HAL
open science

Investigating bracovirus chromosomal integration and inheritance in lepidopteran host and nontarget species

Héloïse Muller, Camille Heisserer, Taiadjana Fortuna, Florence Mougel, Elisabeth Huguet, Laure Kaiser, Clément Gilbert

To cite this version:

Héloïse Muller, Camille Heisserer, Taiadjana Fortuna, Florence Mougel, Elisabeth Huguet, et al.. Investigating bracovirus chromosomal integration and inheritance in lepidopteran host and nontarget species. Molecular Ecology, In press, 10.1111/mec.16685 . hal-03799713

HAL Id: hal-03799713

https://hal.science/hal-03799713

Submitted on 10 Oct 2022

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

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

Investigating bracovirus chromosomal integration and inheritance in lepidopteran host and non-target species

Héloïse Muller[1], Camille Heisserer[1,2], Taiadjana Fortuna[1], Florence Mougel[1], Elisabeth Huguet[2], Laure Kaiser[1], Clément Gilbert[1*]

[1]Université Paris-Saclay, CNRS, IRD, UMR Évolution, Génomes, Comportement et Écologie, 91190 Gif-sur-Yvette, France.

[2]UMR 7261 CNRS, Institut de Recherche sur la Biologie de l'Insecte, Faculté des Sciences et Techniques, Université de Tours, Tours, France

*Author for Correspondence: Clément Gilbert, Université Paris-Saclay, CNRS, IRD, UMR Évolution, Génomes, Comportement et Écologie, 91190 Gif-sur-Yvette, France; email: clement.gilbert@egce.cnrs-gif.fr

Keywords: bio-control | Braconidae | bracoviruses | evolutionnary genomics | horizontal transfer | parasitoid wasps

Abstract

Bracoviruses (BVs) are domesticated viruses found in braconid parasitoid wasp genomes. They are composed of domesticated genes from a nudivrius, coding viral particles in which wasp DNA circles are packaged. BVs are viewed as possible vectors of horizontal transfer of genetic material (HT) from wasp to their hosts because they are injected, together with wasp eggs, by female wasps in their host larvae, and because they undergo massive chromosomal integration in multiple host tissues. Here, we show that chromosomal integrations of the *Cotesia typhae* BV (CtBV) persist up to the adult stage in individuals of its natural host, *S. nonagrioides*, that survived parasitism. However, while reproducing host adults can bear an average of nearly two CtBV integrations per haploid genome, we were unable to retrieve any of these integrations in 500 of their offspring using Illumina sequencing. This suggests either that host gametes are less targeted by CtBVs than somatic cells or that gametes bearing BV integrations are non-functional. We further show that CtBV can massively integrate into the chromosomes of other lepidopteran species that are not normally targeted by the wasp in the wild, including one which is at least 100 million years divergent from the natural host. Cell entry and chromosomal integration of BVs are thus unlikely to be major factors shaping wasp host range. Together, our results shed new light on the conditions under which BV-mediated wasp-to-host HT may occur and provide information that may be helpful to evaluate the potential risks of uncontrolled HT associated with the use of parasitoid wasps as biocontrol agents.

Introduction

Horizontal transfer (HT) of genetic material involves transmission of DNA by means other than reproduction (Keeling & Palmer, 2008). Widespread in prokaryotes, HT is also increasingly recognized as an important evolutionary process in eukaryotes (Aubin et al., 2021; Boto, 2014; Van Etten & Bhattacharya, 2020). Many studies report remarkable examples of HT of genes with important consequences in multiple lineages (Danchin et al., 2016; Simion et al., 2021; Wybouw et al., 2016; Xia et al., 2021; Li et al. 2022). In addition, large scale studies of transposable elements (TEs) inferred dozens to thousands of horizontal transfers of transposable elements in various multicellular eukaryotes such as plants (Baidouri et al., 2014), insects (Peccoud et al., 2017; Heringer et al. 2022; Wallau et al. 2016), and vertebrates (Zhang et al., 2020). Given the profound impact TEs have on genome evolution, HT of TEs can be viewed as a key process shaping eukaryote genomes (Gilbert & Feschotte, 2018). Contrasting with our detailed knowledge of HT in prokaryotes, the mechanisms and possible vectors of HT remain elusive in eukaryotes. The finding that viruses and extracellular vesicles can carry host genetic material suggests that they may accidentally act as vectors of HT (Gilbert & Cordaux, 2017; Ono et al., 2019). Here we further evaluate the conditions and frequency at which a specific type of viruses, the bracoviruses (BVs) of Braconidae endoparasitoid wasps (Hymenoptera), may mediate HT between wasps and lepidopterans.

Braconidae endoparasitoid wasps lay their eggs in insect hosts, mostly lepidopterans (see the picture of Fig. 1 as an example), which are used by wasp larvae as substrate and food source. In addition to their eggs, females of many Braconidae wasp species inject BVs into their host, which interfere with host developmental and immunological pathways, facilitating the development of wasp embryos (Beckage & Drezen, 2011; Strand & Burke, 2014). Genes

encoding structural components of these viruses originate from domestication of viral genes that became integrated into the genome of braconid wasp ancestors about 100 million years ago (labeled 1c in Fig. 1) (Gauthier et al., 2018; Herniou et al., 2013). Bracovirus structural genes, of nudiviral origin, are specifically expressed in the calyx of the ovaries, resulting in the production of viral particles into which circular double stranded DNA molecules are packaged (labeled 2 in Fig. 1). These packaged DNA circles result from amplification, excision and circularization of so-called "proviral segments" (labeled 1a and 1b in Fig. 1), which contain wasp genes and other genes of unknown evolutionary origin (Bézier et al., 2013; Drezen et al., 2014; Herniou et al., 2013). Circularization takes place in calyx cells *via* site-specific recombination at Direct Repeat Junction (DRJ) motifs, which are conserved between proviral segments and between Braconidae species (Beck et al., 2011; Desjardins et al., 2008; Gauthier et al., 2021; Gruber et al., 1996; Muller, Chebbi, et al., 2021; Pasquier-Barre et al., 2002; Savary et al., 1997). A remarkable feature of BVs is that once released in host's cells, DNA circles containing a specific motif called Host Integration Motif (HIM) integrate massively into the host genome (labeled 3 in Fig. 1) (Beck et al., 2011; Chevignon et al., 2018). We use the term "HIM-containing circles" to refer to these circles and we call other circles, which do not contain HIM "circles devoid of HIM". Integration involves site-specific recombination between HIMs and the host genome, likely mediated by BV-encoded enzymes. In the wasp *Cotesia typhae*, it was shown that the 16 HIM-containing circles (out of 27 circles) undergo chromosomal integration in at least four host tissues/organs (hemocytes, fat body, ganglionic chain, and head), yielding an average of 12 to 85 integrated circles per host haploid genome (Muller, Chebbi, et al., 2021). Integration involves linearization of DNA circles through two DNA double strand breaks within HIMs and loss of a small ≃50-bp region lying between the two breaks (labeled 3 in Fig. 1). The

role played by BV DNA circles integration in wasp-host interactions remains to be characterized. So far, it appears that integration does not enhance expression of circle-borne genes (Chevignon et al., 2014, 2018), nor does it enhance persistence of DNA circles throughout wasp development (Muller, Chebbi, et al., 2021).

Systematic and massive DNA circle integration in host genomes during parasitism may facilitate wasp-to-host HT with long-term consequences. Some of BV circle sequences have been found in various lepidopteran genomes that clearly result from HT from wasp to lepidopterans (Gasmi et al., 2015; Schneider & Thomas, 2014). Remarkably, some BV genes acquired by *Spodoptera* species have been domesticated and now play a role in anti-bacterial immune response (Di Lelio et al., 2019; Gasmi et al., 2015). Thus, BV DNA circles can integrate into the genome of the host germline and some host individuals surviving parasitism can transmit these integrations to their offspring (Drezen et al., 2017). These studies also demonstrated that horizontally transferred BV genes can be an important source of genetic novelty and adaptation in lepidopterans.

Integration in the host germline is consistent with the known tropism of BV particles for a large range of tissues (Beck et al., 2011; Wyder et al., 2003) and the capacity of BV DNA circles to integrate in the genome of various cell types (Muller, Chebbi, et al., 2021), although at different levels depending on the species and tissues. Survival of parasitized lepidopterans is a rather common phenomenon in nature, as the success of endoparasitoid wasps' development often strongly varies depending on the geographical origin of host populations (Gitau et al., 2007). To which extent resistant hosts retain BV DNA integrated in their genome at the adult stage, and in which proportion these wasp sequences can be transmitted vertically to the next host

generation has never been investigated. Yet, given that some endoparasitoid wasps are used as biocontrol agents (Polaszek & Walker, 1992), measuring the frequency of such possible HT is relevant not only from a basic science point of view, but also to evaluate the risks of uncontrolled HT potentially associated with strategies that use exotic wasps bearing BV.

Another possible way chromosomally integrated BV circles may be vertically transmitted in host populations is through accidental parasitism of a non-permissive host species, in which by definition parasitism fails. In laboratory conditions, endoparasitoid wasps are indeed able to lay their eggs into species not known to be targeted in the wild, with varying degrees of success (Harwood et al., 1998). Whether BV DNA circle integration occurs in non-natural hosts, supporting such hosts as possible facilitators of wasp-to-host HT of BV DNA, has never been investigated so far.

In this study, we first evaluate the frequency under which integration of BV DNA circles in host germline followed by transmission to the next generation may occur in the wasp-host system involving the wasp *Cotesia typhae* and its host, the Mediterranean corn borer *Sesamia nonagrioides* (Lepidoptera: Noctuidae). S*. nonagrioides* is a major pest of corn in Mediterranean regions (Bosque-Perez & Schulthess, 1998; Kankonda et al., 2018; Moyal et al., 2011). *C. typhae* (Fernandez-Triana) (Hymenoptera: Braconidae) is a newly described species closely related to *Cotesia sesamiae* found in east Africa (Kaiser et al., 2015, 2017). The two wasp species are reproductively isolated and although the former is a generalist endoparasitoid, *C. typhae* is a specialist at both host (*S. nonagrioides*) and plant (*Typha domingensis*) levels. Due to such a specialization, in addition to inducing a high host death rate during parasitism, *C. typhae* is currently under study as a potential biological control agent

against *S. nonagrioides* in Europe (Kaiser et al., 2015, 2017). We then assessed whether *C. typhae* BV DNA circles integrate into the genome of three other stemborer lepidopteran species: two Noctuidae (*Globia sparganii* and *Nonagria typhae*) and one Crambidae (*Chilo phragmitella*). These species are normally not targeted in the wild by *C. typhae* but since their distribution overlaps with that of *S. nonagrioides* in France and they share similar ecological niches (Galichet et al., 1992; Teder & Tammaru, 2002; Tewksbury et al., 2002), they might suffer non-target effects during a putative biocontrol campaign, the consequences of which are important to assess (Louda et al., 2003; Sands & Driesche, 2000).

