A comparison of the adaptations of strains of *Lymantria dispar* multiple nucleopolyhedrovirus to hosts from spatially isolated populations


Laboratory of Ecological Parasitology, Institute of Systematics and Ecology of Animals SB RAS, Frunze str. 11, 630091 Novosibirsk, Russia

A. Biological Institute, Irkutsk State University, Karl Marx str. 1, Irkutsk 664003, Russia

Genomics Core Facility, Institute of Chemical Biology and Fundamental Medicine SB RAS, Lavrentyev av. 8, 630090 Novosibirsk, Russia

Institute of Biology, Irkutsk State University, Karl Marx str. 1, Irkutsk 664003, Russia

USDA Forest Service, Northern Research Station, 359 Main Rd., Delaware, OH 43014, United States

USDA Forest Service, Northern Research Station, 51 Mill Pond Rd., Hamden, CT 06515, United States

Institute of Systematics and Ecology of Animals SB RAS, Frunze str. 11, 630091 Novosibirsk, Russia

Biological Institute, National Research Tomsk State University, Lenin av. 36, 634050 Tomsk, Russia

Corresponding author at: Laboratory of Ecological Parasitology, Institute of Systematics and Ecology of Animals SB RAS, Frunze str. 11, 630091 Novosibirsk, Russia.

⁎ martemyanov79@yahoo.com (V.V. Martemyanov), jpodgwaite@fs.fed.us (J.D. Podgwaite), puza.ia@mail.ru (I.A. Belousova)

E-mail addresses: martemyanov79@yahoo.com (V.V. Martemyanov), jpodgwaite@fs.fed.us (J.D. Podgwaite), puza.ia@mail.ru (I.A. Belousova), sergey-pavlushin@mail.ru (S.V. Pavlushin), jslavicek@fs.fed.us (J.M. Slavicek), baturina@niboch.nsc.ru (O.A. Baturina), kabilov@niboch.nsc.ru (M.R. Kabilov), avilyinykh@mail.ru (A.V. Ilyinykh).

Keywords: *Lymantria dispar*, nucleopolyhedrovirus, Potency, Phylogeography, Phylogenetic analysis, Local adaptation

ABSTRACT

The adaptation of pathogens to either their hosts or to environmental conditions is the focus of many current ecological studies. In this work we compared the ability of six spatially-distant Lymantria dispar (gypsy moth) nucleopolyhedrovirus (LdMNPV) strains (three from eastern North America and three from central Asia) to induce acute infection in gypsy moth larvae. We also sequenced the complete genome of one Asian (LdMNPV-27/0) and one North American (LdMNPV-45/0) strain which were used for bioassay. We found that all of the North American virus strains, with the exception of one, demonstrated higher potency than the Asian virus strains, either in North American (Lymantria dispar) larvae or, in Asian (Lymantria dispar asiatica) larvae. Complete genome sequencing revealed two gene deletions in the LdMNPV-27/0 strain: the virus enhancin factor gene (vef-1) and the baculovirus repeated orf gene (bro-p). These deletions were not seen in the LdMNPV-45/0 strain nor in other American strains available in archiving systems. We also found deletions of the bro-e and bro-o genes in LdMNPV-45/0 strain but not in the LdMNPV-27/0 strain. The phylogenetic inference with an alignment of the 37 core gene nucleotide sequences revealed the close relationship of the LdMNPV-45/0 strain with other American strains accessed in GenBank (Ab-a624 and 5-6) while the LdMNPV-27/0 strain was clustered together with the LdMNPV-3054 strain (isolated in Spain) instead of predicted clustering with LdMNPV-3029 (isolated in Asia). Our study demonstrated that first, different LdMNPV isolates from the same metapopulations of *L. dispar* exhibit little or no difference in the degree of virulence towards host larvae and second, that locality of host population is not an important driver of LdMNPV virulence. Virulence of LdMNPV is determined only by viral genetics. The genetic differences between North American and Central Asian virus strains are discussed.

