TECHNISCHE UNIVERSITÄT DARMSTADT

Use of Next Generation Sequencing to Improve Identification and Quantification of Baculovirus Diversity

Vom Fachbereich Biologie der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines

Doctor rerum naturalium

(Dr. rer. nat)

Dissertation

von Dipl. Biol. (Uni) Gianpiero Gueli Alletti

aus Mannheim

1. Referent: Prof. Dr. Johannes A. Jehle

2. Referent: Prof. Dr. Gerhard Thiel

Tag der Einreichung: 24.09.2018

Tag der mündlichen Prüfung: 07.12.2018

Darmstadt 2018D 17

Gueli Alletti, Gianpiero: Use of Next Generation Sequencing to Improve Identification and Quantification of Baculovirus Diversity
Darmstadt, Technische Universität Darmstadt,
Jahr der Veröffentlichung der Dissertation auf TUprints: 2018
URN: urn:nbn:de:tuda-tuprints-82879
Tag der mündlichen Prüfung: 07.12.2018

Für meine Frau, Inese Vieško.

CONTENT

LIST OF ABBREVIATIONS

%	per cent	ml	mililiter
× *g*	multiple of *g*	mM	milimolar
°C	degree Celsius	MNPV	multiple nucleopolyhedrovirus
µg	microgram	n	number of tested individuals
µl	microliter	N	number of independent replicates
µm	micrometer	n.d.	not determined
aa	amino acid	NGS	Next Generation Sequencing
bp	base pair	nm	nanometer
Bt	*Bacillus thuringiensis*	NPV	nucleopolyhedrovirus
BV	budded virus	nt	nucleotide
cm	centimeter	OECD	Organisation for Economic Co-operation and Development
ddH_2O	bidistilled water	OB	occlusion body
dpi	days post infection	ODV	occlusion derived virion
DNA	2'deoxyribonucleic acid	ORF	open reading frame
dsDNA	double stranded DNA	p.i.	post infection
e. g.	for example	PCR	polymerase chain reaction
et *al.*	and others	pH	measure of acidity of basicityof aqueous solutions
etc	et cetera	PM	peritrophic membrane
EU	European Union	RFLP	restriction fragment length polymorphism
g	gram	RNA	ribonucleic acid
GV	granulovirus	sec	seconds
h	hour	SNPV	single nucleopolyhedrovirus
ICTV	International Commitee on Taxonomy of Viruses	spp.	species pluralis; multiple species
JKI	Julius Kühn Institute	TAE	tris base, acetic acid and EDTA buffer solution
kb	kilobase pair	TBE	tris base, boric acid and EDTA buffer solution
LD_{50}	median lethal dose	UV	ultraviolet light
m	meter	V	Volt
mg	miligram	v/v	volume in volume
min	minute	w/v	weight in volume

Chapter I: General Introduction

Baculoviruses based insecticides: A brief retrospective

"Of all viruses known to mankind, baculoviruses are the most beneficial from an anthropocentric [sic!] viewpoint." (Miller, 1997).

Lois Kathryn Miller, always reminds us with her statement that an insect disease can be useful for mankind, as long as the insect host is considered as a pest to our human-centered spheres (Passarelli et al., 2014). In fact, the first observation of baculovirus infection were documented in the silk production in the poem "De Bombyce" by the Italian bishop Marco Vida of Cremona at the end of the Renaissance (Benz, 1986). In the rearing of silk worm larvae (*Bombyx mori*), diseased caterpillars were accounted with symptoms which were referred to as "melting" or "wilting". Taken together these two examples represent the negative outcome of baculovirus infection in commercial aspects. However, the use of baculovirus based insecticides can be dated back to late 19th century. Albeit, the knowledge about virus infection was limited, baculovirus mixtures were used to dam up outbreaks of the nun moths (*Lymantria monacha*) in pines in Europe (Huber, 1986). Since then, baculovirus based insecticides have emerged in numbers with over 60 biocontrol products and a commercial turnover of 49.2 million US$ in 2010, representing about 12 % of the whole biocontrol product sales (Beas-Catena et al., 2014; Lacey et al., 2015). However, the development of baculovirus based products always has to succeed several hurdles, as reviewed in the OECD consensus document on the assessment of baculoviruses used as insecticides: "*Baculoviruses…are endowed with a variety of application constraints. Among these, the narrow host range, a limited life time and the slow speed of action resulting in demands on application strategies, are important factors in their failure to effectively compete with chemicals…*" (OECD, 2002). In this context, the commercial production of baculoviruses either depends on their hosts – which dictates rearing of a pest species – or. Nevertheless, once developed baculovirus based insecticides benefit from the durability their morphology and contribute to an environment friendly and sustainable pest control.

Baculoviruses: Morphology, infection cycle and taxonomy

The insect virus family of *Baculoviridae* comprises dsDNA viruses with rod-shaped, enveloped virions infecting larval stages of the insect orders Lepidoptera, Diptera and Hymenoptera. During their life-cycle baculoviruses produce an occluded phenotype of their virions, the so-called occlusion derived virion (ODV). The ODVs are embedded in crystalline protein matrixes, termed occlusion bodies (OBs), which naturally protect the virions from hazardous environmental conditions as exposure to UV radiation, draught or excess humidity as well as enzymatic degradation (Eberle et al., 2012; Herniou et al., 2011). Based on their phylogenetic relationship, baculoviruses have been classified recently into four genera. This classification also reflects their host range and to certain extend the OB morphology (Figure 1). Viruses of the genera *Alphabaculovirus* and *Betabaculovirus* only infect species of the insect order Lepidoptera, whereas those from the genera *Gammabaculovirus* and *Deltabaculovirus* infect species from the orders Hymenoptera and Diptera, respectively (Herniou and Jehle, 2007). ODVs of alpha-, gamma- and deltabaculoviruses are occluded into polyhedral OBs. OBs of these genera are termed nucleopolyhedra which besides the morphology also implicates the main structural matrix protein: polyhedrin. In case of alphabaculoviruses nucleocapsids, can be

further enveloped singly or as multiple units in one envelope; historically, these observations where used to term corresponding viruses single nucleopolyhedrovirus or multiple nucleopolyhedrovirus, like for *Autographa californica multiple nucleopolyhedrovirus* (AcMNPV) (Ayres et al., 1994; Pang et al., 2001). The OB of a nucleopolyhedrovirus is generally 1 to 2 µm (or up to 15 µm) in size and contains few to several hundreds of virions. In contrast to alpha-, gamma- and deltabaculoviruses, that can occlude multiple enveloped virions into a polyhedral OB, the single-nucleocapsid virions of betabaculoviruses are occluded into an ovo-cylindrical, granule-shaped OB (Gati et al., 2017; Paillot, 1926). Species of the genus *Betabaculovirus* are termed granulovirus after their granule-shaped OBs. The granulovirus OB itself is relatively small with about 0.2 to 0.4 µm in diameter and about 0.5 µm in length, when compared to nucleopolyhedroviruses.

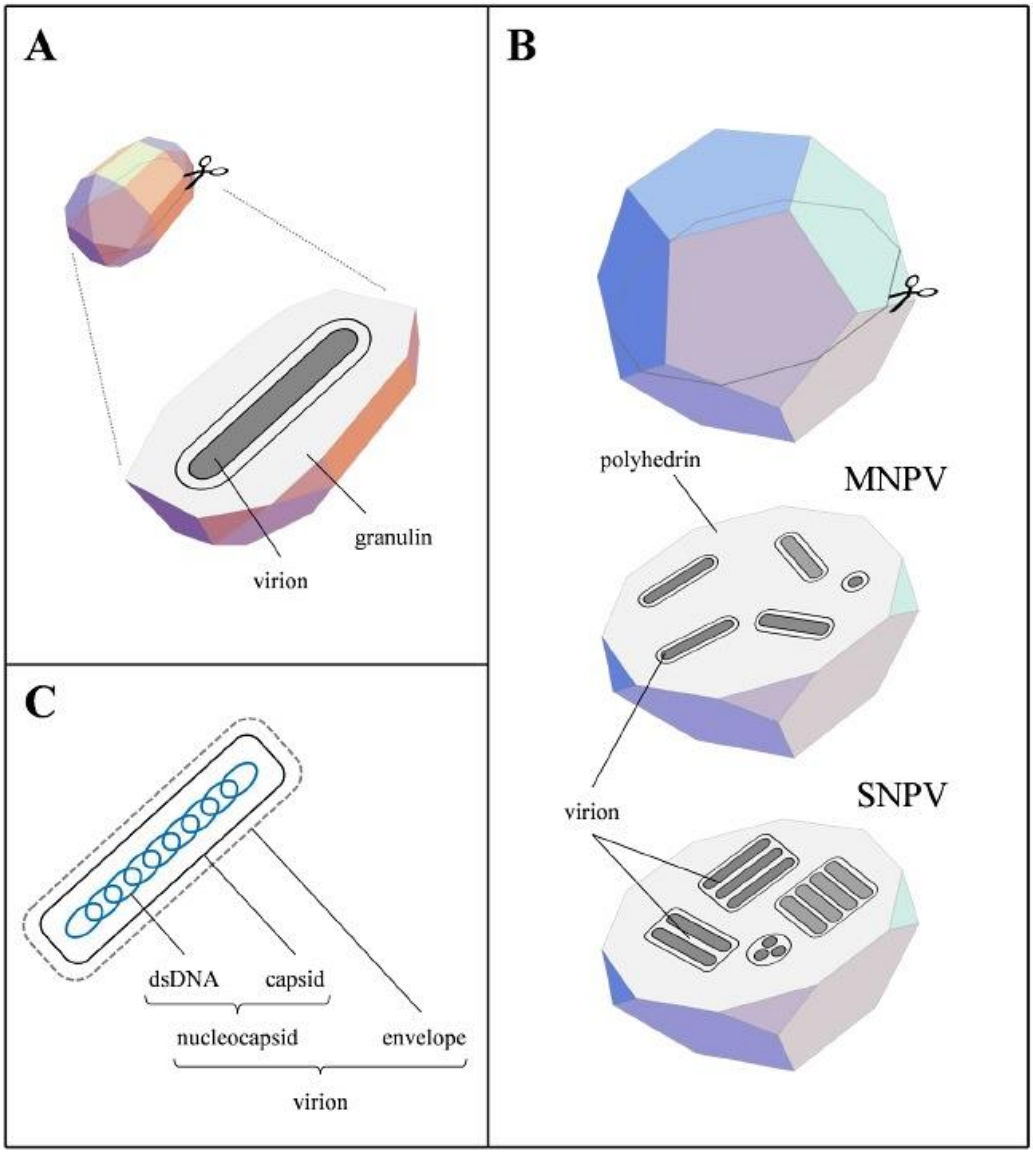

Figure 1 Schematic illustration of baculovirus occlusion bodies (OBs) of (A) granulovirus (GV) and (B) nucleopolyhedrovirus (NPV). Structure of a baculovirus virion with a single nucleocapsid (C). GV have singly embedded virions surrounded by Granulin in a ovo-cylindrical matrix. The OBs of NPV are shaped polyhedral and embed up to many virions. In MNPV multiple nucleocapsids are enveloped, while SNPV virions only harbor one single nucleocapsid. Sizes of OBs are given only schematic, original sizes are mentioned in the paragraph. Originally published in Wennmann (2014).

The second baculovirus virion phenotype is the so-called budded virus, or seldomly also referred to as budded virion (BV). The BV phenotype is generated when the nucleocapsids bud through the plasma membrane at the surface of infected cells. BVs contain typically only one nucleocapsid and is responsible for the systemic cell to cell infection in the host (Harrison et al., 2018a). Latest native electron-microscopic (cryo-EM) investigations of the BV morphology of group I and group II *Alphabaculovirus* have revealed that BV share a common elongated ovoid shape (Figure 2) (Wang et al., 2016). The spikes of BVs from both groups, namely GP64 in case of group I and F-protein in case of group II alphabaculoviruses, respectively, are incorporated in a 6-7 nm thick layer. Although both types of spikes are more or less distributed evenly over the surface, GP64 is located in higher densities at both apical ends of the BV. Prior findings of the BV morphology deriving from negative-staining electron microscopy describe this phenotype instead as long rod-shaped virions with apical ends containing the fusion proteins GP64 and/or F-protein (Figure 3) (Fraser, 1986; Harrap, 1972).

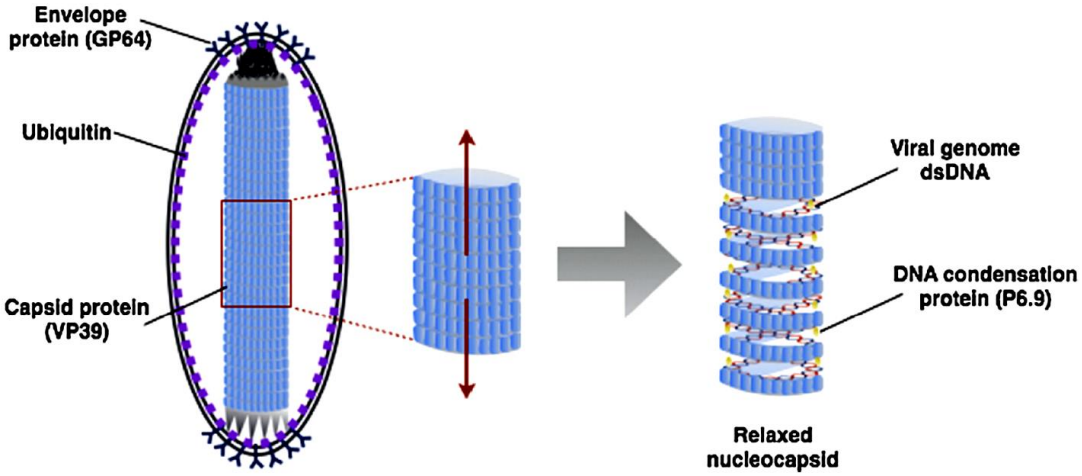

Figure 2 Structural model of the baculovirus budded virus (BV) represented by AcMNPV. The major envelope protein GP64 is distributed at the polar ends of the ovoid shaped enveloped. The DNA genome is condensed by the protein P6.9 and organized as super coil in the nucleocapsid. The major capsid protein VP39 forms a spiral structure visualized by its relaxed state. Originally published in Wang et al. (2016)

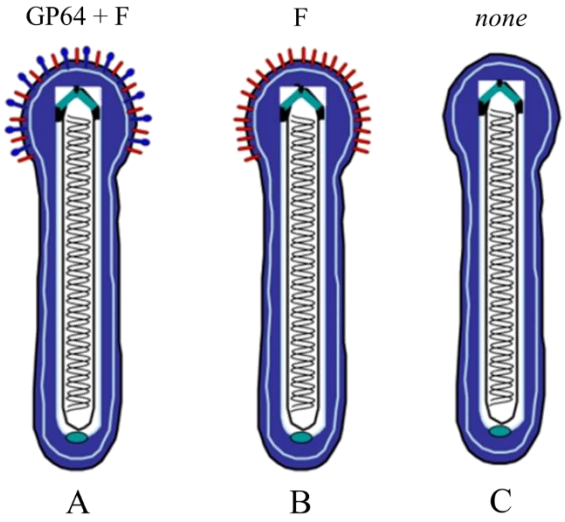

Figure 3 Schematic structural modell of the baculovirus budded virus (BV) deriving from negative-staining electron microscopy. The BVs show apical ends and belong to group I alphabaculoviruses (A), group II alphabaculoviruses, beta- and deltabaculoviruses (B) and gammabaculoviruses (C) Rohrmann (2013)

In host larvae, baculovirus infection undergoes a biphasic cycle connected to these distinct virion phenotypes. Readily after the ingestion of OB, the protein matrix is dissolved in the alkaline midguts of larvae and ODVs are released (Figure 4). As first obstacle, these virions must pass the peritrophic membrane (PM), which mainly consist of glycoproteins and chitin. This layer forms a natural physical barrier which protects the midgut epithelium from physical damage and pathogens. Baculoviruses may encode proteins which help to disintegrate the PM of the midgut. These proteins are described as enhancin proteins, enhancins, viral enhancing factors (vef), or after their first observations "synergistic factors" (Tanada, 1959). Synergistic factors, later referred as enhancins, were first reported for *Pseudaletia unipuncta granulovirus* (PsunGV). This granulovirus encodes for a synergistic factor that increases the susceptibility of *Pseudaletia unipuncta nucleopolyhedrovirus* to larvae of the armyworm, *P. unipuncta*, in co-infections (Tanada, 1959; Tanada and Hukuhara, 1971). In general, baculovirus enhancins belong to a metallopeptidase family (Lepore et al., 1996a). This family of peptidase contains a zinc binding motif of the amino acids histidine, glutamine, two variable amino acids and a posterior histidine (HEXXHX(8,28)E), a motif which is also found in bacterial enhancin-like peptidases (Galloway et al., 2005). In baculoviruses, these metalloproteases can be found either way as part of the occlusion body matrix, like in PsunGV, or also as components of the ODV, as demonstrated for *Lymantria dispar multiple nucleopolyhedrovirus* (LdMNPV) (Slavicek and Popham, 2005). Enhancins can further appear in multiple copies in the baculovirus genome, as shown for *Xestia c-nigrum granulovirus* (XcenGV), which encodes four copies of an enhancin. Although the function of enhancins is well-described, they are not necessarily essential for baculovirus infection in general, as not all baculoviruses encode these proteins. For these baculoviruses, infection is not fully understood.

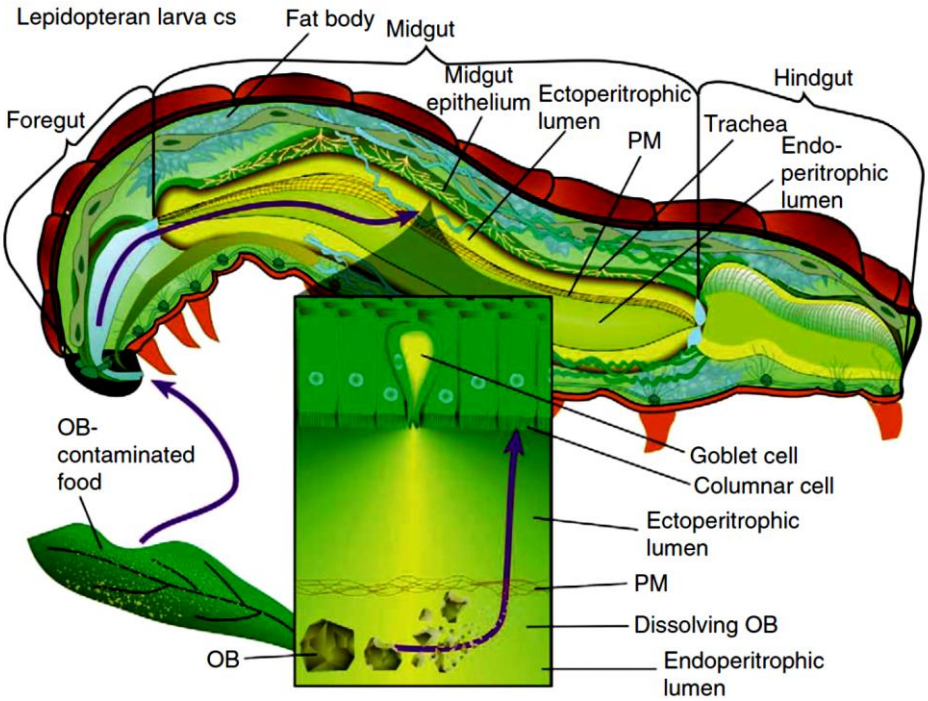

Figure 4 *Per os* infection by baculoviruses. A cross-sectional (cs) representation of the anatomy of an insect larva is depicted. A baculovirus occlusion body (OB) enters by oral uptake of contaminated food. OBs pass through the foregut end enter the alkaline midgut where they are dissolved in the lumen and release ODVs. The midgut is surrounded by the peritrophic membrane (PM). Originally published by Slack and Arif (2006)

In all baculovirus infections, the ODV fuse with cell surface of the epithelium cells and thusly release the nucleocapsids into the cells. This fusion is administered with the help of specific proteins linked to the ODV envelope: the *per os* infectivity factors (pifs). Eight genes, *pif-0/p74* (*Ac138*), *pif-1* (*Ac119*), *pif-2* (*Ac22*), *pif-3* (*Ac115*), *pif-4* (*Ac96*), *pif-5* (*Ac148/odv-e56*), *pif-6* (*Ac68*), and *vp91*/*p95* (*Ac83*) have been identified to encode for factors that are essential in the primary infection process (Fang et al., 2009; Javed et al., 2017; Nie et al., 2012; Sparks et al., 2011; Zhu et al., 2013). In first experiments with deletion mutants of AcMNPV it has been demonstrated that at least the four proteins P74 (PIF-0), PIF-1, PIF-2 and PIF-3 form a stable complex on the surface of the ODV (Peng et al., 2010), however protein interaction assays (fluorescence complementation (BiFC) assays) demonstrated that six protein-protein interactions form the *per os* infectivity complex (PIF complex) (Zheng et al., 2017). As demonstrated by microscopic investigations of a genetically modified AcMNPV isolate, the nucleocapsids are delivered to the nuclei of infected cells in a cork-screw-drive manner managed by the manipulation of the cell structure protein actin. The baculovirus nuclear import has been lately reviewed, but in general for nucleopolyhedroviruses (Au et al., 2013). The transcription of baculovirus genes is linked to the infction cycle and can be roughly divided into three temporal stages during the infection: an early, late and very late phase (Friesen, 1997; Lu and Miller, 1997). During the early stage of infection, the host RNA polymerase II is used for baculovirus transcription, while with the beginning of the late stage baculoviruses encode for their own RNA polymerase. As mentioned, in alphabaculovirus infections the nuclear membrane remains intact, which leads to hypertrophied nuclei that may fill the entire cell as reviewed by Rohrmann (2013). In betabaculovirus infections, the nuclear membrane is degraded at an early point during infection and the nucleocapsids are assembled in a nuclear-cytoplasmic milieu (Federici, 1997). After this first step of infection, the nucleocapsids bud through the plasma membrane and form the above described phenotype budded virus (BV).

As transported via the hemolymph and the tracheal system, BVs soon cause infections in multiple tissues of the host (Engelhard et al., 1994). Essential for the systemic infection is the modified viral envelope of BVs, that differs in its components form ODVs. The main envelope fusion protein of group I alphabaculoviruses is GP64, while group II alphabaculoviruses and betabaculoviruses encode the homologous F-protein (Monsma et al., 1996; Pearson et al., 2000). These proteins facilitate the virus entry as demonstrated in experiments with *gp64* deleted AcMNPV. In these experiments, the loss of GP64 was recovered with the F-protein deriving from either group II alphabaculoviruses or betabaculoviruses, which lead to a recovery of AcMNPV BV infectivity (Lung et al., 2002; Yin et al., 2003). Group I alphabaculoviruses can also contain genes for *f-proteins*, which may play also an essential role in virus infections of this group, however *gp64* is restricted to alphabaculovirus lineage (Wang et al., 2014). The *gp64* and *f-protein* genes are both transcribed during the early and late (Washburn et al., 2003) stage of infection and both proteins are incorporated into the cytoplasm membrane where they become part of the BV envelope during the event of budding. Here, the GP64 and F-protein interact with receptors of the host cell membrane that initiate the endocytosis of BVs. The responsible host receptors of this process are unknown. First, the BV becomes enclosed within a coated vesicle, the endosome, which later releases the nucleocapsid (Figure 5). In the final phase of infection, the assembly of ODVs, the second virus phenotype generated during the baculovirus replication cycle, is initiated. Nucleocapsids remain within the host cell and become enclosed by a *de novo* synthesized envelope that carries all ODV envelope proteins. At this time

of infection, the gene of the major OB matrix protein, *polh*/*gran*, reaches a very high level of transcription and expression. The Polyhedrin and/or Granulin constitutes the crystalline protein matrix of the OBs. The OBs are released to the environment from dying or dead liquefied larvae.

Baculovirus taxonomy and molecular phylogeny

In general, baculoviruses are normally named after the host, from which they have been initially isolated in followed by "nucleopolyhedrovirus" for alpha-, gamma- and deltabaculoviruses, e.g. *Rachioplusia ou nucleopolyhedrovirus* (*Alphabaculovirus*), *Neodiprion lecontei nucleopolyhedrovirus* (*Gammabaculovirus*) and *Culex nigripalpus nucleopolyhedfrovirus* (*Deltabaculovirus*), as well as the host plus "granulovirus" for betabaculoviruses (e.g. *Cydia pomonella granulovirus*). In order to counter the proliferation of different abbreviations of virus names, first authors of baculoviruses nowadays tend to follow certain rules, in which the first two letters of genus and species of a host followed by NPV (nucleopolyhedrovirus) or GV (granulovirus) are used. Examples are *Agrotis ipsilon nucleopolyhedrovirus* (AgipNPV), *Plutella xylostella nucleopolyhedrovirus* (PlxyNPV), or *Phthorimaea operculella granulovirus* (PhopGV). However, baculoviruses with historical references retained their traditional abbreviations, like *Autographa californica multiple nucleopolyhedrovirus* (AcMNPV), *Lymantria dispar multiple nucleopolyhedrovirus* (LdMNPV), *Rachioplusia ou multiple nucleopolyhedrovirus* (RoMNPV), or *Cydia pomonella granulovirus* (CpGV), as their original abbreviations have been extensively used in the literatrue. In numbers, the majority of baculoviruses has been isolated from lepidopteran species, namely 456 alphabaculoviruses and 148 betabaculoviruses, while only 30 gamma- and 27 deltabaculoviruses have been isolated from hymenopteran and dipteran species, respectively. Whereas microscopic differentiation of genera can be readily performed for large the number of alpha- and betabaculoviruses, the intra-generic species differentiation is difficult, if not impossible, by microscopy methods, as all baculoviruses of a distinct genus have a common morphology. For this reason, species classification highly depends on molecular methods like molecular phylogeny of baculovirus genomes and conserved genes. All baculoviruses share double-stranded, circular supercoiled DNA genomes, with sizes varying from about 80 to 180 kilobase pairs (kb), encoding 90 to 180 genes. Since the early 1990s, genomes of more than 60 baculoviruses have been fully sequenced and deposited to GenBank; including 155 genomes from isolates and/or strains of the genus *Alphabaculovirus*, 46 from *Betabaculovirus*, 2 from *Gammabaculovirus* and 1 from *Deltabaculovirus*, respectively (see Table 1 in the appendix of this thesis). However, these deposited complete genome sequences only contribute in parts to the taxonomic species list of *Baculoviridae* according to the ICTV. The reasons for these discrepancies are partial sequences which have been used in the past to recognize novel baculovirus species as well as missing data in case of recent complete sequences. The complete list of recognized baculovirus species can be found on https://talk.ictvonline.org/files/master-species-lists/m/msl/7185 (dated March 2018). This list comprises all background information as proposals for the recognition and representing isolates. With the rising number of full genome sequences, baculovirus genomics has progressed in identifying highly conserved genes based on the predicted amino-acid sequences of the putative open reading frames. In all four baculovirus genera a set of 38 homologous genes termed "baculovirus core genes" has been identified so far (Javed et al., 2017; Wennmann et al., 2018). Strikingly, 19 of these baculovirus core genes are also conserved in the genomes of closely related nudiviruses, 11 in hytrosaviruses and 20 are present in polydnaviruses (Rohrmann, 2013). Additional conserved genes have been described for alpha-

and betabaculoviruses, as well as for some gammabaculoviruses (see Table 1). Phylogenetic analyses based on the putative amino-acid sequences of these 38 genes and evaluating of the nucleotide substitutions according to the Kimura two parameter substitution model (K2P) have emerged in assessment of evolutionary trees of baculoviruses (Garavaglia et al., 2012; Miele et al., 2011).

Prior to the phylogenetics of the 38 core genes, evolutionary traits of alpha- and betabaculoviruses in particular can be based on concatenated alignments of the partial amino-acid sequences of the three genes *late expression factor (lef)* 8, *lef*-9 and *polyhedrin* for alphabaculoviruses and/or *granulin* for betabaculoviruses, respectively (Jehle et al., 2006b). Albeit, *polyhedrin* homologs are also present in the two remaining baculovirus genera, this approach is driven by the fact that baculovirus infecting Lepidoptera only comprise alpha- and betabaculoviruses. Thus, the background information on the host range is important for supporting baculovirus phylogeny. The three genes *polyhedrin/granulin*, *lef*-8 and *lef*-9 harbor sufficiently enough variation, that the species demarcation can be based on the K2P nucleotide distances of these partial sequences to a certain extent. More than a decade ago, that particular species criterion was set arbitrary according to observations of resulting resolutions of the phylogenetic tree. In this context, additionally to other criteria, two baculoviruses are recognized to belong to different species when the K2P distance of these loci is larger than 0.05 substitutions/site. K2P distances smaller than 0.015 between two isolates have been proposed to indicate isolates are definitely members of the same species. This species demarcation has been recently confirmed in comparisons of phylogenetic trees based on the 37 core genes with trees deriving from the concatenated *polyhedrin/granulin*, *lef*-8 and *lef*-9 sequences (Wennmann et al., 2018). For this reason, rules for identifying and including baculoviruses as novel species/isolates in the catalog of the ICTV are still valid. The species description comprises in addition to the phylogeny based on these K2P distances the host range as well as a description of the morphology and pathology.

Table 1 Baculovirus core genes according to Garavaglia et al. (2012) and Javed et al. (2017) with designated Autographa californica multiple nucleopolyhedrovirus (AcMNPV) ORF number and gene function according to Rohrmann (2013)

Gene designation	AcMNPV ORF	Description
DNA Replication/processing:		
lef-2	6	DNA replication/primase-associated factor
lef-1	14	DNA primase
DNA polymerase	65	DNA replication
helicase	95	Unwinding DNA
alk-exo	133	Involved in DNA recombination and replication
Transcription/RNA polymerase:		
p47	40	RNA polymerase subunit
lef-8	50	RNA polymerase subunit
lef-9	62	RNA polymerase subunit
vlf-1	77	Involved in expression of the p10 and polh genes
lef-4	90	RNA polymerase subunit/capping enzyme
lef-5	99	Transcription initiation factor
Structure		
vp1054	54	Nucleocapsid protein
desmoplakin	66	Present in nucleocapsid
gp41	80	Tegument protein

Table 1 Baculovirus core genes according to Garavaglia et al. (2012) and Javed et al. (2017) with designated Autographa californica multiple nucleopolyhedrovirus (AcMNPV) ORF number and gene function according to Rohrmann (2013)

vp39	89	Major capsid protein
odv-e25	94	ODV envelope protein
p6.9	100	Nucleocapsid protein
odv-ec43	109	Associated with ODV 72 69
49k	142	Required for BV production
odv-e18	143	ODV envelope protein
odv-e27	144	ODV envelope protein
Per os infectivity factors/ODV structure:		
pif-2	22	Required for per os infection (PIF-2)
ac68	68	Required for per os infection (PIF-6)
vp91/p95	83	Viral capsid-associated protein
ac96	96	Required for per os infection (PIF-4)
pif-3	115	Required for per os infection (PIF-3)
pif-1	119	Mediates binding of ODV to midgut (PIF-1)
p74	138	Mediates binding of ODV to midgut (PIF-0)
odv-e56	148	ODV envelope protein (PIF-5)
Predicted enzymes:		
ac53	53	Likely involved in nucleocapsid assembly
p33	92	Sulfhydryl oxidase
38k	98	Required for nucleocapsid assembly
Other:		
ac78	78	Unknown function/transmembrane domain
ac81	81	Unknown function
p18	93	Egress of nucleocapsids
p40	101	Subunit of protein complex
p48	103	BV production and ODV envelopment

"Cutworms" and baculoviruses infecting *Agrotis* spp.

The non-taxonomic term "cutworms" comprises several species from the lepidopteran family of Noctuidae. Among cutworms, the genus *Agrotis* forms the largest group with 739 reported species (Hill, 1983). The common cutworm or turnip moth, *Agrotis segetum* (Denis & Schiffermüller, 1775) (Figure 6), occurs all over Africa, Eurasia, South Asia and Japan, but has not been reported on the American continent (Hill, 1983) (Figure 7). Adult moths are of 30 to 40 mm in wingspan and are of dark-brown color with black markings. As many other members of Noctuidae, *A. segetum* moths are nocturnal. Flight seasons have been reported to take from May to June and from August to September. Female moths can lay up to 1000 legs which are preferably placed on crops, or on the soil around. Depending on temperature, first instar larvae hatch after 3 (at +29-30 °C) to 24 days (at +10-12 °C) (Hill, 1983). They live on leaves but cause little to no harm to crops. With increasing age, larvae migrate into lower soil levels where they feed on root and stems of the vegetation. Larval development undergoes five to six instars within 24-40 days in which the larvae can reach 40-50 mm in size. The feeding behavior comes to a resting stage right before pupation starts. The pupae stage can last up to three weeks and adults once hatch emerge to the surface.

Agrotis ipsilon (Hufnagel, 1766), or the black and greasy cutworm (Figure 8), has a very similar morphology, but can be easily differentiated by the Y-shaped black markings on the foreside of the wings, its genitalia or on a molecular level using the lepidopteran genetic marker for the *cytochrome c oxidase* gene. In contrast to the turnip moth, *A. ipsilon* is almost cosmopolitan distributed in the Northern hemisphere within 60 °N latitude and equator, as well as in South Africa, New Zealand, East Australia and parts of South America (Hill, 1983) (Figure 9). The flight behavior and larval development is similar to *A. segetum*. Members of both species are considered to be significant agricultural pests to many crops due to their polyphagous behavior. For example, *A. ipsilon* has been reported as pest organism in corn in the United States (Clement and McCartney, 1982; Engelken et al., 1990; Showers et al., 1983), or in turf grass (Potter, 1998). Current pest management includes methods such as tillage, hand-collection, flooding, ploughing, the use of diatomaceous earth and chemical insecticides (mostly based on pyrethrin) or the use of entomopathogenic nematodes, bacteria (e.g. *Bacillus thuringiensis*) and fungi (Caballero et al., 1993; Gokce et al., 2013; Ignoffo and Garcia, 1979; Wraight et al., 2010).

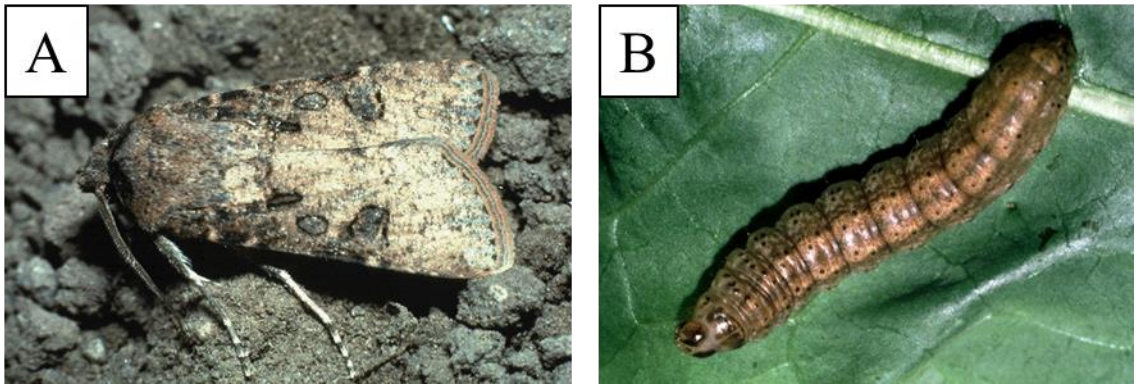

Figure 6 *Agrotis segetum* adult moth (A) and larva (B). Pictures downloaded from www.cabi.org, copyrights: ©Crown Copy (A) and ©J. Porter (B).

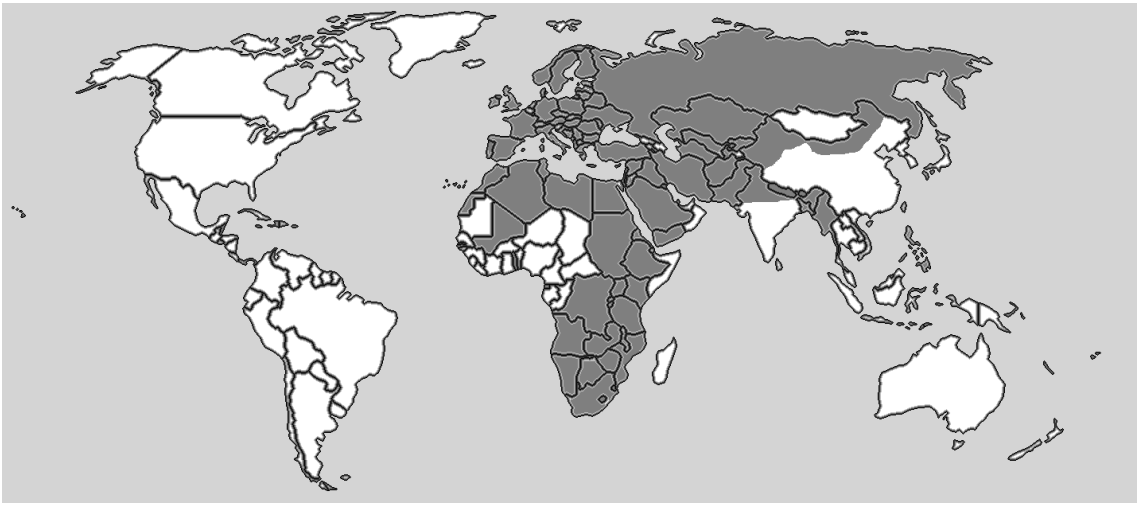

Figure 8 Distribution map of the turnip moth, *Agrotis segetum* based on the EPPO database of quarantine pests available online at www.eppo.int.

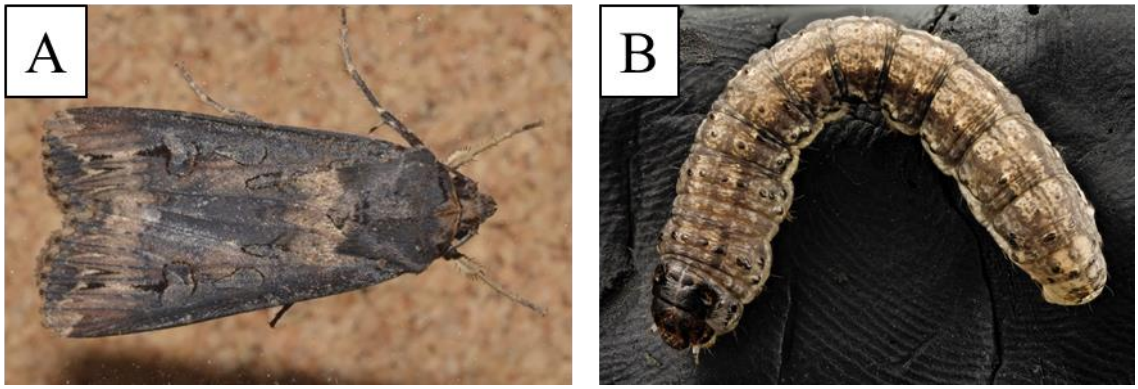

Figure 7 *Agrotis ipsilon* adult moth (A) and larva (B). Pictures downloaded from www.discoverlife.org, copyrights: ©Cameron Prybol (A) and ©Discover Life (B).

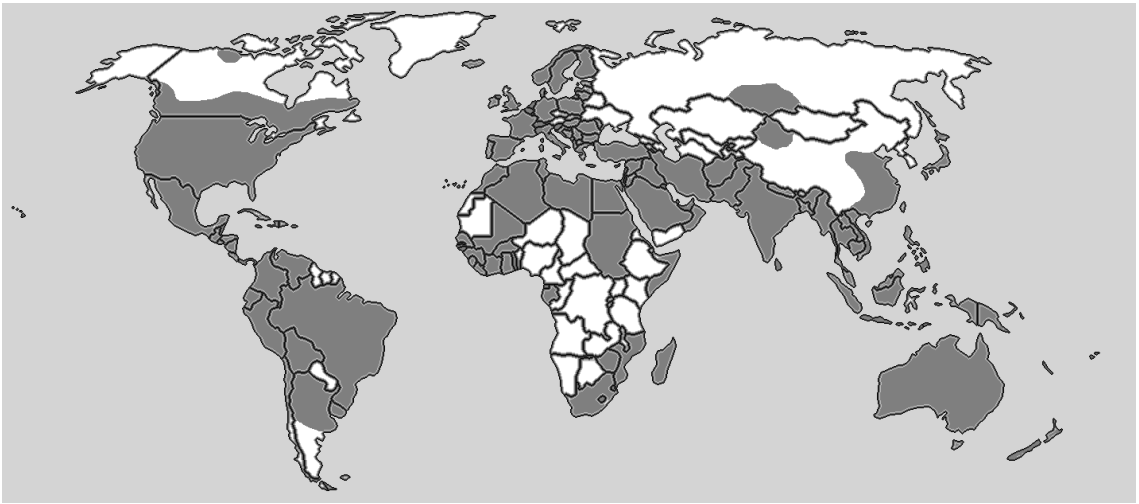

Figure 9 Distribution map of the dark sword grass moth, *Agrotis ipsilon* based on the datasheet provided by CABI available online at www.cabi.org.

As for cutworms, the non-taxonomic term "*Agrotis* baculovirus complex" describes a group of baculovirus species, which have been isolated from the noctuid species *A. segetum* and *A. ipsilon*. So far, the *Agrotis* baculovirus complex comprises three species of the genus *Alphabaculovirus*, namely *Agrotis ipsilon nucleopolyhedrovirus*, *Agrotis segetum nucleopolyhedrovirus A* and *Agrotis segetum nucleopolyhedrovirus B* as well as the *Betabaculovirus* species *Agrotis segetum granulovirus* (Boughton et al., 1999; Jakubowska et al., 2006; Wennmann et al., 2015a; Zethner et al., 1987). Formerly, *Agrotis segetum nucleopolyhedrovirus* has been recognized as a single species, however based on the genetic distance between AgseNPV-A and AgseNPV-B they are now considered as two different baculovirus species (Wennmann et al., 2015a). Isolates from the four baculovirus species have been tested in bioassays with lepidopteran larvae for their specific host ranges. As presented by recent publications A. ipsilon nucleopolyhedrovirus (AgipNPV), A. segetum A and B nucleopolyhedrovirus (AgseNPV-A and AgseNPV-B, respectively) and A. segetum granulovirus (AgseGV) are able infect larvae of both *A. segetum* and *A. ipsilon* (El-Salamouny et al., 2003; Harrison, 2009). All four viruses are further able to simultaneously infect single larvae as demonstrated in particular in co-infections of AgseNPV-B and AgseGV in larvae of *A. segetum* (Wennmann and Jehle, 2014; Wennmann et al., 2015c). Furtherly, AgseNPV-A, AgseNPV-B and AgipNPV share homologues of different copies of the *viral enhancing factor* (*vef*) gene and a homologue of that, the *enhancin*, has also been identified in sequences of two Chinese isolates of AgseGV, AgseGV-L1 and AgseGV-XJ (Figure 10) (Wennmann et al., 2015a).

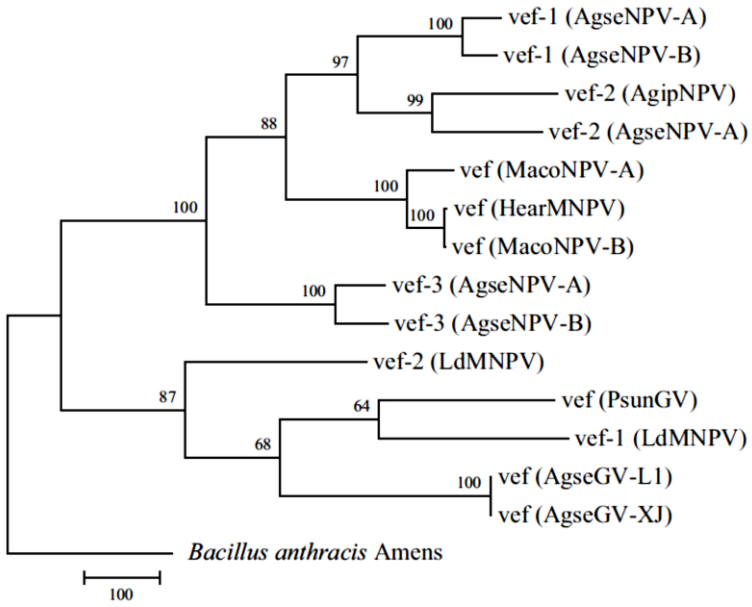

Figure 10 Phylogenetic analysis based on aligned amino acid sequence of enhancin/viral enhancing factors (vef) of *Agrotis* baculoviruses. Maximum Parsimony tree was calculated by using the tree-bisection-reconnection (TBR) algorithm. Percentage bootstrap values (1,000 replicates) are given at each node. The bacterial enhancin sequence of *Bacillus anthracis* Amens (GenBank accession no. AE017034) was used as outgroup. Originally published in Wennmann et al. (2015a).

In addition to AgseGV-L1 and AgseGV-XJ, there are also Spanish and Danish isolates of AgseGV, which have been formerly used in biological control of cutworms (Zethner, 1980; Zethner et al., 1987). Whereas these two isolates have been characterized by RFLP analyses (Bourner et al., 1992) with indistinguishable profiles, complete genome sequences have been available only for AgseGV-L1 and AgseGV-XJ (Zhang et al., 2014), both showing only

negligible nucleotide differences. According to historic records from the Julius Kühn Institute (JKI), specimens of a European isolate of AgseGV, which has been termed AgseGV-DA, have been distributed from Germany to several research institutes in Europe, including the Horticulture Research International (Warwick, UK) and the Laboratory of Virology (Wageningen, the Netherlands). All these laboratory stocks apparently derived from the same original material: infected *A. segetum* larvae collected in Austria in 1964 and sent to Federal Biological Research Institute (now the JKI) in Darmstadt, Germany. AgseGV-DA was used as representative isolate for the species assignment of *Agrotis segetum granulovirus* (Gueli Alletti et al., 2015) and is characterized further as subject of this thesis.

The "codling moth" and Cydia pomonella granulovirus

The codling moth, *Cydia pomonella* (Linnaeus, 1758), is one of the most important pests in apple orchards. The greyish moth with a wing-spread of about 2 cm and a characteristic cross band of chocolate brown (Figure 10) deposits lays 50 to 75 eggs on leaves, twigs and fruits (Blomfield and Giliomee, 2011). Eggs hatch in about a week and after a few days the young larvae enter the fruit and eat their way to the core. Larvae complete development in three to five weeks. Then they leave the fruit and seek suitable places for hiding, such as underneath bits of loose bark and other protected places mainly on the tree and sometimes in the debris on the ground (Börner, 1997). Here cocoons are spun and pupation follows. Depending on climate, one, two or even more generations (in warm regions) are possible. The larvae of the codling moth injure and contaminate the fruits by feeding and feces: the literally worm in the apple is familiar to everyone. Fruits very often drop prematurely, those remaining are not marketable. Codling moth attacks mainly apples, but to lower extent also pears, quinces and walnuts can be affected.

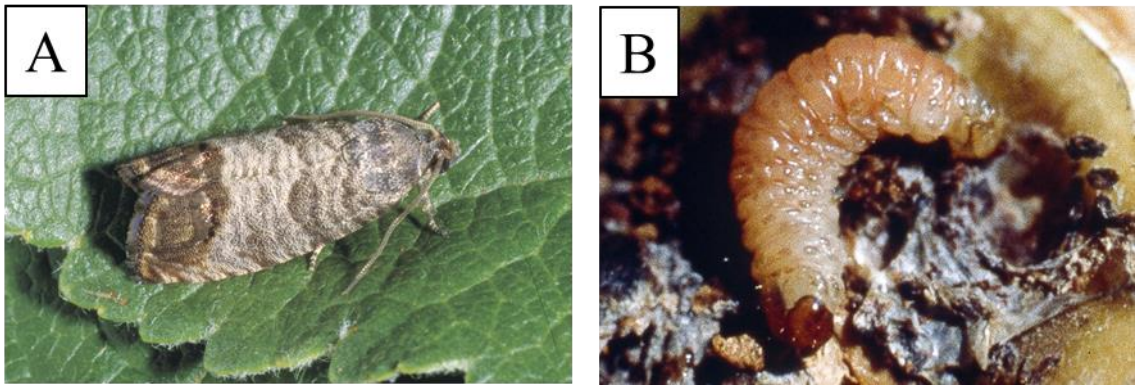

Figure 10 *Cydia pomonella* adult moth (A) and larva on discected apple (B). Pictures downloaded from www.agric.wa.gov.au, copyrights: ©Government of Western Australia, 2018.

Figure 11 Distribution map of the codling moth, *Cydia pomonella* based on the EPPO database of quarantine pests available online at www.eppo.int.

Codling moth is generally controlled by chemical insecticides and/or biological control agents, such as Cydia pomonella granulovirus (CpGV), *B. thuringiensis* (Bt) preparations and/or pheromone-based mating disruption. Codling moth control with commercial CpGV preparations has been applied for 30 years and was mainly based on first CpGV isolatediscovered in Mexico, therefore termed "Mexican isolate" or CpGV-M (Tanada, 1964). The genome sequence of CpGV-M has been one of the first complete baculovirus sequences (Luque et al., 2001). According to phylogenetic analyses of both the partial sequences of *lef*-8, *lef*-9 and *granulin* as well as based on the analyses of 36 out of 38 baculovirus core genes, the next common relative to CpGV is *Cryptophlebia leucetrata granulovirus*, CrleGV (Figure 12) (Harrison et al., 2016; Lange and Jehle, 2003).

In the last years reduced efficacy against Cydia pomonella granulovirus based commercial products has been reported in selected *C. pomonella* populations. In the past, all commercial CpGV preparation, registered in Europe, were based on the Mexican isolate CpGV-M. First observations on reduced susceptibility of *C. pomonella* populations to CpGV-M were made in Southern Germany since 2002 (Fritsch et al., 2005; Sauphanor et al., 2006). It was demonstrated that the observed failure of CpGV application was caused by genetically inherited resistance. The genetic basis of this particular resistance, termed type I resistance, was based on a Z-linked inheritance (Asser-Kaiser et al., 2010). Phenotypically, the isolate CpGV-M and similar isolates show little to no effect in larvae with type I resistance. Although the virus host interaction is not fully understood, the reduced infectivity is correlated by the presence of a distinct 24 bp insertions into gene *pe*38, coding for a supposed transcription factor of CpGV (Gebhardt et al., 2014). Other CpGV isolates lacking this insertion were shown to be resistance-breaking (Eberle et al., 2008; Eberle et al., 2009; Gebhardt et al., 2014). Recently two novel types of resistance have been reported from two German field populations, NRW-WE and SA-GO, found in North Rhine-Westphalia and Saxony, respectively (Jehle et al., 2017; Sauer et al., 2017a). Type II resistance was described for laboratory codling moth strains CpR5M/CpR5S selected from NRW-WE-08. Inheritance of this resistance is determined as being dominant and autosomal (Sauer et al., 2017b). Larvae with this trait are resistant against isolates from the CpGV genome group A, like CpGV-M and from the CpGV genome group E (Jehle et al., 2017; Sauer et al., 2017a).

The third type of resistance, termed as type III resistance, found in SA-GO showed mixed characteristics of type I and type II resistance (Sauer et al., 2017c). However, similar to type II resistance, larvae with type III resistancewere susceptible for isolates with the CpGV genome group B and are resistant for isolates from with the genome type A, C, D and E (Sauer et al., 2017c). It has to be pointed out that resistance to CpGV isolates is still restricted to relatively few orchards where the majority of populations came from Germany, while eight populations originated from Italy, six from France, two from Austria and one from the Netherlands and Switzerland, respectively (Schmitt et al., 2013). Thus, strategies are necessary to effectively control these populations and to prevent the spread of the resistance. Different new CpGV isolates are available that are able to break the resistance the CpGV isolates reported (Jehle et al., 2017).

In the past lineages of *Cydia pomonella granulovirus* have been characterized either by their restriction endonucleases profiles (REN) (Crook et al., 1985; Harvey and Volkman, 1983; Rezapanah et al., 2008) or by a combination of these profiles with nucleotide sequence data deriving from partial sequencing of the genes *granulin*, *late expression factor* (*lef*) 8 and *lef*-9 (Eberle et al., 2009). These methods assessed the characterization of isolates from different geographic origins and comprised: the Mexican isolate CpGV-M and its cloned isolate CpGV-M1 (Crook et al., 1997), the English isolate CpGV-E (Crook et al., 1985), the Russian isolate CpGV-R (Harvey and Volkman, 1983), twelve Iranian isolates (i.a. CpGV-I01, -I07, -I08, -I12, -I66, and -I68 and) (Eberle et al., 2008; Eberle et al., 2009; Rezapanah et al., 2008) and the Georgian isolates CpGV-G01 and -G02 (Eberle et al., 2009). The broad collection of Iranian isolates was first grouped into three CpGV lineages according to the comparison with CpGV-M, -E2 and -R. Thus, these isolates either belonged to a CpGV-M-like, a CpGV-E-like or a CpGV-R-like lineage, while mixtures of different genotype lineages could also be readiliy identified by so called “sub-molar” bands in the restriction profiles (Rezapanah et al., 2008). However, only due to the addition of the partial sequence data, a fourth lineage could be introduced (Eberle et al., 2009). The fifth CpGV lineage was identified by analyses of the above described insertion in *pe*38 in the Candian isolate CpGV-S and by the molecular phylogeny of 35 out of 38 baculovirus core gene of selected CpGV isolates from different genome lineages (Gebhardt et al., 2014). Recently, the alignment of six fully sequenced CpGV isolates, namely CpGV-M, CpGV -M1, CpGV -E2, CpGV -I07, CpGV-I12 and CpGV-S confirmed the division into five phylogenetic lineages of CpGV: the genome groups A to E (Figure 13) (Wennmann et al., 2017). The genomes of these representative isolates range from 120,816 bp to 124,269 bp, for which CpGV-M and CpGV-I12 are the most derived lineages, followed by CpGV-E2 and CpGV-S, whereas CpGV-I12 is the most ancestral lineage. From the most basal to the most derived lineages an overall trend of increasing genome sizes to certain extend also decreasing %GC-content was observed. Furthermore, a total number of 788 single nucleotide polymorphisms (SNPs) was detected compared to the CpGV-M sequence, of which 534 SNPs were specific for exclusively one representative isolate of the lineages. Therefore, besides to other mutations, genome group C (CpGV-I07) harbored 356 specific SNPs, genome group E (CpGV-S) 101, genome group B (CpGV-E2) 54, genome group D (CpGV-I12) 21 and finally genome group A (CpGV-M) two specific SNPs. The analyses of these isolate group-specific SNPs are useful for the identification of present genome groups in newly discovered CpGV isolates.

