Polydnaviruses: Evolution and Function

Michael R. Strand* and Gaeleen R. Burke

Department of Entomology, University of Georgia, Athens, GA, USA.

*Correspondence: mrstrand@uga.edu

https://doi.org/10.21775/cimb.034.163

Abstract
Polydnaviruses (PDVs) were originally viewed as large DNA viruses that are beneficial symbionts of parasitoid wasps. Two groups of PDVs were also recognized: bracoviruses (BVs), which are associated with wasps in the family Braconidae, and ichnoviruses (IVs), which are associated with wasps in the family Ichneumonidae. Results to date indicate that BVs are endogenous virus elements (EVEs) that evolved from an ancient betanudivirus. IVs are also likely EVEs but are unrelated to BVs. BVs and IVs are very unusual relative to most known EVEs because they retain many viral functions that benefit wasps in parasitizing hosts. However, BVs and IVs cannot be considered beneficial symbionts because all components of their genomes are fixed in wasps. Recent studies indicate that other nudiviruses have endogenized in insects. Each exhibits a different functional fate from BVs but shares certain architectural features. We discuss options for classifying BVs and other endogenized nudiviruses. We also discuss future directions.

Introduction
Parasitoids are defined as free living insects as adults that feed parasitically on other arthropods as larvae (Pennacchio and Strand, 2006). Adult females usually lay their eggs in (endoparasitoids) or on (ectoparasitoids) a host, which is usually another insect species, and hosts usually die after offspring complete their development. The vast majority of parasitoids are wasps (order Hymenoptera) that have specialized to parasitize only one or a few host species. Host range specialization in turn is associated with high speciation rates; estimates suggest more than 25% of all insect species are parasitic Hymenoptera (Godfray, 1994; Forbes, et al., 2018).

The Polydnaviridae was recognized in 1991 as a family of large double-stranded DNA viruses associated with endoparasitoid wasps in the Braconidae and Ichneumonidae families (Francki, 1991). Braconid-associated polydnaviruses (PDVs) were assigned to the genus Bracovirus (BV) and ichneumonid-associated PDVs were assigned to the genus Ichnovirus (IV). Subsequent studies indicated that all BV-carrying braconids reside in six subfamilies that form a monophyletic and hyperdiverse (≈ 50,000 species) assemblage named the microgastroid complex, while all IV-carrying ichneumonids reside in two subfamilies named the Campopleginae and Banchinae (≈ 14,000 species) (Whitfield, 2002; Quicke et al., 2009). Both BV and IV-carrying wasps primarily parasitize larval stage Lepidoptera, with several also being used as biological control agents for management of pest species. However, the primary interest in these insects stemmed from the finding that each wasp species relies on its PDV to successfully parasitize hosts, while each PDV relies on its wasp for transmission. Thus, the association between PDVs and wasps was viewed as the first example of viruses evolving into beneficial symbionts (Edson et al., 1981).

We wrote a contribution for the first edition of Insect Virology that summarized the PDV literature through 2009 (Strand, 2010). Several findings, reviews and essays have since been published that substantially advanced understanding of PDVs (Beckage and Drezen, 2012; Strand and Burke, 2013, 2014, 2015; Gundersen-Rindal et al., 2013;
Herniou et al., 2013; Drezen et al., 2017; Gauthier et al., 2018). In this chapter, we first summarize the early literature that led to the view of PDVs as beneficial symbionts of wasps. We next discuss results that shed important light on PDV evolution and function. We conclude by summarizing recent results showing that viruses related to the ancestor of BVs have independently associated themselves with other insects, including parasitoids outside of the microgastrine complex.

**The early literature concluded that PDVs are beneficial symbionts of wasps**

PDVs were first discovered in the late 1960s and early 1970s through electron microscopy (EM) studies that identified large numbers of particles in the ovaries and lateral oviducts of certain braconid and ichneumonid wasps (Rotheram, 1967; Vinson and Scott, 1975; Stoltz et al., 1976). Mature particles in some braconid wasp species had one cylindrical capsid of similar width ($\approx 40$ nm) but variable length ($25–100$ nm) surrounded by a single envelope, while in others multiple capsids were surrounded by a single envelope (Stoltz and Vinson, 1979). Mature particles in ichneumonid wasps had biconvex capsids ($\approx 85–330$ nm) that were individually surrounded by two envelopes (Stoltz and Vinson, 1979). While morphologically distinct, braconid and ichneumonid particles both exhibited morphological features that suggested they were viruses. Molecular data further indicated that particles from both braconids and ichneumonids contained multiple, circular, dsDNAs that were non-equimolar in abundance (Webb and Strand, 2005). The estimated number of DNA segments in particles (20 to more than 100), their individual sizes (2 to more than 30 kb) and their estimated aggregate sizes (150 to more than 700 kb) varied between wasp species, with particles from braconids usually contain a smaller number of larger segments than ichneumonids (Krell and Stoltz, 1979; Krell et al., 1980, 1982; Espagne et al., 2004; Webb et al., 2006; Lapointe et al., 2007). Each capsid in mature particles from braconids further contained only one segment (Allbrecht et al., 1994; Beck et al., 2007). This indicated that the total number of DNA segments detected in particles from a given wasp species are distributed across a population of particles and that the non-equimolar abundance of segments is due to particles containing certain segments being more abundant than others (Beck et al., 2007).

