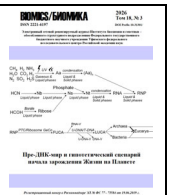
IBG UFRC RAS

BIOMICS

ISSN 2221-6197 https://biomicsj.ru



# Phylogenetic analysis of the NS1 protein of the *Clinch Densovirus* of Bivalve mollusks and its closest homologs using the unweighted pair group method with arithmetic mean (UPGMA)

[1,2]Bulavina M.K.*, [3]Lu Y.-H.,[1]Suzina N.E., [1]Zimin A.A.**

[1]Pushchino Scientific Center for Biological Research Russian Academy of Sciences, Pushchino, Russian Federation
[2]Federal State Budgetary Institution "National Medical Research Centre of Cardiology Named After Academician E.I. Chazov" of the Ministry of Health of the Russian Federation, Moscow, Russian Federation
[3]School of Life Sciences, Shanghai Normal University, Shanghai, China
*E-mail: ruffusramone@yandex.ru, **E-mail: dr.zimin8@yandex.ru

**Resume**

Densoviruses (*Densovirinae*) are a subfamily of small, non-enveloped, single-stranded DNA viruses which predominantly infect invertebrates, including insects, crustaceans, and echinoderms. Despite being structurally simple, they exhibita significant genomic diversity and are widespread in nature. The key nonstructural protein of these viruses is NS1, which plays a central role in viral DNA replication, regulation of gene expression, and induction of apoptosis in the infected cells. NS1 contains conserved domains characteristic of parvoviruses.

In the present study, we have performed a phylogenetic analysis of the NS1 protein of *Clinch densovirus* 1 (ClDV1) isolated from the freshwater mussel *Actinonaias pectorosa* during a mass die-off of the mollusks in the Clinch River. On the basis of UPGMA analysis of amino acid sequences, we have constructed a phylogenetic tree comprising 59 NS1 proteins isolated from various sources, including insects, crustaceans, echinoderms, birds, and bats.

The analysis has revealed that *Clinch densovirus* 1 clusters with unclassified *Ambidensoviruses*, confirming its assignment to the genus *Ambidensovirus.* In some cases, closely related viruses are detected in the organisms that are not their typical hosts (birds, bats), suggesting a possibility of passive accumulation of the viruses through food chains rather than their active replication in these animals. NS1 is found to be a highly conserved marker suitable for phylogenetic classification of densoviruses.

The practical significance of this work lies in the fact that some of the studied viruses infect commercially important aquaculture species (crayfish *Cherax quadricarinatus*) or insect pests, this opening prospects for using densoviruses as bioinsecticides. At the same time, viruses infecting mass-reared feeder insects (*Zophobas morio*) pose an economic threat. The study underscores the need for further investigation into the ecology and transmission mechanisms of densoviruses in natural ecosystems.

**Key words:** *Densovirinae, Clinch densovirus* 1, UPGMA, NS1 protein, phylogenetic analysis

**Citation:** Bulavina M.K., Lu Y.-H., Suzina N.E., Zimin A.A. Phylogenetic analysis of the NS1 protein of the *Clinch Densovirus* of Bivalve mollusks and its closest homologs using the unweighted pair group method with arithmetic mean (UPGMA). Biomics. 2026. V.18(3). P. 250-255. DOI: 10.31301/2221-6197.bmcs.2026-17

© **TheAuthors,** Bulavina M.K., Lu Y.-H., Suzina N.E., Zimin A.A.

## Introduction

Parvoviruses infecting vertebrates and invertebrates represent a family of single-stranded DNA viruses with small, non-enveloped capsids possessing icosahedral symmetry. The subfamily *Densovirinae* includes arthropod parvoviruses and causative agents of

natural epizootics in insect and crustacean populations. Structurally simple, these small DNA viruses demonstrate a great diversity of genomic sequences and organizations. They are characterized by a unique variety of hosts, including crustaceans, dipterans, mollusks, and even sea stars. This unique host diversity allowed them to spread widely across the globe [Grenet et al., 2015].

The NS1 protein of densoviruses is a key nonstructural protein, which plays a central role in the replication of viral genome and regulation of gene expression [Kapelinskaya et al., 2008]. NS1 forms a hexamer with an open, barrel-like shape. Three dimers connect side-by-side to create a wide, lipid-rich central channel [Gutsche et al., 2011].

