



**HAL**  
open science

## Genome sequence of a crustacean iridovirus, IIV31, isolated from the pill bug, [*i*]Armadillidium vulgare[/*i*]

Benoit Piegu, Sébastien Guizard, Tan Yeping, Corinne Cruaud, Sassan  
Asgari, Dennis K. Bideshi, Brian A. Federici, Yves Bigot

### ► To cite this version:

Benoit Piegu, Sébastien Guizard, Tan Yeping, Corinne Cruaud, Sassan Asgari, et al.. Genome sequence of a crustacean iridovirus, IIV31, isolated from the pill bug, [*i*]Armadillidium vulgare[/*i*]. Journal of General Virology, 2014, 95 (7), pp.1585-1590. 10.1099/vir.0.066076-0 . hal-01129876

**HAL Id: hal-01129876**

**<https://hal.science/hal-01129876v1>**

Submitted on 15 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### **Genome sequence of a crustacean iridovirus, IIV31, isolated from the pill bug, *Armadillidium vulgare***

**Citation for published version:**

Piegu, B, Guizard, S, Tan, YP, Cruaud, C, Asgari, S, Bideshi, DK, Federici, BA & Bigot, Y 2014, 'Genome sequence of a crustacean iridovirus, IIV31, isolated from the pill bug, *Armadillidium vulgare*', *Journal of General Virology*, vol. 95, pp. 1585-1590. <https://doi.org/10.1099/vir.0.066076-0>

**Digital Object Identifier (DOI):**

[10.1099/vir.0.066076-0](https://doi.org/10.1099/vir.0.066076-0)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Journal of General Virology

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Short Communication

# Genome sequence of a crustacean iridovirus, IIV31, isolated from the pill bug, *Armadillidium vulgare*

Benoît Piégu,<sup>1</sup> Sébastien Guizard,<sup>1</sup> Tan Yeping,<sup>2,3</sup> Corinne Cruaud,<sup>4</sup> Sassan Asgari,<sup>5</sup> Dennis K. Bideshi,<sup>2,6</sup> Brian A. Federici<sup>2,3</sup> and Yves Bigot<sup>1</sup>

### Correspondence

Yves Bigot  
yves.bigot@tours.inra.fr

<sup>1</sup>UMR INRA-CNRS 7247, PRC, Centre INRA Val de Loire, 37380 Nouzilly, France

<sup>2</sup>Department of Entomology, Molecular and Developmental Biology, University of California, Riverside, CA 92521, USA

<sup>3</sup>Interdepartmental Graduate Programs in Cell, Molecular and Developmental Biology, University of California, Riverside, CA 92521, USA

<sup>4</sup>CEA/Institut de Génomique GENOSCOPE, 2 Rue Gaston Crémieux, CP 5706, 91057 Evry CEDEX, France

<sup>5</sup>School of Biological Sciences, University of Queensland, St Lucia, QLD 4072, Australia

<sup>6</sup>California Baptist University, Department of Natural and Mathematical Sciences, 8432 Magnolia Avenue, Riverside, CA 92504, USA

Members of the family *Iridoviridae* are animal viruses that infect only invertebrates and poikilothermic vertebrates. The invertebrate iridovirus 31 (IIV31) was originally isolated from adult pill bugs, *Armadillidium vulgare* (class Crustacea, order Isopoda, suborder Oniscidea), found in southern California on the campus of the University of California, Riverside, USA. IIV31 virions are icosahedral, have a diameter of about 135 nm, and contain a dsDNA genome 220.222 kbp in length, with 35.09 mol % G + C content and 203 ORFs. Here, we describe the complete genome sequence of this virus and its annotation. This is the eighth genome sequence of an IIV reported.