Early studies also identified several shared life cycle features. First, sequences homologous to the DNA segments in particles were detected in the genomes of both male and female wasps across all life stages and cell types including the germ line (Stoltz, 1990; Fleming and Summers, 1991). These results suggested the DNAs in braconid and ichneumonid particles are integrated into the genome of wasps. They further suggested these entities persist as proviruses that can be transmitted vertically through the germline of wasps (Fleming, 1992). EM studies also provided strong, visual evidence that braconid and ichneumonid particles replicate in pupal and adult stage females in specialized cells of the proximal ovaries named calyx cells (Stoltz and Vinson, 1979; Wyler and Lanzrein, 2003). Braconid and ichneumonid-associated particles both assembled in calyx cell nuclei. Braconid particles were released by lysis of calyx cells. In contrast, ichneumonid particles were released by budding, which also resulted in particles acquiring a second envelope (Stoltz and Vinson, 1979; Norton and Vinson, 1983). Mature particles accumulated to high densities in the lumen of the proximal ovaries and lateral oviducts of female wasps, where they formed a thick paste called calyx fluid (Stoltz and Vinson, 1979). Female wasps were also shown to inject calyx fluid into the hosts they parasitized, and that particles rapidly infect haemocytes, fat bodies and other tissues (Stoltz and Vinson, 1979; Strand et al., 1992; Strand, 1994; Li and Webb, 1994). After infecting a host cell, capsids migrated to nuclear pores followed by discharge of encapsidated DNAs into the nucleus and expression of genes encoded on encapsidated DNAs (Stoltz and Vinson, 1979; Fleming et al., 1983; Blissard et al., 1986). Functional assays further indicated that infection of hosts disabled immune defences that would otherwise kill wasp offspring and altered growth and metabolism in ways that prevented hosts from surviving (Edson et al., 1981; Stoltz and Cook, 1983; Davies et al., 1987; Dover et al., 1987). In contrast, no replication of braconid and ichneumonid particles occurred in infected hosts (Theilmann and Summers, 1986).

Altogether, particle morphology, the packaging
of gene-encoding dsDNAs into particles, and the ability of particles to infect and express genes in the hosts of wasps were identified as key virus-like features. The life cycle also showed parallels with known viruses that can persist as proviruses (Retroviridae, several bacteriophages) or that can be transmitted by insects to other hosts (Togaviridae, Bunyaviridae, Reoviridae). However, no other known viruses had large segmented circular dsDNA genomes or exhibited traits consistent with being beneficial symbionts of one host (wasps) and pathogens of another (hosts of wasps). Furthermore, no other viruses were known to only be transmitted vertically through one host (wasps), replicate in the same host to produce particles that infect but never replicate in another host (lepidopteran larvae). Thus, the family Polydnaviridae, genus Bracovirus (BV) and genus Ichnovirus (IV) were proposed (Stoltz et al., 1984) and later recognized (Francki, 1991) as mutualistic viruses in braconid and ichneumonid wasps that have large ‘polyDNA’ genomes.

Subsequent phylogenetic studies further refined the boundaries of BV and IV wasp associations by showing that: (1) BVs are restricted to species in six subfamilies that form the microgastrid complex, (2) the microgastrid complex diverged approximately 100 million years ago from other subfamilies of braconids that lack BVs, and (3) IVs are restricted to species in two subfamilies of the Ichneumonidae: the Campopleginae and Banchinae (Whitfield, 2002; Murphy et al., 2008; Thézé et al., 2011; Quicke et al., 2009) (Fig. 8.1).

Higher order data also indicated that BV-carrying braconids and IV-carrying ichneumonids belong to the same superfamily (Ichneumonoidea) but are distantly related (Quicke et al., 2009; Heraty et al., 2011) (Fig. 8.1). Taken together, these results strongly suggested that all BV-carrying species evolved from a common ancestor, the BV-microgastrid association is ancient, and BVs likely evolved independently from IVs (Fig. 8.1). In contrast, these data do not address what BVs and IVs evolved from, or whether the ancestors of BVs

![Figure 8.1](caister.com/cimb)
and IVs are related. The uncertain phylogenetic relationship between campoplegine and banchine ichneumonids also made it unclear whether all IVs share a common ancestor.

**Key insights into PDV evolution and function required genomic data from wasps**

Microbes within the Eubacteria, Archaea, and Eukarya each share molecular machinery and genes that can be used as ‘barcodes’ for discerning ancestry and assigning species/isolates to particular taxa. In contrast, multiple evolutionary origins result in no molecular machinery or genes that are universally shared among viruses (Forterre, 2006; Holmes, 2011a). Different lineages of viruses can be classified into families and sometimes higher order associations through core genes that are shared among certain isolates. Most RNA and DNA viruses with small genomes (<20 kb) have gene inventories consisting entirely or mostly of core genes (Payne, 2017). Large DNA viruses, on the other hand, encode a small proportion of core genes of ancient and usually uncertain origin and a much larger proportion of non-core genes that are present in only some isolates or are unique to particular isolates (McGeoch and Davison, 1999; Iyer et al., 2006; Holmes, 2011a; Garavaglia et al., 2012). This is because large DNA viruses, like prokaryotes, have over evolutionary time acquired a diversity of genes from hosts and other organisms by horizontal gene transfer (HGT) (Shackelton and Holmes, 2004). In addition, large DNA viruses have sufficient coding capacity to maintain genes acquired by HGT through selection if they confer advantageous functions (Holmes, 2011a). Core genes in large DNA viruses often include structural and other genes with essential functions in replication, while non-core genes are often virulence-related, enabling isolates to infect and successfully propagate in particular hosts (McGeoch and Davison, 1999; Iyer et al., 2006; Garavaglia et al., 2012). Very little sequence data existed for any PDV when they were first classified as mutualistic viruses due to several technical challenges (Strand and Burke, 2015). Furthermore, the data needed to discern ancestry and key functions of PDVs took many years to generate. Acquisition of this information occurred in three major steps.