Several functions are associated with this protein. For example, it initiates replication of viral DNA, which is carried out by DNA polymerase B of the host cell. The NS1 protein binds to a specific sequence $(GAC)_4$ in the double-stranded region of densoviral DNA and introduces a single-strand break at a specific site within inverted terminal repeats (ITR). This break serves as the starting point for the synthesis of a new viral DNA strand by the host DNA polymerase B via the "rolling circle" mechanism. NS1 also regulates gene expression by activating both its own promoter and the promoters of capsid genes. This creates a positive feedback loop that drives the replicative cycle forward. NS1 is able to covalently attach to the 5' end of DNA, which is necessary for packaging the genome into a new capsid. In addition, the NS1 protein has the ability to induce programmed cell death (apoptosis) in the infected cells. NS1 functions as a dimer.

*Clinch densovirus* 1 (ClDV1) is a novel virus from the family *Parvoviridae*, subfamily *Densovirinae*, which was identified by metagenomics in the freshwater mussel *Actinonaias pectorosa* (family *Unionidae*) inhabitating the Clinch River (USA). The virus was epidemiologically linked to a mass mortality event of mollusks, during which the *A. pectorosa* population declined by 85.4% in certain sections of the river. *Densoviruses* are well known as causative agents of epizootics in invertebrates, including shrimp, cockroaches, crickets, moths, crayfish, and sea stars. They cause lethargy, anorexia, flaccidity, and rapid death [Richard et al., 2020].

## Materials and Methods

The amino acid sequence of the NS1 protein was selected as a marker for *Clinch densovirus* 1. The database of nonredundant protein sequences (nr database) was searched for homologous proteins using the BLASTp algorithm. Phylogenetic analysis was performed using the UPGMA method [Sneath, Sokal, 1973]. Evolutionary distances were calculated using the JTT matrix-based method [Jones et al., 1992]. Evolutionary analysis was conducted with MEGA X software [Kumar et al., 2018].

## Results and Discussion

The optimal tree with a sum of branch lengths equal to 20.70524891 is presented in Figure 1. The percentage of replicates, in which the associated taxa were clustered together in the bootstrap test (1000 replicates), is shown next to the branches [Felsenstein, 1985] and is expressed as the number of amino acid substitutions per site. The analysis included 59 amino acid sequences. All positions containing gaps and ambiguous data were eliminated (complete deletion option).The final dataset contained 304 positions.

The amino acid sequence of the target protein (indicated by a black diamond in the figure) is almost identical to the putative sequence of a virus from the genus *Ambidensovirus*; perhaps this protein is a nearly complete NS1 homolog. These viruses are very small, non-enveloped particles consisting of 60 copies of the capsid protein VP1. Their genome is represented by a linear single-stranded DNA with a size of approximately 5.3–6.0 kb [Erlandson, 2008; Bochow et al., 2015]. According to the ICTV database, their key feature is the ambisense organization: both positive and negative DNA strands are packaged in equal amounts in the virions. Viruses of the genus *Ambidensovirus* infect a wide range of invertebrates, predominantly insects, but also crustaceans and echinoderms. ClDV1 clusters with unclassified ambidensoviruses, confirming its assignment to the genus *Ambidensovirus*.

The neighboring branch contains three nonstructural proteins associated with the genus *Ambidensovirus*, the subfamily *Densovirinae*, and the family *Parvoviridae*. Their names reflect the organisms from which they were isolated: *Periparus ater* (coal tit) and *Periplaneta fuliginosa* (smokybrown cockroach). According to the GenBank database, the genome of the first virus is formed by a linear DNA consisting of 5483 nucleotides (nt). It was isolated from a bird's cloacal swab and studied using metagenomics and next-generation sequencing technologies [Dai et al., 2022]. The genome of the *Periplaneta fuliginosa densovirus* is 5454 nucleotides long and contains distal inverted terminal repeats (ITRs). The ORF2 polypeptide contains a highly conserved NTP-binding domain of parvovirus NS proteins [Guo et al., 2001; Hu et al., 2007]. Adjacent to these viruses is the NS1 protein of an unclassified *Parvovirus*.