Received 21 March 2014

Accepted 8 April 2014

The family *Iridoviridae* consists of large dsDNA viruses that infect species of both poikilothermic vertebrates (fishes, amphibians and reptiles) and invertebrates (arachnids, cephalopods, crustaceans, insects, molluscs, nematodes and polychaetes; Williams, 2008). These viruses are members of the order *Megavirales* (Colson *et al.*, 2013), so-called nucleocytoplasmic large DNA viruses (NCLDV; Iyer *et al.*, 2001). The dsDNA genomes of iridoviruses are circularly permuted with terminal redundancy. As a consequence, the map of their genomes is represented as a circular molecule. Only one linear molecule is encapsidated in each virion, with the ends of individual encapsidated genomes being located at different positions on the map of different virions (Bigot *et al.*, 2000). The genome of vertebrate iridoviruses is highly methylated, whereas little or no methylation occurs in the genomes of the invertebrate iridoviruses (IIV). Replication of the iridoviral genome includes distinct nuclear and cytoplasmic phases, but virions only assemble in the cytoplasm (Jancovich *et al.*, 2011). The genomes are encapsidated within an icosahedral capsid ranging between 120 and 180 nm in diameter. Capsids are composed predominantly of a 50 kDa major capsid protein. The IIV studied by

cryo-electron microscopy have surface fibrils 2 nm in diameter (Yan *et al.*, 2000).

The taxonomic structure of the family *Iridoviridae* is currently organized into five genera: *Chloriridovirus*, *Iridovirus*, *Lymphocystivirus*, *Megalocytivirus* and *Ranavirus*. Members of the two first genera have a host range restricted to invertebrate species, whereas the three others infect only poikilothermic vertebrates. The model species for the genus *Chloriridovirus* is IIV3 (Delhon *et al.*, 2006; Jancovich *et al.*, 2011), the only species reported in this genus. The type species for the genus *Iridovirus* is IIV6, and only two species, IIV1 and IIV6, have been recognized by the International Committee for Virus Taxonomy (ICTV) as representatives of this genus. Ten other related viruses that may be *Iridovirus* species await biological and genomic data before it can be determined whether they are valid species or variants of existing species. Interestingly, the phylogenetic analyses of proteins encoded by IIV3, IIV6 (Jakob *et al.*, 2001) and IIV9 (Wong *et al.*, 2011) have revealed that IIV9 is more closely related to IIV3 than to IIV6. This has been confirmed with IIV22, IIV22A, IIV25 and IIV30, four close relatives of IIV9 (Piégu *et al.*, 2013a, b, c, d), indicating that some species of the genus *Iridovirus* are more closely related to members of the genus *Chloriridovirus* than to other *Iridovirus* species from insects. Consequently, this suggests that the genus

The GenBank/EMBL/DDBJ accession number for the complete genome sequence of IIV31 is HF920637.

A supplementary table is available with the online version of this paper.

*Iridovirus* contains several diverse species or species complexes. Division of the *Iridovirus* genus into three species complexes, the Polyiridovirus (type species IIV9), the Oligoiridovirus (type species IIV6) and the Crustaceoiridovirus (type species IIV31), has been proposed (Williams, 1994; Williams & Cory, 1994), among which members of the Polyiridovirus complex would share a common iridovirus ancestor with those of the genus *Chloriridovirus*. This proposal has not been accepted formally by the ICTV, but in agreement with it are data showing that members of the Oligoiridovirus complex are the closest IIV relatives of the family *Ascoviridae* (Bigot *et al.*, 2011), and share a common IIV ancestor (Stasiak *et al.*, 2003; Bigot *et al.*, 2009). Moreover, ascoviruses and IIVs are more closely related to each other than to vertebrate iridoviruses. Although first determined through phylogenetic analyses, this classification is strongly supported by data showing that these invertebrate viruses share 26 core genes, of which only 19 are found in vertebrate iridoviruses (Eaton *et al.*, 2007; Bigot *et al.*, 2009). A precise definition of genera among the IIVs therefore remains unresolved at present, and further investigations are required to elucidate whether or not species complexes should be elevated to genus rank.

To date, six genomes of IIVs have been sequenced: IIV3, the mosquito iridescent virus (Delhon *et al.*, 2006), IIV6, the *Chilo* iridescent virus (Jakob *et al.*, 2001), IIV9, the *Wiseana* iridovirus (Wong *et al.*, 2011), IIV22 (Piégu *et al.*, 2013a), IIV22A (Piégu *et al.*, 2013b), IIV25 (Piégu *et al.*, 2013c) and IIV30 (Piégu *et al.*, 2013d). Here, we present a summary classification and a set of features for IIV31, the eighth IIV sequenced, together with the description of the sequencing and annotation of its genome. To date, the classification status of IIV31 reveals that it is a tentative species belonging to the genus *Iridovirus*, in which it is related to the Crustaceoiridovirus complex (Williams, 1994; Williams & Cory, 1994; Jancovich *et al.*, 2011).