**The DNA segments in BV and IV virions contain virulence genes with functions in parasitizing the hosts of wasps**

The first step was accomplished in the 2000s when methodology progressed to the point that the DNA segments in BV virions from five species of braconid wasps showed strong evidence of being wasp specific BVs (Desjardins et al., 2006; Webb et al., 2006; Desjardins et al., 2008) and IV virions from four species of ichneumonid wasps (Webb et al., 2006; Tanaka et al., 2007; Lapointe et al., 2007) could be sequenced and assembled. The DNAs in select other PDVs were also later partially or fully sequenced (Chen et al., 2011; Yu et al., 2016). These results showed that coding densities for the DNAs in BV and IV virions were low and were more similar to eukaryotic genomic regions than to the genomes of other known large DNA viruses. Identified genes were either single copy or formed multimember families; some contained small introns, while others were intronless. Several of these genes shared homology with known eukaryotic genes but no genes in either BV or IV virions share homology with any known viral core gene. BVs from closely related wasps shared several genes with one another as did IVs from closely related wasps. BVs and IVs from more distantly related wasps shared fewer genes with one another (Webb et al., 2006; Tanaka et al., 2007; Lapointe et al., 2007; Pruijssers and Strand, 2007; Desjardins et al., 2008; Serbielle et al., 2012). However, only a few types of genes were shared between BVs and campoplegine IVs [ankyrin repeat genes (ank)] or banchine ichneumonids [ank and protein tyrosine phosphatase genes (ptps)]. Analysis of the genes in BVs and IVs further support historical acquisition at different times and from different sources (Huguet et al., 2012). For example, some genes unique to particular BVs show strong evidence of being wasp genes (Desjardins et al., 2008; Burke and Strand, 2014). Other genes present in some or most isolates of BVs or IVs show evidence of acquisition from other viruses or other types of organisms or are of ancient enough lineage to be of indeterminate origin (Huguet et al., 2012; Burke and Strand, 2012a; Serbielle et al., 2012; Herniou et al., 2013). Genes like the ank families in BVs and campoplegine IVs, and ptps in BVs and banchine IVs show evidence of independent acquisition
rather than common ancestry (Lapointe et al., 2007; Huguet et al., 2012).

From a functional perspective these results indicate that: (1) BVs and IVs do not replicate in the hosts of wasps because no genes with discernible functions in replication are present on the DNAs in virions, and (2) a number of genes likely function as virulence factors with roles in parasitism. Further support for the latter came from studies showing that most genes on DNAs in virions are transcribed in hosts but are not transcribed in wasps (Rasoolizadeh et al., 2009; Bitra et al., 2011; Provost et al., 2011; Burke and Strand, 2012b; Chevignon et al., 2014; Doremus et al., 2014; Bitra et al., 2015). Experimental studies also showed that genes/microRNAs in BV or IV virions disrupt specific host immune defences against wasp offspring (Beck and Strand, 2005; Theoetkattikul et al., 2005; Beck and Strand, 2007; Prijssers and Strand, 2007; Ibrahim and Kim, 2008; Labropoulou et al., 2008; Kwon and Kim, 2008; Cooper et al., 2011; Magkrioti et al., 2011; Bitra et al., 2012; Gueguen et al., 2013), alter host growth (Provost et al., 2004; Falabella et al., 2006; Prijssers et al., 2009; Kim et al., 2013; Prasad et al., 2014; Ignesti et al., 2018; Wang et al., 2018), and potentially affect host interactions with food plants and other organisms (Zhu et al., 2018; Tan et al., 2018).

From an evolutionary perspective, the paucity of shared genes between BVs and IVs was consistent with phylogenetic studies of wasps that suggested BVs and IVs arose independently. The diverse origins of genes in BV and IV virions were also similar to the variation in non-core genes that other families of large DNA viruses exhibit. In contrast, the absence of any recognizable core genes suggested that BVs and IVs are either not of viral origin or that core genes with functions in producing virions reside in the genomes of wasps (Stoltz, 1993; Federici and Bigot, 2003).

**PDV core genes reside in the genomes of wasps**

The second step in characterizing PDV genomes was analysing proteins in virions and determining whether the corresponding genes were present in the genomes of wasps and if they were related to core genes in known viruses. The first studies were conducted in the campoplegine ichneumonid wasp *Campoletis sonorensis* (Fig. 8.1), which produces C. *sonorensis ichnovirus* (CsIV). N-terminal sequencing of a protein (p44) in CsIV virions enabled identification of the corresponding cDNA, which was detected in calyx cells and hybridized to wasp genomic DNA but did not hybridize to DNAs in CsIV virions (Deng et al., 2000). These results indicated that p44 resides in the *C. sonorensis* genome but did not provide any evidence for viral ancestry because the genes share no homology with any known gene.

Scaling by proteomic analysis of virions, Sanger sequencing of ovary cDNAs, and sequencing bacterial artificial chromosome (BAC) genomic libraries yielded the first evidence that viral core genes reside in the genomes of two microgastrine braconids, *Cotesia congregata* and *Chelonus inanitus* (Fig. 8.1), that produce the BVs CcBV and CiBV (Bézier et al., 2009; Wetterwald et al., 2010). Proteomic analysis of MdBV virions from *Microplitis demolitor* (Fig. 8.1) together with deep sequencing of wasp ovaries using the Illumina platform further identified the full set of viral core genes identifiable by homology (Burke and Strand, 2012b; Burke et al., 2013). Altogether, these studies identified 42 genes that are specifically upregulated in wasp ovaries when BV virions are being produced and that share significant homology with core genes in two families of large DNA viruses: the Nudiviridae and Baculoviridae (see Chapter 9). This finding was intuitively appealing because nudiviruses and baculoviruses both infect insects (Rohrmann, 2013) and because nudiviruses produce virions that are morphologically similar to BVs (Stoltz and Vinson, 1979).

