RHODES UNIVERSITY
Where leaders learn

Genetic and biological characterisation of a novel South African *Cydia pomonella* granulovirus (CpGV-SA) isolate

A thesis submitted in fulfilment of the requirements for the degree of

MASTER OF SCIENCE IN MICROBIOLOGY

At

RHODES UNIVERSITY

By

BOITUMELO MADIKA MOTSOENENG

January 2016

Abstract

The codling moth, *Cydia pomonella* (L.) (Lepidoptera: Tortricidae), is the primary pest of pome fruit cultivated worldwide. The control of this insect pest has been dependent on the frequent use of broad-spectrum chemical pesticides, which has led to the development of resistance in pest populations and negative effects on human health and the environment. The *Betabaculovirus* of *C. pomonella* has successfully been applied as a biological control agent in integrated pest management (IPM) programmes for the suppression of pest populations worldwide. Previously, all *Cydia pomonella* granulovirus (CpGV) biopesticides were based on a Mexican isolate (CpGV-M) and although these products are highly efficient at controlling *C. pomonella,* resistance cases have been reported across Europe. The identification of novel CpGV isolates as additional or alternative control agents to manage resistance is therefore necessary. This study aimed to genetically and biologically characterise a novel South African *C. pomonella* granulovirus isolate and to test its virulence against neonate larvae.

Based on the morphology of the occlusion bodies observed using transmission electron microscopy, granuloviruses were recovered from diseased and dead larvae collected from an orchard in South Africa where no virus applications had been made. DNA was extracted and the identification of the isolated granulovirus was achieved through the PCR amplification and sequencing of the *lef*-8, *lef*-9, *granulin* and *egt* genes. Submission of the gene sequences to BLAST revealed high percentage identities to sequences from various CpGV isolates, resulting in the naming of the isolate in this study as the South African *Cydia pomonella* granulovirus (CpGV-SA) isolate. Phylogenetic analysis based on the single nucleotide polymorphisms (SNPs) detected in the *lef*-8, *lef*-9 and *granulin* nucleotide sequences grouped the South African isolate with CpGV-E2 (genome type B) and CpGV-S (genome type E).

The CpGV-SA isolate was further genetically characterised by restriction endonuclease analysis and complete sequencing of the genomic DNA. Differences were observed for the *Bam*HI, *Eco*RI, *Pst*I and *Xho*I profiles of CpGV-SA in comparison to the respective profiles generated for CpGV-M extracted from a biopesticide, Carpovirusine® (Arysta Lifescience, France). Several genetic variations between the complete genome sequence of CpGV-SA and the reference isolate, CpGV-M1, as well as a recent genome submission of CpGV-M, both representing genome type A were observed. The complete genome analysis

confirmed that CpGV-SA is genetically different from the Mexican CpGV isolate, used in the development of most biopesticides.

In silico restriction profiles of the genome sequence obtained for CpGV-SA and genome sequences of genetically different CpGV isolates originating from Mexico (M1 and M), England (E2), Canada (S) and Iran (I12 and I07), available on the NCBI's GenBank database confirmed that CpGV-SA is of mixed genotypes. Furthermore, the South African isolate shared the single common difference found in the *pe38* gene of resistance overcoming isolates, which was the absence of an internal 24 nucleotide repeat present in CpGV-M1. In addition to the common difference, SNPs detected in the *pe38* gene grouped the isolate with the CpGV-S isolate, suggesting that the CpGV-SA isolate is predominantly of genome type E.

To determine the biological activity of CpGV-SA against neonate *C. pomonella* larvae, surface bioassays were conducted alongside CpGV-M (Carpovirusine®) bioassays. The LC_{50} and LC_{90} values for the South African isolate were 1.6×10^3 and 1.2×10^5 OBs/ml respectively. The LT_{50} was determined to be 135 hours. These values were similar to the values obtained for CpGV-M (Carpovirusine®). The results in this study suggest that a novel South African CpGV isolate of mixed genotypes, potentially able to overcome resistance in *C. pomonella,* with biological activity similar to CpGV-M (Carpovirusine®) and important for the control of *C. pomonella* was recovered. The CpGV-SA isolate could therefore potentially be developed into a biopesticide for use in resistance management strategies against *C. pomonella* populations in South Africa.

Table of contents

List of figures

Chapter 1 **Page**

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of tables

Chapter 5

List of abbreviations

General

Bt	–	*Bacillus thuringiensis*
BLAST	–	Basic Local Alignment Search Tool
BV	–	budded virus
CTAB	–	Cetyltrimethyl ammonium bromide
DNA	–	deoxyribonucleic acid
D1	–	Dilution 1
D2	–	Dilution 2
D3	–	Dilution 3
D4	–	Dilution 4
D5	–	Dilution 5
D6	–	Dilution 6
ddH_2O	–	Double distilled water
egt	–	Ecdysteroid UDP-glucoslytransferase
GV	–	Granulovirus
IPM	–	Integrated pest management programme
LC_{50}	–	Lethal concentration (50%)
LC_{90}	–	Lethal concentration (90%)
LT_{50}	–	Median lethal time
lef-8	–	Late expression factor-8
lef-9	–	Late expression factor-8
Ltd	–	Limited
MNPV	–	Multiple nucleopolyhedrovirus
NCBI	–	National Center for Biotechnology Information
NGS	–	Next generation sequencing
NPV	–	Nucleopolyhedrovirus
OB	–	Occlusion body
ODV	–	Occlusion derived viruses
ORF	–	Open reading frame
PCR	–	Polymerase chain reaction
qPCR	–	Quantitative real-time polymerase chain reaction
REN	–	Restriction endonuclease
RNA	–	Ribonucleic acid
SEM	–	Scanning electron microscopy
SNPV	–	Single nucleopolyhedrovirus
SNP	–	Single nucleotide polymorphism
SDS	–	Sodium dodecyl sulphate
SE	–	Standard error
SIR	–	Sterile insect release
TEM	–	Transmission electron microscopy
UV	–	Ultraviolet

UK	–	United Kingdom
USA	–	United States of America

Units and symbols

bp	–	Base pairs
χ^2	–	Chi-square
°C	–	Degrees Celcius
df	–	Degrees of freedom
g	–	Gram
kb	–	Kilobase pairs
µl	–	Microlitres
µM	–	Micromolar
mM	–	Millimolar
mg	–	Milligrams
mm	–	Millimetre
ml	–	Millilitre
min	–	Minute
M	–	Molar
ng	–	Nanograms
nm	–	Nanometre
nt	–	nucleotide
%	–	Percentage
p	–	Test level
xg	–	Times gravity
V	–	Volts
v/v	–	Volume per volume
w/v	–	Weight per volume

Viruses

AcMNPV	–	*Autographa californica* MNPV
CpGV	–	*Cydia pomonella* GV
CrleGV	–	*Cryptophlebia leucotreta* GV

Research outputs

Conferences:

Motsoeneng B.M., Knox C.M., Hill M.P., Moore S.D., (2015). The genetic characterisation of a novel South African *Cydia pomonella* granulovirus. Oral presentation ESSA 2015 conference held at Rhodes University, Grahamstown, Eastern Cape Province, South Africa, 12-15 July.

Acknowledgements

I would like to thank the following people for their support and assistance in the completion of this work:

- My supervisor Prof. Caroline Knox, I greatly appreciate all the guidance, advice, support and opportunities that you have given me which have enabled me to learn more and grow professionally throughout the years.
- My co-supervisors Prof. Martin Hill and Dr. Sean Moore for the criticism and encouragement you have both given and the enthusiasm you have both shown at all the meetings.
- Dr Craig Chambers from River Bioscience (Pty) Ltd. for supplying the virus samples and materials needed for the bioassays.
- Dr Daleen Stenekamp from Hortgro Science for supplying egg sheets for the completion of the bioassays.
- Tamryn Marsberg for your willingness to help me with the preparation of the bioassays and for your words of encouragement.
- Michael Jukes and Fatima Abdulkadir for advice and giving me all the necessary help in the lab.
- Marvin Randal and Shirley Pinchuck for their assistance with electron microscopy.
- Mellissa Peyper and Candice Coombes for their assistance given in the entomology lab.
- My lab colleagues Nicole Upfold, Oikwathaile Onosi, Kudzai Mtambanengwe and the honours students for all the lively conversations and making the laboratory a pleasant work space.
- Sibusiso Mbandazayo for your presence throughout and words of motivation.
- National Research Foundation for funding.
- My siblings and entire family for all their support.
- To my parents words cannot not express my appreciation for all the love, patience, encouragement, advice and support that you have both given me. A special thanks to you both for trusting and believing in me, Mr MD and Mrs ND Motsoeneng.

God is great!

Chapter 1

Literature review: *Cydia pomonella* and the use of baculoviruses to control insect pests

1.1 Introduction

The pome fruit industry in South Africa made a turnover of R7.2 billion from the cultivation of 22925 ha of apples and 12211 ha of pears in 2014. These fruits are amongst the most important crops consumed locally and exported, with the largest export markets being the European market and the rest of the African continent (Hortgro, 2014). A constant major threat to the pome fruit industry worldwide is an insect pest, *Cydia pomonella* (L.), (Lepidoptera: Tortricidae), commonly known as the codling moth. Since it was first introduced in South Africa in 1885 and the first records of fruit infestation were reported in 1898 it has become a key pest in the country (Lounsbury, 1898; Blomefield & Giliomee, 2014).

In the past 50 years the need for environmentally sustainable pest management strategies has increased. The use of broad-spectrum chemical pesticides has led to the development of pest resistance, depletion of beneficial organisms, as well as consumer health and food safety concerns (Charleston *et al.*, 2003; Reyes *et al.*, 2007; Pajač *et al.*, 2011). Biological control offers a cost efficient, environmentally safe and sustainable solution to control insect pests. Biological control involves the use of natural enemies such as predators, parasitoids or entomopathogens, which consist of bacteria, nematodes, fungi and viruses, to suppress pest populations (Lazarovits *et al.*, 2007; Hoddle & Van Driesche, 2009).