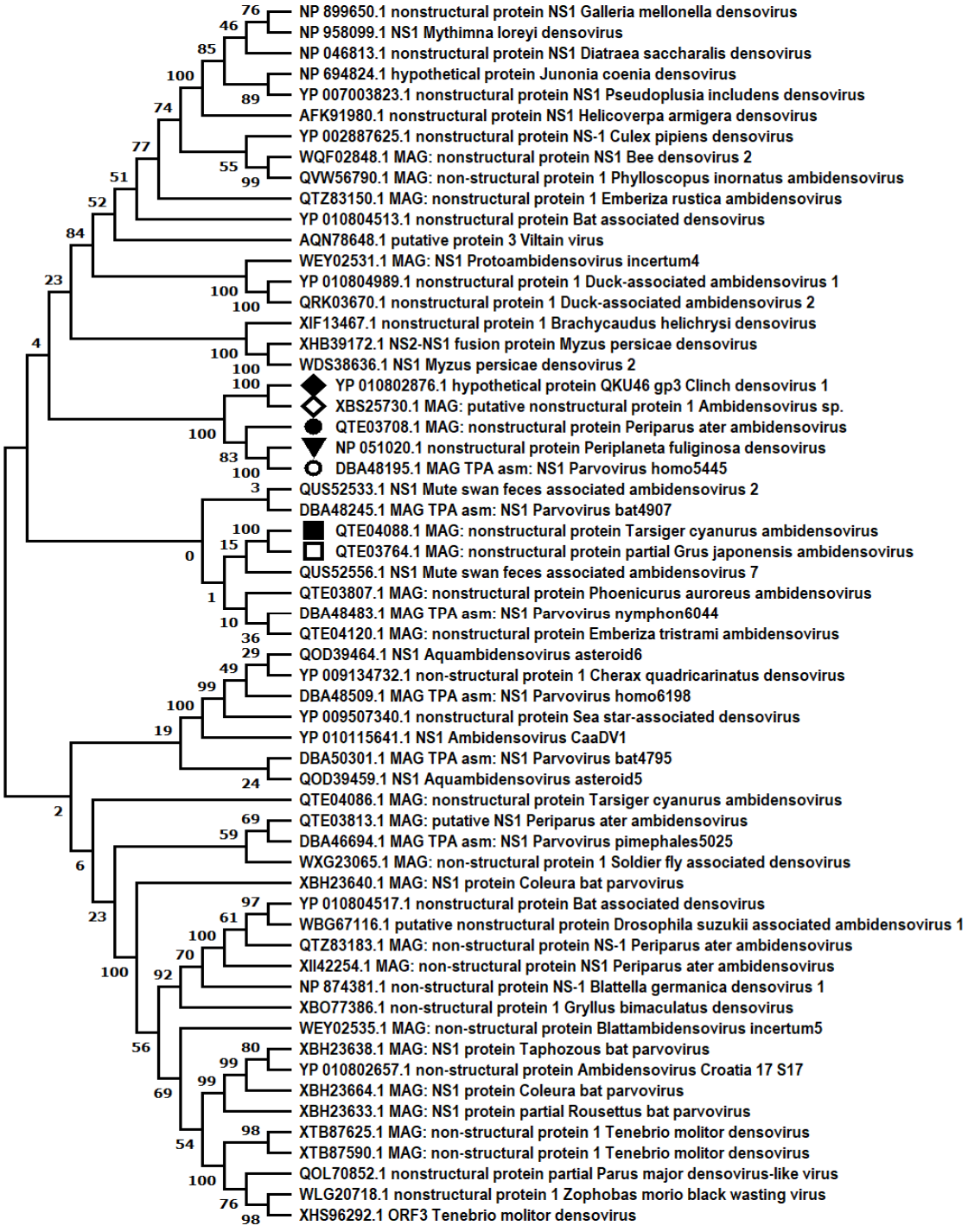



Fig. 1. Evolutionary relationships of the amino acid sequences of nonstructural NS1 proteins of *Clinch densovirus* 1 (indicated by a black diamond) and other viruses. The putative *Ambidensovirus* protein is indicated by a white diamond. The *Periparus ater Ambidensovirus ambidensovirus* protein is marked by a black circle. The black inverted triangle marks the densovirus of the smokybrown cockroach (*Periplaneta fuliginosa*). The white circle denotes the NS1 protein of an unclassified *Parvovirus*. The black and white squares mark *Ambidensovirus* proteins infecting red-flanked bluetails (*Tarsiger cyanurus*) and red-crowned cranes (*Grus japonensis*), respectively.