All baculoviruses and nudiviruses have genomes that consist of a single large circular dsDNA. Six nudiviruses have been sequenced which share 33 genes with one another, forming a candidate core gene set (Wang and Jehle, 2009; Jehle et al., 2013; Bézier et al., 2015). Results further place these nudiviruses into two groups: the alpha- and betanudivirus genera (Bézier et al., 2015) (Fig. 8.2). Of the 33 candidate nudivirus core genes, 21 are also shared with all baculoviruses (Garavaglia et al., 2012; Rohrmann, 2013), which supports that the Nudiviridae and Baculoviridae are sister families (Jehle et al., 2013; Bézier et al., 2015; Burke et al., 2018a) (Fig. 8.2). None of the candidate core genes in nudiviruses have been functionally studied, but those shared with baculoviruses have
been functionally studied in model isolates like Autographa californica multiple nucleopolyhedrovirus (AcMNPV) (Rohrmann, 2013) (Fig. 8.2). These include a DNA polymerase (dnapol) and helicase (helicase) required for replicating the viral genome, four subunits (lef-4, lef-8, lef-9, p47) of a DNA dependent RNA polymerase, and structural genes such as vp39, vlf-1, p74 and pif-1 present in virions and containing promoter sequences specifically recognized by viral RNA polymerase (Rohrmann, 2013).

Orthologues to 21 of the core genes shared between nudiviruses and baculoviruses were identified in the M. demolitor ovary transcriptome data set and partially identified in C. congregata (Bézier et al., 2009; Burke and Strand, 2012b). The only
omission in *M. demolitor* was *dnapol*, which suggests that amplification of the DNAs packaged into BV virions requires wasp replication machinery (Burke and Strand, 2012b). In contrast, the presence of all of the RNA polymerase subunits and structural genes suggested that BV virion assembly is conserved with nudiviruses and baculoviruses. Further support for this latter conclusion comes from: (1) detection of all nudivirus-like structural gene products in MdBV virions and partial identification of these products in CcBV and CiBV (Bézier et al., 2009; Wetterwald et al., 2010; Burke et al., 2013), and (2) RNAi experiments which show that the RNA polymerase subunit genes produce a functional enzyme in *M. demolitor* that transcribes BV structural genes but not wasp genes (Burke et al., 2013). RNAi knockdown of *vp39*, *p74*, and *pif-1* also indicates each is required for producing normal virions (Burke et al., 2013). A tyrosine recombinase named *integrate* (*int*) related to baculovirus *vlf-1* is present in all nudiviruses and BV-carrying wasps, which suggests *int* genes are also part of a BV core gene set. Functional studies further implicate *int-1* and *vlf-1* in *M. demolitor* in processing of the DNAs that are packaged into MdBV virions (Burke et al., 2013, 2015).

Proteomic analysis of virions, sequencing of ovary cDNAs and sequencing of genomic BAC clones has also been conducted in two campoplegine (*Hyposoter didymatator* and *Transoema rostrale*) and one banchine ichneumonid wasp (*Glypta fumiferanae*) that produce HdIV, TrIV, and GfIV, respectively (Volkoff et al., 2010; Bélieuves et al., 2015) (Fig. 8.1). Volkoff et al. (2010) identified several structural genes in HdIV and TrIV. None shared homology with structural genes in any known virus, but in *H. didymatator* these genes reside in clusters of 9–16 genes in three regions of genomic DNA, ichnovirus structural protein encoding regions (IVSPERs), that suggest a viral origin (Volkoff et al., 2010). Comparisons to GfIV further indicated the inventory of IVSPER-associated structural genes were very similar to HdIV and TrIV, which supports the presumption that campoplegine and banchine IVs have similar ancestors (Béliveau et al., 2015). Altogether, these findings strongly support the conclusion that IVs are not related to BVs, but whether the IV ancestor is an extinct or undiscovered virus remains unknown.

### The genes and DNA segments required to produce BVs are dispersed in the genomes of wasps

The third step in genome analysis of PDVs was determining how the DNA segments are packaged into virions and how the core genes required for virion formation are organized. Currently, most insights derive from BV-carrying braconids. Sequencing of bacterial artificial chromosome (BAC) libraries against four species of braconids (*Glypta pantalees indensis, G. flavicoxis, C. congregata, C. sesamiae*), plus whole genome sequencing of *M. demolitor*, show that the DNA segments in corresponding virions (GiBV, GfBV, CcBV, CsBV, and MdBV) reside in 6–8 ‘proviral’ loci as integrated, linear DNAs (Desjardens et al., 2007, 2008; Bézier et al., 2013; Burke et al., 2014, 2018b). In each wasp species, two of these loci contain multiple segments arranged in tandem, while the others contain only one or two segments. Within species comparisons indicate adjoining segments in a given locus are more similar to one another than to segments in other loci, which suggests the DNAs packaged into virions primarily evolve through tandem duplication events followed by sequence divergence (Desjardens et al., 2008; Bézier et al., 2013; Burke et al., 2014).