Baculoviruses are a group of DNA viruses that infect insects and have successfully been applied as biological control agents for the suppression of insect pest populations worldwide (Szewczyk *et al.*, 2006). *Cydia pomonella* granulovirus (*Betabaculovirus*) specifically infects *C. pomonella* larvae and most of the commercially available biopesticides developed using this virus, are based on a strain isolated in Mexico (CpGV-M) (Tanada, 1964). Although, CpGV-M has been used to control *C. pomonella,* several cases of host resistance to the virus in orchards in Germany, France and across Europe have been reported, which has led to the development and increased demand for resistance management strategies (Fritsch *et al.*, 2005; Eberle & Jehle, 2006; Sauphanor *et al.,* 2006;

Asser-Kaiser *et al.*, 2007; Schmitt *et al.*, 2013). Studies suggest that different virus isolates from the same species of virus could be used as alternatives to control pests that are resistant to a specific isolate (Eberle *et al.*, 2008; Eberle *et al.,* 2009; Berling *et al.*, 2009a; Berling *et al.*, 2009b). In South Africa, resistance cases to CpGV-M based biopesticides have not been reported. However, *C. pomonella* remains a significant economic pest in South Africa and it is therefore important to bioprospect for novel CpGV isolates as additional or alternative control agents to be included in resistance management programmes.

1.2 The pest: *Cydia pomonella*

1.2.1 Origin, taxonomy and distribution

Cydia pomonella acquired its common name, codling moth, when the larvae were discovered in green, elongated, English cooking apples referred to as codling apples. The insect originates from Eurasia and was accidently introduced into other countries through human migration, seedling dispersion and the transportation and cultivation of apples and pears (Franck *et al.*, 2007; Welter, 2009). More specifically, the distribution of *C. pomonella* is linked to the distribution of apples in Western Asia, Eastern Europe and South Western Siberia, pears in Caucasus and walnuts in Caucasus and Turkestan, as it is closely associated with these fruits (Barnes, 1991).

The taxonomic history of *C. pomonella* is complex (Wearing *et al.*, 2001). The moth was first described from Europe by Linnaeus and is classified under the order Lepidoptera and family Tortricidae and was given the species name *C. pomonella* (Linnaeus, 1758). Currently, *Cydia* and related genera are included in the tribe Grapholitini of the sub-family Olethreutinae (Brown, 1979; Pajač *et al.*, 2011).

The insect pest currently occurs in the temperate regions of all major continents, and is considered to be one of the most successful pests due to its global distribution (Figure 1.1) (Barnes, 1991; Thaler *et al.*, 2008; Pajač *et al.*, 2011). *Cydia pomonella* is present in both the eastern and western states of the United States of America, Canada, Mexico, and in the pome fruit production areas of most South American countries. This insect pest has spread across Europe but also occurs in South Africa, Australia and New Zealand (Wearing *et al*., 2001; Franck *et al.,* 2007).

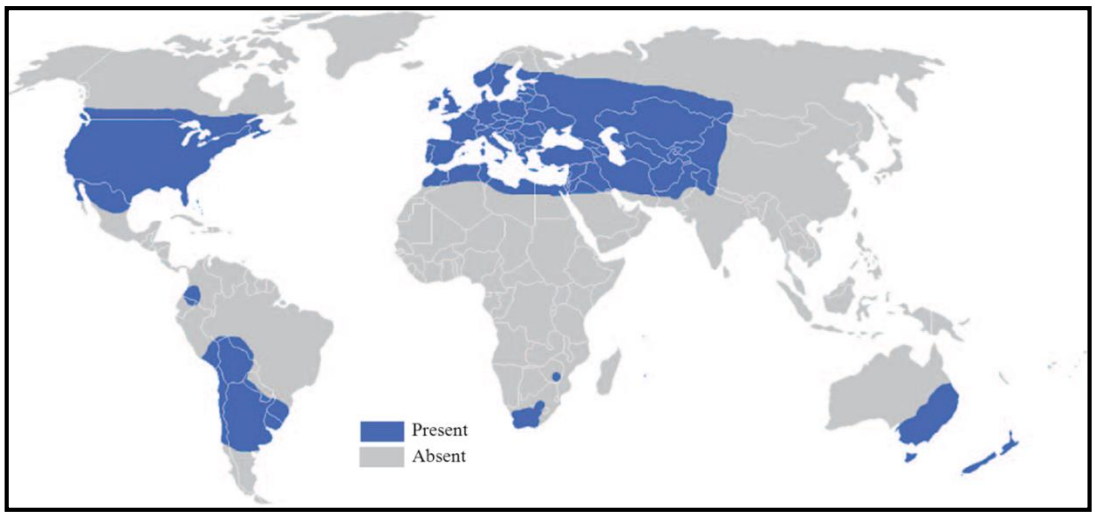


Figure 1.1: The global distribution of *Cydia pomonella* (Pajač *et al.*, 2011)

1.2.2 Life cycle of *Cydia pomonella*

The number of generations of *C. pomonella* that may occur in a year is dependent mainly on the temperature of the region (Riedl, 1983; Aghdam *et al.*, 2009). South Africa's hot climate conditions has resulted in it having one of the highest infestation potentials, 3-4 generations of *C. pomonella* occur per growing season as opposed to 1-2 generations in regions with colder climate conditions such as Canada, Northwestern USA and parts of Russia and the UK (Setyobudi, 1989; Pringle *et al.*, 2003; Blomefield & Giliomee, 2012). During the day or at dusk, when the weather conditions are favourable, the moths become active, mate, and the females may lay more than 100 eggs on the surfaces of the fruit or on the leaves (Agnello & Kain, 1996; Blomefield and Giliomee, 2012). The eggs are either laid singly or in groups of two or three. When the eggs are laid, each is disc-shaped, clear to creamy and approximately 1 mm in diameter (Figure 1.2A). A distinct red ring occurs around the egg, during its development (Figure 1.2B) (Pajač *et al.*, 2011). Before hatching the dark head capsule of the neonate larva becomes visible (Figure 1.2C). Eggs take approximately 6-20 days to hatch depending on the temperature in the orchards (Agnello & Kain, 1996).

The larvae go through five instars which are considered to be the destructive stages where the fruit is damaged (Wearing, 1979). After the eggs have hatched, neonate larvae feed on foliage while searching for fruit. Once the fruit is found the larvae will bore into the fruit through the calyx or on the side of large fruit, and the 2nd to 5th instars will develop in the

fruit over 3 to 4 weeks (Wearing, 1979; Agnello & Kain, 1996; Pajač *et al.*, 2011). The neonate larvae develop from 2 mm in size to 20 mm and are white with black heads (Figure 1.2C). When fully grown, the larvae are pinkish-white with brown heads (Figure 1.2D). *Cydia pomonella* larvae have no anal combs; this distinguishes the larvae from other apple insect pests such as *Grapholita molesta*, commonly known as the oriental fruit moth (Agnello & Kain, 1996; Pajač *et al.*, 2011; Wearing *et al.,* 2001).

Once the larvae have reached the 5th instar, they exit the fruit, find dry hidden bark, and begin to pupate. Cocoons are between 8-13 mm long and pupal development occurs over 7-30 days depending on temperature (Figure 1.2E), with relatively small adult moths emerging from the cocoons (Agnello & Kain, 1996; Pajač *et al.*, 2011). Adult moths are 10-12 mm long and have a forewing span of approximately 14-22 mm (Agnello & Kain, 1996). The moths are grey-brown with forewings that are scaled with fine white lines and brown-gold markings on the tip, and hind wings that are copper-brown and cannot be seen when the moth is at rest (Figure 1.2F) (Pajač *et al.*, 2011).

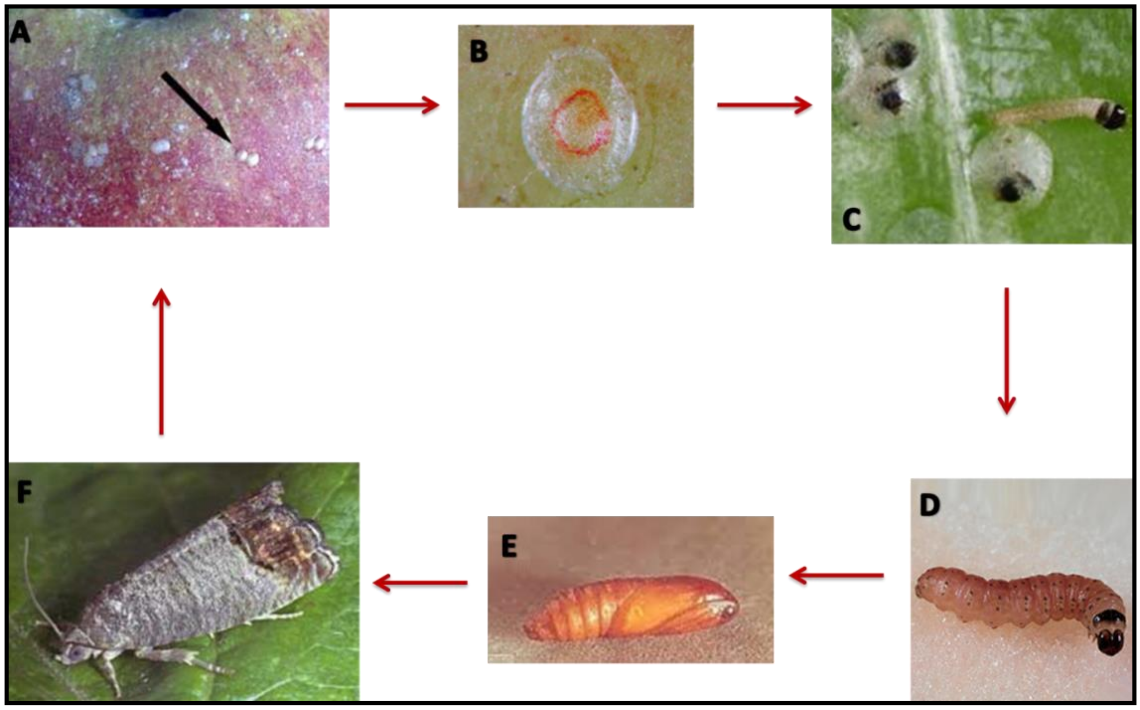


Figure 1.2: *Cydia pomonella* life cycle. A) Disc-shaped eggs are laid on the fruit surface; B) Red ring visible on the egg, indicates embryonic development; C) Head capsule of neonate larvae visible before eggs hatch; when eggs hatch neonate larvae burrow into fruit and exit fruit as a fully developed larva; D) 5th instar larva finds pupation site and produces silk; E) Cocoon is formed; F) Adult moth emerges from cocoon, mates and the cycle is repeated.

1.2.3 Economic importance

1.2.3.1 Pest status

The pest status of an insect can be determined through the use of monitoring systems and economic thresholds. Apples and pears are high value crops that need to be protected from insect pests in order to export undamaged and non-decaying fruit to the European Union and the United States, which are South Africa's main foreign markets. The extent to which the pest is controlled per season is dependent on the pest status (Pringle, 2006). An economic threshold is a point at which the population density of an insect pest has the potential to increase and cause economic injury, and is used to refer to fruit damage that justifies the need for control measures. Economic thresholds are used for pests that occur sporadically or perennially (Stern, 1973).