All the proteins described above are connected to another branch, which branches repeatedly. The closest branch includes proteins of three viruses; the first two proteins are similar to each other and form their own group. One of these proteins, the NS2-NS1 fusion protein of *Myzus persicae densovirus*, is a chimeric protein formed by alternative mRNA splicing in the virus [Sukhikh et al., 2024]. It infects *Myzus persicae* (green peach aphid). The second is a nonstructural protein of the MpDV2 virus. These two viruses are close relatives. Studies show that MpDV2 forms a separate evolutionary branch within the genus *Hemiambidensovirus*. The MpDV and MpDV2 viruses have differences in the organization of their genomes and in protein sequences. Phylogenetic analysis based on the conserved NS1 protein sequences showed that aphid densoviruses form a separate branch within the genus *Hemiambidensovirus* [Guo et al., 2023]. Adjacent to them is the nonstructural protein of a densovirus infecting *Brachycaudus helichrysi* (leaf-curl plum aphid), from which the virus was first isolated. In terms of amino acid sequence, the NS1 of BhDV shows similarity to the homologous proteins of both aphid-infecting densoviruses and densoviruses of other hosts. Phylogenetic analysis indicates that BhDV represents a new species in the genus *Hemiambidensovirus* [Li et al., 2025].

The next branch contains nonstructural proteins of an unclassified *Protoambidensovirus* virus and two other viruses, first discovered in wild ducks of North America. Phylogenetic analysis of the NS1 protein showed that the Duck-associated *Ambidensovirus* is evolutionary proximal to the viruses found in both vertebrates (including birds) and invertebrates [Canuti et al., 2021].

Two separate branches are occupied by single viral proteins. One of them is the putative protein 3 of the *Viltain virus*. The genome of this virus contains several putative reading frames; one of the frames has anS03 segment which encodes a protein homologous to DNA polymerase B (DNA PolB) [Wallace et al., 2020]. The other was found in insectivorous bats. Viruses of this type are similar in their genome structure and phylogenetically close to densoviruses found in insects and birds. This may indicate that the viruses enter the bat's body with food and are not its specific pathogens [Popov et al., 2025]. The third virus was isolated from a reed bunting (*Emberiza rustica*), but was not classified or described.

Next, the tree branches, forming many small groups of NS1 proteins isolated from the microbiota of various animals. Among them are phylloscopus (*Phylloscopus inornatus*), bees, common mosquito (*Culex pipiens*), cotton bollworm (*Helicoverpa armigera*), butterflies *Junonia coenia* and *Mythimna loreyi,* moths *Diatraea saccharalis* and *Galleria mellonella*.

It's worth to return to the root of the tree. Initially, it splits into two large branches, one of which contains the proteins described above, along with several others. The nearest branch is represented by eight nonstructural proteins divided into two groups. The first group includes the protein of a virus found in the feces of mute swan (*Cygnus olor*). The name of the virus reflects the place of its discovery (in swan droppings) and its belonging to the genus *Ambidensovirus* [McCallum et al., 2026]. Adjacent to it is a virus isolated from the bat isolate and mentioned in the context of horizontal gene transfer between animals [Buck et al., 2024]. The secondgroup combines nonstructural proteins of various viruses. *Tarsiger cyanurus ambidensovirus*, marked with a black square, was found in red-flanked bluetail (*Tarsiger cyanurus*); a white square indicates a virus found in red-crowned crane. These viruses were identified by a metagenomic analysis of the virome of wild birds [Dai et al., 2022]. The adjacent group also contains viruses found in birds.