Materials and methods

Parasitism and sequencing of 500 *S. nonagrioides* larvae and two adults

Sixty-two French *S. nonagrioides* larvae at the L4/L5 stage were each parasitized by a different *C. typhae* female. *S. nonagrioides* larvae used for this experiment came either from the wild (collected in France at Sorbets and Coublucq) or from a strain reared at EGCE since 2010 which initially comes from the southwest of France and which is refreshed yearly with French individuals from the wild. In order to maximize the number of resistant *S. nonagrioides*, we chose to parasitized them with the *C. typhae* strain “Makindu”, which shows low virulence on French *S. nonagrioides* larvae, with about 60% of individuals resisting to parasitism (Benoist et al., 2017). This “Makindu” wasp strain is reared at EGCE since 2015, but it initially comes from 10 individuals who emerged from *S. nonagrioides* collected in 2010-2011 from the wild in the Makindu locality (2.278S, 37.825E; South-East Kenya). An isofemale line was then produced and maintained by parasiting at each generation three-weeks-old larvae from the Makindu *S.*

149 *nonagrioides* strain, on which the wasp virulence is higher than on the french strain used for
150 the experiments of this article (Benoist et al., 2020).

151 Oviposition events were all confirmed by visual observation. Here, 37 larvae (59.7%) resisted
152 to parasitism, *i.e.* no wasp pupae emerged from them and they reached adult stage. We
153 obtained a total of 20 males and 17 females, but because of important delays in the
154 development of some individuals, we were able to form 15 couples only. A total of 13 females
155 involved in these couples produced some eggs (sometimes less than 10 eggs), but most of
156 these eggs were either unfertilized or non-viable. Numerous and viable eggs leading to the
157 development of offspring could be recovered for only five females (Suppl. Fig. 1). The names
158 of these F1 offspring refer to their mothers: F1-Adult1, F1-Adult2, F1-Adult3, F1-Adult4, and
159 F1-Adult5. We let all these five progenies develop until the larval stage L3/L4 and we sacrificed
160 them by freezing at -80°C. We then randomly picked a total of 500 caterpillars for the
161 experiment. For technical reasons (high mortality of eggs in some progenies because of mold
162 in our rearing tubes), we obtained less than 100 caterpillars for some couples, impeding us to
163 sample identical numbers of individuals from each crossing. We ended up with 110 offspring
164 individuals for F1-Adult1, 80 for F1-Adult2, 115 for F1-Adult3, 70 for F1-Adult4, and 125 for F1-
165 Adult5 (Suppl. Fig. 1).

166 A piece from each of these caterpillars was then cut with a razor blade and weighed. Pieces of
167 equal weight were pooled five by five for each set of offspring, *i.e.* 100 pools of 5 offspring.
168 The rest of the bodies were individually frozen at -80°C to be able to identify which individuals
169 would bear the eventual BV integration identified in pools of F1 (see below). We extracted
170 DNA from the 100 pools with the kit DNeasy Purification (QIAGEN). We then pooled these DNA

extractions 20 by 20 with equimolar ratio, leading to five libraries composed of 100 caterpillars each (Suppl. Fig. 1): libraries PoolF1-A to PoolF1-E. We sub-contracted Novogen to build a paired-end library (2 x 150 bp; insert size = 350 bp) for each of these five libraries. Each library was then sequenced on an Illumina platform to produce a targeted amount of 100 Gbp each. In addition, the DNA of two resistant *S. nonagrioides* adults was also sequenced. One of these adults (Adult-1) is a female who produced offspring (progeny included in this experiment under the name F1-Adult1), whereas the other (Adult-6) is a female who did not produce any offspring. DNA was extracted from the thorax of these two resistant *S. nonagrioides* females, using the DNeasy Purification kit (QIAGEN) after their natural death. Library construction and sequencing was done by Novogen following the same protocol as the one used for the five other libraries, except that the targeted number of bases produced was 50 Gbp (*i.e.*, a depth of about 50X on the *S. nonagrioides* genome).

Parasitism of non-target species

C. typhae females from the strain “Kobodo” reared at EGCE since 2019, also as an isofemale line but from six parasitized caterpillars collected in the field at Kobodo locality (0.679S, 34.412E; West Kenya), were set to parasitize two caterpillars from the bulrush wainscot *Nonagria typhae* (Lepidoptera: Noctuidae), five from the Webb's wainscot *Globia sparganii* (Lepidoptera: Noctuidae), and eight from the Reed Veneer *Chilo phragmitella* (Lepidoptera: Crambidae). Each of these caterpillars was parasitized by a different wasp, and ovipositon was confirmed by visual observation. These non-natural host species were sampled in the wild in different regions in France in June 2020. The caterpillars of *N. typhae* and *G. sparganii* were collected in Saint Michel en l’Herm (46.354N, 1.260W), and the caterpillars of *C. phragmitella* in Sacy-le-Grand (49.347N, 2.537E). *C. typhae* larvae emerged from four out of these 15

194 parasitized non-natural hosts, all from *C. phragmitella*. Irrespective of whether parasitism was
195 successful or not, all parasitized individuals died between five and 26 days post parasitism, all
196 at late larval or pre-pupal stage. Using the kit NucleoBond AXG20 (Macherey-Nagel), we
197 extracted DNA from the whole body of three larvae, from which no wasp larvae emerged: one
198 *N. typhae* larva, one *G. sparganii* larva, and one *C. phragmitella,* which died 17, 19, and 25
199 days post-parasitism, respectively. We sub-contracted Novogen to build a paired-end library
200 (2 x 150 bp; insert size = 350 bp) for these three samples for sequencing on an Illumina
201 platform to produce a targeted amount of 100 Gb each.

202 **Genome assemblies**

203 The genome of *C. typhae* and *S. nonagrioides* were retrieved from GenBank under accession
204 numbers JAAOIC000000000 (Muller, Chebbi, et al., 2021) and JADWQK000000000 (Muller,
205 Ogereau et al., 2021). The genomes of *N. typhae, G. sparganii,* and *C. phragmitell*a were *de*
206 *novo* assembled for this study with the following procedure: (i) raw Illumina reads produced
207 from parasitized larvae were mapped against the *C. typhae* genome with Bowtie2 in order to
208 remove all reads from *C. typhae*, (ii) the filtered reads were assembled with MaSuRCA v4.0.1
209 with all parameters set to default, except for jellyfish hash size (JF_SIZE = 20000000000), and
210 we activated USE_LINKING_MATES since we did not use any long-reads, (iii) any trace of *C.*
211 *typhae* genome remaining in the assembly were identified by similarity search (blastn)
212 between the newly assembled genomes and *C. typhae* genome. We filtered out from the
213 assemblies all contigs with more than 99% identity with a *C. typhae* contig on more than 90%
214 of its length. This last step led to remove four contigs from *N. typhae,* none from *G. sparganii*,
215 and 24 from *C. phragmitella*. We also ran purge_dups on *C. phragmitella*, whose assembly

216 initially harbored 9.5% of duplicated BUSCO (Guan et al., 2020). We assessed the quality of
217 these three new genomes with Busco v5.0.0 using 5286 lepidopterean core genes.

218 **Sequencing data processing**

219 To search for the presence of BV sequences integrated into *G. sparganii*, *N. typhae*, *C.*
220 *phragmitella* and *S. nonagrioides* genomes, we developed the workflow WorkflowBowBlast
221 (github/HeloiseMuller/WorkflowBowBlast). This workflow proceeds all the data the same way
222 as described in Muller, Chebbi, et al. (2021), in which we characterized and quantified the
223 genome-wide patterns of *C. typhae* BV (CtBV) integration in various tissues of parasitized *S.*
224 *nonagrioides*. We activated all the tools of the workflow (integrity with fastqc v.0.11.8,
225 trimmomatic v.0.38, bowtie2 v2-2.3.4.2, coverage with bedtools v.2.26.0 and blastn v.2.6.0)
226 with all default parameters. We used the appropriate lepidopteran genomes (*S. nonagrioides*,
227 *N. typhae*, *G. sparganii,* or *C. phragmitell*a) for the first reference genome and *C. typhae* for
228 the second optional genome. For the analyses, we used two outputs of WorkflowBowBlast.
229 First, we used the bedtools genomecov output of WorkflowBowBlast that indicates the
230 sequencing depth at each position of the *C. typhae* genome. We then used CovWindows
231 (github/HeloiseMuller/CovWindows) to get sequencing depth on each CtBV DNA
232 (corresponding to any of the three CtBV forms) from this file. WorkflowBowBlast also
233 automatically gives the average sequencing depth on both genomes, and the percentage of
234 the genomes covered at least once by Illumina reads. Second, we used the two blastn outputs
235 of WorkflowBowBlast, one on each genome (Lepidoptera and wasp), to identify chimeric reads
236 involving *C. typhae* proviral segments and the lepidopteran DNA using an R pipeline (Peccoud
237 et al., 2018). This pipeline, specifically adapted to find chromosomally integrated BV in Muller,
238 Chebbi, et al. (2021), identifies a read as chimeric if (i) at least 16 bases align only on the first

reference genome, and a minimum of 16 other bases align only on the second reference genome; (ii) less than 10% of the read length maps to neither reference genome; (iii) no more than 20 bases align simultaneously on both reference genomes; and (iv) no more than 5 bases are inserted between the two genomes at the integration point. The two latter filters imply that aligned read regions are allowed to overlap by up to 20 bp or to be separated by at most 5 bp. After this pipeline, we filtered out chimeric reads resulting from PCR duplicates, except when comparing the number of chimeric reads to the sequencing depth, since PCR duplicates are not filtered out in the latest.

In addition to our ten samples (five libraries of 100 pooled *S. nonagrioides* offspring, two adult *S. nonagrioides* females, one *G. sparganii* larva, one *N. typhae* larva, and one *C. phragmitella* larva), we also processed in the exact same way two samples that we generated in earlier studies and used here for comparison: Caterpillar and Control. Both samples were sequenced in the same conditions as our ten samples. Caterpillar corresponds to a *S. nonagrioides* caterpillar parasitized by *C. typhae* and sequenced seven days post parasitism, for which no wasp larvae were present (sample called “whole body” in the original publication (Muller, Chebbi, et al., 2021)). This sample was included to compare the number of CtBV chromosomal integrations in the natural host to those in the non-natural hosts. Control corresponds to the Illumina trimmed reads used to assemble the genome of non-parasitized *S. nonagrioides* (Muller, Ogereau, et al., 2021). We included this last sample in order to estimate the proportion of reads that would map on *C. typhae* genome when *S. nonagrioides* did not undergo any contact with C. *typhae*. In addition, this control also allowed us to determine that our pipeline returned no false positive *S. nonagrioides*-*C. typhae* chimeras. All analyses and figures were done with R v4.0.5.