Cydia pomonella is a chronic pest that requires regular control measures and therefore the concept of an economic threshold cannot be applied to it (Pringle, 2006). However, with the development of pest resistance to insecticides, the infestation potential and pest status of *C. pomonella* continues to increase and it therefore remains a key economic pest worldwide (Barnes & Bloemfield, 1997; Blomefield & Giliomee, 2014). In 1996, *C. pomonella* was ranked to have the third highest status of lepidopteran pests of cultivated plants in South Africa; *Helicoverpa armigera* (Hübner, 1808) and *Agrotis segetum* (Denis and Schiffermüller, 1775) were ranked first and second respectively (Bell & McGeoch, 1996).

1.2.3.2 Host range and extent of damage on fruit

Apples and pears are not the only crops damaged by *C. pomonella*. Damage by *C. pomonella* larvae to walnuts, quinces and some stone fruit such as plums, apricots, peaches and nectarines has been recorded. Some larvae feed on the surface of the fruit, which causes scarring and blemishes (Agnello & Kain, 1996). However, most larvae invade one fruit causing deeper injury to it by feeding under the skin, burrowing through to the core of the fruit and feeding on the flesh through to the seeds (Figure 1.3A). Indirect contamination is caused by the larval frass, which fills the entry hole caused by the pest (Figure 1.3B) (Welter, 2009). The larva will then exit the fruit and the fruit then ripens faster and drops early in the growing season. The damage caused by *C. pomonella* results in unmarketable fruit produce but depending on the severity of the damage, the fruit may be used for the production of juice (Pajač *et al.*, 2011). The damage caused by the second

generation is more severe than the first as the eggs are laid on fruit that are ripening (Capinera, 2008). Infestations do not only occur on the fruit hanging in the trees but can also occur when ripe fruit have been sent to pack houses and stores. *Cydia pomonella* is persistent and, even when all the fruit are destroyed and none are developing, the larvae feed on the leaves of the plant (Hansen *et al.*, 2006).

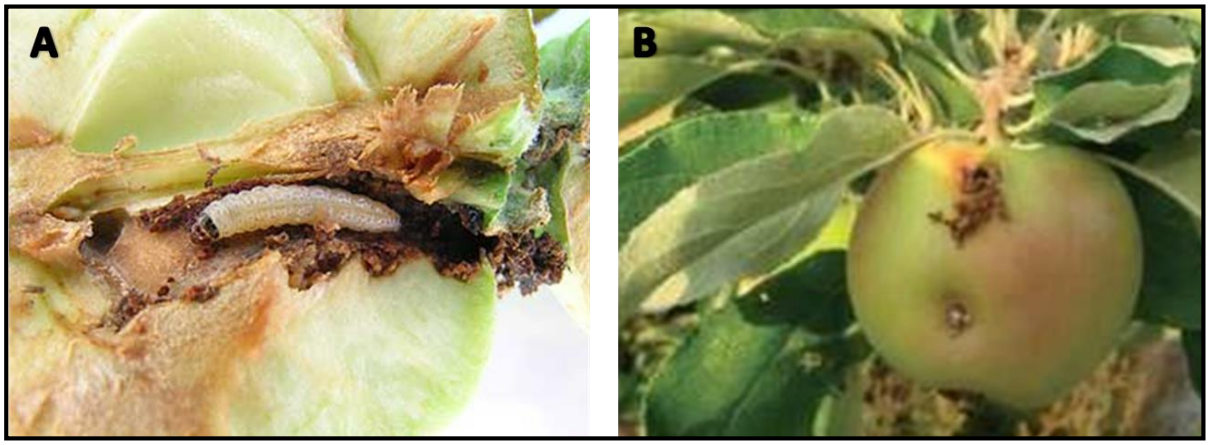


Figure 1.3: Extent of damage to fruit caused by the insect pest. A: *Cydia pomonella* larva feeding on the flesh of an apple; B: Indirect contamination (Welter, 2009).

1.2.4 Controlling *Cydia pomonella*

The deciduous fruit industry in South Africa is one of the leaders in the application of integrated pest management (IPM) due to international markets demanding fresh produce that has minimal insecticide residue and the realisation that for sustainable fruit production, environmentally safe control methods should be used (Charleston *et al.,* 2003). IPM is the application of a combination of various environmentally safe control methods with chemical control, based on the state of the environment, to reduce insect pest survival and the risk of resistance occurring (Flint & Bosch, 2012; Rodriguez *et al.*, 2012). The emphasis in IPM is on developing and using biological control methods (Pajač *et al.*, 2011). IPM strategies aim to protect crops at low cost and low risks to humans and the environment (Rodriguez *et al.*, 2012). Combined control methods have been shown to promote the population growth of beneficial organisms in fruit orchards (Charleston *et al.,* 2003). South Africa uses a wide range of IPM strategies that include chemical insecticides in combination with alternative methods such as pheromone traps, cultural practices, sterile insect release and the use of natural enemies and microbes, such as viruses to control *C. pomonella* (Swezey, 2000; Charleston *et al.,* 2003).

These methods are described in the following sections:

1.2.4.1 Chemical control

Strategies used to control *C. pomonella* have always been based on increasing the mortality of the insect pest at certain stages, increasing the efficiency of natural enemies or disrupting the development or the behaviour of the insect (Audemard, 1991). The primary method of control has always been to control the pest through the use of broad-spectrum insecticides such as carbamates, chlorinated hydrocarbons and pyrethroids and particularly organophosphate (OP) insecticides (Reyes *et al.,* 2007). Most insecticides are based on OP insecticides such as acephate and azinophs-methyl as chemical ingredients (Quinn *et al.*, 2011). In the past, these insecticides were sprayed frequently without monitoring the populations of *C. pomonella* present in the orchards. However, the frequent use of insecticides has had a negative impact on the environment and has led to insects developing resistance and cross-resistance to multiple insecticide classes and the depletion of beneficial organisms (Knight, 1994; Pajač *et al.*, 2011; Swezey, 2000). Some insecticides have been banned while others may still be used and are still the primary control method for *C. pomonella*. In South Africa, over 20 insecticides are currently registered for use in controlling *C. pomonella* (Quinn *et al.,* 2011). These are categorised according to environmental impact to allow farmers to choose safer insecticides (Charleston *et al.,* 2003). Moreover, some of the *C. pomonella* insecticides used also kill other insect pests such as the African bollworm, weevil and leaf rollers (Quinn *et al.,* 2011). Similar insecticides are also registered in other parts of the world such as the United States and Israel for the control of *C. pomonella* (Dunley & Welter, 2000; Reuveny & Cohen, 2004).

1.2.4.2 Monitoring and Cultural Practices

The identification of pests and the detection of crop damage and natural enemies are the initial steps of monitoring programmes, which are essential in pest management. Monitoring provides information on the effectiveness of previous applications of control methods and indicates any need for further control. There are several methods used to monitor pest activity in orchards; most involve the use of traps (Swezey, 2000).

Insects find mates and communicate through the use of chemical sex attractants, pheromones (Myers *et al.*, 2000). Prior to the 1970s, synthetic pheromones were used as baits to trap and monitor the growth or reduction in the *C. pomonella* populations, and

monitoring systems allowed fruit growers to make informed decisions. The pheromone baited traps generally improved the use of the insecticides as the timing of the sprays was relative to the infestation potential and the use of the insecticides was reduced (Madsen *et al.*, 1974; Barnes, 1990). Thus, traps are currently used where possible.

Cultural practices are performed to keep pest infestations to a minimum (Road, 1991; Riedl *et al.*, 1998). Orchard sanitation is maintained by discarding damaged fruit from trees and from the ground, and tree banding to capture overwintering larvae (Judd *et al.*, 1997). Potential sites for pupation to occur such as dry loose bark and leaf litter are also removed from the orchards and wooden fruit bins used to transport the fruits need to be disinfected (de Waal *et al.*, 2010). Although this is a simple method of control, it is as important as the other control methods and enhances the efficiency of alternative control methods but it cannot be used alone (Judd *et al*., 1997). Pheromone traps, the application of biopesticides and sterile insect release are more successfully applied in orchards with low pest infestation (Judd *et al*., 1997; Witzgall *et al.*, 2008).

1.2.4.3 Mating disruption

Pheromones are also used for a control method referred to as mating disruption (Riedl *et al*., 1998; Pringle *et al*., 2003; Witzgall *et al*., 2008). Mating disruption has been successfully used in the management of *C. pomonella* worldwide for two decades (Witzgall *et al.*, 2008). The male moths are attracted to the pheromones and this reduces the chances of mating occurring with the female moths, thus reducing future moth infestations (Pringle *et al.,* 2003). The main disadvantage of this method is that the female moths are not drawn to the pheromones and therefore the traps are only effective where there are low population densities (Yan *et al.*, 1999). In South Africa, decisions regarding spraying in orchards where pheromones were used was proven less reliable than in orchards where pheromones were not used (Pringle *et al.,* 2003). This is due to limiting factors such as the geographical structure of the production area, wind and open spaces between the orchards, which are all concerns in the Western Cape Province of South Africa where apples and pears are cultivated (Carde & Minks, 1995). Local studies have shown that the efficiency of the pheromones is also dependent on weather conditions and the density of the insect population (Lacey *et al.,* 2007; Bloemfield and Giliomee, 2014).

1.2.4.4 Sterile Insect Release (SIR)

Sterile insect release involves rearing, sterilising and releasing a large number of males into the orchards to mate with wild females, which results in a reduction of the population (Addison, 2005). The first *C. pomonella* sterile insect release programme was initiated in the Okanagan Valley of British Columbia (Myers *et al.*, 1998; Winston, 1999). It is currently and successfully used in Western Canada, the United States and several other countries (Calkins *et al.,* 1998; Bloem *et al.*, 2000). South Africa's *C. pomonella* SIR pilot project was initiated in 2003 in the Western Cape but was recently closed in 2014 due to economic considerations (Addison, 2005; Barnes *et al.*, 2015). The main advantages to this technique are that there is no environmental contamination and non-target organisms are not affected (Myers *et al.,* 2000). The main disadvantages are that SIR programmes require good maintenance and management and are expensive initially (Addison, 2005).