IIV31 was originally isolated from a sample of adult pill bugs of the species *Armadillidium vulgare* (Crustacea, Isopoda, Oniscidea) in southern California on the campus of the University of California, Riverside, USA (Federici, 1980; Williams, 1994). IIV31 has been found to have a wide host range in populations of crustacean species (Wijnhoven & Berg, 1999). Large quantities of virions can be directly purified from adults recovered from natural habitats such as leaf litter. However, it can also be produced *per os* and by injection of healthy isopods, or larvae of a coleopteran (Cole & Morris, 1980). Here, virions of iridovirus type IIV31 were harvested in April 2010 from isopods collected from a bed of ivy on the campus of the University of California, Riverside, and frozen at  $-80^{\circ}\text{C}$ . IIV31-infected isopod adults are recognizable by a characteristic blue discoloration of the normally grey cuticle. IIV31 virions and their genomic DNA (gDNA) were purified as described elsewhere (Federici, 1980; Bigot *et al.*, 2009). The purity of our IIV31 sample was determined by digesting virion gDNA with restriction enzymes using published data as a reference (Federici, 1980; Williams, 1994).

In 2009, the scientific committee of GENOSCOPE selected the IIV31 genome for sequencing. The complete genome sequence and annotation are now available in the EMBL database (accession no. HF920637). A summary of the project results is shown in Table 1. The genome of IIV31 was sequenced using the 454 FLX pyrosequencing platform (Roche/454). Library construction and sequencing were performed as previously described elsewhere (Henn *et al.*, 2010). *De novo* genome assembly was performed using the Newbler v2.3 assembly software package as previously described elsewhere (Henn *et al.*, 2010). Assembly metrics are described in Table 1. The assembled contig representing the entire IIV30 genome sequence was confirmed by comparing five predicted restriction fragment profiles from the genome, for *Bam*HI, *Eco*RI, *Hind*III, *Pst*I and *Sal*I, with the matching fragment profiles produced by actual restriction digestions of the IIV31 genome (Federici, 1980; Williams, 1994).

Genes were identified using the Broad Institute Automated Phage Annotation Protocol as described elsewhere (Ashburner *et al.*, 2000; Henn *et al.*, 2010). Briefly, evidence-based and *ab initio* gene prediction algorithms were used to identify putative genes, followed by construction of a consensus gene model using a rules-based evidence approach. Gene models were manually checked for errors such as in-frame stops, very short peptides, splits and merges. Additional gene prediction analysis and functional annotation were performed as previously described (Bigot *et al.*, 2009).

General features of the IIV31 genome sequence (Table 2) include a nucleotide composition of 35.09 mol% G + C (Fig. 1). Pair alignment using BLASTN of the IIV31 genome with those of the IIV3, IIV6, IIV9, IIV22, IIV22A, IIV25 and IIV30 genomes did not reveal any identity at the level of nucleotide sequences. Similarly, we did not find any conserved cluster of collinear genes between IIV31 and the other IIVs.

A total of 203 genes encoding proteins were predicted. No genes coding for tRNAs were found. Of the 203 coding DNA sequences (CDSs), 125 were in forward orientation and 78 in reverse orientation. Four gene pairs were found to overlap: 103R/104L, 114R/155L, 129L/130R and 162L/163R. Eighty-four CDSs (41.4%) have been annotated with functional product predictions. The annotation of the 203 genes, described in Table S1 (available in the online Supplementary Material), revealed that 184 of the 203 CDSs have a related gene in databases, with *e* values below  $10^{-3}$ . One hundred and seventy-two genes have an orthologue in the IIV6 genome, three genes (004R, 040R, 067L) have an orthologue in the IIV9 genome that did not occur in the IIV6 genome, four genes (015L, 024R, 084R, 119R) have a viral orthologue that did not occur in one of the currently sequenced IIV genomes, and five genes (013L, 022L, 082L, 128L, 136L) have no orthologue in any currently sequenced viral genomes, but have an orthologue in a prokaryotic or a eukaryotic genome. Finally, 19 genes have no orthologue in databases and putatively correspond