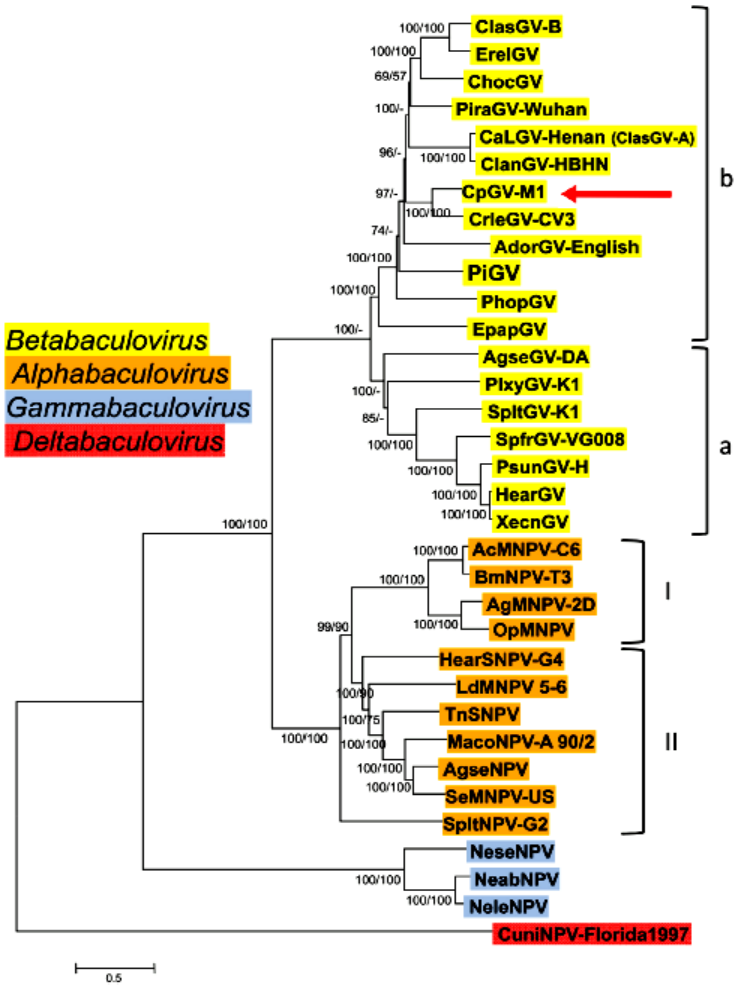

Figure 12 Relationships of CpGV and isolates of other baculovirus species, inferred from the predicted amino acid sequences of 36 baculovirus core genes. The phylogenetic tree was constructed using the minimum-evolution (ME) method (P6.9 was excluded). Different genera are indicated with colored text background. Both the group I and II clades of genus *Alphabaculovirus* and the a and b clades of *Betabaculovirus* are indicated. Bootstrap values >50% for both ME and maximum likelihood (ML) analysis are indicated for each interior branch (ME/ML). Originally published in Harrison et al. (2016).

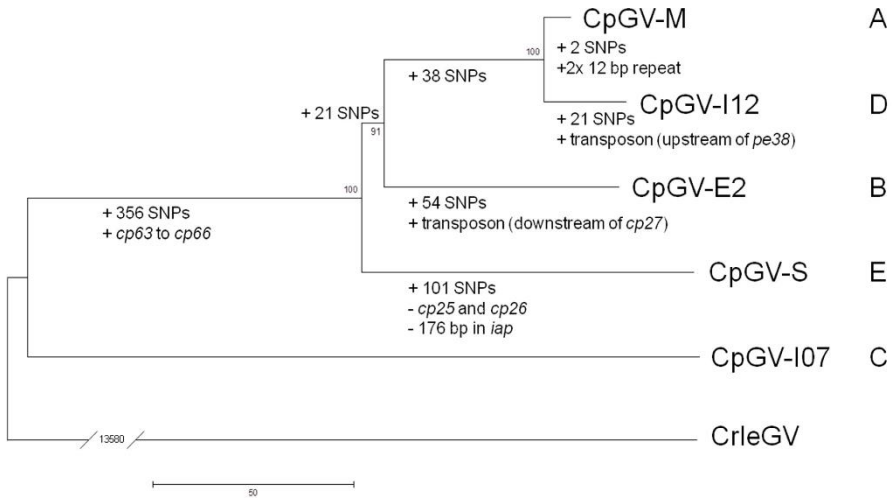

Figure 13 CpGV phylogeny and evolutionary trends of CpGV-M, -I12, -E2, -S and -I07. Phylogeny is based on the minimum evolution method of the alignment of whole genome nucleotide sequences with 1000 bootstrap replicates. CpGV genome groups A to E are given to the right. Cryptophlebia leucotreta granulovirus (CrleGV) was set as outgroup. Group and isolate specific single nucleotide polymorphisms (SNPs) and additional specific genomic features are given below each branch. Originally published in Wennmann et al. (2017)

Next generation sequencing (NGS) in baculovirus virology and aims of the thesis

Characterization of baculovirus, besides the pathology and common genera-specific morphology, undergoes also an extensive evaluation on a molecular level. As described for the molecular phylogeny, full genome sequences support the assessment of genomes by either the three baculovirus marker genes *polyhedrin/granulin*, *lef-8* and *lef-9* or by the so far 38 baculovirus core genes, used for baculovirus classification (Jehle et al., 2006b; Wennmann et al., 2018). However, the genetic information given by fully sequenced genomes aside of these genes, e. g. in terms of insertions, deletions, inversions, duplications or occurrence of SNPs, also allows following species-specific micro-evolution and deciphering the complexity of baculovirus populations and co-evolution with the hosts (Gilbert et al., 2014; Gilbert et al., 2016). One example for the fast development from classical sequencing of genes of interest (e.g. *me53*) to whole genome NGS sequencing is the comparison of Australian isolates of Helicoverpa armigera single nucleopolyhedrovirus (HaSNPV)(Baillie and Bouwer, 2012; Noune and Hauxwell, 2016). While the first attempts of characterizing two of these geographically distinct isolates concentrated on the two genes *dbp1* and *me53*, the second approach from the same laboratory institute relied on the evaluation of the genomic composition of 22 isolates either originating from the Australian HaSNPV-AC53 or coming from Eurasia, North- and South-America. Although all isolates shared a generally high sequence identity to HaSNPV-AC53 (~94% - 99%) sequence differences were identified in several open reading frames (e.g. *hoar*, *DNA-polymerase*) and in *homologous repeat sequences* (*hrs*).

Furthermore, the genetic stability of virus isolates has been also analyzed for a South African Cryptophlebia leucotreta granulovirus (viz. CrleGV-SA) spanning a time period of 15 years (van der Merwe et al., 2017). Again, all sequenced isolates, in this case isolated at different time-points, showed a high identity to each other only differing up to 5% in SNPs. Taking together, the combination of the read-depth generated by the large datasets of NGS, in particular Solexa Illumina, can be assessed to identify and quantify the genomic composition of population of single baculovirus isolates (Chateigner et al., 2015). In case of an *in vivo* amplified isolate of *Autographa californica multiple nucleopolyhedrovirus* (viz. AcMNPV-WP10) it was postulated that not only the population of physically separate OBs can be considered as diverse, within one single OB nucleocapsids possess different genetic traits. Thus, even single OBs are considered as presumably diverse in their genomic composition (Chateigner et al., 2015).

The NGS sequencing was used in the following five chapters to (i) characterize the European isolate AgseGV-DA; (ii) to assess the virus stability of AgseNPV-B in continuous passages in insect cell culture and (iii) to quantify the genomic compositions of different commercially available CpGV isolates. In chapter II, the fully sequenced genome of AgseGV-DA is compared to the two Chinese isolates, AgseGV-L1 and AgseGV-XJ (Zhang et al., 2014), as these full genome sequences lacked details on putative open reading frames (ORFs). Furthermore, the molecular phylogeny based on 37 core genes for this particular isolate was compared to the phylogeny based on *granulin*, *lef-8* and *lef-9*. To achieve a complete description qualifying for requirements for species inclusion by the ICTV this data was connected to the morphology of AgseGV-DA as well as to the pathology in host larvae. In chapter III the permissivity of the insect cell line AiE1611T was evaluated for AgseGV-DA and AgseNPV-B. The cell line AiE1611T had been generated from *A. ipsilon* eggs which are susceptible for AgipNPV

(Harrison and Lynn, 2008; Lynn and Harrison, 2016). While AgseGV and AgseNPV-B are able to co-infect susceptible larvae (Wennmann et al., 2015c), it has been demonstrated that AiE1611T is only permissive for AgseNPV-B.

Several plaque purified clones of AgseNPV-B were characterized based on restriction patterns generated in digestions with *Eco*RI and *Hin*dIII and one particular clone, namely AgseNPV-B PP2 was used in chapter IV for the continuous passages in AiE1611T in order to assess the genomic stability in cell culture amplifications. The isolates from the continuous passages of PP2 were characterized by the *in vivo* activity in host larvae as well as by Solexa Illumina sequencing (NGS) and comparison of the fully sequenced genomes. AgseNPV PP2 showed endured a high genomic stability during the passages, while the responses in larvae were scattering over ten passages. In chapter V the construction of an *in vitro* cloned bacmid of AgseNPV-B, termed bAgseNPV-B, is described This bacmid is contains a deletion of about 43 kb in the AgseNPV-B genome but still possesses the full bacterial cassette, which had been inserted by homologous recombination into *hr6* of AgseNPV-B. The deletion and SNPs were characterized by NGS of bAgseNPV-B DNA isolated from *Escherichia coli* DH5-α clones carrying the bacmid as extra-chromosomal DNA. Finally, chapter VI combines data-sets generated of NGS of commercially available CpGV isolates with their responses in three codling moth populations with different susceptibility to CpGV-M and CpGV-S (Asser-Kaiser et al., 2011; Sauer et al., 2017a; Sauer et al., 2017b; Sauer et al., 2017c). It is shown in this chapter, that the differences in virulence of CpGV isolates is correlated to different genomic compositions of the commercial isolates, which has been assessed by the comparing the different SNP compositions to those of the CpGV genome groups A to E (Wennmann et al., 2017).

Chapter II: The genome sequence of Agrotis segetum granulovirus, isolate AgseGV-DA, reveals a new *Betabaculovirus* species of a slow killing granulovirus

This chapter is published with few modifications in:
Gueli Alletti, G., Eigenbrod, M., Carstens, E. B., Kleespies, R. G., Jehle, J. A. (2017), The genome sequence of Agrotis segetum granulovirus, isolate AgseGV-DA, reveals a new *Betabaculovirus* species of a slow killing granulovirus, J. Invert. Pathol., 146:48-68

Abstract

The European isolate Agrotis segetum granulovirus DA (AgseGV-DA) is a slow killing, type I granulovirus due to low dose-mortality responses within seven days post infection and a tissue tropism of infection restricted solely to the fat body of infected *Agrotis segetum* host larvae. The genome of AgseGV-DA was completely sequenced and compared to the whole genome sequences of the Chinese isolates AgseGV-XJ and AgseGV-L1. All three isolates share highly conserved genomes. The AgseGV-DA genome is 131,557 bp in length and encodes for 149 putative open reading frames, including 37 baculovirus core genes and the *per os* infectivity factor *ac110*. Comprehensive investigations of repeat regions identified one putative non-*hr* like origin of replication in AgseGV-DA. Phylogenetic analysis based on concatenated amino acid alignments of 37 baculovirus core genes as well as pairwise distances based on the nucleotide alignments of partial *granulin, lef-8* and lef-*9* sequences with deposited betabaculoviruses confirmed AgseGV-DA, AgseGV-XJ and AgseGV-L1 as representative isolates of the same *Betabaculovirus* species. AgseGV encodes for a distinct *enhancin* gene, with a distant relation to the enhancins from the genus *Betabaculovirus*.

Introduction

The family of *Baculoviridae* comprises occluded dsDNA viruses with rod-shaped, enveloped virions infecting larval stages of the insect orders Lepidoptera, Diptera and Hymoptera. Infections of insect larvae with these viruses have been described long before any knowledge of structure and mode of infection arose (Benz, 1986). Based on their phylogenetic relationship, baculoviruses are classified into four genera, which also reflect their host range and to a certain extent, occlusion body (OB) morphology. Members of the genera *Alphabaculovirus* and *Betabaculovirus* only infect Lepidoptera, whereas viruses from *Gammabaculovirus* and *Deltabaculovirus* are specific for Hymenopteran and Dipteran species, respectively (Herniou et al., 2003; Jehle et al., 2006a). In contrast to alpha-, gamma- and deltabaculoviruses, that can occlude multiple enveloped virions into a polyhedral occlusion body (OB), the single-nucleocapsid virions of betabaculoviruses are occluded into an ovo-cylindrical, granule-shaped OB (Gati et al., 2017). The term granulovirus is derived from that well-described OB morphology, previously used in baculovirus taxonomy (Theilmann et al., 2005). Currently, all granuloviruses are classified in the genus *Betabaculovirus* (Herniou and Jehle, 2007). Baculovirus phylogeny has been under intensive revision with the advent of molecular markers, based on conserved genes such as *polyhedrin*/*granulin*, *lef-8*, *lef-9* and/or *pif-2* (Herniou et al., 2003; Jehle et al., 2006b). Extensive enhancement of baculovirus molecular phylogeny is based on a set of baculovirus core genes, predictably present in all baculovirus species. Until now, 37 genes have been accepted as baculovirus core genes (Garavaglia et al., 2012) and their predicted amino-acid sequences are used for phylogenetic analyses. Aside from phylogeny-based

classification, three types of granuloviruses have been identified by their pathogenesis in infected host larvae (Federici, 1997). Slow-killing type I granuloviruses include Trichoplusia ni granulovirus (TnGV), Pseudaletia unipuncta granulovirus (PsunGV), Xestia c-nigrum granulovirus (XecnGV) and Adoxophyes orana granulovirus (AdorGV) (Hilton and Winstanley, 2008a; Mukawa and Goto, 2008). In type I granulovirus infections, the production of OBs is restricted to the fat body of host larvae, and larval development remains largely unaffected, allowing infected larvae to grow to their final larval stages. In contrast, fast-killing type II granuloviruses such as Cydia pomonella granulovirus (CpGV) (Tanada and Leutenegger, 1968), cause killing within a few days after infection and formation of OBs is not limited to the fat body but occurs systemically in multiple larval tissues, such as trachea, neurons, and others. Finally, infections caused by the type III granulovirus Harrisina brillians granulovirus (HbGV) have been so far observed only in larvae of *Harrisina brillians.* Their tissue tropism is limited to the midgut epithelium cells and, unlike other types of granuloviruses, the infected larvae discharge infectious granules and cells due to a diarrheic effect until they die within a week post infection (Federici, 1997).

Infections of *Agrotis* species with isolates of Agrotis segetum granulovirus (AgseGV) have been well described since the late 1960s from field collected infected larvae (Zethner, 1980; Zethner et al., 1987). Since then, AgseGV has been considered as a candidate for the biological control of the so-called cutworm larvae of the turnip moth, *Agrotis segetum* and the ipsilon moth, *Agrotis ipsilon*, which are severe pests of numerous crops in the world (Dugdale, 1995). According to historic records from the Julius Kühn Institute (JKI), specimens of a European isolate of AgseGV, which we termed AgseGV-DA, have been distributed from Germany to several research institutes in Europe, including the Horticulture Research International (Warwick, UK) and the Laboratory of Virology (Wageningen, the Netherlands). All these laboratory stocks apparently derived from the same original material: infected *A. segetum* larvae collected in Austria in 1964 and sent to Federal Biological Research Institute (now the JKI) in Darmstadt, Germany. Chinese AgseGV isolates have also been identified: AgseGV-XJ with the first complete AgseGV genome sequence, and AgseGV-L1, an isolate from Shanghai (Zhang et al., 2014). According to recent reports, the turnip moth, *A. segetum*, is invasive even to isolated areas of Chinese islands and considered as a pest species to various crops (Guo et al., 2015). With this background, the interest in AgseGV as a biological control agent has increased during the last decade, particularly in China (Yang et al., 2012).

In the current study, whole genome sequencing with Solexa Illumina techniques was used to complement genome characterizations of AgseGV and to identify geographical differences between the two Chinese isolates AgseGV-XJ and AgseGV-L1 and the European isolate AgseGV-DA. As little is known about the pathogenesis of AgseGV in general, histopathological investigations were performed and infectivity parameters (LD_{50}) were determined to assert the infection mode of AgseGV-DA in larvae of the turnip moth. These findings will help to facilitate the use of this betabaculovirus in biological control, irrespectively of its presence in mixed infections with Agrotis segetum nucleopolyhedrovirus B as described previously (Wennmann et al., 2015b) or used as single agent as in China (Yang et al., 2012).

Materials and Methods

Insects

Mass-rearing of *A. segetum* was performed at the Institute for Biological Control (JKI) in Darmstadt as described previously (Wennmann and Jehle, 2014). In brief, neonate larvae were kept on semi-artificial diet (Ivaldi-Sender, 1974) at 22 °C with a 16/8 h light/dark photoperiod until they reached the fourth instar (L4). For pupation, groups of up to 30 individuals were then transferred to plastic boxes (18.3 cm × 13.6 cm × 6.4 cm) containing 3 cm thick layers of vermiculite (<0.5 mm grain size). Additional diet was provided shortly before pupation. Pupae were collected every second day and incubated at 25 °C until the adults hatched. Adult moths were kept for two weeks in groups of about 30 – 50 individuals in transparent plastic cylinders (20 cm diameter, 25 cm height) that were covered inside with rough surfaced paper. Eggs were collected three times a week by replacing paper tissues and incubation at 25 °C in moist boxes for several days until hatching.

AgseGV lineage and virus propagation

A virus stock of AgseGV was provided by Doreen Winstanley, Horticulture Research International (HRI) collection, Warwick (UK). This virus originated from 112 infected *A. segetum* larvae that were collected in Vienna (Austria) by Otto Muhr in 1964. Occlusion bodies (OB) of this AgseGV isolate (termed AgseGV-DA) had been purified and sent to the Institute for Biological Control in Darmstadt in the late 1960s, from where it was handed over to the HRI by Ole Zethner in 1974. For the experiments presented in this study, virus was propagated with late third instars to early fourth instars of *A. segetum* fed with small pieces of artificial diet (8 mm^3) that were overlaid with 10^6 OB of AgseGV-DA. Larvae that had consumed the entire diet cube within 12 h were transferred to normal diet and incubated individually under standard rearing conditions. Dead larvae were collected on a daily basis and stored at -20 °C for OB purification. Frozen larvae were homogenized in 0.5% sodium dodecyl sulfate (SDS) and treated with an ultra-sonic pulse before being filtered through a sandwich-filter consisting of layers of gaze with mesh cotton. The filtrate was repeatedly washed with water and centrifuged at 12,000 *g* for 15 min until the pellet had a whitish/light grey appearance. The pellet was resuspended in water and 2 ml of this suspension were loaded on a 55% - 80% glycerol gradient [gradient steps 80% / 70% / 60% / 55% (v/v) glycerol/water] and centrifuged at 4,000 *g* at 12 °C in a swing-out rotor for 45 min to separate OB from cadaver debris. The OB band between 60% and 70% glycerol was recovered with a sterile plastic Pasteur pipette and washed twice in water by spinning down in an Eppendorf 5418 R tabletop centrifuge at 12,000 *g* for 2 min and resuspending the pellet in water. Finally, the OB were recovered in sterile water and stored at -20 °C. OB concentration was enumerated with a Petroff-Hauser hemocytometer (2.5×10^{-3} mm^2 × 0.02 mm depth) and dark-field microscopy (Leica DM RBE).

Determination of median lethal dose (LD_{50})

To determine the median lethal viral dose-mortality response (LD_{50}) of AgseGV-DA, full range bioassays were performed with late third instars to early fourth instars (L3-L4) of *A. segetum*. The larvae were starved overnight and subsequently fed with small cubic pieces of artificial diet (8 mm^3) overlaid with serial 1:10 dilutions of AgseGV-DA suspensions of 83 to 8.3×10^6 OB/cube. Each treatment consisted of triplicates of 25 – 30 tested larvae and 50 – 60 uninfected control larvae. Larvae that did not ingest the diet overnight were excluded from the experiment. The tested larvae were then transferred individually into 50 well boxes with normal diet and

mortality was scored 7 and 14 days post infection (dpi). Dose-response rates were estimated by probit analysis with ToxRat 3.2 software (ToxRat® Solutions).

Histopathological studies

To characterize the course of infection of AgseGV-DA in *A. segetum* larvae, selected L4 larvae from the full range bioassays were dissected for light microscopy, whereas L3 larvae of *A. segetum* fed with 2×10^3 OB of AgseGV-DA were used for transmission electron microscopy (TEM). In both cases, larvae were collected 7 dpi and anesthetized with ethyl acetate before embedding. For light microscopy, specimens were fixed in in Bouin´s Gendre solution, embedded in paraffin (Histosec®) and a series of 6 µm longitudinal sections were produced with a SK4 rotary microtome (Leitz, Wetzlar). The serial sections were stained following the Heidenhain´s iron hematoxylin technique (Eberle et al., 2012) and were evaluated on a DMRB light microscope (Leica Microsystems). For transmission electron microscopy, a modified protocol was used as described previously (Rose et al., 2013). For this purpose, fat-bodies and midguts of infected larvae were dissected and either fixed in 3% glutaraldehyde or in Karnovsky solution at 4 °C for 24 h. After post-fixation with 2% osmium tetroxide for 5 h, samples were washed three times for 30 min each with 2.5% sucrose solution in 0.1 M cacodylic acid and then stained twice with 1% uranyl acetate and 2% wolfram phosphoric acid in 70% ethanol for 30 min. Specimens were stepwise dehydrated by incubating in increasing concentrations of ethanol (70%, 80%, 96% and 100% v/v ethanol in water) for at least 1 hour per step, followed by embedding in Spurr solution (4.1 g/L 4-vinylcyclohexene dioxide, 1.43 g/L diglycidyl ether of polypropylene glycol, 5.9 g/L nonenyl succinic anhydrine, 0.1 g/L dimethylethanolamine). Ultra-thin sections of 50 – 100 µm were prepared using a LKB 8800A Ultratome III (L.K.B. Produkter Fabriks AB). After a final post-staining with 1% uranyl acetate and lead citrate, specimens were examined with a Zeiss 902 transmission electron microscope.

DNA extraction and whole genome sequencing

For purposes of whole genome sequencing, genomic DNA was isolated from AgseGV-DA OB as described previously (Wennmann and Jehle, 2014). The viral OB matrix was solubilized in 0.1 M Na_2CO_3 at 60 °C for 1 h. The suspension was adjusted to pH 8 by titrating with 1 M HCl, then treated with RNaseA (90 µg/ml) at 37 °C for 10 min and then with Proteinase K (250 µg/ml) and 1% SDS at 50 °C for a further 60 min. DNA was separated from protein debris by phenol/chloroform/ isoamylalcohol (25:24:1, v/v) extraction (O'Reilly et al., 1994) using Phase Lock Gel Tubes (5 PRIME) in order to avoid phenol/protein contamination. Finally, the viral DNA was ethanol precipitated (Sambrook, 2001). DNA concentration and purity were estimated by UV-VIS absorbance with a NanoDrop 2000c spectrophotometer. A total amount of 5 µg of ultra-pure genomic AgseGV-DA DNA was submitted for paired-end next generation sequencing by GATC Biotech plc (Konstanz, Germany). There, Solexa Illumina standard genomic libraries were produced and applied to Illumina HiSeq sequencing with up to 5×10^6 100 bp read pairs. FastQ-files were delivered and additionally quality-filtered using the tool shed of the JKI Galaxy server (Afgan et al., 2016b). Read pairs with 50% consecutive bases below an average Phred quality score of 30 per read cycle (99.9% base call accuracy) were excluded from the analysis (Gordon, 2009). The conducted read pairs were used in a *de novo* assembly using a CLC Assembly Cell on the Galaxy server. The genome sequence generated for the isolate AgseGV-DA has been deposited in GenBank (accession number KR58466). In order to detect possible mixtures of genetic variations, including single nucleotide

polymorphisms (SNPs), insertions and deletions (InDels), the read-pairs were re-mapped against the de-novo genome sequence of AgseGV-DA with Bowtie2 aligner for short reads using standard parameters for very sensitive local alignment (Langmead and Salzberg, 2012b). Consecutively, the presence of SNPs and InDels < 10 bp was searched with SAMtools Mpileup. Regions of low coverage in the alignment were used to localize deletions that expanded more than 10 bp in the genome.

Open reading frame and homologous repeat (*hr*) region annotation

Open reading frames (ORFs) were annotated using criteria previously described (Wennmann et al., 2015a): they should not overlap by more than 100 bp and were >50 codons in length. The putative amino acid sequences from *in silico* predictions (standard codon usage) were used to identify viral and non-viral ORFs, as ascertained by PSI-BLAST against the RefSeq protein database from all organisms deposited in GenBank (http://blast.ncbi.nlm.nih.gov/Blast.cgi). ORFs that were identified within annotated baculovirus genes were only annotated if they had significant homolog matches (with e-values > 0.010) with other annotated baculovirus ORFs. In accordance with convention, the ORF coding for granulin was designated as the first ORF and the start codon adenine of the granulin ORF was designated as first nucleotide (bp) position of the genome. Homologous repeat (*hr*) regions were identified using the online tools REPuter (Kurtz et al., 2001) and the Tandem Repeat Finder (Benson, 1999). Additionally, inverted repeat sequences were identified with the EMBOSS palindrome tool (http://emboss.bioinformatics.nl/cgi-bin/emboss/palindrome). In cases of palindromes, secondary structures were predicted using the Vienna RNA secondary structure server (Hofacker, 2003). All annotated ORFs and repeat regions were compared to the genome sequences of AgseGV-XJ and AgseGV-L1 using the Mauve whole genome aligner of Geneious R10 (Darling et al., 2004).

Molecular phylogeny and sequence comparison

The relationship of AgseGV-DA, AgseGV-XJ and AgseGV-L1 to other baculoviruses was inferred by phylogenetic analysis based on the concatenated amino acid alignments of 37 putative baculovirus core genes (Garavaglia et al., 2012) from the genomes of 23 betabaculoviruses, 4 alphabaculoviruses, and a representative of gammabaculovirus and deltabaculovirus, respectively (Table S1). In order to avoid undesirable bias in the amino acid sequence alignments (Harrison et al., 2016), all core genes were identified and confirmed for their presence by PSI-BLAST and BLASTp. Thus, inferred amino acid sequences were aligned by ClustalW using Geneious R10 with default parameters. Pairwise genetic distances to each other were computed by the number of substitutions per site using the Kimura-2-parameter model (K2P) for base substitutions with MEGA7 (Kimura, 1980; Kumar et al., 2016).

Results and Discussion

Determination of dose-mortality response and histopathological studies of AgseGV-DA infections in *A. segetum* larvae

The dose-mortality response of L3/L4 larvae per orally infected with a dilution series of AgseGV-DA did not achieve higher mortality than 61.4% at 7 dpi at the highest applied virus concentration of 8.3×10^9 OB/ml. Based on these data, LD_{50} was calculate to be 6.29×10^6 OB/larva (Table 1) but the confidence limits of the cumulated mortality were broadened in the fit curve for higher OB doses. This effect also influenced a solid estimation of the LD_{90} at 7 dpi, resulting in overlapping 95% confidence limits with the LD_{50} observed at the same time

(Table). At 14 dpi, almost 100% mortality was reached with the applied concentration of 8.3×10^9 OB/ml. The resulting dose-mortality response showed homogeneity of variance of the applied doses with defined LD_{10}, LD_{50} and LD_{90} and divisible 95 % confidence limits (Table). As expected, more larvae died within two weeks and compared to the LD_{50} of 7 dpi, the LD_{50} decreased by a factor of 80 to 2.88×10^4 OB/larva at 14 dpi. The LT_{50} of infections was not calculated, as the calculated LD_{80} for 14 dpi exceeded the highest possible dose for application. However, the majority of the larvae died in their last larval stage, indicating a key characteristic of slow-killing granuloviruses.

Table 1 Computed dose-mortality responses of AgseGV-DA in L3/L4 larvae of *Agrotis segetum*. Computed doses shown: LD_{10}, LD_{50} and LD_{90} at 7 and 14 days post infection (dpi) with their corresponding 95% confidence limits (CL), the slope of the probit line and the goodness of the probit fit (χ^2). Mortality rates of 3 independent replicates were corrected for the control mortality rates (2.1% at 7 dpi and 2.8% at 14 dpi) according to Abbott (Abbott, 1925).

dpi	n	LD_{10} (95% CL)[a] OB/larva	LD_{50} (95% CL)[a] OB/larva	LD_{90} (95% CL)[a] OB/larva	Slope	χ^2	d(f)
7	455	2.02×10^3 ($12.5 – 1.63x10^4$)	6.00×10^6 ($7.14x10^5 – 1.15x10^9$)	1.78×10^{10} ($2.11x10^8 – .57x10^{16}$)	0.37	9.34	4
14		1.61×10^2 ($3.24 – 1.06x10^3$)	2.88×10^4 ($6.54x10^3 – 6.33x10^5$)	5.16×10^7 ($7.41x10^5 – 2.56x10^9$)	0.57	12.79	4

[a] CL (confidence limits) calculated by Fieller´s Theorem; d(f) degrees of freedom

The early stage of infections with AgseGV-DA in *A. segetum* larvae was followed by histopathological analysis of infected larvae at 7 dpi. In contrast to uninfected larvae, infected larvae showed hypertrophic fat body at low doses of AgseGV-DA as shown for 8.3×10^4 OB/larva. The correlation between infection and disintegrating fat body tissues was shown by the increasing number of OB applied and subsequently swelling of fat body, which culminated in disintegrated and highly hypertrophic fat body cells with the highest doses applied. The formation of OB was restricted only to the fat body and absent in other tissues like e.g. epidermis and hemocytes in light microscopy. This observation was confirmed by transmission electron microscopy (TEM) of fat body and midgut tissues (Figure 1). Fat body ultra-thin sections revealed formation of nucleocapsids and their occlusion (Figure 1 A), and foci of virogenic stroma (Xeros, 1956) (Figure 1 B). Around the virogenic stroma, large numbers of OB accumulated (Figure 1 B and C). Higher resolution confirmed the AgseGV-DA structures in detail, such as individually enveloped nucleocapsids occluded in the ovo-cylindrical granulin OB matrix, of approximately 450 × 150-200 nm (Figure 1 C & D). AgseGV-DA OBs were absent in midgut ultra-thin sections with few exceptions of migrated OBs at 7 dpi. These sections also showed intact tracheae and microvilli with no signs of infection. The absence of OBs in all tissues other than fat body indicated that virus infection, with high production of progeny OB, was restricted to the fat body at 7 dpi. Visual evaluation of the larvae of the bioassays, that were also the basis of this histopathological study, revealed that the infected larvae did not necessarily die in the infected instar but some infected larvae were able to develop to L4/L5 instars before dying from virus infection. Similar to the slow-killing granulovirus AdorGV (Wormleaton and Winstanley, 2001), AgseGV-DA revealed that the low mortality in the first week post infection, the tissue tropism of infection restrained to the fat body, and the

rather unaffected development of larvae until death are three key characteristics for type I granuloviruses (Federici, 1997).

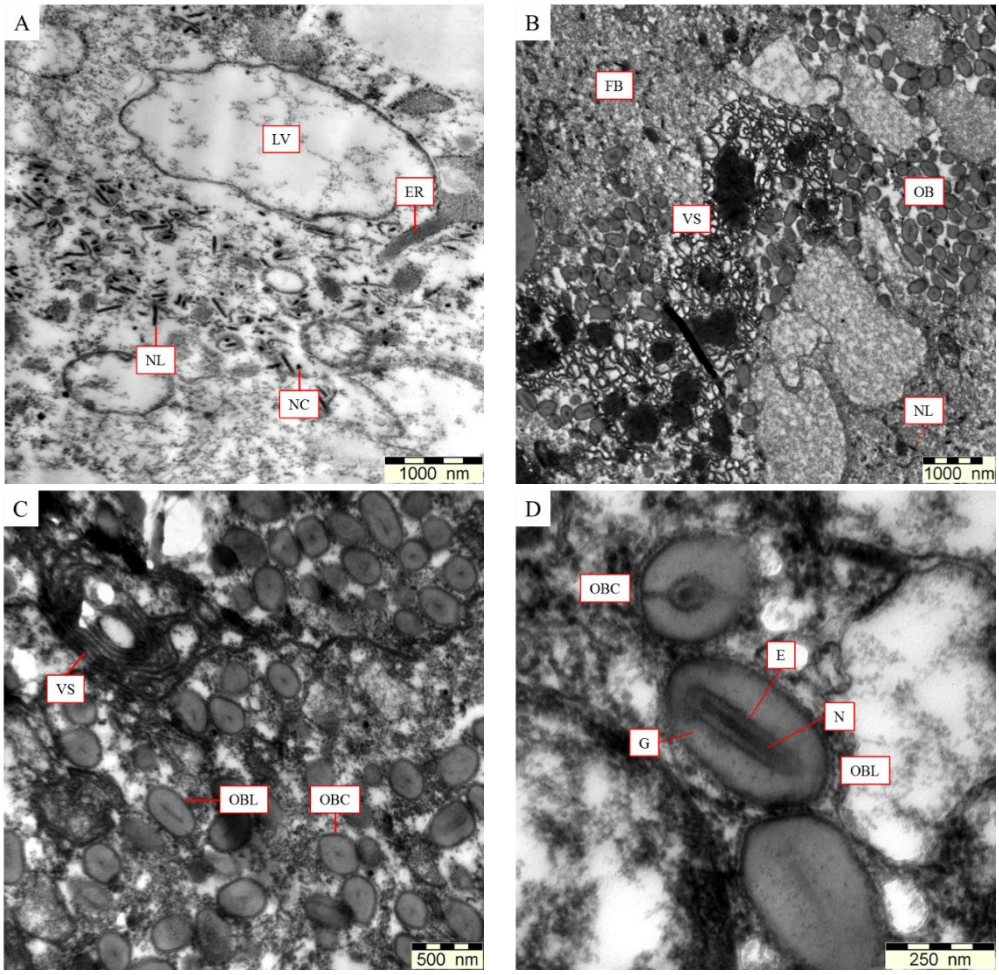

Figure 1 Transmission electron-micrographs of ultra-thin sections of the fat body of infected larvae at 7 dpi. (A) Nucleocapsid formation with longitudinal sections and cross-sections of nucleocapsids in cells of fat body; (B) large numbers of fully developed occlusion bodies (OB) and virogenic stroma; (C) longitudinal and cross-sections of OBs around virogenic stroma, and (D) cross-sections and longitudinal sections of AgseGV OB showing the granulin matrix, ODV envelope and nucleocapsid (D). (LV = lipid vesicle, ER = endoplasmic reticulum, NL = nucleocapsid longitudinal, NC = nucleocapsid cross-section, VS = virogenic stroma, OBL = occlusion body longitudinal section, OBC = occlusion body transversal section, G = granulin matrix, E = ODV envelope, N = nucleocapsid). EM photographs provided by M. Eigenbrod.

Genome composition and molecular phylogeny of AgseGV-DA

Solexa Illumina HiSeq sequencing resulted in 5,532,755 read pairs with an average Phred quality score above 30. The *de novo* assembly of the reads with the CLC assembly cell yielded five large contigs, the largest being 131,557 bp in length with a GC-content of 37.3%. This contig shared 99.8% and 99.6% sequence identity with AgseGV-XJ and AgseGV-L1, respectively. The quality filtered reads were re-mapped against the AgseGV-DA sequence using the Bowtie2 local aligner with default parameters for highest sensitivity. A total number of 5,439,487 (98.3% of all) read pairs mapped against the AgseGV-DA sequence yielded an 8,323-fold average genome coverage. Neither single nucleotide polymorphisms (SNPs), nor small insertion/deletion (InDel) mutations were detected in the alignment with MPileup, indicating that the AgseGV-DA isolate consisted of a single and highly pure genotype. The annotated sequence revealed that AgseGV-DA encodes 149 putative open reading frames

(ORFs), of which 72 are orientated clock-wise and 77 anti-clock-wise relative to the granulin ORF (Table S2). All 37 baculovirus core genes (Garavaglia et al., 2012), 18 granulovirus conserved genes, 72 genes with orthologues in other baculoviruses, and 21 putative ORFs that are unique in AgseGV-DA, AgseGV-XJ and AgseGV-L1 were identified (Figure 2 and Table 4). With a few exceptions, the genome of AgseGV-DA can be physically divided into two main clusters of ORFs. One cluster spanning almost 40% of the circular genome from position 57,301 to 104,845 nt harbors most of the baculovirus core genes. The second cluster includes less conserved genes, such as the betabaculovirus conserved genes and orthologues, and is interspersed by 19 of 21 unique AgseGV-DA genes. Similar to AgseGV-L1 (Zhang et al., 2014), one *bro* gene was found in AgseGV-DA. A PSI-BLAST search of the putative amino acid (aa) sequence of the *bro* gene showed, that it shared 34% aa identity with Mamestra brassicae nucleopolyhedrovirus BRO-D and 30% aa identity with Mamestra configurata nucleopolyhedrobvirus BRO-E. The presence of 36 baculovirus core genes was confirmed by PSI-BLAST searches against all deposited baculovirus protein RefSeq. The baculovirus core gene *p6.9* was identified by comparison of the three AgseGV *p6.9* homologues. The BLASTp analysis showed that the three isolates AgseGV-DA, AgseGV -XJ and AgseGV -L1 shared an identical P6.9. Given the standard parameters for BLASTp and PSI-BLAST searches, this particular P6.9 did not show high homology to other baculovirus proteins. A comparison of a visually adjusted ClustalW alignment with relaxed penalties for gap-opening (5) and gap-extension (0.1), however, confirmed the presence of this baculovirus core gene in AgseGV-DA (Figure 3). Nine conserved amino acids, present in all baculovirus P6.9, were also identified in the respective positions in AgseGV-DA P6.9. AgseGV-DA P6.9 further shares the highest amino acid identities with P6.9 sequences of the clade “a” granuloviruses HearGV, MoGV, PsunGV, XecnGV and SpfrGV (Garavaglia et al., 2012).

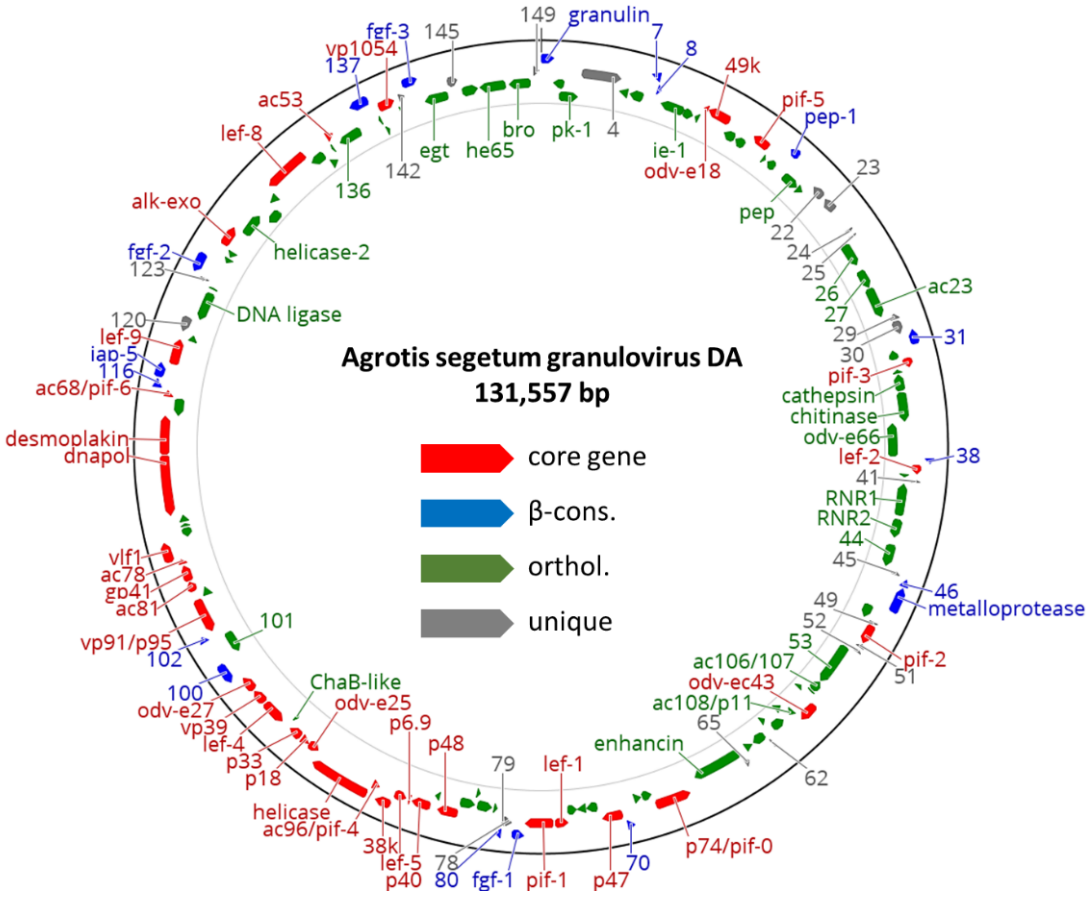

Figure 2 Map of the open reading frames (ORFs) of AgseGV-DA. ORF arrows indicate the ORF direction and position relative to *granulin*. Core gene = baculovirus core gene (red), β-cons. = conserved gene in *Betabaculovirus* (blue), orthol. = gene with orthologues in *Baculoviridae* (green), unique = unique gene in *Agrotis segetum granulovirus* (grey).

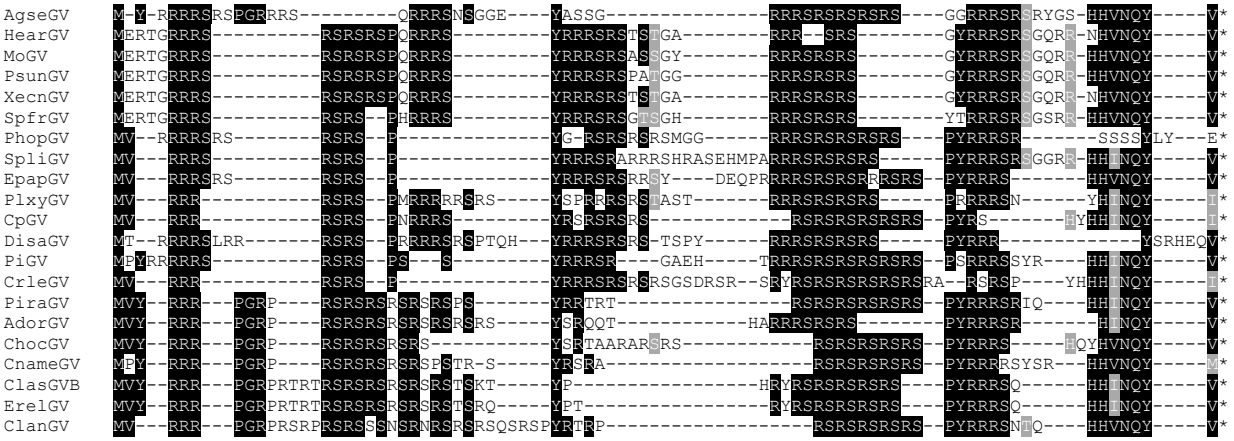

Figure 3 Visually adjusted ClustalW alignment of putative p6.9 amino acid sequences of 23 *Betabaculovirus* species used in the phylogenetic analysis (Table S1). Fully conserved amino acids are shaded in black, grey indicates 80% identity.

The analysis of the molecular phylogeny of AgseGV-DA with representative isolates of other baculovirus species (Table 3) was based on the concatenated alignments of the predicted amino acid sequences of 37 core genes, produced with ClustalW and default parameters. The minimum evolution phylogenetic tree (Rzhetsky and Nei, 1992) is given in Figure 4. All three isolates, AgseGV-DA, AgseGV-XJ and AgseGV-L1, form a unique branch within the betabaculovirus clade "a", which clusters betabaculovirus species infecting larvae of the

Lepidopteran family Noctuidae. Within this clade, they share PlxyGV as their nearest neighbor (Figure 4). Strikingly, PlxyGV showed a closer relationship to AgseGV-DA, AgseGV-XJ and AgseGV-L1 in the phylogenetic tree based on the amino acid alignments of 37 core genes than given by the substitution rates of nucleotides based on K2P-distances of partial *granulin*, *lef-8* and *lef-9* sequences, which are used as species demarcation criterion (Table 2). The classification into slow-killing type I granuloviruses and fast-killing type II granuloviruses is, however, not reflected by the molecular phylogeny of granuloviruses with the two clades "a" and "b", since both types of granuloviruses are represented in the two clades.

Table 2 Estimates of evolutionary divergence by pairwise distances of concatenated partial *granulin*, *lef-8* and *lef-9*. The number of base substitutions per site between betabaculovirus isolates representing their species are shown. Distances of the three Agrotis segetum granulovirus isolates are highlighted. Analyses were conducted using the Kimura 2-parameter (K2P) model with 1000 bootstrap replicates.

	AgseGV-DA	AgseGV-L1	AgseGV-XJ
AgseGV-L1	**0.0000**		
AgseGV-XJ	**0.0014**	**0.0014**	
PlxyGV	0.3932	0.3955	0.3932
XecnGV	0.3972	0.3970	0.3972
CpGV	0.3553	0.3575	0.3575

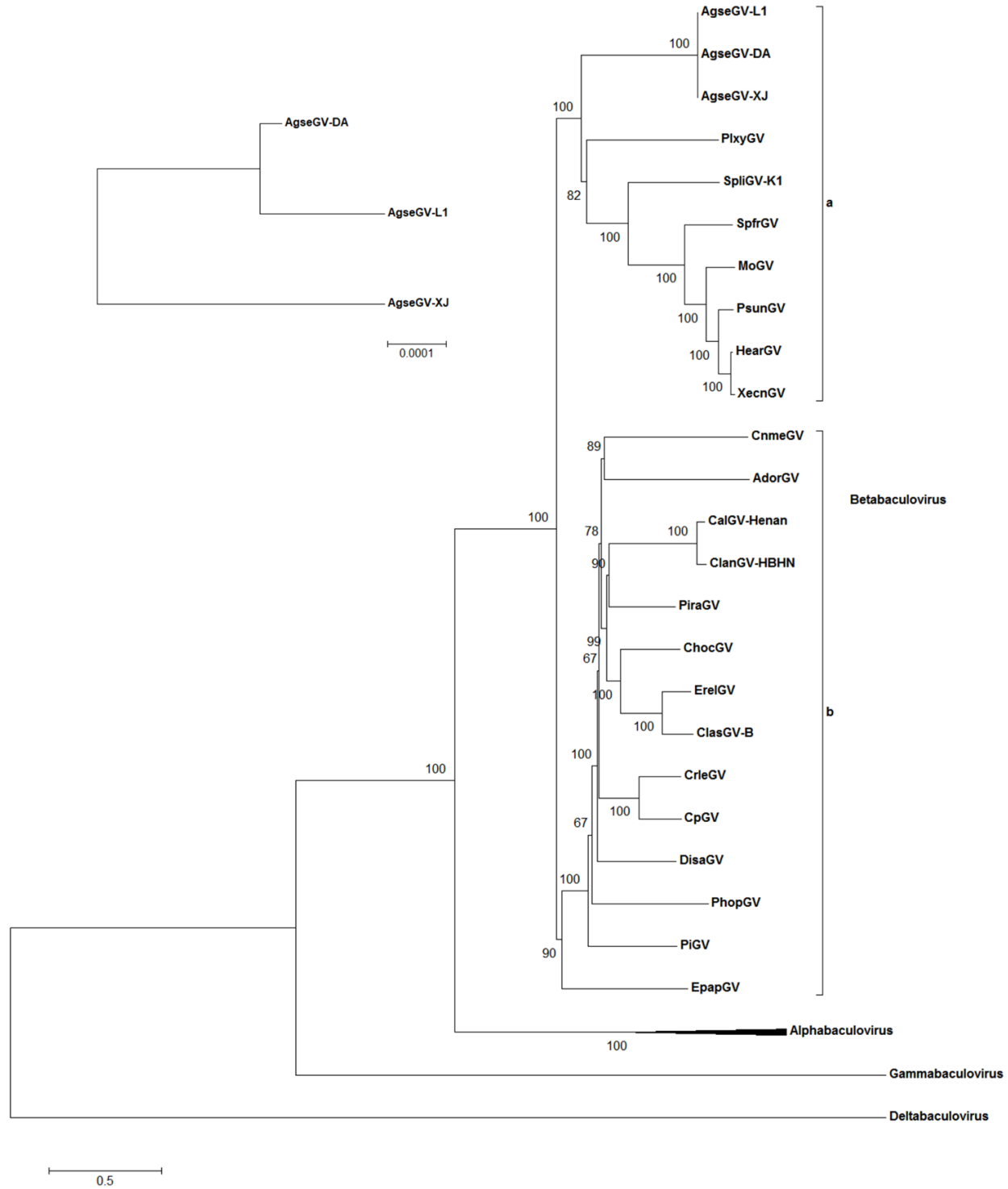

Figure 4 Minimum evolution (ME) phylogenetic tree of AgseGV-DA and representative isolates of other baculoviruses (see Table S1) inferred from concatenated predicted amino acid sequences of 37 baculovirus core genes (standard genetic code). Close-Neighbor-Interchange (CNI) algorithm and Neighbor-joining (NJ) algorithm as initial tree were used. Ambiguous positions were removed for each sequence pair. There was a total of 19,904 positions in the final dataset. The optimal tree with the sum of branch lengths = 19.37 is shown. Taxon genera are indicated behind corresponding branches. Both betabaculovirus clades "a" and "b" are indicated in brackets. The percentage of replicates in bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths as evolutionary distances (Dayhoff matrix). Insert top left: Midpoint rooted ME subtree of AgseGV-DA, AgseGV-L1 and AgseGV-XJ based on whole genome nucleotide alignment.

Pairwise distances were computed for the concatenated alignments of the partial *granulin*, *lef-8* and *lef-9* sequences using the Kimura-2-parameter (K2P) model (Kimura, 1980) and 1000 bootstrap replicates. According to the propositions of Jehle et al. (Jehle et al., 2006b), two baculoviruses were recognized to belong to different species when the K2P distances of these loci were larger than 0.05 substitutions/site. Baculoviruses with K2P distances smaller than 0.015 substitutions/site were proposed to belong to the same species. All three isolates AgseGV-DA, AgseGV-XJ and AgseGV-L1 showed pairwise K2P distances of maximum 0.014 to each other, indicating that they should be considered as different isolates of the same baculovirus species. In contrast, pairwise distances to all other baculovirus species were above 0.05 as shown, for example, in a subset of the pairwise distances of AgseGV-DA, AgseGV-XJ, AgseGV-L1, PlxyGV and XecnGV as representatives of the betabaculovirus clade "a" and CpGV from the clade "b" (Table 2). It is therefore suggested that all three AgseGV-DA, -XJ, -L1 are isolates of a single species, termed *Agrotis segetum granulovirus,* which was proposed to and accepted by the International Committee on Taxonomy of Viruses (ICTV) (Adams et al., 2016; Gueli Alletti et al., 2015). The sequence of AgseGV-DA sequence was chosen as representative of the *Agrotis segetum granulovirus* type species sequence for the following reasons. Although the complete genome sequence of AgseGV-XJ became available almost a decade ago, this sequence lacked 17 ORFs that were not annotated but identified in AgseGV-L1 and AgseGV-DA. Also, numerous ORFs in the AgseGV-XJ sequence were not originally annotated as baculovirus core genes. Further, studies on the genome sequence of AgseGV-L1 (Zhang et al., 2014) did not include biological data such as granule structure or investigations on the pathology. In addition, characteristics such as the isolation of AgseGV-DA from infected larval material, the host range, and infections of larvae were previously described (Zethner, 1980).