1.2.4.5 Biological control

Biological control involves the use of natural enemies such as predators, parasitoids and entomopathogens to suppress pests and maintain populations of beneficial organisms (Hoddle & Van Driesche, 2009). Predators and parasitoids range from spiders and mites to *Trichogrammatoidae* species (Lacey *et al.*, 2003; Lacey & Unruh, 2005). In South Africa, the *Trichogrammatoidea* species have been shown to attack the eggs of *C. pomonella* but have not been well-studied (Lacey & Unruh, 2005; Wahner, 2008). Various entomopathogens have been recovered from *C. pomonella*, but the granulovirus is the most commonly used biological agent to develop microbial insecticides to control *C. pomonella* (Lacey & Unruh, 2005). *Cydia pomonella* larvae are susceptible to *Bacillus thuringiensis* (Bt) biopesticides, which are environmentally safe (Andermatt *et al.*, 1988; Falcon & Huber, 1991; Cross *et al.*, 1999). However, Bt is more efficient in controlling other insects such as leaf rollers. *Beauveria bassiana* is a fungal species that has been extensively studied as a potential biological control agent for *C. pomonella* (Falcon & Huber, 1991; Cross *et al.,* 1999, Garcia-Gutierrez *et al.,* 2004). However, Garcia-Gutierrez *et al.* (2004) reported that fruit damage was significantly higher for the *B. bassiana* treatments than for an azinphos-methyl treatment. Therefore this fungal species was not considered suitable for *C. pomonella* control. Biological control agents have an important role in integrated pest management (IPM) programmes for the control of insect pests. A highly virulent baculovirus

has been successfully used to control *C. pomonella* and will be discussed in the sections below.

1.3 Baculoviruses

1.3.1 Classification and structure

The discovery of baculoviruses originated in the discovery of diseases of silkworms, specifically 'wilting disease' discovered in the 16th century (Rohrmann, 2013). In the 19th century, polyhedral crystals were detected from diseased larvae and the presence of viruses in these polyhedral structures was later discovered in the 20th century (Benz, 1986). In the 1940s rod-shaped virions in polyhedral structures, now referred to as occlusion bodies (OBs), were observed by electron microscopy (Bergold, 1948). These discoveries led to the investigation of baculoviruses as natural control agents of insects (Miller, 2013).

Baculoviruses are a group of arthropod viruses with rod-shaped nucleocapsids of ± 300 nm (Miller, 1996). This group of highly specific insect viruses has been re-classified several times. Baculoviruses were initially classified into two virion phenotypes, namely occlusion-derived virions (ODV) and budded virions (BV) (Figure 1.4). ODV were defined as virions enclosed in OBs, in a crystalline protein matrix, that infected insects through epithelium cells of the midgut. Virions that budded through the plasma membrane of infected cells and therefore usually contained a single nucleocapsid within an envelope were referred to as BV. The two virion morphologies are genetically identical and indicated that BV was responsible for cell-to-cell transmission while ODV was responsible for insect-to-insect transmission of baculovirus infection (Rohrmann, 2013).

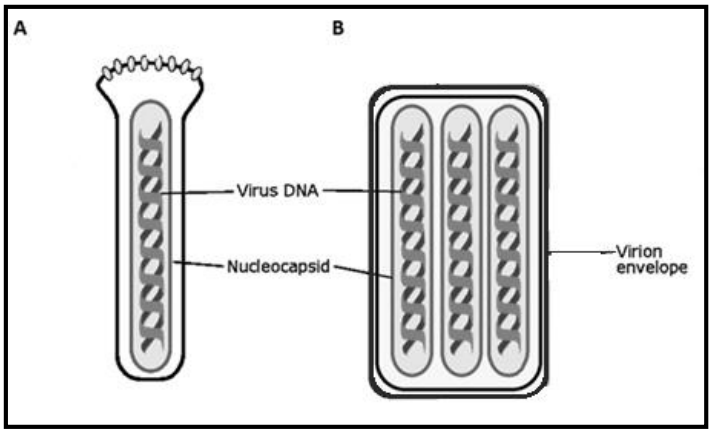


Figure 1.4: Viral phenotypes. A) Budded virus with one nucleocapsid; B) Occluded derived virus with multiple nucleocapsids (Adapted from Kalmakoff & Ward, 2003).

Two OB morphologies were then reported (Figure 1.5) and the baculoviruses were classified into two genera *Nucleopolyhedrovirus* (NPV) and *Granulovirus* (GV), of which NPVs may contain either many or single ODVs, whilst GVs contain a single ODV in an OB (Ackermann & Smirnoff, 1983; Jehle *et al.*, 2006a). It was discovered that an ODV may contain one or more nucleocapsids and this led to a grouping of the NPVs as single nucleopolyhedroviruses (SNPVs) containing one enveloped nucleocapsid per virion (Figure 1.5B) and multiple nucleopolyhedroviruses (MNPVs) with multiple nucleocapsids per virion (Figure 1.5C). Granuloviruses have a granular appearance that can be observed under high magnification under a light microscope. The granules are also referred to as OBs, and consist of a viral encoded protein matrix in which a single rod-shaped, enveloped virion is occluded (Figure 1.5A) (Tanada & Hess, 1991). The nucleocapsid consists of a protein coat containing the viral DNA genome. Occlusion bodies are highly stable structures that consist of a crystalline matrix composed of a protein called polyhedrin in NPVs and granulin in GVs. Theses occlusions range in size from 0.15 µm to 15 µm (Fauquet *et al.*, 2005). More than 600 baculoviruses have been described from various insect species (Rodriguez *et al.,* 2012). NPVs have been reported from the insect orders Lepidoptera, Diptera and Hymenoptera, whilst GVs have only been recovered from Lepidoptera (Jehle *et al.,* 2006a).

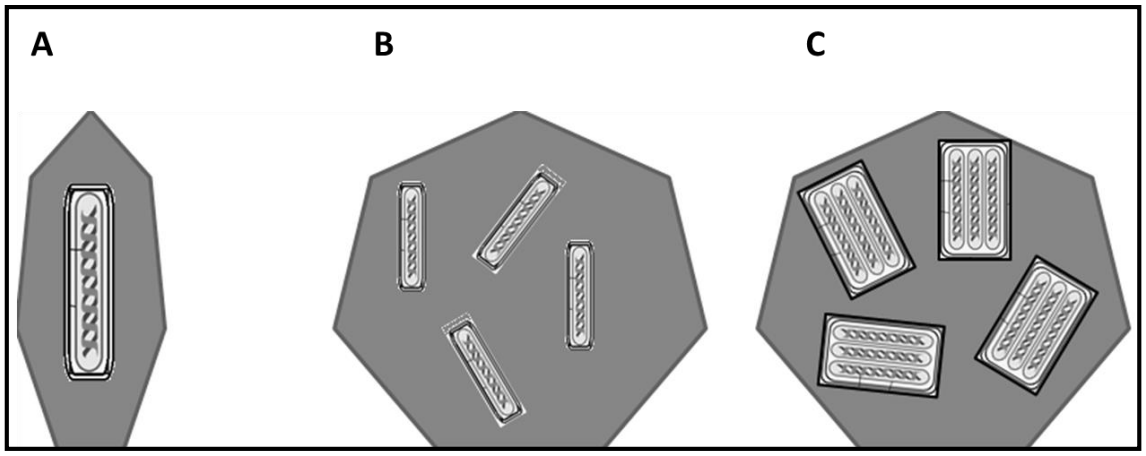


Figure 1.5: Baculovirus occlusion bodies. A) Granulovirus; B) Single nucleopolyhedrovirus; C) Multiple nucleopolyhedrovirus (Adapted from Kalmakoff & Ward, 2003).

Advances in genome analysis have improved the accuracy of classifying baculoviruses (Jehle *et al.,* 2006a). Baculovirus genomes vary in size depending on the species and consist of a circular double-stranded DNA molecule approximately 80–180 kbp that encodes between 90-180 genes (Rohrmann, 2013). The International Committee on

Taxonomy of Viruses (ICTV) has accepted a new classification on *Baculoviridae* based on DNA sequence data. It preserves correlation with both OB morphology and host taxonomic classification. The family *Baculoviridae* is now subdivided into four genera: *Alphabaculovirus* (nucleopolyhedroviruses isolated from Lepidoptera), *Betabaculovirus* (Granuloviruses isolated from Lepidoptera), *Gammabaculovirus* (nucleopolyhedroviruses isolated from Hymenoptera) and *Deltabaculoviruses* (nucleopolyhedroviruses isolated from Diptera) (Jehle *et al.,* 2006a). Baculovirus isolates are named based on the host species and the morphology of the OBs isolated (Hunter-Fujita *et al.,* 1998).

1.3.2 Infection cycle

Baculovirus infection occurs in two phases shown in Figure 1.6 (Rohrmann, 2013). Insect larvae must ingest OBs in order for baculovirus infection to occur (Figure 1.6A). The genomes of baculoviruses are packaged into nucleocapsids that are enclosed in OBs which protect virions from harsh environmental elements such as UV light and aid in delivering the virion to-the midgut of the host (Rodriguez *et al.,* 2012). Once the virus reaches the alkaline (pH 8.5-11) midgut OBs are dissolved and the ODVs are released (Hu *et al.*, 2003; Rohrmann, 2013). The released ODVs then need to pass through the peritrophic membrane of the gut and fuse with the microvilli on the columnar epithelial cells in order to enter the cells (Haas-Stapleton *et al.*, 2004). Two mechanisms have been discussed for the manner in which the ODV pass through the peritrophic membrane; one is by virus-encoded metalloproteinases and the other could be through bacterial proteinases associated with the OB (Rubinstein & Poison, 1983; Rohrmann, 2013). Once the virions have entered the epithelial cells the nucleocapsids are transported to the nuclear membrane where the viral DNA genome enters the nucleus through the nuclear pore and replicates near nuclei in virogenic stroma (Figure 1.6B). Virus replication will eventually occur in both the nuclei and cytoplasm when the nuclear membrane ruptures and budded virus particles will be produced (Hess & Falcon, 1987; Miller, 2013; Tanada & Kaya, 1993). BVs are produced when nucleocapsids move from the basal membrane to the haemocoel and acquire host membrane and viral proteins (Washburn *et al.*, 2003). BV infect other tissues such as fat body, endodermis, muscle sarcolemma and nerve ganglia (Figure 1.6C). At this stage the first phase of viral replication is complete (Washburn *et al.,* 2003; Rohrmann, 2013).

The second phase of the viral replication cycle involves the production of OBs (Figure 1.6D). This phase occurs in the cells infected with the budded viruses. The membrane of

nucleocapsids that are embedded in polyhedrin or granulin OBs is synthesised *de novo*. Cells produce large amounts of OBs and upon cell lysis the insect dies. The OBs are then spread on the food material and the viral life cycle begins again (Rohrmann, 2013). Larvae become creamy in colour and swell when infected with a baculovirus; limited movement and a pause in feeding are also typical symptoms of baculovirus infections (Rodriguez *et al.,* 2012).

