Recurrent DNA virus domestication leading to different parasite virulence strategies

Apolline Pichon, Annie Bézier, Serge Urbach, Jean-Marc Aury, Véronique Jouan, Marc Ravallec, Julie Guy, François Cousserans, Julien Thézé, Jérémy Gauthier, et al.

LIFE SCIENCES

10.1126/sciadv.1501150

Recurrent DNA virus domestication leading to different parasite virulence strategies

Apolline Pichon,[1] Annie Bézier,[2] Serge Urbach,[3] Jean-Marc Aury,[4] Véronique Jouan,[1] Marc Ravallec,[1] Julie Guy,[4] François Cousserans,[1] Julien Thézé,[2] Jérémy Gauthier,[2] Edith Demettre,[3] Sandra Schmieder,[5] François Wurmser,[6] Vonick Sibut,[2] Marylène Poirié,[5] Dominique Colinet,[5] Corinne da Silva,[4] Arnaud Couloux,[4] Valérie Barbe,[4] Jean-Michel Drezen,[2*†] Anne-Nathalie Volkoff[1*†]

Relics of ancient infections are abundant in eukaryote genomes, but little is known about how they evolve when they confer a functional benefit on their host. We show here, for the first time, that the virus-like particles shown to protect *Venturia canescens* eggs against host immunity are derived from a nudivirus genome incorporated by the parasitic wasp into its own genetic material. Nudivirus hijacking was also at the origin of protective particles from braconid wasps. However, we show here that the viral genes produce "liposomes" that wrap and deliver *V. canescens* virulence proteins, whereas the particles are used as gene transfer agents in braconid wasps. Our findings indicate that virus domestication has occurred repeatedly during parasitic wasp evolution but with different evolutionary trajectories after endogenization, resulting in different virulence molecule delivery strategies.

INTRODUCTION

Virus endogenization appears to be a frequent, ongoing process in the evolution of eukaryotic genomes (*1*, *2*). Polydnaviruses constitute a unique example in which eukaryotes maintain a complex viral machinery in their genomes, providing a function essential for their life cycle. Polydnaviruses are carried by thousands of species of parasitic wasps that develop within living insect larvae (*3*). The wasp makes use of these viruses to establish successful parasitism in hosts with competent immune defenses. Polydnaviruses include the bracoviruses and ichnoviruses associated with braconid and ichneumonid wasps, respectively. Both groups of viruses mediate the transfer of virulence genes (*4*) and have evolved by convergence, through two genome integration events involving the insertion of DNA from unrelated viruses into wasp chromosomes (*5*, *6*). However, an enigma remains regarding the nature of the first particles from parasitic wasps to be described, produced in a species from the group of ichnovirus-carrying wasps, *Venturia canescens*. This wasp uses "virus-like particles" (VLPs) to protect its eggs against the host immune response (*7*–*10*). VLPs are produced in the ovaries and released into the ovary lumen, where they become associated with the eggs. In the parasitized host, they protect the eggs against encapsulation, an insect immune response involving the engulfment of foreign bodies in a sheath of immune cells. Unlike polydnaviruses, which convey virulence genes, VLPs are devoid of DNA (*10*, *11*) and enclose virulence proteins (*12*). VLPs do not resemble the particles of known viruses, and a debate about whether these particles are truly of viral origin has raged since the early 1970s (*12*). We show here that VLPs are clearly of viral origin and consist of virulence proteins wrapped in viral envelopes. They are produced by a machinery derived from the domestication of a nudivirus [a sister group of baculoviruses (*13*)]. This domestication occurred through an event independent of the capture of a nudivirus at the origin of bracoviruses (*14*). Thus, similar viruses have been acquired in different wasp lineages but have followed different evolutionary trajectories after endogenization.

RESULTS

A nudiviral machinery produces *V. canescens* VLPs

We first investigated the nature of the VLPs using a comprehensive deep sequencing analysis of the mRNAs synthesized in the tissues producing them. An analysis of the calyx transcriptome led to the identification of 51 abundant transcripts displaying significant similarities to nudivirus genes (Table 1 and table S1). Polymerase chain reaction (PCR) amplification with DNA from various sources showed that all of the wasp specimens analyzed contained nudiviral genes (fig. S1A), which were therefore probably fixed in *V. canescens* and did not correspond to a pathogenic virus infecting the laboratory strain.

The hypothesis on a role for these nudiviral genes in VLP production was further supported by the nature of the transcripts. Indeed, nudiviruses have 21 conserved genes involved in essential viral functions in common with baculoviruses (*15*), and transcripts corresponding to 16 of these genes were identified in *V. canescens* ovaries. Their function in the baculovirus life cycle has been characterized [reviewed by Rohrmann (*16*)], and we can infer that nudiviral homologs may have a similar role. The VLP-producing tissue expressed homologs of the seven genes encoding viral envelope proteins [named *per os* infectivity factors (PIFs)] mediating the entry of particles into target cells. We also detected the expression of four genes (*lef-4*, *lef-8*, *lef-9*, and *p47*) encoding the subunits of the viral RNA polymerase controlling the transcription of a series of viral genes. Conversely, no viral DNA polymerase or capsid protein (for example, VP39, 38K, and VLF-1) transcripts were detected, consistent with the absence of nucleocapsids

[1]Microorganism and Insect Diversity, Genomes, and Interactions (DGIMI) Laboratory, UMR 1333 INRA, Université de Montpellier, Place Eugène Bataillon, CC101, Montpellier Cedex 34095, France. [2]Institut de Recherche sur la Biologie de l'Insecte (IRBI), UMR 7261, CNRS-Université François Rabelais de Tours, Parc de Grandmont, Tours 37200, France. [3]Functional Proteomics Platform, BioCampus Montpellier, UMS CNRS 3426, INSERM US009, Institut de Génomique Fonctionnelle, UMR CNRS 5203, INSERM U661, Université de Montpellier, Montpellier 34094, France. [4]Commissariat à l'Energie Atomique (CEA), Institut de Génomique (IG), Genoscope, 2 rue Gaston Crémieux, BP5706, Evry 91057, France. [5]Institut Sophia Agrobiotech (ISA), UMR INRA 1355, CNRS 7254, Université Nice Sophia Antipolis, 400 route des Chappes, Sophia Antipolis 06903, France. [6]PPF Analyse des systèmes biologiques, Université François Rabelais de Tours, 3 Boulevard Tonnellé, Tours 37000, France.
*Corresponding author. E-mail: volkoff@supagro.inra.fr (A.-N.V.); drezen@univ-tours.fr (J.-M.D.)
†These authors contributed equally to this work.

(DNA packaged in structural proteins) in the VLPs. A role for nudivirus genes in VLP production was also supported by their temporal and tissue-specific patterns of expression, as assessed by quantitative reverse transcription PCR (RT-PCR), which were strongly correlated with VLP production (fig. S1). Finally, transmission electron microscopy (TEM) studies of VLP production in wasp ovary cells showed that the morphogenesis of VLPs resembled that of nudiviruses (*17*) and bracoviruses (*18*). Indeed, VLPs form in the nucleus (Fig. 1A) from a "virogenic stroma" (the "viral factory" in which virus particles are generally assembled) and a network of viral envelopes (Fig. 1, B and C) typical of these viruses. This results in the production of enveloped particles, which leave the nucleus through the nuclear envelope and are released from the cell by budding from the plasma membrane (Fig. 1D).