AgseGV Enhancins and taxonomic representation

A major sequence difference among the three isolates AgseGV-DA, AgseGV-XJ, AgseGV-L1 was identified in the putative AgseGV *enhancin* gene. Baculoviruses enhancins alternatively termed viral enhancing factors (VEF), have accounted for a synergistic effect between two heterologous baculoviruses, as described for co-infections of the armyworm *Pseudaletia unipuncta* in co-infections with Pseudaletia unipuncta granulovirus and Pseudaletia unipuncta nucleopolyhedrovirus (Tanada and Hukuhara, 1971).They play an active role in the degradation of the peritrophic membrane (PM) (Lepore et al., 1996b). This mucin-rich intestinal layer functions as a primary barrier between midgut lumen and midgut epithelium that has to be penetrated by occlusion derived virions (ODVs) originating from dissolved OB to initiate primary infection of midgut epithelium cells. A possible mode of enhancin function derives from a metalloprotease-like zinc-binding domain HEXXH that is found in many baculovirus Enhancins (Bischoff and Slavicek, 1995; Rawlings and Barret, 1995; Toprak et al., 2012) and also some *Bacillus* bacteria. However, this particular HEXXH motif is missing in the putative Enhancins of AgseGV-XJ and AgseGV-L1 (Wennmann et al., 2015a). All three Enhancins from AgseGV-DA, AgseGV-XJ and AgseGV-L1 were aligned based on their predicted amino acid sequence using ClustalW with default parameters. They share a stretch of 865 conserved amino acids at their N-termini, which is followed by 28 identical amino acid positions in AgseGV-XJ and AgseGV-L1. In one region of the *enhancin* sequence, namely between the genome positions 56,401 and 56,725, a statistically significant decrease based on the mean coverage and its standard deviation (s.d.) at each position, from an average coverage

of 8323-fold (s.d. = 620, n = 131,557) to only 4,851-fold coverage (s.d. = 1767, n = 324) was observed (t-test, p = 0.0001). This particular decline indicated some deletions or false alignments of the short sequencing reads in this particular region. A thorough analysis of the sequence assembly

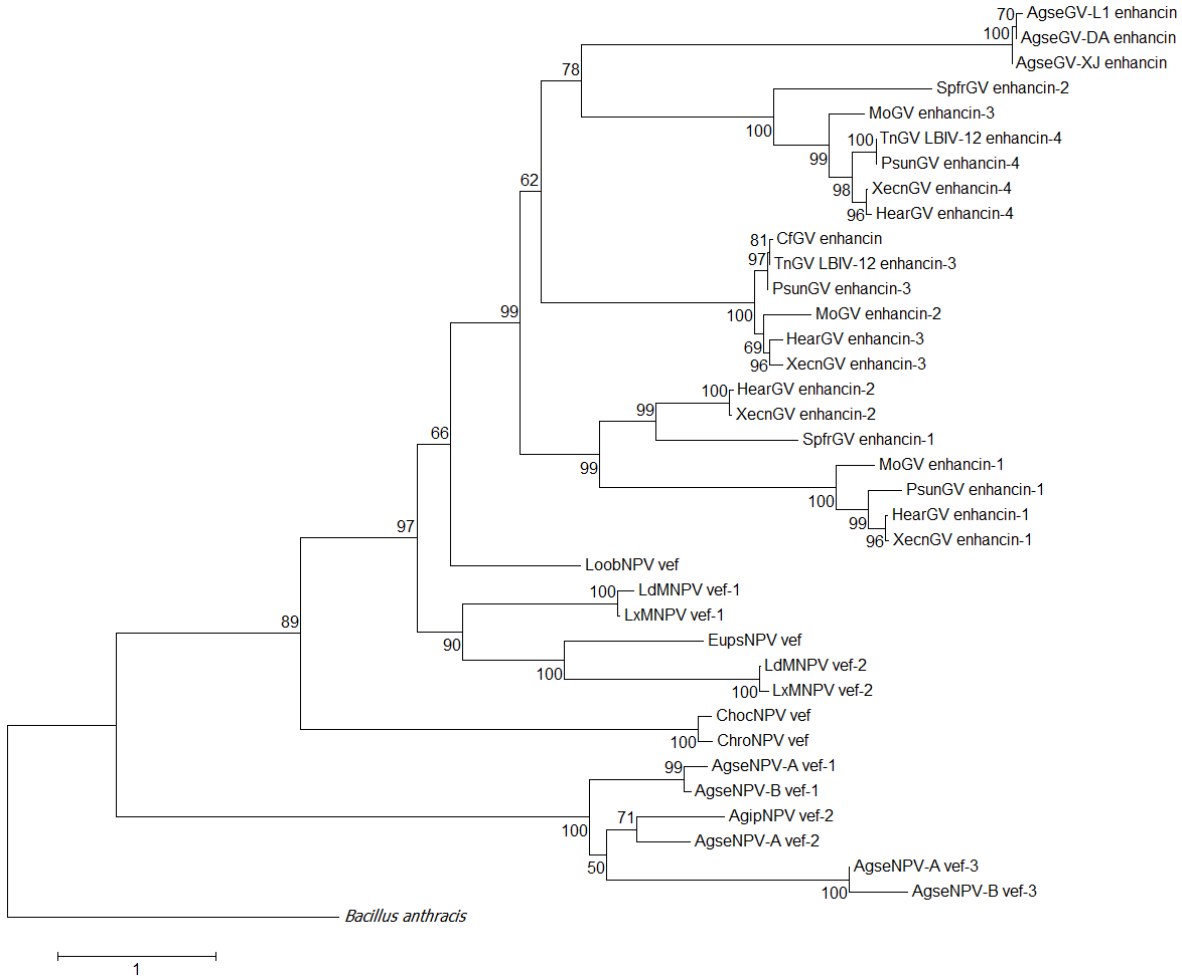

Figure 5 Minimum Evolution (ME) tree of baculovirus Enhancins, inferred from the alignment of Enhancin amino acid sequences by using CNI with NJ as initial tree. Bootstrap percentages (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths as evolutionary distances (Dayhoff matrix). The bacterial Enhancin sequence of *B. anthracis* Amens (Acc-N°. AE017034) was used as outgroup

files covering this region revealed numerous repeat sequences, which were eventually resolved by PCR amplification and Sanger sequencing of the amplicons. This repeat region in the AgseGV-DA *enhancin* consists of 12 iterative repeats of a 24-mer 5´-ACTCCTGAACCTWCTCCYCCACCT-´3. Strikingly, this repeat is also present as 11-fold repeat in AgseGV-XJ but is completely absent in AgseGV-L1. The mentioned repeats in AgseGV-DA and AgseGV-XJ are responsible for a 12- and 11-fold repetition of a predicted amino acid motif TPEPTPPP and lead to a putative enrichment of the amino acid proline at the C-termini of the two Enhancins. Because of these repetitions, the putative Enhancins of AgseGV-DA and AgseGV-XJ share 96% aa identity, whereas the aa identity of the AgseGV-DA and AgseGV-L1 Enhancins is only 85%. The amino acid consensus sequence of the AgseGV Enhancin alignment was used in PSI-BLAST searches in order to identify 37 Enhancin

homologues from different species. These sequences also included a predicted bacterial Enhancin from *Bacillus anthracis* Amens (GenBank Acc.-N° AE017034), the Viral Enhancin Factors (VEF) of Agrotis segetum nucleopolyhedrovirus A (AgseNPV-A), Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) and Agrotis ipsilon nucleopolyhedrovirus (AgipNPV). All 37 identified homologues were aligned to the Enhancins of AgseGV-DA, AgseGV-L1 and AgseGV-XJ using ClustalW. This alignment was used to infer a Minimum Evolution (ME) phylogenetic tree using the *B. anthracis* Enhancin as an outgroup (Figure 5). The ME tree could be divided into a single clade of Enhancins from the genus *Betabaculovirus* and a number of paralogues from the genus *Alphabaculovirus*. The closest alphabaculovirus Enhancin neighbor was represented by the VEF from Lonomia obliqua nucleopolyhedrovirus (LoobNPV). The Enhancins from the three AgseGV isolates showed a high genetic distance to those of all other granuloviruses as well as to the VEFs from AgseNPV-A, AgseNPV-B and AgipNPV, emphasizing the absence of a horizontal gene transfer of the *enhancins* between *Agrotis* NPVs and the AgseGVs (Wennmann et al., 2015a). Most of the nucleotide variations between AgseGV-DA, AgseGV-XJ and AgseGV-L1 were located in their distinct *enhancin* genes but were little in the remaining parts of the genomes. The very opposite of genetic divergence is reflected by this relation; although there is geographic distance between the European AgseGV-DA and the Chinese AgseGV-XJ and AgseGV-L1, all three viruses share evolutionarily highly stable and similar genomes. Thus, AgseGV isolates appear to be considerably homogenous compared to other granuloviruses, e.g. geographic isolates of CpGV, for which differences in the ORF content and in their virulence to different codling moth populations were reported (Eberle et al., 2009; Gebhardt et al., 2014; Rezapanah et al., 2008). The highly-conserved genome sequences in the three AgseGV isolates presumably indicate a similar pathology in larvae of *A. segetum*. Thus, our findings on their close phylogenetic relationship may facilitate future efforts in developing a biological control agent based on AgseGV, as the high sequence similarity suggests a similar pathogenicity of the three isolates.

Intergenic repeats, palindromic sequences and putative origins of replication

Analysis of repeats and palindromes in AgseGV-DA also identified the presence of four direct repeats and one palindromic sequence (Figure 6 and Figure 7). Two of these direct repeats have already been reported for AgseGV-L1 (Figure 6) as homologous regions (hrs) (Zhang et al., 2014) and are present in AgseGV-DA within the genome positions 19,895 – 19,943 (*hr1*) the intergenic region between the ORF22 and ORF23, and at the genome positions 85,045 – 85,108 (*hr2*) in the intergenic region between odve27 and ORF100. The third repetition unit of *hr2* overlaps with the 3´end of ORF100 in five nucleotides. Another short two-fold direct repeat of 24 identical nucleotides, namely dr3a/b was found in the intergenic region between pp34 and ORF21 at the genome positions 16,862 – 16,885 (Figure 7). Further direct repeats, dr4a/b/c, consisting of up to 52 bp long repetitions were located in the same intergenic region as *hr2*, between odve27 and ORF100.

One inverted repeat, Pal-1, was found in the intergenic region of ORF23/ORF24. The inverted repeats were 42 nt in length and inclosed a non-coding 444-bp region containing three direct repeats dr5a/b/c of 5'-YAGATGGCGY-3' in its center (Figure 8). Pal-1was not confined by direct repeats, which would have been a signal for a possible transposition of the whole element. Blastn searches for whole region with both palindromic ends did not reveal any significant hits with an E-value < 0.01, except for homologous sequences present in AgseGV-XJ and AgseGV-

L1. The function of such palindromic sequences in baculoviruses is unknown, though similar structures are present in numerous non-coding RNAs, such as tRNA cloverleaf structure. Blastn searches did not identify any hit for non-coding RNA, nor did the structure resemble any palindromic sequences reported for other baculoviruses (Hilton and Winstanley, 2008b). However, similar structures with inverted repeats around a region with a duplication of restriction endonuclease target sides have been reported for the Cydia pomonella granulovirus isolates CpGV-I01 and CpGV-E2 (Eberle et al., 2009). Homologous regions (*hrs*) of many alphabaculoviruses consist of short palindromic sequences which are present in multiple locations within the genome (Rohrmann, 2013). They were characterized as enhancers as well as origins of replication (Kool et al., 1993; Pearson and Rohrmann, 1995). Although the *hrs* of betabaculoviruses are more variable and often lack palindromes in contrast to *hrs* from NPVs (Hilton and Winstanley, 2008b), they are also present in several positions within the genome. The short direct repeats observed in AgseGVs, however, do not exhibit any homology to alphabaculovirus *ori*s. These findings are in accordance with previous investigations of granulovirus *hr*-like origins of replication (Harrison et al., 2016; Hilton and Winstanley, 2008b) that did not discover any homologous region in analysis of the AgseGV genomes submitted to GenBank.

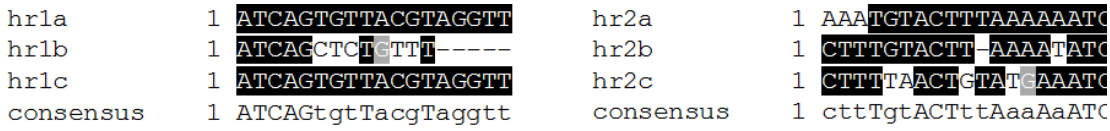

Figure 6 ClustalW alignment of the direct repeats hr1 and hr2 in the genome of AgseGV-DA. Positions of at least 60% shared identity are shaded black, ambiguities are shaded grey.

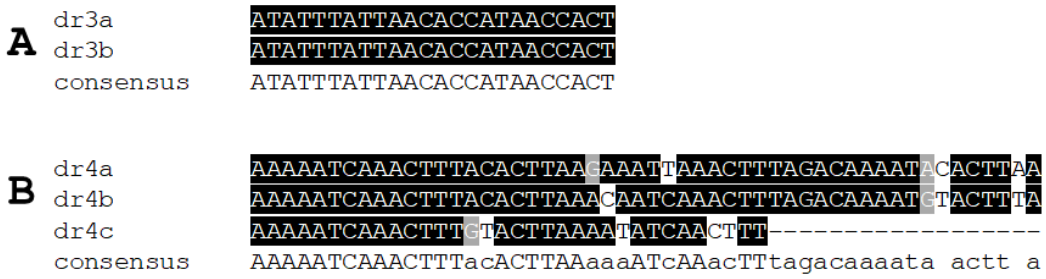

Figure 7 ClustalW alignment of two direct repeats dr3 (A) and dr4 (B) conducted by genome analysis with REPuter and Tandem Repeat Finder. Positions of at least 60% shared identity are shaded black, ambiguities are shaded grey.

Figure 8 Intergenic region (genome nucleotide position 19,274 – 19,800 in AgseGV-DA) with inverted repeat sequences Pal-1a and Pal-1b at each end of 444 bp and the threefold direct repeats of 5'-YAGATGGCGY-3' (dr5a, dr5b, dr5c) in its center.

Conclusion

Our pathological studies classify AgseGV-DA as a type I granulovirus based on the typical characteristics of slow infection rates in *A. segetum* larvae and a tissue tropism restricted to the fat body of infected caterpillars. Illumina genome sequencing of AgseGV-DA providing a more than 8000-fold coverage did not reveal any single nucleotide polymorphisms, insertion or deletions, defining AgseGV-DA as a genetically highly homogenous genotype. The genome of the European isolate AgseGV-DA shows a close relationship to the two Chinese isolates AgseGV-XJ and AgseGV-L1, as demonstrated by fully co-linear genomes, high nucleotide conservation and a small genetic distance. The three isolates are therefore considered as representatives of the same *Betabaculovirus* species *Agrotis segetum granulovirus.*

Table 3 Baculovirus isolates (n = 30) used to infer the molecular phylogeny based on concatenated alignments of 37 predicted amino acid sequences of baculovirus core genes

Organism/Isolate/Strain	Acc. N°	Abbreviation	Size (bp)	% GC
Autographa californica multiple nucleopolyhedrovirus	NC_001623	AcMNPV	133,894	40.7
Agrotis ipsilon multiple nucleopolyhedrovirus	NC_011345	AgipNPV	155,122	48.6
Agrotis segetum nucleopolyhedrovirus	NC_007921	AgseNPV-A	147,544	45.7
Agrotis segetum nucleopolyhedrovirus B	KM102981	AgseNPV-B	148,981	45.7
Adoxophyes orana granulovirus	NC_005038	AdorGV	99,657	34.5
Agrotis segetum granulovirus DA	KR584663	AgseGV-DA	131,557	37.3
Agrotis segetum granulovirus L1	KC994902	AgseGV-L1	131,442	37.3
Agrotis segetum granulovirus XJ	NC_005839	AgseGV-XJ	131,680	37.3
Clostera anastomosis granulovirus CaLGV-Henan	NC_022646	CaLGV-Henan	101,818	46.7
Choristoneura occidentalis granulovirus	NC_008168	ChocGV	104,710	32.7
Clostera anachoreta granulovirus ClanGV-HBHN	HQ116624	ClanGV-HBHN	101,487	44.4
Clostera anastomosis granulovirus ClasGV-B	KR091910	ClasGV-B	107,439	37.8
Cnaphalocrocis medinalis granulovirus	KP658210	CnmeGV	112,060	35.2
Cydia pomonella granulovirus	NC_002816	CpGV-M1	123,500	45.3
Cryptophlebia leucotreta granulovirus	NC_005068	CrleGV	110,907	32.4
Diatraea saccharalis granulovirus	KP296186	DisaGV	98,392	34.9
Epinotia aporema granulovirus	NC_018875	EpapGV	119,082	41.5
Erinnyis ello granulovirus	KJ406702	ErelGV	102,759	38.7
Helicoverpa armigera granulovirus	NC_010240	HearGV	169,794	40.8
Mocis sp. granulovirus	KR011718	MoGV	134,272	38.3
Phthorimaea operculella granulovirus	NC_004062	PhopGV	119,217	35.7
Plodia interpunctella granulovirus	KP864638	PiGV	112,536	44.2
Pieris rapae granulovirus	NC_013797	PiraGV	108,592	33.2
Plutella xylostella granulovirus	NC_002593	PlxyGV	100,999	40.7
Pseudaletia unipuncta granulovirus	NC_013772	PsunGV	176.677	39.8
Spodoptera frugiperda granulovirus isolate VG008	KM371112	SpfrGV	140,913	46.2
Spodoptera litura granulovirus	NC_009503	SpliGV-K1	124,121	38.8
Xestia c-nigrum granulovirus	NC_002331	XecnGV	178,733	40.7
Neodiprion lecontei nucleopolyhedrovirus	AY349019	NeleNPV	81,755	33.4
Culex nigripalpus nucleopolyhedrovirus	NC_003084	CuniNPV	108,252	50.9

Table 4 Annotation of Open reading frames (ORF) of AgseGV-DA, AgseGV-L1 and AgseGV-XJ as annotated (columns 1 and 2), the annotated ORF names (3), the conserved occurrence in other baculovirus genera (baculovirus core genes = α+β+γ+δ, conserved in *Alpha-*, *Beta-* and *Gammabaculovirus* = α+β+γ, conserved in *Alpha-*, *Beta-* and *Deltabaculovirus* = α+β+δ, conserved in *Alpha-* and *Betabaculovirus* = α+β, conserved *Betabaculovirus* genes = β, orthologues in other baculoviruses = bac. ortho., AgseGV-DA/L1 unique genes = unique) (4); the genome sequence position in AgseGV-DA and orientation relative to *granulin* (5); the amino acid (aa) number of predicted gene products (6), the next baculovirus homolog (7), regarding to PSI-BLAST/BLASTp with shared amino acid identity (%) (8)

(1) AgseGV-DA/-L1[1]	(2) AgseGV-XJ[1]	(3) Name	(4) Conservation	(5) Position	(6) aa	(7) Homolog	(8) % Identity
1	1	granulin	α+β+γ	1 → 747	249	ErelGV	90
2	2	p78/83	bac. ortho.	744 ← 1,412	223	ChocGV	50
3	3	pk-1	α+β	1,084 → 2,238	385	PlxyGV	63
4	4	4	unique	2,290 → 4,557	756	*none*	--
5	5	5	bac. ortho.	4,591 ← 5,100	170	AdhV	38
6	--	6	bac. ortho.	5,195 ← 6,007	271	AgipNPV	40
7	6	7	β	6,103 ← 6,657	185	PhopGV	51
8	7	8	β	6,647 → 6,910	88	PhopGV	40
9	8	ie-1	bac. ortho.	7,073 ← 8,464	464	XecnGV	41
10	9	ac146	α+β	8,484 → 9,086	201	DisaGV	40
11	10	ac145	α+β+γ	9,126 ← 9,425	100	SpliGV	58
12	11	odv-e18	α+β+γ+δ	9,438 ← 9,716	93	SpliGV	51
13	12	49k	α+β+γ+δ	9,717 ← 11,057	447	DisaGV	52
14	13	14	bac. ortho.	11,073 ← 11,798	242	SpliGV	27
15	14	15	bac. ortho.	11,839 ← 12,474	212	PhopGV	43
16	15	pif-5	α+β+γ+δ	12,599 ← 13,657	353	ChocGV	58
17	16	17	bac. ortho.	13,716 → 13,943	76	HearGV	42
18	17	dUTPase	bac. ortho.	14,026 ← 14,604	193	EpapGV	44
19	18	pep-1	β	14,622 ← 15,242	207	CrleGV	57
20	19	Baculo PEP N	bac. ortho.	15,299 → 16,351	351	ErelGV	60
21	20	pe/pp34	bac. ortho.	16,366 → 16,806	147	ClanGV	60
22	21	22	unique	16,960 ← 17,763	268	*none*	--
23	22	23	unique	17,896 ← 18,723	276	*none*	--
24	--	24	unique	20,030 → 20,245	72	*none*	--
25	--	25	unique	20,403 → 20,576	58	*none*	--
26	23	26	bac. ortho.	20,649 → 22,004	452	PiraGV	25
27	24	27	bac. ortho.	22,377 → 23,468	364	PlxyGV	27

Table 4 Annotation of Open reading frames (ORF) of AgseGV-DA, AgseGV-L1 and AgseGV-XJ as annotated (columns 1 and 2), the annotated ORF names (3), the conserved occurrence in other baculovirus genera (baculovirus core genes = α+β+γ+δ, conserved in *Alpha-*, *Beta-* and *Gammabaculovirus* = α+β+γ, conserved in *Alpha-*, *Beta-* and *Deltabaculovirus* = α+β+δ, conserved in *Alpha-* and *Betabaculovirus* = α+β, conserved *Betabaculovirus* genes = β, orthologues in other baculoviruses = bac. ortho., AgseGV-DA/L1 unique genes = unique) (4); the genome sequence position in AgseGV-DA and orientation relative to *granulin* (5); the amino acid (aa) number of predicted gene products (6), the next baculovirus homolog (7), regarding to PSI-BLAST/BLASTp with shared amino acid identity (%) (8)

28	25	ac23	α+β+δ	23,507 → 25,303	599	PiraGV	55
29	--	29	unique	25,342 ← 25,668	109	*none*	--
30	26	30	unique	25,803 → 26,525	241	*none*	--
31	27	31	β	26,488 ← 27,312	275	CpGV	32
32	28	32	bac. ortho.	27,324 ← 27,899	192	SpfrGV	52
33	29	pif-3	α+β+γ+δ	27,911 → 28,486	192	ErelGV	49
34	30	34	bac. ortho.	28,481 ← 28,720	80	CpGV	34
35	31	cathepsin	bac. ortho.	28,723 ← 29,706	328	PiraGV	62
36	32	chitinase	bac. ortho.	29,752 → 31,509	586	PiraGV	62
37	33	odv-e66	α+β	31,481 ← 33,517	679	OpMNPV	37
38	34	38	β	33,545 → 33,880	112	CpGV	49
39	35	lef-2	α+β+γ+δ	33,927 → 34,472	182	SpliGV	50
40	36	40	bac. ortho.	34,475 → 34,726	84	SpliGV	50
41	--	41	unique	34,775 ← 34,981	69	*none*	--
42	37	RNR1	bac. ortho.	35,074 ← 36,921	616	EpapGV	63
43	38	RNR2	bac. ortho.	37,117 → 38,226	370	EpapGV	58
44	39	44	bac. ortho.	38,639 → 39,907	423	AgipNPV	48
45	--	45	unique	39,974 ← 40,186	71	*none*	--
46	40	46	β	40,243 ← 40,671	143	ClanGV	37
47	41	metalloprotease	β	40,734 ← 42,179	482	TiniGV	40
48	42	p13	bac. ortho.	42,272 → 43,069	266	ChocGV	63
49	--	49	unique	43,075 → 43,260	62	*none*	--
50	43	pif-2	α+β+γ+δ	43,277 → 44,425	383	DisaGV	57
51	--	51	unique	44,456 ← 44,728	91	*none*	--
52	44	52	unique	44,827 ← 45,021	65	*none*	--
53	45	53	bac. ortho.	45,042 → 47,564	841	MoGV	32
54	46	ac106/107	α+β+γ	47,585 ← 48,244	220	SpliGV	70
55	--	ac110	α+β (+γ+δ)	48,265 → 48,417	51	DisaGV	61
56	47	ubiquitin	α+β	48,414 ← 48,707	98	SpliGV	86
57	48	odv-ec43	α+β+γ+δ	48,747 → 49,868	374	EpapGV	59

Table 4 Annotation of Open reading frames (ORF) of AgseGV-DA, AgseGV-L1 and AgseGV-XJ as annotated (columns 1 and 2), the annotated ORF names (3), the conserved occurrence in other baculovirus genera (baculovirus core genes = α+β+γ+δ, conserved in *Alpha-*, *Beta-* and *Gammabaculovirus* = α+β+γ, conserved in *Alpha-*, *Beta-* and *Deltabaculovirus* = α+β+δ, conserved in *Alpha-* and *Betabaculovirus* = α+β, conserved *Betabaculovirus* genes = β, orthologues in other baculoviruses = bac. ortho., AgseGV-DA/L1 unique genes = unique) (4); the genome sequence position in AgseGV-DA and orientation relative to *granulin* (5); the amino acid (aa) number of predicted gene products (6), the next baculovirus homolog (7), regarding to PSI-BLAST/BLASTp with shared amino acid identity (%) (8)

58	49	ac108/p11	α+β+γ	49,881 → 50,192	104	CpGV	58
59	50	59	bac. ortho.	50,168 ← 50,572	135	AdorGV	30
60	51	pp31	bac. ortho.	50,600 ← 51,424	275	DisaGV	45
61	52	lef-11	α+β+γ	51,408 ← 51,728	107	PsunGV	60
62	--	62	unique	51,834 ← 52,001	56	*none*	--
63	53	iap-3	bac. ortho.	52,032 → 52,841	270	AngeNPV	40
64	54	sod	bac. ortho.	52,893 ← 53,363	157	ClanGV	67
65	--	65	unique	53,529 ← 53,804	92	*none*	--
66	55	enhancin	bac. ortho.	53,826 → 56,768	981	PsunGV	27
67	56	p74/pif-0	α+β+γ+δ	57,301 ← 59,328	676	CpGV	50
68	57	acetyltransferase-like	bac. ortho.	59,382 ← 59,981	200	PhopGV	59
69	58	69	bac. ortho.	60,033 ← 60,440	136	MabrNPV	30
70	59	70	β	60,674 ← 61,141	156	CrleGV	64
71	60	p47	α+β+γ+δ	61,233 → 62,444	404	EpapGV	59
72	61	ac38	α+β	62,564 → 63,226	221	SpliGV	77
73	62	p24	α+β	63,242 → 63,775	178	EpapGV	60
74	63	38.7	bac. ortho.	63,791 ← 64,372	194	PlxyGV	48
75	64	lef-1	α+β+γ+δ	64,303 ← 65,064	254	ClanGV	58
76	65	pif-1	α+β+γ+δ	65,108 → 66,751	548	DisaGV	38
77	66	fgf-1	β	66,736 ← 67,413	226	SpliGV	37
78	--	78	unique	67,445 ← 67,843	133	*none*	--
79	--	79	unique	67,848 → 68,060	71	*none*	--
80	67	80	β	68,011 → 68,337	109	SpliGV	32
81	68	lef-6	α+β	68,334 ← 68,582	83	CpGV	44
82	69	dbp	α+β+γ	68,633 ← 69,523	297	EpapGV	35
83	70	83	bac. ortho.	69,411 ← 69,713	101	PiraGV	46
84	71	84	bac. ortho.	69,664 ← 70,542	293	EpapGV	24
85	72	p48	α+β+γ+δ	70,529 → 71,713	395	CpGV	62
86	73	p12	α+β	71,718 → 72,038	107	CpGV	53
87	74	p40	α+β+γ+δ	72,102 → 73,223	374	SpliGV	59

Table 4 Annotation of Open reading frames (ORF) of AgseGV-DA, AgseGV-L1 and AgseGV-XJ as annotated (columns 1 and 2), the annotated ORF names (3), the conserved occurrence in other baculovirus genera (baculovirus core genes = α+β+γ+δ, conserved in *Alpha-*, *Beta-* and *Gammabaculovirus* = α+β+γ, conserved in *Alpha-*, *Beta-* and *Deltabaculovirus* = α+β+δ, conserved in *Alpha-* and *Betabaculovirus* = α+β, conserved *Betabaculovirus* genes = β, orthologues in other baculoviruses = bac. ortho., AgseGV-DA/L1 unique genes = unique) (4); the genome sequence position in AgseGV-DA and orientation relative to *granulin* (5); the amino acid (aa) number of predicted gene products (6), the next baculovirus homolog (7), regarding to PSI-BLAST/BLASTp with shared amino acid identity (%) (8)

88	75	p6.9	α+β+γ+δ	73,256 → 73,441	62	*none*	--
89	76	lef-5	α+β+γ+δ	73,474 ← 74,229	252	SpliGV	55
90	77	38k	α+β+γ+δ	74,182 → 75,075	298	DisaGV	53
91	78	ac96/pif-4	α+β+γ+δ	75,078 ← 75,563	162	EpapGV	49
92	79	helicase	α+β+γ+δ	75,550 → 79,020	1157	MoGV	54
93	81	odv-e25	α+β+γ+δ	79,065 ← 79,718	218	PhopGV	69
94	82	p18	α+β+γ+δ	79,750 ← 80,253	168	EpapGV	51
95	83	p33	α+β+γ+δ	80,320 → 81,078	253	PiraGV	60
96	84	ChaB-like	bac. ortho.	81,075 ← 81,320	82	SpliGV	57
97	85	lef-4	α+β+γ+δ	81,657 ← 83,063	469	PiraGV	49
98	86	vp39	α+β+γ+δ	83,102 → 83,980	293	SpliGV	50
99	87	odv-e27	α+β+γ+δ	84,046 → 84,942	299	SpliGV	60
100	88	100	β	85,104 ← 86,249	382	EpapGV	34
101	89	101	bac. ortho.	86,262 ← 87,569	436	AdhV	25
102	90	102	β	87,656 → 88,030	125	CrleGV	67
103	91	vp91/p95	α+β+γ+δ	88,027 ← 89,997	657	ClanGV	37
104	93	tlp	α+β	89,966 → 90,499	178	SpfrGV	34
105	94	ac81	α+β+γ+δ	90,480 → 91,073	198	SpliGV	60
106	95	gp41	α+β+γ+δ	91,105 → 92,034	310	CpGV	64
107	96	ac78	α+β+γ+δ	92,103 → 92,420	106	XecnGV	33
108	97	vlf1	α+β+γ+δ	92,389 → 93,537	383	PhopGV	63
109	98	109	bac. ortho.	93,520 ← 94,068	183	SpliGV	41
110	99	110	bac. ortho.	94,119 → 94,376	86	CpGV	60
111	100	ac75	α+β+γ	94,417 → 94,863	149	DisaGV	53
112	101	dnapol	α+β+γ+δ	94,869 ← 98,306	1146	EpapGV	59
113	102	desmoplakin	α+β+γ+δ	98,308 → 100,515	736	SpfrGV	26
114	103	lef-3	α+β	100,512 ← 101,540	343	XecnGV	27
115	104	ac68/pif-6	α+β+γ+δ	101,512 → 101,901	130	ChocGV	57
116	105	116	β	101,968 → 102,501	178	ErelGV	36
117	106	iap-5	β	102,520 → 103,371	284	PhopGV	46

Table 4 Annotation of Open reading frames (ORF) of AgseGV-DA, AgseGV-L1 and AgseGV-XJ as annotated (columns 1 and 2), the annotated ORF names (3), the conserved occurrence in other baculovirus genera (baculovirus core genes = α+β+γ+δ, conserved in *Alpha-*, *Beta-* and *Gammabaculovirus* = α+β+γ, conserved in *Alpha-*, *Beta-* and *Deltabaculovirus* = α+β+δ, conserved in *Alpha-* and *Betabaculovirus* = α+β, conserved *Betabaculovirus* genes = β, orthologues in other baculoviruses = bac. ortho., AgseGV-DA/L1 unique genes = unique) (4); the genome sequence position in AgseGV-DA and orientation relative to *granulin* (5); the amino acid (aa) number of predicted gene products (6), the next baculovirus homolog (7), regarding to PSI-BLAST/BLASTp with shared amino acid identity (%) (8)

118	107	lef-9	α+β+γ+δ	103,355 → 104,845	497	EpapGV	71
119	108	f-protein	bac. ortho.	104,851 → 105,297	149	EpapGV	66
120	109	120	unique	105,281 ← 106,129	283	*none*	--
121	110	DNA ligase	bac. ortho.	106,171 ← 107,868	566	PiraGV	58
122	111	122	bac. ortho.	108,049 → 108,225	59	ChocGV	42
123	112	123	unique	108,261 ← 108,485	75	*none*	--
124	113	fgf-2	β	108,540 ← 109,721	394	PiraGV	35
125	114	p10	bac. ortho.	109,789 ← 110,160	124	ClanGV	35
126	--	126	bac. ortho.	110,168 ← 110,578	137	ErelGV	29
127	115	alk-exo	α+β+γ+δ	110,597 → 111,790	398	PiraGV	51
128	116	helicase-2	bac. ortho.	111,768 → 113,183	472	SpfrGV	55
129	117	129	bac. ortho.	113,176 ← 114,081	302	SpliGV	33
130	--	130	bac. ortho.	114,150 ← 114,647	166	SpliGV	25
131	118	lef-8	α+β+γ+δ	114,756 ← 117,407	884	CpGV	67
132	119	132	bac. ortho.	117,411 ← 118,310	300	XecnGV	31
133	120	133	bac. ortho.	118,305 → 118,775	157	SpfrGV	33
134	121	134	bac. ortho.	118,789 ← 118,992	68	EpapGV	44
135	122	ac53	α+β+γ+δ	118,961 → 119,371	137	PhopGV	50
136	123	136	bac. ortho.	119,366 ← 120,799	478	PsunGV	45
137	124	137	β	120,861 ← 121,976	372	SpliGV	36
138	125	138	bac. ortho.	121,979 ← 122,164	62	PsunGV	44
139	126	lef-10	bac. ortho.	122,163 → 122,393	77	PiraGV	49
140	127	vp1054	α+β+γ+δ	122,254 → 123,249	332	DisaGV	50
141	--	141	bac. ortho.	123,328 → 123,513	62	MoGV	52
142	--	142	unique	123,510 → 123,887	126	*none*	--
143	128	fgf-3	β	123,937 → 124,851	305	ChocGV	37
144	129	egt	bac. ortho.	124,868 ← 126,256	463	LacoGV	60
145	130	145	unique	126,361 → 126,984	208	*none*	--
146	131	me53	α+β	127,085 → 128,014	310	PlxyGV	50
147	132	he65	bac. ortho.	128,064 ← 129,638	525	PsunGV	46

Table 4 Annotation of Open reading frames (ORF) of AgseGV-DA, AgseGV-L1 and AgseGV-XJ as annotated (columns 1 and 2), the annotated ORF names (3), the conserved occurrence in other baculovirus genera (baculovirus core genes = α+β+γ+δ, conserved in *Alpha-*, *Beta-* and *Gammabaculovirus* = α+β+γ, conserved in *Alpha-*, *Beta-* and *Deltabaculovirus* = α+β+δ, conserved in *Alpha-* and *Betabaculovirus* = α+β, conserved *Betabaculovirus* genes = β, orthologues in other baculoviruses = bac. ortho., AgseGV-DA/L1 unique genes = unique) (4); the genome sequence position in AgseGV-DA and orientation relative to *granulin* (5); the amino acid (aa) number of predicted gene products (6), the next baculovirus homolog (7), regarding to PSI-BLAST/BLASTp with shared amino acid identity (%) (8)

148	--	bro	bac. ortho.	129,769 ← 131,058	430	MabrNPV	34
149	--	149	unique	131,180 → 131,530	117	*none*	--

Chapter III: Agrotis segetum nucleopolyhedrovirus but not Agrotis segetum granulovirus replicate in AiE1611T cell line of Agrotis ipsilon

This chapter is published with few modifications in:
Gueli Alletti, G., Carstens, E. B., Weihrauch, B., Jehle, J. A. (2018),
Agrotis segetum nucleopolyhedrovirus but not Agrotis segetum granulovirus replicate in AiE1611T cell line of *Agrotis ipsilon*, J. Invert. Pathol., 151:7-13

Abstract

Both Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) and Agrotis segetum granulovirus (AgseGV) belong to a cluster of four baculoviruses that are infective for different *Agrotis* species. Belonging further to different baculovirus genera, namely *Alphabaculovirus* and *Betabaculovirus*, respectively, AgseNPV-B and AgseGV are candidates to investigate virus interactions in co-infections. However, for the investigation of virus interactions on a cellular level, permissive insect cell-lines are needed. The cell line AiE1611T deriving from *Agrotis ipsilon* eggs has been shown to be permissive for several *Alphabaculovirus* isolates. In this study, virus replication was followed based on microscopic analysis of infected and transfected cells, as well as on a molecular level by PCR of DNA and cDNA of selected baculovirus transcripts. While the permissivity was not verified for AgseGV, AgseNPV-B produced occlusion bodies in both infection with hemolymph of infected larvae and Lipofectamin transfection with AgseNPV-B genomic DNA. In addition to the possibility to investigate virus interaction of AgseNPV-B with other alphabaculoviruses, the permissivity of AiE1611T for AgseNPV-B further offers the possibility a biological selection to separate AgseNPV-B from AgseGV.

Introduction

The family of *Baculoviridae* comprises occluded dsDNA viruses with rod-shaped virions infecting larval stages of the insect orders Lepidoptera, Diptera and Hymoptera (Herniou et al., 2011; Jehle et al., 2006a). Based on their phylogenetic relationship, baculoviruses are classified in four genera, which also reflect their host association and to a certain extent the morphology of their occlusion body (OB), a proteinaceous matrix covering the occlusion derived virion (ODV) phenotype. Members of the genera *Alphabaculovirus* and *Betabaculovirus* are specific for Lepidoptera, whereas viruses form *Gamma*- and *Deltabaculovirus* infect Hymopteran and Dipteran species, respectively (Herniou et al., 2003; Jehle et al., 2006a). While betabaculoviruses normally occlude a single virion in an ovoid OB, the polyhedral protein matrices of alpha-, gamma- and deltabaculoviruses contain few to multiple ODVs. These differences in the OB morphology were recognized in an earlier classification, namely granuloviruses and nucleopolyhedroviruses (Theilmann et al., 2005). ODVs are the phenotype responsible for *per os* infections of larvae, the second virion phenotype, the budded virus (BV) is responsible for the intercellular transmission of infection (Federici, 1997). Baculoviruses are highly virulent for the larval stages of their insect hosts and generally exhibit a narrow host range to a single or a few host species (Cory and Bishop, 1997; Cory and Myers, 2003; Herniou et al., 2004). This characteristics have been utilized for the development and commercialization of baculovirus biocontrol agents of pest insects in agriculture, horticulture and forestry (Black et al., 1997; Moscardi, 1999).

Four baculoviruses have been isolated from larvae of the cutworm species *Agrotis segetum* and *A. ipsilon* (Lepidoptera: Noctuidae), namely Agrotis segetum nucleopolyhedrovirus A (AgseNPV-A), Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B), Agrotis ipsilon nucleopolyhedrovirus (AgipNPV) and Agrotis segetum granulovirus (AgseGV), belonging to four distinct species of the genera *Alphabaculovirus* and *Betabaculovirus*, respectively (Gueli Alletti et al., 2015; Jakubowska et al., 2006; Wennmann et al., 2015a). Recently, AgseNPV-B has been shown to infect *A. segetum* larvae and its potential role as a biological control agent of *Agrotis* species has been proposed (Wennmann et al., 2015a; Wennmann et al., 2015c). Under *in vivo* conditions, infections with AgseNPV-B are often associated with AgseGV in co-infections fueling the interest in studying mechanism of alphabaculovirus and betabaculovirus co-infections in Lepidoptera (Wennmann et al., 2015c). On the other hand, cell lines derived from *A. ipsilon* eggs have been shown to be permissive for the closely related AgipNPV (Harrison and Lynn, 2008). These cell lines, namely AiE1611T and AiEd6T, exhibit favorable growth conditions, form monolayers of spherical cells and produce OBs when infected with AgipNPV. As such, a cell line system would be an ideal tool to study baculovirus mixed infections at the cellular level so we investigated whether the cell line AiE1611T is permissive to AgseNPV-B and AgseGV. AgseNPV-B but not AgseGV was also able to replicate AiE1611T cells.

Material and Methods

Insects

Mass-rearing of *A. segetum* was performed as described previously (Wennmann and Jehle, 2014). In brief, neonate larvae were kept on semi-artificial diet (Ivaldi-Sender, 1974) at 22 °C with a 16/8 h light/dark photoperiod until they reached the fourth larval stage. Pupae were collected and kept at 25 °C until the adults hatched. Adult moths were kept in plastic cylinders. Eggs were incubated at 25 °C in moist boxes until neonates hatched.

Cell line and maintenance

Stocks of the *A. ipsilon* cell line AiE1611T were provided by Robert L. Harrison from the Agricultural Research Service of the United States Department for Agriculture (Harrison and Lynn, 2008). AiE1611T cells were maintained in ExCell420 (Sigma Aldrich) serum free insect cell culture medium supplemented with 3% fetal bovine serum (Fisher Scientific), 10 U/ml penicillin and 0.1 mg/ml streptomycin. The AiE1611T cells were maintained in 25 cm² (T25) Greiner® cell culture flasks at 26 °C and split on a weekly basis into fresh T25 cell culture flasks (4×10^6 cells per flask) (Lynn, 2002).

Viruses

Virus stocks of Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) and Agrotis segetum granulovirus (AgseGV, isolate DA) were originally provided by Doreen Winstanley, Horticulture Research International (HRI) collection, Warwick (UK) as previously described (Gueli Alletti et al., 2017b; Wennmann et al., 2015a). AgseNPV-B and AgseGV were propagated in late third or early fourth instar larvae by feeding them with diet plugs (8 mm³) overlaid with either 1000 OBs of AgseNPV-B (1×10^6 OBs/ml), or with 10^6 OBs of AgseGV (1×10^9 OBs/ml). AgseNPV-B and AgseGV OBs were purified out of diseased cadavers as previously described (Wennmann and Jehle, 2014). AgseNPV-B OB concentration was measured in refractive phase-contrast microscopy using a Neubauer-improved hemocytometer.

AgseGV OB concentration was measured with a Petroff-Hauser hemocytometer (0.02 mm depth) in dark-field microscopy (Leica DM RBE).

Transfection AiE1611T cells with genomic DNA from either AgseGV or AgseNPV-B

Genomic DNA was isolated from either AgseNPV-B OBs or AgseGV OBs by dissolving the OB matrix with 0.1 M Na_2CO_3 followed by pH adjustment to pH 7 with 0.1 M HCl, protein and RNA digestion, a standard phenol/chloroform extraction and ethanol precipitation (Sambrook and Russell, 2001; Wennmann et al., 2015c). DNA concentration and purity were measured with a NanoDrop 2000c spectrophotometer. About 1×10^6 AiE1611T cells were transfected 2 µg genomic DNA with the aid of Lipofectamin® (Invitrogene) treated with from either AgseNPV-B or AgseGV (Felgner et al., 1987). Due to their different refractions in light microscopy, putatively infected cells were identified by the formation of OBs (occ^+) for AgseNPV-B and cytopathological effects (CPE) for AgseGV compared to control wells treated with Lipofectamin® and water only.

Cell debris and OBs were separated from BV and suspended in TEK-buffer (1 M Tris-EDTA pH 7.5, 0.15 M KCl). After an incubation at 65 °C for 30 min, sodium dodecyl sulfate and proteinase K were added in final concentrations of 3% and 0.2 M, respectively, followed by a further incubation at 65 °C for 30 min. DNA was purified by phenol/chloroform extraction and ethanol precipitation (Gross-Bellard et al., 1973; Sambrook and Russell, 2001). Supernatants from transfections with AgseNPV-B DNA and AgseGV DNA, respectively, were used to infect a second round of cells. DNA from these infections was extracted as described and all samples were subjected to PCR reactions with specific primers for AgseNPV-B *polyhedrin* (UKf_2.2: GCCGAGGATCCATTTTTTG , UKr_2: CGCAGAG CGTGTTGAGCTAAA, 260 bp fragment) and AgseGV *granulin* (GVf_spez03 GACAGGCGTATA TCGGAAGC, GVr_spez TGAGCGACGTAATCTGGATG, 347 bp fragment), respectively. PCR conditions were identical as below.

RNA purification from transfected cells and RT-PCR of selected AgseNPV-B and AgseGV genes

Laboratory equipment and surfaces were treated with 0.1% (v/v) diethyl pyro carbonate/water (DEPC Carl Roth®) or RNase AWAY® (Carl Roth®) to avoid RNase contamination. Cells and supernatants were harvested from transfections. Cells and possible OBs were separated from supernatants by centrifugation at 300 *g* and 10 °C for 10 min. Supernatants were removed and the pelleted cells derived from transfections with genomic DNA of AgseNPV-B or AgseGV were resuspended in 1.2 ml of RLT Plus (QIAGEN®) before purification of RNA using a QIAGEN® RNeasy® Plus Mini Kit. Total RNA was measured by UV-VIS absorbance. Reverse transcription of 1 µg of RNA was performed using the iScript™ Reverse Transcription Supermix for RT-qPCR kit (BIO-RAD©). PCR reactions of cDNA samples were conducted in order to detect transcription of selected AgseNPV-B and AgseGV genes (Table 1). Amplicons were generated in separate PCR reactions (0.2 µM primer pairs, 1 µl of cDNA added to Axon®1 × reaction buffer BD detergent and Mg^{2+} free buffer, 2 mmol $MgCl_2$, 0.2 mmol of a dNTP mix (each nucleotide with equimolar quantities of 10 mM) and 2.5 U *Taq* polymerase (Axon®). PCR regime: 3 min initial denaturation at 95 °C, followed by 35 cycles of 30 sec denaturation at 95 °C, 30 sec primer annealing at 58 °C and DNA elongation for 30 sec at 72 °C. PCR amplicons were electrophoresed through 3% (w/v) agarose gels at 8 V/cm in TBE buffer for 120 min

(Sambrook and Russell, 2001). Separated amplicons were visualized by staining the gels with 6 µl of Midori Green Advance (Biozym) per 100 ml agarose solution and documented under UV light.

Comparison of budded virus of AgseNPV-B from infected *A. segetum* larvae with DNA from AgseNPV-B transfected cells

Late third to early fourth instars were fed with small pieces of artificial diet (~8 mm³) overlaid with 1 – 2 µl of AgseNPV-B OB suspension (10^6 OBs/ml) At day 3 post infection, larvae were anesthetized with hemolymph suspension was carefully collected clipping one of the anterior legs. The hemolymph was sterile-filtered through a 0.2 µm Whatman® membrane. For the initial amplification of BV, 0.5 ml of the pooled hemolymph was used to infect 2×10^6 cells. For the comparison, about 1×10^6 AiE1611T cells were transfected with 1 µg of genomic DNA from AgseNPV-B with the aid of Lipofectamin. Supernatant of occ^+ was filtered and used for a second round of infection (2×10^6 AiE1611T cells). In both cases, infection was checked visually on a daily basis until 75% were occ^+. Genomic DNA was subsequently digested with *Eco*RI and *Hin*dIII (FastDigest, Thermo Scientific) in order to conclude possible differences of intracellular DNA that derived either from transfections of cells (AgseNPV-B TRA) or from infections with *A. segetum* hemolymph (AgseNPV-B HEM). The fragments were separated by 0.8% agarose gel electrophoresis at 8 V/cm in TAE buffer and were visualized by staining the gels with 6 µl of Midori Green Advance (Biozym) per 100 ml. The pattern and size of the restriction fragments of the clones were compared to *Eco*RI and *Hin*dIII digestions of the AgseNPV-B (KM102981) (Wennmann et al., 2015a).

Determination of the median tissue culture infective dose ($TCID_{50}$)

The concentration of plaque forming units (pfu/ml) of budded virus (BV) suspensions was determined by end-point dilution titration with 3×10^3 cells/well (O'Reilly et al., 1994; Reed and Muench, 1938). Occ^+ wells were scored on a daily basis by refractive phase-contrast for seven days. Concentration of plaque forming/infectious units (pfu) were calculated based on Poisson-distributed estimation of infection (O'Reilly et al., 1994).

Table 1 Oligonucleotide primers specific for Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) and Agrotis segetum granulovirus (AgseGV) used for reverse transcription PCR. Given are their target genes, their names, positions of amplicons in the genome, the expected amplicon sizes and the primer sequences.

NPV	Gene	Primer name	Pos. in genome	Size (bp)	Primer sequence (5‘– 3‘)
	vef-1	prAsBvef1-f prAsBvef1-r	72,627 – 72,746	120	TCGAGCTGTTGGACAACGAC TCCACCTGTATTGCAGCTCG
	efp	prAsBefp-f prAsBefp-r	12,160 – 12,379	220	CGAACACACCAATCTGCACG TTTAACTACGGCCAGCGTCA
	polyhedrin	prAsBpolh-f prAsBpolh-r	467 – 606	140	TCGTCGAACCTGTGTACGTC TTCCCATATGACGCGGTTGA
	ie-1	prAsBie1-f prAsBie1-r	140,868 – 141,067	200	ACCAAACGGGCCTTGTACTC GGAGTCGATATTCCGCTGCA
	lef-8	prAsBlef8-f prAsBlef8-r	119,969 – 120,148	180	TCGCGACAGTCCAACTTTGA AACCTTTTCGTGGCCAGTGT
	vp39	prAsBvp39-f prAsBvp39-r	83,849 – 84,008	160	ATTCCGTGGCAACTCTGTGG GGTGACTTCTTCTTCGCGGG

Table 1 Oligonucleotide primers specific for Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) and Agrotis segetum granulovirus (AgseGV) used for reverse transcription PCR. Given are their target genes, their names, positions of amplicons in the genome, the expected amplicon sizes and the primer sequences.

GV	Gene	Primer name	Pos. in genome	Size (bp)	Primer sequence (5‘– 3‘)
	enhancin	prAsGVenh-f prAsGVenh-r	54,623 – 54,772	150	GACGTGGGCGAGCATATTTG GGAAAAGGAGTTCTGGCCGA
	f-protein	prAsGVfpro-f prAsGVfpro-r	23,927 – 24,046	120	AGCCTTTTCGGAGGAGCTTT GTCGGTGCTGTTTTCATGCT
	granulin	prAsGVgran-f prAsGVgran-r	112 – 311	200	GACAGGCGTATATCGGAAGCT CGTGTCCAAGTTTCGCGAAG
	ie-1	prAsGVie1-f prAsGVie1-r	7,690 – 7,839	150	ACAACTCTTGCACCGAGGAC TCGCTGAAATTCACGCCAAC
	lef-8	prAsGVlef8-f prAsGVlef8-r	115,297 – 115,466	170	TCCACTTTTTGTCCCCTCACC GACGTGGAGGGATGGAATCG
	vp39	prAsGVvp39-f prAsGVvp39-r	83,343 – 83,542	200	AACAGAACACCGAAGACGCT CTGTAATTTGGCGAGCACGT

Plaque purified AgseNPV-B clones

Plaque purified clones of AgseNPV-B were isolated from amplified BV suspensions. The virus was adsorbed to the cells for 90 min at 26 °C with gentle rocking every 15 min. BV suspensions were removed by aspiration and the cells were overlaid with 2 ml of 0.8% low melt agarose (Roth) dissolved in ExCell420 with 3% FBS, 10 U/ml penicillin and 0.1 mg/ml streptomycin (Brown and Faulkner, 1978; Cooper, 1961; O'Reilly et al., 1994). The cells were monitored on a daily basis and a randomized set of occ^+ (0.5 – 2 mm) plaques was picked at 7 dpi. Each isolated plaque was used to infect 1×10^6 cells. Infected cells were harvested at 7 dpi Cell debris and OBs were separated from BV. The DNA was extracted from the pelleted cells and OBs and subjected to digestion with *Hin*dIII. Fragments were separated in a 0.8% agarose gel electrophoresis at 8 V/cm in TAE buffer for 4 h followed by staining with Midori Green Advance and documentation under UV light. Restriction patterns of the clones were size-compared to a *Hin*dIII *in silico* digestion of the AgseNPV-B genome sequence (KM102981).

Full-range bioassays for determination of the median lethal viral dose (LD_{50})

Full range bioassays were performed with OBs of the original in vivo produced virus stock of AgseNPV-B and the in vitro produced AgseNPV-B PP 2. Serial 1:10 dilutions of purified OBs were prepared and fed to L2/L3 *A. segetum* larvae with small cubic pieces of artificial diet (8 mm^3) overlaid each with 1 µl of virus treatment, or water in cases of control groups. Each treatment consisted of 25 – 30 tested animals and 50 uninfected control animals, three to four independent replicates were performed. Larvae that did not ingest the offered diet within 12 h were excluded from the experiment. The groups of tested larvae were transferred into individual 50-well boxes containing 50 ml of premixed Stonefly *Heliothis* diet (Ward´s Science). Mortality rates were scored at 1 dpi (to exclude larvae killed by handling) and at 7 dpi to score mortality caused by virus treatment. The mortality rates at 7 dpi were corrected for control mortality according to Abbott (1925). Probit analysis was performed with software ToxRat 3.0; the infectivity was compared in a parallel line assay (ToxRat Solutions GmbH, Alsdorf, Germany).

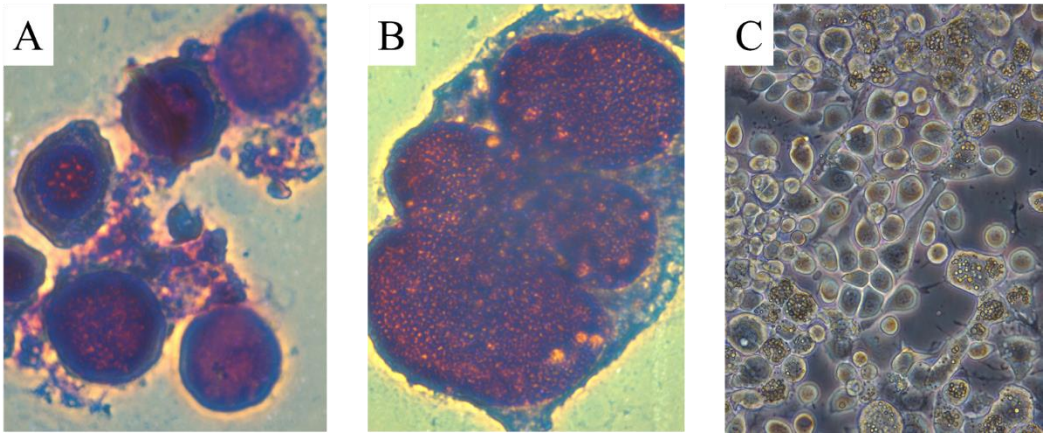

Figure 1 Light micrograph at (A) 200-fold and (B) 400-fold resolution of Giemsa stained AiE1611T cells transfected with 2 µg genomic DNA from AgseGV. Cells show signs of bloated nuclei and disordered cell division. (C) Phase-contrast micrograph at 200-fold resolution of AiE1611T cells transfected with 2 µg AgseNPV-B DNA. Foci of occlusion bodies are concentrated in regions of cell nuclei at 10 days after transfection.

Results

Infection and transfection experiments using budded virus and DNA

From a total of six wells of AiE1611T cells transfected with genomic DNA of AgseGV, after 10 days, only one well showed any cells with cytopathological effects (CPE). In Giemsa-stained micrographs of these particular candidate cells, nuclei of the cells were enlarged as shown by the enrichment of the purple red DNA complexed with Giemsa-stain. The enlarged nuclei were embedded in a thin layer of bluish stained cytoplasm. At higher resolution, cells were spotted that had stopped cell division, which resulted in bloated cells with more than one enlarged nucleus (Figure 1A, 1B). However, the other wells did not show any signs of CPE or signs of AgseGV infection. When AiE1611T cells were transfected with purified AgseNPV-B DNA, the cells produced OBs within 10 days (Figure 1C). In addition, infection with hemolymph derived from *A. segetum* larvae infected with AgseNPV-B resulted in visible OB production within 3 days post infection (data not shown). Phase-contrast microscopy showed three different kinds of cells: cells without OBs, cells with few OBs and cells with large numbers OBs, often densely packed, in the nuclei. The formation of OBs was a clear evidence for virus replication and completion of the infection with AgseNPV-B.

RFLP and PCR analyses of AgseNPV-B clones and AgseGV transfected cells

When total cellular DNA was isolated from transfected or infected cells and subjected to PCR using granulin/polyhedrin-specific primers, specific products were confirmed for AgseNPV-B but not for AgseGV (data not shown). When supernatants of these transfections, were used to infect fresh cells, OBs (data not shown) and positive PCR signals from cell pellets (Figure 2) was only observed with AgsNPV-B but not with AgseGV supernatants, indicating that AgsNPV-B but not AgseGV had produced infective BV. Restriction analysis of the isolated total intracellular DNA, originating from transfection with AgseNPV-B (TRA) of from infection with hemolymph derived BV of AgseNPV-B (HEM) provided identical *Eco*RI and *Hin*dIII digests (Figure 3), which were typical for AgseNPV-B (Wennmann et al., 2015). Both infections of AgseNPV-B (TRA) and AgseNPV-B (HEM) produced budded viruses in the cell culture supernatant, with titers of 1.9×10^5 pfu/ml and 1.5×10^6 pfu/ml, respectively.

RT-PCR

To provide further evidence that AgseGV was not replicating in AiE1611T cells, transcription patterns of selected ORFs were determined by RT-PCR. Total intracellular RNA from cells showing CPE following transfection with AgseGV-DNA was analyzed for a selection of early, late and very late virus genes from both AgseGV and AgseNPV-B (*vef-1/enhancin*, efp/*f-protein*, *polyhedrin*/*granulin*, *ie-1*, *lef-8*, *vp39*). Only two faint bands, i.e. *granulin* and *ie-1*, were identified in the cDNA samples of AgseGV transfected cells (Figure 4 A). In contrast, AgseNPV-B RNA produced positive cDNA signals for all tested ORFs (Figure 4B).

Plaque purification of AgseNPV-B and bioassays

Plaque purification was performed for both AgseNPV-B (TRA) and AgseNPV-B (HEM). In total ten isolates were identified and purified from plaques at 7 dpi. Six plaques derived from AgseNPV-B (HEM) (AgseNPV-B PP1 - PP6) and four plaque purified genotypes derived from AgseNPV-B (TRA) (AgseNPV-B PP7 – PP10). Total intracellular DNA was extracted from AiE1611T cells infected with the plaque purified material in order to assess restriction fragment analysis of these isolates (Figure 5). All isolates showed similar RFLP patterns and were considered as genetically identical isolates of AgseNPV-B.

OBs of AgseNPV-B PP2 were used for the bioassays and compared to *in vivo* produced AgseNPV-B. The LD_{50} of this clone was 296 OB/larva compared to the LD_{50} of 150 OB/larva of the *in vivo* produced AgseNPV-B OB (Table 2). Based on the criterion of overlapping confidence intervals of the LC_{50} value (Robertson and Preisler, 1992), the *in vivo* and *in vitro* produced OBs did not differ in their activity in *A. segetum* larvae.