**Table 1.** Genome sequencing project information

MIGS ID	Property	Information
MIGS-31	Finishing quality	Finished (>99 %)
	No. of contigs	1
	Assembly size	220 222 bp
	Assembly coverage	30 ×
	Total number of reads used	33 821
MIGS-29	Sequencing plate form	454
MIGS-30	Assemblers	Newbler version 2.3, post-release – 19.11.2009
	Gene calling method	Annotation protocol – Bigot <i>et al.</i> (2009)
	EMBL accession no.	HF920637

to novel genes (003R, 010L, 025L, 034R, 035L, 072L, 074R, 080R, 085R, 091R, 098R, 101R, 102R, 105L, 109R, 114R, 153L, 183R, 192R). Due to its close relatedness to IIV6, the functions of 84 of the proteins, assigned by Jakob *et al.* (2001), are considered the same for this isopod IIV.

With regard to repeats, three families of gene paralogues occur in the IIV31 genome. The first contains 10 members that are related to CIV genes 006L, 019R, 029R, 146R, 148R, 211L, 212L, 238R, 313L, 388R, 420R and 468L. The second contains three members related to CIV261R, 396L and 443R. The third contains eight members that belong to the family *bro*-like genes, a widespread family of repeated genes in NCLDV (Bideshi *et al.*, 2003).

Three gene fossils, i.e. genes containing stop codons or frameshifts, were found in IIV31. The status of these genes was confirmed by PCR and sequencing, and therefore we decided to annotate them as fossil genes. The first is located between nucleotides 27 211 and 27 755. It is a derivative of a mimivirus gene, MIMI\_R865, that belongs in the *Acanthamoeba polyphaga* mimivirus genome to a family of gene paralogues to which also belong genes MIMI\_L17, MIMI\_R298 and MIMI\_L754. The second fossil is located between nucleotides 55 260 and 55 883. It is related to the

gene CIV261R and has a functional paralogue in the IIV31 genome, 051R. The third fossil is located between nucleotides 190 307 and 191 099. It was determined as a member of the family *bro*-like genes.

The presence of certain mobile genetic elements that occur in some NCLDVs belonging to the families *Phycodnaviridae*, *Mimiviridae* and *Ascoviridae* was searched for in the IIV31 genome (Desnues *et al.*, 2012; Bao & Jurka, 2013). No transpovirons, group I introns and *Fanzor1* or *Fanzor2* DNA transposons were found. In contrast to IIV6, IIV9, IIV22, IIV22A, IIV25 and IIV30, no inteins were found inserted into the ORF 097R of IIV31, which encodes the alpha subunit of the ribonucleotide-diphosphate reductase. However, one intein was found to be specifically inserted in-frame in 001R, which encodes the delta DNA polymerase, as reported previously (Bigot *et al.*, 2013).

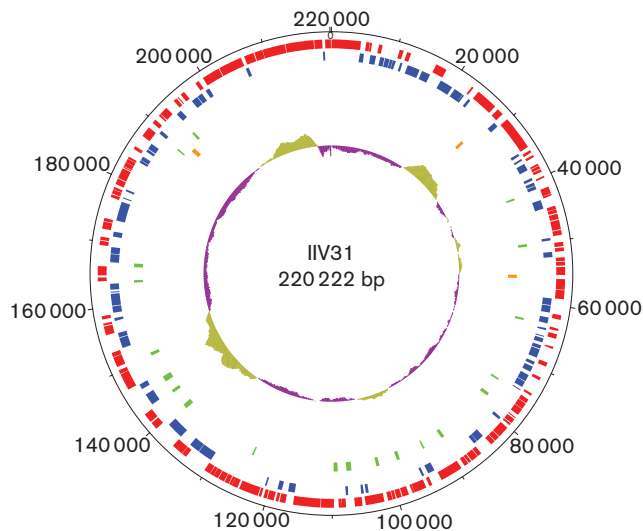
A miniature transposon (MITE) was found between nucleotides 38 829 and 39 878. This 1049 bp MITE was named IIV31-MITE. We think that this is a derivative of a class II transposon (Wicker *et al.*, 2007). Indeed, both of its extremities correspond to inverted terminal repeats (ITRs) about 371 bp in length that are juxtaposed at their outer ends by a CTAG tetranucleotide that corresponds to a

**Table 2.** Genome statistics of IIV31

Attribute	Value	% of total*
Size (bp)	220 222	100
G + C content (bp)	77 276	35.09
Coding region (bp)	189 465	86.0
Pseudogenes	3	100
Total genes (putatively functional)	203	100
Protein-encoding genes with function prediction	84	41.3
Protein-encoding genes with orthologues in databases	–	–
Family of gene paralogues	3	–
Genes in families of paralogous genes	21	10.4
Non-coding regions over 200 bp in length	10 788 (20 segments)	4.9

\*The total is based on either the size of the genome in bp or the total number of protein-encoding genes in the annotated genome, where applicable.