We checked whether the nudiviral genes produced VLPs by analyzing proteins from purified particles using liquid chromatography tandem mass spectrometry (LC-MS/MS). We identified 26 nudiviral gene products, including the whole set of PIF envelope proteins, as components of the

Table 1. Nudiviral and virulence genes expressed in *V. canescens* ovaries and the presence of their products in VLPs. Nudiviral genes of unknown function are detailed in table S1. RPKM, number of Illumina reads per kilobase per million mapped reads. PSM, number of peptide spectrum matches.

Protein putative function	Gene name	Transcript abundance (RPKM)	Peptide abundance (PSM)	OrNV homolog	Accession number	BLASTp *e* value
Nudiviral genes encoding VLP components						
Envelope component	*pif-0* (*p74*)	12,900	54	P74 protein (OrNV_gp126)	YP_002321437	1.0×10^{-85}
	pif-1	5,196	28	PIF-1 (OrNV_gp060)	YP_002321371	2.0×10^{-94}
	pif-2	1,980	34	PIF-2 (OrNV_gp017)	YP_002321328	9.8×10^{-112}
	pif-3	2,655	12	PIF-3 (OrNV_gp107)	YP_002321418	2.0×10^{-59}
	pif-4	7,731	5	19-kDa protein (OrNV_gp033)	YP_002321344	5.0×10^{-44}
	pif-5-1	527	5	ODV-E56 (OrNV_gp115)	YP_002321426	1.0×10^{-34}
	pif-5-2	863	21	ODV-E56 (OrNV_gp115)	YP_002321426	4.8×10^{-38}
	pif-5-3	1,048	9	ODV-E56 (OrNV_gp115)	YP_002321426	1.6×10^{-29}
	pif-6	3,011	2	Ac68-like (OrNV_gp072)	YP_002321383	4.0×10^{-40}
	vp91	2,966	23	VP91 (OrNV_gp106)	YP_002321417	1.4×10^{-84}
	p33	5,553	5	Ac92-like (OrNV_gp113)	YP_002321424	8.6×10^{-48}
Unknown	15 other genes related to OrNV					
***V. canescens* genes encoding VLP components**						
Virulence	*PHGPx VLP1*	1,115	134			
	RhoGAP VLP2	23,134	400			
	Neprylisin VLP3	3,540	138			
Nudiviral genes transcribed in VLP-producing tissue not encoding VLP components						
Replication	*Helicase*	27	—	DNAHEL (OrNV_gp034)	YP_002321345	0.0
Transcription	*RNA polymerase lef-4*	488	—	LEF-4 (OrNV_gp042)	YP_002321353	2.8×10^{-92}
	RNA polymerase lef-8	125	—	LEF-8 (OrNV_gp064)	YP_002321375	0.0
	RNA polymerase lef-9	724	—	LEF-9 (OrNV_gp096)	YP_002321407	0.0
	RNA polymerase p47	916	—	P47 (OrNV_gp020)	YP_002321331	1.6×10^{-97}
	Initiation factor lef-5	245	—	LEF-5 (OrNV_gp052)	YP_002321363	5.5×10^{-18}
Envelope component	*pif-5-4*	2,286	—	ODV-E56 (OrNV_gp115)	YP_002321426	9.1×10^{-22}
	Ac81-1	272	—	Ac81-like protein (OrNV_gp004)	YP_002321315	2.0×10^{-34}
	Ac81-2	1,777	—	Ac81-like protein (OrNV_gp004)	YP_002321315	1.0×10^{-39}
	Ac81-3	1,080	—	Ac81-like protein (OrNV_gp004)	YP_002321315	9.3×10^{-34}
Unknown	18 other genes related to OrNV					