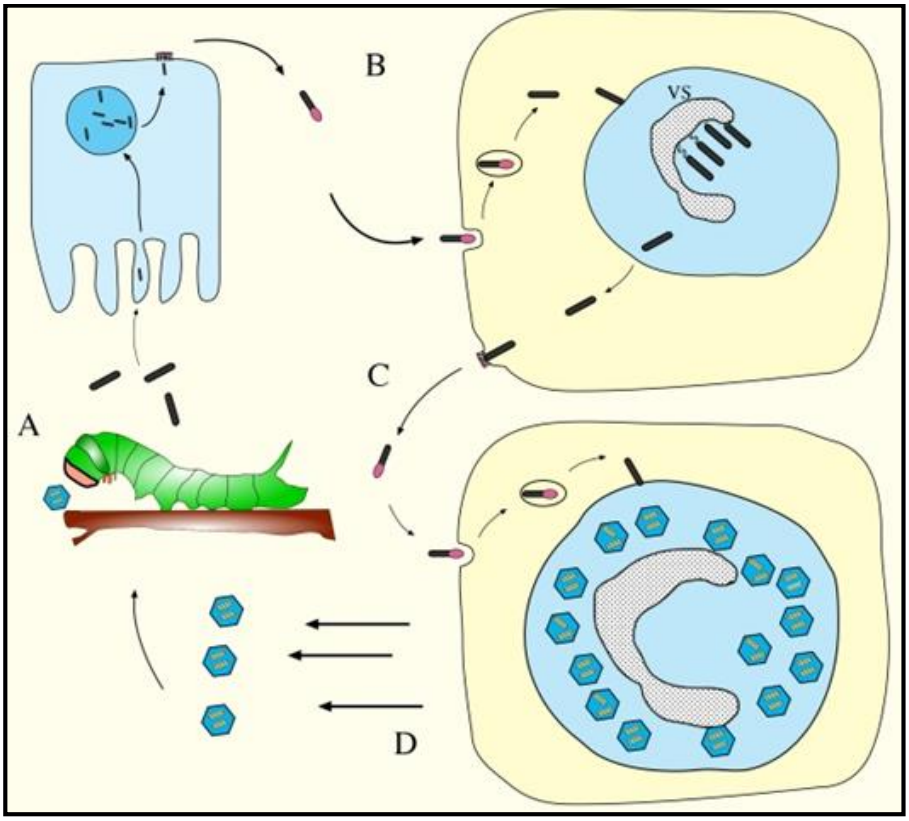


Figure 1.6: A life cycle of a baculovirus infection. A) OB ingested by an insect; B) ODV are released in the midgut, bind to epithelial cells and replicate near the nucleus; C) BV produced spread the infection throughout the insect; D) Occluded virions are produced, and the cell then dies releasing the OBs and the cycle is repeated (Rohrmann, 2013)

1.3.3 Genome and gene expression

Over 78 baculovirus genomes have been fully sequenced and are available on the GenBank database (Zhu *et al.*, 2014). Baculoviruses encode 80 to 180 predicted open reading frames in both strands (Ferelli *et al*., 2012). The wide range of genome sizes indicates that some baculoviruses lack genes present in other baculoviruses. The genome consists of 31 core genes which are a distinctive characteristic of the virus family (Miele *et al.*, 2011). The products of these genes are involved in different functions such as

transcription, replication, cell cycle arrest or interaction with host proteins, virus packaging, assembly and release, and oral infectivity (Table 1.1) and proteins for which functions have not yet been determined (Herniou *et al.,* 2003; Miele *et al.,* 2011). In 1994, the complete genome of *Autographa californica* MNPV (AcMNPV) was published (Ayres *et al.,* 1994). It is the most well-studied baculovirus genome and is often used as a model to describe baculovirus genome structure (Possee and Rohrmann, 1997).

Table 1.1: *Baculoviridae* core genes and product functions (Miele *et al.,* 2011; Ferrilli *et al.,* 2012)

	Core Genes	Function
Replication genes	*DNA pol*	Polymerization activity
	DNA helicase	DNA unwinding
	lef-1	Primase activity
	lef-2	Primase accessory factor
Transcription genes	*lef-4, lef-8, lef-9, p47*	Code for four subunits that form the viral RNA polymerase
	lef-5	Acts as an initiation factor
Structural genes: Packaging, assembly and release	*vlf-1*	Involved in the expression of very late genes such as polyhedron and granulin
	p6.9	Facilitates condensation of DNA within nucleocapsid
	vp39	Main nucleocapsid protein
	vp1054	Associated with BV
	vp91	Associated with ODV
	gp41, 38k, p49	Associated with nucloecapsid of both BV and ODV
	p33, odv-ec43, odv-nc42, odv-e18, alk-exo	Unknown function
Cell cycle arrest and/or interaction with host	*odv-e27*	Virion structure
	ac81	Unknown function
Oral infectivity	*pif-1, pif-2, pif-3, p74*	Form stable complex essential for oral infectivity
	pif-4	Essential
	odv-e56	ODV specific envelope protein

The genome replication cycle of baculoviruses is not well understood. In addition to the 4 core replication genes shown in Table 1, *lef-3* and *ie-1* are genes that are also essential in the replication of most baculoviruses, however, these genes are less conserved (Luque *et*

al., 2001). The gene *lef-3* encodes for a protein that binds to single stranded DNA during the replication process and ie-1 encodes for an activator of early transcription (Mikhaĭlov, 2003; Rohrmann, 2011; Ferrilli *et al.,* 2012). The transcription of baculovirus genes occurs in four stages, immediate early, delayed early, late and very late. The host cells transcription factors and RNA polymerase II are responsible for the transcription of early genes. Genes expressed in the late and very late stages are transcribed by the viral RNA polymerase encoded by the core transcription genes (Fuchs *et al.*, 1983; Guarino *et al.*, 1998). Proteins LEF-8 and LEF-9 have motifs common to the subunits of the DNA-dependent RNA polymerases of prokaryotes and eukaryotes. LEF-8 contains the C-terminal and LEF-9 contains the Mg^{2+} binding site of the catalytic centre. Both these proteins are highly conserved in the baculovirus family. LEF-4 is a RNA capping enzyme and P47 is a subunit with an unknown function (van Oers & Vlak, 2007). Viral transcription regulators, such as *ie-0*, *ie-1*, *ie-2* and *pe38* are not found in all baculoviruses but are also involved in the transcription of late genes. After genome replication and transcription the genomic DNA interacts with proteins encoded by the core structural genes to form nucleocapsids. Polyhedrin and granulin are also produced and are the most conserved proteins found in baculoviruses (Rohrmann, 2013). Most baculovirus genomes have an ecdysteroid-UDP-glycosytransferase (*egt*) gene, although this is not a core gene. EGT is an important protein as it causes the inactivation of moulting hormones (ecdysone) in insects. The virus benefits from the presence of this gene product as larvae continue to feed and this allows higher virus progeny yields (O'Reilly *et al.*, 1989). Cathespin and chitinase are virus encoded enzymes responsible for the cuticle rupture and dead larvae liquefaction that occurs in the final stage of baculovirus infections (Hawtin *et al.*, 1997).

Molecular studies have resulted in baculoviruses being used for a variety of applications such as the expression of recombinant proteins in insect cells and as a tool for gene transduction in mammalian cells (Kost *et al.*, 2005). The main application of baculoviruses is as a biological control agent for insect pests (Szweczyk *et al.*, 2006). The advantages and disadvantages of using baculoviruses as pest control agents are described below.

1.3.4 Baculoviruses as biocontrol agents

Baculoviruses found naturally in pest populations are important biological control agents and are used to control lepidopteran pests of economically important crops (Rodriguez *et al.,* 2012). Some of these insect viruses have narrow host ranges and can therefore not

cause harm to non-target organisms in the environment (Szewczyk *et al.,* 2006). Various baculoviruses have been tested against non-target organisms and so far no adverse effects have been reported, due to the high specificity of the viruses (McWilliam, 2007). Human and food safety concerns have increased over the years and the European Commission has reviewed and removed over two thirds of pesticides that were used to control pests and plant diseases (European Commission, 2009). Baculoviruses have been studied as an alternative method of control and have been proven safe for humans to work with and are safe for the environment as no toxic residues are released in the environment (Rodriguez *et al.,* 2012). Many chemical insecticides are banned due to the harmful residues that remain in the fields and can be found on the crops (Quin *et al.,* 2011). The viruses are more frequently used in IPM strategies due to their advantages. Baculoviruses are low risk control agents that are stable in the environment for long periods when conditions are favourable and most are applied using simple spray methods (Rodriguez *et al.,* 2012). Another advantage is that the viral activity of baculoviruses on the host is not affected when mixed with other pesticides (Copping *et al.*, 1992). This means that if an IPM strategy requires some chemical pesticide usage, usually for other pests, the efficacy of the virus will not be reduced.

The use of baculoviruses as a pest control method is not without limitations and disadvantages. The high specificity of the virus means that chemical pesticides may still be required to control other insect pests present in the orchards. This may be costly for farmers, as farmers prefer broad spectrum control agents (Szewczyk *et al.,* 2011). The costs involved in the production of biological control agents is usually high (Ignoffo *et al.,* 1977). This cost will directly affect the cost of the products and will make it difficult for private companies to invest in the production and sales of the biopesticides (Canningham, 1995). This does not allow biopesticides to be competitive with chemicals (Black *et al*. 1997). Another disadvantage of baculoviruses is that the late larval stages are more resistant to infections, for insects which do not burrow into the fruit. The neonate larvae need to be targeted, as done for the control of *C. pomonella*, which do burrow into the fruit, or higher concentrations of virus will need to be applied to achieve control over the insect pest (Washburn *et al*., 2003). The persistence of the virus in the field is affected by solar radiation, which is known to degrade the virus particles (Moscadi, 1999; Rodriguez *et al.,* 2012). This would mean that more applications of the virus would be needed and would increase the cost of the control method. However, the biggest disadvantages of

baculoviruses are that insects must ingest the virus in order to get infected and the infection has a slow speed of kill. Infected larvae continue feeding throughout the period of infection causing damage to crops due to the slow speed of kill (Ignoffo *et al*, 1992). Although there are several disadvantages to the use of baculoviruses in controlling insect pests, CpGV, used to control *C. pomonella* has been developed into a successful control agent. Field trial results have shown that the reduction in the damage caused to apples is the same as with chemical insecticide use (Huber & Dickler, 2009).