Early studies suggested that replication of proviral segments begins with excision from the wasp genome in calyx cells followed by circularization and amplification by a rolling-circle mechanism (Gruber et al., 1996; Savary et al., 1997; Pasquier-Barre et al., 2002). However, subsequent studies showed that unknown factors amplify proviral loci at discrete boundaries, which results in the formation of large, concatenated DNA intermediates that are then processed and circularized before packaging into nucleocapsids (Louis et al., 2013; Burke et al., 2015). Specific nucleotides in repeat domains identify the boundaries of amplification for proviral loci, while *int/vlf-1*-dependent processing occurs in direct repeats containing the tetramer AGCT, which flanks each segment. This results in formation of circularized DNAs that are packaged into virions (Annaheim and Lanzrein, 2007; Louis et al., 2013; Burke et al., 2013, 2015). Wasp integration motifs (WIMs) or direct repeat junctions (DRJs) are the sites of processing and circularization that flank each segment (Beck et al., 2011; Louis et al., 2013). Most circularized DNA segments in BV virions also
contain a single conserved inverted repeat domain, the host integration motif (HIM), where circularized segments integrate into the genome of infected host cells (Beck et al., 2011; Chevignon et al., 2018). This enables segments to persist and express associated genes in infected hosts for long periods in the absence of replication. The integration of BV DNAs into hosts has also potentially contributed to horizontal transfer of BV genes to Lepidoptera through presumptive infection of germ cells (Beck et al., 2011; Gasmi et al., 2015; Drezen et al., 2017).

Sequencing BAC genomic library clones from C. congregata showed that some nudivirus-like genes reside in an 18 kb domain, with one gene (odv-e66-like) located between proviral loci 1 and 2 (Bézier et al., 2009). However, most insights about the locations of BV proviral loci and nudivirus-like replication genes in the genomes of wasps derive from sequencing of M. demolitor (Burke et al., 2014, 2018b). The most current assembly (Mdem2) indicates the M. demolitor genome is 241 Mb and has an N50 scaffold size of 1.1 Mb (Burke et al., 2018b). The 8 Mb DV proviral loci reside on 5 scaffolds, while all of the nudivirus-like replication genes reside on 24 scaffolds (Fig. 8.3A). Proviral locus 2, located on a 2.4 Mb scaffold, is ~75 kb away from one nudivirus-like gene (p74) and 323 kb away from several others (35a-8 to 35a-14; odv-e66–9 to odv-e66–20; 35a-6 and 35a-7; and helicase) (Fig. 8.3A). Another nudivirus-like gene, HZNVorf93-like, is located close to DNA segment T (Fig. 8.3A). An additional 20 nudivirus-like replication genes reside in a 75 kb region of another scaffold with a subset of these genes present in the same order and orientation in the C. congregata genome (Fig. 8.3B). These data also indicate this nudivirus-like gene cluster has remained stable since divergence of M. demolitor and C. congregata approximately 53 million years ago, which suggests dispersal of the other nudivirus-like genes occurred relatively early in BV evolution (Burke et al., 2014). However, this cluster likely resides more than 1 Mb away from any proviral locus. All remaining nudivirus-like genes, which include the 4 essential RNA polymerase subunits and several virion components, are flanked by insect genes on different scaffolds and thus are also distantly located from: (1) all proviral loci, (2) the scaffold on which the cluster of 20 nudivirus-like genes are located, and (3) one another except for local gene family expansion events (Burke et al., 2018b) (Fig. 8.3A). No nudivirus-like genes are flanked by the conserved repeats that identify the amplification boundaries for proviral loci and processing boundaries for segment circularization. This architectural feature is why no DNA regions containing nudivirus-like genes are amplified and packaged into virions. Lastly, while all nudivirus and baculovirus genes are intronless, four nudivirus-like genes in M. demolitor (lef-8, lef-9, HZNVorf128-like, and K425_12) contain small introns, which, as previously noted, is also the case for some but not all virulence genes in proviral loci. Thus, acquisition of introns by some MdBV genes is likely a feature linked to transcription by wasp or host machinery. Overall, results from M. demolitor suggest proviral loci and nudivirus-like replication genes are likely dispersed in all microgastrine braconids. However, confirming this and assessing whether MdBV proviral loci and replication genes are fixed in the same or different locations across microgastrine genomes will require data from other species.

**BV**s are endogenous virus elements (**EVEs**) that retain many ancestral functions

Endogenous virus elements (**EVEs**) are DNA sequences from viruses that have integrated and become fixed in the germline of a host resulting in strict vertical inheritance (Katzourkakis and Gifford, 2010; see also Chapter 2). Most known EVEs derive from retroviruses (eukaryotes) and phages (prokaryotes), which is consistent with the former obligately integrating into hosts as part of replication and the latter either obligately or facultatively integrating into a broad range of bacteria and archaea (Calendar, 2005; Greenwood et al., 2018). Other types of viruses, including several that are not known to integrate into the genomes of hosts, have also endogenized in different organisms through unknown mechanisms (Katzourkakis and Gifford, 2010; Holmes, 2011b; Aswad and Katzourkakis, 2012). Most EVEs are ancient and persist as non-functional fragments that have decayed through accumulation of neutral mutations. However, a few individual genes or regulatory elements of viral origin have been coopted by hosts for novel functions that provide fitness benefits in development and control of particular genes (Feschotte and Gilbert, 2012; Stoye, 2012).
Current results strongly support that BVs evolved from a betanudivirus that infected the common ancestor of microgastroid braconids approximately 100 million years ago; IVs also likely evolved from a virus ancestor (Volkoff et al., 2010; Bézier et al., 2015). Fixation of all DNA components required to produce BVs in the genomes of wasps indicate these entities are now EVEs that have been coopted by wasps for beneficial functions in parasitism (Strand and Burke, 2013; Herniou et al., 2013; Burke et al., 2015).
In turn, BVs and IVs can no longer be viewed as beneficial symbionts of wasps because no BV traits exist outside of the wasp genome.