1. Introduction

Gypsy moth, *Lymantria dispar* L. (Lepidoptera: Erebidae) nucleopolyhedrovirus (LdMNPV) is a member of the family Baculoviridae, genus Alphabaculovirus, and is specific for larval stages of the gypsy moth (Barber et al., 1993). The virus is periodically responsible for widespread natural-epizootics in *L. dispar* populations and is the basis for commercial products to control the pest (van Frankenhuyzen et al., 2007).

*L. dispar* is a widely distributed forest defoliator that inhabits temperate forests of Asia, Europe, and Northern Africa. At the end of nineteenth century the European strain of the insect (*L. dispar*) was introduced into the eastern part of North America where it continues to successfully adapt and produce regular outbreaks while expanding the area of infestation in a westerly direction (Liebhold et al., 1989). *L. dispers* is a very plastic species that possesses a high level of genetic variation between its populations (Bogdanowicz et al., 2000; de Waard et al., 2010; Wu et al., 2015). There are also significant physiological and behavioral differences between European (= North American) and Asian populations (*L. dispers asiatica*) of the insect; Asian...
female-moth flight ability is one of these and is of particular importance in the spread of the pest (Keena et al., 2007). Moreover, larvae in some L. dispar populations possess the ability to consume plant species that are not preferable to larvae in others (Matsuki et al., 2011). This may be due to variability in larval biochemistry between populations (Lindroth and Weisbrod, 1991). Thus, it might be hypothesized that the dramatic level of host intraspecific variation will likely lead to the divergence of LdMNPV strains that are tightly adapted to the host population in order to facilitate viral-reproductive success. The high variation of virulence of geographically distant LdMNPV isolates tested against the same host population (Duan et al., 2012) lends support to this concept. In spite of what may appear obvious (adaptation of a local parasite to its “native” population), there could be different scenarios of an evolving host–parasite system (Kawecki and Ebert, 2004) and environmental factors, such as host plant species of the herbivore, could effect the adaptation of the pathogen to its host (Cory and Myers, 2004). In the current study we hypothesize that geographically disconnected LdMNPV strains will possess less virulence to their spatially (and consequently, evolutionarily) distant host populations than to their native host population.

In an earlier publication (Martemyanov et al., 2015) we reported on potency and gene variations of two geographically distinct isolates of LdMNPV: LdMNPV-45/0 from Massachusetts, USA; and LdMNPV-27/0 from Tatarsk, Novosibirsk, Russia (Martemyanov et al., 2015). In that study, bioassays of these two LdMNPV isolates revealed large differences in their virulence towards two populations of European and Asian gypsy moth larvae. Further, reads mapping to LdMNPV-27/2 genome (Kabilov et al., 2015) revealed that most of the genome copies of LdMNPV-27/0 were missing sequences corresponding to a baculovirus repeated ORF (bro) gene and to an enhancin-encoding gene both of which are present in the consensus sequence for this gypsy moth virus. In the present study, we extended these results by assembly of complete genomes of strains LdMNPV-27/0 and LdMNPV-45/0 resulting from in vivo passages of the original virus strains. Also, bioassays of four additional LdMNPV isolates from two geographically close areas in North America and in Russia, were carried out against European and Asian gypsy moth larvae.

Here we pose two main questions. First, does the virulence of LdMNPV isolates link to tight spatial development of both virus and host populations, and second, what genetic differences are associated with LdMNPV isolates from spatially different locations?

2. Materials and methods

2.1. Insects, viral strains and experimental design

We tested six isolated LdMNPV strains (three from North America and three from Asia) against two populations of L. dispar: North American (Pennsylvania, L. dispar dispar) and Asian (Novosibirsk region, L. dispar asiatica). Thus we tested four additional LdMNPV strains to compare with results published earlier (Martemyanov et al., 2015) and to estimate within-group variation of viruses’ potency. Strains were isolated from dead larvae collected from forests in New York (42.70°N, 74.49°W), New Jersey (40.64°N, 74.92°W), and Massachusetts (42.16°N, 72.48°W) (LdMNPV-45/0) in the United States, and from forests in Tatarsk (55. 24°N, 75.82°E) (LdMNPV-27/0), Karasuk (53.70°N, 77.99°E), and Chistoozernyi (54.73°N, 76.65°E) in Western Siberia, Russia. Each viral strain was refreshed before the study by passaging once in an appropriate population of gypsy moth, i.e., Siberian LdMNPV strains were passed through groups of Siberian larvae and American LdMNPV strains were passed through groups of American larvae. Viruses were extracted from larval cadavers and lyophilized to minimize any loss of activity prior to starting bioassays (Lewis and Rollinson, 1978). Egg masses of the host were collected in autumn and held at 4°C during the winter to use as a stock for the experiments the following spring.