**Fig. 1.** Circular map of the 220 222 bp IIV31 genome. The outer scale is numbered clockwise in bp. Circles 1 and 2 (from outside to inside) denote CDSs (forward strands in red and reverse strands in blue). Green boxes in circle 3 represent ORF-free regions with a size of over 200 bp. The orange boxes in circle 4 represent the three fossil genes. Circle 5 represents the local variations of G+C content along the genome sequence (green, regions with a GC content above the average; purple, regions with a GC content below the average).

target site duplication (TSD), which occurred at the time of the MITE insertion (Fig. 2). Its inner regions contain two repeat motifs that are 39 and 10 bp in length, each repeated four times. In the IIV31 genome, this MITE overlapped two small paralogous CDSs, 034R and 035L (Table S1), that had no orthologue in other sequenced iridoviruses.

The nucleotide sequence comparison of IIV31-MITE with IIV9-MITE and IIV22-MITE revealed that they were not related. Interestingly, the TSD and the ITR length of

the IIV31-MITE match those of a prokaryotic family of insertion sequence, IS5. An IS5-like transposon with similar TSD and ITR features, IS5\_Av, has already been described in the genome of a bdelloid rotifer, *Adineta vaga* (Gladyshev & Arkhipova, 2009). Here, we propose that IIV31-MITE and IS5\_Av could represent the only two members of a currently unknown eukaryotic family of transposable elements that would be related to the IS5 prokaryotic family. However, no sequence data for the transposase that mediated the transposition of IIV31-MITE are available to confirm this proposition.

IIV31 is, to our knowledge, the first iridovirus genome infecting a crustacean species to be sequenced and reported. This genome revealed 19 new putative proteins with sizes varying from 105 to 435 aa residues. Many (184) of the CDSs identified displayed high conservation with their counterparts in other IIVs, insect and bacterial genomes. Further sequencing of related strains will reveal more about the genetic and functional diversity of these interesting viruses.

The discovery of IIV31-MITE, together with that of IIV09-MITE and IIV22-MITE, and of *Fanzor1* and 2 DNA transposons (Bao & Jurka, 2013) in two ascoviruses, SfAV1a and HvAV3e, suggests that these closely related invertebrate viruses (Stasiak *et al.*, 2003; Bigot *et al.*, 2009) might be, a priori, vectors of horizontal transfers of transposable elements between host species. However, their presence in these genomes must be considered as unexpected for two reasons. First, the cellular niche of these viruses is mostly cytoplasmic, whereas that of the DNA transposons is only nuclear. Second, the genome of these viruses has a molecular configuration in which the DNA is not negatively supercoiled or is negatively supercoiled to only a small extent (Bigot *et al.*, 2000). Such genome configurations are incompatible with some needs of DNA transposons. Indeed, DNA transposons need a negatively supercoiled DNA environment for an efficient mobility at the excision and the insertion sites (Sinzelle