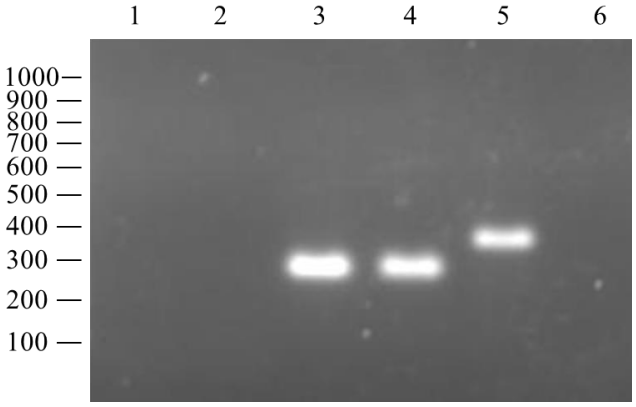

Figure 2 PCR of total intracellular DNA after infections with supernatants of AgseNPV-B and AgseGV transfected cells; (lane 1) non target control (NTC) with water and specific primers for AgseNPV-B *polyhedrin* and AgseGV *granulin*; (lane 2) uninfected cells with both specific primer pairs; (lane 3) genomic AgseNPV-B DNA with polyhedrin primers (positive control); (lane 4) cells from AgseNPV-B infection with specific primers; (lane 5) genomic AgseGV DNA with granulin primers (positive control) ; (lane 6) cells from AgseGV infection with specific primers. Size marker (bp) is given to the left.

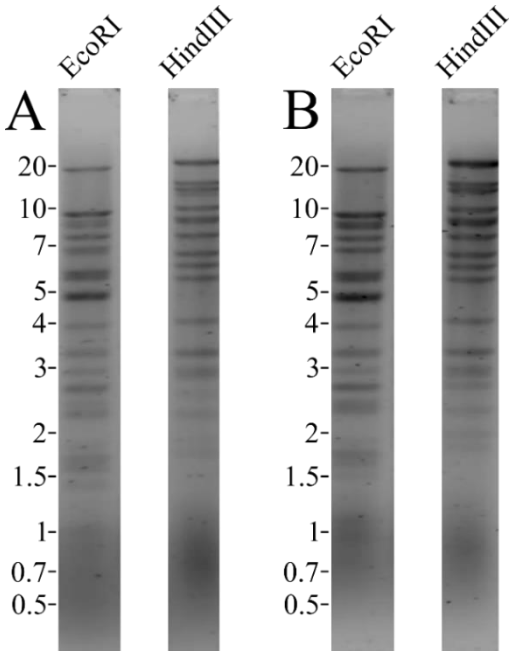

Figure 3 *Eco*RI and *Hin*dIII restriction analysis of 1 µg total intracellular DNA isolated from infected cells with AgseNPV-B TRA (A), AgseNPV-B HEM (B). Fragments were separated on a 0.8% agarose gel in 1 × TAE at 8 V/cm for 4 h. Size marker (kbp) is given to the left; inverse photograph.

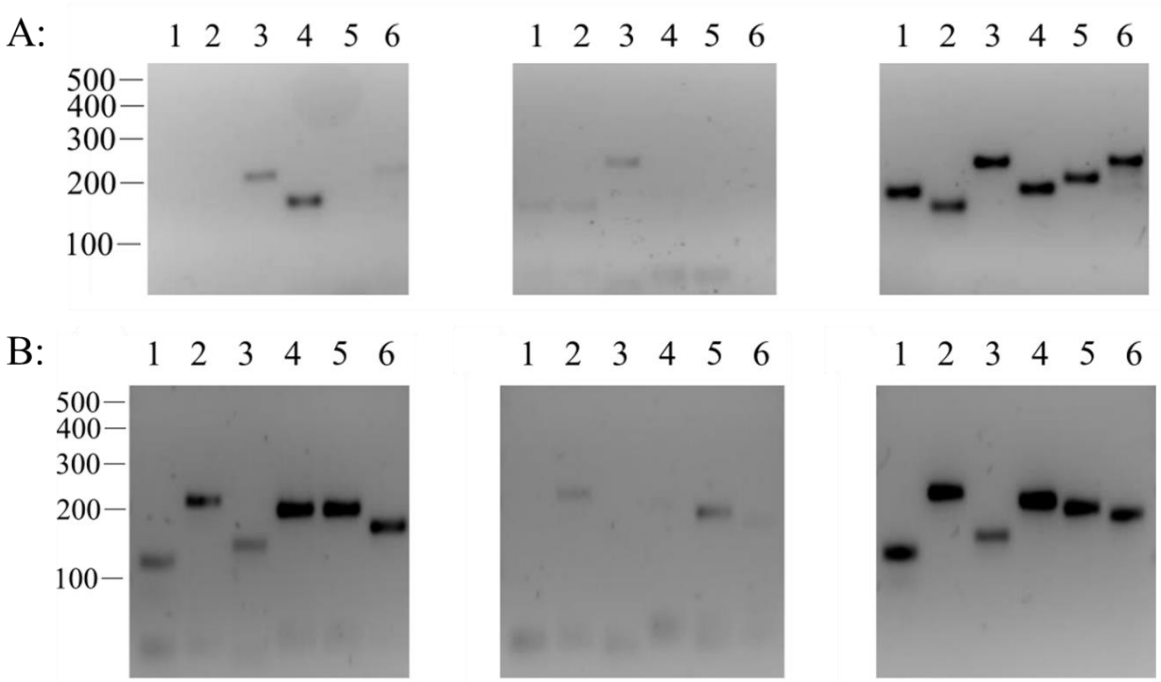

Figure 4 PCR with specific primers for selected (A) AgseGV and (B) AgseNPV-B genes. Amplicons were produced in PCRs with cDNA (left), no-RT controls (middle) from RNA of infected cells and with genomic DNA of AgseGV and AgseNPV-B (right). AgseGV: (1) *enhancin*, (2) *f-protein*, (3) *granulin*, (4) *ie-1*, (5) *lef-8*, and (6) *vp39*. AgseNPV-B: (1) *vef-1*, (2) *efp*, (3) *polyhedrin*, (4) *ie-1*, (5) *lef-8*, and (6) *vp39*. For expected fragment sizes see Table 1. Amplicons were separated on a 3% agarose gel in 1 × TBE at 8 V/cm for 120 min. Size marker (bp) is given to the left; inverse photograph.

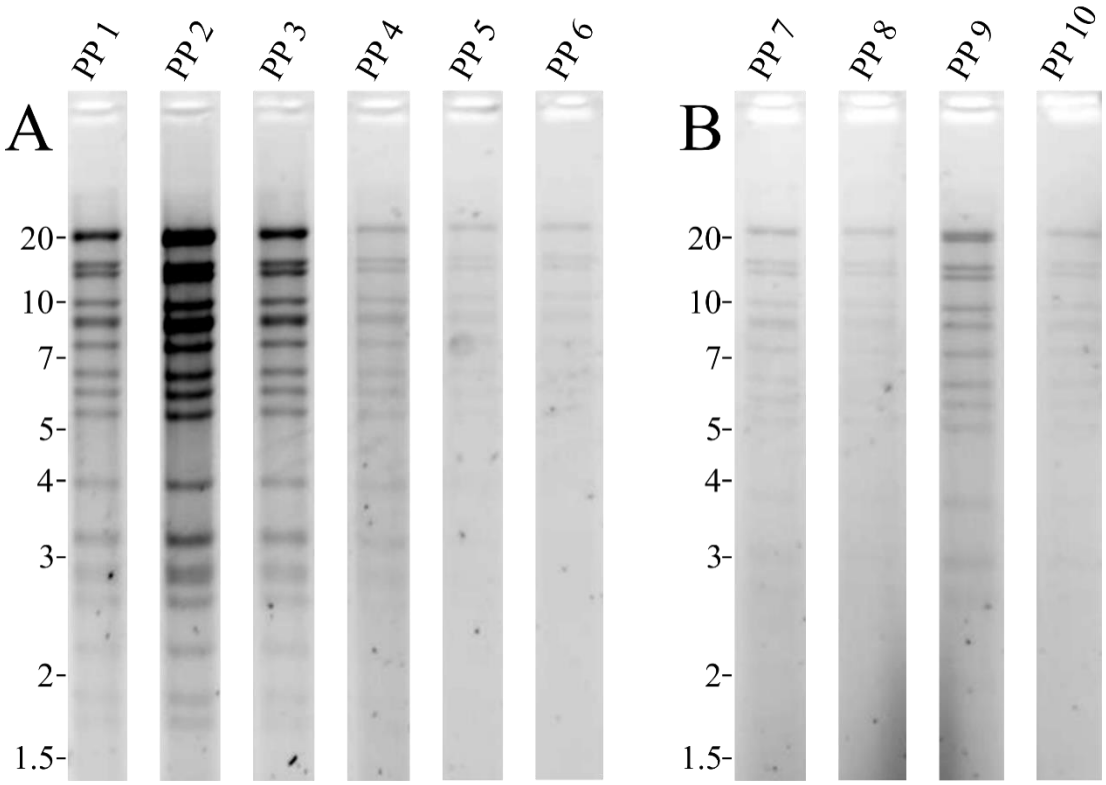

Figure 5 *Hin*dIII restriction analysis of 5 µg total intracellular DNA isolated from cells infected with plaque purified isolates (PP1 – PP10). (A) PP1 – PP6 derived from hemolymph of AgseNPV infected larvae, (B) PP7 – PP10 from cells transfected with AgseBPV-B DNA. Fragments were separated on a 0.8% agarose gel in 1 × TAE at 8 V/cm for 4 h. Size markers (kbp) are given to the left; inverse photograph.

Table 2 Median lethal dose (LD_{50}) of AgseNPV-B produced *in vivo* and *in vitro* with L2/L3 instars of *A. segetum*[a]

AgseNPV-B	No. individuals per bioassay / No replicates	LD_{50} (95% CL) OB/larva	Slope[b]	p(χ^2)[c]	Potency
in vivo	1120 / 4	150 (35 – 424)	0.68	0.002	1
in vitro	430 / 3	296 (170 – 515)	0.89	0.146	0.43

[a] Mortality rates were scored on day 7 post infection and corrected for control mortality (Abbott, 1925). CL (confidence limits) calculated by Fieller´s Theorem; df degrees of freedom; Relative Potency according to dose-response of AgseNPV-B ORI

[b] The validity criterion of parallelism was confirmed as p(F) values were < 0.05.

[c] Heterogenity correction of the fitted confidence limits was performed when p(χ^2) values of the fit curves were < 0.05.

Discussion

Both AgseNPV-B and AgseGV belong to a cluster of four baculoviruses that infect different *Agrotis* species (Jakubowska et al., 2006; Wennmann et al., 2015a; Zethner, 1980; Zethner et al., 1987; Zhang et al., 2014). In this study, we demonstrated that AiE1611T cells are fully permissive for AgseNPV-B infection leading to production of viral OBs. This result was not restricted to transfections with genomic AgseNPV-B DNA, but also to infection with naturally obtained BV from infected *Agrotis segetum* larvae. It appears that AiE1611T cells are permissive for a wide range of Alphabaculoviruses including AgipNPV (Harrison and Lynn, 2008), AgseNPV-A (El-Menofy and Jehle, unpublished results), AgseNPV-B (this study),

Autographa californica multiple nucleopolyhedrovirus (AcMNPV), Anagrapha falcifera multiple nucleopolyhedrovirus (AfMNPV), Anticarsia gemmatalis multiple nucleopolyhedrovirus (AgMNPV), Galleria mellonella nucleopolyhedrovirus (GmNPV), Heliothis armigera nucleopolyhedrovirus (HearNPV), Plutella xylostella nucleopolyhedrovirus (PlxyNPV) and Rachiplusia ou nucleopolyhedrovirus (RoMNPV) (Harrison and Lynn, 2008). In addition to OB production, AgseNPV-B infections were also confirmed by RT-PCR and PCR or REN analyses. AgseNPV-B OBs produced *in vitro* in AiE1611T cells were infective to *A. segetum* larvae as confirmed with bioassays. Although AiE1611T cells have a broad host range for many alphabaculoviruses, it was not permissive for the betabaculovirus AgseGV. In RT-PCR, only faint bands for *ie-1* and *granulin* were detected. In general, *ie-1* possesses an early promoter with a canonic CAGT motif recognized by host RNA polymerase (Blissard et al., 1992; Pullen and Friesen, 1995a; Pullen and Friesen, 1995b). Recognition of *ie-1* promoter is host independent and has been reported even in mammalian cell lines (Pfeifer et al., 1997; Theilmann and Stewart, 1991). Additionally to the host independent transcription of *ie-1*, it also appears to be transcribed irrespectively from complete infection cycles (Kovacs et al., 1992). The *granulin* gene is a homologue to the polyhedrin gene of alphabaculoviruses. It is considered as a very late gene transcribed by a baculovirus RNA polymerase. Its expression is thought to appear only after virus replication. This generalized picture of polyhedrin/granulin transcription maybe challenged by some betabaculoviruses, because recent reverse transcriptionanalyses and micro-array studies of Cydia pomonella granulovirus (CpGV) revealed some early transcription of *granulin* under *in vivo* infection conditions (Schneider and Jehle, unpublished results). The lack of transcription of other late AgseGV ORFs, such as *enhancin*, *f-protein*, *lef-8*, *vp39*, provide clear evidence that AgseGV is not able to replicate and that the transcription is presumably blocked at an early stage of infection. For this reason, despite to some cytopathological effects in transfections with AgseGV DNA, neither DNA replication (data not shown) nor RNA transcription could be demonstrated. In general, only a very few cell cultures permissive for betabaculoviruses have been reported; the best studied system is the cell line CpDW1 which replicates CpGV (Gebhardt et al., 2014; Hilton and Winstanley, 2007; Winstanley and Crook, 1993).

Although AgseNPV-B and AgseGV belong to different baculovirus genera, namely *Alphabaculovirus* and *Betabaculovirus*, they are both able to cross- and co-infect larvae of the hosts *A. segetum* and *A. ipsilon* (El-Salamouny et al., 2003; Wennmann and Jehle, 2014; Wennmann et al., 2015c; Zethner, 1980; Zethner et al., 1987). Hence, a combined infection of these two viruses, could serve as a promising model to investigate virus-virus interaction of two baculoviruses from different genera (Wennmann et al., 2015c). Due to the lack of replication of AgseGV in AiE1611T, these studies, however, cannot be extended to cell culture. However, AiE1611T cells allow replication of AgseNPV-B under *in vitro* conditions, which will provide new opportunities for studying its pathology on molecular level. Its permissivity to AgseNPV-B and many other alphabaculoviruses allows mixed infection studies using virus isolates from different alphabaculovirus species. Under natural infection condition of *A. segetum* larvae, AgseNPV-B and AgseGV often occur mixed because of co-infections. Since we could not find any hint that AgseGV can replicate in AiE1611T, this cell line may also be used as a biological selection system to purify AgseNPV-B from AgseGV.

Chapter IV: Agrotis segetum nucleopolyhedrovirus B shows high genome stability during serial *in vitro* passages in AiE1611T cells

This chapter is intended for submission.

Abstract

Ten serial infections of the plaque purified clone PP2 of Agrotis segetum nucleopolyhedrovirus (AgseNPV-B) were performed in the permissive cell line AiE1611T to study the stability of the viral genome during cell culture passage. A loss of virulence was observed right after the first round of passaging, then PP2 remained relatively stable over ten passages. This was observed by the absence of *few polyhedra* phenotypes in phase contrast microscopy and by next-generation sequencing (NGS) of five selected passages. NGS sequencing revealed that *defective particles* were absent over ten passages and that only very few mutations, in total only 19 single nucleotide polymorphisms (SNPs), occurred throughout the experiments. Two of these SNPs were located in the baculovirus core gene *vp91*. A moderate multiplicity of infection (MOI = 1) was applied. The results show that despite the decreased virulence, AgseNPV-B (PP2) is therefore a tentative candidate for the production of a microbial pest control agent in AiE1611T cells due to its genome stability observed in ten serial passages.

Introduction

The Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) Agrotis segetum nucleopolyhedrovirus A (AgseNPV-A) and Agrotis ipsilon nucleopolyhedrovirus (AgipNPV), as well as the Agrotis segetum granulovirus (AgseGV), form a cluster of baculoviruses infecting different *Agrotis* species (Wennmann et al., 2015a). These viruses have some potential of being used as biocontrol agent of different cutworm species, such as *Agrotis segetum*, *A. ipsilon* or *A. exlamationis* (El-Salamouny et al., 2003; Jakubowska et al., 2006). The genome of AgseNPV-B is 148,981 bp in size and encodes 150 putative open reading frames (orf), as well as six putative homologous repeat sequences (hrs). Based on the phylogenetic analyses of the amino acid (aa) sequences of 37 core genes from 43 baculovirus genome sequences it was found that AgseNPV-B is most closely related to AgipNPV but both viruses belong to different *Alphabaculovirus* species (Wennmann et al., 2015a). In addition, infection experiments of *Agrotis* larvae revealed that AgseNPV-B had a similar virulence as AgipNPV (El-Salamouny et al., 2003). As for all baculoviruses, AgseNPV-B has two virion phenotypes: (1) occlusion derived virions (ODVs), which are embedded in a proteinaceous occlusion body (OB) are responsible for *per os* infections of larvae and for horizontal transmission from larvae to larvae, and (2) budded viruses or budded virions (BVs), which are responsible for a cell-to-cell transmission of infection within larvae (Gueli Alletti et al., 2017a; Herniou et al., 2011). In general, BVs are also infectious for susceptible insect cell lines (Volkman and Summers, 1977), though suspensions of purified ODVs may also initiate infections of cell culture, but to a far lower extend (Lynn, 2003).

Whereas BVs contain only one nucleocapsid, the OBs of alphabaculoviruses contain numerous ODVs with single or multiple enveloped nucleocapsids. Recent concepts of baculovirus population genetics, that were based on ultra-deep sequencing approaches, suggested that single OBs of Autographa californica multiple nucleopolyhedrovirus (AcMNPV) can be considered as genetically diverse populations of genotypes due to the large number of diverse nucleocapsids (Chateigner et al., 2015). In infections of larvae, competition between these

different genotypes of an isolate results in a stabilizing effect on genomes of baculovirus populations. Consequently, genotypes with infective phenotypes for larvae are maintained as the predominant form in serial *in vivo* passages. Opposing to serial *in vivo* infections, serial *in vitro* passages of BVs through susceptible cell lines, can act as a bottleneck for the BV phenotype and even peril the loss of virus activity in host larvae. Recently, the insect cell line AiE1611T deriving from *Agrotis ipsilon* eggs has been identified to be permissive for several alphabaculoviruses, including AgseNPV-B (Gueli Alletti et al., 2017a; Harrison and Lynn, 2008). The cell line exhibit favorable growth conditions, forms monolayers of spherical cells and produces large numbers of AgseNPV-B OBs, when infected with AgseNPV-B BV. Thus, AgseNPV-B is a candidate for the *in vitro* production of a biological control agent against *Agrotis* species. However, any production of virus in permissive cell lines is affected by limitations that alter yield, activity and performance of virus progeny. Of great importance is the passage effect in serial infections in cell culture (Krell, 1996), which describes an alternation of the selection pressure in favor of the BV phenotype. As a consequence, mutant genotypes lacking ODV specific genes that are necessary for *in vivo* infection might be accumulated over several rounds of *in vitro* infections.

Two main virus mutants have been observed in serial *in vitro* passages: *few polyhedra* (*fp*) mutants that are characterized by low numbers of newly produced OBs per infected cell (Fraser et al., 1983; Potter et al., 1978), and defective interfering particles (DPIs) that have large deletions (Rohrmann, 2013). One of the reasons for DPI accumulation is their replication advantage during *in vitro* infection because of the smaller genome sizes. The accumulation of DPIs depends highly on the multiplicity of infection employed (Kool et al., 1991). As shown by electron microscopy and molecular studies, low multiplicities of infection assist selection against *fp* and DPI mutants in serial passages (Kool et al., 1991; Zwart et al., 2008). Classical approaches to identify DPI production include the analyses of DNA endonuclease restriction fragment patterns deriving from digestions of DNA isolated from infected cell cultures. Since large parts of the baculovirus genome can be missing in DPI mutants, restriction patterns of DPIs differ to those of wild-type genomes in these analyses. However, small mutations, such as single nucleotide polymorphisms (SNPs), small insertions and deletions of few base pairs cannot be easily identified by these methods.

By applying Solexa Illumina sequencing we investigated the genomic stability of AgseNPV-B over ten serial passages in the cell line AiE1611T when a moderate multiplicity of infection was applied. The *in vivo* activity of cell culture produced OBs was further tested in bioassays using larvae of *A. segetum*. It was found that the biological activity and the genome compositions of AgseNPV-B remained rather stable during the passage experiment.

Material and Methods

Insects, cell line and maintenance

Mass-rearing of *A. segetum* was performed as described previously (Wennmann and Jehle, 2014). Stocks of the *A. ipsilon* cell line AiE1611T were provided by Robert L. Harrison from the Agricultural Research Service of the United States Department for Agriculture (Harrison and Lynn, 2008). AiE1611T cells were maintained as previously described (Gueli Alletti et al., 2017a) and split on a weekly basis into fresh T25 cell culture flasks (4×10^6 cells per flask) (Lynn, 2002).

Viruses

An occlusion body (OB) stock of Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) [species *Agrotis segetum nucleopolyhedrovirus B,* genus *Alphabaculovirus*] was provided by Doreen Winstanley, Horticulture Research International (HRI) collection, Warwick (UK). After budded viruses (BV) were prepared from larvae, suspensions of this virus were plaque purified *in vitro* in cell cultures of AiE1611T as previously described (Gueli Alletti et al., 2017a). One of these purified clones, namely AgseNPV-B PP2 (abbreviated PP2 in the following) showed a high virus titer in determinations of the $TCID_{50}$, as well as typical OB formation (Occ^+) in infections of AiE1611T cells. PP2 was therefore subjected to experiments in serial infections of the cell culture AiE1611T.

Determination of the median tissue culture infective dose ($TCID_{50}$)

The concentration of plaque forming units (pfu/ml) of BV suspensions was determined by end-point dilution titration with 3×10^3 cells/well (O'Reilly et al., 1994; Reed and Muench, 1938). Occ^+ wells were scored on a daily basis by refractive phase-contrast for seven days. Concentration of plaque forming/infectious units (pfu) were calculated based on Poisson distributed estimation of infection (O'Reilly et al., 1994).

Serial infections of AiE1611T cells

The initial density of seeded cells was 4×10^4 cells/cm². The cells were infected with a multiplicity of infection (MOI) of 1 pfu per cell for the initial infection with PP2 and incubated at 26 °C for one week. All consecutive passages (PP2 #1 – PP2 #10) included a determination of the $TCID_{50}$ as described above, followed by an infection with a MOI = 1 for one week under the same conditions as in the initial infection. For DNA sequencing, the first, third, fifth, seventh and tenth passage (PP2 #1, #3, #5, #7 and #10) were seeded in T75 cell culture flasks to yield higher amounts of OBs, for all other passages cells were seeded in T25 for easier maintenance. Samples of the passages #1, #3, #5, #7 and - #10 were divided equally and used either for total DNA extraction or for OB purification, as previously described (Gueli Alletti et al., 2017a). The obtained OBs were separated from cells by disintegrating the cell membrane in 0.1% SDS and ultra-sonic treatment for 10 min at room temperature, before they were pelleted at 20,000 *g*. Concentration of OBs was measured in refractive phase-contrast microscopy with a Neubauer improved hemocytometer.

Full-range bioassays for determination of the median lethal viral dose (LD_{50})

Full range bioassays were performed with OBs of the original virus stock of AgseNPV-B (in the following: ORI), and the passages #1, #3, #5, #7 and #10 of PP2. Second to third instars (L2/L3) of *A. segetum* larvae were starved overnight and then fed with a small cube (2 × 2 × 2 mm) of artificial diet (Wennmann et al., 2015c), each overlaid with 1 µl of serial tenfold OB dilutions resulting in final doses of 1 to 10^5 OBs per cube. For untreated control groups water instead of OB suspensions were applied. Each treatment group consisted of 25 – 30 larvae, whereas 50 additional larvae were used for the untreated control group. Animals that did not ingest the inoculum within 12 h were excluded from the experiment. Then, the larvae were individually placed in each well of a bioassay tray (Licefa, Germany), each well contained about 1 ml of premixed Stonefly *Heliothis* diet (Ward´s Science). Larval mortality was recorded at 7 days post infection (dpi) and corrected for mortality of untreated controls (Abbott, 1925). The dose-responses were subjected to Probit analysis with linear maximum likelihood regression (ToxRat 3.0) (Lehmann et al., 2016). Confidence limits (CL) of 95% were computed based on

Fieller´s Theorem and mortality within the passages was compared based on parallel line assays (PLA) of quantal responses and 95% CL. In cases, that corresponding dose-response lines did not fulfill the criterion of linearity in PLA, F-statistics were used to investigate parallelism of lines based on standard operating procedures of the European Pharmacopeia Commission (https://www.edqm.eu/ en/european-pharmacopoeia-commission).

Sequencing and consensus genome assembly

About 50 ng purified DNA samples of AgseNPV-B ORI and the selected passages PP2 #1, #3, #5, #7, #10 were subjected to NexteraXT library preparation and an Illumina NextSeq500 paired-end sequencing (StarSEQ Ltd., Mainz Germany). Each sequencing produced approximately 1.0 – 2.5 Mio reads with an average length of 151 nucleotides (nt), resulting in an average of about 1000- to 2500-fold genome coverage when a genome size of 148,981 bp for AgseNPV-B is assumed (Wennmann et al., 2015a). A heuristic assembly strategy combining *de novo* assembly, re-mapping and Sanger sequencing was followed for the correct consensus genome assembly and detection of mutations. *De novo* assembly was conducted using the Edena assembly program with standard parameters (Hernandez et al., 2008; Hernandez et al., 2014). The contigs produced in that assembly were aligned against the AgseNPV-B reference genome using the native Geneious assembler for mapping (Geneious R9.1 Biomatters Ltd.). Large contigs (> 1000 bp) that did not match to the reference genome were compared to the nucleotide collection database (nr/nt) using the basic local alignment search tool for nucleic acids (blastn) in order to identify contaminating nucleic acids (Altschul et al., 1990; Johnson et al., 2008). The consensus sequences were aligned to the AgseNPV-B reference sequence using the Mauve whole genome aligner of Geneious (Darling et al., 2004). Re-arrangements in the alignment were further investigated by designing flanking sequencing primers and submitting the amplicons for Sanger sequencing (StarSEQ Ltd., Mainz Germany).

Detection of deletions, insertions and single nucleotide polymorphisms to AgseNPV-B

Read-pairs with 50% consecutive bases below an average Phred quality score of 30 per read cycle (99.9% base call accuracy) were excluded from the analyses (Gordon, 2009). Obtained read-pairs were re-mapped against AgseNPV-B reference genome (KM102981) using the Bowtie2 aligner for short reads with standard parameters for very-sensitive local alignment of reads on the JKI Galaxy server (Afgan et al., 2016a; Langmead and Salzberg, 2012a; Mielczarek and Szyda, 2016; Ye et al., 2015; Ziemann, 2016). Insertions and deletions (InDel) were identified by their breakpoints using Pindel (Ye et al., 2009). Single nucleotide polymorphisms (SNPs) were detected with the SAMtools MPileup tool (Li et al., 2009). The average threshold of the quality of mutations in the alignment was set to a Phred-scaled probability of 20 (99% base call accuracy in the alignment) or higher in order to take into account possible errors due to Illumina sequencing and Bowtie2(Li, 2011a; Li, 2011b; Li and Durbin, 2009). The statistical analyses were assessed using the R-package “Bioconductor” (MacQueen, 1966; R Development Core Team, 2015; Robinson, 2016).

Results

Influence of serial infections to OB production and infectivity in host larvae

Infections of AiE1611T cells with PP2 produced large numbers of OBs within 7 days post infection (dpi) as observed in phase-contrast microscopy of these infected cells (Figure 1A). In infected cells, OBs appeared as large, refractive polyhedrons within the Giemsa-stained nuclei and were aggregated within the nuclei (Figure1B). The formation of numerous OBs per cell did not decrease after 10 rounds of serial passaging, as shown exemplarily for the eighth passage, PP2 #8 (Figure C, D). OB formation was generally visible starting at 3 dpi and resulted in highly packed cells at 7 dpi (Figure C, D). No visual evidence for *fp* phenotypes was observed in any serial passaging step. The OBs of the passages PP2 #1, #3, #5, #7 and #10 were harvested and used for both the determination of the lethal dose-responses and to investigate the genomic stability. The employed MOI and infected cell numbers were identical for each passage, also the harvested OB amounts were similar, ranging from 1.12×10^5 OB/µl for PP2 #7 to 3.29×10^5 OB/µl for PP2 #1 ($P < 0.05$, Tukey-HSD) (Table 1). As the OB production in different passages was in the same range of magnitude, a mean concentration of $2.38 \times 10^5 \pm 8.27 \times 10^4$ OB/µl for all passages and thus an OB productivity of 397 OB per infected cell was estimated.

In bioassays with ORI a LD_{50} of 150 OB/larva was observed, the LD_{50} of PP2 #1 was 296 OB/larva. For PP2 #3 to #7 the LD_{50} were about 10-fold reduced whereas PP2 #10 showed a similar LD_{50} as PP2 #1. This corresponded to a change of the potency from 1.97 in passage #1, to 16.7 (#7) to 1.8 (#10). As the number of test animals and the biological replicates were not uniform within the single passages, statistical tests for the comparison slopes was based on the F distribution. Here, no statistical differences of slopes of the probit lines were found according to F-test statistics, although the slope of the probit line of PP2 #7 was flatter than those of all other passages. Furthermore, the *in vivo* virulence of the OBs decreased during the passage experiments until PP2 #7, but in the end this decline was reverted with passage #10. Therefore, the trend of a declining *in vivo* virulence with increasing passage number could not be fully concluded from the experiments.

Table 1 Production and biological activity of occlusion bodies (OBs) of different AgseNPV-B passages. Given are the OB concentration obtained and the,median lethal dose (LD_{50}) of AgseNPV-B ORI and the passages #1, #3, #5, #7, #10 of AgseNPV-B PP2 with L2/L3 instars of *A. segetum*[a].

Passage	OB conc. ± SD[b] [10^5 × OB/µl]	No. individuals per bioassay	LD_{50} (95% CL) OB/larva	Slope[c]	p(χ^2)[d]	df	Potency[e]
ORI	1.8 ± n.d.[A]	1120	150 (35 – 424)	0.68	0.002	4	1
#1	3.29 ± 0.76[B]	430	296 (170 – 515)	0.89	0.146	3	1.97
#3	2.16 ± 0.27[C]	725	1921 (1350 – 2727)	0.94	0.525	3	12.8
#5	2.44 ± 0.33[C]	752	2131 (768 – 6053)	0.87	0.042	3	14.2
#7	1.12 ± 0.27[D]	1167	2507 (736 – 14108)	0.51	0.002	4	16.7
#10	2.90 ± 0.49[BC]	726	265 (56 – 879)	0.86	0.012	3	1.8

[a] Mortality rates were scored on day 7 post infection and corrected for control mortality (Abbott, 1925). CL (confidence limits) calculated by Fieller´s Theorem; df degrees of freedom; Relative Potency according to dose-response of AgseNPV-B ORI

[b] Same letters indicate overlapping 95% family-wise CL in multiple comparisons of means in Tukey-HSD-Test

[c] The validity criterion of parallelism was confirmed as p(F) values were < 0.05.

[d] Heterogenity correction of the fitted confidence limits was performed when p(χ^2) values of the fit curves were < 0.05.

[e] Potencies are calculated as the quotient of LD_{50} values from #1, #3, #5, #7 and #10 to ORI

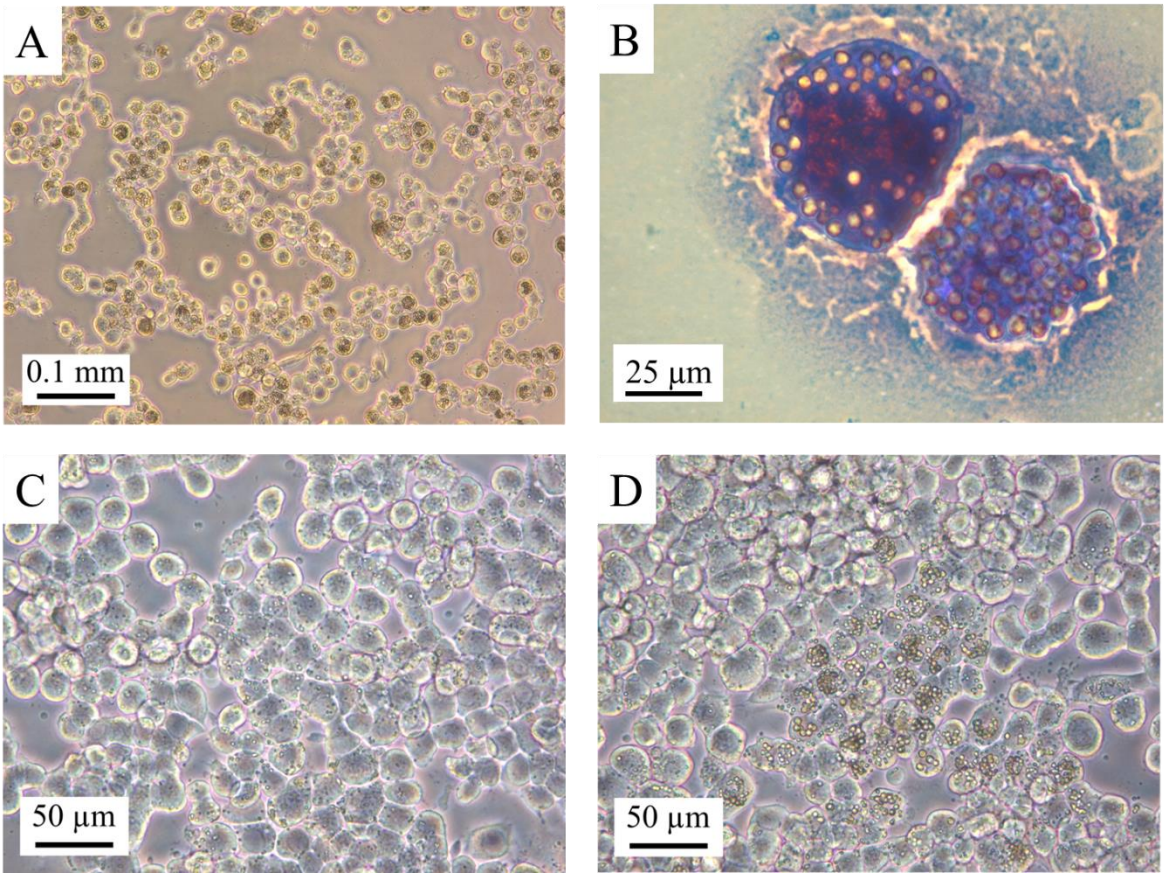

Figure 1 Phase-contrast micrograph of AiE1611T cells infected with PP2 at 7 dpi. Cells show intact structure and signs of late infection as high numbers of occlusion bodies are formed (A). Light-micrograph of Giemsa-stained cells infected with PP2. Cell structure is disrupted; occlusion bodies are assembled in the centers of the cells (B). Phase-contrast micrographs of a high passage of AgseNPV-B PP2 (#8) at 3 dpi (C) and 7 dpi (D).

Table 2 Assembly reports of reads with an average Phred-quality score ≥ Q30 generated by NextSeq500 NGS assembled against the AgseNPV-B reference genome sequence (KM102981) using the Bowtie2 mapper.

AgseNPV-B Sample	Reads pairs ≥ Q30	No of read pairs assembled to RefSeq	No. of read pairs not assembled to RefSeq	Mean coverage ± SD	%GC content[a]
ORI	193,138	158,136 (90%)	35,002 (18%)	316 ± 323	43.7
PP2 #1	1,160,163	695,277 (60%)	464,886 (40%)	1370 ± 243	46.3
PP2 #3	1,196,424	1,104,009 (92%)	92,415 (8%)	2162 ± 377	45.3
PP2 #5	1,120,376	1,008,500 (90%)	111,876 (10%)	1959 ± 930	43.5
PP2 #7	1,204,968	778,537 (65%)	426,431 (35%)	1532 ± 458	46.6
PP2 #10	956,470	954,650 (> 99%)	1,820 (< 1%)	1812 ± 360	45.0

[a]%GC content refers to the content of all mapped reads (not to the %GC-content of the genome sequence)

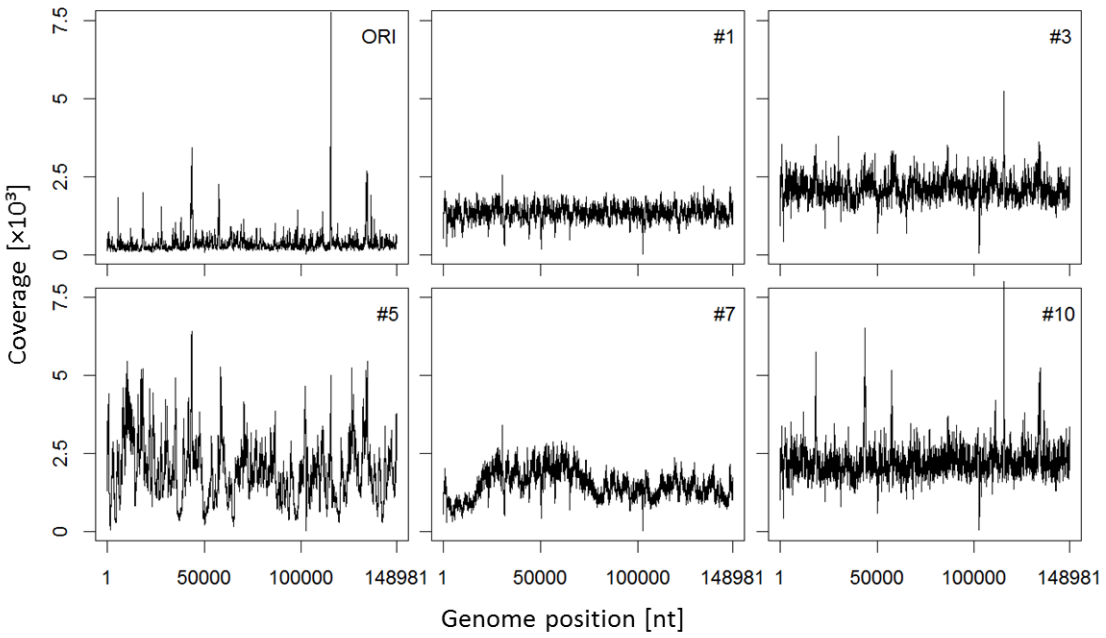

Figure 2 Coverages of Bowtie2 assemblies of AgseNPV-B (ORI) and the passages #1, #3, #5, #7 and #10 of PP2 according to their genome position [nt] in the reference AgseNPV-B sequence (KM102981)

Comparison of AgseNPV-B ORI and the AgseNPV-B reference genome sequence

The original AgseNPV-B virus stock (Wennmann et al., 2015a), termed ORI in this study, was re-sequenced in this NGS approach and compared to the reference genome sequence (Acc. No. KM102981) that was determined by 454 sequencing. The Solexa Illumina data resulted in 316-fold average global coverage of the genome. One reason for this low coverage compared to the sequenced clones from serial infections (Table 2) was a low yield of high-quality reads due to problematic clustering in the sequencing flow-cell. The coverage of the reads fluctuated over the reference genome sequence (Figure 2). The presence of large deletions was evaluated by a de novo assembly of the obtained reads and by aligning the contigs to the AgseNPV-B reference genome. A large deletion would have resulted in a decrease of sequencing coverage to zero of an expanded genome region in comparison to the reference sequence. However, such extended gaps in the coverage could not be observed in the obtained of AgseNPV-B ORI indicating a lack of major deletions in the passages. The presence of smaller deletions, insertions or duplications (up to 10 nt) was analyzed using the detection tool PinDel. Although two insertions were detected (data not shown), they were not considered for the further analyses as the Phred quality score even fell below the quality criterion of 13 (95% base call accuracy).

In the analyses of single nucleotide polymorphisms (SNPs), SNPs were assessed when the Phred quality score was ≥ 20 (99% base call accuracy). The obtained sequencing reads of the initial AgseNPV-B ORI already comprised single nucleotide polymorphisms in 10 nucleotide positions, when mapped with Bowtie2 and in 5 positions when mapped with BWA-MEM, respectively (Table 3 and 4). Only on position, namely a proposed transversion from G to T at nucleotide (nt) position 972 in *orf1629* was identified with a variant frequency of 100% in both assembly strategies. As this SNP was transmitted to all following passages in the serial infections with 100% variant frequency, it rather reflects an annotation error in the reference genome of AgseNPV-B, which was sequenced on an Abi Solid 454 platform in 2014 (Wennmann et al., 2015a). Re-evaluation of the original 454 raw data confirmed that this the

reference genome sequence has indeed a thymine residue at this position and not a guanin. The fourteen remaining SNPs were identified in *homologous repeat* sequences (11 in *hr1* and 3 in *hr3*) and in *asb018*, causing a silent mutation due to a change from guanine to adenine in this AgseNPV-B specific *orf*. All but one identified SNPs in ORI were supported by read coverages above the mean global read coverages per nucleotide position of 316 ± 323 reads. The exception is the SNP at nt position 972, with only 440 read coverage.

Table 2 Single nucleotide polymorphisms observed between the AgseNPV-B reference genome (KM102981) and the sequences passages in MPileup analyses of Bowtie2 assemblies. The variations are based on the original genotype at one of the 16 nucleotide positions on the reference genome. SNPs were quality filtered by their mapping quality (Phred quality ≥ 20) with a resulting base call accuracy of 99% or greater.

Position	Variant frequency [%] Total coverage [reads][b]						Change	Variations	CDS/hr	AA-Effect
	ORI	#1	#3	#5	#7	#10				
972[a]	100 440	100 1,685	100 2,964	100 2,674	100 1,658	100 2,961	G → T	Transversion	*orf1629*	none
18,450	40 763	40 1,226	30.8 2,794	42.3 4,536	50 1,339	50 3,974	C → T	Transition	*hr1*	none
18,454	75 827	40 1,226	30.8 2,886	48 4,680	50 1,426	47.4 4,126	G → T	Transversion	*hr1*	none
18,462	50 699	40 1,226	30.8 2,793	47.8 4,684	50 1,442	62.5 3,740	C → A	Transversion	*hr1*	none
18,493	50 895	50 1,217	28.6 2,772	45 4,537	60 1,391	35.3 3,952	T → G	Transversion	*hr1*	none
18,501	57.1 1,504	50 1,219	28.6 2,973	45 4,596	60 1,399	33.3 5,268	A → C	Transversion	*hr1*	none
18,505	57.1 1,509	40 1,224	26.7 3,000	47.4 4,637	60 1,411	31.6 5,304	A → G	Transition	*hr1*	none
18,657	77.8 670	62.5 1,159	52.4 2,915	50 3,912	60 1,336	80 2,982	A → T	Transversion	*hr1*	none
18,669	80 705	50 1,277	57.1 3,162	50 4,221	50 1,453	72.7 3,164	T → G	Transversion	*hr1*	none
18,670	77.8 701	25 1,217	33.3 3,013	34.5 4,039	16.7 1,388	50 3,085	T → C	Transition	*hr1*	none
18,676	0 532	37.5 1,117	25 3,751	14.3 3,813	33.3 1,319	30.4 2,934	G → T	Transversion	*hr1*	none
18,678	0 533	37.5 1,173	25 2,720	14.3 3,731	33.3 1,309	25 2,888	A → C	Transversion	*hr1*	none
86,291	0 250	0 1,435	30.8 2,849	12.5 2,396	66.7 1,493	25 1,724	A → G	Transition	*vp91*	D→E
86,313	0 282	0 1,377	46.2 2,874	25 2,343	66.7 1,459	25 1,720	T → C	Transition	*vp91*	D→G
128,603	0 423	50 1,349	66.7 2,084	75 1,674	66.7 1,552	0 1,959	G → A	Transition	*vef-3*	none
128,625	0 541	50 1,387	60 2,099	28.6 1,769	50 1,630	0 2,150	C → T	Transition	*vef-3*	P→S

[a]The single nucleotide polymorphism can be considered as an artefact deriving from false *ad initio* nucleotide information in the reference sequence.

[b]The total coverage is the sum of reads of the reference and the variant nucleotide at the SNP position.

Table 3 Single nucleotide polymorphisms observed between the AgseNPV-B reference genome (KM102981) and the sequenced passages in MPileup of BWA-MEM assemblies. The variations are based on the original genotype at one of the 5 nucleotide positions on the reference genome. SNPs were quality filtered by their mapping quality (Phred quality ≥ 20) with a resulting base call accuracy of 99% or greater.

Position	Variant frequency [%] Total coverage [reads][b]						Change	Variations	CDS/hr	AA-Effect
	ORI	#1	#3	#5	#7	#10				
972[a]	99.7	100	100	100	100	99.9	G → T	Transversion	*orf1629*	none
	292	771	1,137	1,101	687	1,247				
17,740	2.6	2.6	2	2.7	3.2	20.5	G → A	Transition	*asb018*	none
	78	610	1,131	1,278	504	1,035				
57,725	27.7	0	6.8	12.8	0.7	11.1	T → A	Transversion	*hr3*	none
	719	345	1,015	842	458	1,502				
57,739	29.4	0.6	6.3	11.2	0.4	12.5	T → A	Transversion	*hr3*	none
	714	348	932	758	471	1,467				
57,746	51.3	0.8	5.1	7.7	0.9	10.7	T → A	Transversion	*hr3*	none
	772	470	1,138	977	634	1,649				

[a]The single nucleotide polymorphism can be considered as an artefact deriving from false *ad initio* nucleotide information in the reference sequence.

[b]The total coverage is the sum of reads of the reference and the variant nucleotide at the SNP position.

Genome analyses of serially passaged AgseNPV-B PP2

Illumina sequencing of different passages was performed to study the dynamic of genomic changes of AgseNPV-B through ten serial passages in cell culture. Like for AgseNPV-B ORI, larger deletions were not detected in the serial passages as no extended break-in was observed in coverages of the *de novo* assembled contigs aligned to the AgseNPV-B reference genome (KM102981). Some larger contigs that did not map against the reference sequence were assigned as homologous to *Agrotis ipsilon* mitochondrial sequences (GeneBank: KF163965) and were considered to be derived from contamination of the samples with host DNA. Bowtie2 re-mapping of reads with an average Phred quality score of 30 or higher concluded in mean global coverages in the plaque purified clones that ranged from 1370-fold (PP2 #1) to 2162-fold (PP2 #3) nucleotide coverages (Figure 2). The high fluctuations might be caused by the method how sequenced reads were initially generated. The approach presented here utilized a transposase-like enzyme. Therefore, sequencing inserts from this transposase reaction ranged from several hundred bp up to 2 kb which led to an unevenly distribution in the Solexa Illumina reads. However, certain positions were identified in all serial passaged with very low coverages in the sequencing reads. Such an exemplary position spans in the genome from nucleotide position 102,540 to 102,600 in an AgseNPV-B specific *orf* (*asb099*). This region is region is characterized by a repetition of the 14 nucleotides GAT TTP CTT ACP GY. Although fluctuations of the sequencing reads in the AgseNPV-B ORI mapping do not allow an identification of the low coverage, the in-break at this particular position in PP2 #1 to #10 derives from an error in the reference sequence. The PinDel analyses did not allow any assessment of insertions or deletions at this position (Phred quality score below the criterion of Q20), another possible explanation could be that beginning with the first passage a deletion of roughly 40 nucleotides was introduced at this position. Following the coverages in figure 1, several other positions could be identified with coverages converging to zero, e.g. around nucleotide position 50,000. However, the coverages around those positions were all within the range of the mean coverage and the standard deviation.

In variant calling with the multi-sample variant caller MPileup 20 SNPs with a base calling accuracy of 99% were identified using data from Bowtie2 or BWA-MEM (Table 3 and 4) from passage #3 to #7. Two SNPs were identified in each *vp91* and *vef-3*, which encodes a virion capsid protein associated with the *per os* infectivity factors and a viral enhancing factor, respectively. The transitions in *vp91* presumably caused in both cases an amino acid change from aspartic acid to glutamic acid at nt position 86,291, and from aspartic acid to glycine at nucleotide position 86,313. While in *vef-3* one transition concluded in a silent mutation, whereas the transition at nt position 128,625 may have caused an amino-acid change from proline to serine. The transitions in *vp91* were not detected in ORI and #1, and in addition the transitions in *vef-3* were not detected in ORI and #10.

The variant frequencies of identified SNPs fluctuated within the passages with no evident pattern. For example, the frequency of the single SNP in *asb018* increased only in the last passage (#10) to 20%, whereas it ranged between 2% and 3% in the remaining passages, including ORI. The mean frequency of SNPs in *hr1* dropped from initial 51% in ORI to 43 and 34% in #1 and #3, respectively and increased in the following passages again to almost 50%. An almost similar development was observed in *hr3*, with exception for PP2 #7. Here the mean frequency dropped from initial 36% to 0% in #1 which was followed by an increasing frequency in #3, #5 and #10. The two SNPs in *vef-3* were introduced in the first passage (PP2 #1) with 50% and inherited to all following passages. The frequencies almost increased simultaneously and ranged between 28.6 and 75% in the passages #3, #5 and #7. In the last passage of PP2, #10, the frequency of the SNPs dropped back to 0%. The two SNPs in *vp91* were introduced in the third passage (PP2 #3) and inherited to all following passages. The highest SNP frequencies and therefore putative point mutations were observed in PP2 #7 with 66.7%, whereas the frequencies dropped back to 25% in the last passage of PP2 #10.

Discussion

The aim of this study was to determine the *in vivo* virulence and the genetic stability of AgseNPV-B OBs that derived from serial infection of BV through the cell line AiE1611T, which was evaluated using BVs from each passage for passing on the infection. In the serial passages high numbers of OBs were produced during all passage steps. Formation of *fp* mutants) was not observed during the passage experiment. Previously, *fp* mutants have been observed for several baculoviruses passaged in cell culture, often related to aberrations of the *fp/25k* gene. In passages of AcMNPV) insertions of host-derived transposable elements or large deletions caused the *fp* phenotype (Beames and Summers, 1989; Harrison and Summers, 1995). The deletion of transposon sites has been shown to delay the insertion of transposable elements in that particular gene but it does not prevent from insertion (Giri et al., 2012; Giri et al., 2010). In contrast, the *fp* mutants from Lymantria dispar multiple nucleopolyhedrovirus (LdMNPV) and Helicoverpa armigera nucleopolyhedrovirus (HearNPV) were caused by small indel mutations. In LdMNPV, insertions of 1 bp or small deletions of 8 to 24 bp (Bischoff and Slavicek, 1997) whereas in HearNPV alternation of only few nucleotides as (Lua et al., 2002) were responsible. *Fp* mutants of AcMNPV and LdMNPV were in particular observed when viruses were employed at a high infection dose(MOI), whereas with HearNPV *fp* mutants were also produced at a low MOI (0.5 pfu/cell) (Lua et al., 2002). All *fp* mutants observed so far share the same phenotypic key characteristics, which are (i) a reduced number of OBs, with a reduced number of virions, (ii) an overproduction of BV, and (iii) a decreased *in vivo* virulence

of OBs (Beames and Summers, 1989; Bischoff and Slavicek, 1997; Bull et al., 2003; Giri et al., 2012; Harrison and Summers, 1995; Lua et al., 2002). AgseNPV-B PP2 may be less affected by these mutations or the moderate MOI (= 1) employed may have reduced this effect. Similar to HearNPV, where a 28-fold reduced *in vivo* activity within six consecutive passages was observed (Lua et al., 2002), the calculated LD_{50} values of AgseNPV-B PP2 varied during the passages. The LD_{50} values of ORI and PP2 #1 OBs appeared to be similar, hence the cell culture production of PP2 had no immediate effect on the *in vivo* activity. However, a 10-fold increase of the LD_{50} was noted for the passages PP2 #3 – #7, whereas PP2 #10 showed a similar lethal dose response as ORI and PP2 #1. A logic explanation of this observation cannot be given and may be a result of the experimental limitations. Due these limits neither technical nor biological replicates could be obtained and each bioassay was carried out only once. A clear trend of change of virulence over time could not be observed. Therefore, these results suggest that the observed differences are in the variability of the chosen *in vivo* system and may not reflect a significant decrease or increase of virulence during the passage experiment. A possible reason for the different responses might also be that tested *A. segetum* larvae ranged between 2nd and 3rd instars in their larval stages. In general, infection of only L1 larvae may provide more homogenous results due to their higher susceptibility to virus infection. In contrast for infecting in L2/L3 larvae more infective virus particles are deemed necessary (Evans, 1983; Smits and Vlak, 1988). The scattering age of tested animals in this approach is reflected by the large and partially overlapping 95% confidence limits of the determined median lethal doses. However, LD_{50} determination of L1 larvae is also very difficult to achieve because they feed very little medium resulting in an extremely high variability of virus uptake. Therefore, the use of L2/L3 larvae scattering in age was a compromise dictated by the biological system and needed to access a full uptake of virus medium.

On the other hand, there are also examples, that the *in vivo* activity of i*n vitro* cloned baculovirus isolates can be increased by serial *in vivo* passages (Arrizubieta et al., 2014). Another point of concern when producing baculoviruses *in vitro* is the yield of OBs. In case of AgseNPV-B PP2, a mean production of 397 OBs per cell were observed over all passages. This amount of OBs per cell is in yield range comparable to other NPVs, for example HearNPV produced in early *in vitro* passages (~225 OBs/cell) (Lua et al., 2002) and those from selected clones of HearNPV strain H25EA1(~200-600 OBs/cell) (Nguyen et al., 2011). Furthermore, baculovirus yields normally decrease with growing passage numbers as demonstrated for several other alphabaculoviruses, e.g. AcMNPV, LdMNPV, AgMNPV (Anticarsia gemmatalis multiple nucleopolyhedrovirus), HearNPV (Castro et al., 1997; Chakraborty and Reid, 1999; Huynh et al., 2015).