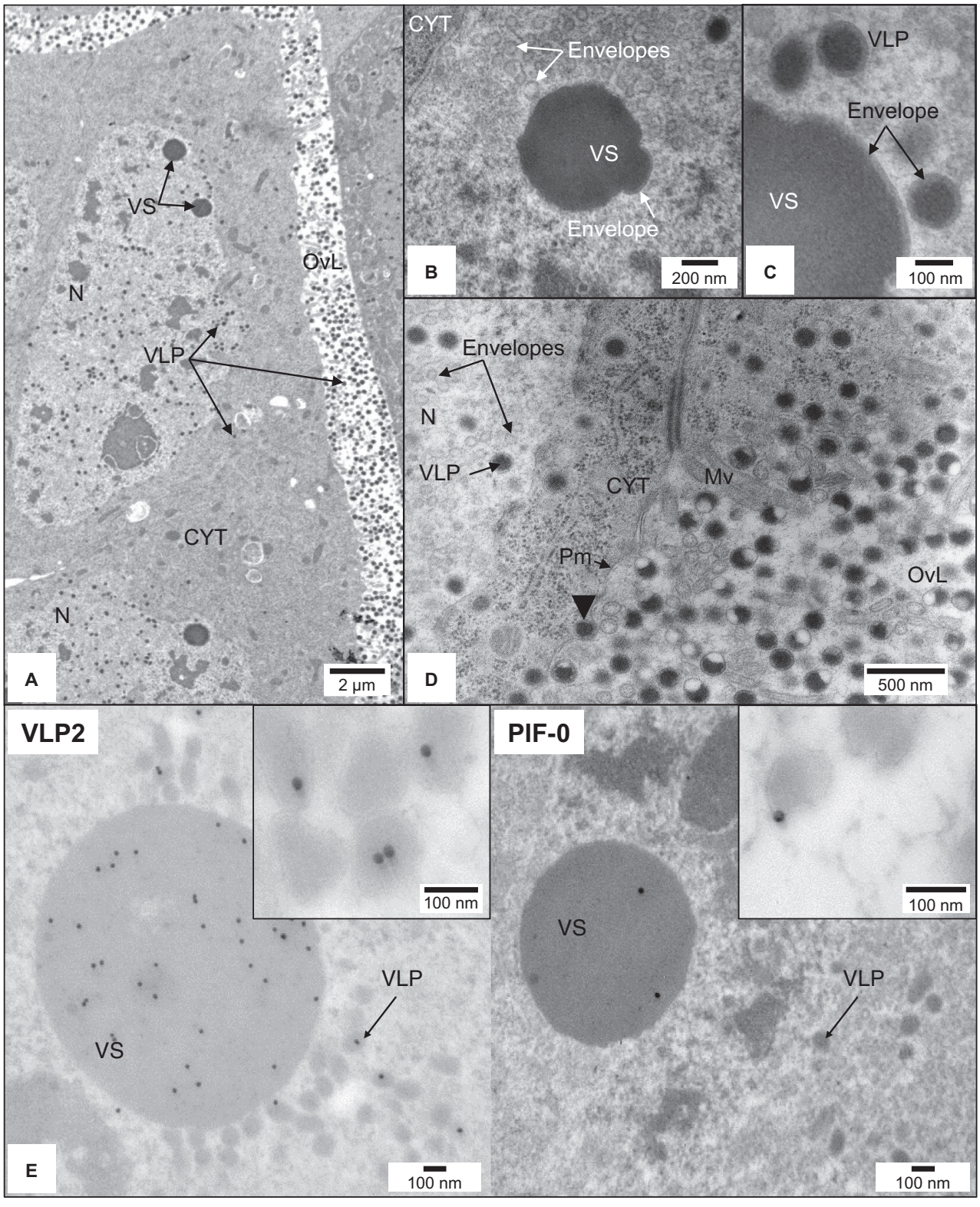

Fig. 1. VLP morphogenesis in *V. canescens* calyx cells. (**A** to **E**) Transmission electron micrographs of (A) calyx cells displaying hypertrophic nuclei (N) containing virogenic stroma (VS) and VLPs; (B) calyx cell nuclei with empty envelopes close to the VS; (C) envelopes filled with VS material giving rise to VLPs; (D) VLPs budding from the calyx cell membrane; (E) immunogold labeling of VS and VLPs in the nucleus with specific antibodies against VLP2 (**left**) and PIF-0 (**right**). (**Insets**) Larger magnifications of VLPs. CYT, cytoplasm; Mv, microvilli of the cytoplasmic membrane; OvL, oviduct lumen; Pm, plasma membrane.

VLPs (Table 1). We also identified the virulence proteins of wasp origin thought to ensure immune protection: a peroxidase (VLP1), a Rho GTPase (guanosine triphosphatase)–activating protein (VLP2), and a metalloprotease (VLP3) (*19*–*21*) (Table 1)—the most abundant components of VLPs (Table 1). The presence of VLP2 and a nudiviral PIF protein in the VLPs was also confirmed by their direct visualization in immunogold labeling experiments (Fig. 1E). We conclude that VLPs consist of wasp virulence proteins wrapped in nudiviral envelopes and that they are produced by a nudiviral machinery that has lost the ability to package DNA. VLPs may thus be considered as naturally produced viral liposomes.

Recurrent domestications of nudiviruses in parasitic wasps evolve differently

We investigated whether the nudiviral genes were present in the wasp DNA by sequencing the whole genome of *V. canescens*. We found that all of the nudiviral genes were grouped into six clusters, flanked by regions containing wasp genes (Fig. 2A and table S1), confirming that they formed an integral part of the genetic material of the wasp. Phylogenetic analyses indicated that nudiviral genes from the different clusters all belonged to the clade of *Oryctes rhinoceros* nudivirus (OrNV) (Fig. 2B). This suggests that VLPs are derived from the chromosomal integration of a nudivirus related to OrNV, the genome of which has undergone rearrangements leading to fragmentation and gene loss with respect to pathogenic nudiviruses (*15*).

The polydnaviruses associated with braconid wasps (bracoviruses) are also derived from the acquisition of a nudivirus genome (*5*). However, phylogenetic analyses indicated that the nudiviral genes giving rise to VLPs and bracoviruses came from different viral clades, the *Alphanudiviruses* and the *Betanudiviruses*, respectively (Fig. 2B). Moreover, *V. canescens* nudiviral genes are more closely related to those of pathogenic nudiviruses and are not widely distributed throughout the wasp genome, unlike most of the genes involved in bracovirus particle production (*5*, *22*). This is consistent with a more recent acquisition in