2.2. Viral potency assay

In the spring, egg masses from both gypsy moth populations were cleaned of setae, and surface sterilized by sodium hypochlorite to inactivate any LdMNPV on the eggs (Doane, 1969). Eggs from each population were pooled and transferred to 28° growth chambers for hatching and parallel larval rearing in the U.S. Forest Service Laboratory in Hamden, CT and in the Laboratory of Insect Pathology of the Institute of Systematics and Ecology of Animals, Novosibirsk, Russia. Larvae were reared on artificial diet within Petri dishes and were challenged with viruses when they reached the second instar (Lewis and Rollinson, 1978). We reared insects on artificial diet to exclude potential host plant effects such as those shown in a study with *Malacosoma californicum pluviale* nucleopolyhedrovirus (Cory and Myers, 2004).

For viral challenges we used a series of ten-fold dilutions of viral occlusion body (OBs) suspensions in sterile physiological saline incorporated into artificial diet. This was the same protocol used for the bioassay whose partial results were published earlier (Martemyanov et al., 2015). Unfortunately the protocol described in Martemyanov et al. (2015) was incorrectly described as a diet surface OB application. In fact it was a diet incorporation protocol and identical to that described for this study. The doses ranged from 10⁶ to 10² OBs per ml of diet. Fifty larvae from the Siberian *L. dispar asiatica* population and thirty larvae from the Pennsylvania *L. dispar dispar* population (ten larvae per Petri dish for both populations) were used for each dilution. Sterile physiological saline was used as a control. Larval mortality was recorded daily until the fifteenth day post challenge. Mortality data was subjected to probit analysis (Robertson et al., 2003) and results presented as lethal concentration values (LC₅₀) for each viral strain. Ninety-five percent confidence limits around the LC₅₀ were used to detect statistically significant potency differences between the virus strains.

2.3. The sequencing of LdMNPV strains

It is necessary to mention here that in the current study we used the strain LdMNPV-27/0 that varies from strain LdMNPV-27/2 (accession number KP027546) on which we published earlier (Kabilov et al., 2015). The strain LdMNPV-27/0 was obtained after one passage (refreshing) of stock strain LdMNPV-27 in the appropriate *L. dispar asiatica* population (as described above) while strain LdMNPV-27/2 (Kabilov et al., 2015) was obtained after two additional passages in the same *L. dispar asiatica* population. Viral DNA from LdMNPV-45/0 and LdMNPV-27/0 strains was phenol-chloroform extracted from purified viral polyhedral inclusion bodies (Zhu et al., 2009). DNA libraries prepared by Nextera DNA Sample Prep Kits (Illumina) were analyzed using a MiSeq genome sequencer (2 × 150 Cycles, Illumina) in the SB RAS Genomics Core Facility (ICBFM SB RAS, Novosibirsk, Russia), yielding an average of 250-fold coverage with QV > 20. Some regions were verified by Sanger sequencing. The full-length genome was assembled de novo with CLC GW 8.0 (Qiagen) and SPAdes softwares (Bankevich et al., 2012). Annotation of genomes was done as described previously (Kabilov et al., 2015). The genome sequences of LdMNPV-45/0 and LdMNPV-27/0 have been deposited in GenBank under the accession numbers KU862282 and KX249580 respectively.

2.4. Phylogenetic analysis of LdMNPV strains

For the eight LdMNPV strains with known whole genomes, we used the nucleotide sequences for the 37 core genes described for all members of family Baculoviridae (Garavaglia et al., 2012) and aligned them by MAFFT (Katoh and Standley, 2013) with default parameters. Sequences from LyxyMNPV-5 were used as an outgroup. Phylogenetic trees were inferred from the alignments in MEGA 7.0.14 (Kumar et al., 2016). Maximum likelihood (ML) tree construction methods were used.
are not signifi-

cantly different. 

with 1000 bootstrap re-sampling. The GTR + G + I substitu-
tion model was used for ML analysis.