In addition to the *in vivo* activity, the genetic stability of different viral passages was compared in mappings of Solexa Illumina reads to the AgseNPV-B reference genome sequence (KM102981), the only AgseNPV-B genome sequence available deriving from *de novo* assemblies of genomic AgseNPV-B sequenced with an ABi Solid 454 platform. Large deletions or insertions were neither observed in the genome sequence of AgsNPV-B ORI (original isolate), nor in any of the obtained genome sequences of the serially passaged clone PP2. Although SNPs were identified in only 20 positions of AgseNPV-B ORI, a high identity to the reference genome sequence was demonstrated. In general, different SNPs were identified when the reads were mapped with Bowtie2 or with BWA-MEM. Also, BWA-MEM resulted in lower

coverages than Bowtie2. This is caused by the two different algorithms behind the mapping strategies. BWA-MEM excludes more reads from the mapping, when the base-call accuracy at their ends is below the threshold. Although this rigorous mapping may lead to detection of less SNPs, as a lower number of variant reads are mapped against the reference sequence, it also guarantees high quality in detected SNPs. In contrast Bowtie2, maps more reads at certain positions, as the ends of reads with low quality are clipped. On the one hand, this increases the coverage artificially. However, it also harbors the possibility of detecting false positive SNPs. The differences in the two strategies demonstrated that a clear SNP detection and variant quantification could not be performed with the data given for AgseNPV-B ORI and the passages PP2 #1 to #10. Furthermore, almost all SNPs were localized in *hr1*. In contrast to other hrs in the AgseNPV-B genome *hr1* is relatively short, but still has been reported as a similar hotspot for recombination as other *hrs* in AgseNPV-B, also for flanking open reading frames (Wennmann et al., 2015a). Biologically, this underlines the role of hrs and their lax sequence structure consisting of several repeat units to promote mutation as recombination events and point mutations, besides being described as possible origins of replication and enhancers of transcription in baculoviruses (Guarino et al., 1986; Pearson and Rohrmann, 1995). However, the repeat sequences within one single *hr* and the identity of different *hrs* also contributes to misplacing sequencing reads when mapping against the reference sequence. In case of this sequencing approach, the inserts were generated on random base by a Nextera transposase kit prior sequencing (pers communication StarSeq). Therefore, the resulting sequencing reads could not be attributed to a certain region in the genome sequence, e.g. clearly to the region of *hr1* or a different *hr*. Moreover, the detected variant frequencies, e.g. in *vp91*, cannot be explained by a biological background. In theory, when considering a uniform MOI of 1 plaque forming unit per cell, every cell is infected by one budded virus. Thus, a mutation in *vp91* would have to occur independently in the first round of replication in every infected cell. However, the variant frequencies in *vp91* and *vef-3* but also in the *hrs* again decreased in the consecutive passages. Still, as only a small number of SNPs was detected and, it can be generally concluded that AgseNPV-B ORI and PP2 consist of homogenous genomes and passaging of PP2 has rather no effect on the occurrence of point mutations when passed in a MOI of 1 in AiE1611T.Alongside the absence of *fp*-mutants and the lack of SNPs also major genomic deletions were not observed. Such deletions would have led to the accumulation of defective interfering particles (DIP). Normally, DIPs lack large portions of the genome, which feature a benefit in replication, but they depend on the presence of complete genomes as helpers in replication or represent genomes lacking genes not essential for *in vitro* replication (Heldens et al., 1996; Kool et al., 1991). The generation of DIPs is enhanced when high MOIs (>5) are employed because the likely-hood of co-replication of DPIs and full genomes in single cells is favored compared to low MOIs when replication of DPIs cannot be rescued by trans factors provided by full viral genomes (Kool et al., 1991; Zwart et al., 2008).

In terms of NGS quantification, such accumulation of DIPs would have been evident, when sequencing coverages of larger genomic regions have dropped from PP2 #1 to #10 compared to global mean coverage. Although highly scattering coverage lines were noticed, in particular in #5 and #7, no systematic change in the coverage could be detected through the series of passages of PP2. Dropping sequence coverages as an indicator for a large consecutive deletion was observed in a deletion mutant recombinant AgseNPV-B bacmid: bAgseNPV-B (Chapter V). In bAgseNPV-B 42,443 bp are deleted between genome position 6,290 and 48,733,

resulting in a sequence coverage converging to zero when mapped against the AgseNPV-B genome sequence. In the serial passages, the observed genome stability might have been a consequence of a moderate MOI (MOI = 1) employed. A similar effect is described in passages with lower MOI for the production of a LdMNPV clone that does not accumulate fp-mutants in serial infections of the host cell culture Ld652Y (Slavicek et al., 2001). In this system, extensive plaque purification assays were assessed at a MOI of 0.1 in order to reach a stable LdMNPV clone. Apparently, the system of PP2 in serial infections in AiE1611T cells achieved a similar equilibrium with a ten times higher employed MOI.

The focus of this study was on the development of a data-based pipeline to detect crucial mutations, without depending on RFLP analyses or cloning and sequencing of genes of interest. In this context NGS methods, such as Solexa Illumina, which relies on the 'sequencing by synthesis' method, or even latest state of the art methods as nanopore sequencing (e.g. MinION sequencing), have introduced sequence data-based methods to investigate baculovirus diversity, as well as micro-evolutionary and evolutionary processes (Chateigner et al., 2015; Gueli Alletti et al., 2017c; van der Merwe et al., 2017). Solexa Illumina generates data with reliable read depths, which means that even SNP distribution can be resolved in high genome fold coverages. Thus, genetic lineages of mutations and their relative abundances can be directly correlated the sequencing reads of a virus population, as shown for the genomic constitution of different Cydia pomonella granulovirus (CpGV) isolates (Gueli Alletti et al., 2017c). One advantage of MinION sequencing over Solexa Illumina is that the DNA molecule does not need to be fragmented in small pieces of sequencing inserts. However, this sequencing method has been so far used only in combination with Solexa Illumina sequencing, as demonstrated for geographically and genetically diverse Cryptophlebia leucotreta granulovirus (CrleGV) isolates from South Africa (van der Merwe et al., 2017) and did not allow a quantification of SNPs in different isolates. The advantage of the NGS approach used in this study was to detect and to quantify small variations, caused from serial passaging steps at moderated MOIs. Combined with the stable production of OBs in cell culture and its relatively high genome stability, AgseNPV-B is a tentative candidate for the production of a biological control agent in AiE1611T cells.

Chapter V: Solexa Illumina Sequencing of bAgseNPV-B reveals a deletion mutant bacmid of Agrotis segetum nucleopolyhedrovirus B

Abstract

To make the AgseNPV-B genome accessible to genetic modifications, the construction of a bacmid carrying the Geneome of AgseNPV-B was envisaged. This bacmid, termed bAgseNPV-B was generated by homologous recombination in the *homologous repeat sequence (hrs) hr*6. Though bAgseNPV-B DNA replicated in *E. coli*, it lacked infectivity and viral replication in cell lines. Solexa Illumina Sequencing revealed that bAgseNPV-B represented a defective virus of AgseNPV-B with a deletion of 42,443 bp in the AgseNPV-B genome but a correctly inserted bacterial cassette. The deletion affected 42 *orfs* from the ORFs *asb005* to *adb046* and two *hrs*. AiE1611T cells can be transfected with DNA of bAgseNPV-B and show cytopathological effects, however the infection is blocked at an early stage with missing DNA replication correlated to the deletion of *lef-1*, *lef-2* and *me53* and no spreading of virus infection was observed.

Introduction

In the terminology of baculovirus molecular biology, bacmids are defined as recombinant DNA deriving from baculoviruses incorporated with bacterial plasmids in their genomes that replicate either in insect cell cultures and in bacteria (O'Reilly et al., 1994). First generations of recombinants consisted of deletion mutants of *p10* and/or *polyhedrin*, both genes with strong baculovirus promotors in insect cell cultures. Although *p10* is not a major occlusion body (OB) protein, deletion of this gene also alters the correct expression of the major occlusion body protein Polyhedrin. Phenotypes of such deletion mutants are defined as *occ*$^{-}$, according to their lack to produce OBs, but importantly the insertion of genes of interest into these sequences is consequently regulated by the *polyhedrin* and/or *p10* promotors. Both promotors are active at the terminal stage during infection and are regulated by transcriptional activators, like the very late expression factor 1 (VLF-1). Furthermore, at the very late stage of infection, genes are transcribed at a high rate due to efficiency of the baculovirus RNA polymerase and LEF4, which ensures an efficient capping, transport and translation. Thus, replacing the *polyhedrin* or *p10* ORF with a heterologous ORF of interest under the *polyhedrin* or *p10* promoter, results in a high and efficient expression of that ORF of interest (Merryweather-Clarke et al., 1994). A bacmid is therefore a circular covalently closed DNA molecule with origins of replication (ORIs) allowing to replicate in *E. coli* as well as in insect cells. This facilitates the manipulation of baculovirus genomes as the bacmid itself can function as acceptor of heterologous ORFs in *E. coli*, but the expression can be perfomed in insect cell culture. Such bacmids therefore consist of a (complete) baculovirus genome with a bacterial cassette containg an ori, regulatory sequences as well as an antibiotic resistance gene as a selection marker inserted ideally in an intergenic region (Hilton et al., 2008). Such plasmids must contain a bacterial ori, which allows a complete replication of the recombinant DNA, exceeding the size of >100 kb and a selection marker, such as genes for antibiotic resistance. If a gene of interest should be inserted into the plasmid, a system of the introduction of that gene is needed. This can either be managed by directed restriction of the bacmid in that region and ligation with the gene of interest, by homologous recombination or by inserting the gene of interest by transposition. Within the available Bac-to-Bac system (©ThermoFisher Scientific) a transposase helper plasmid, a donor

transfer plasmid and a recognition site for the transposition in the bacmid can be used readily for the transposition of a gene of interest (Figure 1). Correct insertion should be followed by reporter genes such as the common bacterial beta-galactosidase gene *lacZ* (Pijlman et al., 2002). If the gene of interest is not inserted with a baculovirus promotor, additional insertion of such a promotor is deemed necessary for the latter expression in insect cell culture. Such a plasmid can be of course introduced itself into the baculovirus genome by restriction and ligation with specific enzymes, however in that case both the baculovirus genome and the plasmid need unique restriction site that allow a directed ligation. In case of bAgseNPV-B a different approach was followed, although two restriction enzymes, namely *Sbf*I and *Avr*II were identified as unique cutters in the baculovirus genome of AgseNPV-B (KM102981) (Wennmann et al., 2015a). The recognition sites for *Sbf*I and *Avr*II are localized in the genes *chitinase* and *vef*-3, respectively. Albeit these genes might not be essential for AgseNPV-B, a restriction/ligation approach would have included the recovery of these genes if the deletion and/or truncation should be avoided. However, as previously demonstrated transfection of the insect cell line AiE1611T with genomic AgseNPV-B DNA concluded in both production of budded virus and OBs which can be assessed for the infection of more AiE1611T cells or larvae, respectively. Furthermore, continuous passaging of AgseNPV-B in the cell line AiE1611T does not accumulate defective genomes as demonstrated in ten passages of AgseNPV-B plaque purified clone (Chapter IV). This allowed an approach including the insertion of the plasmid pBAC-asb134/135 (Figure 2) into the genome of AgseNPV-B by homologous recombination in the laboratory workhorse *Escherichia coli* strain DH5α.

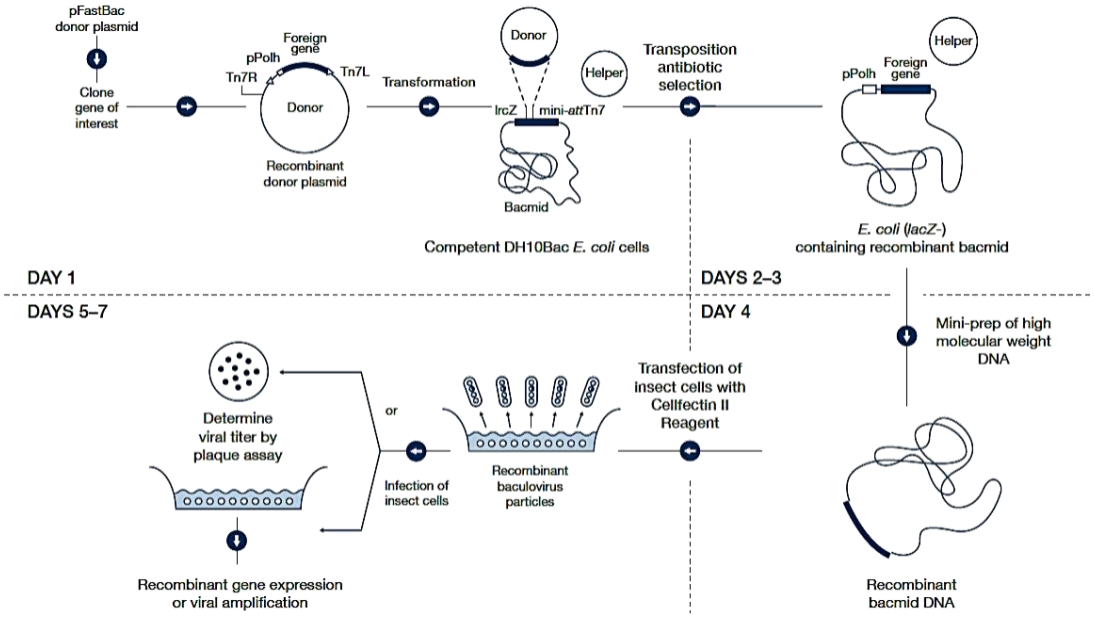

Figure 1 Bac-to-Bac system as commercially available by ThermoFisher Scientific. A gene of interest is cloned into the pFastBac donor plasmid from which it is moved into the bacmid by transposition with the aid of the helper plasmid expressing Tn7-transposase. The bacmid containing the inserted gene is amplified in *E. coli* been selected by blue/white selection due to the *lacZ* reporter gene. Amplified bacmids are extracted and used to transfect insect-cell cultures, which produce recombinant virus and can be readily used for the recombinant gene expression.

Material and Methods

Construction of the plamid pBAC-asb134/135

The plasmid pBAC-asb134/135 was constructed by inserting the genomic region of AgseNPV-B spanning over the two open reading frames *asb134* and *asb135* as well as the homologous region *hr6* which is flanked by the ORFs utilizing directed ligation into the unique restriction site for *Bsu*36I of the plasmid pBAC-BB which was kindly provided by Monique van Oers from the University of Wageningen, Netherlands (Pijlman et al., 2002), comprising a mini-F replicon for autonomous replication in *E. coli*, a kanamycin resistance gene as selection marker and a Tn7 recognition site for transposition within the reporter gene lacZ.

The two AgseNPV-B specific ORFs *asb134* and *asb135* flanking *hr6* in the AgseNPV-B genome were PCR amplified individually using primers up- and downstream from these ORFs. In second step, the two individual amplicons were fused in a second PCR amplification, which generated a fragment of 1,251 bp containg *asb134* and *asb135* connected to a 309 bp spacer sequence, but not *hr6*. The primer pairs of this second PCR amplifaction were constructed containing the recognition site for the restriction endonuclease *Bsu*36I. This fragment was consequently column-purified and inserted into the plasmid pBAC-Bsu36I (Pijlman et al., 2002) by restriction with *Bsu*36I and inserted using T4 DNA ligase (Figure 2) (Thermo Fisher Scientific). Bacterial cultures of *Eschericha coli* DH 5α were transformed by electroporation (Sambrook and Russell, 2001) and positive clones were identified by selection on kanamycine and PCR amplification of *asb134*. Positive clones were amplified in LB medium overnight and plasmid DNA was extracted using a column-based midi prepation kit (Promega).

Generation of bAgseNPV-B in AiE1611T and amplification in *E. coli* DH5 α cells

Each 1 µg of DNA of AgseNPV-B and pBAC-asb134/135 were used to transfect about 1×10^6 AiE1611T cells with the aid of Lipofectamin® (Invitrogene) (Felgner et al., 1987). Infection was checked visually on a daily for cytopathological effects (CPE) and the formation of occlusion bodies (*occ*$^+$) by comparing to control wells treated with Lipofectamin® and water only. Cell debris and from supernatant and suspended in TEK-buffer (1 M Tris-EDTA pH 7.5, 0.15 M KCl). After an incubation at 65 °C for 30 min, sodium dodecyl sulfate and proteinase K were added in final concentrations of 3% and 0.2 M, respectively, followed by a further incubation at 65 °C for 30 min. DNA was purified by phenol/chloroform extraction and ethanol precipitation (Gross-Bellard et al., 1973; Sambrook and Russell, 2001). The extracted DNA was used to transform *E. coli* DH5 α cells by electroporation (Sambrook and Russell, 2001). Clones carrying bAgseNPV-B were selected by growth on kanamycine as well as by the PCR of *asb134* and *polyhedrin*. DNA of these clones was isolated and used for the comparison of restriction digests with *Eco*RI and *Hin*dIII with *in silico* deduced digests of the putative bAgseNPV-B sequence.

Detection of deletions, insertions and single nucleotide polymorphisms to bAgseNPV-B

DNA was isolated from bAgseNPV-B positive clones using a column-based midi prepation kit (Promega). About 50 ng purified DNA samples of bAgseNPV-B were subjected to NexteraXT library preparation and an Illumina NextSeq500 paired-end sequencing (StarSEQ Ltd., Mainz Germany). The sequencing produced approximately 1.0 – 2.5 Mio reads with an average length of 151 nucleotides (nt), resulting in an average of about 1000- to 2000-fold genome coverage when a genome size of 157,042 bp for nAgseNPV-B is assumed. Read-pairs with 50%

consecutive bases below an average Phred quality score of 30 per read cycle (99.9% base call accuracy) were excluded from the analyses (Gordon, 2009). Obtained read-pairs were re-mapped against the *in silico* deduced bAgseNPV-B sequence using the Bowtie2 aligner for short reads with standard parameters for very-sensitive local alignment of reads on the JKI Galaxy server (Afgan et al., 2016a; Langmead and Salzberg, 2012a; Mielczarek and Szyda, 2016; Ye et al., 2015; Ziemann, 2016). Insertions and deletions (InDel) were identified by evaluating the coverage of read pairs in the Bowtie2-mapping. Single nucleotide polymorphisms (SNPs) and InDels of few base pairs were detected with the SAMtools MPileup tool (Li et al., 2009). The average threshold of the quality of mutations in the alignment was set to a Phred-scaled probability of 20 (99% base call accuracy in the alignment) or higher in order to take into account possible errors due to Illumina sequencing and Bowtie2 (Li, 2011a; Li, 2011b; Li and Durbin, 2009). The statistical analyses were assessed using the R-package "Bioconductor" (MacQueen, 1966; R Development Core Team, 2015; Robinson, 2016).

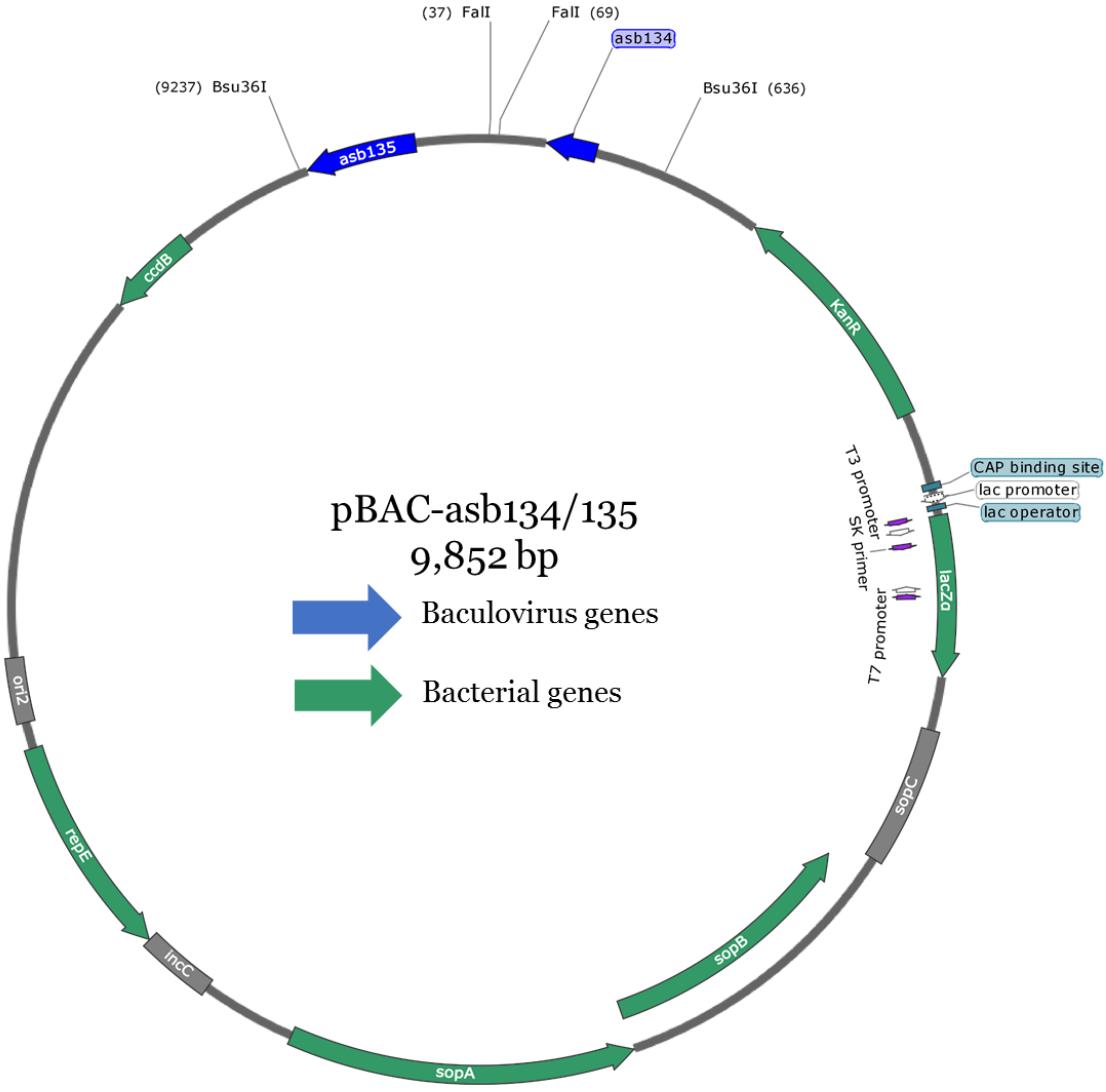

Figure 2 Bacmid-Cassette – pBAC-asb134/135: Plasmid backbone derived from pBAC-BB (Pijlman et al., 2002). Insert consisting of *asb134* and *asb135* with 309 bp spacer in between was added by restriction with *Bsu*36I and successive ligation. The bacterial backbone comprises a mini-F replicon (replication in *E. coli*) and a Kanamycin resistance gene (selection on Kanamycin).

Results

Generation of bAsgeNPV-B

The successful homologous recombination was screened by PCR of the baculovirus gene *polyhedrin* in kanamycin-resistant clones and DNA of these clones was extracted and submitted to digestions with the restriction enzymes *Hind*III and *Eco*RI (data not shown) to size compare the fragments to *in silico* digestions of AgseNPV-B (KM102981) as well as the theoretical sequence of bAgseNPV-B (Figure 3) in case of a successful homologous recombination. The correct insertion of the plasmid was checked by PCR and subsequent Sanger sequencing with primers flanking the regions of homologous recombination. In that way, the plasmid pBAC-asb134/135 was inserted in the AgseNPV-B genome at the predicted position, which also concluded in deletion of *hr*6. The initial euphoria for the homologous recombination was dulled, when transfected AiE1611T cells showed CPE without formation of OBs, but above all when the restriction patterns did not match to either the sequence of AgseNPV-B nor to the theoretical sequence of bAgseNPV-B. By comparison with the *in silico* digestions the resulting bAgseNPV-B was lacking a certain amount of its genome size. However, the restriction analyses did not make clear whether this aberration was caused by a single deletion, or by several deletions. In order to gain information on the deleted region and which part of the AgesNPV-B genome sequence (or even pBAC-asb134/135) was affected, the DNA samples of one clone were submitted for sequencing with Solexa Illumina (StarSeq GmbH, Mainz, Germany) with rigorous successive analyses of the obtained sequencing products.

Deleted *open reading frames* (*orfs*) and *homologous regions* (*hrs*) in bAgseNPV-B

Sequencing of the bAgseNPV-B sample generated a total number of 1,998,835 reads with a base call accuracy of 99.9 (Q30) or above. In total 1,158,923 of these reads (58%) mapped against the theoretical genome sequence of bAgseNPV-B (Figure 3), 839,912 reads remained unmapped (42%). The mapping consulted in a global 1,520-fold mean coverage. As observed in the analyses of digestions of bAgseNPV-B DNA with *Eco*RI and *Hind*III (data not shown), a deletion of 42,443 bp was identified between the genome positions 6,290 and 48,733 (Figure 4). This deletion spanned over 27% of the full theoretical genome sequence and included the deletion of 42 *orfs* as well as the two *homologous regions hr1* and *hr2* (Table 1). The deleted *orfs* included the six baculovirus core genes *odv-e56*, *lef-1*, *lef-2*, *pif-1*, *pif-2* and *alk-exo*, three *orfs* specific for the genera *Alpha-* and *Betabaculovirus* (*me53*, *p24*, *gp37*), 12 *orfs* with orthologues in other baculovirus genomes as well as 21 unique AgseNPV-B *orfs*. Furthermore, as observed in the cell culture passages of AgseNPV-B PP2 (see chapter IV), a deletion of three nucleotides, namely of one CGA repetition, was identified between the genome positions 102,240 and 102,242.

Detection of single nucleotides polymorphisms (SNPs) and deletions of few nucleotides

The mapped reads were subsequently used in analyses with MPileup to detect single nucleotide polymorphisms (SNPs) and deletions of few nucleotides. For quality reasons the mutations were filtered according to their mapping quality and only mutations with a Phred quality score of 20 or above (99% accuracy) were analyzed. In total five transitions, three transversions and one deletion were detected (Table 2). In *hr3*, three transitions and three transversions were detected ranging in their variant frequency from 48% to 68%. The single deletion of two nucleotides was detected in the AgseNPV-B unique orf *asb096* with 100% variant frequency.

The remaining transitions were identified in the bacterial cassette in intergenic regions with a variant frequency of 100%.

Table 4 Deleted open reading frames (*orf*) and homologous regions (*hrs*) in bAgseNPV-B. Given are the annotated *orf* (1), the conserved occurrence (baculovirus core genes = α+β+γ+δ, *Alpha-* and *Betabaculovirus* = α+β, *Alpha-*, *Beta-* and *Deltabaculovirus* = α+β+δ, orthologues in other baculoviruses = bac. ortho. and AgseNPV-B unique genes = unique) (2), the genome sequence position in AgseNPV-B and orientation relative to *polyhedrin* (3) and the amino acid (aa) number of predicted gene products (4)

(1) *orf*/*hrs*	(2) Conservation	(3) Position	(4) aa
asb005[A]	unique	5877 ← 6329	151
asb006	unique	6350 → 7114	255
odv-e56	α+β+γ+δ	7162 → 8268	369
me53	α+β	8592 → 9653	354
efp	α+β+ δ	10,535 ← 12,589	685
asb010	unique	12,740 ← 13,699	320
gp16	bac. ortho.	13,787 ← 14,080	98
p24	α+β	14,107 ← 14,817	237
asb013	unique	14,896 → 15,255	120
lef-2	α+β+γ+δ	15,212 → 15,856	215
asb015	unique	15,870 ← 16,043	58
asb016	unique	16,064 → 16,540	159
asb017	unique	16,738 → 17,493	252
asb018	unique	17,650 ← 18,333	228
hr1	n.a.	18,428 .. 18,683	n.a.
38.7k	bac. ortho.	18,750 ← 19,856	369
lef-1	α+β+γ+δ	19,858 ← 20,517	220
cathepsin	bac. ortho.	20,779 ← 21,927	383
he65	bac. ortho.	22,022 → 23,644	541
asb023	unique	23,724 ← 24,155	144
chitinase	bac. ortho.	23,346 → 26,091	582
asb025	unique	26,149 ← 26,451	101
gp37	α+β	26,492 → 27,274	261
ptp-2	bac. ortho.	27,271 ← 27,765	165
egt	bac. ortho.	27,926 → 29,530	535
asb029	unique	29,753 → 30,289	179
asb030	unique	30,195 → 30,968	258
asb031	unique	31,040 ← 33,757	906
asb032	unique	34,068 → 34,610	181
pkip-1	bac. ortho.	34,718 → 35,215	166
asb034	unique	35,409 ← 35,741	111
arif-1	bac. ortho.	35,746 ← 36,549	268
pif-2	α+β+γ+δ	36,464 → 37,699	412
pif-1	α+β+γ+δ	37,726 → 39,369	548
asb038	unique	39,425 → 39,688	88
fgf	bac. ortho.	39,725 ← 40,828	368
asb040	unique	41,173 → 41,901	243
alk-exo	α+β+γ+δ	41,956 ← 43,173	406
hr2	n.a.	43,213 .. 43,593	n.a.
asb042	unique	44,327 ← 44,653	109
asb043	unique	44,667 → 45,839	391
asb044	unique	45,882 ← 46,283	134
rr2B	bac. ortho.	46,380 → 47,321	314
asb046[B]	unique	47,498 → 48,892	465

Table 4 Deleted open reading frames (*orf*) and homologous regions (*hrs*) in bAgseNPV-B. Given are the annotated *orf* (1), the conserved occurrence (baculovirus core genes = α+β+γ+δ, *Alpha-* and *Betabaculovirus* = α+β, *Alpha-*, *Beta-* and *Deltabaculovirus* = α+β+δ, orthologues in other baculoviruses = bac. ortho. and AgseNPV-B unique genes = unique) (2), the genome sequence position in AgseNPV-B and orientation relative to *polyhedrin* (3) and the amino acid (aa) number of predicted gene products (4)

[A] asb005 is truncated by a deletion of 40 bp
[B] asb046 is truncated by a deletion of 1,236 bp

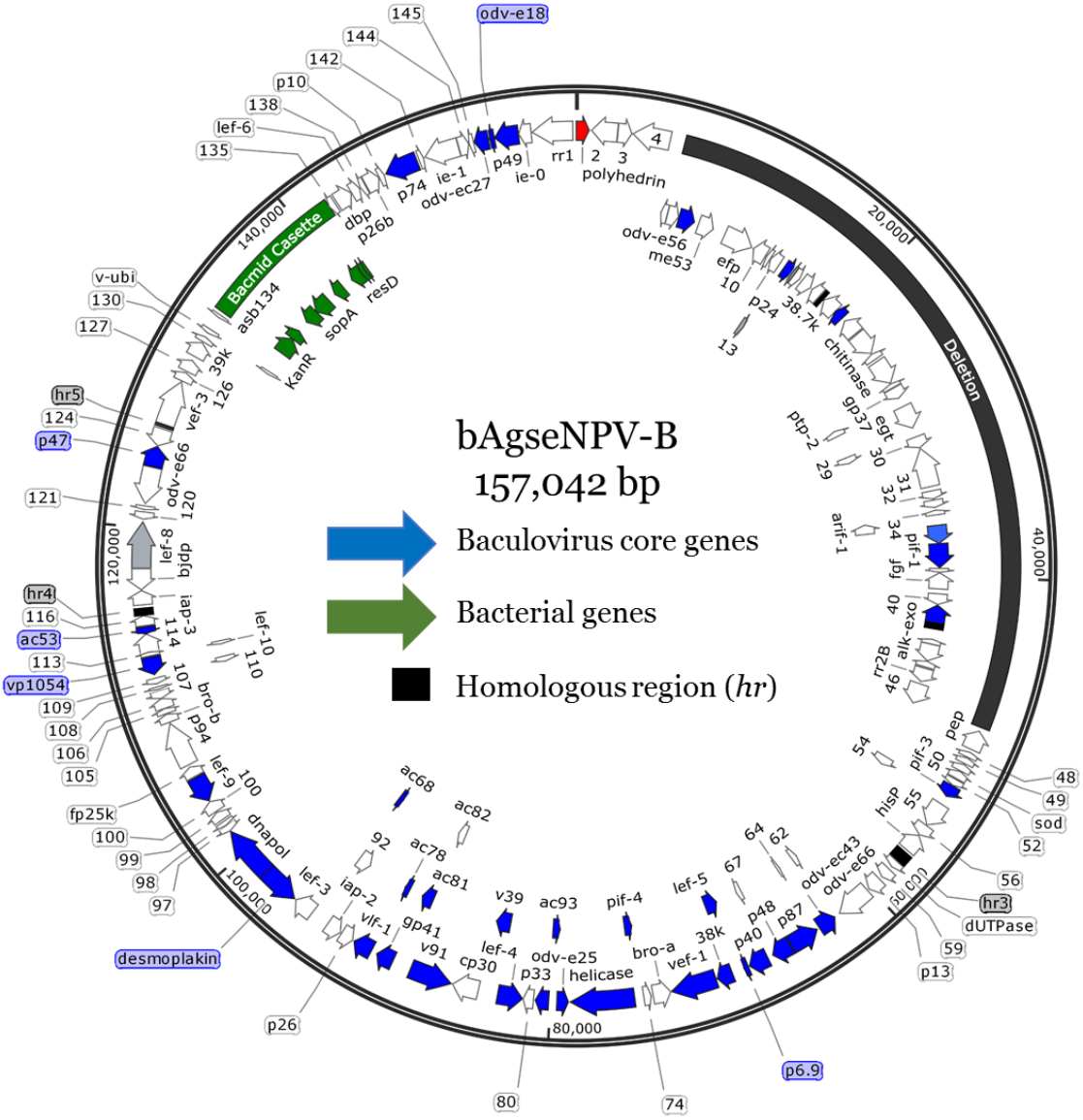

Figure 3 Genome sequence of bAgseNPV-B consisting of baculovirus genome and pBAC-asb134/135 (Bacmid Cassette) inserted into the genome by homologous recombination. The deletion of 42,443 bp of the AgseNPV-B genome is highlighted as black bar in the genome sequence.

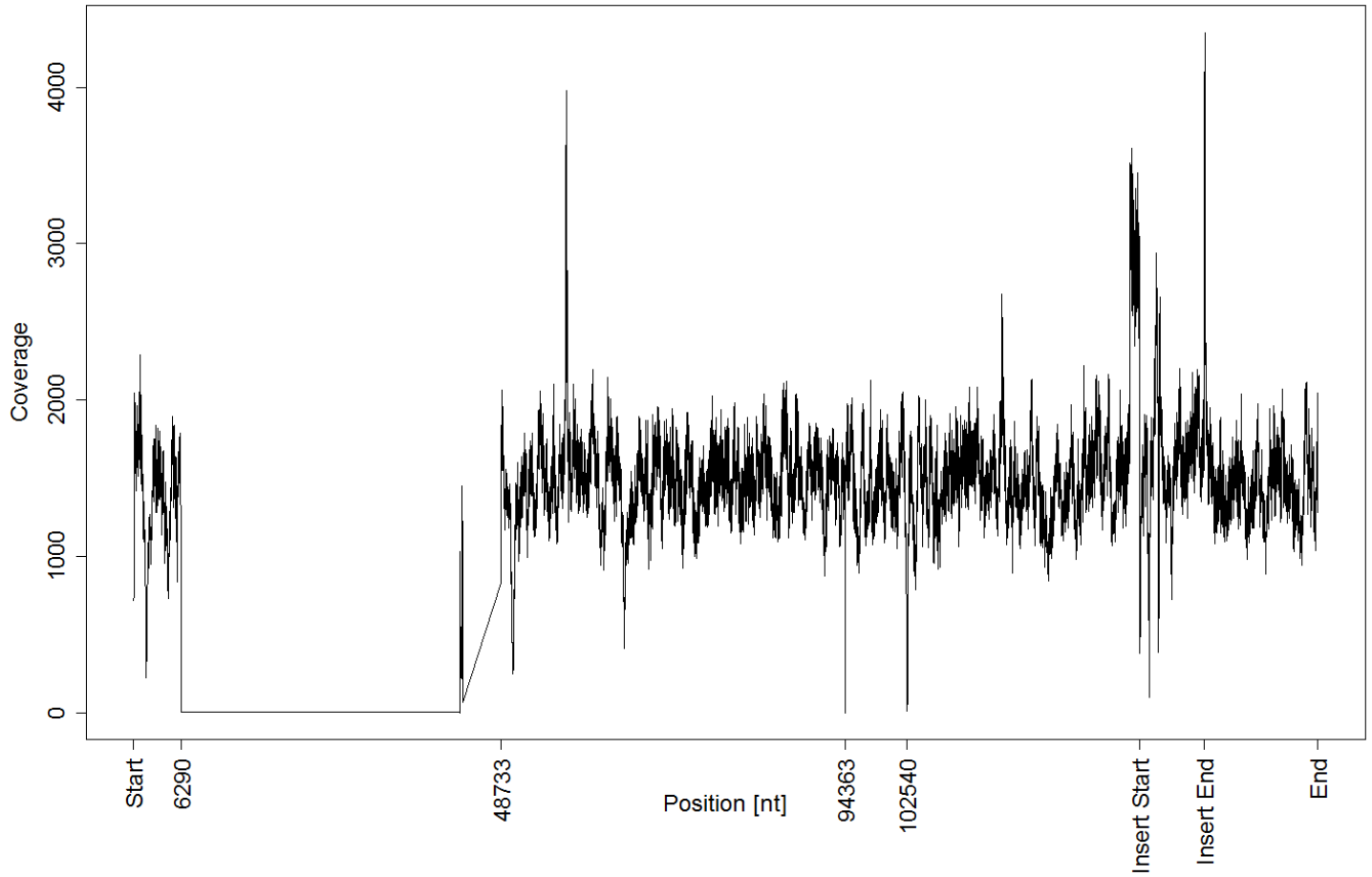

Figure 4 Coverage Bowtie2-Alignment against bAgseNPV-B genome sequence. In total 42,443 bp (6,290 .. 48,733) were deleted from the genome (inbreak of coverage). One additional dinucleotide deletion with a Phred-Quality score >30 was identified between 94,363 .. 94,634 in analyses of the Bowtie2-BAM file using MPileup. Additionally, one deletion of three nucleotides was identified between 102,240 .. 102,242 by analyses of the coverage. The Bacmid-Cassette was inserted between 133,338..141,984 bp.

Table 5 Single nucleotide polymorphisms and deletions observed between the bAgseNPV-B reference genome and the Bowtie2 assemblies identified with MPileup. The mutations were quality filtered by their mapping quality (Phred quality ≥ 20) with a resulting base call accuracy of 99% or greater. Given are the nucleotide position in bAgseNPV-B (1), the affected open reading frame (*orf*) or homologous region (*hrs*) (2), the mutation type (3), the nucleotide change (4) and the variant frequency of the mutation in % (6).

(1) Position	(2) *orf/hrs*	(3) Type	(4) Change	(6) Variant Frequency [%]
57,185	*hr3*	Transition	A → G	68
57,188	*hr3*	Transition	A → G	57
57,195	*hr3*	Transversion	A → T	68
57,196	*hr3*	Transversion	T → A	50
57,203	*hr3*	Transversion	T → A	48
57,208	*hr3*	Transition	T → C	65
94,363	*asb096*	Deletion	CGCCGCAC → CCGCAC	100
139,649	n.a.	Transition	T → C	100
140,220	n.a.	Transition	A → G	100

Discussion

The selection for bAgseNPV-B, as being a recombinant AgseNPV-B isolate with an insertion of the bacterial cassette pBAC-asb134/135, was based on a PCR based detection of amplicons of the AgseNPV-B *polyhedrin* gene as well on screening of kanamycin resistant clones on selective medium. As *polyhedrin* was not affected by the deletion of consecutive 42 kilo base-pairs (kbp) in the AgseNPV-B genome, the selection method mentioned above was not sensitive enough to detect possible deletions of that size. The first signs of such a large deletion during RFLP analyses (data not shown) were confirmed by full genome sequencing of the selected bAgseNPV-B clone. NNGS data were further used to confirm the correct insertion of pBAC-asb134/135 into the AgseNPV-B genome between the nucleotide positions 133,338 and 141,984 bp. The insertion of the bacterial cassette already replaced one homologous region in the AgseNPV-B genome, namely *hr6*. In addition, *hr1* and *hr2 (position angeben)* were additionally deleted, as shown bythe NGS. Thus, bAgseNPV-B comprised deletions of three out of six recognized *hrs*, which are generally assumedas both possible origins of replications and transcriptional enhancers in baculoviruses (Kool et al., 1993; Pearson et al., 1992; Pijlman et al., 2002). However, as described for recombinant isolates of Autographa californica multiple nucleopolyhedrovirus (AcMNPV), deletion of individual *hrs* or cohorts of these sequences does not automatically alter the virus replication of these AcMNPV isolates in cell culture (Wu et al., 1999).

More importantly than the deletion of these three *hrs*, the deletion of 42 kbp also affected genes essential for the baculovirus replication and infection cycle. Baculovirus replication is characterized by the expression of a baculovirus specific DNA polymerase and its accessory proteins at a early stage of infection (Tomalski et al., 1986). Likewise to other DNA polymerases, DNA polymerization in baculoviruses requires free hydroxyl groups which are supplied by RNA primers (Watson et al., 2004). These RNA primers are catalyzed by enzymes referred to as “primase”, which are represented by LEF-1 in baculoviruses (Mikhailov and Rohrmann, 2002). The deletion of *lef-1* therefore concludes with fatal consequences for baculovirus replication, and deletion mutants can be considered as non-viable. While *lef-1* encodes for the primase, *lef-2* encodes for a DNA primase accessory factor, which interacts with both, the DNA matrix and LEF-1 (Kool et al., 1994). Although the deletion of *lef-2* can be

compensated in the replication, e.g. through DNA nicking, it also has an essential role in the transcription of late genes (Wu et al., 2010). For this reason, *lef-2* deleted AcMNPV isolates are considered as non-viable. The selected bAgseNPV-B isolate also lacks the gene *me53*, in addition to deletions of *lef-1* and *lef-2*. ME53 is involved in both the transcription of early and late genes (de Jong et al., 2011). Furthermore, cells infected with AcMNPV mutants lacking me53 show preliminary signs of infections, such as the occurrence of cytopathological effects (CPE) but do not produce viable virus progeny (Xi et al., 2007). These observations are consistent to those of bAgseNPV-B: AiE1611T cells that were transfected with DNA of bAgseNPV-B showed CPE after 10 to 14 days post transfection. However, neither did the infection spread to the surrounding cells, nor did transfected cells produce occlusion bodies (OBs). Therefore, deletion of *me53* may already block the early stage of infection, which is characterized by the formation of budded virus (BV) and infections spreading from cell to cell. However, the missing infectivity in cell culture cannot be correlated to *me*53 only, as 42 *orfs* and two *hrs* had been deleted. Beyond the deletions of genes essential in replication and regulation of genes, possible bAgseNPV-B OBs also might lack essential structural proteins, so called "*per os* infectivity factors" (pifs) as *pif-1, pif-2* and *odv-e56* (*pif-5*) were deleted in the genome. These structural proteins are associated with occlusion derived virion (ODV) and are responsible for an infection of host larvae. Summarized, we found no evidence that bAgseNPV-B was capable for replication in insect cells and the infection might be blocked inter-cellular in AiE1611T cells and in host larvae.

In this context bAgseNPV-B represents an isolate similar to defective interfering viruses (DIs). DIs are commonly produced in baculovirus populations. They have been classically observed in continuous passaging through permissive cells, where they can appear rapidly when high multiplicities of infection are employed (Pijlman et al., 2001). DIs are characterized by large deletions in their genomes, which can affect regulatory and structural genes. Therefore, DIs may lack the ability to replicate on their own and always require fully functional genomes to compensate the deleted genes. DIs have also been observed in natural populations of baculoviruses like for Spodoptera frugiperda nucleopolyhedrovirus (SfMNPV) isolates (Simon et al., 2010). Here, larvae of the fall armyworm Spodoptera frugiperda were infected with mixed ratios of SfMNPV full genomes and SfMNPV DIs and the virus progenies produced were compared to ratios of DIs to full genomes observed in natural samples of SfMNPV. Interestingly, the ratio of DIs to full genomes always reached the identical numbers after several passages through the host larvae as in natural populations of of SfMNPV. However, the co-replication with a fully functional virus as helper has still to be tested for bAgseNPV-B. This could be readily performed in AiE1611T, since this cell culture is permissive for AgseNPV-B and a broad range of alphabaculoviruses. The bacterial cassette inserted in bAgseNPV-B allows further to assess the function of deleted genes, e.g. *me*53 more in detail for AgseNPV-B or to insert the deleted sequences since with pBAC-asb134/135 a fully functional Tn7 insertion sites is integrated into bAgseNPV-B..

Chapter VI: Next Generation Sequencing to identify and quantify the genetic composition of resistance-breaking commercial isolates of *Cydia pomonella granulovirus*

This chapter is published with modifications in:
Gueli Alletti, G., Sauer, A. J., Weihrauch, B., Fritsch E., Undorf-Spahn, K., Wennmann, J. T., Jehle, J. A. (2017), Using Next Generation Sequencing to Identify and Quantify the Genetic Composition of Resistance-Breaking Commercial Isolates of Cydia pomonella Granulovirus, Viruses, 9(9): e250

Abstract

The use of Cydia pomonella granulovirus (CpGV) isolates as biological control agents of codling moth (CM) larvae is important in organic and integrated pome fruit production worldwide. The commercially available isolates CpGV-0006, CpGV-R5, and CpGV-V15 have been selected for the control of CpGV resistant CM populations in Europe. In infection experiments, CpGV-0006 and CpGV-R5 were able to break type I resistance and to a lower extent also type III resistance, whereas CpGV-V15 overcame type I and the rarely occurring type II and type III resistance. The genetic background of the three isolates was investigated with next generation sequencing (NGS) tools by comparing their nucleotide compositions to whole genome alignments of five CpGV isolates representing the known genetic diversity of the CpGV genome groups A to E. Based on the distribution of single nucleotide polymorphisms (SNPs) in Illumina sequencing reads, we found that the two isolates CpGV-0006 and CpGV-R5 have highly similar genome group compositions, consisting of about two thirds of the CpGV genome group E and one third of genome group A. In contrast, CpGV-V15 is composed of equal parts of CpGV genome group B and E. According to the identified genetic composition of these isolates, their efficacy towards different resistance types can be explained and predictions on the success of resistance management strategies in resistant CM populations can be made.

Introduction

Cydia pomonella granulovirus (CpGV) belongs to the genus *Betabaculovirus* in the family of *Baculoviridae* (Herniou et al., 2011). A number of isolates from different geographic regions have been found since its first detection in Mexico (Mexican isolate, CpGV-M) (Tanada, 1964; Tanada and Leutenegger, 1968). Depending on the isolate, the circular dsDNA genome of CpGV ranges from 120.8 to 124.3 kbp, encoding 137 to 142 open reading frames (ORFs) (Gebhardt et al., 2014; Wennmann et al., 2017; Wormleaton and Winstanley, 2001). Based on phylogenetic analyses of their genome sequences, all known CpGV isolates can be classified into five genome groups A to E, describing different phylogenetic lineages. Representative isolates are CpGV-M (genome group A), CpGV-E2 (B), CpGV-I07 (C), CpGV-I12 (D), and CpGV-S (E) (Eberle et al., 2009; Gebhardt et al., 2014). Although these isolates share a highly conserved genome architecture with similar sizes, collinear ORF arrangements, and %GC contents (Gebhardt et al., 2014), differences have been noted by restriction length polymorphisms (RFLPs) (Crook et al., 1985; Eberle et al., 2009; Rezapanah et al., 2008), single nucleotide polymorphisms (SNPs) in conserved genes (Eberle et al., 2009), and finally by their genome sequences (Crook et al., 1985; Eberle et al., 2009; Garavaglia et al., 2012; Gebhardt et al., 2014; Miele et al., 2011). Cydia pomonella granulovirus has a narrow host range and is

highly virulent against the Lepidopteran pest species *Cydia pomonella* (codling moth, CM) and to a lower extent to a very few closely related Tortricids (Lacey et al., 2005; Zingg et al., 2012). Known as a fast-killing granulovirus, neonate CM larvae succumb within four to six days after infection with CpGV. Because of these characteristics, CpGV has been developed and intensively used as a commercial biocontrol agent of CM in virtually all pome fruit production areas since its first registration in Switzerland in 1989. Most of these commercial products were based on the isolate CpGV-M (Asser-Kaiser et al., 2007).

Cases of laboratory selected resistance of insects to baculoviruses have only rarely been described, e.g., for *Phthorimaea operculella*/PhopGV (Briese, 1982; Briese and Mende, 1983), *Anticarsia gemmatalis*/Anticarsia gemmatalis multiple nucleopolyhedrovirus (AgMNPV) (Abot et al., 1996), *Trichoplusia ni*/Trichoplusia ni single nucleopolyhedrovirus (TnSNPV) (Milks and Myers, 2000; Milks and Theilmann, 2000), and *Adoxophyes honmai*/Adoxophyes honmai nucleopolyhedrovius (AdhoNPV) (Iwata et al., 2017; Nakai et al., 2017). In 2005, however, the first cases of CM field populations with a more than 1000-fold reduced susceptibility to commercial CpGV products were reported from Germany and France (Fritsch et al., 2005; Sauphanor et al., 2006; Schmitt et al., 2013). For CpRR1, a genetically homogenous inbred strain which derived from a resistant CM field population from South Germany, it was shown that CpGV resistance is inherited by an incomplete dominant and monogenic mode that is linked to the Z chromosome (Asser-Kaiser et al., 2010; Asser-Kaiser et al., 2011; Zichová et al., 2013). Laboratory CM strains which derived from the French CM population RGV showed similar responses in full-range bioassays as CpRR1, although these strains have shown inheritance patterns that could not be fully explained by a Z-chromosmal linkage (Berling et al., 2009; Berling et al., 2013; Graillot et al., 2013). However, it was proposed to term this form of CpGV resistance as type I resistance (Jehle et al., 2017). Strikingly, type I resistance appeared to only be targeted against CpGV-M (genome group A), since CpGV isolates from genome groups B to E were shown to be resistance-breaking (Berling et al., 2009; Eberle et al., 2008; Gebhardt et al., 2014). Using molecular analyses, it was further shown that type I resistance in CpRR1 was targeted against the viral gene *pe38* of CpGV-M (Gebhardt et al., 2014; Jehle et al., 2017). Recently, two novel types of field resistances have been discovered: A proposed type II resistance appeared in a field population, termed NRW-WE, in North-West Germany (Jehle et al., 2017). Larvae of this population showed resistance not only against CpGV-M (genome group A), but also against CpGV-I07, -I12, and -S representing genome groups C to E, respectively; only CpGV-E2 (genome group B) appeared to overcome type II resistance (Jehle et al., 2017). Two laboratory strains, CpR5M and CpR5S, were selected from NRW-WE by exposing the offspring of five consecutive inbred mass crosses to either CpGV-M or CpGV-S, respectively (Sauer et al., 2017a; Sauer et al., 2017b). It was demonstrated that type II resistance followed a dominant, monogenic but autosomal inheritance pattern. Furthermore, a cross-resistance to at least two CpGV isolates, CpGV-M and CpGV-S, was observed (Sauer et al., 2017a). In addition, a further field population from Germany, termed SA-GO, possessed a third resistance type (type III), which is directed against CpGV isolates from the genome types A and E. Selection and crossing experiments indicated a highly complex polygenic inheritance pattern with some mixed characteristics of type I and type II resistance (Sauer et al., 2017c).

Because CpGV resistances are isolate dependent, novel commercial products based on diverse CpGV isolates have been tested in the field and were eventually registered (Berling et al., 2009; Brand et al., 2017; Zingg, 2010). Three different isolates, namely CpGV-0006, CpGV-R5, and CpGV-V15, are currently commercially available in Europe. To ensure their field efficacy and to develop optimum resistance management strategies (Union, 2009; Union, 2013), knowledge of their activity against different types of CpGV resistance, as well as their composition of genome groups, are essential. In the present study, we tested the activity of the commercial isolates CpGV-R5, CpGV-0006, and CpGV-V15 against different laboratory selected CM strains representing different known resistance types I to III. The currently known genetic diversity of CpGV can be differentiated by SNPs that are unique for single CpGV genome groups A to E and which are distributed across the genome (Wennmann et al., 2017). In combination with next generation sequencing techniques, the SNP distribution was used to identify and quantify the putative CpGV genome groups present in the commercial isolates.

Materials and Methods

Insects

Five different strains of *Cydia pomonella* (codling moth, CM) were maintained for experimental purposes at the Institute for Biological Control in Darmstadt; one strain being susceptible to all CpGV isolates, termed CpS, and four resistant strains termed CpRR1 (Asser-Kaiser et al., 2007), CpR5M, and CpR5S, as well as CpRGO (Sauer et al., 2017a). While CpRR1 exhibits type I resistance, CpR5M and CpR5S possess a type II resistance, and CpRGO a type III resistance. All CM strains were reared under the same laboratory conditions (Asser-Kaiser et al., 2007; Asser-Kaiser et al., 2011; Gebhardt et al., 2014). In brief, adult moths were kept at 26 °C, 60% relative humidity, and 16/8 h light/dark photoperiod for 10–14 days in groups of about 80–100 individuals in transparent plastic cylinders (14 cm diameter, 25 cm height), lined with transparent plastic sheets for oviposition. Egg sheets were incubated at 26 °C and immediately after hatching, neonate larvae were transferred to autoclavable 50-well plates containing a semi-artificial diet (Asser-Kaiser et al., 2011; Ivaldi-Sender, 1974). Insects of the last larval stage were allowed to pupate in corrugated cardboard stripes.

Viruses

All CpGV isolates were isolated from commercial CpGV products: The isolate CpGV-0006 derived from the product MadexMAX and CpGV-V15 from MadexTOP (both Andermatt Biocontrol, Stahlermatten, Switzerland). Isolate CpGV-R5 is the active ingredient of Carpovirusine EVO2 (Arysta Lifescience, Noguères, France) (Commission, 2008; EFSA, 2012). The CpGV occlusion bodies (OB) contained in the products were purified as described previously (Jehle et al., 1992) in order to avoid residues of pesticide formulations.

Resistance Testing

The mortality responses of the neonate larvae of each strain CpS, CpRR1, CpR5M, CpR5S, and CpRGO were tested by incorporating the purified OB of either CpGV-0006, CpGV-R6, or CpGV-V15 at the final discriminating concentration of 5.8×10^4 OB/mL into the semi-artificial diet. This concentration causes >95% mortality in susceptible CpS neonates after seven days (Asser-Kaiser et al., 2007). CpGV-V15 OBs derived from an unformulated test product were diluted according to the OB concentration determined by counting with a Petroff-Hauser counting chamber (2.5×10^{-3} mm^2 × 0.02 mm depth) in dark-field microscopy (Leica DM

RBE) in at least three independent replications. The isolates CpGV-V006 and CpGV-R5 derived from formulated commercial products were diluted according to the OB concentrations given on the label. Larvae that did not survive handling within the first 24 h were excluded from the analysis. The mortality rates of larvae were determined at seven and 14 days post infection (dpi) in three to five independent repetitions for each CpGV treatment. Each treatment included control groups with untreated neonates of CpS, CpRR1, CpR5M, CpR5S, and CpRGO. The mortality rates of the CpGV treatments were corrected for the corresponding control mortality according to Abbott (Abbott, 1925). The corrected mortality rates were used to compute the arithmetic mean mortality and the standard deviation (SD) of each treatment. Differences in the mean mortality responses were evaluated for significance ($p < 0.05$) using analysis of variance (ANOVA) and the Tukey's Honestly Significant Difference test (Tukey-HSD) comparison of means with standard R code (R version 3.3.1 in RStudio 1.0).

Alignment of CpGV Isolates

For the detection of SNPs, a ClustalW alignment of genome nucleotide sequences of five CpGV isolates, namely CpGV-M (KM217575), CpGV-E2 (KM217577), CpGV-I07 (KM217574), CpGV-I12 (KM217576), and CpGV-S (KM217573), was used. This alignment was used to infer SNPs specific for the CpGV genome groups A, B, C, D, and E (Eberle et al., 2009; Gebhardt et al., 2014), which were represented by the mentioned CpGV isolates.