1.4 *Cydia pomonella* granulovirus (CpGV)

1.4.1 Description, pathogenesis, and transmission of CpGV

Cydia pomonella granulovirus (*Betabaculovirus*) was first isolated from diseased *C. pomonella* larvae collected in Valle Chihuahua, Mexico (Tanada, 1964). The virus has a circular, double-stranded DNA genome and both budded and occluded virus phenotypes have been observed (Tanada and Hess, 1991; Federici, 1997). CpGV OBs are ovo-cylindrical and approximately 360 nm in length and 190 nm in width (Tanada, 1964).

CpGV has a faster killing speed than most baculoviruses and has the ability to infect a broader range of tissues (Tanada and Hess, 1991; Federici, 1997). CpGV infects fat body, tracheal matrix cells, hypodermis, and malpighian tubules (Tanada & Leutenegger, 1968). Fat body cells produce the highest numbers of OBs that are infectious upon ingestion by other larvae. The infection cycle is the same as that of other baculoviruses, the OBs must be ingested by the larvae to allow infection. Studies have shown that when lethal concentrations of virus have been ingested, neonate larvae are killed within 3 days (Ballard *et al.*, 2000). Swelling and a glossy, moribund appearance are the symptoms of baculovirus infections. Closer to death, the larvae become milky and liquefy (Figure 1.7) (Lacey *et al.,* 2008).

CpGV has been recovered in small amounts in naturally infected larvae in different parts of the world. This indicates that vertical or horizontal transmission is necessary for maintenance of the virus in nature. Horizontal transmission occurs when the larvae feed on fruit, contaminated eggs and leaf surfaces (Steineke & Jehle, 2004). Transmission of the virus also occurs between instars and also from females that survive light infections. These females have the ability to transmit virus to eggs (Etzel & Falcon, 1976).

Figure 1.7: Diseased larvae infected with granulovirus (Lacey *et al.,* 2008)

1.4.2 Specificity

In addition to the high virulence of CpGV, it also has little effect on species other than *C. pomonella*, even if closely related. This means that as a biological control agent the virus would be safe to use, as non-target organisms would not be affected by its application. This has been documented by several researchers (Lacey *et al.,* 2008). Although CpGV can infect other *Cydia* species and species in closely related genera in the family Tortricidae, a higher dosage would be necessary to kill these insects (Falcon *et al.,* 1968). Due to its specificity, the use of CpGV aids in the conservation of other natural enemies in the orchards as the need for broad spectrum insecticides will be reduced.

1.4.3 Commercial development and field applications

Cydia pomonella granulovirus is the most effective microbial insecticide used to control *C. pomonella*. It was first commercialised by Sandoz Corporation (Switzerland) as a product named SAN 406. The EPA granted an experimental use permit in 1981 and as a consequence, the product was successfully tested worldwide between 1981 and 1984. Although, the results showed the efficacy of the product, Sandoz Corporation was terminated and the commercial development of CpGV products in the USA has also ceased. However, in 1979, the Commission of European Communities (CEC) had established a 'Biological Control in Apple Orchards' programme which supported further research on the use of CpGV in orchards throughout Europe (Falcon & Huber, 1991). Collaborative work was done between the European government agencies, Swiss scientists

and companies to develop and commercialise CpGV products. Madex™ (Andermatt Biocontrol, Switzerland), Granupom™ (Hoechst, Germany) and Carpovirusine™ (Calliope, France) were then produced and since then CpGV products have been used on approximately 100000 ha in Europe annually (Eberle & Jehle 2006). More CpGV products have been developed and are registered by different companies, shown in Table 1.2 (Arthurs & Lacey, 2004; Vincent *et al.*, 2007). All the commercially available products, registered before the year 2000 were based on CpGV-M and even though other isolates of CpGV had been discovered in England, Russia and Iran none of these replaced CpGV-M in commercial products until resistance occurred (Crook *et al.,* 1985; Rezapanah *et al.*, 2002).

Table 1.2: Commercially available CpGV based products

Manufacturer	Country of origin	Product	Isolate used	Obs/litre
Andermatt Biocontrol AG	Switzerland	Madex	CpGV-M	3×10^{13}
		Madex Max, Madex Top	Laboratory selected isolates	
		Madex Plus	CpGV genotype mixture	
		Madex Twin	Laboratory selected isolate used for both *Cydia pomonella* and *Grapholita molesta*.	
		Madex I12	CpGV-I12 (Iran)	
Arysta Lifescience Corporation	France	Carpovirusine, Carpovirusine 2000	CpGV-M	1×10^{13}
		Carpovirusine evo2, Carpovirusine Super Sc	Laboratory selected isolates	
Probis GmbH	Germany	Granupom	CpGV-M	2.2×10^{13}
Certis	United States of America	Cyd-X	CpGV-M	3×10^{13}
BioTepp	Canada	Virosoft	CpGV-S (Quebec region, Canadian isolate)	4×10^{13}
		virosoft CP4	CpGV isolated from CP4 region, Canadian	

The most common method of applying baculoviruses for the control of insect pests is through spraying (Hunter-Fujita *et al*., 1998). All formulations of CpGV are manufactured and sold as suspension concentrates. Commercial products available have ultra violet (UV) protectants due to the virus' sensitivity to solar radiation (UVB, 180-320 nm). However, UV radiation is still one of the main contributing factors that limits persistence of virus in the field (Jacques, 1985; Ignoffo, 1992). Successful trials of experimentally and commercially produced virus have been recorded across Europe, North America, Argentina, New

Zealand, Australia and South Africa (Lacey *et al.,* 2008). Factors that play a role in the success of the applications are the density of the insect population, the dosage, frequency and timing of virus application, the number of generations, and the environmental conditions. The number of applications varies and depends on the number of generations. The success of the virus applications is determined by assessing fruit damage and scouting for over-wintering larvae in tree bands (Lacey *et al.,* 2007).

1.4.4 Resistance

Cydia pomonella resistance to a variety of chemical pesticides has been reported in Europe and North America (Dunley & Welter 2000; Reyes *et al.,* 2007). CpGV has been developed as an alternative control method to manage *C. pomonella* populations in orchards. However, with the frequent use of CpGV, resistance has been reported in several European countries (Fritsch *et al.,* 2005; Eberle & Jehle 2006; Sauphanor *et al.,* 2006). Studies suggest that the resistance may have occurred due to sex-linked inheritance of dominant-resistant genes (Asser-Kaiser *et al.,* 2007). The mode of inheritance of CpGV resistance was determined by mass crossing experiments and single pair crossing experiments (Eberle & Jehle, 2006; Asser-Kaiser *et al*., 2007) between a susceptible and a resistant *C. pomonella* strain, followed by back-crossings and susceptibility tests of the offspring. A European project (SustainCpGV) was established to identify and characterise more virulent CpGV isolates in order to understand how the various isolates can be used in resistance management strategies. Currently most CpGV biopesticides are based on the Mexican isolate (CpGV-M). At least four other related CpGV strains had been reported from broadly separated countries, before cases of resistance were recorded (England, Canada, Russia, and Iran) (Crook *et al.,* 1985; Vincent *et al.,* 2007). Recently, more CpGV isolates from Iran, Georgia, Argentina and China have been recovered (Eberle *et al.*, 2009; Arneodo *et al.*, 2015; Fan & Wu, unpublished). A number of authors have compared variations between CpGV isolates (Harvey & Volkman, 1983; Crook *et al*., 1985; Voudouris *et al.*, 2011). Two isolates, one from Russia (CpGV-R) and another from England (CpGV-E), showed small genotypic differences. CpGV-E was found to be equally infectious to the Mexican strain. Estimates for the infectivity of CpGV-R have varied significantly. This presents incentive for more isolates to be recovered and characterised. To determine the potential for resistance to develop in a *C. pomonella* population, the frequency of genes responsible for resistance in *C. pomonella* populations, the number and frequency of CpGV application per season

and the size of the treated areas and their proximity to untreated *C. pomonella* populations need to be assessed (Lacey *et al*., 2008).

1.5 Morphological and genetic characterisation of baculoviruses

Transmission electron microscopy (TEM) can be used to classify baculoviruses as NPVs or GVs based on the morphology of the OBs. Molecular studies on CpGV are performed to determine genetic differences amongst isolates from different geographical regions. Restriction endonuclease (REN) analysis can be used to determine genetic variation among different baculovirus isolates (Goto *et al.*, 1992). This technique also allows the estimation of the genome size as the total genomic DNA is extracted from the virus and digested with a range of restriction enzymes. The resulting DNA fragments can be separated according to their sizes by agarose gel electrophoresis to form specific restriction patterns known as DNA fingerprints. The characterisation and comparison of seven CpGV isolates from Europe, North America and New Zealand using restriction enzyme analysis showed only small genotypic differences among the isolates (Crook *et al.,* 1985). Isolates from Russia and England have been shown to vary greatly to the CpGV-M strain (Harvey & Volkman, 1983). Polymerase chain reaction (PCR) is an alternative method used to detect the virus in laboratory colonies and field populations of *C. pomonella*. Highly conserved baculovirus genes*, lef-8, lef-9, granulin* are usually PCR amplified and used for phylogenetic analysis (Jehle *et al.,* 2006b).

Another method of identifying and distinguishing between different baculovirus isolates is by full genome sequencing. The nucleotide sequence obtained is subjected to BLAST (Basic local alignment search tool), for comparison and alignment with other related organisms. Single nucleotide polymorphisms (SNPs) can be used to track polymorphisms between organisms. Full genome sequencing produces accurate and reliable data but is expensive and the data analysis is time consuming. Luque *et al.* (2001) reported the complete sequencing of the CpGV genome and the genetic similarities with granuloviruses of two other lepidopterans.

1.6 Determining the biological activity of CpGV

Several bioassays for CpGV preparations and products have been developed in order to determine the biological activity of the isolate before the production process (Laing & Jaques, 1980; Huber, 1981; Glen & Payne, 1984; Lacey *et al.,* 2008). Neonate larvae are

placed in individual vials and are fed artificial medium treated with CpGV suspensions of various concentrations. *Cydia pomonella* bioassays can be performed only by using neonate larvae as the larvae burrow into the fruit and will therefore only be exposed to the virus for a short period (Lacey *et al.,* 2008). The median lethal concentration (LC_{50}) which is the concentration required to kill 50% of the test insect population is then determined using five to seven concentrations of virus. The aim of using several concentrations is to obtain mortalities between 10 and 90% in order to determine two viral concentrations below and above the LC_{50}. Higher concentrations of virus are required when attempting to produce a range of mortalities for late larval stages, as the larvae are less susceptible to infections and the number of larvae and replicate tests vary amongst researchers (Lacey *et al.,* 2008). Bioassays are useful tools which aid in determining CpGV product quality, stability and resistance in *C. pomonella* populations. Laboratory bioassays have also been developed, where the neonate larvae are fed virus-treated leaf disks and fruit (Ballard *et al.,* 2000; Lacey *et al.,* 2007). A technique that permits even virus coverage of fruit and exposure to simulated solar radiation has been developed (Lacey *et al.,* 2007). Several other bioassays have been described for other insect viruses that can be developed into CpGV bioassays. Some of the LC_{50} values obtained from CpGV bioassays are 17 OBs per neonate larva per 28 mm^2, 2.4 x 10^3 OBs per ml of diet and 35 granules per mm^2 from bioassays developed by Laing & Jaques (1980), Glen & Payne (1984) and Lacey *et al.* (2005) respectively. CpGV has been shown to be highly virulent and has been shown to have a LD_{50} value of 1.2 granules for neonates (Vincent *et al.,* 2007).