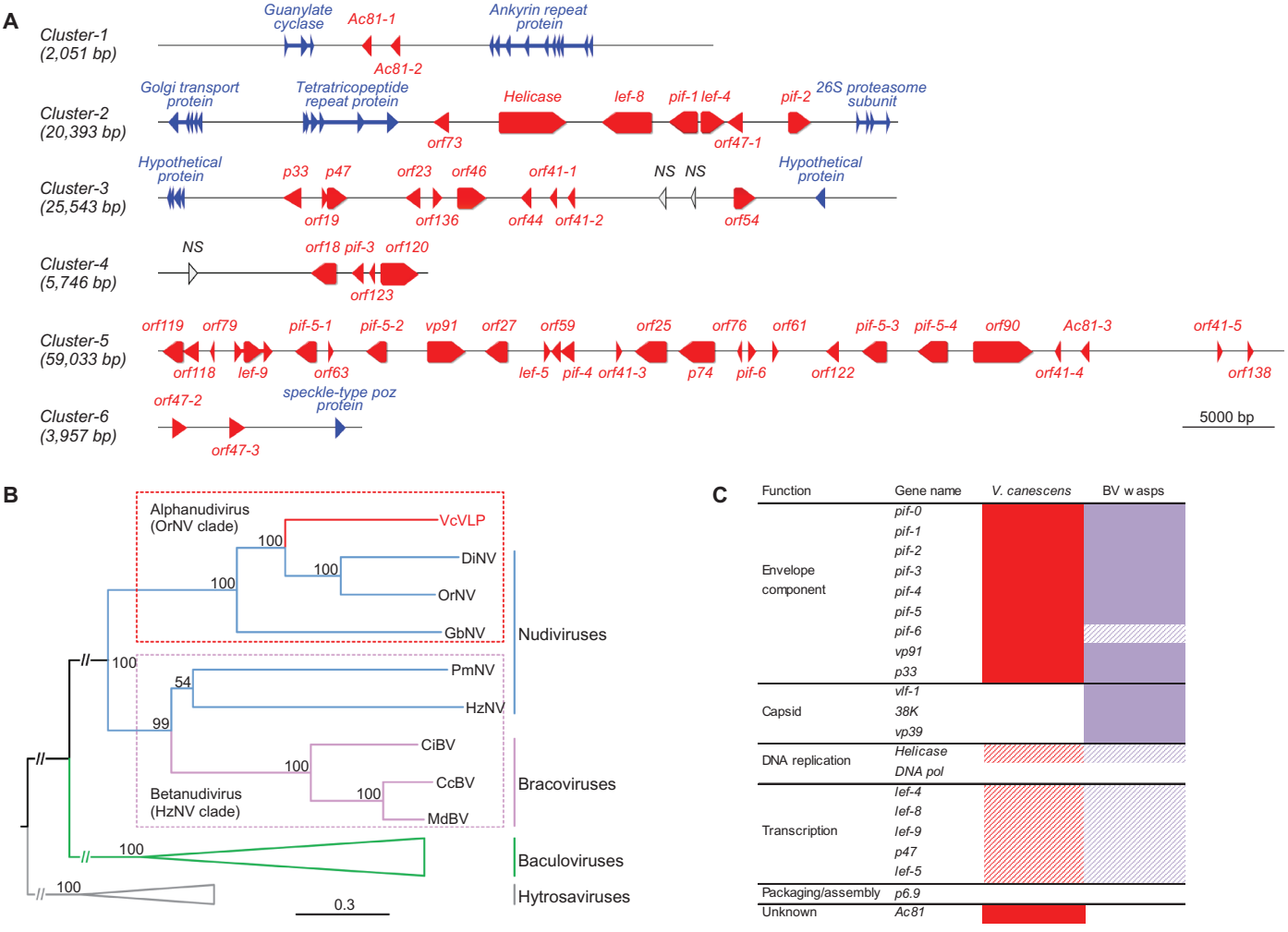

Fig. 2. Organization, origin, and putative function of the nudiviral genes of the *V. canescens* genome. (**A**) Nudiviral genes (in red; named after their OrNV homologs) are grouped into six clusters flanked by regions containing conserved wasp genes (in blue). (**B**) Phylogenetic tree of insect large double-stranded DNA viruses obtained by maximum likelihood analysis with a concatenated multiple alignment for 37 genes showing that the nudiviruses captured by the *V. canescens* and braconid wasp lineages belonged to different clades. VcVLP, *V. canescens* VLP. (**C**) Genes common to nudiviruses and baculoviruses conserved in *V. canescens* (red) or wasps carrying bracoviruses (BV) (purple). The products of the *helicase* gene and of the genes mediating viral gene transcription are not incorporated into the particles (hatched motif). Note that *V. canescens* lacks capsid genes, consistent with the absence of DNA in the particles.

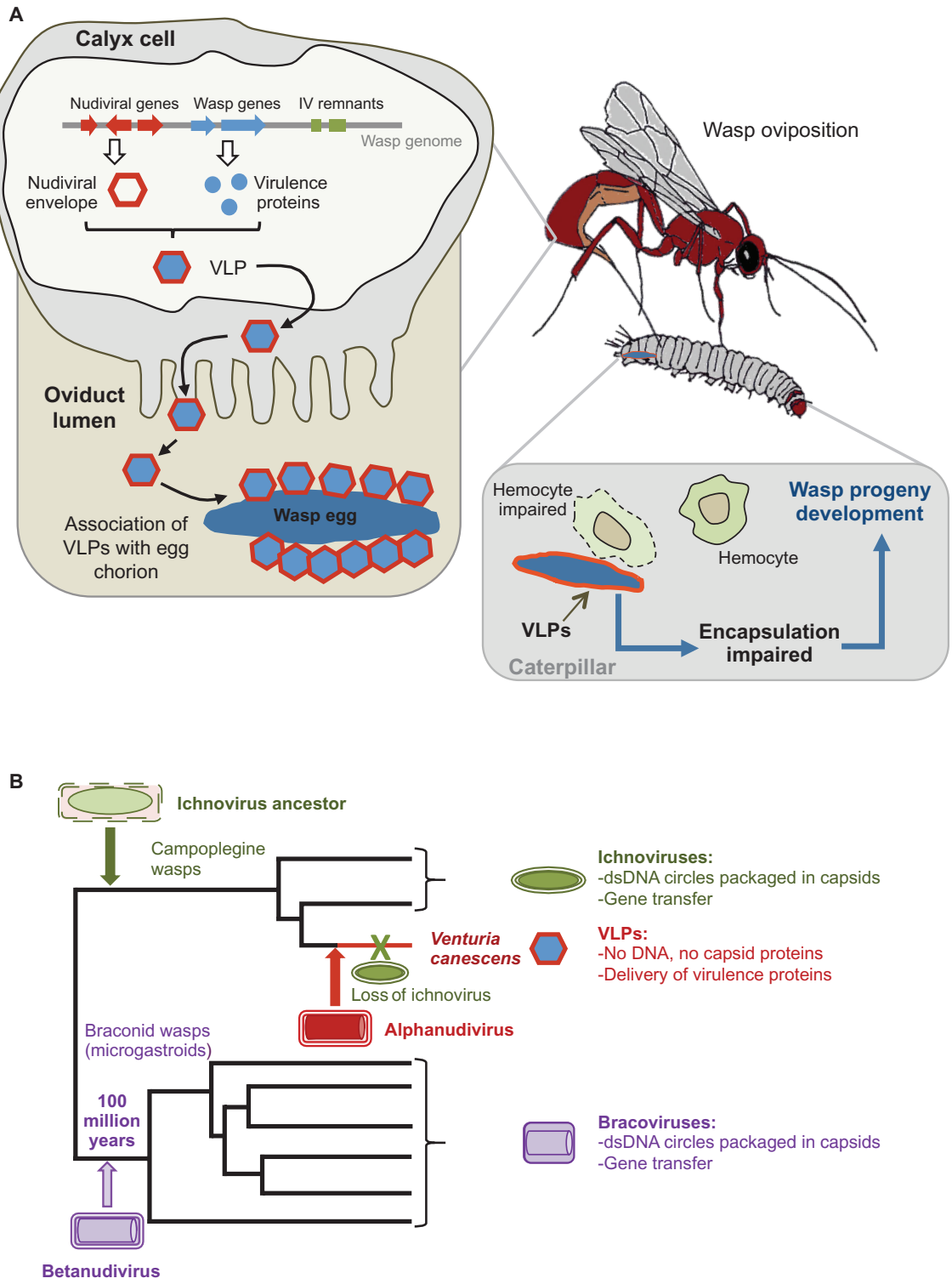

Fig. 3. Overview of the production, function, and origin of *V. canescens* VLPs. (**A**) VLPs consist of virulence proteins of wasp origin wrapped in nudiviral envelopes. VLPs are released into the oviduct lumen, where they become attached to the wasp egg. Once in the parasitized caterpillar host, VLPs confer immune protection to the egg by impairing capsule formation by the host immune cells. (**B**) Recurrent virus domestication during parasitic wasp evolution. Two virus genome integration events have occurred during the evolution of campoplegine wasps: the first involved a genome, making it possible to produce ichnoviruses (mediating gene transfer), and the second involved an alphanudivirus genome used for VLP production (mediating the delivery of virulence proteins), which has replaced the ichnovirus in *V. canescens*. A betanudivirus was independently acquired by a braconid wasp ancestor and gave rise to bracoviruses (DNA delivery).

V. canescens than in braconid wasps. Indeed, VLPs have been described only in the species *V. canescens*, suggesting that alphanudivirus genome integration occurred, at most, a few million years ago, whereas betanudivirus genome integration occurred ~100 million years ago (*14*). Overall, our data indicate that the hijacking of nudiviruses has contributed at least twice to virulence strategies in parasitic wasps. However, the integrated nudivirus genomes have evolved differently. In particular, the structural components of the nucleocapsid, encoded by *vlf-1*, *38K*, and *vp39*, have been specifically lost from *V. canescens*, despite the relatively recent nature of the alphanudivirus integration (Fig. 2C). This study thus provides the first documented example of related endogenous viruses evolving differently within eukaryote host genomes according to the function conferred on the host (gene transfer versus protein delivery). This is of particular interest because these viruses initially had very similar genetic potentials. Their different uses by the parasitic wasps suggest that stochastic phenomena may have played a role in endogenous virus evolution. Indeed, we hypothesize that, shortly after the initial integration of the nudivirus ancestors into the genome, regulatory sequences allowing the packaging of viral DNA into nucleocapsids may have been lost in the case of *V. canescens* and translocated in the case of braconid wasps, resulting in the absence of viral DNA in the particles, or its replacement by wasp DNA, conferring an advantage during parasitism.