DNA Extraction & Whole Genome Sequencing

For purposes of whole genome sequencing of commercial isolates CpGV-0006, CpGV-R6, or CpGV-V15, genomic DNA was isolated from CpGV OB as described previously (Arends and Jehle, 2006). The viral OB matrix was solubilized in 0.1 M Na_2CO_3 at 60 °C for 1 h. The suspension was adjusted to pH 8 by titrating with 1 M HCl, treated with RNaseA (90 µg/mL) at 37 °C for 10 min, and then with Proteinase K (250 µg/mL) and 1% SDS at 50 °C for a further 60 min. DNA was separated from protein debris by phenol/chloroform/isoamylalcohol (25:24:1, *v*/*v*) extraction (O'Reilly et al., 1994) using Phase Lock Gel Tubes (all purchased from, Carl Roth GmbH + Co. KG, Karlsruhe, Germany) in order to avoid phenol/protein contamination. The viral DNA was precipitated with ethanol and finally dissolved in ultra-pure water (Sambrook and Russell, 2001). DNA concentration and DNA purity was estimated by ultraviolet-visible absorbance spectroscopy (UV-Vis) with a NanoDrop 2000c spectrophotometer. Paired-end next-generation sequencing of 50 ng purified DNA each was performed by using a NexteraXT library preparation and an Illumina NextSeq500 sequencing system (StarSEQ Ltd., Mainz, Germany). The sequencing approach produced approximately 2.5 million paired end read-pairs per sample of 151 nucleotides in length.

Detection of Single Nucleotide Polymorphisms

For the sequence assembly of CpGV-0006, CpGV-R5, and CpGV-V15, the conducted paired-end reads were quality filtered excluding reads with less than a 50% average Phred quality score below 30 (base-call accuracy 99.9%) per read cycle (Gordon, 2009). The quality-filtered reads were re-mapped against the CpGV alignment consensus sequence using the Bowtie2 aligner (ver. 2.3.0, source code downloaded from http://bowtie-bio.sourceforge.net/bowtie2/index.shtml and compiled last in January 2017) for short sequencing reads with standard parameters for very-sensitive local alignments (Langmead and Salzberg, 2012a). Each re-mapping was used in order to infer genome type specific SNPs in

CpGV-0006, CpGV-R5, and CpGV-V15, respectively, using the Geneious RC10 native SNP prediction tool with default parameters (Kearse et al., 2012). SNPs within the assembled sequences with an average mapping quality above 30 (base-call accuracy 99.9%) were identified and quantified by comparing the nucleotide sequences with the SNP map generated for the five genome groups of CpGV (Wennmann et al., 2017). SNPs with lineage specific nucleotide information as described previously (Wennmann et al., 2017) were carried out for the further analyses. The distinct nucleotide frequencies of these SNPs were quantified using the Geneious RC10 native SNP prediction tool. The SNP information was subsequently extracted by evaluating the nucleotide frequencies specific for a genome group at each particular SNP using descriptive statistics in R (ver. 3.0).

Results

Resistance Testing with Commercial CpGV Products

Three commercial CpGV isolates, namely CpGV-R5, CpGV-0006, and CpGV-V15, were tested for their infectivity of neonates of susceptible CpS and resistant CpRR1, CpR5M, CpR5S, and CpRGO. Virus-induced mortality of CpS neonates was between 91% (CpGV-V15) and 100% (CpGV-0006) after seven days and between 99% (CpGV-R5) and 100% (CpGV-0006 and CpGV-V15) after 14 days for all treatments and did not show any significant differences (ANOVA, post-hoc Tukey HSD test, $p < 0.05$) (Figure 1). For type I resistant CpRR1, mortality ranged between 62% (CpGV-V15) and 98% (CpGV-0006) after seven days and increased to 86% (CpGV-V15) to >98% (CpGV-0006, CpGV-R5) after 14 days. Though statistically not significant, CpGV-V15 caused a tendentiously lower mortality in CpRR1 than the two other isolates. For type II resistant strains CpR5M and CpR5S, mortality was 8 to 17% for the isolates CpGV-R5 and CpGV-0006 after seven days and increased to no more than 37% mortality after 14 days of exposure. In contrast, the mortality of CpR5M and CpR5S larvae caused by CpGV-V15 was between 48% and 49% after seven days and increased to >85% after 14 days, which was statistically different from treatments with R5 and 0006 (ANOVA, post-hoc Tukey HSD test, $p < 0.05$). Type III resistant CpRGO neonates showed a mortality of 53% for CpGV-R5, 64% for 0006, and 91% for CpGV-V15 after seven days. The mortality increased to 80% for both CpGV-R5 and CpGV-0006, and to 100% for CpGV-V15 after 14 days. The different treatments did not differ statistically (Figure 1).

The results clearly indicate a very similar activity of CpGV-R5 and CpGV-0006. Both isolates showed a high virulence against the strains CpS and CpRR1, and also an effect on CpRGO. However, these two isolates did not cause the high mortality of neonates of CpR5M and CpR5S, neither after seven nor 14 days. Only the isolate CpGV-V15 was able to cause >85% mortality for all tested CM strains after 14 days.

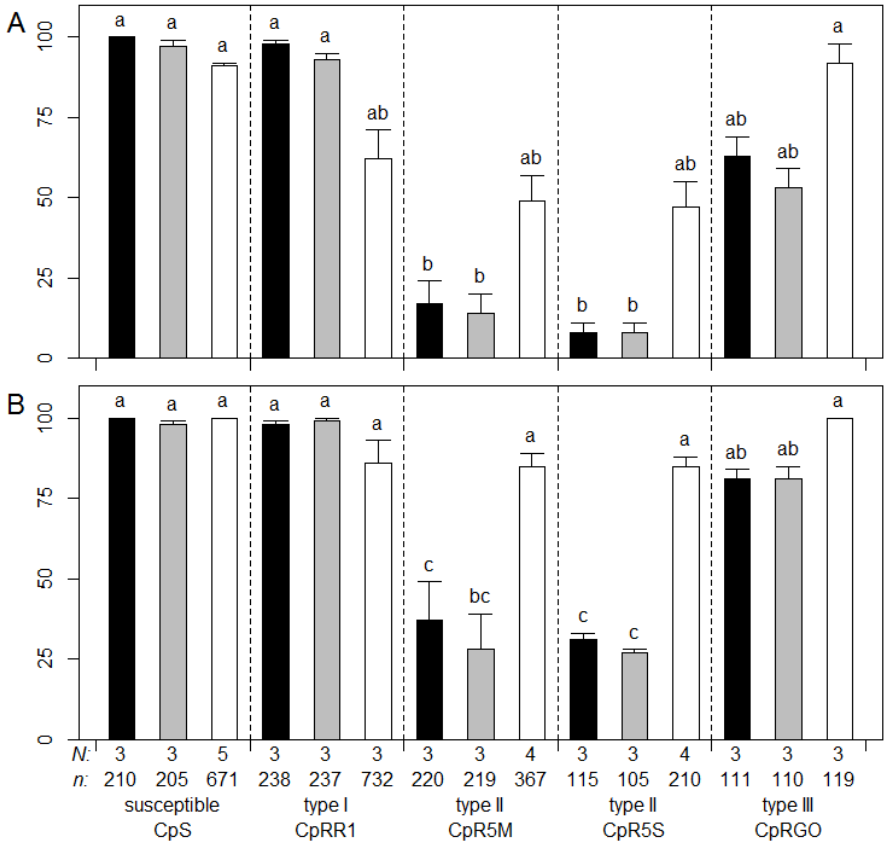

Figure 1 Resistance testing in different *Cydia pomonella* strains with novel Cydia pomonella granulovirus isolates. Mortality of neonates of CpS, CpRR1, CpR5M, CpR5S, and CpRGO (type I to III resistance) tested for resistance on artificial diet containing the commercial CpGV isolates 0006 (black bars), R5 (grey bars), or V15 (white bars) at the discriminating concentration of 5.8×10^4 occlusion bodies per mL. Abbott (1925) corrected mean mortality and standard error of mean (error bars) were determined at seven days (**A**) and 14 days (**B**) post infection. The total number of tested individuals (*n*) and independent replicates (*N*) are given below the chart. Columns marked by different letters differed significantly (analysis of variance (ANOVA), post-hoc Tukey´s Honestly Significant Differences test, $p < 0.05$).

Detection of Genome Type Specific SNPs in the CpGV Alignment

The genome sequences of five CpGV isolates, each one representing a CpGV genome group (CpGV-M: A, CpGV-E2: B, CpGV-I07: C, CpGV-I12: D, CpGV-S: E), were aligned against each other (for details see [6]). By considering each nucleotide position in the alignment and ignoring possible gap positions, a CpGV consensus sequence with a theoretical genome size of 126,225 bp was produced. According to the whole genome assembly, 650 positions with possible SNPs were identified in the Solexa Illumina reads (Wennmann et al., 2017). These SNPs were either specific for one genome group, or specific for a combination of at least two genome groups (Figure 2). One additional position, encoding for three possible lineage specific nucleotides as assessed by Wennmann et al. (Graillot et al., 2016), was excluded in the analysis of the Solexa Illumina reads. The vast majority (534 positions or 82%) were specific for a single genome group. The fewest SNPs were assigned to genome group A (CpGV-M) with only two unique SNPs located close to each other on the genome (Figures 2 and 3). In contrast, genome group C (CpGV-I07) contained the largest number of 356 SNPs, which were distributed almost evenly over the genome. Genome group E (CpGV-S) contained 101 SNPs, whereas 54 SNPs and 21 SNPs were identified for genome group B (CpGV-E2) and genome group D (CpGV-

I12), respectively. A smaller fraction of identified SNPs, namely 117, were found in two genome groups.

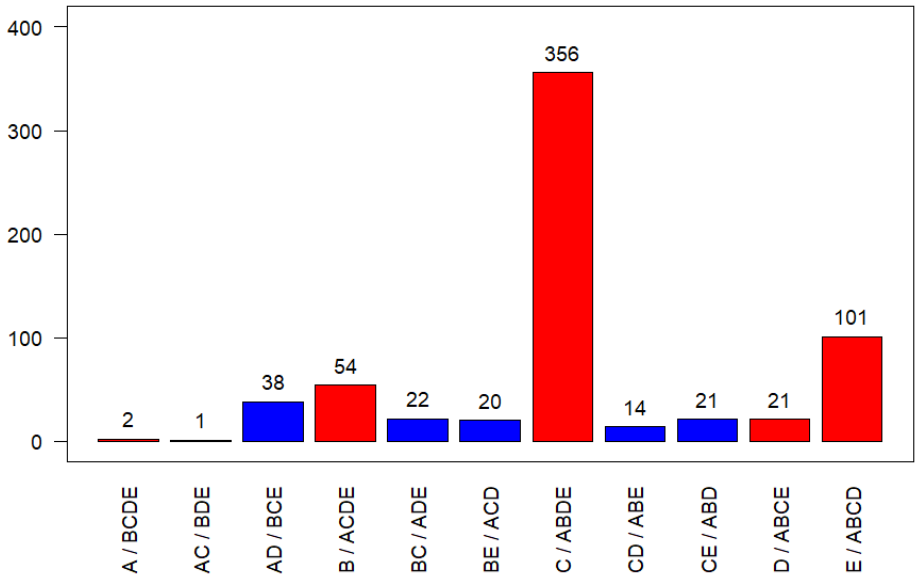

Figure 2 Number of genome positions in the CpGV whole genome alignment consensus sequence that encode for alleles with single nucleotide polymorphisms (SNPs) either specific for single genome groups (red columns) or two genome groups (blue columns). No common SNPs were detected for the combined genome groups AB, BD, or AE. The corresponding alleles are listed after the slashes.

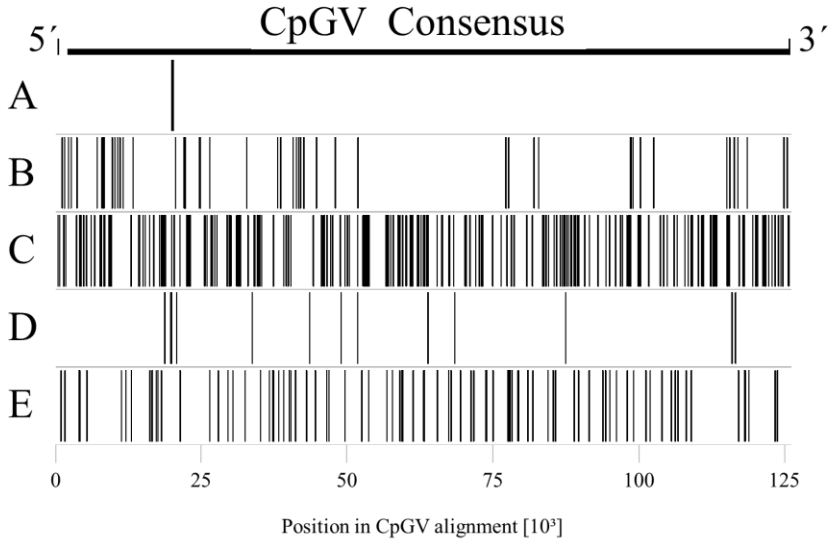

Figure 3. Distribution of SNPs specific to one of the five genome groups A to E of the CpGV whole genome alignment. Genome groups are indicated to the left, SNPs are represented by bars along the relative genome positions in the CpGV consensus sequence. Genome group A carries two SNPs, genome group B 54 SNPs, genome group C 356 SNPs, genome group D 21 SNPs, and genome group E 101 SNPs.

Genome group composition of CpGV-0006, CpGV-R5 and CpGV-V15

In order to evaluate the genome group composition of CpGV-0006, CpGV-R5, and CpGV-V15, DNA samples of these isolates were purified from commercial CpGV products and subjected to Solexa Illumina sequencing. Due to different formulations and possible contaminants, such as phenolic buffers, in the DNA samples, the sequencing of CpGV-0006 and CpGV-V15 was based on NexteraXT libraries, while classical Solexa Illumina libraries had to be produced for CpGV-R5 prior to sequencing. These circumstances affected the sequencing efficiency, resulting in various amounts of 151 bp read pairs being yielded (Table 1).

Table 1 Assembly reports of read pairs with an average Phred-quality score ≥Q30 generated by NextSeq500 next-generation sequencing and assembled against the CpGV alignment consensus sequence using the Bowtie2 mapper.

CpGV Isolate	Reads Pairs ≥ Q30	Number and Percentage of Read Pairs Assembled to CpGV Consensus	Number and Percentage of Read Pairs Not Assembled to CpGV Consensus	Mean Coverage ± SD
0006	736,927	729,391 (99%)	7536 (1%)	1690 ± 515
R5	1,517,362	240,501 (16%)	1,273,861 (84%) *	565 ± 180
V15	1,232,807	1,231,217 (>99%)	1590 (<1%)	2848 ± 557

*Read pairs in CpGV-R5 that did not map against the CpGV alignment consensus sequence were identified to derive mostly from contamination with DNA from *Cydia pomonella*. SD: standard deviation.

To determine the composition of the isolates, the frequency of each of the 534 genome group specific SNPs, which vary from two (CpGV-M, group A) to 356 (CpGV-I07, group C) positions, was determined [6]. Apparently, the very few specific SNPs of group A make the identification and quantification of genome group A much more difficult than that of groups with higher numbers of specific SNPs. Therefore, group A can hardly be quantified by itself; a good estimation of genome group A frequency, however, is given by determining the complement of the frequencies of SNPs specific for the genome groups B, C, D, and E. A further estimator for the presence of a given genotype is not only the frequency of a given SNP, but also the presence of the group-specific SNPs at all. In the sequencing approach of CpGV-0006, 729,391 read pairs were mapped against the CpGV consensus sequence. Consequently, a 1690-fold average genome coverage was achieved. In this assembly, a total number of 210 SNPs in the reads were assigned to lineage specific alleles according to their position and nucleotides (Figures 4–6). The SNPs distributed in CpGV-0006 could be grouped into four different frequency levels throughout the consensus sequence (Figure 4). With few exceptions, the SNPs of B, C, and D were assorted to frequencies below 10%, SNPs of group A and ABCD between 20–40%, and SNPs of group E between 60–80%. Exceptions to this assortment were observed at very few positions in the consensus sequence. The number of lineage specific SNPs was critical for the assessment of the genome group composition. Variation was detected at only 19 of the 356 alleles identified in genome group C (Figure 5), suggesting that isolates of lineage C were likely not a constituent of CpGV-0006. This was also observed for the genome groups B with four of the 19 variations detected and D with 12 of the 21 variations detected, respectively. Further, nucleotides representing these SNPs only appeared in very low median frequencies of about 1%. In contrast, SNPs with nucleotide variations identical to A and E alleles occurred in 100% of the Illumina reads. That indicated that CpGV-0006 was presumably only composed of SNPs from the lineages A or E, e.g., CpGV-M and CpGV-S. This hypothesis was supported by the high number of 160 SNPs identified for these genome groups divided into

nucleotides representing either group A or E. The median frequency of SNPs typical for genome group A reached 32%, while SNPs specific for genome group E reached a 67% median frequency. Therefore, the SNPs were subdivided into four combined (genome) groups, representing either genome group A or E, or the combination of A and E (AE), as well as the corresponding combination of B, C, and D (BCD). Once more, by observing the aspect of symmetry (Figure 6) for points in datasets of distinct genome groups, the following pattern was observed: The frequencies of genome groups A, AD, ACD, ABD, and ABCD (representing A) were all mirror inversed to the frequencies of genome groups BCDE, BCE, BE, CE, and E (representing E). Further, the genome groups ABCE, ABDE, ABE, ACDE, and ADE (representing A or E) all reached almost 100% SNP frequencies, while the genome groups B, BC, C, CD, and D acted here as symmetric counterparts and only achieved ca. 1% SNP frequencies. Thus, CpGV-0006 only consisted of isolates from the genome group A and E. As observed in the geometric dot-plots (Figure 6), the combined genome groups in all investigated isolates had overlapping 95% confidence limits. As a consequence, the frequency calculations were based on the median frequencies. The total number of data points of these combined groups was set up as a weighting factor for the median frequency. In that way, the initial overestimated frequency of BCD was corrected to 0%, leaving a genome group composition of CpGV-0006 consisting of 32% genome type A and 68% genome type E, and the differences between these weighted medians were evaluated by a post-hoc Tukey HSD test (Table 2).

A similar picture was observed for the isolate CpGV-R5. Here, the sequencing approach had to be performed with classic Solexa Illumina DNA libraries and only 487,002 read pairs (16%) could be mapped against the CpGV consensus sequence (Table 1). The remaining unmapped read pairs were de novo assembled using the Geneious native assembler. The largest contigs were submitted to Blastn searches and hits were allotted to *Bacillus cereus*, *Escherichia coli*, and *C. pomonella*. As the majority of these read pairs mapped against the mitochondrial genome sequence of *C. pomonella* (Acc.-N° JX407107) with an average 90-fold genome coverage, there was enough evidence that the sequenced non-CpGV DNA mostly derived from host larvae. Albeit this resulted in a reduced genome coverage, a total number of 192 lineage specific alleles were identified in the assembled reads (Figure 4 and Figure 5). In contrast to CpGV-0006, more variation was observed in the frequency levels (Figure 4). In general, the SNPs of C and D were grouped together according to their rather low frequencies, A and ABCD to their range of 20–40% SNP frequencies, and E to their range of 60–80%. Exceptions to this general classification were observed in the nucleotide positions of about ca. 37,000 to ca. 44,000 nt, and ca. 60,000 and 65,000 nt, as well as, in particular, between ca. 87,000 and ca. 100,000 nt (Figure 4). At these positions, several SNPs that were specific for either E or ABCD showed inversed frequencies compared to the general observation. Inversions as observed in these positions might indicate possible events of recombination. As for CpGV-0006, only very few SNPs specific for the genome groups C or D (12 in total) were detected in CpGV-R5. In particular, none of the 54 SNPs specific for the lineage of the genome type B were detected in CpGV-R5 (Figure 5). This suggested that isolates of lineage B were not a constituent of CpGV-R5. Again, SNPs representing the genome types C and D showed a low median frequency of 5%. Hence, CpGV-R5 was composed of 95% of SNPs from the lineages of genome type A or E. In CpGV-R5, nucleotides of these SNPs showed more variation in their frequencies, compared to CpGV-0006, e.g., E/ABCD compensating for E or A. Again, the initial median frequencies were

corrected by weighting with the number of data points of each group (Table 2). Similar to CpGV-0006, CpGV-R5 therefore contained SNPs from the lineages of genome type A to 36% and from genome type E to 64%.

In the case of CpGV-V15, a different composition was observed. Here, by the Bowtie2 mapping of 1,231,217 read pairs (>99% of total reads) against the CpGV consensus sequence, a 2848-fold mean coverage was achieved and 253 lineage specific alleles were identified. In CpGV-V15, SNPs specific to the genome types B reached almost 50% frequency, as much as SNPs specific to the genome type E reached almost 50% frequency over the consensus sequence (Figure 4). Similar to CpGV-0006 and CpGV-R5, the genome types C and D were almost absent as indicated by the very low numbers of SNPs specific for these genome groups (Figure 5). SNPs specific for alleles of A, C, and D yielded together median frequencies of only 6%. However, due to the low number of SNPs specific for genome group A in general, it was impossible to resolve whether an isolate of this genome group was present at all. Therefore, SNPs from these lineages were grouped together as ACD representing any isolate different from those deriving from the genome types B and E. Variations from these two lineages formed the two main components of CpGV-V15 with almost equal quantities. Amongst others, this was observed by the high frequencies of SNPs containing nucleotides for both B (50 out of 54) and E (97 out of 101). When considering the SNPs that encoded nucleotides for either B or E, slight variations in median frequencies were observed. However, both groups showed similar frequencies of about 50% (Figure 6). After weighting the median frequencies with the number of data points, CpGV-V15 contained 48% of isolates from genome group B and 50% from genome group E, while this difference in the frequencies was not confirmed by the post-hoc Tukey HSD test ($p = 0.95$). In contrast, only 2% of the composition was assigned to isolates from the combined genome groups A, C, and/or D (Table 2). Therefore, CpGV-V15 has a qualitatively and quantitatively different genome group composition when compared to CpGV-0006 and CpGV-R5. As in CpGV-0006 and CpGV-R5, two group E specific SNPs, namely at the positions 17,007 nt and 118,508 nt, were not found in CpGV-V15 (Figure 4).

Table 2 Genome group compositions of CpGV-0006, CpGV-R5, and CpGV-V15 mean and median frequencies of SNP variants with correspondent 95% confidence limits and five to 95 percentiles, respectively. Weighted medians are calculated by weighting the medians with the number of individual points from a dataset (data size = *n*). Post-hoc Tukey´s Honestly Significant Differences test (Tukey-HSD) was performed to evaluate significant differences in medians ($p = 0.05$)

CpGV Isolate	Genome Groups	Combined Group	Data Size (*n*)	Mean (95%-CL) (%)	Median (5–95%) (%)	Weight. Median (%)
0006	A, AD, ACD, ABD, ABCD	A	160	33 (32–34)	32 (27–38)	32
	BCDE; BCE; BE, CE, E	E	160	65 (63–67)	67 (55–72)	68
	B, BC, C, CD, D	BCD	50	3 (2–4)	1 (0.6–7)	0
R5	A, AD, ABD, ACD, ABCD	A	177	41 (39–43)	36 (26–77)	36
	BCDE, BCE, BE, CE, E	E	177	57 (54–60)	64 (18–74)	64
	BC, C, CD, D	CD [1]	15	9 (5–13)	5 (1–26)	0
V15	ABD, ABCD, B, BC	B	188	48 (47–50)	49 (36–57)	48 *
	ACDE, ADE, CE, E	E	188	52 (51–54)	51 (43–64)	50 *
	A, AD, ACD, C, D	ACD	65	15 (11–20)	6 (2–50)	2

1: None of the 54 possible SNPs specific for genome type B were detected in CpGV-R5.
*: identical medians (by Tukey-HSD) are marked with asterisks

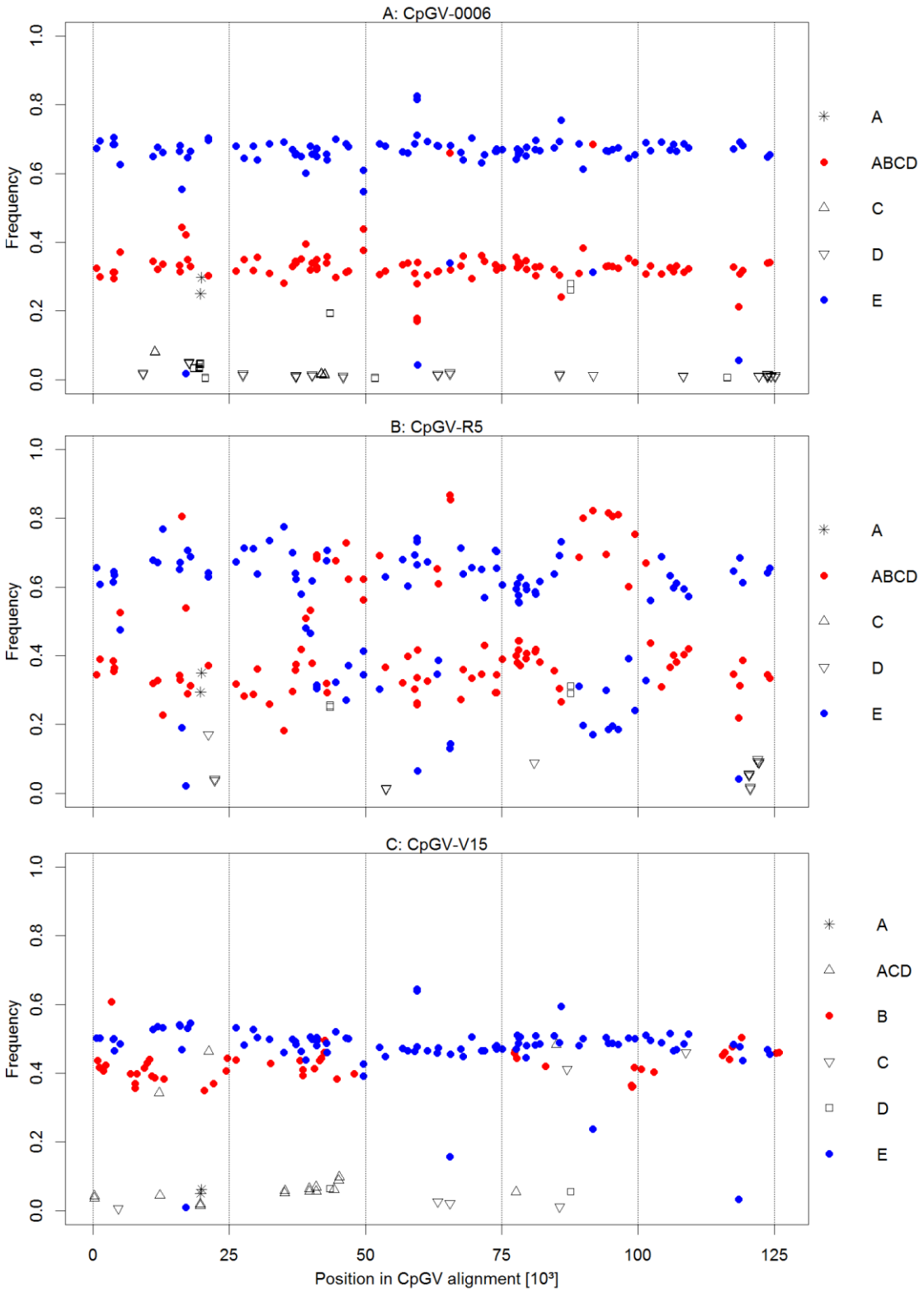

Figure 4 Frequencies of SNP variants specific for one of the CpGV genome types (A, B, C, D, E) or the combination ABCD (complement to E) in CpGV-0006 (**A**); CpGV-R5 (**B**) and CpGV-V15 (**C**) according to their nucleotide position in the CpGV consensus sequence. Blue: genome type E in CpGV-0006, -R5 and -V15; red: genome type A in CpGV-0006 and -R5, as well as genome type B in CpGV-V15; open symbols: other genome groups as indicated.

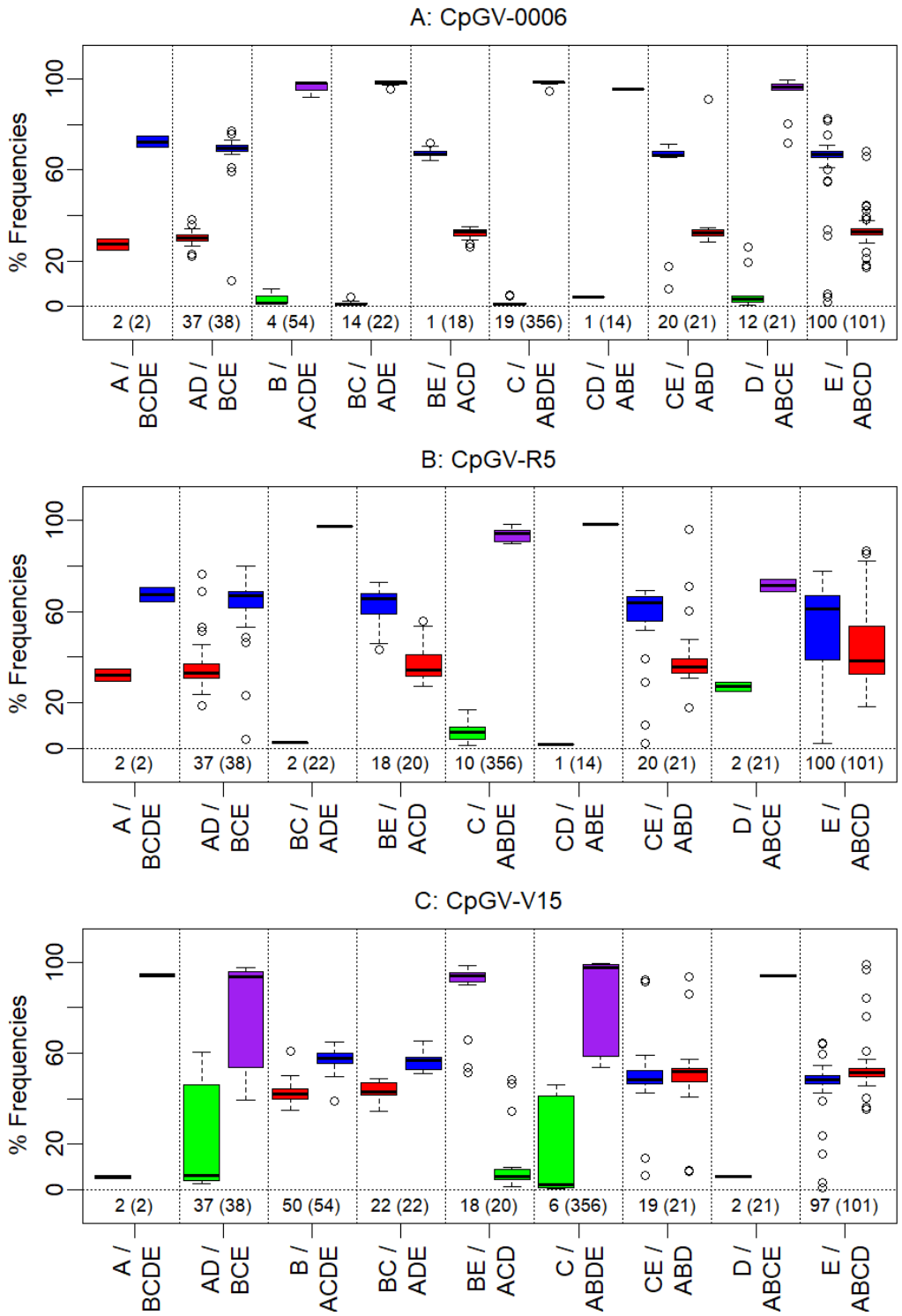

Figure 5 Frequencies (%) of SNP variants specific for one of the CpGV genome groups (A, B, C, D, E) or a combination of the genome groups in CpGV-0006 (**A**); CpGV-R5 (**B**) and CpGV-V15 (**C**). Given are the number of identified SNPs at the bottom line with all distinct possible SNPs in brackets, as well as box-whisker-plots with the median frequency of SNP variants given as lines and outliers as circles.

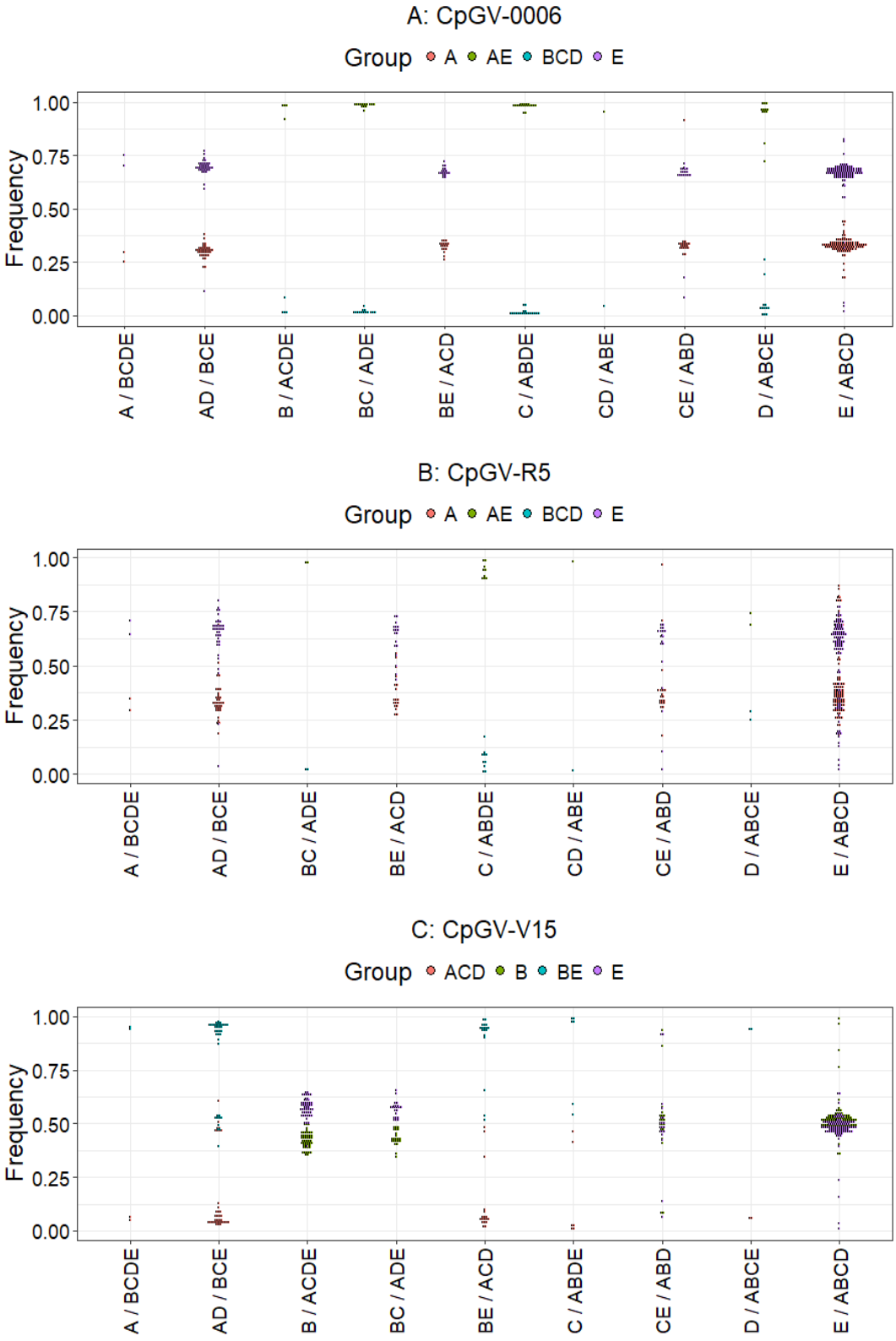

Figure 6 Geometric dot-plot charts of SNPs found in the Bowtie2 mappings of CpGV-0006 (**A**); CpGV-R5 (**B**); and CpGV-V15 (**C**). Given are the frequencies of nucleotide reads at positions assigned to the genome groups A, B, C, D, and E, as well as combinations of these. Data points are grouped together by genome groups A, E, and BCD for CpGV-0006 and CpGV-R5, as well as by genome B, E, and ACD for CpGV-V15, respectively.

Discussion

Ever since the first appearance of CM field populations resistant against commercial CpGV products, investigations have successfully identified resistance-breaking CpGV isolates (Berling et al., 2009; Eberle et al., 2008). Some of these isolates have been eventually used in new commercial CpGV formulations. However, CM populations did not only develop type I resistance against genome group A (Gebhardt et al., 2014), but also resistance against other genome groups, such as C, D, and E (type II and type III resistance) (Jehle et al., 2017), hampering early attempts to overcome resistance by using alternative CpGV isolates. So far, the natural genetic diversity of CpGV seems to be sufficient to break type I to III resistance, by applying CpGV isolates from different genome groups (Sauer et al., 2017a; Sauer et al., 2017b; Sauer et al., 2017c). In the present study, three novel commercial isolates, CpGV-0006, CpGV-R5, and CpGV-V15, were tested for their efficacy against different resistance types. The isolates CpGV-0006 and CpGV-R5 showed a similar activity against, i.e., high mortality in CpRR1, low mortality in CpR5M and CpR5S, and a somewhat intermediate mortality in CpRGO. CpGV-V15, in contrast, was the only commercial isolate with good activity in all resistant CM strains. Apparently, the activity of these isolates correlated well with their genetic composition, which we determined by a genome-wide identification of single nucleotide polymorphisms (SNPs) from next-generation sequencing (NGS) data. The high redundancy of NGS sequencing data does not only allow identification, but also the quantification of genetic variability within a given CpGV sample.

The frequencies of SNPs being characteristic for one of the genome groups were distributed evenly over the genome sequence, suggesting that CpGV-0006 and CpGV-V15 likely consisted of mixtures of separate CpGV lineages. In the case of CpGV-R5, the distribution was inversed in three genome regions, i.e., between ca. 37,000–44,000, ca. 60,000–65,000, and ca. 87,000–100,000 nt (Figure 4). This finding may indicate regions of recombination of the lineages A (e.g., CpGV-M) and E (e.g., CpGV-S) at these positions. On the other hand, the complete lack of two group E specific SNPs, namely at the positions 17,007 nt and 118,508 nt, in all three isolates hints to a sequencing error in the underlying original genome sequence of CpGV-S rather than a mutation or recombination at these two sites. Similar to the approach presented in this study, NGS data has been used to observe and evaluate occurring mutations over several passages, e.g., for several passages of Autographa californica multiple nucleopolyhedrovirus (AcMNPV) in the alfalfa looper (Chateigner et al., 2015) or in serial passages of Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) in permissive cell cultures (Gueli Alletti, unpublished data). In these studies, the NGS data was used either to compare the differences within AcMNPV at the beginning and the end of the experiment, or in order to demonstrate genomic stability in the case of AgseNPV-B. The NGS approach for the detection of genome group-specific SNPs in this study resembles a heuristic method for the detection and quantitation of genotype variation in mixtures. It highly depends on the nucleotide information given by the genome sequence alignment of presumably pure CpGV isolates. In the same way, this apparent key prerequisite also makes it highly adaptable when novel CpGV isolates will be described in the future. Albeit the mutation dynamics not being evaluated, the correlation of genome compositions and biological activities already displayed a possible origination of CpGV-0006, CpGV-R5, and CpGV-V15.

The analyses of the NGS data set revealed that all commercial products contain mixtures of at least two genome groups, though of different identity: Both CpGV-R5 and CpGV-0006 consist of a mixture of genome groups A (ca. 33%) and E (67%), whereas CpGV-V15 is composed of genome groups B (49%) and E (51%). These different compositions correlate with the differences observed in the resistance testing. Type I resistance can be broken by isolates from genome group B to E (Gebhardt et al., 2014; Jehle et al., 2017). The presence of genome group E in CpGV-0006 and CpGV-R5 explains the activity against CpRR1. The mean mortality of CpRR1 infected with either CpGV-R5 or CpGV-0006 (98% mortality) appeared to be higher, though statistically not significant, than obtained with single genome groups A and E (11% and 72%, respectively) (Sauer et al., 2017c), suggesting some synergistic effect, when A and E are combined. Similar effects were also identified in the larvae of the resistant CM strain RGV from France, where some synergy of mixed application of CpGV-M and the resistant-breaking CpGV-R5 in infection experiments were noticed (Graillot et al., 2016). We assume that CpGV-0006 and CpGV-R5 were selected in CM populations with type I resistance. Strikingly, both isolates showed a highly similar composition suggesting an independently selected and stabilized composition of CpGV populations when selected in resistant CMs. In experiments with Spodoptera frugiperda nucleopolyhedrovirus (SfMNPV), it has been demonstrated that laboratory virus genotype mixtures achieved an almost identical composition of genotypes of wild type isolates (Clavijo et al., 2009). Other than effects which presumably derive from virus-virus interactions, the limiting factor of virus selection was the host, *Spodoptera frugiperda*, which was more susceptible to mixtures of genotypes than to single genotypes of SfMNPV (Lopez-Ferber et al., 2003; Simon et al., 2005; Simon et al., 2010). Given these similar observations, the prevalence of partially resistant insect populations represents a bottleneck in the dynamics of virus isolates which attenuate to a certain virus composition.

Nevertheless, no synergism of group A and E CpGV isolates of CpGV-R5 or CpGV-0006 was observed in CMs with type II resistance; corroborating recent findings which demonstrated the lack of efficacy of single and combined applications of genome group A and E CpGVs in CpR5M and CpR5S (Sauer et al., 2017c). Isolate CpGV-V15 (mixture of genome group B and E) induced a mortality of 50% and 85% after seven and 14 days, respectively. It was shown for CpR5M and CpR5S that the single application of CpGV-E2 (genome group B) caused 85% and >99% mortality after seven and 14 days, respectively (Sauer et al., 2017a) Hence, the synergism of group B and E in overcoming type II resistance is not supported. From a laboratory view, the application of pure genome group B instead of mixtures of group B and E appears to be promising to maximize efficacy on CM populations with type II resistance.

It was proposed that the development of type I resistance could be the consequence of the exclusive use of CpGV-M in previous commercial products. For this reason, resistance management strategies are essential to restore and sustain the efficacy of CpGV products. Such strategies should exploit the full genetic diversity of CpGV, including resistance breaking isolates. The rotation of different CpGV isolates could be useful to avoid the continued selection of resistance in the field. Our findings, however, indicate that CpGV-0006 and CpGV-R5 have a highly similar composition of CpGV isolates from genome groups A and E; additionally, their effect on different resistance types is more or less equal. In conclusion, a rotation of these two commercial isolates in the field would not have the desired effect of alternating the selective agent in populations with type I resistance. For type II and type III resistance, their efficacy is

not considered to be sufficient. Any resistance management should therefore include the optimized use of CpGV isolates from different genome groups, as wells non-virus alternatives, such as the application of bacteria, fungi, nematodes, and beneficial arthropods as biological control tools.

Chapter VII: General Discussion

When defining virus diversity, it important to clarify the terms “strain” and “isolate”, which are often mistaken for one of the other. Strains of baculovirus species are characterized as genetically homogenous and their generation often requires extensive laboratory work; e.g. plaque purification of clones and selection *in vitro* in susceptible insect cell culture, or *in vivo* limiting dilution series in host larvae. A prominent example is the cloned Cydia pomonella granulovirus (CpGV) strain CpGV-M1 (Crook et al., 1997), which has been considered as the type strain for *Cydia pomonella granulovirus* species. In contrast, baculovirus isolates can contain different genotypes of one species and hence are considered genetically more diverse. Even occlusion bodies (OBs) of betabaculovirus isolates, which contain only one nucleocapsid per OB, can be composed of genotype mixtures as demonstrated for example for CpGV infection of single larvae (Rezapanah et al., 2008). A good example for this genotypical diversity is the Autographa californica multiple nucleo-polyhedrovirus (AcMNPV) isolate AcMNPV-WP10 (Chateigner et al., 2015). AcMNPV-WP10 derived from a natural population of AcMNPV and was amplified in a single infection in 500 highly susceptible cabbage looper (*Trichopolusia ni*) larvae. As a result, AcMNPV-WP10 presumably contained a mixture of several different AcMPNV genotypes compared to the type strain AcMNPV-C6, which were classified by their unique variant frequency pattern of mutations. Still, given this molecular background information isolates can still be distinguished from each other and characterized based on their genetic lineages (Gueli Alletti et al., 2017c; van der Merwe et al., 2017).

In context of the registration of baculoviruses as biological control agents, a clear definition of the biological identity is inevitable as for example defined in the European regulations 1107/2009[1] and 283/2013[2]. Baculoviruses benefit from shared characteristics regarding the impact on human health and non-target organisms (OECD, 2002). In consequence, in the EU baculoviruses used as plant protection agents are approved on their species level as active ingredient, and novel isolates of these species can be included in the list of active substances based on the common identity. Furthermore species, but also isolates, may consist of mixtures of different, but sufficientsimilar genotypes. Specific isolates may also contain more than one species, be it by contamination or co-infection as reported for example for AgseNPV-B (Wennmann and Jehle, 2014). However, the use of isolates consisting of more than one species as baculovirus based plant protection products is not feasible also in the aspect of identity information requirements for the registration. Nevertheless, such differences in the genotypes do not alter the risk assessment towards non-target organisms or the environment, specific lineages may influence some biological properties such as the virulence in hosts, as observed for CpGV (Gueli Alletti et al., 2017c; Jehle et al., 2017). This emphasizes the utmost necessity for describing baculovirus diversity on a detailed molecular level, be it on species or strain level and demands for tools for unequivocal identification, such as RFLP or whole genome sequencing. Within the scope of this thesis, it was elucidated how next generation sequencing

[1] Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC

[2] Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market

can contribute on describing the diversity of baculovirus isolates, based on three different isolates and scenarios: the comparison of geographically distant virus isolates with high genetic similarity in case of AgseGV-DA and the two Chinese isolates AgseGV-XJ and –L1; the attempt to describe the genetic stability of a in vitro cloned AgseNPV-B genotype under the selection pressure of consecutive amplifications in cell culture; and the genomic diversity within each of the commercial CpGV isolates, CpGV-0006, CpGV-R5 and CpGV-V15.

Recently, two isolates of AgseGV were isolated in China, namely AgseGV-XJ in Xinjiang in North-West China and in a distance of 3,390 km AgseGV-L1 in Shanghai on the East coast of China. Their fully sequenced genomes were provided in comparison to each other by Zhang et al. (2014). The European isolate presented in this thesis, AgseGV-DA, originally derived from field collections in Vienna, Austria (Chapter II). All three isolates have their origins in geographically distant and confined spots and one would assume that this separation may have led to distinct genetic lineages. However, all three isolates show 99.6 to 99.8% nucleotide identity with most of the differences in the *enhancin* gene (Chapter II). The speed of killing and the gross pathology as observed microscopically in tissue of infected larvae defines AgseGV-DA as a "slow"-killing granulovirus. Slow-killing baculoviruses are common within betabaculoviruses infecting larvae of Noctuidae, e.g. as for Helicoverpa armigera granulovirus (HearGV), Trichoplusia ni granulovirus (TnGV), Adoxophyes orana granulovirus (AdorGV) and Xestia c-nigrum granulovirus (XecnGV) and occassionly also in alphabaculoviruses, e.g. Adoxophyes honmai nucleopolyhedrovirus (AdhoNPV) (Federici, 1997; Harrison and Popham, 2008; Jehle et al., 2006b; Takahashi et al., 2015; Wormleaton and Winstanley, 2001). Pathological traits, like the speed of kill, are not monophyletic traits in betabaculoviruses, but rather evolved independently. (Jehle et al., 2006b). In case of AgseGV, one reason for the slow killing behavior is their restricted infection to midgut epithelium and fat body cells. Thus, infected larvae can develop further and do not stop feeding. In consequence, they decease and liquify more later. Unfortunately, there is no data available on the infectivity of AgseGV-XJ or AgseGV-L1 in *A. segetum* larvae. Hence, it may only be assumed that the slow killing pathology also applies to the two Chinese isolates. In contrast, infections with alphabaculoviruses and fast killing betabaculoviruses, like CpGV, are able to infect all larval tissues and cause a fast endemic spread of the infection, resulting in a relatively quick death and complete disintegration of the larval body. Albeit full genome sequences of AgseGV have been available previously, the full description of the gross pathology has not been described yet. For this reason, AgseGV-DA was proposed as the type isolate for the species AgseGV, as the investigations to this isolate provided all necessary information on identity, host range and pathology, as well as the morphology (Chapter II, Gueli Alletti et al., 2017b).

AgseGV was classified following two different approaches; based on the comparison of the Kimura-two-parameter (K2P) nucleotide distances of concatenated *granulin*/*polyhedrin*, *late expression factor* (*lef*) 8 and *lef*-9 alignments of *Alpha*- and *Betabaculovirus* species and by molecular phylogeny of alignments of the 38 baculovirus core genes (Jehle et al., 2006b; Wennmann et al., 2018). The advantage of the K2P-approach derives from the large number of sequences from different *Alpha*- and *Betabaculovirus* species due to the development of degenerated primers for these genes (Jehle et al., 2006b). Thus, the collection of taxon-specific sequences of these marker genes exceeds the number of fully sequenced baculovirus genomes

by far. And furthermore, the phylogenetic analyses based on K2P-distances resolves the relationship between alpha- and betabaculoviruses as adequate as the molecular phylogeny based on the alignments of the baculovirus core genes (Wennmann et al., 2018). However, the limitation of this method is the scarce conservation of *polyhedrin* in gammabaculoviruses and the absence of its homologue in deltabaculviruses as well as the restriction that the K2P-distances based method has been developed primarily for alpha- and betabaculoviruses (Afonso et al., 2001; Duffy et al., 2006; Garcia-Maruniak et al., 2004; Jehle et al., 2006b; Lauzon et al., 2004; Miele et al., 2011; Perera et al., 2007; Wennmann et al., 2018). Therefore, in order to assess a full picture of baculovirus evolution, baculovirus core gene-based phylogeny is deemed necessary. Another benefit of analyzing full genome sequences is the possibility to follow evolutionary lineages beyond nucleotide-based analyses, e.g. for the assessment of inversions, reiterations/duplications and translocations of whole genes as demonstrated for the *viral enhancin factors* (*vef*) of AgseNPV-B, AgseNPV-A and AgipNPV (Wennmann et al., 2015a).

Similar to AgseGV, AgseNPV-B has shown no genotype variation when sequenced on an ABI SOLiD next generation sequencing platform as described by Wennmann et al. (2015a). This homogenous genome sequence was confirmed by the Solexa Illumina approach used in this thesis, with only one adaption to the published ABI SOLiD genome sequence (a change of guanine to thymine at nucleotide position 972) (Chapter IV). In contrast to this genomic homogeneity, a variation within the gene for the baculovirus ecdysteroid UDP-glucosyltransferase (*egt*) was observed in field isolates of the North American AgipNPV Illinois strain (Harrison, 2009). Plaque purified clones of AgipNPV with a non-functional *egt* could be separated within three consecutive infections in insect cell culture. On the one hand, these isolates showed an increased speed of kill, but on the other hand they also showed a reduced pathogenicity (Harrison, 2013). The deletion of of *egt* was also investigated for the model baculovirus AcMNPV and it was confirmed that this enzyme catalyses the transfer of glucose from UDP-glucose to ecdysteroids, which are insect molting hormones (O´Reilly and Miller, 1989; O´Reilly and Miller, 1990). The expressed EGT is translocated to the hymolymph of infected larvae and forms an ecdysteroid-UDP-sugar-enzyme ternary complex (Evans and O´Reilly, 1998; Evans and O´Reilly, 1999) and its active role in interfering with molting was demonstrated by *in vivo* assays with *egt* deletion mutants of AcMNPV been inserted host-owned juvenile hormone esterase genes (Eldridge et al., 1992). By interfering with the insect molting, infected larvae die at an earlier larval instar stage (O´Reilly and Miller, 1989) und thus the speed of kill is increased. However, histopathological investigations revealed that the altered sped of kill is rather influenced by the degeneration of the Malpighian tubules due to EGT, and not necessarily to the disrupted molting (Flipsen et al., 1995). As the whole larval development is influenced by EGT, this may also affect their climbing behavior to some extend, which is referred to as “tree-top” disease (Ros et al., 2015; van Houte et al., 2014a; van Houte et al., 2014b). Although, such features as a reduced speed of kill may be of special interest in pest control by baculoviruses, the low number of selection rounds in insect cell culture needed to produce *egt* mutants in AgipNPV, however, has shown that there is certain instability in the distribution of AgipNPV genotypes in this isolate.

After all, baculovirus based products demand virus isolates with high genome stability for quality control, be it as information on mutation rate of traits related to the mode of action under

environmental conditions or during the production process. Here, it is shown that AgseNPV-B features high genome stability in ten consecutive infection rounds in a susceptible insect cell culture. This has been tested in the insect cell line AiE1611T, which has been tested before for its susceptibility to AgseGV and AgseNPV-B (Chapter III). Not only did infections and transfections with AgseNPV-B produce numerous OBs, moreover, the AiE1611T cells were only permissive for AgseNPV-B but not for AgseGV. Hence, infection of this cell culture could be used to separate mixtures of AgseNPV-B and AgseGV which may occur during *in vivo* propagation of the virus, but different mixtures e.g. AgipNPV and AgseGV have still to be tested in AiE1611T (Wennmann and Jehle, 2014; Wennmann et al., 2015c). Although the *in vivo* virulence of the tested isolate AgseNPV-B PP2 (deriving from plaque purified clone N° 2) scattered in *A. segetum* larvae during the ten consecutive infections, the genome of AgseNPV-B PP2 remained stable. Mutations such as few polyhedra (*fp*) mutants, defective interfering particles (DIP), but also as observed for the *egt* mutant AgipNPV isolate, were absent in AgseNPV-B PP2. Furthermore, AgseNPV-B PP2 showed an average ~6,700-fold increased virulence in its LD_{50} in *A. segetum* larvae compared to AgseGV-DA, although such comparison formally requires parallelism of the probit lines (REF). Thus, the robust genome stability and the comparably high virulence make AgseNPV-B an excellent candidate for a baculovirus control agent against *A. segetum*. The Solexa Illumina NGS used in evaluating the genome stability could not satisfactorily resolve single nucleotide polymorphisms (SNPs), as most of the SNPs were identified within homologous repeat sequences (*hrs*). AgseNPV-B carries six of these *hrs* sequences which range in number of repetition units and length. *Hrs* sequences presumably act as replication origins and/or enhancers of transcription (Kool et al., 1993; Leisy and Rohrmann, 1993; Pearson et al., 1993). However, as demonstrated in the sequential deletion of all *hrs* in AcMNPV, these sequences are rather involved in the production of budded virus and the expression of late proteins, as the *hrs*-deleted AcMNPV mutants still produced functional occlusion bodies in the *Spodoptera frugiperda* cell culture Sf21 (Bossert and Carstens, 2018)One of the *hrs (hr6)* of AgseNPV was used as target for homologous recombination in AiE1611T cells to generate a recombinant baculovirus consisting of a *hr6* deleted AgseNPV-B with the bacterial plasmid pBAC-asb134/135 inserted containing a *lacZ* and kanamycin-resistance gene, as well as a bacterial mini-F origin of replication. Unfortunately, amplification of this so-called AgseNPV-B bacmid, bAgseNPV-B, resulted in a deletion of 42,443 bp between ORF *asb005* and ORF *asb046*, which was confirmed by Solexa Illumina sequencing and mapping against the *in silico* deduced sequence of bAgseNPV-B (Chapter V), as well as by the endonuclease restriction analysis using the enzymes *Hin*dIII and *Eco*RI (data not shown). Again, most of the detected SNPs in bAgseNPV-B were identified in *hr*3. But as the methodology included the *in vitro* production in *E. coli* DH5α, it can be assumed that bAgseNPV-B was a genetically homogenous genotype. In consequence, the SNP detection with Solexa Illumina was either not satisfactory, or impossible as as single nucleotide mutations might have been absent in the genome. Thus, the detected SNPs rather resemble artefacts from false mapping against the *in silico* deduced genome sequence of bAgseNPV-B. However, this method has shown on the other hand that large deletions can be detected easily by their coverage in mappings of Solexa Illumina reads against a reference sequence (here bAgseNPV-B) and confirmed the methodology used to detect large deletions in the serial passages of AgseNPV-B PP2 in AiE1611T.

