNPV. It also provides additional data towards the evaluation of the use of LdMNPV-based pesticides against the gypsy moth, especially the Asian gypsy moth, a particularly dire threat to trees and forests in North America. Additional LdMNPV genomes are currently being determined to refine our picture of the relationships of LdMNPV isolates. In addition, more bioassays with additional isolates and host strains are planned to confirm the trends observed in this study and extend our knowledge of the susceptibilies of different populations of gypsy moth to different geographic isolates of LdMNPV.

Disclosures

The authors report no conflicts of interest to be declared.

Acknowledgments

We wish to thank Daniel Barakh and Pallavi Thapa (USDA-ARS, Beltsville, MD) and Paul Moore (USDA-FS, Ansonia, CT) for assistance with bioassays, and Renée Lapointe (Northern Forestry Centre, Natural Resources Canada) for providing a sample of Gypchek produced by Sylvar Technologies Inc. This work was supported by the U.S. Department of Agriculture. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jip.2016.03.014.

References

Ahrens, C.H., Russell, R.L., Funk, C.J., Evans, J.T., Harwood, S.H., Rohrmann, G.F., 1997. The sequence of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome. Virology 229, 381–399.

Alyoshina, O.A., 1980. Study of entomopathogenic viruses in the USSR. In: Ignoffo, C. M., Martignoni, M.E., Vaughn, J.L. (Eds.), Characterization, Production and Utilization of Entomopathogenic Viruses, Proceedings of the Second Conference of Project V, Microbiological Control of Insect Pests of the US/USSR. Clearwater Beach, Florida, USA, pp. 1–16.

Ayres, M.D., Howard, S.C., Kuzio, J., Lopez-Ferber, M., Possee, R.D., 1994. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. Virology 202, 586–605.

Bakhvalov, S.A., Martemyanov, V.V., Podgwaite, J.D., 2005. Comparative biological activities of two nucleopolyhedrovirus preparations: Virin NSH and Gypchek. Euroas. Entomol. J. 4, 183–186.

Barrera, G., Simón, O., Villamizar, L., Williams, T., Caballero, P., 2011. Spodoptera frugiperda multiple nucleopolyhedrovirus as a potential biological insecticide: genetic and phenotypic comparison of field isolates from Colombia. Biol. Control 58, 113–120.

Chen, Y.R., Zhong, S., Fei, Z., Hashimoto, Y., Xiang, J.Z., Zhang, S., Blissard, G.W., 2013. The transcriptome of the baculovirus *Autographa californica* multiple nucleopolyhedrovirus in Trichoplusia ni cells. J. Virol. 87, 6391–6405.

Chen, F., Luo, Y., Keena, M.A., Wu, Y., Wu, P., Shi, J., 2016. DNA barcoding of gypsy moths from China (Lepidoptera, Erebidae) reveals new haplotypes and divergence patterns within gypsy moth subspecies. J. Econ. Entomol. 109, 366–374.

Darling, A.C., Mau, B., Blattner, F.R., Perna, N.T., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14, 1394–1403.

deWaard, J.R., Mitchell, A., Keena, M.A., Gopurenko, D., Boykin, L.M., Armstrong, K.F., Pogue, M.G., Lima, J., Floyd, R., Hanner, R.H., Humble, L.M., 2010. Towards a global barcode library for *Lymantria* (Lepidoptera: Lymantriinae) tussock moths of biosecurity concern. PLoS ONE 5, e14280.

Duan, L.Q., Otvos, I.S., Xu, L.B., Conder, N., Wang, Y., 2011. Comparison of the activities of three LdMNPV isolates in the laboratory against the Chinese strain of Asian gypsy moth. Open Entomol. J. 5, 24–30.

Duan, L.Q., Otvos, I.S., Xu, L.B., Conder, N., Wang, Y., 2012. Field testing Chinese and Japanese gypsy moth nucleopolyhedrovirus and disparvirus against a Chinese population of Lymantria dispar asiatica in Huhhot, Inner Mongolia, People's Republic of China. J. Econ. Entomol. 105, 344–353.