Yet, BVs and IVs also differ from most known EVEs because they retain a remarkably large number of ancestral features. This begins with overall genome composition where nudiviruses encode a small proportion of conserved core genes and a large proportion of non-core genes from diverse sources with functions in infecting hosts. Most of the same core genes in nudiviruses and baculoviruses remain in BV-producing wasps while experimental studies indicate that BV core genes also have similar functions as determined primarily in experimental studies of baculoviruses. BV proviral loci contain virulence genes, which have functions in parasitism of hosts by wasps, that also potentially derive from the nudiivirus ancestor. A crucial difference between nudiviruses and BVs is that all components of nudiivirus genomes reside on a single, large dsDNA that can be amplified, packaged into virions, and horizontally transferred to another host, while BV components are dispersed in the genomes of wasps. Given results showing that the titres of MdBV virions produced in M. demolitor calyx cells greatly exceed baculovirus titres produced in permissive cell lines, dispersal of the ancestral genome is not a barrier for function (Burke et al., 2012a, 2014). This is because the MdBV RNA polymerase subunits, once expressed and assembled into a holoenzyme, specifically transcribe other core genes required for virion formation through promoter recognition regardless of their location in the wasp genome (Burke et al., 2013, 2014; Strand and Burke, 2015). In turn, incompletely characterized DNA replication machinery plus two nudiavirus-like integrases (int-1, vlf-1) use the repeat boundaries that flank proviral segments for proper amplification, processing and packaging into virions. That motifs like WIMs are required for proper processing further suggests these flanking motifs derive from the nudiivirus ancestor. In contrast, genome dispersal is a key factor in explaining how BVs amplify and package a specific population of DNAs containing virulence genes into infectious virions, but package no DNAs containing core genes, which results in BVs being replication defective (Strand and Burke, 2015).

A challenge then is what to call and how to classify PDVs. Some have suggested PDVs are a type of organelle (Federici and Bigot, 2003; Whitfield and Asgari, 2003) or could be named using a novel term like ‘suppresson’, which refers to their generally inhibitory effects on parasitized hosts (Kroemer and Webb, 2004). Evolution of BVs clearly involved gene transfer from an ancestral nudiivirus to the germ line of a wasp, as occurred in the symbiogenesis of mitochondria from bacteria (Margulis, 1981). However, the word organelle does not accurately reflect what PDVs are because mitochondria still retain small autonomously replicating genomes that provide the potential for evolving separately from the rest of the host genome. Kroemer and Webb (2004) also noted that ‘a derivative terminology’, like suppresson, ‘whether based on viruses or organelles, is likely to prove insufficient to describe PDVs, which clearly have a unique set of genetic, evolutionary, and physiological attributes.’

The closest parallel to PDVs outside of findings discussed below are gene transfer agents (GTAs), which are phage-like entities in certain proteobacteria, spirochaetes and archaea (Lang et al., 2012). Like PDVs, GTAs produce particles from ancestral phage genes that are fixed in the bacterial genome and package DNA into particles that the GTA can transfer to another cell. However, GTAs differ from PDVs in that the DNAs they package are random, small (≈5 kb or less) pieces of the host genome that sometimes contain GTA genes (Lang et al., 2012). Whether GTAs have been coopted by their hosts for beneficial functions is also unclear.

Taken together, the PDV community continues to use the terms polydnavirus, bracovirus and ichnovirus but has largely coalesced in recognizing that BVs and IVs are complex EVEs that retain many functions of their virus ancestors (Strand and Burke, 2013, 2014; Bézier et al., 2013; Drezen et al., 2017). This is perhaps most analogous to endogenous retroviruses (ERVs), which share many traits with exogenous retroviruses (XRVs) (Greenwood et al., 2018). ERVs and XRVs are both used in phylogenies for the Retroviridae (Greenwood et al., 2018). Recent phylogenies for the Nudiviridae do so as well (Bézier et al., 2015; Pichon et al., 2015; Burke et al., 2018a) (see Fig. 8.2 and below). Thus, endogenous nudiivirus (ENV) is likely the most parsimonious name for BVs. If adopted, M. demolitor bracovirus (MdBV) would become MdENV or M. demolitor endogenous betanudivirus (MdEBENV) while other endogenized nudiviruses would be named by
the host species and potential genus in which they reside. This approach is not possible for IVs because their exogenous virus ancestor remains unknown.

**Other nudivirus-derived EVEs exist**

Nudiviruses and baculoviruses infect a variety of insects in several orders but, similar to parasitoids, most isolates have narrow ranges (Rohrmann, 2013). Most isolates also cause systemic fatal infections. Baculovirus genomes remain episomal in infected host cells and productively replicate through the coordinated expression of early genes like dnapol that amplify the genome and RNA polymerase subunits that transcribe the structural genes needed to produce infectious virions. Nudiviruses replicate similarly, but two very similar betanudiviruses, HzNV-1 and HzNV-2, preferentially infect the reproductive system of hosts rather than establish systemic infections (Lu and Burand, 2001; Burand et al., 2012). These isolates can also establish persistent infections in vitro by integrating into the genome of infected cells (Wu et al., 2010, 2011). This suggests the betanudivirus ancestor of BVs initially established a persistent infection in the progenitor of the microgastrid complex by infecting its reproductive system and integrating into the germ line (Herniou et al., 2013; Strand and Burke, 2013). Initial integration of multiple copies of the ancestral genome or duplication of the ancestral genome was then followed by rearrangements and mutation that resulted in fixation and dispersal in the genomes of wasps. Other alterations presumably restricted nudivirus-like replication gene expression to calyx cells in pupal and adult stage female parasitic wasps, which enabled the formation of replication defective virions that could be stored in the lumen of the reproductive tract and deliver virulence genes to the wasps’ hosts.