Estimations and normalization of integration events

Integration events of CtBV DNA into host genomes were counted for each sample, as each position in the host genome supported by at least one read (*i.e.*, CtBV-host junctions supported by more than one read were counted as one integration event). We estimated the average number of integration events per haploid genome with IPMH*G/k, where IMPH is the normalized number of Integration events Per Million of Host reads, G is the genome size of the host in megabases (1,021Mb for *S. nonagrioides*), and k is the read length (150bp).

When assessing the transmission of CtBV from resistant parasitized moths to their offspring, we did not sequence 100% of the 1000 gametes we sampled (we sampled 500 offspring, and each inherited 2 gametes from their resistant parents). Thus, we had to estimate the number of gametes we sequenced (noted S). To do this, we calculated an upper limit, which is the maximum number of gametes we might have sequenced, with: $S = \sum_{k=1}^{n}(dk * pk)$, where d is the sequencing depth in library k (n=5 libraries), and p is the proportion of the genome sequenced on average in one offspring for library k. To estimate p, we randomly sampled one hundredth reads of the library (because each library contains 100 offspring), and we calculated the proportion of the host genome covered when looking only at this subset of reads. This proportion is an estimation of the coverage of one of the 100 offspring of the library. We repeated this sampling ten times and we calculated the average, *i.e.* the proportion of genome sequenced for one individual on average in library k (p_k).

Results

Wasp cells do not persist in adult *S. nonagrioides* surviving to parasitism by *C. typhae*

In a previous study, we showed that seven days after a female *C. typhae* wasp oviposited into larvae of its host *S. nonagrioides*, CtBV DNA persisted under three forms in parasitized caterpillars: the circular form, the chromosomally integrated form, and the proviral segments located in the wasp genome of wasp teratocyte cells present in the caterpillars (Muller, Chebbi, et al., 2021). Teratocytes are cells released from the wasp embryonic membranes into the host when eggs hatch. They are involved in various functions, including suppressing the host immune response (Buron & Beckage, 1997). Since insects undergo important cellular changes during metamorphosis (Tettamanti & Casartelli, 2019), we wondered whether some CtBV DNA and/or other wasp DNA can persist at the adult stage in host individuals surviving to parasitism, in one form or another. Addressing this question is important to assess the extent to which transmission of germline-integrated CtBV and/or free CtBV to host offspring can occur. For this end, we Illumina-sequenced the whole genome of two adult *S. nonagrioides* out of the 37 that resisted to parasitism in our experiment. Both sequenced adult *S. nonagrioides* were female, but while Adult-1 produced many offspring, Adult-6 did not produce any.

To assess the extent to which wasp cells may persist in adult hosts resisting to parasitism, we looked whether the wasp genome was covered by the Illumina reads. For this, we mapped trimmed Illumina reads obtained from the two *S. nonagrioides* females on the whole genomes of both *C. typhae* and *S. nonagrioides*. Average sequencing depths over the *S. nonagrioides* genome were in the targeted range (41X and 49X) and 97% of the moth genome was sequenced at least once for the two individuals (Fig. 2). By contrast, average sequencing

depths over the *C. typhae* genome were very low (0.5X and 0.86X) and reads mapped to only a small fraction of the wasp genome (<1% and 7.4%) (Fig. 2). To assess whether these sequencing depth and coverage, although low, indicate the presence of residual amounts of wasp cells, we compared these statistics to a control dataset corresponding to trimmed Illumina reads generated in an earlier study aiming at sequencing the genome of non-parasitized *S. nonagrioides* individuals that were never in contact with *C. typhae (Muller, Ogereau, et al., 2021).* For this dataset, we obtained an average sequencing depth of 0.43X over the *C. typhae* genome with less than 1% of the wasp genome being covered at least once. The average sequencing depth on the genome of *S. nonagrioides* was 38.9X. This control indicates that in a *S. nonagrioides* individual devoid of any *C. typhae* DNA, we can expect an average sequencing depth on the latter at about 1.1% of the former, and less than 1% of the wasp genome is expected to be covered at least once. For Adult-1, the average depth over the *C. typhae* genome was 1.23% of the one on *S. nonagrioides* and less than 1% of the wasp genome was covered, which is close to what we obtained for the control dataset. This suggests that teratocytes and other wasp cells were absent in this individual. However, Adult-6 clearly had higher amounts of wasp DNA than expected in the absence of *C. typhae* DNA, with an average depth on the *C. typhae* genome at 1.77% of the one computed on the *S. nonagrioides* genome, and 7.4% of the wasp genome covered at least once. Although these figures could be interpreted as being due to the presence of small amounts of wasp cells in this *S. nonagrioides* adult individual, we found that the wasp genome regions covered by sequencing reads were very small, although *C. typhae* scaffolds are large (83% of these regions were between 100 and 300 bp). This suggests that the wasp DNA sequenced from this adult individual is highly fragmented and unlikely to originate from complete whole wasp genomes. Thus, while the

origin of this wasp DNA remains to be determined, our results tend to indicate that teratocytes and other wasp cells are unlikely to persist in host adult individuals surviving to parasitism.

Persistence of CtBV DNA in adults *S. nonagrioides* surviving to parasitism by *C. typhae*

We next focused on evaluating whether CtBV DNA may persist in adult *S. nonagrioides* surviving to parasitism, in one form or another. We found that the average sequencing depth over the 27 CtBV circles was higher than the average depth on the entire wasp genome in the two *S. nonagrioides* adult individuals (3.7X *versus* 0.5X in Adult-1 and 12.46X *versus* 0.86X in Adult-6). This suggests that contrary to wasp cells, CtBV persisted in adult hosts, in one form or another. To assess whether some of the CtBV circles persisted in their chromosomally integrated form, we looked for chimeric reads between CtBV and *S. nonagrioides.* We found 43 and 303 chimeric reads corresponding to HIM-mediated integration of the 16 HIM-containing CtBV segments in Adult-1 and Adult-6, respectively. Thus, the 16 HIM-containing CtBV segments were able to persist at least in their integrated forms at the adult stage in both adult hosts (Fig. 3). More precisely, if we consider only independent integration events (IEs), *i.e.* counting only once chimeric reads with the exact same integration coordinate, we found 41 HIM-mediated IEs in Adult-1 and 297 HIM-mediated IEs in Adult-6. In order to compare samples, we normalized the number of IEs with the number of host reads (Integration events Per Million reads of Host, or IPMH). Adult-1 had 0.29 IPMH and Adult-6 had 1.77 IPMH (Fig. 3). From this normalized number of IEs, we can also express the average number of IE per haploid genome (see materials and methods), which allowed us to estimate 1.97 and 12.03 IEs per haploid genome for Adult-1 and Adult-6, respectively.

We then compared the normalized number of IEs of both adult *S. nonagrioides* to the one found in the whole body of a parasitized caterpillar in which no wasp embryos were observed seven days after oviposition (called “Caterpillar”) (Muller, Chebbi, et al., 2021). We observed that the number of integrated CtBV circles can be quite high in adults, since Adult-6 is in the same order of magnitude as in a caterpillar (Fig. 3b). This raises the question of whether some CtBV circles might be integrated into the germline genome of *S. nonagrioides* and transmitted to the next generation of moth.

Before addressing this question, we investigated whether CtBV could also persist under their circular forms. Based on the current knowledge of the system, factors involved in bracovirus replication are not injected in the host (Louis et al., 2013), so bracoviruses are only known to replicate in the calyx of the wasp, not in the caterpillar. This is why, one can expect bracovirus circles to be degraded after some times in the caterpillar. However, Muller, Chebbi, et al. (2021) found that these circles still persist in high amounts in *S. nonagrioides* caterpillars seven days post-parasitism, and Chevignon et al. (2018) found a persistence of bracovirus circles in *Manducta sexta* 12 days post-parasitism. To determine whether CtBV can persist under circular form in adult moths, we measured the sequencing depth over circles devoid of HIM, for which no or very few integrations have been found in the present study (zero and one IE in total in Adult-1 and Adult-6, respectively). While in Adult-1 all CtBV circles devoid of HIM had a sequencing depth close to the one of the *C. typhae* genome (0.5X), in Adult-6 at least two CtBV devoid of HIM were sequenced at a depth markedly higher than the *C. typhae* genome (0.86X): CtBV S20/33 (4.8X) and S23 (5.9X) (see the two red diamonds in Fig. 4b). Although limited, the quantity of persisting S20/33 and S23 is as high as that of three HIM-containing CtBV (S10, S11,

and S14), indicating that at least some CtBV circles can persist up to the adult stage of the host and during its entire lifespan (Adult-6 died 46 days post-oviposition).

No vertical transmission of CtBV circles by *S. nonagrioides* adults surviving to parasitism by *C. typhae*.

To assess the extent to which CtBV DNA persisting in adults *S. nonagrioides* can be transmitted vertically to the next generation of moth, we screened for the presence of CtBV in F1 individuals resulting from crosses of *S. nonagrioides* individuals that had survived parasitism. In total we Illumina-sequenced 500 offspring individuals produced by five couples. Each couple was made of a male and a female *S. nonagrioides* that both survived parasitism. The DNA of the 500 individuals was distributed in five libraries each containing 100 individuals (Suppl. Fig. 1).

The obtained sequencing depth on the *S. nonagrioides* genome was between 105 and 143X depending on the library, and 98.7% to 99.5% of the moth genome was covered at least once in the five libraries (Fig. 2). As expected, the sequencing depth and coverage of the *C. typhae* genome were very low for the five libraries (1.2% to 1.3% of the sequencing coverage on the host, and <1% of the wasp genome covered at least once). These figures are similar to those obtained for the control dataset of reads produced from *S. nonagrioides* larvae that never encounter *C. typhae* (1.1% of the sequencing coverage on the host, and <1% of the wasp genome covered). Thus, as expected, no wasp whole genome was transmitted to moth F1 offspring individuals.

Importantly, the average sequencing depth on the 27 CtBV circles was close to zero (between 0 and 0.01X) for the five libraries of offspring DNA, suggesting that no CtBV DNA, in any forms,

were sequenced (Fig. 2A). In agreement with this result, we could not identify any chimeric read between CtBV and *S. nonagrioides.* Thus, we could not identify any vertical transmission of CtBV, neither in their circular nor integrated forms, to any of the 500 offspring individuals we sequenced. This result provides a way to estimate an upper limit of the frequency at which vertical transmission of bracoviruses may be transmitted in host populations. We estimated that among the 1000 gametes transmitted to these offspring, we did not sequence more than 374 gametes (see methods). This gives a probability to detect the vertical transmission of a CtBV lower than 1/374 gametes. This probability might have been different, if we had chosen a different sampling strategy. Here we chose to sequence 100 offspring individuals from five couples but we cannot tell whether sequencing more offspring from less couples, or sequencing less offspring from more couples could have increased our chances to find a transmission.