1.7 Statement of research problem

The codling moth, *C. pomonella,* is a serious insect pest responsible for severe damage caused to pome fruit produced world-wide. *Cydia pomonella* granulovirus (CpGV) is highly pathogenic for *C. pomonella* and has been developed into one of the most important biopesticides used for the control of the insect pest. Most commercially available CpGV biopesticides are based on a Mexican isolate (CpGV-M) discovered in 1963. These have been applied in pome fruit production areas in most European countries, New Zealand, North and South America and in South Africa. Although, CpGV products are regarded as the most successful baculovirus insecticides, reduced susceptibility to the biopesticides has been observed, and resistance has been reported in some *C. pomonella* populations found in commercial apple orchards across Europe. In order to address the problem associated

with the occurrence of reduced susceptibility and resistance to products developed using the Mexican isolate, bioprospecting for novel CpGV isolates is essential in the development of resistance management strategies.

In 2013, a novel CpGV isolate was recovered from a field in the Free State Province of South Africa. Preliminary work suggests that the isolate is genetically different from the Mexican isolate (Chambers, 2014). The full genome sequences of genetically different CpGV isolates as well as partial sequences of CpGV isolates from Iran, Georgia, and England and recently from the Northwest of China are available on the GenBank® database for comparison. The purpose of this study is to characterise the South African isolate of *C. pomonella* granulovirus in terms of morphology, genotype and virulence.

1.8 Chapter outline: aims and objectives

The overall aim of this study was to genetically and biologically characterise a South African *C. pomonella* granulovirus isolate and to test its virulence against laboratory-reared insect hosts.

The specific objectives were:

1. To perform transmission electron microscopy on viral samples in order to determine the size and morphology of the OBs.
2. To genetically identify the virus through PCR amplification and sequencing of selected genes.
3. To characterise the virus genome by restriction endonuclease analysis of genomic DNA and full genome sequencing.
4. To conduct comparative genomic analysis of the viral genome assembled in this study, with genomes of geographically and genetically different CpGV isolates, available on the NCBI's GenBank database.
5. To determine the virulence of the virus against neonate *C. pomonella* larvae.

Chapter 2 describes the morphological and genetic identification of virus isolated from field collected *C. pomonella* larvae in South Africa. The virus particles were detected and analysed using a transmission electron microscope. Through PCR amplification, sequencing and BLAST analysis of specific viral genes the virus was identified as CpGV-

SA. Sequence data was examined for the presence of SNPs and used in the construction of phylogenetic trees.

Furthermore, the characterisation of the CpGV-SA genome was achieved through restriction endonuclease analysis and complete genome sequencing as discussed in Chapter 3. The restriction profiles generated for CpGV-SA were compared to that of CpGV-M extracted from Carpovirusine® (Arysta Lifescience, France). A comprehensive comparison between the CpGV-SA genome sequence obtained and the reference isolate, CpGV-M1, as well as a recent genome submission of CpGV-M, both representing genome type A was conducted. The differences recorded in this chapter revealed that the isolate is genetically different from the reference isolate, CpGV-M1 and CpGV-M.

In Chapter 4, a comparative genomic analysis of the CpGV-SA genome and the genomes of CpGV isolates, available on the NCBI's GenBank database, representing genome types A-E was performed. *In silico* restriction digests, whole genome alignments and a multiple alignment of the *pe38* gene were performed to verify that CpGV-SA is a novel isolate and to determine if it could be considered a resistance overcoming strain.

The biological activity of CpGV-SA against the host neonate larvae was determined and is shown in Chapter 5. To evaluate the virulence of CpGV-SA surface concentration-mortality response and time-mortality response biological assays were carried out alongside CpGV-M extracted from Carpovirusine®.

Chapter 6 is a general discussion of the results from the previous chapters focussing on the potential development of CpGV-SA for resistance management strategies and general control of *C. pomonella* in South Africa.

Chapter 2

Morphological and genetic identification of a granulovirus isolated from field collected *Cydia pomonella* larvae in South Africa

2.1 Introduction

Cydia pomonella is a major pest of pome fruit worldwide, including in South Africa. The current control of *C. pomonella* in South Africa is dependent on the application of chemical insecticides in combination with alternative control methods such as mating disruption and the use of biopesticides (Pringle *et al.*, 2003; Addison, 2005; Lacey & Shapiro-Ilan, 2008). The commercially available biopesticides used in South Africa, namely Madex® (Andermatt Biocontrol, Switzerland) and Carpovirusine® (Arysta Lifescience, France), are based on a Mexican isolate of *Cydia pomonella* granulovirus (CpGV-M). This isolate had been successfully used for the control of *C. pomonella* for over 15 years before the first cases of resistance and reduced susceptibility to the virus developed in insect populations across Europe (Fritsch *et al.*, 2005; Sauphanor *et al.,* 2006; Schmitt *et al.*, 2013). Novel virus isolates from the same species of virus could be used to overcome resistance, due to genetic variations resulting in different virulence properties (Eberle *et al.,* 2008; Berling *et al.*, 2009b; Eberle *et al.*, 2009; Gebhardt *et al.*, 2014). It is therefore important to identify and genetically characterise novel isolates found in different geographical regions as it may provide alternative control agents that can be used in resistance management strategies.

Transmission electron microscopy (TEM) or scanning electron microscopy (SEM) is commonly used to detect and morphologically characterise purified virus particles (Lacey *et al.*, 2008; Kumar *et al.*, 2011; Abdulkadir *et al.*, 2013). It is particularly important to be able to view baculoviruses to determine whether the occlusion bodies (OBs) are granuloviruses (GV) or nucleopolyhedroviruses (NPV), as the viral phenotype is used in the initial taxonomic classification of the virus (Rohrmann, 2013). Electron microscopy has been used as a technique in diagnostic studies for the identification of viruses in *C. pomonella* populations and studying the replication cycle of the virus (Hess & Falcon, 1987; Zimmermann *et al.*, 2013). The comparison of morphological characteristics such as shape and size between unidentified and known viruses can aid in the identification of unknown viruses. However, genetic characterisation is required and is more accurate in determining the species identity of the virus.

Various molecular techniques are used in the identification and genetic characterisation of baculoviruses. These include PCR amplification and sequencing of selected viral genes or, more comprehensively, sequencing and analysis of the complete viral genome (Luque *et al.*, 2001; Lange & Jehle, 2003; Lange *et al.*, 2004; Jehle *et al.*, 2006b). The *granulin*, *late expression factor* 8 *(lef*-8) and *late expression factor* 9 (*lef*-9) genes are the primary targets for PCR amplification from granulovirus genomic DNA (Jehle *et al.*, 2006b). These genes are present in all members of the virus family and have been shown to be highly conserved amongst baculoviruses and can therefore be used to identify isolates and analyse the phylogenetic relationships between isolates (Herniou *et al.*, 2003; Lange *et al.*, 2004; Jehle *et al.*, 2006b). Concatenation of the gene sequences is performed and is a reliable method of constructing phylogenetic trees as each gene contributes to the overall phylogenetic analysis (Eberle *et al.*, 2009; Arneodo *et al.*, 2015). The *ecdysteroid UDP-glucosyltransferase* (*egt*) gene is another gene which is conserved amongst baculoviruses and has been successfully used in phylogenetic studies (Carpio *et al.*, 2013; Jukes *et al.*, 2014).

Prior to phylogenetic studies the sequence data of complete or partial gene sequences is submitted to the NCBI's Basic Local Alignment Search Tool (BLAST) to search for similar sequences in the NCBI's GenBank database to identify the purified virus and closely related isolates (Altschul *et al.*, 1990; Benson *et al.*, 2013). A list of sequences with a high percentage of identity obtained from the NCBI's GenBank database are aligned with the sequence data obtained from the purified virus and analysed for individual nucleotide changes referred to as single nucleotide polymorphisms (SNPs) which may be synonymous (silent substitutions), or non-synonymous resulting in amino acid changes in the protein. Genetically distinct CpGV isolates have been isolated in countries such as England, Russia, Iran, Georgia, Canada, Argentina and China (Vincent *et al.*, 2007; Eberle *et al.*, 2009; Gebhardt *et al.*, 2014; Arneodo *et al.*, 2015; Fan & Wu, unpublished). The full genome sequence of the *in vivo* cloned strain of CpGV-M, referred to as CpGV-M1 is available on the NCBI's GenBank database and is used as the reference sequence in CpGV studies (Luque *et al.*, 2001). The full genome sequences of five other isolates (M, I12, S, E2 and I07) and partial *granulin*, *lef*-8 and *lef*-9 sequences are available for the other geographically different isolates (Eberle *et al*., 2009; Gebhardt *et al.,* 2014; Arneodo *et al.,* 2015). Genotyping into genome types, A to D, based on SNPs present in these gene sequences in comparison to CpGV-M1 has been proposed (Eberle *et al*., 2009). The

nucleotide sequences of the reference isolate, CpGV-M1 (Luque *et al.,* 2001) and an additional CpGV-M (Gebhardt *et al.,* 2014) genome sequence are both classified under genome type A and the combination of specific variations in the *granulin*, *lef*-8 and *lef*-9 sequences allows for the classification of the genetically different isolates into the other genome types.

The overall aim of this chapter was to morphologically and genetically identify purified virus samples isolated from *C. pomonella* larvae collected from a farm in the Free State, South Africa and supplied by River Bioscience (Pty) Ltd. South Africa. The first specific objective involved the use of transmission electron microscopy to detect the virus particles and determine the size and morphology. The second objective was to purify viral OBs and extract genomic DNA in order to PCR amplify the *granulin*, *egt*, *lef*-8 and *lef*-9 viral gene sequences. The PCR amplicons were sequenced and submitted to BLAST in order to identify the novel baculovirus isolate and analyse potential SNPs within the selected genes. Sequence alignments were also used to construct a phylogenetic tree based on concatenated *granulin, lef*-8 and *lef*-9 nucleotide sequences to determine the genome type of the South African isolate.