It is impossible to know the precise series of events that led to the evolution of BVs given the ancestral infection event occurred approximately 100 million years ago (see above). Because all microgastrine braconids studied to date exhibit these traits, it is likely that genome dispersal, the architecture of proviral loci, and the restriction of replication to calyx cells occurred relatively early in the association. If nudiviruses more generally exhibit tropisms for infecting reproductive organs and/or the capacity to establish persistent infections, it is also possible other endogenization events have occurred that could provide insights into whether nudiviruses exhibit similar or different fates as BVs.

Notably, recent studies identify three examples of endogenized nudiviruses evolving independently of BVs. The first was discovered by sequencing the brown planthopper Nilaparvata lugens (Hemiptera), which identified 32 of the 33 known nudivirus core genes in the genome (Cheng et al., 2014). Analysis of concatenated amino acid alignments for these genes indicate the ancestor of N. lugens endogenous nudivirus (NlENV) was an alphanudivirus (Bézier et al., 2015) (Fig. 8.2). qPCR data show that some of the nudivirus core genes are transcribed in the gut and fat body of N. lugens, but EM studies detected no virions in any tissue including the reproductive tract (Cheng et al., 2014).

The second example was identified in the parasitoid wasp Venturia canescens (Pichon et al., 2015). This example is notable because this species is a campoplegine ichneumonid, which were all thought to carry IVs. V. canescens was one of the wasp species where virus-like particles were first observed (Rotheram, 1967). Like other campoplegine ichneumonids, V. canescens particles assemble in calyx cell nuclei and accumulate in the lumen of the reproductive tract. Particles also bind to the surface of wasp eggs and impede encapsulation by host immune cells after parasitism by wasps (Rotheram, 1967; Salt, 1973). The particles contain no nucleic acid and differ from IVs morphologically by having an amorphous electron dense centre surrounded by a single envelope (Rotheram, 1967; Bedwin, 1979). Subsequent studies identified three proteins (VeVLP1–3) in particles encoded by putative intron-containing wasp genes, suggesting they may not derive from any virus (Hellers et al., 1996; Reineke et al., 2002; Asgari et al., 2003). Recent sequencing of the V. canescens genome, in contrast, identified 16 nudivirus-like genes in six dispersed clusters most closely related to alphanudiviruses (Figs. 8.2 and 8.4). Among these are nudivirus core genes including the four RNA polymerase subunit genes and structural genes, *per os* infectivity factors (*pifs*), that are envelope components in baculoviruses (Boogaard et al., 2018) (Fig. 8.4). Immunolabeling studies confirmed the presence
of some *pif* gene encoded proteins in the envelope and VLP2 in the electron dense centre of *V. canescens* particles (Pichon et al., 2015). In contrast, the *V. canescens* genome did not yield nudivirus-like genes with predicted functions in capsid formation (Pichon et al., 2015). Sequence data also indicated that the IVSPER genes present in other campoplegine ichneumonids have been either pseudogenized or lost in *V. canescens* (Pichon et al., 2015). These results collectively suggest alpha-nudivirus endogenization in *V. canescens* occurred subsequently to endogenization of the virus that gave rise to IVs (Pichon et al., 2015). Whether VcENV is directly or indirectly responsible for the loss of IVSPERs is unclear.

The third example also derives from a parasitoid wasp, *Fopius arisanus*, which resides in the Braconidae in a subfamily (Opiniinae) outside the microgastroid complex (Burke et al., 2018a) (Fig. 8.1). Sequencing of *F. arisanus* identified 55 nudivirus-like genes in 9 clusters and several other dispersed locations also derived from an alphanudivirus ancestor (Geib et al., 2017; Burke et al., 2018b) (Figs. 8.2 and 8.5). EM studies showed that particles assemble specifically in calyx cell nuclei of female wasps, which accumulate in the oviducts of ovaries (Burke et al., 2018a). These particles consisting of long open-ended capsids that are individually surrounded by an envelope are strikingly different from BVs and the particles in *V. canescens*. The nudivirus genes present in *F. arisanus* include all of the nudivirus core genes except *dnapol, pif-5, p6.9*, and *vlf-1* (Fig. 8.5). Proteomic analysis confirmed that nudivirus envelope and capsid gene products are present in particles, but they do not contain nucleic acid (Burke et al., 2018b).

Since *lef-8* is the most conserved gene among sequenced alphanudiviruses, Burke et al. (2018b) screened five other species of *Fopius* by PCR to determine if potentially other species in this genus contain an endogenized nudiviruses. Results identified *lef-8* orthologues in three species whose identity were confirmed by sequencing while no *lef-8* orthologues were identified in two others (Burke et al., 2018a). Thus, detection of *lef-8* in only some *Fopius* species suggests endogenization likely occurred relatively recently in this genus of parasitoids (Burke et al., 2018b). Whether the particles produced by *F. arisanus* have functions in parasitism of hosts in contrast is unknown.