IIV31-MITE  
 38820 - gaaga CTAGAAATGGCAGATACAAAGATTGAAGATTTAAAGGGTGTGGCCCTTATG  
 ACACCAGAAGAGTCAAAAATAATATTATAAGAAAGGAAAAGGATATCTCGCTAGCATAGT  
 GGTAGGAGGAAACAGTCTCCGAAAGAAATAGATGAAGCCAACATTAAGACATGATATACT  
 TTTTCTTTTCTTTGGAGGAGATGGGGATCTCAGAACATGTTAAGGATCCACGCCATAAA  
 AACACATTAATTAATACTCTCATTTGACACCGTTGTCAAATCGCTAAAAATGGATAAAGACAG  
 CGATTACATTCAAATGCTACaAAACAACAGAAAACCCACTCTTTAAGGCACAGAGTAAGAA  
 TGCCGATTAACGACCTCGGACTCCGTCTTTTACTGGACTACGACGTTTAGGACTCGGGCTCC  
 TCCTTCTAACAGGACTACGACGTTTAGGACTCGGACTTCTCCTTCTTACGGGACTACGACGT  
 TTAGGACTCGGACTTCTCCTTCTTACGGGACTTCCCGTTTAGGACTCGGACTACGCTTTTT  
 AATTCTGTATAAAAATTGGCATTATTTATAAAAATTAATTTAAAAATTTAATGGATCGATC  
 CATTAAAAACTCGTTAATGCATTCTAACGATACTATCTCATTATCTTTGATGGCTTTATGT  
 ACCTTTTCTTCTCGTTAATCGGCATTCTTACTCTGTGCCTTAAAGAGTGGGTTTTCTGTT  
 GTTTGTAGCATTGAAATGTAATCGCTGTCTTTATCCAATTTTAGCGATTTGACAACGGTGT  
 CAATGAGAGTTTTAATTAATGTGTTTTTATGGCGTGGATCCTTAACATGTTCTGAGATCCC  
 CCATCTCCTCAAAGAAAAGAAAAAaTATATCATGTCTTTAATGTTGGCTTCATCTATTTC  
 TTCGGGAGCACTGTTTCTCCTACCCTATGCTAGCGAGATATCCTTTTCCCTTCTTAATAA  
 TATTATTTTGACTTCTTCTGGTGTCAATAAGGGCAACACCCTTTAAATCTTCAATCTTTGTA  
 TCTGCCATTTCTGA aatgc - 39887

**Fig. 2.** Nucleotide sequence of the IIV31-MITE that was found in the region spanning nucleotides 38 829 and 39 878. ITRs are highlighted in black with white text, and the duplicated CTGA tetranucleotide at the insertion site is in bold type. The four 39 bp motifs that are tandemly repeated in the inner region are in italic type and alternately highlighted in light or dark grey. The four TATAAAAATT motifs are underlined.

*et al.*, 2008; Crénès *et al.*, 2009; Claeys Bouuaert & Chalmers, 2013). Given these requirements, DNA transposons are expected to be found in the genomes of nuclear viruses such as baculoviruses (van Oers & Vlak, 2007), nudiviruses and hytrosaviruses, but not in those of cytoplasmic viruses. As a consequence, an alternative interpretation is to suggest that the presence of the DNA transposon in these viruses would have a function under a selection pressure, which would be advantageous for the viruses bearing them in their genome. This might find support based on two observations. First, DNA transposons and MITEs are repeated elements that are interspersed in eukaryotic genomes, including introns, and 5' and 3' regions of genes that are transcribed in mRNA, but are not translated. Due to the presence of ITRs at their ends, DNA transposon and MITE transcripts can assemble in intra-strand dyad structures (Petit *et al.*, 2007), i.e. short hairpin RNA (shRNA) molecules that are good substrates for the RNA interference machineries. Second, IIVs and ascoviruses were found to encode enzymes able to manipulate the host RNA interference machineries (Bigot *et al.*, 2009; Hussain *et al.*, 2010; Wong *et al.*, 2011). Since all the genome of IIVs and ascoviruses seems to be transcribed during the viral cycle (D'Costa *et al.*, 2004), our hypothesis is that shRNA transcripts from DNA transposon and MITE provide another way used by these viruses to modify the expression of some host genes by RNA interference. Further sequencing of the genome of their hosts will be required to verify the consistency of this hypothesis.

## Acknowledgements

This work was supported by the C.N.R.S. and GENOSCOPE, and funded by the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie and the Groupement de Recherche CNRS 3546.