Taken together, both baculovirus examples from AgseGV and AgseNPV-B showed highly homogenous genomes of a given isolate. In contrast a different picture was elucidated by the genomic analyses of the three CpGV isolates, CpGV-0006, CpGV-R5 and CpGV-V15. When speaking of genotypical diversity in CpGV it is important to mention that based on phylogenetic and RFLP analyses of their genome sequences, all known CpGV isolates can be classified into five genome groups (A to E), describing different phylogenetic lineages. Representative isolates are CpGV-M (genome group A), CpGV-E2 (B), CpGV-I07 (C), CpGV-I12 (D), and CpGV-S (E) (Eberle et al., 2009; Gebhardt et al., 2014). Great efforts have been made in characterizing these representative isolates in comparison to CpGV-M (genome group A), e.g. in terms of the presence of a 24 nt iteration within the gene *pe*38 in CpGV-M (Gebhardt et al., 2014), differences in the restriction patterns (Eberle et al., 2008; Eberle et al., 2009) or lineage-specific SNP distribution (Wennmann et al., 2017). In the approach presented in Chapter VI of this thesis all three tested isolates were mixture of at least two different genotypes, namely to 32-36% genome group A and to 64-68% genome group B for the two isolates CpGV-0006 and CpGV-R5, and to 48% genome group B and 50% genome group E for CpGV-V15. These different genotype compositions could be attributed to specific mortality responses in codling moth larvae with different types of resistance, which of course has to be tested in dose-response bioassays whenever products based on these isolates are sought to be commercially marketed. In terms of baculovirus diversity, two different evolutionary trends are presented by CpGV and the two *Agrotis* specific baculoviruses (AgseNPV-B/AgseGV). In case of CpGV, the genotypically mixed isolates can be correlated to different responses in resistant codling moth populations. So far all laboratory examined resistant codling moths populations could be summarized into three different resistance patterns in which codling moth with type I resistance are not susceptible to CpGV-M-like isolates (genome group A), type II and III resistant codling moths are resistant to CpGV-M and CpGV-S like isolates, but with different inheritance patterns of these resistances (Asser-Kaiser et al., 2007; Asser-Kaiser et al., 2010; Berling et al., 2013; Sauer et al., 2017a; Sauer et al., 2017b; Sauer et al., 2017c). While these different populations were investigated extensively in laboratory experiments with "pure" imbred individuals, it is important to note that in nature codling moth populations may occur as mixtures of differently resistant individuals. Furthermore, pure imbred populations often suffer from reduced fertility. The successful propagation of viruses on the other hand the possibility to infect its host and also the access to large numbers of host individuals in order to spread horizontally. Mixtures as observed in the three commercial isolates therefore resemble an adaption to the circumstances in laboratory selections: susceptible larvae get infected either way and decease and while the consecutive infection in resistant larvae may influence the heterogeneity of the virus (Graillot et al., 2014) and *vice versa* the selection pressure given by the resistance-breaking genotypes may reduce the size of resistant individuals due to the lowered fecundancy in the laboratory selections. Genotype mixtures were absent in AgseGV and AgseNPV-B, but mixtures of both virus species were obtained from infected larvae due to their ability to co-infect single larvae (Wennmann et al., 2015c). Mixed infections of betabaculoviruses and alphabaculoviruses have also been observed in different host-virus complexes, such as in *Pseudaletia unipuncta* in mixed infections with Pseudaletia unipuncta granulovirus (PsunGV) and Pseudaletia unipuncta nucleopolyhedrosis virus (now classified as Mythmia unipuncta nucleopolyhedrovirus; MyunNPV) as well as for *Trichoplusia ni* (with no

detailed characterization of the infecting viruses) (Harrison et al., 2018b; Lowe and Paschke, 1968; Tanada and Hukuhara, 1971). The sources of recombination between the baculovirus genomes have inter-generic hurdles in mixed infections of nucleopolyhedroviruses and granuloviruses. One possible obstacle in mixed infection is the case of super-infection or co-occlusion of different viruses, as for example observed in simultaneous superinfections of invertebrate cells by Lymantria dispar nucleopolyhedrovirus (LdMNPV) and AcMNPV (McClintock and Dougherty, 1987). However, superinfections of two different baculoviruses are constrained by the infection cycle of baculoviruses *per se*. During baculovirus infection the RNA II polymerase of the host in inactivated and soon replaced by the baculovirus-own RNA polymerase, thus transcription of two different baculoviruses and consequently a successful superinfection is limited to the period before inactivation of the host RNA polymerase II (Beperet et al., 2014). Nevertheless, superinfection between different virus families does also occur with baculoviruses and the transmitted virus progeny often contains mixed or co-occluded viruses. One example is the superinfection of Spodoptera exigua nucleopolyhedrovirus (SeMNPV) with the iflavirus of *Spodoptera exigua* SeIV1 (Jakubowska et al., 2016; Virto et al., 2014).

But what efforts have been made to detect baculovirus diversity? In the past, approaches to detect the diversity and evolution of CpGV included extensive analyses of restriction fragment length polymorphisms and hence a broad background knowledge of presumably pure CpGV isolates/strains (Eberle et al., 2009). Analyses of mixed infections of AgseNPV-B and AgseGV was limited to quantitative PCR (qPCR) reactions, which again demanded knowledge about the genome sequences as highly specific primers had to be developed (Wennmann and Jehle, 2014; Wennmann et al., 2015c). In the same manner mixtures of SeMNPV and SeIV1 were exclusively quantified by qPCR (Jakubowska et al., 2016). Alternatives to visualize the genetic variation of baculoviruses also were limited on few genes, like in the laboratory intensive analyses by denaturating gradient gel electrophoresis (DGGE) of the two Helicoverpa armigera nucleopolyhedrovirus genes *dbp*1 and *me*53 from an Australian isolate (Baillie and Bouwer, 2012). In contrast, next generation sequencing of such samples generates large datasets, which can be analyzed by automated protocols, which simply need some background information on the sequences and elaborated coding. The sequence analyses of AgseNPV-B, as bacmid or passed over several rounds in cell culture, have shown that Solexa Illumina sequencing can be readily used to evaluate the presence of large genomic deletions. A deeper application is given by the analyses of the CpGV isolates, where background information on the appearance of SNPs in the genome was already available. Here, the presented sequencing approach readily provided the composition of the three CpGV isolates. In conclusion, next generation sequencing (NGS) of baculovirus isolates provides a comprehensive genome comparison of whole genomes and a detailed evaluation of genotype mixtures. With the ongoing changes in legislation on methods to provide the identity of baculoviruses, NGS methods represent a contemporary tool to address not only scientific questions but also regulatory needs.

Refernces

Abbott, W. S., 1925. A Method of Computing the Effectiveness of an Insecticide. J. Econ. Entomol. 18, 265-267.

Abot, A. R., Moscardi, F., Fuxa, J. R., Sosa-Gomez, D. R., Richter, A. R., 1996. Development of Resistance by Anticarsia gemmatalis from Brazil and the United States to a Nuclear Polyhedrosis Virus under Laboratory Selection Pressure. Biol. Control. 7, 126-130.

Adams, M. J., Lefkowitz, E. J., King, A. M. Q., Harrach, B., Harrison, R. L., Knowles, N. J., Kropinski, A. M., Krupovic, M., Kuhn, J. H., Mushegian, A. R., Nibert, M., Sabanadzovic, S., Sanfaçon, H., Siddell, S. G., Simmonds, P., Varsani, A., Zerbini, F. M., Gorbalenya, A. E., Davison, A. J., 2016. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2016). Arch. Virol. 161, 2921-2949.

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Gruning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., Goecks, J., 2016a. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. 44, W3-W10.

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Gruning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., Goecks, J., 2016b. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucl. Acid. Res. 44, W3-W10.

Afonso, C. L., Tulman, E. R., Lu, Z., Balinsky, C. A., Moser, B. A., Becnel, J. J., Rock, D. L., Kutish, G. F., 2001. Genome sequence of a baculovirus pathogenic for Culex nigripalpus. J Virol. 75, 11157-65.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., 1990. Basic Local Alignment Search Tool. J. Mol. Biol. 215, 403-410.

Arends, H. M., Jehle, J. A., 2006. Sequence Analysis and Quantification of Transposase cDNAs of Transposon TCp3.2 in Cydia pomonella Larvae. Arch. Insect. Biochem. Physiol. 63, 135-145.

Arrizubieta, M., Williams, T., Caballero, P., Simon, O., 2014. Selection of a nucleopolyhedrovirus isolate from Helicoverpa armigera as the basis for a biological insecticide. Pest Manag Sci. 70, 967-76.

Asser-Kaiser, S., Fritsch, E., Undorf-Spahn, K., Kienzle, J., Eberle, K. E., Gund, N. A., Reineke, A., Zebitz, C. P., Heckel, D. G., Huber, J., Jehle, J. A., 2007. Rapid emergence of baculovirus resistance in codling moth due to dominant, sex-linked inheritance. Science. 317, 1916-8.

Asser-Kaiser, S., Heckel, D. G., Jehle, J. A., 2010. Sex linkage of CpGV resistance in a heterogeneous field strain of the codling moth Cydia pomonella (L.). J Invertebr Pathol. 103, 59-64.

Asser-Kaiser, S., Radtke, P., El-Salamouny, S., Winstanley, D., Jehle, J. A., 2011. Baculovirus resistance in codling moth (Cydia pomonella L.) caused by early block of virus replication. Virology. 410, 360-7.

Au, S., Wu, W., Panté, N., 2013. Baculovirus Nuclear Import: Open, Nuclear Pore Complex (NPC) Sesame. Viruses. 5, 1885-1900.

Ayres, M. D., Howard, S. C., Kuzio, J., Lopez-Ferber, M., Possee, R. D., 1994. The Complete DNA Sequence of Autographa californica Nuclear Polyhedrosis Virus. Virology. 202, 586-605.

Baillie, V. L., Bouwer, G., 2012. High levels of genetic variation within Helicoverpa armigera nucleopolyhedrovirus populations in individual host insects. Arch Virol. 157, 2281-2289.

Beames, B., Summers, M. D., 1989. Location and nucleotide sequence of the 25K protein missing from baculovirus few polyhedra (FP) muntants. Virology. 168, 344-353.

Beas-Catena, A., Miron, A. S., Camacho, F. G., Contreras-Gómez, A., Molina-Grima, E., 2014. Baculovirus biopesticides: An overview. Journal of Animal and Plant Science. 24, 362-373.

Benson, G., 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucl. Acid. Res. 27, 573-580.

Benz, G. A., The Preclassic Period of Light Microscopy. In: R. R. Granados, B. A. Federici, Eds.), The Biology of Baculoviruses. CRC Press, Inc., Boca Raton, Florida, 1986, pp. 1-34.

Beperet, I., Irons, S. L., Simon, O., King, L. A., Williams, T., Possee, R. D., Lopez-Ferber, M., Caballero, P., 2014. Superinfection exclusion in alphabaculovirus infections is concomitant with actin reorganization. J Virol. 88.

Berling, M., Blachere-Lopez, C., Soubabere, O., Lery, X., Bonhomme, A., Sauphanor, B., Lopez-Ferber, M., 2009. Cydia pomonella granulovirus genotypes overcome virus resistance in the codling moth and improve virus efficiency by selection against resistant hosts. Appl. Environ. Microbiol. 75, 925-30.

Berling, M., Sauphanor, B., Bonhomme, A., Siegwart, M., Lopez-Ferber, M., 2013. A single sex-linked dominant gene does not fully explain the codling moth's resistance to granulovirus. Pest Manag Sci. 69, 1261-6.

Bischoff, D. S., Slavicek, J. M., 1995. Molecular Analysis of an enhancin Gene in the *Lymantria dispar* Nuclear Polyhedrosis Virus. J. Virol. 71, 8133–8140.

Bischoff, D. S., Slavicek, J. M., 1997. Phenotypic and genetic analysis of Lymantria dispar nucleopolyhedrovirus few polyhedra mutants: mutations in the 25K FP gene may be caused by DNA replication errors. J Virol. 71, 1097-1106.

Black, B. C., Lynn, A. B., Dierks, P. M., Gard, I. E., Commercialization of Baculoviral Insecticides. In: L. K. Miller, (Ed.), The Baculoviruses. Spinger US, 1997, pp. 341-387.

Blissard, G. W., Kogan, P. H., Wei, R., Rohrmann, G. F., 1992. A synthetic early promoter from a baculovirus: roles of the TATA box and conserved start site CAGT sequence in basal levels of transcription. Virology. 190, 783-793.

Blomfield, T. L., Giliomee, J. H., 2011. Effect of temperature on the oviposition, longevity and mating of codling moth, *Cydia pomonella* (L.) (Lepidoptera: Tortricidae). African Entomology. 19, 42-60.

Börner, H., Krankheitserreger und Schädlinge: Lepidoptera. Pflanzenkrankheiten und Pflanzenschutz, Vol. 8. Ulmer, Stuttgart, 1997, pp. 394-401.

Bossert, M., Carstens, E. B., 2018. Sequential deletion of AcMNPV homologous regions leads to reductions in budded virus production and late protein expression. Virus Res. 256, 125-133.

Boughton, A. J., Harrison, R. L., Lewis, L. C., Bonning, B. C., 1999. Characterization of a Nucleopolyhedrovirus from the Black Cutworm, Agrotis ipsilon (Lepidoptera: Noctuidae). J Invertebr Pathol. 74, 289-294.

Bourner, T. C., Vargas-Osuna, E., Williams, T., Santiago-Alvarez, C., Cory, J. S., 1992. A comparison of the efficacy of nuclear polyhedrosis and granulosis viruses in spray and bait formulations for the control of *Agrotis segetum* (Lepidoptera: Noctuidae) in Maize. Biocontrol Sci Tech. 2, 315-326.

Brand, G., Zingg, D., Querrioux, A., 2017. Virus de la granulose du carpocapse: le bal des isolats. Phytoma. 702, 47-49.

Briese, D. T., 1982. Genetic Basis for Resistance to a Granulosis Virus in the Potato Moth *Phthorimaea operculella*. J Invertebr Pathol. 39, 215-218.

Briese, D. T., Mende, H. A., 1983. Selection for increased resistance to a granulosis virus in the potato moth, *Phthorimaea operculella* (Zeller) (Lepidoptera, Gelichiidae). Bull. Entomol. Res. 73, 1-9.

Brown, M., Faulkner, P., 1978. Plaque assay of nuclear polyhedrosis viruses in cell culture. Appl. Environ. Microbiol. 36, 31-35.

Bull, J. C., Godfray, H. C. J., O'Reilly, D. R., 2003. A Few-Polyhedra Mutant and Wild-Type Nucleopolyhedrovirus Remain as a Stable Polymorphism during Serial Coinfection in Trichoplusia ni. Appl. Environ. Microbiol. 69, 2052-2057.

Caballero, P., Alamo, S., Vargas-Osuna, E., Santiago-Alvarez, C., 1993. Biology of *Cotesia* (= *Apanteles*) *telengai* (Hymenoptera: Braconidae) on its natural host *Agrotis segetum* (Lepidoptera: Noctuidae). Biocontrol Sci. Tech. 3, 481-489.

Castro, M. E. B., Souza, M. L., Araujo, S., Bilimoria, S. L., 1997. Replication of Anticarsia gemmatalis Nuclear Polyhedrosis Virus in Four Lepidopteran Cell Lines. J Invertebr Pathol. 69, 40-45.

Chakraborty, S., Reid, S., 1999. Serial passage of a Helicoverpa armigera nucleopolyhedrovirus in Helicoverpa zea cell cultures. J Invertebr Pathol. 73, 303-308.

Chateigner, A., Bezier, A., Labrousse, C., Jiolle, D., Barbe, V., Herniou, E. A., 2015. Ultra Deep Sequencing of a Baculovirus Population Reveals Widespread Genomic Variations. Viruses. 7, 3625-46.

Clavijo, G., Williams, T., Munoz, D., Caballero, P., Lopez-Ferber, M., 2009. Mixed genotype transmission bodies and virions contribute to the mainenance of diversity in an insect virus. Proc. R. Soc. B. 277, 943-951.

Clement, S. L., McCartney, 1982. Black cutworm (Lepidoptera: Noctuidae) Measurement of larval feeding parameters on field corn in the greenhouse. J Econ Entomol. 75.

Commission, E., 2008. Review Report for the active substance *Cydia pomonella* Granulovirus (Mexican isolate) finalised in the Standing Committee on the Food Chain and Animal Health at its meeting on 11 July 2008 in view of the inclusion of *Cydia pomonella* Granulovirus (Mexican isolate) in Annex I of Directive 91/414/EEC.SANCO/1548/08 – rev.5, 11 July 2014.

Cooper, P. D., 1961. The plaque assay of animal viruses. Adv. Virus. Res. 8, 319-378.

Cory, J. S., Bishop, D. H., 1997. Use of baculoviruses as Biological Insecticides. Mol. Biotechnol. 7, 303-312.

Cory, J. S., Myers, J. H., 2003. The Ecology And Evolution of Insect Baculovirus. Annu. Rev. Ecol. Syst. 34, 239-272.

Crook, N. E., James, J. D., Smith, I. R. L., Winstanley, D., 1997. Comprehensive physical map of the Cydia pomonella granulovirus genome and sequence analysis of the granulin gene region. J Gen Virol. 78, 965-974.

Crook, N. E., Spencer, R. A., Payne, C. C., Leisy, D. J., 1985. Variation in Cydia pomonella Granulovirus Isolates and Physical Maps of the NDA from Three Variants. J. Gen. Virol. 66, 2423-2430.

Darling, A. C. E., Mau, B., Blattner, F. R., Perna, N. T., 2004. Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. Genome Res. 14, 1394-1403.

de Jong, J., Theilmann, D. A., Arif, B. M., Krell, P. J., 2011. Immediate-Early Protein ME53 Forms Foci and Colocalizes with GP64 and the Major Capsid Protein VP39 at the Cell Membranes of Autographa californica Multiple Nucleopolyhedrovirus-Infected Cells. Journal of Virology. 85, 9696-9707.

Duffy, S. P., Young, A. M., Morin, B., Lucarotti, C. J., Koop, B. F., Levin, D. B., 2006. Sequence analysis and organization of the Neodiprion abietis nucleopolyhedrovirus genome. J Virol. 80, 6952-63.

Dugdale, J. S., 1995. Index of economically important Lepidoptera. Compiled by Bin-Cheng Zhang. (Wallingford: CAB INTERNATIONAL, 1994). Bull. Entomol. Res. 85, 302-303.

Eberle, K. E., Asser-Kaiser, S., Sayed, S. M., Nguyen, H. T., Jehle, J. A., 2008. Overcoming the resistance of codling moth against conventional Cydia pomonella granulovirus (CpGV-M) by a new isolate CpGV-I12. J Invertebr Pathol. 98, 293-8.

Eberle, K. E., Sayed, S., Rezapanah, M., Shojai-Estabragh, S., Jehle, J. A., 2009. Diversity and evolution of the *Cydia pomonella* granulovirus. J. Gen. Virol. 90, 662-671.

Eberle, K. E., Wennmann, J. T., Kleespies, R. G., Jehle, J. A., Chapter II: Basic techniques in insect virology In: L. A. Lacey, (Ed.), Manual of Techniques in Invertebrate Pathology Academic Press, Amsterdam, Boston, Heidelberg, New York, Oxford, Paris, San Diego, San Francisco, Sydney, Tokyo, 2012, pp. 15-74.

EFSA, 2012. Conclusion on the peer review of the pesticide risk assessment of the active substance *Cydia pomonella granulovirus* (CpGV). EFSA Journal. 10, 2655.

El-Salamouny, S., Lange, M., Jutzi, M., Huber, J., Jehle, J. A., 2003. Comparative study on the susceptibility of cutworms (Lepidoptera: Noctuidae) to Agrotis segetum nucleopolyhedrovirus and Agrotis ipsilon nucleopolyhedrovirus. J Invertebr Pathol. 84, 75-82.

Eldridge, R., O´Reilly, D. R., Hammock, B. D., Miller, L. K., 1992. Insecticidal Properties of Genetically Engineered Baculoviruses Expressing an Insect Juvenile Hormone Esterase Gene. Appl Environ Microbiol. 58, 1583-1591.

Engelhard, E. K., Kam-Morgan, L. N., Washburn, J. O., Volkman, L. E., 1994. The insect tracheal system: a conduit for the systemic spread of Autographa californica M nuclear polyhedrosis virus. Proceedings of the National Academy of Sciences. 91, 3224-3227.

Engelken, L. K., Showers, W. B., Taylor, S. E., 1990. Weed management to minimize black cutworm (Lepidoptera: Noctuidae) damaga in no-till corn. J Econ Entomol. 83, 1058-1063.

Evans, H. F., 1983. The influence of larval maturation on responses ofMamestra brassicae L. (Lepidoptera: Noctuidae) to nuclear polyhedrosis virus infection. Archives of Virology. 75, 163-170.

Evans, O. P., O´Reilly, D. R., 1998. Purification and kinetic analysis of a baculovirus ecdysteroid UDP-glucosyltransferase. Biochem J. 330, 1265-1270.
Evans, O. P., O´Reilly, D. R., 1999. Expression and structural characterization of a baculovirus ecdysteroid UDP-glucosyltransferase. J Gen Virol. 80, 485-492.
Fang, M., Nie, Y., Harris, S., Erlandson, M. A., Theilmann, D. A., 2009. *Autographa californica* multiple nucleopolyhedrovirus core gene *ac96* encodes a *per os* infectivity factor (PIF-4). J Virol. 83, 12569-12578.
Federici, B. A., Baculovirus Pathogenesis. In: L. K. Miller, (Ed.), The Baculoviruses. Plenum Press, New York 1997, pp. 33-59.
Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M., Danielsen, M., 1987. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proc Natl Acad Sci U S A. 84, 7413-7.
Flipsen, J. T. M., Mans, R. M. W., Kleefsman, A. W. F., Knebel-Mörsdorf, D., Vlak, J. M., 1995. Deletion of the Baculovirus Ecdysteroid UDP-Glucosyltransferase Gene Induces Early Degeneration of Malpighian Tubules in Infected Insects. J Virol. 69, 4529-4532.
Fraser, M. J., 1986. Ultrastructural observations of virion maturation in *Autographa californica* nuclear polyhedrosis virus infected *Spodoptera frugiperda* cell cultures. J. Ultra. Mol. Struct. Res. 95.
Fraser, M. J., Smith, G. E., Summers, M. D., 1983. Acquisition of Host Cell DNA Sequences by Baculoviruses: Relationship Between Host DNA Insertions and FP Mutants of Autographa californica and Galleria mellonella Nuclear Polyhedrosis Viruses. J. Virol. 47, 287-300.
Friesen, P. D., Regulation of baculovirus early gene expression. In: L. K. Miller, (Ed.), The Baculoviruses. Plenum Press, New York and London, 1997, pp. 141-170.
Fritsch, E., Undorf-Spahn, K., Kienzle, J., Zebitz, C. P., Huber, J., 2005. Apfelwickler-Granulovirus: Erste Hinweise auf Unterschiede in der Empfindlichkeit lokaler Apfelwickler-Populationen. Nachrichtenbl. Deut. Pflanzenschutzd. 57, 29-34.
Galloway, C. S., Wang, P., Winstanley, D., Jones, I. M., 2005. Comparison of the bacterial Enhancin-like proteins from Yersinia and Bacillus spp. with a baculovirus Enhancin. J Invertebr Pathol. 90, 134–7.
Garavaglia, M. J., Miele, S. A., Iserte, J. A., Belaich, M. N., Ghiringhelli, P. D., 2012. The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family *Baculoviridae*. J. Virol. 86, 12069-12079.
Garcia-Maruniak, A., Maruniak, J. E., Zanotto, P. M., Doumbouya, A. E., Liu, J. C., Merritt, T. M., Lanoie, J. S., 2004. Sequence analysis of the genome of the Neodiprion sertifer nucleopolyhedrovirus. J Virol. 78, 7036-51.
Gati, C., Oberthuer, D., Yefanov, O., Bunker, R. D., Stellato, F., Chiu, E., Yeh, S.-M., Aquila, A., Basu, S., Bean, R., Beyerlein, K. R., Botha, S., Boutet, S., DePonte, D., Doak, R. B., Fromme, R., Galli, L., Grotjohann, I., James, D. R., Kupitz, C., Lomb, L., Messerschmidt, M., Nass, K., Rendek, K., Shoeman, R. L., Wang, D., Weierstall, U., White, T. A., Williams, G. J., Zatsepin, N. A., Fromme, P., Spence, J., Goldie, K. N., Jehle, J. A., Metcalf, P., Barty, A., Chapman, H. N., 2017. Atomic structure of granulin determined from native nanocrystalline granulovirus using an X-ray free-electron laser. Proc. Natl. Acad. Sci. USA *In Press*.
Gebhardt, M. M., Eberle, K. E., Radtke, P., Jehle, J. A., 2014. Baculovirus resistance in codling moth is virus isolate-dependent and the consequence of a mutation in viral gene pe38. Proc. Natl. Acad. Sci. USA. 111, 15711-15716.
Gilbert, C., Chateigner, A., Ernenwein, L., Barbe, V., Bezier, A., Herniou, E. A., Cordaux, R., 2014. Population genomics supports baculoviruses as vectors of horizontal transfer of insect transposons. Nat Commun. 5, 3348.
Gilbert, C., Peccoud, J., Chateigner, A., Moumen, B., Cordaux, R., Herniou, E. A., 2016. Continuous Influx of Genetic Material from Host to Virus Populations. PLoS Genet. 12, e1005838.
Giri, L., Feiss, M. G., Bonning, B. C., Murhammer, D. W., 2012. Production of baculovirus defective interfering particles during serial passage is delayed by removing transposon target sites in fp25k. J Gen Virol. 93, 389-99.
Giri, L., Li, H., Sandgren, D., Feiss, M. G., Roller, R., Bonning, B. C., Murhammer, D. W., 2010. Removal of transposon target sites from the *Autographa californica* multiple nucleopolyhedrovirus fp25k gene delays, but does not prevent, accumulation of the few polyhedra phenotype. J Gen Virol. 91, 3053-3064.

Gokce, C., Yilmaz, H., Erbas, Z., Demirbag, Z., Demir, I., 2013. First record of *Steinernema kraussei* (Rhabditida: Steinernematidae) from Turkey and its virulence against *Agrotis segetum* (Lepidoptera: Noctuidae). J. Nematol. 45**,** 253-259.

Gordon, A., 2009. FASTX Toolkit. http://hannonlab.cshl.edu/fastx_toolkit/.

Graillot, B., Bayle, S., Blachere-Lopez, C., Besse, S., Siegwart, M., Lopez-Ferber, M., 2016. Biological Characteristics of Experimental Genotype Mixtures of Cydia Pomonella Granulovirus (CpGV): Ability to Control Susceptible and Resistant Pest Populations. Viruses. 8.

Graillot, B., Berling, M., Blachere-Lopez, C., Siegwart, M., Besse, S., Lopez-Ferber, M., 2014. Progressive adaptation of a CpGV isolate to codling moth populations resistant to CpGV-M. Viruses. 6**,** 5135-44.

Graillot, B., Blachere-Lopez, C., Besse, S., Siegwart, M., Lopez-Ferber, M., 2013. Variations in the susceptibility to CpGV in populations of the codling moth, *Cydia pomonella*. IOBC Bull. 90**,** 149-153.

Gross-Bellard, M., Oudet, P., Chambon, P., 1973. Isolation of High-Molecular-Weight DNA from Mammalian Cells. Eur. J. Biochem. 36**,** 32-38.

Guarino, L. A., Gonzalez, M. A., Summers, M. D., 1986. Complete sequence and enhancer function of the homologous DNA regions of Autographa californica nuclear polyhedrosis virus. J Virol. 60**,** 224-229.

Gueli Alletti, G., Carstens, E. B., Weihrauch, B., Jehle, J. A., 2017a. Agrotis segetum nucleopolyhedrovirus but not Agrotis segetum granulovirus replicate in AiE1611T cell line of *Agrotis ipsilon*. J Invertebr Pathol. submitted.

Gueli Alletti, G., Eigenbrod, M., Carstens, E. B., Kleespies, R. G., Jehle, J. A., 2017b. The genome sequence of Agrotis segetum granulovirus, isolate AgseGV-DA, reveals a new *Betabaculovirus* species of a slow killing granulovirus. J Invertebr Pathol. 146**,** 58-68.

Gueli Alletti, G., Sauer, A. J., Weihrauch, B., Fritsch, E., Undorf-Spahn, K., Wennmann, J. T., Jehle, J. A., 2017c. Using Next Generation Sequencing to Identify and Quantify the Genetic Composition of Resistance-Breaking Commercial Isolates of Cydia pomonella Granulovirus. Viruses. 9**,** 17.

Gueli Alletti, G., Wennmann, J. T., Carstens, E. B., Jehle, J. A., 2015. 2015.001aD.A.v2.Betabaculovirus_sp. Downloaded from https://talk.ictvonline.org/files/ictv_official_taxonomy_updates_since_the_8th_report/m/animal-dna-viruses-and-retroviruses/5806 on Jan 18, 2017.

Guo, J., Fu, X., Zhao, X., Wu, K., 2015. Annual Migration of *Agrotis segetum* (Lepidoptera: Noctuidae): Observed on a Small Isolated Island in Northern China. PLoS One. 10**,** e0131639.

Harrap, K. A., 1972. The structure of nuclear polyhedrosis viruses: II. The virus particle. Virology. 50.

Harrison, R., Summers, M. D., 1995. Mutations in the *Autographa californica* multinucleocapsid nuclear polyhedrosis virus 25 kDa protein gene result in reduced virion occlusion, altered intranuclear envelopment and enhanced virus production. J Gen Virol. 76**,** 1451-1459.

Harrison, R. L., 2009. Genomic sequence analysis of the Illinois strain of the Agrotis ipsilon multiple nucleopolyhedrovirus. Virus Genes. 38**,** 155-70.

Harrison, R. L., 2013. Concentration- and time-response characteristics of plaque isolates of Agrotis ipsilon multiple nucleopolyhedrovirus derived from a field isolate. J Invertebr Pathol. 112**,** 159-61.

Harrison, R. L., Herniou, E. A., Jehle, J. A., Theilmann, D. A., Burand, J. P., Becnel, J. J., Krell, P. J., van Oers, M. M., Mowery, J. D., Bauchan, G. R., ICTV Report, C., 2018a. ICTV Virus Taxonomy Profile: Baculoviridae. J Gen Virol. 99**,** 1185-1186.

Harrison, R. L., Lynn, D. E., 2008. New cell lines derived from the black cutworm, Agrotis ipsilon, that support replication of the A. ipsilon multiple nucleopolyhedrovirus and several group I nucleopolyhedroviruses. J Invertebr Pathol. 99**,** 28-34.

Harrison, R. L., Mowery, J. D., Rowley, D. L., Bauchan, G. R., Theilmann, D. A., Rohrmann, G. F., Erlandson, M. A., 2018b. The complete genome sequence of a third distinct baculovirus isolated from the true armyworm, Mythimna unipuncta, contains two copies of the lef-7 gene. Virus Genes. 54**,** 297-310.

Harrison, R. L., Popham, H. J., 2008. Genomic sequence analysis of a granulovirus isolated from the Old World bollworm, Helicoverpa armigera. Virus Genes. 36**,** 565-81.

Harrison, R. L., Rowley, D. L., Funk, C. J., 2016. The Complete Genome Sequence of Plodia Interpunctella Granulovirus: Evidence for Horizontal Gene Transfer and Discovery of an Unusual Inhibitor-of-Apoptosis Gene. PLoS One. 11, e0160389.
Harvey, J. P., Volkman, L. E., 1983. Biochemical and biological variation of Cydia pomonella (Codling moth) granulosis. Virology. 124, 21-24.
Heldens, J. G. M., van Strien, E. A., Feldmann, A. M., Kulcsar, P., Munoz, D., Leisy, D. J., Zuidema, D., Goldbach, R. W., Vlak, J. M., 1996. Spodoptera exigua multicapsid nucleopolyhedrovirus deletion mutants generated in cell culture lack virulence in vivo J. Gen Virol. 77, 3127-3134.
Hernandez, D., Francois, P., Farinelli, L., Osteras, M., Schrenzel, J., 2008. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. Genome Res. 18, 802-809.
Hernandez, D., Tewhey, R., Veyrieras, J.-B., Farinelli, L., Osteras, M., Francois, P., Schrenzel, J., 2014. De novo finished 2.8 Mbp Staphylococcus aureus genome assembly from 100 bp short and long range paired-end reads. Bioinformatics. 30, 40–49.
Herniou, E. A., Arif, B. M., Becnel, J. J., Blissard, G. W., Bonning, B., Harrison, R., Jehle, J. A., Theilmann, D. A., Vlak, J. M., *Baculoviridae*. In: A. M. Q. King, et al., Eds.), Virus Taxonomy: Classification and nomenclature of viruses : Ninth report of the International Committee on the Taxonomy of Viruses. Elsevier, 2011, pp. 163-173.
Herniou, E. A., Jehle, J. A., 2007. Baculovirus phylogeny and evolution. Curr. Drug Targets. 8, 1043-1050.
Herniou, E. A., Olszewski, J. A., Cory, J. S., O'Reilly, D. R., 2003. The Genome Sequence and Evolution of Baculoviruses. Annu. Rev. Entomol. 48, 211-234.
Herniou, E. A., Olszewski, J. A., O'Reilly, D. R., Cory, J. S., 2004. Ancient Coevolution of Baculoviruses and Their Insect Hosts. Journal of Virology. 78, 3244-3251.
Hill, D. S., 1983. Agricultural insect pests of tropics and their control. Camebridge University Press.
Hilton, S., Kemp, E., Keane, G., Winstanley, D., 2008. A bacmid approach to the genetic manipulation of granuloviruses. Journal of Virological Methods. 152, 56-62.
Hilton, S., Winstanley, D., 2007. Identification and functional analysis of the origins of DNA replication in the *Cydia pomonella* granulovirus genome. J Gen Virol. 88, 1496-1504.
Hilton, S., Winstanley, D., 2008a. Biological characterization of an English granulovirus from the summer fruit tortrix moth, *Adoxophyes orana*. J. Invertebr. Pathol. 97, 298-305.
Hilton, S., Winstanley, D., 2008b. The origins of replication of granuloviruses. Arch. Virol. 153, 1527-1535.
Hofacker, I. L., 2003. Vienna RNA secondary structure server. Nucl. Acid. Res. 31, 3429-3431.
Huber, J., Use of Baculoviruses in Pest Management Programs. The Biology of Baculoviruses. Vol. II: Practical Applications for Insect Control, 1986, pp. 181-202.
Huynh, H. T., Tran, T. T. B., Chan, L. C. L., Nielsen, L. K., Reid, S., 2015. Decline in Helicoverpa armigera nucleopolyhedrovirus occlusion body yields with increasing infection cell density in vitro is strongly correlated with viral DNA levels. Arch Virol. 160, 2169-2180.
Ignoffo, C. M., Garcia, C., 1979. Susceptibility of larvae of the black cutworm to species of entomopathogenic bacteria, fungi, protozoa, and viruses. J. Econ. Entomol. 72, 767-769.
Ivaldi-Sender, C., 1974. A simple technique for the permanent rearing of the oriental fruit moth, *Grapholita molesta* (Lepidoptera Tortricidae) on an artificial medium. Techniques simples pour un elevage permanent de la tordeuse orientale, *Grapholita molesta* (Lepidoptera Tortricidae) sur milieu artificiel. Annales de Zoologie, Ecologie Animale. 6, 337-343.
Iwata, K., Haas-Stapleton, E. J., Kunimi, Y., Inoue, M. N., Nakai, M., 2017. Midgut based resistance to oral infection by a nucleopolyhedrovirus 1 in the laboratory-selected strain of the small tea tortrix, *Adoxophyses honmai* (Lepidoptera: Totridicidae) after selection by serial viral administration J Invertebr Pathol. 145, 23-30.
Jakubowska, A. K., Murillo, R., Carballo, A., Williams, T., van Lent, J. W., Caballero, P., Herrero, S., 2016. Iflavirus increases its infectivity and physical stability in association with baculovirus. PeerJ. 4, e1687.
Jakubowska, A. K., Peters, S. A., Ziemnicka, J., Vlak, J. M., van Oers, M. M., 2006. Genome sequence of an enhancin gene-rich nucleopolyhedrovirus (NPV) from Agrotis segetum: collinearity with Spodoptera exigua multiple NPV. J Gen Virol. 87, 537-51.
Javed, M. A., Biswas, S., Willis, L. G., Harris, S., Pritchard, C., van Oers, M. M., Donly, B. C., Erlandson, M. A., Hegedus, D. D., Theilmann, D. A., 2017. Autographa californica Multiple

Nucleopolyhedrovirus AC83 is a Per Os Infectivity Factor (PIF) Protein Required for Occlusion-Derived Virus (ODV) and Budded Virus Nucleocapsid Assembly as well as Assembly of the PIF Complex in ODV Envelopes. J Virol. 91.

Jehle, J. A., Backhaus, H., Fritsch, E., Huber, J., 1992. Physical map of the Cryptophlebia leucotreta granulosis virus genome and its relationship to the genome of Cydia pomoneila granulosis virus. J Gen Virol. 73, 1621-1626.

Jehle, J. A., Blissard, G. W., Bonning, B. C., Cory, J. S., Herniou, E. A., Rohrmann, G. F., Theilmann, D. A., Thiem, S. M., Vlak, J. M., 2006a. On the classification and nomenclature of baculoviruses: a proposal for revision. Arch. Virol. 151, 1257-1266.

Jehle, J. A., Lange, M., Wang, H., Hu, Z., Wang, Y., Hauschild, R., 2006b. Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera. Virology. 346, 180-193.

Jehle, J. A., Schulze-Bopp, S., Undorf-Spahn, K., Fritsch, E., 2017. Evidence for a Second Type of Resistance against Cydia pomonella Granulovirus in Field Populations of Codling Moths. Appl. Environ. Microbiol. 83.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., Madden, T. L., 2008. NCBI BLAST: a better web interface. Nucl. Acid Res. 36, W5-9.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28, 1647-9.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111-120.

Kool, M., Ahrens, C. H., Goldbach, R. W., Rohrmann, G. F., Vlak, J. M., 1994. Identification of genes involved in DNA replication of the Autographa californica baculovirus. Proceedings of the National Academy of Sciences of the United States of America. 91, 11212-11216.

Kool, M., van den Berg, P. M. M. M., Tramper, J., Goldbach, R. W., Vlak, J. M., 1993. Location of Two Putative Origins of DNA Replication of *Autographa californica* Nuclear Polyhedrosis Virus. Virology. 192, 94-101.

Kool, M., Voncken, J. W., Van Lier, F. L. J., Tramper, J., Vlak, J. M., 1991. Detection and Analysis of Autographa californica Nuclear Polyhedrosis Virus Mutants with Defective Interfering Properties. J. Virol. 183, 739-746.

Kovacs, G. R., Choi, J., Guarino, L. A., Summers, M. D., 1992. Functional Dissection of the Autographa califomica Nuclear Polyhedrosis Virus Immediate-Early 1 Transcriptional Regulatory Protein. J Virol. 66, 7429-7437.

Krell, P. J., 1996. Passage effect of virus infection in insect cells. Cytotechnology. 20, 125-137.

Kumar, S., Stetcher, G., Tamura, K., 2016. Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870-1874.

Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., Giegerich, R., 2001. REPuter: The Manifold Applications of Repeat Analysis on a Genomic Scale. Nucl. Acid. Res. 29, 4633-4642.

Lacey, L. A., Arthurs, S. P., Headrick, H., 2005. Comparative Activity of the Codling Moth Granulovirus Against *Grapholita molesta* and *Cydia pomonella* (Lepidoptera: Tortricidae). J Entomol Soc Brit Columbia. 102, 79-80.

Lacey, L. A., Grzywacz, D., Shapiro-Ilan, D. I., Frutos, R., Brownbridge, M., Goettel, M. S., 2015. Insect pathogens as biological control agents: Back to the future. Journal of Invertebrate Pathology. 132, 1-41.

Lange, M., Jehle, J. A., 2003. The genome of the Cryptophlebia leucotreta granulovirus. Virology. 317, 220-236.

Langmead, B., Salzberg, S. L., 2012a. Fast gapped-read alignment with Bowtie 2. Nat Meth. 9, 357-359.

Langmead, B., Salzberg, S. L., 2012b. Fast gapped-read alignment with Bowtie 2. Nat. Meth. 9, 357-359.

Lauzon, H. A., Lucarotti, C. J., Krell, P. J., Feng, Q., Retnakaran, A., Arif, B. M., 2004. Sequence and organization of the Neodiprion lecontei nucleopolyhedrovirus genome. J Virol. 78, 7023-35.

Lehmann, R., Bachmann, J., Maletzki, D., Polleichtner, C., Ratte, H. T., Ratte, M., 2016. A new approach to overcome shortcomings with multiple testing of reproduction data in ecotoxicology. Stochastic Environmental Research and Risk Assessment. 30, 871-882.

Leisy, D. J., Rohrmann, G. F., 1993. Characterization of the replication of plasmids containing *hr* sequences in baculovirus-infected *Spodoptera frugiperda* cells. Virology 196, 722-730.
Lepore, L. S., Roelvink, P. R., Granados, R. R., 1996a. Enhancin, the granulosis virus protein that facilitates nucleopolyhedrovirus (NPV) infections, is a metalloprotease. J Invertebr Pathol. 68, 131-140.
Lepore, L. S., Roelvink, P. R., Granados, R. R., 1996b. Enhancin, the Granulosis Virus Protein That Facilitates Nucleopolyhedrovirus (NPV) Infections, Is a Metalloprotease. J. Invertebr. Pathol. 68.
Li, H., 2011a. Improving SNP discovery by base alignment quality. Bioinformatics. 27, 1157-8.
Li, H., 2011b. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 27, 2987-93.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25, 1754-60.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data Processing, S., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25, 2078-9.
Lopez-Ferber, M., Simon, O., Williams, T., Caballero, P., 2003. Defective or effective? Mutualistic interactions between virus genotypes. Proc. R. Soc. B. 270, 2249-2255.
Lowe, R. E., Paschke, J. D., 1968. Simultaneous Infection with the Nucleopolyhedrosis and Granulosis Viruses of *Trichoplusia ni*. J Invertebr Pathol. 12.
Lu, A., Miller, L. K., Regulation of baculovirus late and very late gene expression. In: L. K. Miller, (Ed.), The Baculovirus. Plenum Press, New York and London, 1997, pp. 193-216.
Lua, L. H. L., Pedrini, M. R. S., Reid, S., Robertson, A., Tribe, D. E., 2002. Phenotypic and genotypic analysis of Helicoverpa armigera nucleopolyhedrovirus serially passaged in cell culture. J Gen Virol. 83, 945-955.
Lung, O., Westenberg, M., Vlak, J. M., Zuidema, D., Blissard, G. W., 2002. Pseudotyping Autographa californica Multicapsid Nucleopolyhedrovirus (AcMNPV): F Proteins from Group II NPVs Are Functionally Analogous to AcMNPV GP64. Journal of Virology. 76, 5729-5736.
Luque, T., Finch, R., Crook, N. E., O'Reilly, D. R., Winstanley, D., 2001. The complete sequence of the Cydia pomonella granulovirus genome. J Gen Virol. 82, 2531-2547.
Lynn, D. E., 2002. Methods for maintaining insect cell cultures. J Insect Sci. 2, 9.
Lynn, D. E., 2003. Comparative susceptibilities of insect cell lines to infection by the occlusion-body derived phenotype of baculoviruses. Journal of Invertebrate Pathology. 83, 215-222.
Lynn, D. E., Harrison, R. L., Available Lepidopteran Insect Cell Lines. In: W. D. Murhammer, (Ed.), Baculovirus and Insect Cell Expression Protocols. Springer New York, New York, NY, 2016, pp. 119-142.
MacQueen, J. B., Some Methods for Classification and Analysis of Multivariate Observations. In: L. M. Le Cam, J. Neyman, Eds.), Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, CA, USA, 1966, pp. 17.
McClintock, J. T., Dougherty, E. M., 1987. Superinfection of baculovirus-infected gypsy moth cells with the nuclear polyhedrosis virus of *Autographa californica* and *Lymantra dispar*. Virus Res. 7.
Merryweather-Clarke, A. T., Hirst, M., Possee, R. D., 1994. Recombination between genetically modified and unmodified Autographa californica nuclear polyhedrosis virus in Trichoplusia ni larvae. Acta Virol. 38, 311-315.
Mielczarek, M., Szyda, J., 2016. Review of alignment and SNP calling algorithms for next-generation sequencing data. Journal of Applied Genetics. 57, 71-79.
Miele, S. A., Garavaglia, M. J., Belaich, M. N., Ghiringhelli, P. D., 2011. Baculovirus: molecular insights on their diversity and conservation. Int J Evol Biol. 2011, 379424.
Mikhailov, V. S., Rohrmann, G. F., 2002. The Baculovirus Replication Factor LEF-1 is a DNA Primase. Journal of Virology. 76, 2287-2297.
Milks, M. L., Myers, J. H., 2000. The development of larval resistance to a nucleopolyhedrovirus is not accompanied by an increased virulence in the virus. Evol Ecol. 14, 654-664.
Milks, M. L., Theilmann, D. A., 2000. Serial Selection for Resistance to a Wild-Type and to a Genetically Modified Nucleopolyhedrovirus in Trichoplusia ni. Biol. Control. 19, 283-289.
Miller, L. K., Introduction to the Baculoviruses. In: L. K. Miller, (Ed.), The Baculoviruses. Springer US, 1997, pp. 1-6.

Monsma, S. A., Oomens, A. G. P., Blissard, G. W., 1996. The *gp64* envelope fusion protein is an essential baculovirus protein required for cell-to-cell transmission of infection. J. Virol. 70, 4607-4616.

Moscardi, F., 1999. Assessment of the appilcation of baculoviruses for the control of Lepidoptera. Annu. Rev. Entomol. 44, 257-289.

Mukawa, S., Goto, C., 2008. In vivo characterization of two granuloviruses in larvae of *Mythimna separata* (Lepidoptera: Noctuidae). J. Gen. Virol. 89, 915-921.

Nakai, M., Takahashi, K., Iwata, K., Koyanagi, J., Ookuma, A., Takatsuka, J., Okuno, S., Kunimi, Y., 2017. Acquired resistance to a nucleopolyhedrovirus in the smaller tea tortrix Adoxophyes honmai (Lepidoptera: Tortricidae) after selection by serial viral administration. J. Invertebr. Pathol. 145, 23-30.

Nguyen, Q., Qi, Y. M., Wu, Y., Chan, L. C., Nielsen, L. K., Reid, S., 2011. In vitro production of Helicoverpa baculovirus biopesticides--automated selection of insect cell clones for manufacturing and systems biology studies. J Virol Methods. 175, 197-205.

Nie, Y., Fang, M., Erlandson, M. A., Theilmann, D. A., 2012. Analysis of the *Autographa californica* multiple nucleopolyhedrovirus overlapping gene pair *lef3* and *ac68* reveals that AC68 is a *per os* infectivity factor and that LEF3 is critical, but not essential, for virus replication. J Virol. 86, 3985-3994.

Noune, C., Hauxwell, C., 2016. Comparative Analysis of HaSNPV-AC53 and Derived Strains. Viruses. 8.

O'Reilly, D. R., Miller, L. K., Luckov, V. A., 1994. Baculovirus expression vectors: a laboratory manual. Oxford University Press, New York

O´Reilly, D. R., Miller, L. K., 1989. A Baculovirus Blocks Insect Molting by Producing Ecdysteroid UDP-Glucosyl Transferase. Science. 245, 1110-1112.

O´Reilly, D. R., Miller, L. K., 1990. Regulation of Expression of a Baculovirus Ecdysteroid UDPglucosyltransferase Gene. J Virol. 64, 1321-1328.

OECD, 2002. Consensus Document on Information Used in the Assessment of Environmental Applications Involving Baculoviruses. Series on Harmonization of Regulatory Oversight in Biotechnology. No. 20, ENV/JM/MONO(2002)1.

Paillot, A., 1926. Sur un nouvelle maladie du noyau ou grasserie des chenilles de *P. brassicae* et un nouveau groupe de microorganisms parasites. C. R. Acad. Sci. 182, 180-182.

Pang, Y., Yu, J., Wang, L., Hu, X., Bao, W., Li, G., Chen, C., Han, H., Hu, S., Yang, H., 2001. Sequence Analysis of the Spodoptera litura Multicapsid Nucleopolyhedrovirus Genome. Virology. 287, 391-404.

Passarelli, A. L., Westpheling, J., Espelie, E. M., Espelie, K., 2014. Lois K. Miller - A Biographiocal Memoire. National Academy of Science.

Pearson, M., Bjornson, R., Pearson, G., Rohrmann, G., 1992. The Autographa californica baculovirus genome: evidence for multiple replication origins. Science. 257, 1382-1384.

Pearson, M. N., Bjornson, R. M., Ahrens, C. H., Rohrmann, G. F., 1993. Identification and characterization of a putative origin of DNA replication in the genome of a baculovirus pathogenic for *Orgyia pseudotsugata*. Virology 197, 715-725.

Pearson, M. N., Groten, C., Rohrmann, G. F., 2000. Identification of the *Lymantria dispar* nucleopolyhedrovirus envelope fusion protein provides evidence for a phylogenetic division of the *Baculoviridae*. J. Virol. 74, 6126-6131.

Pearson, M. N., Rohrmann, G. F., 1995. *Lymantria dispar* Nuclear Polyhedrosis Virus Homologous Regions: Characterization of Their Ability To Function as Replication Origins. J. Virol. 69, 213-221.

Peng, K., van Oers, M. M., Hu, Z., van Lent, J. W. M., Vlak, J. M., 2010. Baculovirus Per Os Infectivity Factors Form a Complex on the Surface of Occlusion-Derived Virus. Journal of Virology. 84, 9497-9504.

Perera, O., Green, T. B., Stevens, S. M., Jr., White, S., Becnel, J. J., 2007. Proteins associated with Culex nigripalpus nucleopolyhedrovirus occluded virions. J Virol. 81, 4585-90.

Pfeifer, T. A., Hegedus, D. D., Grigliatti, T. A., Theilmann, D. A., 1997. Baculovirus immediate-early promoter-mediated expression of the Zeocin™ resistance gene for use as a dominant selectable marker in Dipteran and Lepidopteran insect cell lines. Gene. 188, 183-190.

Pijlman, G. P., Dortmans, J. C. F. M., Vermeesch, A. M. G., Yang, K., Martens, D. E., Goldbach, R. W., Vlak, J. M., 2002. Pivotal Role of the Non-hr Origin of DNA Replication in the Genesis of Defective Interfering Baculoviruses. Journal of Virology. 76, 5605-5611.

Pijlman, G. P., van den Born, E., Martens, D. E., Vlak, J. M., 2001. Autographa californica baculoviruses with large genomic deletions are rapidly generated in infected insect cells. Virology. 283, 132-138.

Potter, D. A., 1998. Destructive Turfgrass Insects: Biology, Diagnosis and Control. Wiley, New York.

Potter, K. N., Jaques, R. P., Faulkner, P., 1978. Modification of Trichoplusia ni nuclear polyhedrosis virus passaged in vivo. Intervirology. 9, 76-85.

Pullen, S. S., Friesen, D. R., 1995a. The CAGT Motif Functions as an Initiator Element during Early Transcription of the Baculovirus Transregulator *ie-1*. J Virol. 69, 3575-3583.

Pullen, S. S., Friesen, D. R., 1995b. Early Transcription of the *ie-1* Transregulator Gene of *Autographa californica* Nuclear Polyhedrosis Virus Is Regulated by DNA Sequences within Its 5' Noncoding Leader Region. J Virol. 69, 156-165.

R Development Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2015.

Rawlings, N. D., Barret, A. J., 1995. Evolutionary Families of Metallopeptidases. Methods Enzymol. 248, 183-228.

Reed, L., Muench, H., 1938. A simple method for estimating fifty percent endpoints Am. J. Hyg. 27, 493-497.

Rezapanah, M., Shojai-Estabragh, S., Huber, J., Jehle, J. A., 2008. Molecular and biological characterization of new isolates of *Cydia pomonella* granulovirus from Iran. J. Pest. Sci. 81, 187-191.

Robinson, D., broom: Convert Statistical Analysis Objects into Tidy Data Frames. 2016.

Rohrmann, G. F., 2013. Baculovirus Molecular Biology.

Ros, V. I., van Houte, S., Hemerik, L., van Oers, M. M., 2015. Baculovirus-induced tree-top disease: how extended is the role of egt as a gene for the extended phenotype? Mol Ecol. 24, 249-58.

Rose, J., Kleespies, R. G., Wang, Y., Wennmann, J. T., Jehle, J. A., 2013. On the susceptibility of the box tree moth *Cydalima perspectalis* to Anagrapha falcifera nucleopolyhedrovirus (AnfaNPV). J. Invertebr. Pathol. 113, 191-197.

Rzhetsky, A., Nei, M., 1992. A simple method for estimating and testing minimum evolution trees. Mol. Biol. Evol. 9, 945-967.

Sambrook, J., 2001. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sambrook, J., Russell, D. W., 2001. Molecular cloning : a laboratory manual Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sauer, A. J., Fritsch, E., Undorf-Spahn, K., Nguyen, H. T., Frantisek, M., Heckel, D. G., Jehle, J. A., 2017a. Novel Resistance to Cydia pomonella granulovirus (CpGV) in Codling Moth shows Autosomal and Dominant Inheritance and Confers Cross-Resistance to Different CpGV Genome Groups Submitted to PLoS One.

Sauer, A. J., Gebhardt, M. M., Fritsch, E., Undorf-Spahn, K., Iwata, K., Kleespies, R. G., Nakai, M., Jehle, J. A., 2017b. Cross-resistance of coding moth against different isolates of Cydia pomonella granulovirus is caused by two different but genetically linked resistance mechanisms. Submitted to J Virol.

Sauer, A. J., Schulze-Bopp, S., Fritsch, E., Undorf-Spahn, K., Jehle, J. A., 2017c. A third type of resistance of codling moth against Cydia pomonella granulovirus (CpGV) shows a mixture of a Z-linked and autosomal inheritance pattern. Submitted to AEM.

Sauphanor, B., Berling, M., Toubon, J.-F., Reyes, M., Delnatte, J., 2006. Carpocapse des pommes cas de resistance aux virus de la granulose dans le Sud-Est. Phytoma. 590, 24-27.