Another unresolved question concerns the fate of the former endogenous symbiotic virus in *V. canescens*, which belongs to a group of parasitic wasps (*23*) (fig. S2A) normally associated with a vertically transmitted ichnovirus with a specific set of genes (*6*). We found several remnants of the endogenous ichnovirus (fig. S2B and table S2). However, the few typical ichnovirus virulence genes identified were not expressed in the calyx. Thus, in *V. canescens*, domestication of the nudiviral machinery has led to the replacement of ichnovirus particles with VLPs to ensure progeny survival (Fig. 3). This constitutes the first reported case of endogenous mutualistic virus replacement.

CONCLUSION

Overall, our data indicate that virus integration has occurred repeatedly during the evolution of parasitic wasps, conferring different functions (gene or protein delivery). The domestication of a whole set of virus genes conferring a new function may thus be a widespread but currently underestimated evolutionary mechanism. Thus, endogenous viruses from parasitic wasps are fascinating from an evolutionary perspective. In addition, VLPs can be considered as natural viral "liposomes" transporting virulence proteins into host target cells. They contain the whole set of envelope proteins responsible for the entry of baculovirus (*16*) particles into insect cells. They may, therefore, enter host hemocytes (that adhere to the wasp egg during the encapsulation process) in a similar manner, thereby protecting the egg by transferring proteins suppressing the immune response in these cells. Recent decades have seen the investment of considerable amounts of energy and money into the development of liposomes or Virus-Like-Particles for delivering molecules of therapeutic interest to target cells. We show here that the natural production of such vectors was selected in parasitic wasp lineages long before the first primates evolved and their descendants were able to develop biotechnologies. The study of the natural production of such vectors in parasitic wasps may inspire the design of new vectors for therapeutic purposes.

MATERIALS AND METHODS

Insects

The thelytokous and arrhenotokous laboratory strains of the solitary parasitoid wasp *V. canescens* (Ichneumonidae, Campopleginae) studied here originated from natural populations collected in the Valence area (France). The two strains were mass-reared separately on larvae of a natural host, *Ephestia kuehniella* (Pyralidae), maintained on maize flour in a growth chamber at 25° ± 2°C under a 16-hour light/8-hour dark photoperiod. Wasps from Tokyo (Japan) and France (two populations from the Nice region and three populations from the Valence region) were preserved in ethanol and sent to our laboratory.

Transmission electron microscopy

Samples were treated as previously described by Doremus *et al.* (*24*). Ovaries from thelytokous wasp pupae (13-day-old pupae) were dissected in 1× phosphate-buffered saline (PBS) solution, transferred, and fixed by overnight incubation at 4°C with 2% (v/v) glutaraldehyde in 0.1 M cacodylate buffer (pH 7.4) and then postfixed by incubation with 2% (v/v) osmium tetroxide in the same buffer for 1 hour at room temperature. Samples were dehydrated by passage through a series of concentrations of ethanol and embedded in Epon. Ultrathin sections of calyx, contrast-stained with uranyl acetate and lead citrate, were examined under a Hitachi H7100 transmission electron microscope.

TEM immunogold labeling

Ovaries from 12-hour-old females were dissected and fixed by incubation for 1 hour in 1× PBS–4% paraformaldehyde at 4°C, washed with PBS, and dehydrated by passage through a series of ethanol solutions (30 to 100%) before progressive embedding in 100% Unicryl resin (EMS Sciences). Ultrathin sections were cut with a Diatome knife on an LKB ultramicrotome and collected on nickel grids. Immunogold electron microscopy was performed with polyclonal rabbit antibodies against the VLP2 (the wasp GTPase activating protein) or PIF-0 proteins. Sections were treated for 10 min with 50 mM NH_4Cl and blocked by two incubations (for 10 min each) with 1% bovine serum albumin (BSA). The sections were washed in 1× PBS and incubated with the primary antibody diluted 1:500 in 0.5% BSA in PBS for 2 hours at room temperature. The preimmune serum was used as control. After several washes in 1× PBS, sections were incubated with goat antirabbit immunoglobulin G antibodies conjugated to 20-nm gold particles (Abcam) and diluted 1:50 in 1× PBS supplemented with 0.5% BSA for 45 min at room temperature. Finally, sections were stained with 1% acetate uranyl and examined under a Hitachi H7100 electron microscope at 80 kV.

Sequencing and assembly of the calyx transcriptome

A complementary DNA (cDNA) library was generated from total RNA extracted from the calyx of thirty 24-hour-old females from the thelytokous laboratory strain. We used the RNeasy Mini Kit (Qiagen), according to the manufacturer's protocol, to extract RNA immediately after calyx dissection in 1× PBS. The RNA was treated with the TURBO DNA-free kit (Life Technologies) according to the manufacturer's instructions, and the mRNA (2.17 μg) was converted into cDNA with the TruSeq RNA Sample Preparation Kit (Illumina). Paired-end sequencing [2 × 100–base pair (bp) reads] was performed on a HiSeq 2000 sequencing system (Illumina) at the Genoscope national sequencing center at a depth of 187,964,116 total reads. Paired-end reads were cleaned by a three-step procedure: (i) sequencing adapters and low-quality nucleotides (quality

value <20) were removed; (ii) sequences between the second unknown nucleotide (N) and the end of the read were removed; and (iii) reads shorter than 30 nucleotides after trimming, together with reads and their mates that mapped onto the quality control sequences, were discarded (PhiX genome). We assembled a total of about 380 M cleaned reads de novo with Velvet 1.2.07 (*25*) and Oases 0.2.08 (*26*), using a *k*-mer size of 51 bp. This two-step assembly produced 47,733 contigs of more than 100 bp in length (with a cumulative size of 77 Mb). Finally, the resulting contigs were compared with the sequences in the nonredundant and environmental National Center for Biotechnology Information (NCBI) protein databases with the Blast2Go online tool, using BLASTX ($e < 1 \times 10^{-3}$).