The second large branch is also divided into multiple groups. The first group is interesting for the proteins of the viruses *Aquambidensovirus asteroid6* and *Seastar-associated densovirus* (SSaDV) *Aquambidensovirus asteroid5*, infecting sea stars [Jackson et al., 2020]. Another well-studied virus of this group is *Cherax quadricarinatus densovirus* (CqDV), which infects Australian redclaw crayfish (*Cherax quadricarinatus*). It was the first virus to be studied in detail in these freshwater crustaceans [Rudoy et al., 2025; Bochow et al., 2015]. Standing out among viruses of this group is *Parvovirus bat4795*, which is a pathogen isolated from bats [Buck et al., 2024].

The next branch is formed by a single viral protein found in red-flanked bluetail bird (*Tarsiger cyanurus*). Below this branch is a small group of viruses isolated from various animals: black tit, hamster *Mesocricetus auratus*, and soldier fly larvae. The second member of this group is a provirus (a virus integrated into the host genome); it was found in hamster cells [Buck et al., 2024]. The third virus was proposed to be used to control fly populations [Divekar et al., 2024].

Below this group is a single protein of *Coleura bat parvovirus* found in the African sheath-tailed bat (*Coleura afra*). Phylogenetically, this virus is related to other unclassified rodent-infecting paramyxoviruses: the *J* and *Beilong* viruses. However, the virus was only found in *Coleura afra* and was not detected in other bat species inhabiting the same caves [Maganga et al., 2014].

Beyond that, the tree branches repeatedly, forming several groups of branches and single branches related to the densoviruses isolated from various organisms: black tit, *Drosophila*, red cockroach, two-spotted cricket, bats of the genus *Taphozous* and other bats, greater wax moth, great tit, and larvae of the beetle

*Zophobas morio*. Among them, the ORF3 protein of *Tenebrio molitor densovirus* stands out, since ORF3 denotes a reading frame, but the functions of this protein have not been studied [Lim et al., 2026].

Wrapping up the analysis, one can note that the related nonstructural NS1 proteins infect various invertebrate hosts: mostly insects, but sea stars and crayfish as well. The host attribution of viruses isolated from non-invertebrate sources is controversial, since *Densovirinae* are known to infect only invertebrates [Aş, Eroğlu, 2025] and it is not entirely clear if a non-invertebrate organism is the true host or merely receives viruses with food. However, there has been an interesting report of viruses of the genus *Ambidensovirus*, which had never been found in humans before, being detected in the blood of two Brazilian patients with symptoms resembling dengue fever [Fahsbender et al., 2020].

Some viruses presented on the tree pose a danger to farm animals, such as *Zophobas morio* black wasting virus (ZmBWV), which causes a severe disease in the larvae of the beetle *Zophobas morio*. These larvae are mass-reared as a valuable protein feed for reptiles, birds, and other exotic animals [Penzes et al., 2024; Pénzes, Kaelber, 2022]. Also on the list are viruses that infect commercially important species. For example, *Cherax quadricarinatus densovirus* infects freshwater crayfish, the meat of which is considered a delicacy [Bochow et al., 2015], and *Acheta domesticus mini ambidensovirus* causes significant economic damage in the cricket farming industry [Liu et al., 2011]. Some viruses from this list infect insect pests, making them potentially interesting for pest control (bioinsecticides) – taking into account that similar methods of using *Densovirinae*as a bioinsecticide have already been implemented in practice [Batool et al., 2022]. The NS1 protein can be considered a virulence factor and used as a target for combating densoviruses.

NS1 is a very conserved protein: its sequence is highly homologous in many representatives of *Densovirinae*. The conservation of NS1 is a result of the common mechanism of genome replication shared by all densoviruses regardless of their diverse hosts. The arrangement of the viruses on the phylogenetic tree generally corresponds to their genus, which suggestsa potential use of NS1 as a marker for the classification of the subfamily *Densovirinae*. The presence of these viruses in the organisms that are not their hosts indicates infection through food or the habitat environment, and also raises questions about their transmission pathways and host switching.

## Conclusion

Densoviruses are still an understudied subfamily. Many isolated viruses are insufficiently characterized: the functions of their proteins are not clear, and the pathways of their accumulation and transmission remain poorly investigated. This group has the potential to become a convenient tool in the hands of virologists, entomologists, as well as specialists in aquatic bioresources and aquaculture.