These findings together indicate that alphanudiviruses independently endogenized in different insects. Endogenization can also result in loss of virion formation and infectivity (NIENV), formation of virions containing core gene products but no nucleic acid (VcENV and FaENV), or formation of infectious, replication-defective virions that package large DNAs containing virulence genes.
Polydnaviruses (BVs). A wide range in the age of these associations is suggested, with VcENV and FaENV showing features suggesting more recent endogenization than in microgastrine braconids. Despite varying fates and age, these endogenized nudiviruses share three similar architectural features: genome fragmentation, retention of multiple core genes, and loss of dnapol. NlENV, VcENV, and FaENV have only a single copy of each retained core gene, whereas some core genes in the more ancient association between BVs and microgastrine braconids form multimember families. As evidenced by insect genes flanking clusters of nudivirus genes, the most parsimonious explanation for the patterns observed.

Figure 8.5 Dispersal of FaENV in the genome of Fopius arisanus. (A) The 11 scaffolds in the F. arisanus genome that contain FaENV genes. Wasp genes are indicated in light (forward strand) or dark (reverse strand) blue, while FaENV genes are indicated in pink. Numbers identify domains were multiple FaENV genes reside in close proximity to one another and form a cluster. (B) Higher resolution schematic of the nine FaENV gene clusters shown in (A). The location of the RNA polymerase subunit and pif genes on the forward (upper) or reverse (lower) strand of a given cluster are identified.
in NIENV, VcENV and FaENV is that a single nudi-virus genome integrated into the germline of each host species, followed by rearrangement. Dispersal of FaENV or VcENV is also less pronounced than for BVs, potentially reflecting an older association between BVs and microgastroids. The fact that all ENVs identified to date are dispersed suggests that genome fragmentation occurs relatively soon after endogenization.

Burke (2018) compared core gene inventories between parasitoid-associated ENVs and noted that the four subunits of the viral RNA polymerase (lef-4, lef-8, lef-9, p47) and select pifs are the only genes shared by all (see Figs. 8.3–8.5). However, no syntenic gene location exists between these ENVs (Figs. 8.3–8.5). RNA polymerase is essential for transcription of structural core genes in both baculoviruses and BVs (Rohrmann, 2013; Burke et al., 2013). Selection also likely maintains functional RNA polymerase subunit genes in FaENV and VcENV because they are required for holoenzyme formation and expression of capsid and/or envelope genes. In baculoviruses, seven known pif genes produce envelope proteins that function together in a complex (pif0–4, pif6, pif8) while one (pif-5) encodes an envelope protein that does not interact with the others (Boogaard et al., 2018). Retention of a complete set of PIFs in MdBV, VcENV and FaENV is consistent with each producing enveloped particles. In contrast, functional parallels to baculoviruses are less clear because BV-carrying wasps, V. canescens, and F. arisanus inject particles plus eggs into the hemocoel of hosts, which precludes interaction with midgut cells as per os infectivity factors. Moreover, RNAi knockdown of pif-1 from MdBV has no effect on the number of DNA segments detected in host cells like haemocytes but does affect transcript abundance of virulence genes expressed in haemocytes (Burke et al., 2013). This suggests pif-1 may not be essential for entry of MdBV into haemocytes but instead may affect capsid migration and/or delivery of encapsidated DNAs to the nucleus where associated virulence genes are transcribed. The universal loss of dnapol among currently known ENVs may reflect a selective advantage to insects by preventing viral genomes from re-mobilizing, as previously hypothesized for BVs (Strand and Burke, 2013, 2014).

**Future directions**

When we wrote a chapter on PDVs for the first edition of *Insect Virology*, the first evidence that BVs evolved from a nudi-virus had entered the literature (Bézier et al., 2009). In contrast, the architecture of BV and IV genomes in wasps were largely unknown, no evidence existed that the DNAs in BV virions integrate into the genomes of hosts that wasps parasitize, and no other ENVs had been identified. Results now strongly support that all BVs derive from a betanudivirus and that IVs likely evolved from a virus whose identity remains unknown. Several additional ENVs have also been identified. Future revision of the Polydnaviridae is obviously needed but has been delayed because of issues related to revision of the Baculoviridae and Nudiviridae. Whether to include endogenous nudi-viruses like BVs in classification of the Nudiviridae, as the retrovirus community does, must be decided as well as how to classify IVs in the absence of a known virus ancestor.

Mechanistically, many unanswered questions exist regarding how nudi-viruses integrate to form persistent infections and the factors that regulate preferential infection of reproductive tissues by isolates like HzNV1 and 2. It is unclear whether the mechanism of integration by nudi-viruses is in any way related to the features in BVs that regulate amplification and processing in wasps, or integration of circularized DNAs into the genomes of hosts that wasps parasitize. While both an early gene and a microRNA are implicated in maintenance and reactivation of HzNV1 persistent infections (Wu et al., 2010, 2011), no orthologues of these products have been identified in BVs. The wasp genes required for amplification of BV proviral loci also remain unknown but are essential to understanding how these domains are mobilized in concert with particle formation in calyx cells. A final point of interest is whether the current pattern of most ENVs being associated with parasitic Hymenoptera reflects the biased efforts of individuals with interests in PDVs to identify ENVs, a legitimate bias in nudi-virus infection of Hymenoptera, or a potential bias in parasitoids providing a selective environment that favours cooption of ENVs for roles in parasitism. The rapid increase in available genomes across different orders of insects should be able to address this question in the fairly near future.
Acknowledgements

Work discussed in this chapter includes research supported by grants from the National Institutes of Health (F32AI096552) (GRB), the US National Science Foundation (IOS-12611328) (MRS), (IOS-1748862) (GRB) and (DEB-1622986) (GRB), and the US Department of Agriculture (2009-35302-05250) (MRS).

References


bracoviruses. PLOS Genet. 11, e1005470. https://doi.org/10.1371/journal.pgen.1005470