Genome sequencing and assembly

Genomic DNA was obtained from 18 *V. canescens* males from a single unfertilized mother (laboratory arrhenotokous strain), using the Wizard Genomic DNA Purification Kit (Promega). We sent 9.4 µg of DNA (quantified by Qubit flurometer) to the Genoscope national sequencing center. Paired-end libraries were prepared with NEBNext DNA Library Prep Kit for Illumina. Paired-end (300 to 600 bp) sequencing was conducted on a HiSeq 2000 sequencing system (Illumina) (40 Gb raw data, ~100× estimated coverage). Reads were trimmed and cleaned by the same three-step procedure as used for RNA samples. In total, 195 M cleaned reads were assembled with VelvetOptimiser (http://bioinformatics.net.au/software.velvetoptimiser.shtml). The parameters retained for Velvet 1.2.07 (*25*) were a *k*-mer size of 81 bp and a coverage cutoff at 5.22×. We obtained 62,001 scaffolds, with an N50 of about 115 kb and a cumulative size of 238 Mb. *V. canescens* genome scaffolds were screened with BLASTp or BLASTn to identify the scaffolds with nudivirus and ichnovirus genes. The cDNA reads were mapped onto the wasp genome scaffolds with the short-read aligner BOWTIE 1.1.1 (*27*). The parameters chosen were *–v* 0 and *–a*. The number of reads was determined, and RPKM was calculated for all open reading frames identified with the Artemis interface (http://sanger.ac.uk/resources/software/artemis/#development). Note that the RNA polymerase *lef-9* homolog was found to contain two introns, a result confirmed by RT-PCR and sequencing analyses; comparisons with other *lef-9* sequences devoid of introns [OrNV *lef-9* (YP_002321407) or *Nilaparvata lugens* endogenous nudivirus *lef-9* (AHW98255)] suggested that intron sequences had been inserted and rearrangements had occurred in the second exon region of the *V. canescens* sequence.

Identification of nudiviral proteins in purified VLPs

VLPs were purified from the oviducts of forty 4-day-old females (thelytokous strain) using a modified version of a published protocol (*28*). Briefly, the ovaries were dissected in 1× PBS. Oviducts were cut out with dissection forceps and transferred to 50 µl of 1× PBS in an Eppendorf tube, which was then centrifuged (770*g*, 10 min, 4°C). The supernatant, containing VLPs and oviduct fluid, was transferred to a new tube and recentrifuged (15,400*g*, 10 min, 4°C). The pellet containing purified VLPs was subjected to elution in 20 µl of 1× PBS and frozen at −80°C. Proteins were separated under denaturing conditions by SDS–polyacrylamide gel electrophoresis in a 12% polyacrylamide gel, and the bands were detected with PAGE-Blue protein staining solution (Fermentas). Gel lanes were cut into 24 slices and destained by three washes in 50% acetonitrile and 50 mM ammonium bicarbonate. The proteins were reduced (10 mM dithiothreitol at 56°C for 45 min) and alkylated (55 mM iodoacetamide at room temperature for 30 min), and were then digested in-gel with Trypsin Gold (15 ng/ml) (Promega) in 100 mM ammonium bicarbonate. The peptides generated were extracted from the gel matrix in 50% acetonitrile and 5% formic acid. Peptide solutions were dehydrated in a vacuum centrifuge. Samples were analyzed online by nanoflow high-performance liquid chromatography (HPLC)–nanoelectrospray ionization with an LTQ Orbitrap Xl mass spectrometer (Thermo Fisher Scientific) coupled to an Ultimate 3000 HPLC machine (Dionex). Samples were desalted and preconcentrated online on a Pepmap precolumn (0.3 mm × 10 mm; Dionex). Peptides were subjected to HPLC on a capillary reversed-phase C18 column (0.075 mm × 150 mm; Pepmap, Dionex), from which they were eluted with a gradient of 0 to 40% B over 60 min (A = 0.1% formic acid and 2% acetonitrile in water; B = 0.1% formic acid in acetonitrile) at a flow rate of 300 nl/min. Spectra were recorded with Xcalibur v2.0.7 software (Thermo Fisher Scientific) and were acquired using the instrument operating in the information-dependent acquisition mode over the entire HPLC gradient. The mass scanning range (*m*/*z*) was 400 to 2000, the capillary temperature was 200°C, and the standard mass spectrometric conditions included an ion spray voltage of 2.20 kV, a capillary voltage of 40 V, and a tube lens voltage of 120 V. Cycles of one full-scan mass spectrum at a resolution of 60,000, followed by five data-dependent MS/MS spectra, were repeated continually throughout the nano-LC separation. MS/MS spectra, obtained with the LTQ Orbitrap Xl mass spectrometer, were recorded with a normalized collision energy of 35% (activation $Q = 0.25$; activation time, 30 ms) and an isolation window of 2 *m*/*z*. All MS/MS spectra were used as queries against entries from a restricted database consisting of the transcriptome assembly described above, together with sequences from VLP virulence proteins obtained from NCBI and nudiviral proteins obtained from the genome assembly with Proteome Discoverer v1.4 software (Thermo Fisher Scientific) and Mascot v2.4 (Matrix Science; http://matrixscience.com). Carbamidomethylation was set as a fixed modification, and methionine oxidation was set as a variable modification for searches. Peptide identifications were accepted on the basis of Percolator results, with a false discovery rate of 1%. We considered only proteins identified with two or more peptides. We manually validated peptides aligned with nudiviral protein sequences obtained from the genome assembly and peptides aligned with three virulence proteins retrieved from the genome assembly.

Amplification of nudiviral sequences in *V. canescens* wasps of different origins

Males and females from arrhenotokous laboratory strains of *V. canescens* collected in the Valence area (France) in 2007 (Valence 1) and 2010 (Valence 2) and from a laboratory strain collected in the Nice area (France) (Nice 1) were sent to our laboratory, together with females from a strain collected in Tokyo (Japan). Twenty *V. canescens* females from wild populations in the Valence (Valence 3) and Nice (Nice 2) areas were also used in this experiment. Total DNA was extracted with the QiaAmp kit (Qiagen). *V. canescens*–specific primers for three nudiviral genes—*Ac81-3*, *OrNVorf18-like*, and *pif-6*—were selected for PCR amplification (primer sequences presented in table S3). The wasp *actin* gene was used as a positive control for amplification. A 30-cycle PCR (94°C for 60 s, 58°C for 60 s, and 72°C for 60 s) was performed with 10 pmol of each primer, 0.2 mM deoxynucleotide triphosphate (dNTP; MP Biochemicals), 1.5 mM $MgCl_2$, 0.5 U of Goldstar (Eurogentec), and 1 µl of total DNA. PCR products (8 µl) were run on 2% agarose gels. As control, PCR was also performed with DNA extracted from *Hyposoter dydimator*, a related wasp species, for which no amplicon

was obtained. In all experiments, a PCR was carried out in the absence of a DNA template, as a negative control.