Chromosomal integration of CtBV in non-target species

In the context of biocontrol, it is worth to evaluate the possible consequences, in terms of BV chromosomal integration, of the introduction of an exotic biocontrol agent on autochthonous lepidopteran species that are not present in the native area of the wasp. Since such species have never interacted with the introduced wasp, they may be less permissive hosts and thus more likely to survive parasitism (Mahmoud et al., 2012). In turn, chromosomal integration of BV DNA in such non- or less- permissive species may increase the likelihood that germline-integrated BV DNA are transmitted through host generations (Drezen et al., 2017). Yet, the capacity of BV circles to undergo chromosomal integration in species not known to be naturally targeted by an endoparasitoid wasp has never been studied.

413 Here, we investigated whether CtBV can undergo chromosomal integration in three stemborer
414 lepidopteran species (*Nonagria typha*e [Noctuidae], *Globia sparganii* [Noctuidae], and *Chilo*
415 *phragmitella* [Crambidae]), the geographical distribution of which overlaps with that of *S.*
416 *nonagrioides* in France (Galichet et al., 1992; Teder & Tammaru, 2002; Tewksbury et al., 2002).
417 Among the two *N. typhae*, the five *G. sparganii*, and the eight *C. phragmitella* individuals that
418 were parasitized by *C. typhae* for this experiment, only four *C. phragmitella* yielded wasp
419 cocoons, yet all hosts died before reaching the adult stage. For each species, we Illumina-
420 sequenced one individual in which no wasp cocoon developed. To study CtBV infection in these
421 three caterpillars, we first assembled their reference genomes, which are not available in
422 public databases. Before doing so, we filtered out all reads mapping to the genome of *C. typhae*
423 (see materials & methods). The quality of the assemblies we obtained was heterogenous, with
424 a higher completeness for *N. typhae* (647Mb) and *G. sparganii* (972Mb) (93.6% and 82.1% of
425 complete BUSCO, respectively) than for *C. phragmitella* (942Mb; 67.4% complete BUSCO)
426 (Table 1). These assemblies all turned out to be sufficient to test for chromosomal integration.

427 Looking at sequencing depth and coverage over the whole *C. typhae* genome in the three
428 samples revealed different patterns. While virtually no wasp whole genome was sequenced in
429 *G. sparganii* (0.5X), it was deeply sequenced in *N. typhae* and *C. phragmitella* (88X and 166X,
430 respectively), and almost entirely covered at least once (99.3% in the two last samples) (Fig.
431 2). These figures indicate the absence of wasp cells in *G. sparganii* and their presence in *N.*
432 *typhae* and *C. phragmitella*. This further suggests that although no cocoons emerged from *N.*
433 *typhae* and *C. phragmitella* larvae before they died, wasp embryos were in fact able to start
434 their development in these two species.

435 Interestingly, we found that as observed when sequencing a parasitized *S. nonagrioides* larvae,
436 the total sequencing depth over CtBV DNA (in any form) was higher than that over the whole
437 wasp genome for the three species not normally targeted by *C. typhae* (Fig. 2). Subtracting the
438 whole wasp genome from this total CtBV DNA depth (*i.e.*, subtracting the proviral segment
439 form) provides an estimate of the depth at which circular plus integrated forms of CtBV were
440 sequenced: 11.8X, 20.5X and 107.6X in *G. sparganii*, *N. typhae* and *C. phragmitella*,
441 respectively (Fig. 2a). Moreover, we identified HIM-mediated integration events (IEs) of CtBV
442 in the three samples, with 120, 592 and 1114 IEs in *G. sparganii*, *N. typhae* and *C. phragmitella*,
443 respectively (Fig. 3). The normalized number of integration events was the highest for *C.*
444 *phragmitella,* in which it is even higher than in the natural host *S. nonagrioides* (Fig. 3). It is
445 noteworthy that sequencing depths over non-integrated CtBV circles (circles for which we did
446 not identify any chimeric read) are in the same range as those over integrated CtBV circles for
447 the three non-natural host species (Fig. 4c-e). This indicates that circles can persist in non-
448 natural host species at similar levels in their non-integrated forms and in their integrated
449 forms, even up to 25 days post-oviposition (for *C. phragmitella*).

450 We finally compared the molecular signatures associated with CtBV chromosomal integration
451 in the three non-target species to those characterized in the natural host *S. nonagrioides*. We
452 have shown that almost all integrations in parasitized *S. nonagrioides* involve HIM-containing
453 circles (99%) and are HIM-mediated (96.5%) (Muller, Chebbi, et al., 2021). In non-target
454 species, we found similar signatures in *N. typhae* and *G. sparganii*, with 98.9% and 96.8% of
455 the integration events involving HIM-containing circles, and 92.8% and 97.6% of these
456 integrations occurring within the HIM, respectively. The pattern was markedly different in *C.*
457 *phragmitella*, with only 91.6% of the integration events involving HIM-containing circles, and

only 74.3% of them occurring within the HIM. This high number of integrations outside the HIM motif in *C. phragmitella* is not due to just one or few circles integrating differently, but to a common tendency of most circles: 14 circles out of the 16 HIM-containing circles have significantly more integration events outside HIM in *C. phragmitella* than in *S. nonagrioides* (Fisher test: p.value<0.05; see Fig. 5). Such differences may be explained by the fact that host factors are involved in BV integration and that these factors are expected to diverge substantially in *C. phragmitella*, the non-target species the most distantly related to *S. nonagrioides*. This would be consistent with a recent study showing that host integrases are involved in BV integration (Wang et al., 2021).

Discussion

In this study, we have characterized the presence of chromosomally-integrated and non-integrated forms of BV circles in *S. nonagrioides* adult individuals surviving parasitism by the endoparasitoid wasp *C. typhae*. We found that the two forms of CtBV can persist to the host adult stage in numbers that are similar to those identified in larval stages (Muller, Chebbi, et al., 2021), despite the tissue remodeling and other physiological changes undergone during metamorphosis by *S. nonagrioides* (Tettamanti & Casartelli, 2019). Interestingly, we found that circular forms of HIM-containing CtBV persist in higher amounts at adult stage than CtBV circles devoid of HIM (sequenced at 12.2X and 1.2X on average respectively, see Fig. 4). In our earlier study (Muller, Chebbi, et al., 2021), we did not observe differences of persistence levels between HIM-containing CtBV circles and those devoid of HIM at seven days post oviposition (their sequencing depths were similar). The present results indicate that most DNA circles are lost/degraded, between seven days post-parasitism and the adult stage. A study monitoring

DNA circles persistence in a permissive host from seven days post parasitism to the end of the wasp larvae development would shed additional light on segment copy number and persistence during development of wasp offspring, and therefore to some extent on the role of bracovirus circle integration. It is also striking that Adult-6, the adult with the highest number of CtBV integrations, did not produce any offspring, nor even any eggs, whereas Adult-1, which suffered less CtBV integration events, is the mother of hundreds of caterpillars (some of them being included in our analysis). We cannot conclude here on a possible link between the number of CtBV integrations and host fertility because of a lack of replicates. In the same way, it is possible that the reason why the majority of resisting adult couples (10 out of 15) did not produce any offspring is because of the effect of CtBV integration. Yet, these adults were not crossed as pools of many male and female individuals as routinely implemented in our lab, but as couples (*i.e.*, each made of one male and one female individual), which we previously noticed tends to decrease the likelihood of successful breeding, even when using non-parasitized individuals. It is thus difficult to assess whether parasitism or breeding conditions or both had an effect on breeding success. This question certainly deserves to be addressed in the future as it would help to better understand the possible risks of transmission of integrated CtBV circles to the next host generation, and thus of horizontal transfer with possible long-term consequences on the host.

We tested for the possibility that CtBV DNA persisting in adult *S. nonagrioides* may be transmitted to F1 individuals resulting from crossings between host individuals surviving to parasitism. We were unable to identify any CtBV DNA in 500 of these F1. This indicates that the probability of transmission of CtBV DNA from one host generation to the next is lower than our detection limit. Our previous study showed that CtBV integrated in all *S. nonagrioides*

503 tissues studied, which did not include gonads (Muller, Chebbi, et al., 2021), and we here report
504 an average of 1.97 IEs per haploid genome in Adult-1 (the adult who produced offspring).
505 Based on these findings, one could expect most gametes to harbor at least one IE.
506 Nonetheless, the lack of detection of transmission we obtain in the present study means that
507 at least a significant number of gametes do not harbor any IE. Although we found 1.97 IEs per
508 haploid genome on average, we actually have no way to estimate the variance, that would
509 indicate whether IEs are equitably shared by all cells or whether some cells carry most IEs.
510 Regardless of the variance, gametes might be less subject to integration than other tissues
511 because they may be less accessible. In agreement with this, Wang et al. (2021) measured
512 lower amounts of circular CvBV in testis than in hemocytes of *Plutella xylostella*. Another
513 possibility is that CtBV may integrate in the germline at rates comparable to other tissues but
514 it may impede proper gametogenesis and/or negatively impact the functionality of gametes.
515 Sequencing gonads of parasitized hosts would provide important information on this aspect.

516 Although we can only estimate an upper limit probability of transmission for resistant
517 individuals, the probability of transmission at the population scale (including all parents, non-
518 surviving, and sterile individuals) must be much lower. Indeed, out of the 62 parasitized
519 caterpillars included in this study, 25 died before reaching adult stage, and 27 did not produce
520 any offspring. Thus, the probability of transmission was equal to zero for 52 out of 62
521 parasitized *S. nonagrioides* in our experiment. It is well known that such transmission can occur
522 in some species, as several cases of bracovirus-derived sequences resulting from ancient HT
523 from wasp to lepidopterans have been characterized (Di Lelio et al., 2019; Gasmi et al., 2015;
524 Schneider & Thomas, 2014). Yet, it is interesting to note that in spite of presumably being
525 heavily targeted by *C. typhae* and other endoparasitoid wasps in the wild since many

526 generations, *S. nonagrioides* does not contain any trace of bracovirus sequences in its
527 (germline) genome (Muller, Chebbi, et al., 2021). Thus, the real probability of BV transmission
528 in host populations is likely to be much lower than the upper limit we were able to estimate
529 here. Although our upper limit estimate will be interesting for future studies aiming at
530 investigating this rate, it is better to consider it as not so informative in the context of
531 biocontrol strategies. Indeed, at the scale of biocontrol, even a low probability could lead to
532 numerous BV transmissions since a large number of *S. nonagrioides* caterpillars (and
533 potentially other lepidopteran species not known to be targeted by *C. typhae* in the wild)
534 would be parasitized every year. Future experiments should aim at refining estimates of the
535 frequency of BV transmission to better inform policy makers in the context of wasp-based
536 biocontrol strategies, as well as to better understand what impact wasp-to-lepidopteran HT
537 may have had on the evolution of lepidopteran genomes.