## References

- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S. & other authors (2000). Gene ontology: tool for the unification of biology. *Nat Genet* **25**, 25–29.
- Bao, W. & Jurka, J. (2013). Homologues of bacterial TnpB\_IS605 are widespread in diverse eukaryotic transposable elements. *Mob DNA* **4**, 12.
- Bideshi, D. K., Renault, S., Stasiak, K., Federici, B. A. & Bigot, Y. (2003). Phylogenetic analysis and possible function of *bro*-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria. *J Gen Virol* **84**, 2531–2544.
- Bigot, Y., Stasiak, K., Rouleux-Bonnin, F. & Federici, B. A. (2000). Characterization of repetitive DNA regions and methylated DNA in ascovirus genomes. *J Gen Virol* **81**, 3073–3082.
- Bigot, Y., Renault, S., Nicolas, J., Moundras, C., Demattei, M. V., Samain, S., Bideshi, D. K. & Federici, B. A. (2009). Symbiotic virus at the evolutionary intersection of three types of large DNA viruses; iridoviruses, ascoviruses, and ichnoviruses. *PLoS ONE* **4**, e6397.
- Bigot, Y., Asgari, S., Bideshi, D. K., Cheng, X., Federici, B. A. & Renault, S. (2011). Family Ascoviridae. In *Viral Taxonomy, IX Report of the International Committee on the Taxonomy of Viruses*, 3rd edn.
- Edited by A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz. London: Elsevier–Academic Press.
- Bigot, Y., Piégu, B., Casteret, S., Gavory, F., Bideshi, D. K. & Federici, B. A. (2013). Characteristics of inteins in invertebrate iridoviruses and factors controlling insertion in their viral hosts. *Mol Phylogenet Evol* **67**, 246–254.
- Claeys Bouuaert, C. & Chalmers, R. (2013). *Hsmar1* transposition is sensitive to the topology of the transposon donor and the target. *PLoS ONE* **8**, e53690.
- Cole, A. & Morris, T. J. (1980). A new iridovirus of two species of terrestrial isopods, *Armadillidium vulgare* and *Porcellio scaber*. *Intervirology* **14**, 21–30.
- Colson, P., De Lamballerie, X., Yutin, N., Asgari, S., Bigot, Y., Bideshi, D. K., Cheng, X.-W., Federici, B. A., Van Etten, J. L. & other authors (2013). “Megavirales”, a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch Virol* **158**, 2517–2521.
- Crénès, G., Ivo, D., Hérisson, J., Dion, S., Renault, S., Bigot, Y. & Petit, A. (2009). The bacterial Tn9 chloramphenicol resistance gene: an attractive DNA segment for *Mos1 mariner* insertions. *Mol Genet Genomics* **281**, 315–328.
- D'Costa, S. M., Yao, H. J. & Bilimoria, S. L. (2004). Transcriptional mapping in Chilo iridescent virus infections. *Arch Virol* **149**, 723–742.
- Delhon, G., Tulman, E. R., Afonso, C. L., Lu, Z., Becnel, J. J., Moser, B. A., Kutish, G. F. & Rock, D. L. (2006). Genome of invertebrate iridescent virus type 3 (mosquito iridescent virus). *J Virol* **80**, 8439–8449.
- Desnues, C., La Scola, B., Yutin, N., Fournous, G., Robert, C., Azza, S., Jardot, P., Monteil, S., Campocasso, A. & other authors (2012). Provirophages and transpovirons as the diverse mobilome of giant viruses. *Proc Natl Acad Sci U S A* **109**, 18078–18083.
- Eaton, H. E., Metcalf, J., Penny, E., Tcherepanov, V., Upton, C. & Brunetti, C. R. (2007). Comparative genomic analysis of the family *Iridoviridae*: re-annotating and defining the core set of iridovirus genes. *Virology* **361**, 11.
- Federici, B. (1980). Isolation of an iridovirus from two terrestrial isopods, the pill bug, *Armadillidium vulgare*, and the sow bug, *Porcellio dilatatus*. *J Invertebr Pathol* **36**, 373–381.
- Gladyshev, E. A. & Arkhipova, I. R. (2009). A single-copy IS5-like transposon in the genome of a bdelloid rotifer. *Mol Biol Evol* **26**, 1921–1929.
- Henn, M. R., Sullivan, M. B., Stange-Thomann, N., Osburne, M. S., Berlin, A. M., Kelly, L., Yandava, C., Kodira, C., Zeng, Q. & other authors (2010). Analysis of high-throughput sequencing and annotation strategies for phage genomes. *PLoS ONE* **5**, e9083.
- Hussain, M., Abraham, A. M. & Asgari, S. (2010). An ascovirus-encoded RNase III autoregulates its expression and suppresses RNA interference-mediated gene silencing. *J Virol* **84**, 3624–3630.
- Iyer, L. M., Aravind, L. & Koonin, E. V. (2001). Common origin of four diverse families of large eukaryotic DNA viruses. *J Virol* **75**, 11720–11734.
- Jakob, N. J., Müller, K., Bahr, U. & Darai, G. (2001). Analysis of the first complete DNA sequence of an invertebrate iridovirus: coding strategy of the genome of Chilo iridescent virus. *Virology* **286**, 182–196. (PMID: 11448171).
- Jancovich, J. K., Chinchar, V. G., Hyatt, A., Miyazaki, T., Williams, T. & Zhang, Q. Y. (2011). Family *Iridoviridae*. In *Viral Taxonomy, IX Report of the International Committee on the Taxonomy of Viruses*, 3rd edn. Edited by A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz. London: Elsevier–Academic Press.
- Petit, A., Rouleux-Bonnin, F., Lambelé, M., Pollet, N. & Bigot, Y. (2007). Properties of the various *Botmar1* transcripts in imagoes of the bumble bee, *Bombus terrestris* (Hymenoptera: Apidae). *Gene* **390**, 52–66.