Expression of nudiviral genes

The transcription of nudiviral genes was quantified in *V. canescens* thelytokous female adults and pupae on the sixth [pupal stage 1 (P1)], seventh [pupal stage 2 (P2)], eighth [pupal stage 3 (P3)], and ninth [pupal stage 4 (P4)] days of pupal development. Three biological replicates were used per developmental stage. For each biological replicate and each developmental stage, the head with thorax, calyx, and ovarioles of 12 individuals were dissected in 1× PBS. Total RNA was extracted from each sample with the RNeasy Mini Kit (Qiagen). After DNase treatment with the TURBO DNA-free kit (Life Technologies), RT (200 ng of total RNA) was carried out using the SuperScript III Reverse Transcriptase kit (Life Technologies) with 500 ng of oligo$(dT)_{15}$ primer (Promega) and 1 µl of 10 mM dNTP mix (Promega). Four *V. canescens* nudiviral genes were quantified in this experiment: *OrNVorf41-like.2*, *OrNVorf59-like*, *pif-0* (*p74*), and *RNA polymerase lef-8*. The wasp genes *Elf1*, *actin*, and *GAPDH* were used as reference genes (primer sequences presented in table S3). For each gene, a standard curve was prepared with plasmid DNA. Briefly, genomic DNA was extracted from thelytokous females with the Wizard Genomic DNA Purification Kit (Promega). Each gene was amplified by PCR and inserted into the pGEM-T vector (Promega) according to the manufacturer's protocol. All of the clones generated were used to transform competent *Escherichia coli* TOP10 cells (IBA). Plasmid DNA was extracted with the QIAprep Spin Miniprep kit (Qiagen) and serially diluted (1e2 to 1e7 plasmid copies). LightCycler 480 SYBR Green I Master Mix (Roche) was used for the analysis of RNA expression in 384-well plates (Roche). The total volume of 10 µl in each well included 5 µl of reaction mix, 2.2 or 2.6 µl of nuclease-free water, 0.2 µl of reference gene primers or 0.4 µl of nudiviral gene primers, and 2 µl of cDNA or diluted plasmid DNA. We determined cDNA levels with a LightCycler 480 System (Roche) through heating at 95°C for 10 min, followed by 45 cycles of 95°C for 10 s, 58°C for 10 s, and 72°C for 10 s. Each sample was evaluated in triplicate, and the results were normalized with respect to the genes *Elf1*, *GAPDH*, and *actin*. The absolute values presented are the means of normalized results, and error bars indicate standard deviations.

Phylogenetic analysis of insect large double-stranded DNA viruses

The tree was obtained by maximum likelihood inference analysis of a concatenated multiple alignment (performed with the ClustalOmega program) of amino acid sequences corresponding to 37 viral genes from the bracovirus, nudivirus, and baculovirus lineages previously used for a similar phylogenetic analysis (*15*). The hytrosavirus lineage was used as an outgroup. Support for nodes is indicated by maximum likelihood nonparametric bootstraps (100 replicates). The sequences used for alignment with the nudiviral genes involved in the production of *V. canescens* VLPs were obtained from the following viruses: *Cotesia congregata* bracovirus (CcBV), *Microplitis demolitor* bracovirus (MdBV), *Chelonus inanitus* bracovirus (CiBV), *Heliothis zea* nudivirus 1 (HzNV), *Penaeus monodon* nudivirus (PmNV), *Gryllus bimaculatus* nudivirus (GbNV), OrNV, *Drosophila innubia* nudivirus (DiNV), *Culex nigripalpus* nucleopolyhedrovirus (CuniNPV), *Neodiprion sertifer* nucleopolyhedrovirus (NeseNPV), *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV), *Lymantria dispar* multiple nucleopolyhedrovirus (LdMNPV), *Cydia pomonella* granulovirus (CpGV), *Musca domestica* salivary gland hytrosavirus (MdSGHV), and *Glossina pallidipes* salivary gland hytrosavirus (GpSGHV).

Wasp phylogeny analysis

A PCR product corresponding to a fragment of the cytochrome oxidase 1 (*CO1*) gene was obtained from *V. canescens* DNA, extracted as described above. A MegaBlast search against the GenBank database was performed to retrieve the *CO1* sequences of the ichneumonid wasps *Tranosema rostrale* (HQ107907), *Campoletis sonorensis* (DQ538849), *Hyposoter exiguae* (DQ538855), *Hyposoter fugitivus* (DQ538858), and *Glypta fumiferanae* (HQ107353); the braconid wasp *C. congregata* (DQ538816); and the pteromalid wasp *Nasonia vitripennis* (EU746534), which was used as an outgroup. Sequences were aligned using the MUSCLE algorithm (available from http://phylogeny.lirmm.fr/), and the optimal nucleotide substitution model for each locus was chosen on the basis of the Bayesian information criterion, as implemented in jModeltest 2.1. We performed a Bayesian analysis with MrBayes 3.1.2, with the selected model GTR + δ and the following parameters: five independent Markov chain Monte Carlo iterations, 1 million generations with tree and parameter sampling occurring every 100 generations, and *N. vitripennis* (as the outgroup). The first 25% of trees were discarded as burn-in, leaving 750,000 trees per run, and we determined the maximum clade credibility tree with a posterior probability limit of 0.5. The resulting tree was visualized with MEGA software.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/1/10/e1501150/DC1

Fig. S1. Link between nudiviral genes and *V. canescens* VLPs.

Fig. S2. Vestiges of the ichnovirus in the *V. canescens* genome.

Table S1. Nudiviral and flanking wasp sequences identified in the *V. canescens* genome scaffolds.

Table S2. Selected list of ichnovirus-like sequences identified in the *V. canescens* genome.

Table S3. List of PCR and quantitative PCR primers used to amplify nudiviral genes and wasp housekeeping genes.

REFERENCES AND NOTES

1. E. C. Holmes, The evolution of endogenous viral elements. *Cell Host Microbe* **10**, 368–377 (2011).
2. C. Feschotte, C. Gilbert, Endogenous viruses: Insights into viral evolution and impact on host biology. *Nat. Rev. Genet.* **13**, 283–296 (2012).
3. M. R. Strand, G. R. Burke, Polydnaviruses: Nature's genetic engineers. *Annu. Rev. Virol.* **1**, 333–354 (2014).
4. M. R. Strand, G. R. Burke, Polydnaviruses as symbionts and gene delivery systems. *PLOS Pathog.* **8**, e1002757 (2012).
5. A. Bézier, M. Annaheim, J. Herbinière, C. Wetterwald, G. Gyapay, S. Bernard-Samain, P. Wincker, I. Roditi, M. Heller, M. Belghazi, R. Pfister-Wilhem, G. Periquet, C. Dupuy, E. Huguet, A.-N. Volkoff, B. Lanzrein, J.-M. Drezen, Polydnaviruses of braconid wasps derive from an ancestral nudivirus. *Science* **323**, 926–930 (2009).
6. A.-N. Volkoff, V. Jouan, S. Urbach, S. Samain, M. Bergoin, P. Wincker, E. Demettre, F. Cousserans, B. Provost, F. Coulibaly, F. Legeai, C. Béliveau, M. Cusson, G. Gyapay, J.-M. Drezen, Analysis of virion structural components reveals vestiges of the ancestral ichnovirus genome. *PLOS Pathog.* **6**, e1000923 (2010).
7. G. Salt, Experimental studies in insect parasitism. XIII. The haemocytic reaction of a caterpillar to eggs of its habitual parasite. *Proc. R. Soc. Lond. B Biol. Sci.* **162**, 303–318 (1965).
8. S. Rotheram, Immune surface of eggs of a parasitic insect. *Nature* **214**, 700 (1967).
9. S. Rotheram, The surface of the egg of a parasitic insect. II. The ultrastructure of the particulate coat on the egg of *Nemeritis*. *Proc. R. Soc. Lond. B Biol. Sci.* **183**, 195–204 (1973).
10. I. Feddersen, K. Sander, O. Schmidt, Virus-like particles with host protein–like antigenic determinants protect an insect parasitoid from encapsulation. *Experientia* **42**, 1278–1281 (1986).