538 Our study also shows that CtBV circles can persist as extrachromosomal molecules and
539 undergo chromosomal integration in parasitized larvae of three lepidopteran species (*N.*
540 *typhae*, *C. phragmitella*, *G. sparganii*) that are not naturally targeted by *C. typhae*, at rates
541 comparable to those measured in *S. nonagrioides* (the native host). This shows that CtBV cell
542 entry and chromosomal integration can occur in a wide range of lepidopteran species,
543 including a Crambidae which diverged at least 100 million years ago from *S. nonagrioides*.
544 Thus, these two processes are unlikely to play a major role in shaping the host range of
545 Braconid wasps. This is in line with a study that showed that bracovirus gene expression levels,
546 rather than cell entry and persistence of BV DNA, are more likely to be a key factor underlying
547 host permissiveness (Bitra et al., 2016).

The extent to which the three non-natural host species included here are permissive to *C. typhae* is under study in our laboratory. It depends on ecological, behavioral and chemical factors that condition their attractiveness and acceptance as host. Yet, the outcome of the present study – only 4 host larvae out of 15 yielded wasp cocoons – together with other unpublished preliminary results suggest that these three species are much less permissive than *S. nonagrioides*. This suggests that the capacity of BV to enter in host cells and undergo chromosomal integration may not be used as a good proxy for evaluating wasp permissive host range. In the context of using *C. typhae* as a biocontrol agent, our results are important because they show that autochthonous lepidopteran species occurring in the area in which the wasp could be released can suffer BV chromosomal integration and thus be affected in one way or another. This is, however, considering that interaction between *C. typhae* and these other autochthonous lepidopteran species in the wild are timely and ecologically possible, which remains to be assessed. On a more basic level, the fact that BV DNA can undergo chromosomal integration in a wide range of non-natural hosts, likely to be less permissive than native hosts, supports the idea that parasitizing non-natural hosts may increase the propensity of wasp-to-host HT of BV DNA with long term consequences (Drezen et al., 2017).

Acknowledgments

Research performed in this study was funded by French National Research Agency (ANR) (grants CoteBio ANR17-CE32-0015-02 to L.K. and TranspHorizon ANR-18-CE02-0021-01 to C.G.). We thank David Ogereau and Remi Jeanette for their help with DNA extraction and insect rearing, respectively.

References

Aubin, E., El Baidouri, M., & Panaud, O. (2021). Horizontal Gene Transfers in Plants. *Life*, *11*(8), 857. https://doi.org/10.3390/life11080857

Baidouri, M. E., Carpentier, M.-C., Cooke, R., Gao, D., Lasserre, E., Llauro, C., Mirouze, M., Picault, N., Jackson, S. A., & Panaud, O. (2014). Widespread and frequent horizontal transfers of transposable elements in plants. *Genome Research*, *24*(5), 831–838. https://doi.org/10.1101/gr.164400.113

Beck, M. H., Zhang, S., Bitra, K., Burke, G. R., & Strand, M. R. (2011). The Encapsidated Genome of Microplitis demolitor Bracovirus Integrates into the Host Pseudoplusia includens. *Journal of Virology*, *85*(22), 11685–11696. https://doi.org/10.1128/JVI.05726-11

Beckage, N., & Drezen, J.-M. (Eds.). (2011). *Parasitoid Viruses: Symbionts and Pathogens*. Academic Press. https://doi.org/10.1016/C2009-0-64055-1

Benoist, R., Chantre, C., Capdevielle-Dulac, C., Bodet, M., Mougel, F., Calatayud, P. A., Dupas, S., Huguet, E., Jeannette, R., Obonyo, J., Odorico, C., Silvain, J. F., Le Ru, B., & Kaiser, L. (2017). Relationship between oviposition, virulence gene expression and parasitism success in Cotesia typhae nov. Sp. Parasitoid strains. *Genetica*, *145*(6), 469–479. https://doi.org/10.1007/s10709-017-9987-5

Benoist, R., Paquet, S., Decourcelle, F., Guez, J., Jeannette, R., Calatayud, P.-A., Le Ru, B., Mougel, F., & Kaiser, L. (2020). Role of egg-laying behavior, virulence and local adaptation in a parasitoid's chances of reproducing in a new host. *Journal of Insect Physiology*, *120*, 103987. https://doi.org/10.1016/j.jinsphys.2019.103987

Bézier, A., Louis, F., Jancek, S., Periquet, G., Thézé, J., Gyapay, G., Musset, K., Lesobre, J., Lenoble, P., Dupuy, C., Gundersen-Rindal, D., Herniou, E. A., & Drezen, J.-M. (2013). Functional endogenous viral elements in the genome of the parasitoid wasp Cotesia congregata: Insights into the evolutionary dynamics of bracoviruses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1626). https://doi.org/10.1098/rstb.2013.0047

Bitra, K., Burke, G. R., & Strand, M. R. (2016). Permissiveness of lepidopteran hosts is linked to differential expression of bracovirus genes. *Virology*, *492*, 259–272. https://doi.org/10.1016/j.virol.2016.02.023

Bosque-Perez, & Schulthess. (1998). *African cereal stem borers: Economic importance, taxonomy, natural enemies and control.* CAB INTERNATIONAL. https://www.cabi.org/cpc/abstract/19981108334

Boto, L. (2014). Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1777), 20132450. https://doi.org/10.1098/rspb.2013.2450

Buron, I. de, & Beckage, N. E. (1997). Developmental changes in teratocytes of the braconid wasp Cotesia congregata in larvae of the tobacco hornworm, Manduca sexta. *Journal of Insect Physiology*, *43*(10), 915–930. https://doi.org/10.1016/S0022-1910(97)00056-5

Chevignon, G., Periquet, G., Gyapay, G., Vega-Czarny, N., Musset, K., Drezen, J.-M., & Huguet, E. (2018). Cotesia congregata Bracovirus Circles Encoding PTP and Ankyrin Genes Integrate into the DNA of Parasitized Manduca sexta Hemocytes. *Journal of Virology*, *92*(15). https://doi.org/10.1128/JVI.00438-18

Chevignon, G., Thézé, J., Cambier, S., Poulain, J., Da Silva, C., Bézier, A., Musset, K., Moreau, S. J. M., Drezen, J.-M., & Huguet, E. (2014). Functional Annotation of Cotesia congregata Bracovirus: Identification of Viral Genes Expressed in Parasitized Host Immune Tissues. *Journal of Virology*, *88*(16), 8795–8812. https://doi.org/10.1128/JVI.00209-14

Danchin, E. G. J., Guzeeva, E. A., Mantelin, S., Berepiki, A., & Jones, J. T. (2016). Horizontal Gene Transfer from Bacteria Has Enabled the Plant-Parasitic Nematode Globodera pallida to Feed on Host-Derived Sucrose. *Molecular Biology and Evolution*, *33*(6), 1571–1579. https://doi.org/10.1093/molbev/msw041

Desjardins, C. A., Gundersen-Rindal, D. E., Hostetler, J. B., Tallon, L. J., Fadrosh, D. W., Fuester, R. W., Pedroni, M. J., Haas, B. J., Schatz, M. C., Jones, K. M., Crabtree, J., Forberger, H., & Nene, V. (2008). Comparative genomics of mutualistic viruses of Glyptapanteles parasitic wasps. *Genome Biology*, *9*(12), R183. https://doi.org/10.1186/gb-2008-9-12-r183

Di Lelio, I., Illiano, A., Astarita, F., Gianfranceschi, L., Horner, D., Varricchio, P., Amoresano, A., Pucci, P., Pennacchio, F., & Caccia, S. (2019). Evolution of an insect immune barrier through horizontal gene transfer mediated by a parasitic wasp. *PLOS Genetics*, *15*(3), e1007998. https://doi.org/10.1371/journal.pgen.1007998

Drezen, J.-M., Chevignon, G., Louis, F., & Huguet, E. (2014). Origin and evolution of symbiotic viruses associated with parasitoid wasps. *Current Opinion in Insect Science*, *6*, 35–43. https://doi.org/10.1016/j.cois.2014.09.008

Drezen, J.-M., Josse, T., Bézier, A., Gauthier, J., Huguet, E., & Herniou, E. A. (2017). Impact of Lateral Transfers on the Genomes of Lepidoptera. *Genes*, *8*(11), 315. https://doi.org/10.3390/genes8110315

Galichet, P. F., Cousin, M., & Girard, R. (1992). Egg dormancy and synchronization of larval feeding with host plant development in three noctuid (Lepidoptera) species. *Acta Oecologica*, *13*(6), 701.

Gasmi, L., Boulain, H., Gauthier, J., Hua-Van, A., Musset, K., Jakubowska, A. K., Aury, J.-M., Volkoff, A.-N., Huguet, E., Herrero, S., & Drezen, J.-M. (2015). Recurrent Domestication by Lepidoptera of Genes from Their Parasites Mediated by Bracoviruses. *PLoS Genetics*, *11*(9). https://doi.org/10.1371/journal.pgen.1005470

Gauthier, J., Boulain, H., van Vugt, J. J. F. A., Baudry, L., Persyn, E., Aury, J.-M., Noel, B., Bretaudeau, A., Legeai, F., Warris, S., Chebbi, M. A., Dubreuil, G., Duvic, B., Kremer, N., Gayral, P., Musset, K., Josse, T., Bigot, D., Bressac, C., … Drezen, J.-M. (2021). Chromosomal scale assembly of parasitic wasp genome reveals symbiotic virus

colonization. *Communications Biology*, *4*(1), 1–15. https://doi.org/10.1038/s42003-020-01623-8

Gauthier, J., Drezen, J.-M., & Herniou, E. A. (2018). The recurrent domestication of viruses: Major evolutionary transitions in parasitic wasps. *Parasitology*, *145*(6), 713–723. https://doi.org/10.1017/S0031182017000725

Gilbert, C., & Cordaux, R. (2017). Viruses as vectors of horizontal transfer of genetic material in eukaryotes. *Current Opinion in Virology*, *25*, 16–22. https://doi.org/10.1016/j.coviro.2017.06.005

Gilbert, C., & Feschotte, C. (2018). Horizontal acquisition of transposable elements and viral sequences: Patterns and consequences. *Current Opinion in Genetics & Development*, *49*, 15–24. https://doi.org/10.1016/j.gde.2018.02.007

Gitau, C. W., Gundersen-Rindal, D., Pedroni, M., Mbugi, P. J., & Dupas, S. (2007). Differential expression of the CrV1 haemocyte inactivation-associated polydnavirus gene in the African maize stem borer Busseola fusca (Fuller) parasitized by two biotypes of the endoparasitoid Cotesia sesamiae (Cameron). *Journal of Insect Physiology*, *53*(7), 676–684. https://doi.org/10.1016/j.jinsphys.2007.04.008

Gruber, A., Stettler, P., Heiniger, P., Schümperli, D., & Lanzrein, B. (1996). Polydnavirus DNA of the braconid wasp Chelonus inanitus is integrated in the wasp's genome and excised only in later pupal and adult stages of the female. *The Journal of General Virology*, *77 (Pt 11)*, 2873–2879. https://doi.org/10.1099/0022-1317-77-11-2873

Guan, D., McCarthy, S. A., Wood, J., Howe, K., Wang, Y., & Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*, *36*(9), 2896–2898. https://doi.org/10.1093/bioinformatics/btaa025

Harwood, S. H., McElfresh, J. S., Nguyen, A., Conlan, C. A., & Beckage, N. E. (1998). Production of Early Expressed Parasitism-Specific Proteins in Alternate Sphingid Hosts of the Braconid WaspCotesia congregata. *Journal of Invertebrate Pathology*, *71*(3), 271–279. https://doi.org/10.1006/jipa.1997.4745

Herniou, E. A., Huguet, E., Thézé, J., Bézier, A., Periquet, G., & Drezen, J.-M. (2013). When parasitic wasps hijacked viruses: Genomic and functional evolution of polydnaviruses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1626), 20130051. https://doi.org/10.1098/rstb.2013.0051

Heringer P, Kuhn GCS. (2022). Multiple horizontal transfers of a Helitron transposon associated with a parasitoid wasp. *Mob DNA*. 13(1):20. doi: 10.1186/s13100-022-00278-y.