A global alignment of the LdMNPV genome sequences was carried out with Mauve 2.4.0 (Darling et al., 2004) with default parameters. For genome-scale detection of positive Darwinian selection in groups of homologous coding sequences of LdMNPV POTION (Hongo et al., 2015) was used with modified “standard” parameters: multiple alignment = -muscle (3.8.31), bootstrap = 500, phylogenetic tree_speed = slow.

3. Results

3.1. Comparison of viral potencies

When LdMNPV strains were tested against larvae from the Pennsylvania L. dispar dispar population, all American LdMNPV strains (i.e. native strains) exhibited signifi-
cantly higher potency (10–100 fold) than any of the Siberian LdMNPV strains (i.e. foreign strains) (Table 1). When viral strains were tested against larvae from the Siberian L. dispar asiatica population, the American LdMNPV strains (with the exception of the New Jersey strain) also exhibited significantly higher potency than the Siberian LdMNPV strains (Table 1). It is important to note that all American strains showed similar potencies when they were tested against American L. dispar dispar population, while they showed more than ten-fold differences between each other when they were tested against Siberian L. dispar asiatica populations (Table 1).

3.2. Complete genome sequences of American and Asian LdMNPV strains

The LdMNPV genome sequences of isolates LdMNPV-27/0 (Asian) and LdMNPV-45/0 (American) were determined. The genome length of the Asian strain was 161,727 bp compared with 161,880 bp for the American strain. The average GC content was 57.5% for both strains. We found that the vef-1 and bro-p genes were absent in the LdMNPV-27/0 genome but not in the LdMNPV-45/0 genome (Fig. 1). On the other hand, genes bro-e and bro-o were absent in LdMNPV-45/0 but not in LdMNPV-27/0 (Fig. 1).

3.3. Phylogeny of LdMNPV isolates

The results of phylogenetic analysis which was done using 37 baculovirus core genes showed that Far East strains of LdMNPV diverged from other strains isolated in Europe/North America and Central Asia (Fig. 2). The strain LdMNPV-45/0 was combined in the same subcluster together with other North American strains. The strain LdMNPV-27/0 does not appear to be similar to strain 3029 isolated from the Russian baculovirus product Virin-ENSh (Fig. 2).

Results of the POTION analysis for the detection of genes evolving under positive Darwinian evolution showed 8 genes, found in all of the LdMNPV genomes analyzed, that contained positions under positive selection; they included hoar, DNA polymerase, DNA ligase, VP91, Desmoplakin, Telokin-like protein 20, Apsup and ORF130 (LdVgp130 from NC_001973) (Table 2). The consensuses of proteins with selected positions under positive selection are presented in a supplementary file (S1).

4. Discussion

In this study we report that the potency of LdMNPV strains does not depend upon the previous co-existence between the host and the pathogen in isolated localities. All but one of the American viral strains tested was more potent than the Asian viral strains against both the native and non-native host populations. The same host population-independen
ty of geographic isolates of LdMNPV was recently demonstrated by Harrison and coauthors (Harrison et al., 2016). Moreover our data show that the results of potencies of certain isolates could be approximated for the metapopulation of LdMNPV since the results among internal localities was not different.