Ebling, P.M., Otvos, I.S., Conder, N., 2004. Comparative activity of three isolates of LdMNPV against two strains of *Lymantria dispar*. Can. Entomol. 136, 737–747.

Fuxa, J.R., 2004. Ecology of insect nucleopolyhedroviruses. Agric. Ecosyst. Environ. 103, 27–43.

Garavaglia, M.J., Miele, S.A., Iserte, J.A., Belaich, M.N., Ghiringhelli, P.D., 2012. The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family *Baculoviridae*. J. Virol. 86, 12069–12079.

Gong, Y., Li, Z., Wang, L., Pan, L., Yang, K., Pang, Y., 2003. Characterization of bro-b gene of Spodoptera litura multicapsid nucleopolyhedrovirus. Virus Genes 27, 115–123.

Guo, F.-B., Zhang, C.-T., 2006. ZCURVE_V: a new self-training system for recognizing protein-coding genes in viral and phage genomes. BMC Bioinf. 7, 9.

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

Harrison, R.L., Hoover, K., 2012. Baculoviruses and other occluded insect viruses. In: Vega, F.E., Kaya, H.K. (Eds.), Insect Pathology, second ed. Academic Press, Boston, pp. 73–131.

Harrison, R.L., Lynn, D.E., 2007. Genomic sequence analysis of a nucleopolyhedrovirus isolated from the diamondback moth, *Plutella xylostella*. Virus Genes 35, 857–873.

Harrison, R.L., Rowley, D.L., 2015. Complete genome sequence of the strain of *Lymantria dispar multiple nucleopolyhedrovirus* found in the gypsy moth biopesticide Virin-ENSh. Gen. Announ. 3, e01407-14.

Harrison, R.L., Keena, M.A., Rowley, D.L., 2014. Classification, genetic variation and pathogenicity of Lymantria dispar nucleopolyhedrovirus isolates from Asia, Europe, and North America. J. Invertebr. Pathol. 116, 27–35.

Herniou, E.A., Arif, B.M., Becnel, J.J., Blissard, G.W., Bonning, B., Harrison, R.L., Jehle, J. A., Theilmann, D.A., Vlak, J.M., 2012. Baculoviridae. In: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, Oxford, pp. 163–174.

Hoover, K., Humphries, M.A., Gendron, A.R., Slavicek, J.M., 2010. Impact of viral enhancin genes on potency of Lymantria dispar multiple nucleopolyhedrovirus in *L. dispar* following disruption of the peritrophic matrix. J. Invertebr. Pathol. 104, 150–152.

Hughes, P.R., Wood, H.A., 1981. A synchronous peroral technique for the bioassay of insect viruses. J. Invertebr. Pathol. 37, 154–159.

Kabilov, M.R., Martemyanov, V.V., Tupikin, A.E., Baturina, O.A., Belousova, I.A., Bondar, A.A., Ilyinykh, A.V., 2015. Complete genome sequence of a Western Siberian Lymantria dispar multiple nucleopolyhedrovirus isolate. Gen. Announ. 3, e00335-15.

Kang, W., Suzuki, M., Zemskov, E., Okano, K., Maeda, S., 1999. Characterization of baculovirus repeated open reading frames (bro) in Bombyx mori nucleopolyhedrovirus. J. Virol. 73, 10339–10345.

Kang, W.K., Imai, N., Suzuki, M., Iwanaga, M., Matsumoto, S., Zemskov, E.A., 2003. Interaction of Bombyx mori nucleopolyhedrovirus BRO-A and host cell protein laminin. Arch. Virol. 148, 99–113.

Kang, W., Kurihara, M., Matsumoto, S., 2006. The BRO proteins of Bombyx mori nucleopolyhedrovirus are nucleocytoplasmic shuttling proteins that utilize the CRM1-mediated nuclear export pathway. Virology 350, 184–191.