Kaiser, L., Fernandez-Triana, J., Capdevielle-Dulac, C., Chantre, C., Bodet, M., Kaoula, F., Benoist, R., Calatayud, P.-A., Dupas, S., Herniou, E. A., Jeannette, R., Obonyo, J., Silvain, J.-F., & Ru, B. L. (2017). Systematics and biology of Cotesia typhae sp. N. (Hymenoptera, Braconidae, Microgastrinae), a potential biological control agent against the noctuid Mediterranean corn borer, Sesamia nonagrioides. *ZooKeys*, *682*, 105–136. https://doi.org/10.3897/zookeys.682.13016

Kaiser, L., Ru, B. P. L., Kaoula, F., Paillusson, C., Capdevielle-Dulac, C., Obonyo, J. O., Herniou, E. A., Jancek, S., Branca, A., Calatayud, P.-A., Silvain, J.-F., & Dupas, S. (2015). Ongoing ecological speciation in Cotesia sesamiae, a biological control agent of cereal stem borers. *Evolutionary Applications*, *8*(8), 807–820. https://doi.org/10.1111/eva.12260

Kankonda, O. M., Akaibe, B. D., Sylvain, N. M., & Ru, B.-P. L. (2018). Response of maize stemborers and associated parasitoids to the spread of grasses in the rainforest zone of Kisangani, DR Congo: Effect on stemborers biological control. *Agricultural and Forest Entomology*, *20*(2), 150–161. https://doi.org/10.1111/afe.12238

Keeling, P. J., & Palmer, J. D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nature Reviews. Genetics*, *9*(8), 605–618. https://doi.org/10.1038/nrg2386

Li Y., Liu Z., Liu C., Shi Z., Pang L., Chen C., Chen Y., Pan R., Zhou W., Chen X.X., Rokas A., Huang J., Shen X.X. (2022). HGT is widespread in insects and contributes to male courtship in lepidopterans. *Cell*. 185(16):2975-2987.e10. 10.1016/j.cell.2022.06.014

Louda, S. M., Pemberton, R. W., Johnson, M. T., & Follett, P. A. (2003). Nontarget effects--the Achilles' heel of biological control? Retrospective analyses to reduce risk associated with biocontrol introductions. *Annual Review of Entomology*, *48*, 365–396. https://doi.org/10.1146/annurev.ento.48.060402.102800

Louis, F., Bézier, A., Periquet, G., Ferras, C., Drezen, J.-M., & Dupuy, C. (2013). The Bracovirus Genome of the Parasitoid Wasp Cotesia congregata Is Amplified within 13 Replication Units, Including Sequences Not Packaged in the Particles. *Journal of Virology*, *87*(17), 9649–9660. https://doi.org/10.1128/JVI.00886-13

Mahmoud, A. M. A., De Luna-Santillana, E. J., Guo, X., Reyes-Villanueva, F., & Rodríguez-Pérez, M. A. (2012). Development of the braconid wasp Cotesia flavipes in two Crambids, Diatraea saccharalis and Eoreuma loftini: Evidence of host developmental disruption. *Journal of Asia-Pacific Entomology*, *15*(1), 63–68. https://doi.org/10.1016/j.aspen.2011.07.007

Moyal, P., Tokro, P., Bayram, A., Savopoulou-Soultani, M., Conti, E., Eizaguirre, M., Le Rü, B., Avand-Faghih, A., Frérot, B., & Andreadis, S. (2011). Origin and taxonomic status of the Palearctic population of the stem borer Sesamia nonagrioides (Lefèbvre) (Lepidoptera: Noctuidae). *Biological Journal of the Linnean Society*, *103*(4), 904–922. https://doi.org/10.1111/j.1095-8312.2011.01666.x

Muller, H., Chebbi, M. A., Bouzar, C., Périquet, G., Fortuna, T., Calatayud, P.-A., Le Ru, B., Obonyo, J., Kaiser, L., Drezen, J.-M., Huguet, E., & Gilbert, C. (2021). Genome-Wide Patterns of Bracovirus Chromosomal Integration into Multiple Host Tissues during Parasitism. *Journal of Virology*, *95*(22), e00684-21. https://doi.org/10.1128/JVI.00684-21

Muller, H., Ogereau, D., Da-Lage, J.-L., Capdevielle, C., Pollet, N., Fortuna, T., Jeannette, R., Kaiser, L., & Gilbert, C. (2021). Draft nuclear genome and complete mitogenome of the Mediterranean corn borer, Sesamia nonagrioides, a major pest of maize. *G3 (Bethesda, Md.)*. https://doi.org/10.1093/g3journal/jkab155

Ono, R., Yasuhiko, Y., Aisaki, K., Kitajima, S., Kanno, J., & Hirabayashi, Y. (2019). Exosome-mediated horizontal gene transfer occurs in double-strand break repair during genome editing. *Communications Biology*, *2*(1), 1–8. https://doi.org/10.1038/s42003-019-0300-2

Pasquier-Barre, F., Dupuy, C., Huguet, E., Monteiro, F., Moreau, A., Poirié, M., & Drezen, J.-M. (2002). Polydnavirus replication: The EP1 segment of the parasitoid wasp Cotesia congregata is amplified within a larger precursor molecule. *Journal of General Virology*, *83*(8), 2035–2045. https://doi.org/10.1099/0022-1317-83-8-2035

Peccoud, J., Lequime, S., Moltini-Conclois, I., Giraud, I., Lambrechts, L., & Gilbert, C. (2018). A Survey of Virus Recombination Uncovers Canonical Features of Artificial Chimeras Generated During Deep Sequencing Library Preparation. *G3: Genes, Genomes, Genetics*, *8*(4), 1129–1138. https://doi.org/10.1534/g3.117.300468

Peccoud, J., Loiseau, V., Cordaux, R., & Gilbert, C. (2017). Massive horizontal transfer of transposable elements in insects. *Proceedings of the National Academy of Sciences*, *114*(18), 4721–4726. https://doi.org/10.1073/pnas.1621178114

Polaszek, A., & Walker, A. K. (1992). The Cotesia flavipes species-complex: Parasitoids of cereal stem borers in the tropics. *Proc. 4th Eur. Workshop Insect Parasitoids, F. Bin (Ed.). Perugia 1991. Redia 74,3 Appendix*, 335–341.

Sands, D. P. A., & Driesche, R. G. V. (2000). *Evaluating the Host Range of Agents for Biological Control of Arthropods: Rationale, Methodology and Interpretation*. 15.

Savary, S., Beckage, N., Tan, F., Periquet, G., & Drezen, J. M. (1997). Excision of the polydnavirus chromosomal integrated EP1 sequence of the parasitoid wasp Cotesia congregata (Braconidae, Microgastinae) at potential recombinase binding sites. *Journal of General Virology*, *78*(12), 3125–3134. https://doi.org/10.1099/0022-1317-78-12-3125

Schneider, S. E., & Thomas, J. H. (2014). Accidental Genetic Engineers: Horizontal Sequence Transfer from Parasitoid Wasps to Their Lepidopteran Hosts. *PLOS ONE*, *9*(10), e109446. https://doi.org/10.1371/journal.pone.0109446

Simion, P., Narayan, J., Houtain, A., Derzelle, A., Baudry, L., Nicolas, E., Arora, R., Cariou, M., Cruaud, C., Gaudray, F. R., Gilbert, C., Guiglielmoni, N., Hespeels, B., Kozlowski, D. K. L., Labadie, K., Limasset, A., Llirós, M., Marbouty, M., Terwagne, M., … Van Doninck, K. (2021). Chromosome-level genome assembly reveals homologous chromosomes and recombination in asexual rotifer Adineta vaga. *Science Advances*, *7*(41), eabg4216. https://doi.org/10.1126/sciadv.abg4216

Strand, M. R., & Burke, G. R. (2014). Polydnaviruses: Nature's Genetic Engineers. *Annual Review of Virology*, *1*(1), 333–354. https://doi.org/10.1146/annurev-virology-031413-085451

Teder, T., & Tammaru, T. (2002). Cascading effects of variation in plant vigour on the relative performance of insect herbivores and their parasitoids. *Ecological Entomology*, *27*(1), 94–104. https://doi.org/10.1046/j.0307-6946.2001.00381.x

Tettamanti, G., & Casartelli, M. (2019). Cell death during complete metamorphosis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1783), 20190065. https://doi.org/10.1098/rstb.2019.0065

Tewksbury, L., Casagrande, R., Blossey, B., Häfliger, P., & Schwarzländer, M. (2002). Potential for Biological Control of Phragmites australis in North America. *Biological Control*, *23*(2), 191–212. https://doi.org/10.1006/bcon.2001.0994

Van Etten, J., & Bhattacharya, D. (2020). Horizontal Gene Transfer in Eukaryotes: Not if, but How Much? *Trends in Genetics*, *36*(12), 915–925. https://doi.org/10.1016/j.tig.2020.08.006

Wallau GL, Capy P, Loreto E, Le Rouzic A, Hua-Van A. (2016). VHICA, a New Method to Discriminate between Vertical and Horizontal Transposon Transfer: Application to the

Mariner Family within Drosophila. *Mol Biol Evol*. 33(4):1094-109. doi: 10.1093/molbev/msv341. Epub 2015 Dec 18.