Schmitt, A., Bisutti, I. L., Ladurner, E., Benuzzi, M., Sauphanor, B., Kienzle, J., Zingg, D., Undorf-Spahn, K., Fritsch, E., Huber, J., Jehle, J. A., 2013. The occurrence and distribution of resistance of codling moth to Cydia pomonellagranulovirus in Europe. J Appl Entomol. 137, 641-649.

Showers, W. B., Von Kuster, G., Mulder, P., 1983. Corn seedling growth stage and black cutworm damage. Environ Entomol. 12, 241-244.

Simon, O., Williams, T., Lopez-Ferber, M., Caballero, P., 2005. Functional Importance of Deletion Mutant Genotypes in an Insect Nucleopolyhedrovirus Population. Appl Environ Microbiol. 71, 4254-4262.

Simon, O., Williams, T., Possee, R. D., Lopez-Ferber, M., Caballero, P., 2010. Stability of a *Spodoptera frugiperda* Nucleopolyhedrovirus Deletion Recombinant during Serial Passage in Insects. Appl Environ Microbiol. 76, 803-809.

Slack, J., Arif, B. M., The Baculoviruses Occlusion-Derived Virus: Virion Structure and Function. Adv. Virus Res. Academic Press, 2006, pp. 99-165.

Slavicek, J. M., Hayes-Plazolles, N., Kelly, M. E., 2001. Identification of a Lymantria dispar Nucleopolyhedrovirus Isolate That Does Not Accumulate Few-Polyhedra Mutants during Extended Serial Passage in Cell Culture. Biological Control. 22, 159-168.

Slavicek, J. M., Popham, H. J., 2005. The Lymantria dispar nucleopolyhedrovirus enhancins are components of occlusion-derived virus. J Virol. 79, 10578-88.

Smits, P. H., Vlak, J. M., 1988. Biological activity of Spodoptera exigua nuclear polyhedrosis virus against S. exigua larvae. Journal of Invertebrate Pathology. 51, 107-114.

Sparks, W. O., Harrison, R. L., Bonning, B. C., 2011. *Autographa californica* multiple nucleopolyhedrovirus ODV-E56 is a *per os* infectivity factor, but is not essential for binding and fusion of occlusion-derived virus to the host midgut. Virology. 409, 69-76.

Takahashi, M., Nakai, M., Saito, Y., Sato, Y., Ishijima, C., Kunimi, Y., 2015. Field efficacy and transmission of fast- and slow-killing nucleopolyhedroviruses that are infectious to Adoxophyes honmai (Lepidoptera: Tortricidae). Viruses. 7, 1271-83.

Tanada, Y., 1959. Synergism between two viruses of the armyworm, *Pseudaletia unipuncta* (Haworth) (Lepidoptera: Noctuidae). J Insect Pathol. 1, 215-231.

Tanada, Y., 1964. A granulosis virus of the codling moth, *Carpocapsa pomonella* (Linnaeus) (Olethreutidae, Lepidoptera). J Invertebr Pathol. 6, 378-380.

Tanada, Y., Hukuhara, T., 1971. Enhanced infection of a nuclear-polyhedrosis virus in larvae of the armyworm, Pseudaletia unipuncta, by a factor in the capsule of a granulosis virus. J Invertebr Pathol. 17, 116-126.

Tanada, Y., Leutenegger, R., 1968. Histopathology of a Granulosis-Virus Disease of the Codling Moth, *Carpocapsa pomonella*. J. Invertebr. Pathol. 10, 39-47.

Theilmann, D. A., Blissard, G. W., Bonning, B. C., Jehle, J. A., O´Reilly, D. R., Rohrmann, G. F., Thiem, S. M., Vlak, J. M., *Baculoviridae*. Virus Taxonomy. Elsevier, 2005, pp. 177 - 185.

Theilmann, D. A., Stewart, S., 1991. Identification and characterization of the *ie-1* gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. Virology. 180, 492-508.

Tomalski, M. D., Wu, J., Miller, L. K., 1986. The location, sequence, transcription and regulation of a baculovirus DNA polymerase gene. Virology. 167, 591-600.

Toprak, U., Harris, S., Baldwin, D., Theilmann, D., Gillott, C., Hegedus, D. D., Erlandson, M. A., 2012. Role of enhancin in Mamestra configurata nucleopolyhedrovirus virulence: selective degradation of host peritrophic matrix proteins. J. Gen. Virol. 93, 744-753.

Union, E., 2009. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJ No L309/1. Official Journal of the European Union. L 309/1.

Union, E., 2013. Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. Official Journal of the European Union. L 93/1.

van der Merwe, M., Jukes, M. D., Rabalski, L., Knox, C., Opoku-Debrah, J. K., Moore, S. D., Krejmer-Rabalska, M., Szewczyk, B., Hill, M. P., 2017. Genome Analysis and Genetic Stability of the Cryptophlebia leucotreta Granulovirus (CrleGV-SA) after 15 Years of Commercial Use as a Biopesticide. Int J Mol Sci. 18.

van Houte, S., Ros, V. I., van Oers, M. M., 2014a. Hyperactivity and tree-top disease induced by the baculovirus AcMNPV in Spodoptera exigua larvae are governed by independent mechanisms. Naturwissenschaften. 101, 347-50.

van Houte, S., van Oers, M. M., Han, Y., Vlak, J. M., Ros, V. I., 2014b. Baculovirus infection triggers a positive phototactic response in caterpillars to induce 'tree-top' disease. Biol Lett. 10, 20140680.

Virto, C., Navarro, D., Tellez, M. M., Herrero, S., Williams, T., Murillo, R., Caballero, P., 2014. Natural populations of Spodoptera exigua are infected by multiple viruses that are transmitted to their offspring. Journal of Invertebrate Pathology. 122, 22-27.

Volkman, L. E., Summers, M. D., 1977. Autographa californica nuclear polyhedrosis virus: Comparative infectivity of the occluded, alkali-liberated, and nonoccluded forms. Journal of Invertebrate Pathology. 30, 102-103.

Wang, M., Wang, J., Yin, F., Tan, Y., Deng, F., Chen, X., Jehle, J. A., Vlak, J. M., Hu, Z., Wang, H., 2014. Unraveling the Entry Mechanism of Baculoviruses and Its Evolutionary Implications. Journal of Virology. 88, 2301-2311.

Wang, Q., Bosch, B.-J., Vlak, J. M., van Oers, M. M., Rottier, P. J., van Lent, J. W. M., 2016. Budded baculovirus particle structure revisited. Journal of Invertebrate Pathology. 134, 15-22.

Washburn, J. O., Chan, E. Y., Volkman, L. E., Aumiller, J. J., Jarvis, D. L., 2003. Early Synthesis of Budded Virus Envelope Fusion Protein GP64 Enhances Autographa californica Multicapsid Nucleopolyhedrovirus Virulence in Orally Infected Heliothis virescens. Journal of Virology. 77, 280-290.

Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M., Losick, R. M., 2004. Molecular Biology of the Gene.

Wennmann, J. T., Diversity of Baculovirus Isolated from Cutworms (*Agrotis* spp.). Fachbereich Biologie Vol. Doctor rerum naturalium. Johannes Gutenberg-University, Mainz, 2014, pp. 131.

Wennmann, J. T., Gueli Alletti, G., Jehle, J. A., 2015a. The genome sequence of Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B) reveals a new baculovirus species within the Agrotis baculovirus complex. Virus Genes. 50, 260-76.

Wennmann, J. T., Jehle, J. A., 2014. Detection and quantitation of Agrotis baculoviruses in mixed infections. J. Virol. Methods. 197, 39-46.

Wennmann, J. T., Keilwagen, J., Jehle, J. A., 2018. Baculovirus Kimura two-parameter species demarcation criterion is confirmed by the distances of 38 core gene nucleotide sequences. J Gen Virol.

Wennmann, J. T., Kohler, T., Gueli Alletti, G., Jehle, J. A., 2015b. Mortality of cutworm larvae is not enhanced by Agrotis segetum granulovirus and Agrotis segetum nucleopolyhedrovirus B coinfection relative to single infection by either virus. Appl. Environ. Microbiol. 81, 2893-2899.

Wennmann, J. T., Kohler, T., Gueli Alletti, G., Jehle, J. A., 2015c. Mortality of cutworm larvae is not enhanced by Agrotis segetum granulovirus and Agrotis segetum nucleopolyhedrovirus B coinfection relative to single infection by either virus. Appl Environ Microbiol. 81, 2893-9.

Wennmann, J. T., Radtke, P., Eberle, K. E., Gueli Alletti, G., Jehle, J. A., 2017. Evolutionary Trends in Isolates of the Cydia pomonella granulovirus based on full genome sequence analysis. Submitted to Viruses.

Winstanley, D., Crook, N. E., 1993. Replication of Cydia pomonella granulosis virus in cell cultures. J. Gen Virol. 74, 1599-1609.

Wormleaton, S. L., Winstanley, D., 2001. Phylogenetic analysis of conserved genes within the ecdysteroid UDP-glucosyltransferase gene region of the slow-killing *Adoxophyes orana* granulovirus. J. Gen. Virol. 82, 2295-2305.

Wraight, S. P., Ramos, M. E., Avery, P. B., Jaronski, S. T., Vandenberg, J. D., 2010. Comparative virulence of Beauveria bassiana isolates against lepidopteran pests of vegetable crops. J. Invertebr. Pathol. 103, 186-199.

Wu, C. P., Huang, Y.-J., Wang, J.-Y., Wu, Y.-L., Lo, H.-R., Wang, J.-C., Chao, Y.-C., 2010. Autographa californica Multiple Nucleopolyhedrovirus LEF-2 Is a Capsid Protein Required for Amplification but Not Initiation of Viral DNA Replication. Journal of Virology. 84, 5015-5024.

Wu, Y., Liu, G., Carstens, E. B., 1999. Replication, Integration, and Packaging of Plasmid DNA following Cotransfection with Baculovirus Viral DNA. Journal of Virology. 73, 5473-5480.

Xeros, N., 1956. The Virogenic Stroma in Nuclear and Cytoplasmic Polyhedroses. Nature. 178, 412-413.

Xi, Q., Wang, J., Deng, R., Wang, X., 2007. Characterization of AcMNPV with a deletion of me53 gene. Virus Genes. 34, 223-232.

Yang, M. M., Li, M. L., Zhang, Y. A., Whang, Y. Z., Qu, L. J., Wang, Q. H., Ding, J. Y., 2012. Baculoviruses and insect pests control in China. Afr. J. Microbiol. Res. 6, 214-218.

Ye, H., Meehan, J., Tong, W., Hong, H., 2015. Alignment of Short Reads: A Crucial Step for Application of Next-Generation Sequencing Data in Precision Medicine. Pharmaceutics. 7, 523-41.

Ye, K., Schulz, M. H., Long, Q., Apweiler, R., Ning, Z., 2009. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics. 25, 2865-71.

Yin, C., Yu, J., Wang, L., Li, Z., Zhang, P., Pang, Y., 2003. Identification of a Novel Protein Associated with Envelope of Occlusion-derived Virus in *Spodoptera litura* Multicapsid Nucleopolyhedrovirus. Virus Genes. 26, 5-13.

Zethner, O., 1980. Control of *Agrotis segetum* [Lep.: Noctuidae] root crops by granulosis virus. Entomophaga. 25, 27-35.

Zethner, O., Khan, B. M., Chaudhry, M. I., Boleti, B., Khan, S., Khan, H., Gul, G., Ogaard, L., Zamnan, M., Nawaz, G., 1987. *Agrotis segetum* granulovirus as a control agent against field populations of *Agrotis ipsilon* and *A. segetum* [Lep.: Noctuidae] on tobacco, okra, potato and sugar beet in Northern Pakistan. Entomophaga. 32, 449-455.

Zhang, X., Liang, Z., Yin, X., Wang, J., Shao, X., 2014. Complete genome sequence of Agrotis segetum granulovirus Shanghai strain. Arch. Virol. 159, 1869-1872.

Zheng, Q., Shen, Y., Kon, X., Zhang, J., Feng, M., Wu, X., 2017. Protein–protein interactions of the baculovirus per os infectivity factors (PIFs) in the PIF complex. Journal of General Virology. 98, 853-861.

Zhu, S., Wang, W., Wang, Y., Yuan, M., Yang, K., 2013. The baculovirus core gene ac83 is required for nucleocapsid assembly and per os infectivity of *Autographa californica* nucleopolyhedrovirus. J Virol. 87, 10573-10586.

Zichová, T., Stará, J., Kundu, J. K., Eberle, K. E., Jehle, J. A., 2013. Resistance to Cydia pomonella granulovirus follows a geographically widely distributed inheritance type within Europe. BioControl. 58, 525-534.

Ziemann, M., 2016. Accuracy, speed and error tolerance of short DNA sequence aligners. bioRxiv.

Zingg, D., 2010. Madex Plus and Madex I12 overcome virus resistance of codling moth. In Ecofruit: proceedings to the 13th International Conference on Cultivation Technique and Phytopathological Problems in Organic Fruit-Growing. Föko e.V., Weinsberg, Germany. 256-260.

Zingg, D., Kraaz, I., Wandeler, H., Züger, M., Madex Twin, a new Cydia pomonella granulovirus isolate for the control of both codling moth Cydia pomonella and oriental fruit moth Grapholita molesta. Fördergemeinschaft Ökologischer Obstbau e.V. (FÖKO), Weinsberg, 2012, pp. 187-193.

Zwart, M. P., Erro, E., van Oers, M. M., de Visser, J. A., Vlak, J. M., 2008. Low multiplicity of infection in vivo results in purifying selection against baculovirus deletion mutants. J Gen Virol. 89, 1220-4.

Summary

The family of *Baculoviridae* comprises occluded double-stranded DNA viruses with rod-shaped, enveloped virions infecting larval stages of the insect orders Lepidoptera, Diptera and Hymoptera. Due to their narrow host range, their high virulence to target insects and the absence of adverse effects to human health and the environment, several baculovirus species are already in use as biological control agents in pest control.

One possibility for their use, is for control of cutworms. Cutworms are severe soil pests of many agricultural and horticultural crops that live near or underneath the soil surface and feed polyphagous on seedlings, stems, roots and other parts of the plant and among other species they also comprise the genus *Agrotis*. In the past, four baculovirus species were isolated from *Agrotis segetum* and *A. ipsilon*, namely: Agrotis segetum nucleopolyhedrovirus A (AgseNPV-A), Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B), Agrotis ipsilon nucleopolyhedrovirus (AgipNPV) and Agrotis segetum granulovirus (AgseGV). All these species are promising candidates for a baculovirus based plant protection product for the control of cutworms. The focus here was laid on AgseGV and AgseNPV-B, as these species have already been extensively investigated at the Julius Kühn Institute in Darmstadt.

The European isolate AgseGV-DA was identified as a slow-killing type I granulovirus due to the low dose mortality response and the tissue tropism of infection restricted solely to the fat body of infected larvae. The fully NGS sequenced genome of AgseGV-DA showed no sequence variation and was more than 99% identical to those of the Chinese isolates AgseGV-XJ and AgseGV-L1. The AgseGV-DA genome is 131,557 bp in length and contains 149 putative open reading frames, including 37 baculovirus core genes and the *per os* infectivity factor *ac*110 as well as one putative non-*hr* like origin of replication. AgseGV has a distinct *enhancin* gene, with a distant relation to the enhancins from the genus *Betabaculovirus*. Although all three isolates belong to the species *Agrotis segetum granulovirus*, AgseGV-DA was proposed as the type isolate due to the complete sequence and pathology description to the International Committee for Virus Taxonomy (ICTV). The permissivity of the insect cell culture AiE1611T was evaluated for AgseGV-DA and AgseNPV-B which is the prerequisite for *in vitro* experiments. While AgseGV and AgseNPV-B can co-infect susceptible larvae, it has been demonstrated that AiE1611T is only permissive for AgseNPV-B but not for AgseGV, as evidenced by two experiments with baculovirus derived DNA transfected cells and in infections with hemolymph containing budded virus from infected larvae. AgseNPV-B produced large numbers of occlusion bodies and the virus was selected in one round of plaque purification for additional experiments. Among the twelve genetical and morphological identical isolates, one isolate termed PP2 was used in particular to investigate the virus stability in ten consecutive passages in AiE1611T cells. The cell culture was further used to generate a recombinant AgseNPV-B clone, the bacmid bAgseNPV-B. AgseNPV-B PP2 showed endured a high genomic stability during the passages, while the dose-mortality responses in larvae were scattering over ten passages. A loss of virulence was observed right after the first round of passaging, then the activity of PP2 remained stable over ten passages. This was observed by the absence of few polyhedra phenotypes in phase contrast microscopy and by next-generation sequencing (NGS) of five selected passages. NGS sequencing revealed that defective particles were absent over ten passages. The few single nucleotide polymorphisms (SNPs) detected by

this approach were mostly within homologous repeat sequences, which make a correct mapping of the short sequencing reads difficult due to their repetitive nature. Thus, no conclusions were drawn from these SNPs. The proof of method for detecting large deletions was given by sequencing the bacmid bAgseNPV-B (chapter V). This bacmid was deleted by roughly 43 kb in the AgseNPV-B genome but still possesses the full bacterial recombinant DNA inserted by homologous recombination into *hr*6 of AgseNPV-B. The deletion additionally affected 42 *orfs* and two *hrs*. In consequence, AiE1611T cells could be transfected with DNA of bAgseNPV-B and showed cytopathological effects, however the infection was blocked at an early stage with missing DNA replication and no spreading of virus infection was observed. This is possibly correlated to the deletion of *lef*-1, *lef*-2 and *me*53 in bAgseNPV-B

The second baculovirus example presented here, is the use of Cydia pomonella granulovirus (CpGV). CpGV has a narrow host range and is highly virulent against the Lepidopteran pest species *Cydia pomonella* (codling moth; CM) and to a lower extent to a very few closely related Tortricids and has been developed and intensively used as a commercial biocontrol agent of CM in virtually all pome fruit production areas. CpGV comprises several isolates and extensive phylogenetic has brought evidence that all isolates can be divided into five genome groups or lineages. The basics for this classification are distinct trends on the genome level such as, insertions and deletions, and as focused here genome group specific SNPs. Data-sets generated of NGS of commercially available CpGV isolates were combined with their responses to codling moth types expressing different types (I - III) of CpGV resistance. In infection experiments, CpGV-0006 and CpGV-R5 were able to break type I resistance and to a lower extent also type III resistance, whereas CpGV-V15 overcame type I and the rarely occurring type II and type III resistance. Based on the distribution of SNPs in Illumina sequencing reads it was found that both CpGV-0006 and CpGV-R5 have highly similar genome group compositions, consisting of about two thirds of the CpGV genome group E and one third of genome group A. In contrast, CpGV-V15 is composed of about equal parts of CpGV genome group B and E. According to the identified genetic composition of these isolates, their efficacy towards different resistance types can be explained and predictions on the success of resistance management strategies in resistant CM populations can be made for future CpGV isolates based in this.

Zusammenfassung

Die Familie *Baculoviridae* umfasst doppelsträngige DNA Viren mit stabförmigen, umhüllten Virionen, welche Larvenstadien der Insektenordnungen Lepidoptera, Diptera und Hymenoptera infizieren. Aufgrund ihres oftmals sehr engen Wirtsspektrums und dem Fehlen von negativen Einflüssen auf die menschliche Gesundheit und die Umwelt, werden bereits einige Baculovirusarten in biologischen Pflanzenschutzmitteln verwendet.

Ein möglicher Einsatz ist die Bekämpfung von Erdraupen. Erdraupen stellen ernstzunehmende Bodenschadorganismen im Garten- und Ackerbau dar und umfassen neben anderen Arten auch die Gattung Agrotis. Sie leben nahe oder unterhalb der Oberfläche wo sie sich polyphag von den Setzlingen, Stämmen, Wurzeln und anderen Pflanzenteilen ernähren. In jüngster Vergangenheit wurden aus den Arten *Agrotis segetum* und *A. ipsilon* vier Baculovirusarten isoliert: Agrotis segetum nucleopolyhedrovirus A (AgseNPV-A), Agrotis segetum nucleopolyhedrovirus B (AgseNPV-B), Agrotis ipsilon nucleopolyhedrovirus (AgipNPV) und Agrotis segetum granulovirus (AgseGV). All diese Arten sind vielversprechende Kandidaten für biologische Insektizide gegen Erdraupen auf Basis von Baculoviren. Der Fokus dieser Arbeit lag jedoch auf AgseGV und AgseNPV-B aufgrund der langjährigen Forschungserfahrung mit diesen beiden Arten am Julius-Kühn-Institut in Darmstadt.

Aufgrund seiner geringen Mortalität in *A. segetum* Larven und des beobachteten Zelltropismus, der aufe eine Infektion des Fettkörpers und des Mitteldarmepitheliums beschränkt ist, wurde das europäische Isolat AgseGV-DA als langsam wirkendes Typ I Granulovirus identifiziert. Die vollständige mit NGS sequenzierte Genomsequenz beinhaltete lediglich einen homogenen Genotyp und war mehr als 99% identisch zu den beiden chinesischen Isolaten AgseGV-XJ und AgseGV-L1. Das AgseGV-DA Genom hat eine Größe von 131,557 bp und enthält 149 mögliche offene Leserahmen (*orfs*), unter anderem 37 Baculoviruskerngene und den Infektionsfaktor *ac*110, sowie einen möglichen non-*hr* Replikationsursprung. AgseGV besitzt für ein spezielles *enhancin* Gen, welches eine entfernte Verwandtschaft zu jenen *enhancin* Genen aufweist, die bisher in der Gattung *Betabaculovirus* entdeckt wurden. Wenngleich alle drei AgseGV Isolate derselben Spezies angehören, so wurde AgseGV-DA als das speziesbeschreibende Isolat vom Internationalen Komitee für Virustaxonomie (ICTV) wegen seiner vollständigen Beschreibung der Genomsequenz und Pathologie ausgewählt.

Die Permissivität der Insektenzellkultur AiE1611T gegenüber AgseGV-DA und AgseNPV-B, wurrde untersucht, da sie die Grundvoraussetzung jedwelcher *in vitro* Experimente darstellt. Während AgseGV und AgseNPV-B anfällige Larven koinfizieren könne, zeigte sich in Experimenten mit transfizierter DNA und infektiöser Hämolymphe von infizierten Larven, dass AiE1611T Zellen lediglich für AgseNPV-B permissiv sind, aber jedoch nicht für AgseGV. In AiE1611T Zellen wurden vielfache Okklusionskörper von AgseNPV-B produziert und das Virus wurde in einer Runde eines „Plaque Purification“ Assay für die weiteren Experimente selektiert. Von zwölf genotypisch und morphologisch identischen Isolaten, wurde eines, PP2 genannt, benutzt um die Virusstabilität in zehn aufeinanderfolgenden Passagen in AiE1611T Zellen untersucht. Weiterhin wurde die Zellkultur für die Untersuchung eines rekombinanten Virusklons, dem Bacmid bAgseNPV-B, benutzt. AgseNPV-B PP2 erwies sich als genomisch stabil, wohingegen die Infektionsrate in Larven über die Passagen schwankte. Während in der

ersten Passage ein Verlust der Virulenz zu verzeichnen war, blieb diese im weiteren Verlauf stabil. Gleichzeitig wurde über den Verlauf von zehn Passagen das Auftreten sogenannter „few polyhdera“ Mutanten weder mikroskopisch noch in der Genomanalyse beobachtet. Die wenigen Einzelnukleotidpolymorphismen (SNPs), welche mit diesem Verfahren ermittelt wurden, lagen im Bereich von „homologous repeat“ Sequenzen (*hrs*). Eine exakte Zuordnung der kurzen Sequenzierungs-Reads war deshalb aufgrund der repetitiven Natur dieser *hrs* nicht möglich, weswegen keine weiteren Schlüsse aus diesen SNPs gezogen wurden. Die Methodenvalidierung für die Detektion von Deletionen gelang mithilfe des Bacmids bAgseNPV-B. Dieses Bacmid war um schätzungsweise 43 kb im AgseNPV-B Genom deletiert, besaß aber weiterhin die bakterielle rekombinante DNA, welche durch homologe Rekombination im *hr*6 eingesetzt wurde. Die Deletion umfasste 42 orfs und zwei *hrs*. Als Konsequenz konnten AiE1611T zwar transfiziert werden, was an dem cytopathologischen Effekt ersichtlich war, die Infektion wurde jedoch zu einem frühen Zeitpunkt blockiert und verteilte sich nicht auf benachbarte Zellen. Dies korreliert wahrscheinlich mit dem Fehlen der AgseNPV-B Gene *lef*-1, *lef*-2 und *me*53.

Bei dem zweiten exemplarischen Baculovirus das hier behandelt wurde, handelt es sich um das Cydia pomonella granulovirus (CpGV). CpGV besitzt ein sehr enges Wirtsspektrum und wirkt sehr effizient gegen die lepidoptären Schädlinge *Cydia pomonella* (Apfelwickler, CM) und im geringeren Maße auch gegen nahe artverwandte Wickler. Aufgrund dessen wurde CpGV erfolgreich zu einem biologischen Insektizid entwickelt, welches nahezu im gesamten Obstbau Verwendung findet. CpGV umfasst einige Isolate und die intensive phylogenetische Forschung an CpGV konnte unter anderem beweisen, dass man die Isolate in fünf Genomgruppen/-typen unterteilen kann. Die Unterteilung gibt bestimmte Trends wieder, unter anderem Insertionen und Deletionen und abstammungslinienspezifische SNPs, welche hier im Fokus stehen. Datensätze aus NGS Sequenzierungen der kommerziellen Isolate CpGV-0006, CpGV-R5 und CpGV-V15 wurden bezüglich ihrer genomischen Zusammensetzung mit ihrer Wirkung in verchiedenen Larven Populationen mit den Resistenztypen I - III kombiniert. In Infektionen, waren die Isolate CpGV-0006 und CpGV-R5 resistenzbrechend für die Typ-I Resistenz und ebenso in geringerem Umfang für die Typ-III Resistenz. Aufgrund der SNP Verteilung der Illumina Sequenzen konnte gefolgert werden, dass beide Isolate genomisch beinahe identisch waren und zu zwei Dritteln aus Genomtyp E und zu einem Drittel aus Genomtyp A bestanden. Im Gegensatz dazu besteht CpGV-V15 zu etwa gleichen Teilen aus den Genomytpen B und E und bricht alle drei Resistenztypen. Die unterschiedliche Wirksamkeit lässt sich mit der Zusammensetzung aus verschiedenen Genomtypen erklären, was für zukünftige CpGV Isolate verwendet werden kann um ihren möglichen Erfolg in der Bekämpfung resistenter Apfelwicklerpopulationen vorherzusagen.

Appendix

Table 3 Complete genome sequences of baculovirus isolates deposited at GenBank; including 155 genome sequences from the genus *Alphabaculovirus*, 46 from *Betabaculovirus*, 2 from *Gammabaculovirus* and 1 from *Deltabaculovirus*. Sequences list downloaded in March 2018

Isolate/Strain Name	Genome size [bp]	Acc.-N°
Alphabaculovirus:		
Adoxophyes honmai NPV	113,220	NC_004690
Adoxophyes orana nucleopolyhedrovirus	111,724	NC_011423
Agrotis ipsilon multiple nucleopolyhedrovirus	155,122	NC_011345
Agrotis segetum nucleopolyhedrovirus	147,544	NC_007921
Antheraea pernyi nucleopolyhedrovirus isolate AnpeMNPV-L2	126,246	EF207986
Antheraea pernyi nucleopolyhedrovirus	126,629	DQ486030
Antheraea pernyi nucleopolyhedrovirus	126,629	NC_008035
Antheraea pernyi nucleopolyhedrovirus DNA, strain: Liaoning	126,593	LC194889
Anticarsia gemmatalis multiple nucleopolyhedrovirus isolate AgMNPV-2D	132,239	DQ813662
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-26	131,678	KR815455
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-27	131,172	KR815456
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-28	130,745	KR815457
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-29	130,506	KR815458
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-30	130,741	KR815459
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-31	132,126	KR815460
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-32	131,494	KR815461
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-33	131,059	KR815462
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-34	131,543	KR815463
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-35	132,176	KR815464
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-36	131,216	KR815465
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-37	131,855	KR815466
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-38	130,740	KR815467
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-39	130,698	KR815468
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-40	132,180	KR815469
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-42	130,949	KR815470
Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-43	132,077	KR815471

Table 3 Complete genome sequences of baculovirus isolates deposited at GenBank; including 155 genome sequences from the genus *Alphabaculovirus*, 46 from *Betabaculovirus*, 2 from *Gammabaculovirus* and 1 from *Deltabaculovirus*. Sequences list downloaded in March 2018

Anticarsia gemmatalis multicapsid nucleopolyhedrovirus isolate AgMNPV-37	131,855	NC_031761
Anticarsia gemmatalis nucleopolyhedrovirus	132,239	NC_008520
Apocheima cinerarium nucleopolyhedrovirus	123,876	NC_018504
Autographa californica multiple nucleopolyhedrovirus isolate WP10	133,926	KM609482
Autographa californica nucleopolyhedrovirus clone C6	133,894	L22858
Autographa californica nucleopolyhedrovirus strain E2	133,966	KM667940
Mutant Autographa californica nucleopolyhedrovirus isolate vAcRev-1	118,582	KU697902
Mutant Autographa californica nucleopolyhedrovirus isolate vAcRev-2	138,991	KU697903
Bombyx mandarina nucleopolyhedrovirus isolate S1	126,770	FJ882854
Bombyx mandarina nucleopolyhedrovirus S2	129,646	JQ071499
Bombyx mori nuclear polyhedrosis virus isolate T3	128,413	L33180
Bombyx mori NPV strain Cubic	127,465	JQ991009
Bombyx mori NPV strain Zhejiang	126,125	JQ991008
Bombyx mori NPV strain India	126,879	JQ991010
Bombyx mori NPV strain Guangxi	126,843	JQ991011
Bombyx mori nucleopolyhedrovirus isolate C1	127,901	KF306215
Bombyx mori nucleopolyhedrovirus isolate C2	126,406	KF306216
Bombyx mori nucleopolyhedrovirus isolate C6	125,437	KF306217
Bombyx mori nucleopolyhedrovirus strain Brazilian	126,861	KJ186100
Bombyx mori nucleopolyhedrovirus DNA, isolate: H4	127,459	LC150780
Buzura suppressaria nucleopolyhedrovirus isolate Hubei	120,420	NC_023442
Buzura suppressaria nucleopolyhedrovirus isolate Guangxi	121,268	KM986882
Catopsilia pomona nucleopolyhedrovirus isolate 416	128,058	NC_030240
Choristoneura fumiferana DEF MNPV	131,160	NC_005137
Choristoneura fumiferana multiple nucleopolyhedrovirus	129,593	NC_004778
Choristoneura murinana alphabaculovirus strain Darmstadt	124,688	NC_023177
Choristoneura occidentalis alphabaculovirus	128,446	KC961303
Choristoneura rosaceana alphabaculovirus	129,052	NC_021924
Chrysodeixis chalcites nucleopolyhedrovirus	149,622	NC_007151
Chrysodeixis chalcites nucleopolyhedrovirus genotype ChchSNPV-TF1-C	150,079	JX560539
Chrysodeixis chalcites nucleopolyhedrovirus genotype ChchSNPV-TF1-B	149,080	JX560540
Chrysodeixis chalcites nucleopolyhedrovirus genotype ChchSNPV-TF1-G	149,039	JX560541
Chrysodeixis chalcites nucleopolyhedrovirus genotype ChchSNPV-TF1-H	149,624	JX560542
Chrysodeixis chalcites SNPV TF1-A	149,684	JX535500
Pseudoplusia includens SNPV isolate IA	140,808	KU669289
Pseudoplusia includens SNPV isolate IB	138,869	KU669290
Pseudoplusia includens SNPV isolate IC	140,859	KU669291
Pseudoplusia includens SNPV isolate ID	140,787	KU669292
Pseudoplusia includens SNPV isolate IF	139,181	KU669293

Table 3 Complete genome sequences of baculovirus isolates deposited at GenBank; including 155 genome sequences from the genus *Alphabaculovirus*, 46 from *Betabaculovirus*, 2 from *Gammabaculovirus* and 1 from *Deltabaculovirus*. Sequences list downloaded in March 2018

Pseudoplusia includens SNPV isolate IG	139,116	KU669294
Clanis bilineata nucleopolyhedrosis virus	135,454	NC_008293
Condylorrhiza vestigialis MNPV	125,767	NC_026430
Agrotis segetum nucleopolyhedrovirus B isolate English	148,981	KM102981
Cyclophragma undans nucleopolyhedrovirus isolate Whiov	140,418	KT957089
Dasychira pudibunda nucleopolyhedrovirus isolate ML1	136,761	KP747440
Ecotropis obliqua NPV	131,204	NC_008586
Ectropis obliqua nucleopolyhedrovirus strain unioasis 1	130,145	KC960018
Epiphyas postvittana NPV	118,584	NC_00308
Euproctis pseudoconspersa nucleopolyhedrovirus	141,291	NC_012639
Helicoverpa armigera multiple nucleopolyhedrovirus	154,196	EU730893
Helicoverpa armigera NPV NNg1 DNA	132,425	AP010907
Helicoverpa armigera NPV strain Australia	130,992	JN584482
Helicoverpa armigera NPV	130,759	NC_003094
Helicoverpa armigera nucleopolyhedrovirus strain LB1	131,966	KJ701029
Helicoverpa armigera nucleopolyhedrovirus strain LB3	130,949	KJ701030
Helicoverpa armigera nucleopolyhedrovirus strain LB6	130,992	KJ701031
Helicoverpa armigera nucleopolyhedrovirus strain SP1A	132,481	KJ701032
Helicoverpa armigera nucleopolyhedrovirus strain SP1B	132,265	KJ701033
Helicoverpa armigera nucleopolyhedrovirus strain L1	136,760	KT013224
Heliocoverpa armigera nucleopolyhedrovirus G4	131,405	AF271059
Helicoverpa armigera SNPV isolate H25EA1	130,440	KJ922128
Helicoverpa SNPV AC53	130,442	KJ909666
Helicoverpa SNPV AC53 strain AC53C1	130,460	KU738896
Helicoverpa SNPV AC53 strain AC53C3	130,443	KU738897
Helicoverpa SNPV AC53 strain AC53C5	130,442	KU738898
Helicoverpa SNPV AC53 strain AC53C6	130,435	KU738899
Helicoverpa SNPV AC53 strain AC53T2	130,437	KU738901
Helicoverpa SNPV AC53 strain AC53C9	130,437	KU738900
Helicoverpa SNPV AC53 strain AC53T4.1	130,440	KU738902
Helicoverpa SNPV AC53 strain AC53T4.2	130,436	KU738903
Helicoverpa SNPV AC53 strain AC53T5	130,439	KU738904
Helicoverpa zea single nucleocapsid nucleopolyhedrovirus	130,869	AF334030
Helicoverpa zea single nucleopolyhedrovirus strain HS-18	130,890	KJ004000
Helicoverpa zea single nucleopolyhedrovirus isolate Br/South	129,694	KM596835
Hemileuca sp. nucleopolyhedrovirus	140,633	NC_021923
Hyphantria cunea nucleopolyhedrovirus	132,959	NC_00776
Lambdina fiscellaria nucleopolyhedrovirus isolate GR15	157,977	NC_026922
Leucania separata nuclear polyhedrosis virus strain AH1	168,041	NC_008348
Lonomia obliqua multiple nucleopolyhedrovirus isolate SP/2000	120,023	KP763670
Lymantria dispar MNPV	161,046	NC_001973
Lymantria dispar multiple nucleopolyhedrovirus strain 3029	161,712	KM386655
Lymantria dispar multiple nucleopolyhedrovirus isolate 2161	163,138	KF695050

Table 3 Complete genome sequences of baculovirus isolates deposited at GenBank; including 155 genome sequences from the genus *Alphabaculovirus*, 46 from *Betabaculovirus*, 2 from *Gammabaculovirus* and 1 from *Deltabaculovirus*. Sequences list downloaded in March 2018

Lymantria dispar multiple nucleopolyhedrovirus strain 3054	164,478	KT626570
Lymantria dispar multiple nucleopolyhedrovirus strain 3041	162,658	KT626571
Lymantria dispar multiple nucleopolyhedrovirus strain Ab-a624	161,321	KT626572
Lymantria dispar multiple nucleopolyhedrovirus isolate LdMNPV-45/0	161,006	KU862282
Lymantria dispar multiple nucleopolyhedrovirus isolate LdMNPV-27/2	164,158	KP027546
Lymantria dispar multiple nucleopolyhedrovirus isolate BNP	157,270	KU377538
Lymantria dispar multiple nucleopolyhedrovirus isolate RR01	159,729	KX618634
Lymantria dispar multiple nucleopolyhedrovirus isolate LdMNPV-27/0	161,727	KY249580
Lymantria xylina MNPV isolate LyxyMNPV-5	156,344	GQ202541
Lymantria xylina MNPV	156,344	NC_013953
Mamestra brassicae MNPV strain K1	152,710	NC_023681
Mamestra brassicae multiple nucleopolyhedrovirus isolate CTa	153,890	KJ871680
Mamestra brassicae multiple nucleopolyhedrovirus strain CHb1	154,451	JX138237
Mamestra configurata NPV-A	155,060	NC_003529
Mamestra configurata nucleopolyhedrovirus A 90/4	153,656	AF539999
Mamestra configurata NPV-B	158,482	NC_004117
Maruca vitrata MNPV	111,953	NC_008725
Mythimna unipuncta nucleopolyhedrovirus strain #7	148,482	MF375894
Operophtera brumata nucleopolyhedrovirus isolate OpbuNPV-MA	119,054	MF614691
Orgyia leucostigma NPV	156,179	NC_010276
Orgyia pseudotsugata MNPV	131,995	NC_001875
Peridroma alphabaculovirus isolate GR_167	151,109	NC_024625
Philosamia cynthia ricini nucleopolyhedrovirus virus	125,376	JX404026
Plutella xylostella multiple nucleopolyhedrovirus isolate CL3	134,417	DQ457003
Pseudoplusia includens SNPV IE	139,132	NC_026268
Rachiplusia ou multiple nucleopolyhedrovirus	131,526	AY145471
Spilosoma obliqua nucleopolyhedrosis virus isolate IIPR	136,141	KY550224
Spodoptera exigua nucleopolyhedrovirus complete genome	135,611	NC_002169
Spodoptera exigua multiple nucleopolyhedrovirus isolate VT-SeAl1	135,653	HG425343
Spodoptera exigua multiple nucleopolyhedrovirus isolate VT-SeAl2	134,972	HG425344
Spodoptera exigua multiple nucleopolyhedrovirus isolate VT-SeOx4	142,709	HG425345
Spodoptera exigua multiple nucleopolyhedrovirus isolate HT-SeG24	135,292	HG425346
Spodoptera exigua multiple nucleopolyhedrovirus isolate HT-SeG26	135,718	HG425348
Spodoptera exigua multiple nucleopolyhedrovirus isolate HT-SeSP2A	135,395	HG425349
Spodoptera exigua multiple nucleopolyhedrovirus isolate HT-SeG25	135,556	HG425347
Spodoptera frugiperda MNPV isolate 19	132,565	EU258200
Spodoptera frugiperda MNPV isolate 3AP2	131,331	EF035042
Spodoptera frugiperda MNPV isolate Nicaraguan	132,954	HM595733
Spodoptera frugiperda MNPV genotype SfMNPV-G defective	128,034	JF899325
Spodoptera frugiperda multiple nucleopolyhedrovirus isolate Colombian	134,239	KF891883
Spodoptera littoralis NPV isolate AN1956	137,998	JX454574
Spodoptera litura nucleopolyhedrovirus strain G2	139,342	AF32515
Spodoptera litura nucleopolyhedrovirus II	148,634	EU780426

Table 3 Complete genome sequences of baculovirus isolates deposited at GenBank; including 155 genome sequences from the genus *Alphabaculovirus*, 46 from *Betabaculovirus*, 2 from *Gammabaculovirus* and 1 from *Deltabaculovirus*. Sequences list downloaded in March 2018

Spodoptera litura nucleopolyhedrovirus II	148,634	NC_011616
Sucra jujuba nucleopolyhedrovirus isolate 473	135,952	NC_028636
Thysanoplusia orichalcea NPV isolate p2	132,978	NC_019945
Trichoplusia ni SNPV	134,394	NC_007383
Betabaculovirus:		
Adoxophyes orana granulovirus	99,657	NC_005038
Adoxophyes orana granulovirus strain Miyazaki	99,507	KM226332
Agrotis segetum granulovirus	131,680	NC_005839
Agrotis segetum granulovirus strain L1	131,442	KC994902
Agrotis segetum granulovirus strain DA	131,557	KR584663
Artogeia rapae granulovirus isolate Wuhan	108,592	NC_013797
Choristoneura fumiferana granulovirus	104,710	NC_008168
Clostera anachoreta granulovirus	101,487	NC_015398
Clostera anastomosis granulovirus B	107,439	KR091910
Clostera anastomosis granulovirus Henan	101,818	NC_022646
Cnaphalocrocis medinalis granulovirus	112,060	KP658210
Cnaphalocrocis medinalis granulovirus strain Enping	111,246	NC_029304
Cryptophlebia leucotreta granulovirus	110,907	NC_005068
Cydia pomonella granulovirus	123,500	NC_002816
Cydia pomonella granulovirus isolate CpGV-S	123,193	KM217573
Cydia pomonella granulovirus isolate CpGV-I07	120,816	KM217574
Cydia pomonella granulovirus isolate CpGV-M	123,529	KM217575
Cydia pomonella granulovirus isolate CpGV-I12	124,269	KM217576
Cydia pomonella granulovirus isolate CpGV-E2	123,858	KM217577
Diatraea saccharalis granulovirus	98,392	NC_028491
Epinotia aporema granulovirus	119,082	NC_018875
Erinnyis ello granulovirus	102,759	NC_025257
Erinnyis ello granulovirus isolate ErelGV-94	102,726	KX859079
Erinnyis ello granulovirus isolate ErelGV-98	102,685	KX859080
Erinnyis ello granulovirus isolate ErelGV-99	102,764	KX859081
Erinnyis ello granulovirus isolate ErelGV-00	102,745	KX859082
Erinnyis ello granulovirus isolate ErelGV-AC	102,741	KX859083
Erinnyis ello granulovirus isolate ErelGV-PA	102,616	KX859084
Helicoverpa armigera granulovirus	169,794	NC_01024
Mocis latipes granulovirus	134,272	NC_029996
Mythimna unipuncta granulovirus isolate MyunGV#8	144,673	NC_033780
Phthorimaea operculella granulovirus	119,217	NC_004062
Phthorimaea operculella granulovirus isolate SA	119,004	KU666536
Pieris rapae granulovirus strain E3	108,476	GU111736
Pieris rapae granulovirus	108,658	JX968491
Plodia interpunctella granulovirus isolate Cambridge	112,536	NC_032255
Plutella xylostella granulovirus	100,999	NC_0025
Plutella xylostella granulovirus isolate PxGV_C	100,980	KU529791

Table 3 Complete genome sequences of baculovirus isolates deposited at GenBank; including 155 genome sequences from the genus *Alphabaculovirus*, 46 from *Betabaculovirus*, 2 from *Gammabaculovirus* and 1 from *Deltabaculovirus*. Sequences list downloaded in March 2018

Plutella xylostella granulovirus isolate PxGV_K	101,004	KU529792
Plutella xylostella granulovirus isolate PxGV_M	100,986	KU529793
Plutella xylostella granulovirus isolate PxGV_T	100,978	KU529794
Plutella xylostella granulovirus isolate SA	100,941	KU666537
Pseudaletia unipuncta granulovirus strain Hawaiin	176,677	EU678671
Spodoptera frugiperda Betabaculovirus isolate VG008	140,913	NC_026511
Spodoptera litura granulovirus isolate SlGV-K1	124,121	DQ288858
Xestia c-nigrum granulovirus	178,733	NC_002331
Gammabaculovirus:		
Neodipiron abietes nucleopolyhedrovirus	84,264	NC_008252
Neodipiron lecontei nucleopolyhedrovirus	81,755	NC_005906
Neodipiron sertifer nucleopolyhedrovirus	86,462	NC_005905
Deltabaculovirus:		
Culex nigripalpus nucleopolyhedrovirus	108,252	NC_003084

Eigenanteil der Arbeit (Own work)

Die Experimente, Datenanalyse, sowie das Verfassen der Arbeit wurden ausschließlich von mir durchgeführt, mit Ausnahme von:

Kapitel II, veröffentlicht in: Gueli Alletti, G., Eigenbrod, M., Carstens, E. B., Kleespies, R. G., Jehle J. A. (2017), The genome sequence of Agrotis segetum granulovirus, isolate AgseGV-DA, reveals a new betabaculovirus species of a slow killing granulovirus. Journal of Invertebrate Pathology 146:48-68.:

- Die elektronenmikroskopischen Aufnahmen (histopathologische Untersuchungen) wurden gemeinsam von M. Eigenbrod und R. G. Kleespies erarbeitet.
- E. Carstens war in Teilen bei der Diskussion sowie Auswertung der genomischen Daten beteiligt. Er gab Hinweise zum möglichen Vorgehen bei der Auswertung.

Kapitel III, veröffentlicht in: Gueli Alletti, G., Carstens, E. B., Weihrauch, B., Jehle, J. A. (2018), Agrotis segetum nucleopolyhedrovirus but not Agrotis segetum granulovirus replicate in AiE1611T cell line of Agrotis ipsilon. Journal of Invertebrate Pathology 151:7-13:

- Konzeption sämtlicher Versuche wurde von G. Gueli Alletti in Zusammenarbeit mit E. B. Carstens und J. A. Jehle erarbeitet.
- E. B. Carstens war an der allgemeinen Einführung in molekulare Methoden der Insektenzellkulturtechnik beteiligt.
- B. Ruoff (ehem. Weihrauch) war als technische Assistenz in der Insektenzellkultur beteiligt.

Kapitel VI, veröffe Gueli Alletti, G., Sauer, A. J., Weihrauch, B., Fritsch, E., Undorf-Spahn, K., Wennmann, J. T., Jehle, J. A. (2017), Using next Generation sequencing to identify and quantify the genetic composition of resistance-breaking commercial isolates of Cydia pomonella granulovirus. Viruses 9 (9): e250.:

- Konzeption der Resistenztests mit neonaten Apfelwicklerlarven wurde von A. J. Sauer, J. A. Jehle, E. Fritsch und K. Undorf-Spahn erarbeitet.
- Die Resistenztests wurden von B. Ruoff (ehem. Weihrauch), A. J. Sauer, E. Fritsch und K. Undorf-Spahn durchgeführt.
- Die Auswertung der Resistenztest wurde überwiegend durch G. Gueli Alletti und A. J. Sauer in Zusammenarbeit mit J. A. Jehle, E. Fritsch und K. Undorf-Spahn erarbeitet.

Danksagung

An dieser Stelle möchte ich mich von ganzen Herzen bei Prof. Dr. Johannes Jehle bedanken. Lieber Johannes, Ihre Unterstützung fing lange vor meiner Zeit als Doktorand noch während meiner Diplomarbeit an und in den vielen Jahren ist mir die Freundschaft zu Ihnen ans Herz gewachsen. Ich kann nur bestätigen, was viele vor mir sagten, unser gemeinsamer Weg hat mich nicht nur wissenschaftlich, sondern auch persönlich geprägt. Nicht zuletzt bin ich dafür dankbar, dass sich auch unsere beruflichen Wege in Zukunft noch weiter kreuzen werden.

Mein Dank gilt auch meinem Zweitgutachter Herrn Prof. Dr. G. Thiel für das Lesen und Bewerten meiner Arbeit.

In meinen vielen Jahren am Julius-Kühn-Institut in Darmstadt, gab es eigentlich keinen Zeitpunkt, an dem ich nicht freudig, wenn auch etwas genervt vom ÖPNV, zur Arbeit kam. Nicht zuletzt liegt das auch an den vielen wunderbaren Menschen innerhalb und außerhalb meiner Arbeitsgruppe. Danke für die vielen, vielen heiteren, manchmal auch verrückten Tage liebe Kolleginnen und Kollegen. Ohne eine Person, wäre ich aber niemals auf das JKI, die Virologie und Johannes Jehle aufmerksam geworden…und diese Person ist mein Alterego, Freund und inzwischen Trauzeuge – du, Jörg! Danke! An den vielen extracurriculären Gelegenheiten (natürlich rein wissenschaftlicher Natur) und unzähligen Freudentränen, sei es beim Feierabendbierchen, auf Exkursionen, Tagungen und kleinen Streichen im Labor waren meine Freunde am JKI immer dabei, lieben Dank an Annette, Diana, Manuela, Katharina, Juliana, Karolin, Dietrich, Andreas, Fan und Christian. Auch wenn ich mich letztendlich gegen einen Werdegang in Forschung und Lehre entschieden habe, danke ich den vielen Studierenden, die ich in meiner Zeit betreuen durfte, Jacqueline, Mareike, Laurin und natürlich Marina. Danke für euer Vertrauen. Dann möchte ich noch Birgit Weihrauch danken, denn gerade in den letzten beiden praktischen Jahren meiner Arbeit (aber auch darüber hinaus) sind wir mit der Zellkultur zu einem super Zweiergespann zusammengewachsen. Für ihre tatkräftige Unterstützung in den Insektenzuchten möchte ich Doris El Mazouar und Eckard Gabrys danken. Während meiner Zeit am JKI durfte ich auch eine Menge internationaler Gäste begrüßen, unter ihnen einen bei dem noch ein Gegenbesuch im fernen Kanada aussteht. Dear Eric, the year with you has been quite productive but above all I have grown fond of you and our friendship. Regards to Anita.

Ich danke auch meinen vielen neuen Kollegen und meinem Vorgesetzten Dr. Rüdiger Hauschild für Ihre Geduld mit mir, wenn ich bei meiner Arbeit mal wieder von der Doktorarbeit abgelenkt wurde.

Danke an die starke Frau an meiner Seite. Inese, für deine unermüdliche Unterstützung und dafür, dass Du mir die Kraft gibst Gipfel zu erklimmen, kann ich Dir nicht genug danken.

Zuletzt möchte ich meiner Familie, meinen Geschwistern, allen voran meiner Mutter danken. Ihr habt mich immer begleitet, wart und seid immer für mich da.

Resume: Gianpiero Gueli Alletti

Gianpiero Gueli Alletti

Emil-Heckel-Straße 22, 68163 Mannheim, Germany
E-Mail: gianpiero.guelialletti@gmail.com
Phone: +49 163 2789120
Born: 5th September 1985 in Mannheim
Marital status: married

- https://www.linkedin.com/in/gianpiero-gueli-alletti/

Summary of Qualifications

Regulatory Affairs /Project-Management	Co-working in regulatory affairs of different existing microbial active ingredient and project managing of novel active ingredients according to Regulation (EU) 2017/2009 and Regulation (EU) 528/2012
Languages	Fluent in German, Italian and English; Advanced knowledge in French; Basic knowledge in Latvian
MS-Office	Expert in MS-Word, -Excel, PowerPoint; Advanced knowledge in MS-Access;
Biostatistics	Advanced knowledge in R and ToxRat; Basic knowledge in SAS
Bioinformatics/ Computional Science	Advanced knowledge in Geneious, Galaxy, Linux; Basic knowledge in Batch, Python, JavaScript, HTML, CSS and SQL

Curriculum vitae

2018 – present	Biopesticides Regulatory Affairs Expert at APIS Applied Insect Science GmbH in Mannheim Germany. Member of the founding team of APIS GmbH.
2017 – 2018	Regulatory Affairs Manager at GAB Consulting GmbH in Heidelberg, Germany
2012 – present	PhD Student in Technical Biology at the Technical University of Darmstadt
2012 – 2016	Scientific Co-worker at the Julius Kühn Institute (JKI), Federal Institute for Biological Control Darmstadt, Germany

2014 – 2016	Student and Post-Doc Affairs Committee Chair of the SIP
2013 – 2014	Organization of the 47th international annual meeting of the Society of Invertebrate Pathology, webmaster of congress webpage
2011 – 2012	Diploma thesis "Interaction of *Agrotis segetum granulovirus* and *Agrotis segetum nucleopolyhedrovirus B* in mixed infections" at the JKI
2008 – 2010	Research Assistant at the International Office Johannes Gutenberg University Mainz (Akademisches Auslandsamt)
2008 – 2011	Member and president of the student representation of the student dormitory "Mainz Hechtsheim"
2006 – 2012	Student in Biology at the Johannes Gutenberg University of Mainz
2005 – 2006	Volunteering year (FSJ) as male nurse at the "Theresienkrankenhaus und St. Hedwig-Klinik GmbH" in orthopaedics and surgery

Publications and Previous Presentations (First author)

2018 August	Gueli Alletti G, Wennmann JT, Carstens EB, Keilwagen J, Berner T, Jehle JA (2018) Agrotis segetum nucleopolyhedrovirus B shows high genome stability during serial *in vitro* passages in AiE1611T cells. (in prep)
2018 January	Gueli Alletti G, Carstens EB, Weihrauch B, Jehle JA (2018) Agrotis segetum nucleopolyhedrovirus but not Agrotis segetum granulovirus replicate in AiE1611T cell line of Agrotis ipsilon. J Invert. Pathol., 151:7-13
2017 August	Gueli Alletti G, Sauer AJ, Weihrauch B, Fritsch E, Undorf-Spahn K, Wennmann JT, Jehle JA (2017) Using Next Generation Sequencing to Identify and Quantify the Genetic Composition of Resistance-Breaking Commercial Isolates of Cydia pomonella Granulovirus. Viruses, 9:250.
2017 April	Gueli Alletti G, Eigenbrod M, Carsetns EB, Kleespies RG, Jehle JA (2017) The genome sequence of Agrotis segetum granulovirus, isolate AgseGV-DA, reveals a new *Betabaculovirus* species of a slow killing granulovirus. J. Invert. Pathol., 146:58-68
2016 July	Gueli Alletti G, Carstens EB, Jehle JA (2016) Genome stability of AgseNPV-B after serial *in vitro* passage. Oral Presentation at the International Congress on Invertebrate Pathology and Microbial Control and 49th Annual Meeting of the Society of Invertebrate Pathology Tours, FR

2015 August	Gueli Alletti G, Wennmann JT, Eigenbrod M, Carstens EB, Kleespies RG, Jehle JA (2015) Infection studies of an Agrotis segetum granulosis Virus isolate from Europe in cell cultures of AiE1611T. Oral Presentation at the International Congress on Invertebrate Pathology and Microbial Control and 48th Annual Meeting of the Society of Invertebrate Pathology Vancouver, CA
2015 June	Gueli Alletti G, Wennmann JT, Carstens EB, Jehle JA (2015) Analysis of the genome sequence of an Agrotis segetum granulovirus: New Challenges for Biological Control. Oral presentation at the 15th Meeting of the IOBC-WPRS Working Group "Microbial and Nematode Control of Invertebrate Pests" in Riga, LV.
2013 August	Gueli Alletti G, Wennmann JT, Jehle JA (2012) Developing of a TaqMan quantitative PCR assay for the quantification of *Agrotis* baculoviruses in single and mixed infections. Poster presentation at the 46th Annual Meeting of the Society for Invertebrate Pathology, Pittsburgh, USA

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

Mannheim, den

.....