Complete genome sequencing revealed the crucial differences in genetics between one studied representative of an Asian LdMNPV (strain LdMNPV-27/0) and one studied representative of an American LdMNPV (strain LdMNPV-45/0). The LdMNPV-27/0 strain revealed two full gene deletions, both the enhancin gene vef-1 and the gene bro-p, while the LdMNPV-45/0 strain revealed two full gene deletions of the bro-e and the bro-o genes. We recently published on some of these deletions in the genomes of compared strains where reads were mapped on the sequenced genome of LdMNPV-27/2 strain (Martemyanov et al., 2015). The enhancin genes (vef-1 and vef-2) are the genes directly associated with LdMNPV virulence (Bischoff and Slavicek, 1997). The second enhancin gene, vef-2, was present in the genomes of both strains compared in our study. Studies of the function of vef genes showed that the products of these genes are components of the virion envelope of the virus (Slavicek and Popham, 2005). Their encoded proteins are not involved in the degradation of the peritrophic membrane of the insect gut, unlike the role of the vef genes in closely related granulosis viruses; however, they are presumably involved in the binding of virions to the host gut cell membranes during the initiation of infection (Hoover et al., 2010). Earlier, Popham and co-authors (2001) assessed the influence of vef-1 and vef-2 genes (in the original paper they correspond to E1 and E2) on the virulence of LdMNPV. The deletion of either the vef-1 or the vef-2 gene decreased viral potency by approximately 2-fold, whereas deletion of both enhancin genes decreased potency by approximately 12-fold. This result indicates that the presence of one of these genes able to significantly compensate the absence of the other gene. That compensation was later confirmed by Hoover et al. (2010) and by Harrison et al. (2016). Thus, these facts indicate that the mutation of the vef-1 gene does not fully explain the observed essential differences in the potency between LdMNPV-27/0 and LdMNPV-45/0. The bro-p, bro-e and bro-o genes are representatives of the family of multicityp genes (baculovirus repeated orf, bro) that are present in some baculo-
viral genomes (Bideshi et al., 2003; Gong et al., 2003; Kang et al., 2006; Harrison et al., 2016). One of the described functions of bro genes is their participation in nuclear export (Kang et al., 2006). The number of bro genes in genomes varies significantly (Harrison et al., 2016), and information about their effect on virulence is not in the available literature.

The phylogenetic analysis which was done by using 37 baculovirus core genes shows close homology of all the annotated American LdMNPV strains that have appeared in North America since the virus was first mentioned in the literature (Glaser, 1915). Surprisingly,
Siberian strain LdMNPV-27/0 was close to the strain 3054, isolated in Spain, rather than to strain 3029 isolated from the Russian LdMNPV product virin ENSh. However, according to Bakhvalov (2001), the initial origin of 3029 strain is the Tien Shan Mountain region (Jalal-Abad, Kyrgyzstan) and that could explain the distant relationship between the 3029 and 27/0 strains. In general, the phylogeny tree of LdMNPV strains appears closely aligned with the genealogy tree of host species constructed using MtDNA sequences (Wu et al., 2015). The exception is the position of central Asian LdMNPV strains (i.e. 3029 and 27). In that case, for the L. dispar host, all Asian and Japanese forms were clustered together within the same branch (Wu et al., 2015 Fig. 4) while Asian LdMNPV strains (3029 and 27/0) were clustered within the European branch. One possible explanation of this phenomenon is that Central Asia is the place of maximal diversity of host’s genetic forms (Wu et al., 2015) and consequently, the same could be inferred for LdMNPV strains which are the host’s specific pathogens. Consequently,

Fig. 1. Mauve alignment of the LdMNPV-27/0 and LdMNPV-45/0 genome sequences, showing differences in structure of genomes. Block outlines of the same color correspond to Locally Collinear Blocks (LCBs), which are segments of the sequences that are conserved. Nucleotide sequence positions for each genome are indicated on a line above the LCBs. White boxes below the genomes correspond to annotated ORFs.

Fig. 2. Phylogenetic analysis of concatenated nucleotide sequence alignments of 37 baculovirus core genes showing relationships of the LdMNPV isolates and LyxyMNPV-5 which was used as an outgroup. Bootstrap values for each node are shown when the node occurred in trees inferred by maximum likelihood (ML).
there is the possibility to place within the tree the genotypes related to either western or eastern branches. However, the number of annotated complete genome sequences of Asian LdMNPV on which to base placement is restricted at the present time.

Using available LdMNPV complete genome data we also found that eight viral genes were under positive selection (Table 2), in seven of which were found certain site positions (File S1). It is noted that some of that relates to conservative protein families. We did not which were found certain site positions (File S1). It is noted that some ether western or eastern branches. However, the number of annotated there is the possibility to place within the tree the genotypes related to models evaluated (M1a, M2, M7, M8a, M8), p-values were computed through a LRT for each pair of nested models evaluated, q-values are FDR-corrected p-values, D values, N values are the number of codons under positive selection, and Len is the length of protein.