Wang, Z., Ye, X., Zhou, Y., Wu, X., Hu, R., Zhu, J., Chen, T., Huguet, E., Shi, M., Drezen, J.-M., Huang, J., & Chen, X. (2021). Bracoviruses recruit host integrases for their integration into caterpillar's genome. *PLOS Genetics*, *17*(9), e1009751. https://doi.org/10.1371/journal.pgen.1009751

Wybouw, N., Pauchet, Y., Heckel, D. G., & Van Leeuwen, T. (2016). Horizontal Gene Transfer Contributes to the Evolution of Arthropod Herbivory. *Genome Biology and Evolution*, *8*(6), 1785–1801. https://doi.org/10.1093/gbe/evw119

Wyder, S., Blank, F., & Lanzrein, B. (2003). Fate of polydnavirus DNA of the egg–larval parasitoid Chelonus inanitus in the host Spodoptera littoralis. *Journal of Insect Physiology*, *49*(5), 491–500. https://doi.org/10.1016/S0022-1910(03)00056-8

Xia, J., Guo, Z., Yang, Z., Han, H., Wang, S., Xu, H., Yang, X., Yang, F., Wu, Q., Xie, W., Zhou, X., Dermauw, W., Turlings, T. C. J., & Zhang, Y. (2021). Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. *Cell*, *184*(7), 1693-1705.e17. https://doi.org/10.1016/j.cell.2021.02.014

Zhang, H.-H., Peccoud, J., Xu, M.-R.-X., Zhang, X.-G., & Gilbert, C. (2020). Horizontal transfer and evolution of transposable elements in vertebrates. *Nature Communications*, *11*(1), 1362. https://doi.org/10.1038/s41467-020-15149-4

Data Accessibility and Benefit-Sharing

Data accessibly statement

All raw sequencing reads produced during this study are available in NCBI under accession number PRJNA816685. The two previously published datasets that we used in this study are available under accession numbers SRX9975807 (Illumina trimmed reads used for the assembly of *S. nonagrioides* reference genome (Muller, Ogereau, et al., 2021)) and

SAMN18533545 (“whole body” sample corresponding to a parasitized caterpillar for which no wasp larvae were present seven days post parasitism (Muller, Chebbi, et al., 2021)). Here, we called the former “Control” and the latter “Caterpillar”. The R scripts used for the analyses and to produce the figures of this project are available on GitHub (github/HeloiseMuller/Chimera).

Benefit-Sharing statement

C. typhae individuals used in this study originated from the International Centre of Insect Physiology and Ecology under the juridical framework of material transfer agreement CNRS 072057/IRD 302227/00. Access to genetic resources of the non-target lepidopteran species collected in France and benefit sharing have been done under the juridical framework of the Ministère de la Transition Ecologique (NOR: TRE2103218S/428, decision of March 26th, 2021).

Tables and Figures

Table 1. Characteristics of the three *de novo* assemblies produced in this study.

	#Contigs	Genome size assembly (Mb)	N50 (kb)	Complete BUSCO (duplicated)	Missing BUSCO
N. typhae	72,086	647	29.6	93.6% (1.9%)	2.0%
G. sparganii	304,227	972	8.5	82.1% (3.3%)	7.1%
C. phragmitella	293,788	942	4.7	67.4% (5.6%)	19.0%

Figure Captions

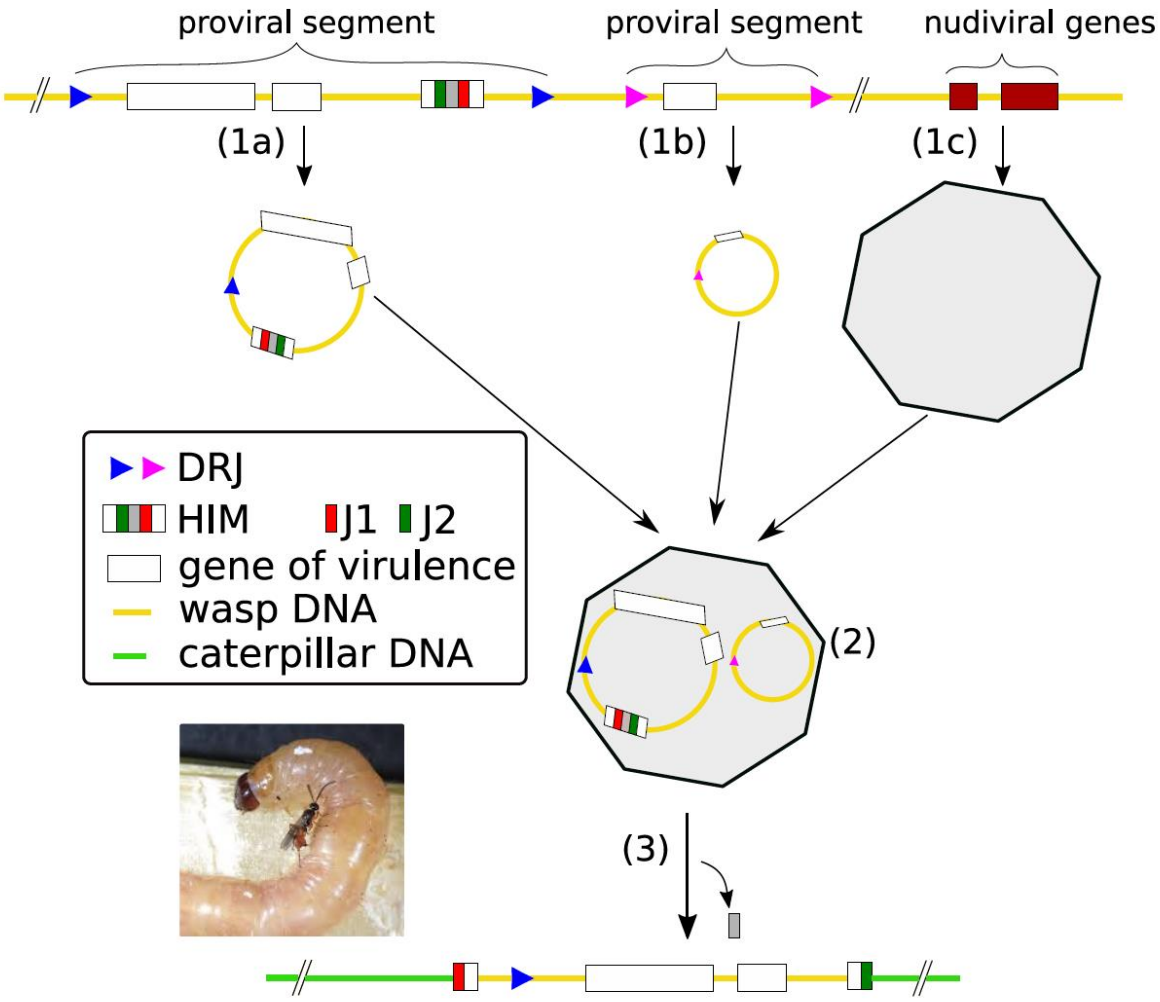

Figure 1. Structure of bracoviruses. In the wasp genome, (1a) is a HIM-containing segment, (1b) is a segment devoid of HIM, and (1c) corresponds to genes of viral origin. White rectangles are virulence genes. In the wasp calyx, (1a) and (1b) form DNA circles thanks to their DRJ, whereas (1c) form viral particles. In (2), DNA circles are packaged into viral particles, which are injected in the caterpillar host at the same time as the wasp eggs (see the picture as an example, where *C. typhae* is parasiting *S. nonagrioides*). In (3), the HIM-containing segment integrates via its HIM into the caterpillar genome, losing a ~50bp region located between J1 and J2.

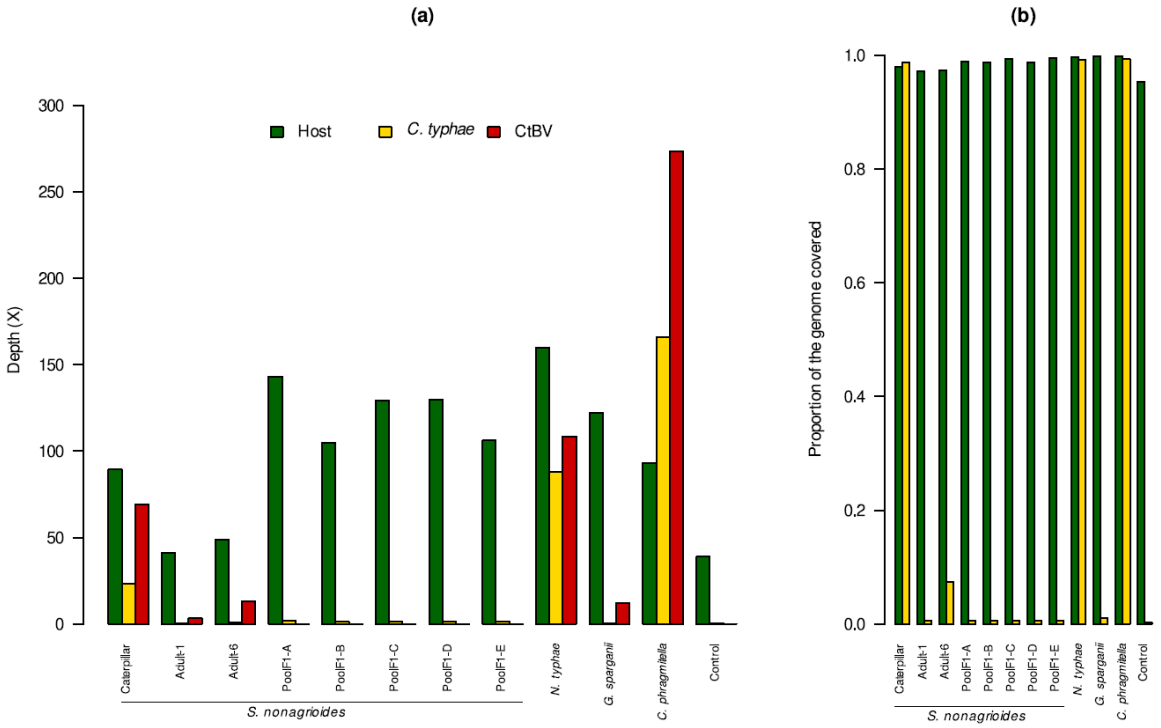

Figure 2. Sequencing depth and coverage over the genome of the wasp (*Cotesia typhae*), the bracovirus (CtBV) and the various moth hosts. (a) Barplot of the average sequencing depths. **(b)** Barplot of the proportion of the genomes of *C. typhae* and the hosts covered at least once. The green, yellow and red colors indicate the sequencing statistics over the whole genome of the host (*S. nonagrioides*, *N. typhae, G. sparganii*, or *C. phragmitella* depending on the samples), the whole genome of *C. typhae*, and the 27 CtBV DNA, respectively. The sequencing metrics are given for the ten samples produced for this study, but also for two other samples: Caterpillar is the sample corresponding to the whole body of a *S. nonagrioides* caterpillar sequenced seven days post-parasitism, and Control corresponds to the Illumina reads used for the assembly of the reference genome of *S. nonagrioides*. This sample has never been in contact with *C. typhae.*