Keena, M.A., Cote, M.J., Grinberg, P.S., Wallner, W.E., 2008. World distribution of female flight and genetic variation in *Lymantria dispar* (Lepidoptera: Lymantriidae). Environ. Entomol. 37, 636–649.

Kuzio, J., Pearson, M.N., Harwood, S.H., Funk, C.J., Evans, J.T., Slavicek, J.M., Rohrmann, G.F., 1999. Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria dispar*. Virology 253, 17–34.

Lewis, F.B., Wallner, W.E., Rollinson, W.D., 1984. Activity of lymantriid NPVs from the People's Republic of China against North American *Lymantria dispar*. Entomophaga 29, 299–302.

Lynn, D.E., Shapiro, M., Dougherty, E.M., 1993. Selection and screening of clonal isolates of the Abington strain of gypsy moth nuclear polyhedrosis virus. J. Invertebr. Pathol. 62, 191–195.

McClintock, J.T., Dougherty, E., Weiner, R.M., 1986. Protein synthesis in gypsy moth cells infected with a nuclear polyhedrosis virus of *Lymantria dispar*. Virus Res. 5, 307–322.

Moscardi, F., 1999. Assessment of the applications of baculoviruses for control of Lepidoptera. Ann. Rev. Entomol. 44, 257–289.

Nai, Y.S., Wu, C.Y., Wang, T.C., Chen, Y.R., Lau, W.H., Lo, C.F., Tsai, M.F., Wang, C.H., 2010. Genomic sequencing and analyses of Lymantria xylina multiple nucleopolyhedrovirus. BMC Genom. 11, 116.

Narang, N., Herard, F., Dougherty, E.M., Chen, K., Vega, F.E., 2001. A gypsy moth (*Lymantria dispar*, Lepidoptera: Lymantriidae) multicapsid nuclear polyhedrosis virus from France: comparison with a North American and a Korean strain. Eur. J. Entomol. 98, 189–194.

Nielsen, C., Keena, M., Hajek, A.E., 2005. Virulence and fitness of the fungal pathogen *Entomophaga maimaiga* in its host *Lymantria dispar*, for pathogen and host strains originating from Asia, Europe, and North America. J. Invertebr. Pathol. 89, 232–242.

Odell, T.M., Keena, M.A., Willis, R.B., 1997. Dietary influence of iron formulation on the development of gypsy moth (Lepidoptera: Lymantriidae) in laboratory colonies. Ann. Entomol. Soc. Am. 90, 149–154.

Ogembo, J.G., Kunjeko, E.C., Sithanantham, S., 2005. A preliminary study on the pathogenicity of two isolates of nucleopolyhedroviruses infecting African bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae). Int. J. Trop. Insect Sci. 25, 218–222.

Oliver, J.L., Marin, A., 1996. A relationship between GC content and coding-sequence length. J. Mol. Evol. 43, 216–223.

Pemberton, R.W., Lee, J.H., Reed, D.K., Carlson, R.W., Han, H.Y., 1993. Natural enemies of the Asian gypsy moth (Lepidoptera: Lymantridae) in South Korea. Ann. Entomol. Soc. Am. 86, 423–440.

Podgwaite, J., Martemyanov, V., Bakhvalov, S., 2006. Pathogenicity of two nucleopolyhedrovirus products, Virin NSh and Gypchek, for Asian and North American gypsy moth larvae. In: Gottschalk, K.W. (Ed.), 17th U.S. Department of Agriculture Interagency Research Forum on Gypsy Moth and Other Invasive Species. USDA Forest Service, Annapolis, Maryland, p. 79.