# Table 2

<table>
<thead>
<tr>
<th>Gene name</th>
<th>lnLm7</th>
<th>lnLm8</th>
<th>lnLm1</th>
<th>lnLm2</th>
<th>D (m78)</th>
<th>D (m12)</th>
<th>p_val (m78)</th>
<th>p_val (m12)</th>
<th>q_val (m78)</th>
<th>q_val (m12)</th>
<th>N (m2)</th>
<th>N (m8)</th>
<th>Len</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hoar</td>
<td>-6135</td>
<td>-6124</td>
<td>-6136</td>
<td>-6124</td>
<td>22</td>
<td>24</td>
<td>1.4E-05</td>
<td>6.0E-06</td>
<td>1.9E-04</td>
<td>1.4E-04</td>
<td>3</td>
<td>4</td>
<td>1195</td>
</tr>
<tr>
<td>DNA ligase</td>
<td>-2355</td>
<td>-2344</td>
<td>-2355</td>
<td>-2344</td>
<td>23</td>
<td>23</td>
<td>9.0E-06</td>
<td>9.0E-06</td>
<td>1.9E-04</td>
<td>1.4E-04</td>
<td>2</td>
<td>2</td>
<td>549</td>
</tr>
<tr>
<td>DNA polymerase</td>
<td>-4099</td>
<td>-4088</td>
<td>-4099</td>
<td>-4088</td>
<td>23</td>
<td>23</td>
<td>1.1E-05</td>
<td>1.0E-05</td>
<td>1.9E-04</td>
<td>1.4E-04</td>
<td>5</td>
<td>5</td>
<td>1003</td>
</tr>
<tr>
<td>VP91</td>
<td>-4062</td>
<td>-4051</td>
<td>-4063</td>
<td>-4051</td>
<td>23</td>
<td>24</td>
<td>1.0E-05</td>
<td>7.0E-06</td>
<td>1.9E-04</td>
<td>1.4E-04</td>
<td>4</td>
<td>4</td>
<td>829</td>
</tr>
<tr>
<td>ORF 130</td>
<td>-3916</td>
<td>-3904</td>
<td>-3916</td>
<td>-3904</td>
<td>24</td>
<td>24</td>
<td>5.0E-06</td>
<td>5.0E-06</td>
<td>1.9E-04</td>
<td>1.4E-04</td>
<td>6</td>
<td>7</td>
<td>856</td>
</tr>
<tr>
<td>Desmolakin</td>
<td>-3631</td>
<td>-3622</td>
<td>-3631</td>
<td>-3622</td>
<td>18</td>
<td>18</td>
<td>1.4E-04</td>
<td>1.3E-04</td>
<td>1.6E-03</td>
<td>1.3E-03</td>
<td>3</td>
<td>5</td>
<td>766</td>
</tr>
<tr>
<td>Telokin-like protein 20</td>
<td>-961</td>
<td>-956</td>
<td>-961</td>
<td>-956</td>
<td>10</td>
<td>9</td>
<td>6.5E-03</td>
<td>9.1E-03</td>
<td>4.8E-02</td>
<td>6.0E-02</td>
<td>1</td>
<td>0</td>
<td>222</td>
</tr>
<tr>
<td>Apsup</td>
<td>-1345</td>
<td>-1341</td>
<td>-1345</td>
<td>-1341</td>
<td>9</td>
<td>10</td>
<td>9.3E-03</td>
<td>8.4E-03</td>
<td>5.1E-02</td>
<td>4.9E-02</td>
<td>0</td>
<td>0</td>
<td>330</td>
</tr>
</tbody>
</table>

# 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.2017.04.004.

# References

Garavaglia, M.J., et al., 2012. The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family baculoviridae. J. Virol. 86, 12069–12079.
Liebold, A., Mastro, V., Schaefer, P.W., 1989. Learning from the legacy of Leopold