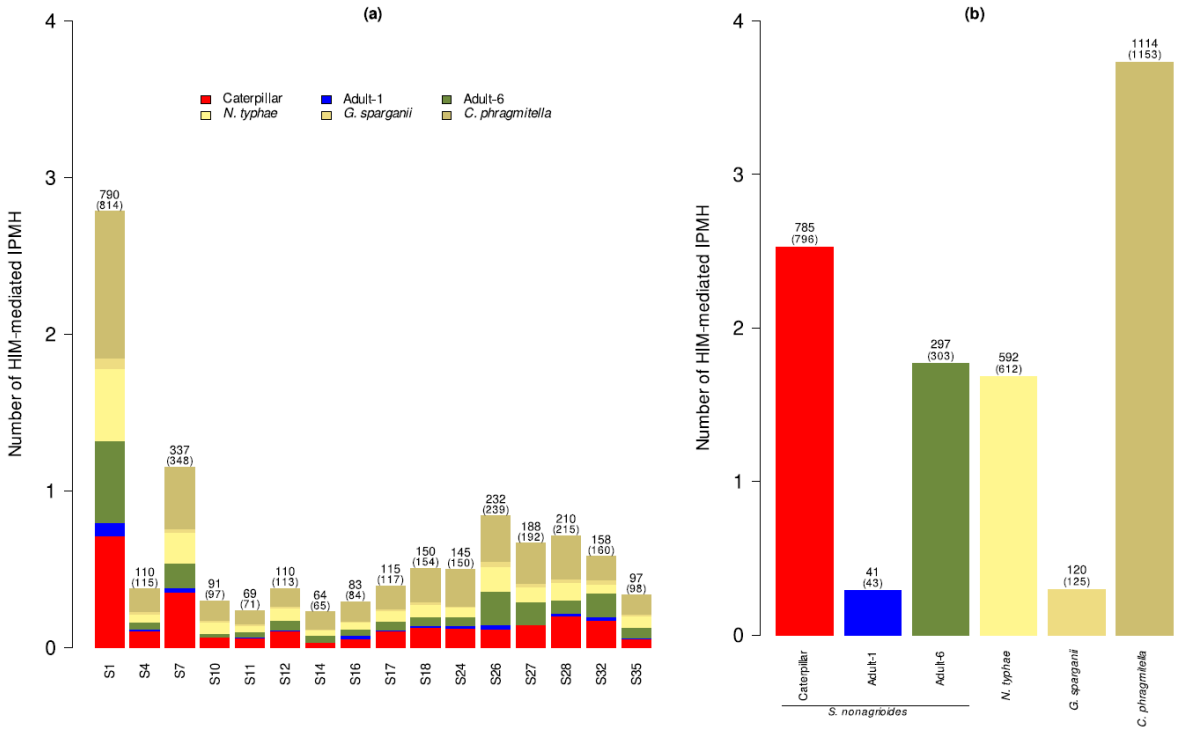

Figure 3. Number of HIM-mediated CtBV integration events per millions reads mapped on the host genome (IPMH) for each sample and CtBV circle. Absolute numbers of HIM-mediated integration events (IEs) and chimeric reads (in parentheses) are shown on top of each bar (without PCR duplicates). **(a)** Barplot comparing HIM-containing segments. **(b)** Barplot comparing samples.

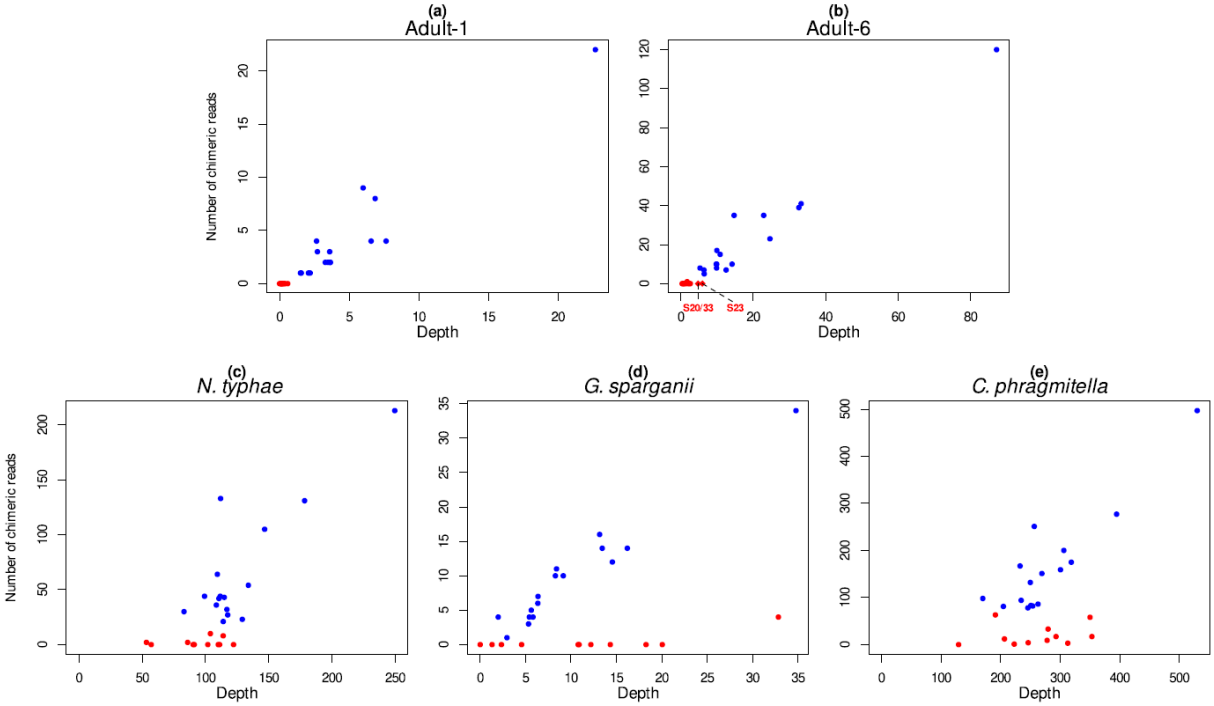

Figure. 4. Sequencing depth over CtBV DNA as a function of the number of chimeric reads. Since all reads are taken into account to calculate the sequencing depth, all chimeric reads are also taken into account in this plot: the ones that result from a HIM-mediated integration or not, and also the PCR duplicates. Each dot shows a CtBV (whose depth can result from sequencing its proviral segment, its circular form or its integrated form), in blue for HIM-containing CtBV, and in red for CtBV devoid of HIM. Results are shown by sample, with *S. nonagrioides* Adult-1 in **(a)**, *S. nonagrioides* Adult-6 in **(b)** where the two red diamonds represent S20/33 and S23, *N. typhae* in **(c)**, *G. sparganii* in **(d)**, and *C. phragmitella* in **(e)**.

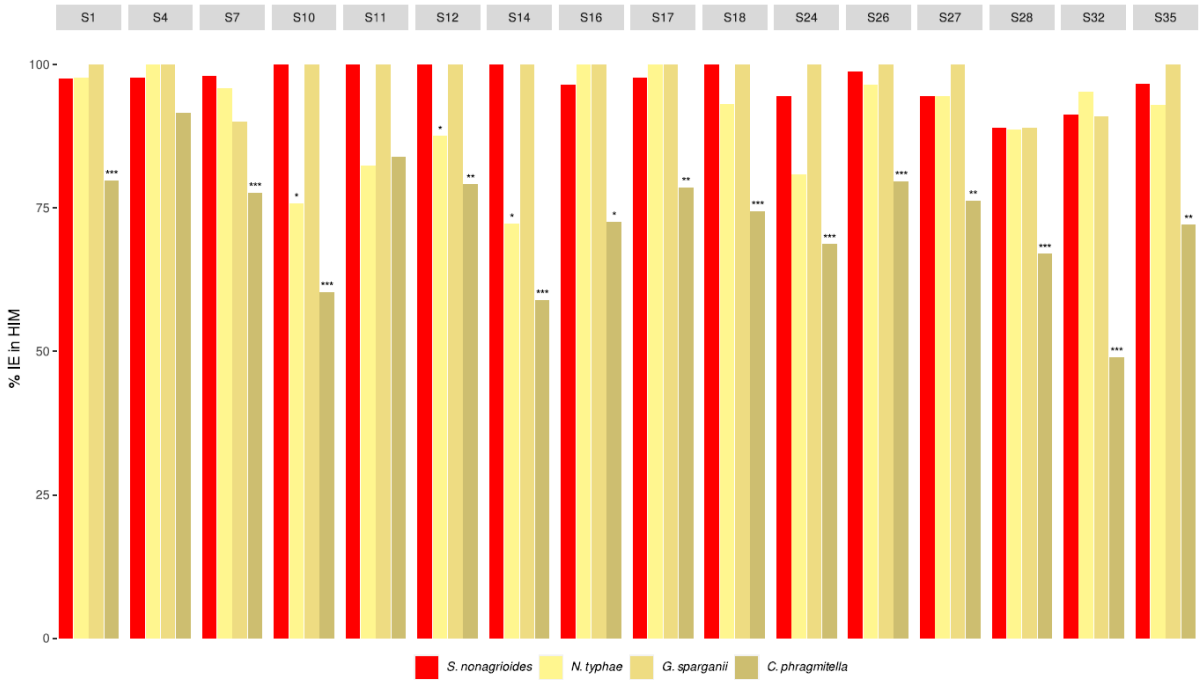

Figure 5. Percentage of HIM-mediated integration events (IEs) for the 16 HIM-containing circles. Data of the two adults *S. nonagrioides* and the sample Caterpillar are merged and compared to the three species not normally targeted by the wasp in the wild. Stars show whether the percentage of integration events which are HIM-mediated for the segment of interest is significantly different between the non-target species and *S. nonagrioides* (Fisher test; *: p.value<0.05, **: p.value<0.01; ***: p.value<0.001).

Supplementary Figure 1. Experimental design to obtain offspring of surviving parasitized caterpillars. In black are represented caterpillars of unidentified sex, in red the female adults, in green the male adults, and in yellow *C. typhae* (the square represents an adult in oviposition, and ovals represent wasp larvae). Although most moth couples produced some eggs, they lead to offspring only for 5 couples. Details of the experiment are shown for the progeny of Adult-1 (F1-Adult1), yet all the offspring were processed the same way. The DNA extractions (composed of 5 caterpillars each) were pooled 20 by 20 with equimolar ratio, leading to five libraries composed of 100 caterpillars each: library PoolF1-A to PoolF1-E. PoolF1-A was composed of 20 pools of F1-Adult5, PoolF1-B (shown on the figure as an example) of 20 pools of F1-Adult1, PoolF1-C of 16 pools of F1-Adult2 plus 1 pool of F1-Adult5 plus 3 pools of F1-Adult3, PoolF1-D of 20 pools of F1-Adult3 and PoolF1-E of 14 pools of F1-Adult4 plus 4 pools of F1-Adult5 plus 2 pools of F1-Adult1.

Author contributions

T.F., C.G., L.K., F.M., and H.M. designed the study; T.F. and H.M. performed lab work; T.F. conducted field work and provided samples; C.H. and H.M performed data analysis with input

from C.G. and F.M.; C.G. and H.M. wrote the manuscript, with input from E.H., F.M., and all other coauthors.