11. O. Bedwin, An insect glycoprotein: A study of the particles responsible for the resistance of a parasitoid's egg to the defence reactions of its insect host. *Proc. R. Soc. Lond. B Biol. Sci.* **205**, 271–286 (1979).
12. A. Reineke, S. Asgari, O. Schmidt, Evolutionary origin of *Venturia canescens* virus-like particles. *Arch. Insect Biochem. Physiol.* **61**, 123–133 (2006).
13. Y. Wang, J. A. Jehle, Nudiviruses and other large, double-stranded circular DNA viruses of invertebrates: New insights on an old topic. *J. Invertebr. Pathol.* **101**, 187–193 (2009).
14. J. Thézé, A. Bézier, G. Periquet, J.-M. Drezen, E. A. Herniou, Paleozoic origin of insect large dsDNA viruses. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15931–15935 (2011).
15. A. Bézier, J. Thézé, F. Gavory, J. Gaillard, J. Poulain, J.-M. Drezen, E. A. Herniou, The genome of the nucleopolyhedrosis-causing virus from *Tipula oleracea* sheds new light on the *Nudiviridae* family. *J. Virol.* **89**, 3008–3025 (2015).
16. G. F. Rohrmann, *Baculovirus Molecular Biology* (National Center for Biotechnology Information, Bethesda, MD, ed. 3, 2013), www.ncbi.nlm.nih.gov/books/NBK114593/
17. A. M. Crawford, C. Sheehan, Replication of *Oryctes* baculovirus in cell culture: Viral morphogenesis, infectivity and protein synthesis. *J. Gen. Virol.* **66**, 529–539 (1985).
18. A. Bézier, J. Herbinière, B. Lanzrein, J. M. Drezen, Polydnavirus hidden face: The genes producing virus particles of parasitic wasps. *J. Invertebr. Pathol.* **101**, 194–203 (2009).
19. S. Asgari, A. Reineke, M. Beck, O. Schmidt, Isolation and characterization of a neprilysin-like protein from *Venturia canescens* virus-like particles. *Insect Mol. Biol.* **11**, 477–485 (2002).
20. U. Theopold, E. Krause, O. Schmidt, Cloning of a VLP-protein coding gene from a parasitoid wasp *Venturia canescens*. *Arch. Insect Biochem. Physiol.* **26**, 137–145 (1994).
21. M. Hellers, M. Beck, U. Theopold, M. Kamei, O. Schmidt, Multiple alleles encoding a virus-like particle protein in the ichneumonid endoparasitoid *Venturia canescens*. *Insect Mol. Biol.* **5**, 239–249 (1996).
22. G. R. Burke, K. K. O. Walden, J. B. Whitfield, H. M. Robertson, M. R. Strand, Widespread genome reorganization of an obligate virus mutualist. *PLOS Genet.* **10**, e1004660 (2014).
23. D. L. J. Quicke, *The Braconid and Ichneumonid Parasitoid Wasps: Biology, Systematics, Evolution and Ecology* (John Wiley and Sons, UK, 2014).
24. T. Dorémus, S. Urbach, V. Jouan, F. Cousserans, M. Ravallec, E. Demettre, E. Wajnberg, J. Poulain, C. Azéma-Dossat, I. Darboux, J.-M. Escoubas, D. Colinet, J.-L. Gatti, M. Poirié, A.-N. Volkoff, Venom gland extract is not required for successful parasitism in the polydnavirus-associated endoparasitoid *Hyposoter didymator* (Hym. Ichneumonidae) despite the presence of numerous novel and conserved venom proteins. *Insect Biochem. Mol. Biol.* **43**, 292–307 (2013).
25. D. R. Zerbino, E. Birney, Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).
26. M. H. Schulz, D. R. Zerbino, M. Vingron, E. Birney, Oases: Robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* **28**, 1086–1092 (2012).
27. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
28. M. Beck, U. Theopold, O. Schmidt, Evidence for serine protease inhibitor activity in the ovarian calyx fluid of the endoparasitoid *Venturia canescens*. *J. Insect Physiol.* **46**, 1275–1283 (2000).

Acknowledgments: We thank E. Desouhant and N. Ris for providing the *V. canescens* wasps, D. Quicke for access to the phylogenetic analysis of campoplegine wasps, and F. Godiard for help and access to the Electron Microscopy Service of Montpellier University. English language usage was revised by A. Edelman and Associates. **Funding:** This work was supported by the Genoscope and the French National Research Agency (Paratoxose ANR-09-BLAN-0243-01). **Author contributions:** A.P. performed and supervised the experiments and participated in the writing of the manuscript. A.B. helped with nudiviral sequences and virus phylogeny analyses and contributed to the writing of the manuscript. S.U. performed the proteomic analyses. J.-M.A. contributed to the Illumina sequence analyses. V.J. participated in the experimental work. M.R. performed the TEM work. F.W. contributed to the bioinformatics work. J. Guy participated in obtention and analyses of Illumina sequencing data. F.C. contributed to the transcriptome sequence analyses. J. Gauthier contributed to the bioinformatics work. E.D. contributed to the VLP protein analyses. S.S. participated in wasp genome analysis. V.S. performed the PCR analyses on wasp populations. J.T. performed the phylogeny analyses. M.P. contributed to the writing of the manuscript. D.C. participated in the work on transcriptome libraries and in the editing of the manuscript. C.d.S. and A.C. helped with sequencing and sequence analyses. V.B. managed the Illumina transcriptome and genome sequencing. J.-M.D. and A.-N.V. directed the study and wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors and GenBank accession numbers for *V. canescens* nucleotide sequences: KP972595-600 and ERA420110.

Submitted 24 August 2015
Accepted 15 October 2015
Published 27 November 2015
10.1126/sciadv.1501150

Citation: A. Pichon, A. Bézier, S. Urbach, J.-M. Aury, V. Jouan, M. Ravallec, J. Guy, F. Cousserans, J. Thézé, J. Gauthier, E. Demettre, S. Schmieder, F. Wurmser, V. Sibut, M. Poirié, D. Colinet, C. da Silva, A. Couloux, V. Barbe, J.-M. Drezen, A.-N. Volkoff, Recurrent DNA virus domestication leading to different parasite virulence strategies. *Sci. Adv.* **1**, e1501150 (2015).

ScienceAdvances
AAAS

Recurrent DNA virus domestication leading to different parasite virulence strategies

Apolline Pichon, Annie Bézier, Serge Urbach, Jean-Marc Aury, Véronique Jouan, Marc Ravallec, Julie Guy, François Cousserans, Julien Thézé, Jérémy Gauthier, Edith Demettre, Sandra Schmieder, François Wurmser, Vonick Sibut, Marylène Poirié, Dominique Colinet, Corinne da Silva, Arnaud Couloux, Valérie Barbe, Jean-Michel Drezen and Anne-Nathalie Volkoff (November 27, 2015)
Sci Adv 2015, 1:.
doi: 10.1126/sciadv.1501150