- Piégu, B., Guizard, S., Spears, T., Cruaud, C., Couloux, A., Bideshi, D. K., Federici, B. A. & Bigot, Y. (2013a). Complete genome sequence of invertebrate iridescent virus 22 isolated from a blackfly larva. *J Gen Virol* **94**, 2112–2116.
- Piégu, B., Guizard, S., Tan, Y., Cruaud, C., Couloux, A., Bideshi, D. K., Federici, B. A. & Bigot, Y. (2013b). Complete genome sequence of invertebrate iridovirus IIV22A, a variant of IIV22, isolated originally from a blackfly larvae. *Stand Genomic Sci* **9**, 3.
- Piégu, B., Guizard, S., Spears, T., Cruaud, C., Couloux, A., Bideshi, D. K., Federici, B. A. & Bigot, Y. (2013c). Complete genome sequence of invertebrate iridovirus IIV-25 isolated from a blackfly larva. *Arch Virol* (in press).
- Piégu, B., Guizard, S., Spears, T., Cruaud, C., Couloux, A., Bideshi, D. K., Federici, B. A. & Bigot, Y. (2013d). Complete genome sequence of invertebrate iridovirus IIV30 isolated from the corn earworm, *Helicoverpa zea*. *J Invertebr Pathol* **116**, 43–47.
- Sinzelle, L., Jégot, G., Brillet, B., Rouleux-Bonnin, F., Bigot, Y. & Augé-Gouillou, C. (2008). Factors acting on *Mos1* transposition efficiency. *BMC Mol Biol* **9**, 106.
- Stasiak, K., Renault, S., Demattei, M. V., Bigot, Y. & Federici, B. A. (2003). Evidence for the evolution of ascoviruses from iridoviruses. *J Gen Virol* **84**, 2999–3009.
- van Oers, M. M. & Vlak, J. M. (2007). Baculovirus genomics. *Curr Drug Targets* **8**, 1051–1068.
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M. & other authors (2007). A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* **8**, 973–982. .
- Wijnhoven, H. & Berg, M. P. (1999). Some notes on the distribution and ecology of Iridovirus (Iridovirus, Iridoviridae) in terrestrial isopods (Isopoda, Oniscidae). *Crustaceana* **72**, 145–156.
- Williams, T. (1994). Comparative studies of iridoviruses: further support for a new classification. *Virus Res* **33**, 99–121.
- Williams, T. (2008). Natural invertebrate hosts of iridoviruses (Iridoviridae). *Neotrop Entomol* **37**, 615–632.
- Williams, T. & Cory, J. S. (1994). Proposals for a new classification of iridescent viruses. *J Gen Virol* **75**, 1291–1301.
- Wong, C. K., Young, V. L., Kleffmann, T. & Ward, V. K. (2011). Genomic and proteomic analysis of invertebrate iridovirus type 9. *J Virol* **85**, 7900–7911.
- Yan, X., Olson, N. H., Van Etten, J. L., Bergoin, M., Rossmann, M. G. & Baker, T. S. (2000). Structure and assembly of large lipid-containing dsDNA viruses. *Nat Struct Biol* **7**, 101–103.