Podgwaite, J.D., Martemyanov, V.V., Slavicek, J.M., Bakhvalova, S.A., Pavlushin, S.V., Hayes-Plazolles, N., Zerillo, R.T., 2013. Potency of nucleopolyhedrovirus genotypes for European and Asian gypsy moth (Lepidoptera: Lymantriidae). J. Entomol. Sci. 48, 332–344.

Pogue, M.G., Schaefer, P.W., 2007. A Review of Selected Species of Lymantria Hubner [1819] (Lepidoptera: Noctuidae: Lymantriinae) From Subtropical and Temperate Regions of Asia, Including the Descriptions of Three New Species, Some Potentially Invasive to North America. Forest Health Technology Enterprise Team.

Possee, R.D., Rohrmann, G.F., 1997. Baculovirus genome organization and evolution. In: Miller, L.K. (Ed.), The Baculoviruses. Plenum, New York, pp. 109–140.

Reardon, R.C., Podgwaite, J., Zerillo, R., 2012. Gypchek-Environmentally Safe Insecticide for Gypsy Moth Control. Forest Health Technology Enterprise Team.
Robertson, J.L., Russell, R.M., Preisler, H.K., Savin, N.E., 2007. Bioassays With Arthropods, second ed. CRC Press, Boca Raton, FL.
Shapiro, M., Robertson, J.L., Injac, M.G., Katagiri, K., Bell, R.A., 1984. Comparative infectivities of gypsy moth (Lepidoptera: Lymantriidae) nucleopolyhedrosis virus isolates from North America, Europe, and Asia. J. Econ. Entomol. 77, 153–156.
Slack, J., Arif, B.M., 2007. The baculoviruses occlusion-derived virus: virion structure and function. Adv. Virus Res. 69, 99–165.
Slavicek, J.M., Hayes-Plazolles, N., 2003. The Lymantria dispar nucleopolyhedrovirus contains the capsid-associated p24 protein gene. Virus Genes 26, 15–18.
Slavicek, J., Podgwaite, J., 1992. Analysis of Lymantria dispar nuclear polyhedrosis viruses isolated from GYPCHEK: purification of high potency LdNPV isolates. In: Gottschalk, K.W., Twery, M.J. (Eds.), U.S. Department of Agriculture Interagency Gypsy Moth Research Forum 1991, U.S. Department of Agriculture, Forest Service, vol. NE-167. Northeastern Forest Experiment Station, Annapolis, MD, pp. 47–48.
Slavicek, J.M., Podgwaite, J., Lanner-Herrera, C., 1992. Properties of two *Lymantria dispar* nuclear polyhedrosis virus isolates obtained from the microbial pesticide Gypchek. J. Invertebr. Pathol. 59, 142–148.
Slavicek, J.M., Hayes-Plazolles, N., Kelly, M.E., 1995. Rapid formation of few polyhedra mutants of *Lymantria dispar* multinucleocapsid nuclear polyhedrosis virus during serial passage in cell culture. Biol. Control 5, 251–261.
Solter, L.F., Hajek, A.E., 2009. Control of gypsy moth, *Lymantria dispar*, in North America since 1878. In: Hajek, A.E., Glare, T.R., O'Callaghan, M. (Eds.), Use of Microbes for Control and Eradication of Invasive Arthropods. Springer, New York, pp. 181–212.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
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. Nucl. Acids Res. 22, 4673–4680.
USDA/APHIS/PPQ, 2015. APHIS Factsheet: Asian Gypsy Moth <https://www.aphis.usda.gov/publications/plant_health/content/printable_version/fs_phasiangm.pdf>.
van Oers, M.M., Vlak, J.M., 2007. Baculovirus genomics. Curr. Drug Targets 8, 1051–1068.
Zhang, J., Lapointe, R., Thumbi, D., Morin, B., Lucarotti, C.J., 2010. Molecular comparisons of alphabaculovirus-based products: Gypchek with Disparvirus (*Lymantria dispar*) and TM BioControl-1 with Virtuss (*Orgyia pseudotsugata*). Can. Entomol. 142, 546–556.