## Acknowledgements

The work was supported by the Russian Science Foundation, grant No. 24-64-00017, https://rscf.ru/project/24-64-00017/

*Received: 03.09.2026.*
*Accepted for publication: 21.09.2026.*

## References

1. Aş Y, Eroğlu GB. Densovirinae: An Eco-Friendly Alternative in Biological Control. *Zenodo*. 2025. doi: 10.5281/zenodo.15877070
2. Batool K, Xiao J, Xu Y et al. Densovirus Oil Suspension Significantly Improves the Efficacy and Duration of Larvicidal Activity against *Aedes albopictus*. Viruses. 2022. 14(3). 475. doi: 10.3390/v14030475
3. Bochow S, Condon K, Elliman J, Owens L. First complete genome of an Ambidensovirus; *Cherax quadricarinatus* densovirus, from freshwater crayfish *Cherax quadricarinatus*. *Marine Genomics*. 2015. 24. 305-312. doi: 10.1016/j.margen.2015.07.009
4. Buck CB, Welch N, Belford AK et al. Widespread Horizontal Gene Transfer Among Animal Viruses. *bioRxiv*. 2024. doi: 10.1101/2024.03.25.586562
5. Canuti M, Verhoeven JTP, Munro HJ et al. Investigating the Diversity and Host Range of Novel Parvoviruses from North American Ducks Using Epidemiology, Phylogenetics, Genome Structure, and Codon Usage Analysis. *Viruses*. 2021. 13(2). 193. doi: 10.3390/v13020193
6. Dai Z, Wang H, Wu H et al. Parvovirus dark matter in the cloaca of wild birds. *GigaScience*. 2022. 12. giad001. doi: 10.1093/gigascience/giad001
7. Divekar G, Colmant AMG, Furlong MJ, Etebari K. Transcriptome Analysis Reveals a Diverse Range of Novel Viruses in Australian Sugarcane Soldier Fly (*Inopus flavus*) Larvae. *Viruses*. 2024. 16(4). 516. doi: 10.3390/v16040516
8. Erlandson MA. Insect Pest Control by Viruses. In: Mahy BWJ, Van Regenmortel MHV (eds.). *Encyclopedia of Virology* (Third Edition). Academic Press. 2008. 125-133. doi: 10.1016/B978-012374410-4.00432-5
9. Fahsbender E, Charlys da-Costa A, Gill DE et al. Plasma virome of 781 Brazilians with unexplained symptoms of arbovirus infection include a novel parvovirus and densovirus. *PLoS One*. 2020. 15(3). e0229993. doi: 10.1371/journal.pone.0229993

10. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*. 1985. 39. 783-791. doi: 10.1111/j.1558-5646.1985.tb00420.x
11. Grenet A-S, Salasc F, Francois S et al. The densovirus: A "massive attack" in the arthropods. *Virologie*. 2015. 19. 19-31. doi: 10.1684/vir.2015.0589
12. Guo H, Zhang J, Hu Y. Complete sequence and organization of *Periplaneta fuliginosa* densovirus genome. *Acta virologica*. 2001. 44. 315-322.
13. Guo Y, Zhao Y, Yang Y et al. Plants affect the horizontal transmission of a new densovirus infecting the green peach aphid *Myzus persicae* by modulating honeydew production. *Insect Science*. 2023. 31. doi: 10.1111/1744-7917.13235
14. Gutsche I, Coulibaly F, Voss JE et al. Secreted dengue virus nonstructural protein NS1 is an atypical barrel-shaped high-density lipoprotein. *Proc Natl Acad Sci USA*. 2011. 108(19). 8003-8008. doi: 10.1073/pnas.1017338108
15. Hu L, Zhang L, Shen C et al. The densovirus of *Periplaneta fuliginosa* (PfDNV) as an insect vector for persistent foreign gene expression *in vivo*. *Biochemical and Biophysical Research Communications*. 2007. 358. 976-982. doi: 10.1016/j.bbrc.2007.04.210
16. Jackson EW, Wilhelm RC, Johnson MR et al. Diversity of Sea Star-Associated Densoviruses and Transcribed Endogenous Viral Elements of Densovirus Origin. *J Virol*. 2020. 95(1). e01594-20. doi: 10.1128/JVI.01594-20
17. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences*. 1992. 8. 275-282. doi: 10.1093/bioinformatics/8.3.275
18. Kapelinskaya TV, Martynova EU, Korolev AL et al. Transcription of the smoky brown cockroach densovirus BgDNV genome: alternative processing of viral RNAs. *Doklady Akademii nauk*. 2008. 421(2). 256-261. doi: 10.1134/S1607672908040042 (In Russian)
19. Kumar S, Stecher G, Li M et al. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*. 2018. 35. 1547-1549. doi: 10.1093/molbev/msy096
20. Li Y, Ma Y, Liu X et al. Genome sequence of a new densovirus infecting the leaf-curl plum aphid *Brachycaudus helichrysi*. *Archives of Virology*. 2025. 170(4). 65. doi: 10.1007/s00705-025-06266-w
21. Lim FS, González-Cabrera J, Jehle JA et al. No longer uncertain: the validation of tenebrionid insects as hosts of *Blattambidensovirus* incertum1 isolates by phylogeny and infection studies. *J Gen Virol*. 2026. 107(2). 002211. doi: 10.1099/jgv.0.002211
22. Liu K, Li Y, Jousset FX et al. The *Acheta domesticus* densovirus, isolated from the European house cricket, has evolved an expression strategy unique among parvoviruses. *J Virol*. 2011. 85(19). 10069-10078. doi: 10.1128/JVI.00625-11
23. Maganga GD, Bourgarel M, Obame-Nkoghe J et al. Identification of an Unclassified Paramyxovirus in *Coleura afra*: A Potential Case of Host Specificity. *PLoS One*. 2014. 9. e115588. doi: 10.1371/journal.pone.0115588
24. McCallum B, Bakkeren G, Wang X et al. Comprehensive analysis of the wheat leaf virome associated with *Puccinia triticina* isolates across Canada reveals high viral diversity and novel mycoviruses. *Phytopathol Res*. 2026. 8. 64. doi: 10.1186/s42483-026-00443-8
25. Penzes JJ, Holm M, Yost SA, Kaelber JT. Cryo-EM-based discovery of a pathogenic parvovirus causing epidemic mortality by black wasting disease in farmed beetles. *Cell.* 2024. 187(20). 5604-5619.e14. doi: 10.1016/j.cell.2024.07.053
26. Pénzes J, Kaelber J. Identification by cryoEM of a densovirus causing mass mortality in mass-reared larval darkling beetles (*Zophobas morio*). *bioRxiv*. 2022. doi: 10.1101/2022.05.14.491968
27. Popov IV, Manakhov AD, Gorobets VE et al. Metagenomic Investigation of Intestinal Microbiota of Insectivorous Synanthropic Bats: Densoviruses, Antibiotic Resistance Genes, and Functional Profiling of Gut Microbial Communities. *Int J Mol Sci*. 2025. 26(13). 5941. doi: 10.3390/ijms26135941
28. Richard JC, Leis E, Dunn CD et al. Mass mortality in freshwater mussels (*Actinonaias pectorosa*) in the Clinch River, USA, linked to a novel densovirus. *Scientific Reports*. 2020. 10. 14498. doi: 10.1038/s41598-020-71459-z
29. Rudoy DV, Olshevskaya AV, Golovko LS et al. The most significant diseases of the Australian redclaw crayfish *Cherax quadricarinatus* (von Martens, 1868) during its cultivation in aquaculture conditions. *Vestnik KrasGAU*. 2025. (9). 123-140. doi: 10.36718/1819-4036-2025-9-123-140 (In Russian)
30. Sneath PHA, Sokal RR. Numerical Taxonomy. *Freeman, San Francisco*. 1973.
31. Sukhikh N, Golyaev V, Laboureau N et al. Deep Sequencing Analysis of Virome Components, Viral Gene Expression and Antiviral RNAi Responses in *Myzus persicae* Aphids. *Int J Mol Sci*. 2024. 25(23). 13199. doi: 10.3390/ijms252313199
32. Wallace M, Coffman K, Gilbert C et al. The discovery, distribution and diversity of DNA viruses associated with *Drosophila melanogaster* in Europe. *bioRxiv*. 2020. doi: 10.1101/2020.10.16.342956