2.2 Material and methods

2.2.1 Morphological identification by transmission electron microscopy

Purified OB samples obtained from River Bioscience (Pty) Ltd., South Africa, were isolated from diseased and dead larvae collected from a farm in the Free State Province, where no biopesticides had ever been applied. The virus samples were purified using the glycerol gradient method described by Opoku-Debrah *et al.* (2013). The extracts were imaged by placing 5 µl of the virus sample onto formvar, carbon coated grids (Wirsam Scientific, South Africa) for 30 seconds. The grids were then negatively stained using 5 µl of 1% uranyl acetate (w/v) applied for 30 seconds. Excess virus sample and stain were removed with sections of filter paper after the respective applications. The grids were left to dry overnight and viewed the following day using the Libra® 120 Plus (Zeiss, Germany) transmission electron microscope. Images were captured using an Olympus MegaView^{G2} CCD camera and the sizes of OBs were measured using iTEM software (Olympus, Japan). The data obtained were further analysed in Microsoft Excel® 2013.

2.2.2 Genomic DNA extraction

A modified cetyltrimethyl ammonium bromide (CTAB) DNA extraction protocol described by Opoku-Debrah *et al*. (2013) was used to extract genomic DNA from the purified OBs. 90 µl of Na_2CO_3 (1 M) was added to a 200 µl aliquot of OBs in a 1.5 ml tube prior to incubation at 37°C in a water bath for 30 min. The mixture was neutralised with 120 µl of Tris-HCl (1 M, pH 6.8) before 50 µl of SDS (10% w/v) and 50 µl of Proteinase K (25 mg/ml) were added. The tube was then incubated at 37°C for 30 min. Thereafter, 10 µl of RNaseA (10 mg/ml) was added and the tube was further incubated at 37°C for 30 min.

Following the incubation, the tube was centrifuged at 12100 ×*g* in an Eppendorf MiniSpin® desktop centrifuge for 3 min. The supernatant was transferred to a 2 ml tube to which 400 µl of CTAB buffer (54 mM CTAB, 0.1 M Tris-HCl pH 8.0, 20 mM Na_2EDTA, 1.4 M NaCl, pre-warmed to 70°C) was added. The tube was incubated at 70°C for 1 hour and inverted several times every 10 min. 400 µl of chloroform (stored at 4°C) was added to the mixture and centrifuged at 6700 ×*g* for 10 min. The upper phase of the supernatant was transferred to a new 2 ml tube and 400 µl of ice-cold isopropanol (-25°C) was added. The tube was stored at -25°C overnight and was centrifuged at 12100 ×*g* for 20 min the following morning. The resulting pellet was washed with 1 ml of ice-cold ethanol (70% v/v, -25°C) and centrifuged at 12100 ×*g* for 5 min. After this final centrifugation step, the ethanol was poured off gently and the pellet left to air-dry to remove traces of ethanol. The DNA pellet was re-suspended in 20 µl Tris-HCl (10 mM, pH 8.0) and stored at -4°C for a few days or at -25°C for longer periods. The resulting DNA could then be used for polymerase chain reaction amplification of selected viral genes, restriction endonuclease analysis and whole genome sequencing.

2.2.3 Determination of DNA concentration and quality

Spectrophotometry and electrophoresis were used to determine the concentration and quality of the DNA extracted from the purified virus samples. 1 µl of the DNA sample was placed on the lower optical surface of a NanoDrop 2000 spectrophotometer (Thermo-Scientific, USA), which measured the DNA concentration across the whole absorption spectrum (220-750 nm). The spectrophotometer was calibrated and cleaned using the manufacturer's instructions.

To observe the quality of DNA extracted, DNA samples were then visualised by 0.6% agarose gel electrophoresis run at 90 V for 30 min in 1X TAE buffer (40 mM Tris-acetate, 20 mM acetic acid, 1 mM EDTA) stained with ethidium bromide. Gel images were captured with the UVIpro chemi (UVItec, UK) UV trans-illuminator.

2.2.4 PCR amplification of the *granulin*, *egt, lef*-8 and *lef*-9 genes

Polymerase chain reaction amplification of the *granulin* gene was performed using universal degenerate primers previously described by Lange *et al.* (2004). The primers were designed by performing ClustalW (Thompson *et al.*, 1994) multiple alignments of the *granulin* genes from 22 fully sequenced baculovirus genomes. The conserved regions within the nucleotide sequences were then used for designing the degenerate primers with target regions for (-21) M13 forward and (-29) M13 reverse standard sequencing primers using Lasergene software (DNASTAR, USA). The degenerate primer target positions were numbered according to positions in the *Autographa californica* multiple nucleopolyhedrin virus (AcMNPV) (GenBank: L22858) and the use of these primers results in partial sequences of the *granulin* gene of lepidopteran-specific baculoviruses (Lange *et al.,* 2004) (Figure 2.1). The degenerate primer pair used for the *granulin* gene amplification reaction is shown in Table 2.1.

The CpGV-M1 reference genome was downloaded from GenBank (NC_002816) and imported into Geneious software (New Zealand) version R7 (Luque *et al.,* 2001; Kearse *et al.*, 2012). Oligonucleotide primers for the amplification of the partial *egt* sequence were designed by targeting regions within the gene, whereas the oligonucleotide primers for the amplification of the *lef-8* and *lef-9* gene sequences were designed to flank the genes in order to obtain the full gene sequences (Figure 2.1). All oligonucleotide primers were synthesised by Inqaba Biotechnical Industries (Pty) Ltd., South Africa (Table 2.1).

Table 2.1: Oligonucleotide primer pairs used in PCR and sequencing reactions of conserved viral genes

Target Gene	Gene Position	Gene Orientation in Genome	Primer name	Position in Genome	Primer Sequence (5' to 3')	Amplicon (nt)
granulin	1-747	*5'-3'*	prPH-1	*42075-42088	#tgtaaaacgacggccagtNRCNGARGAYCCNTT	507-510
			prPH-2	*41373-41389	#caggaaacagctatgaccDGGNGCRAAYTCYTT	
egt	120853-122307	*3'-5'*	egtF	120853-120875	TCATTTACTCCAATATTTATTGC	1455
			egtR	122285-122307	ATGGGACGATACACTCCAAATG	
lef-8	113169-115790	*3'-5'*	lef8F	113078-113099	GATACCCGCATCATGATACACC	2802
			lef8R	115860-115880	CATTTAATCGCTACCGTGCAC	
lef-9	99306-100805	*5'-3'*	lef9F	99207-99225	AATGTGCTCTCACCGTTAC	1679
			lef9R	100867-100885	ACACTCGTCAATGTGGTAG	

*AcMNPV genome positions (GenBank: L22858). #Lowercase letters indicate nucleotides for (-21) M13 forward and (-29) M13 reverse sequencing primers, (N= C, A, T, or G; R= A or G; Y= C or T; D= A, G, or T).

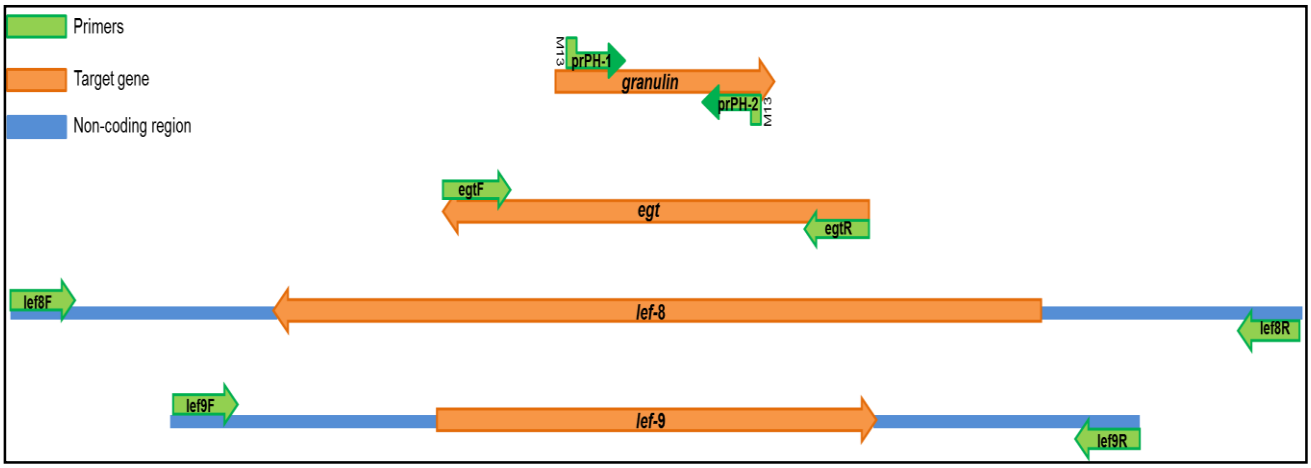


Figure 2.1: Binding sites for the *granulin, egt, lef*-8 and *lef*-9 primer pairs

The *granulin* and *egt* amplification reactions consisted of 1X *Taq* ReadyMix PCR Kit (Kapa Biosystems, USA), 0.8 µM of each of the oligonucleotide primers and between 200 to 400 ng of genomic DNA as the template. Reactions were made up to a total of 25 µl using ddH_2O and with each gene amplification a control reaction with no template was set up. Primer pair prPH-1 and prPH-2 was used for the amplification of the *granulin* gene while primer pair egtF and egtR was used for the *egt* reaction, details of these primer pairs are shown in Table 2.1 and Figure 2.1.

The amplification reactions of the *lef*-8 and *lef*-9 genes were set up as the *granulin* and *egt* reactions but consisted of 1X Maxima™ Hot Start Green PCR Mix (Thermo Scientific, USA) due to the larger size of the genes. For the *lef*-8 gene the lef8F and lef8R primers were

used while the *lef*-9 reaction comprised of primer pair lef9F and lef9R (Table 2.1 and Figure 2.1). The cycling parameters differed in terms of the annealing temperatures, durations of each step and number of cycles for each of the amplification reactions due to the melting temperatures of the oligonucleotides and the expected amplicon sizes. The cycling parameters for the *granulin, egt, lef*-8 and *lef*-9 reactions are shown in Table 2.2. A MJ Mini™ Gradient Thermal Cycler (BIO-RAD, USA) was used and PCR amplicons were visualised by viewing a 0.7% agarose gel run at 90 V for 45 min in 1X TAE stained with ethidium bromide, images were captured as before.

Table 2.2: Cycle parameters for the PCR amplification of *granulin*, *egt*, *lef*-8 and *lef*-9 gene sequence

Stage	Ta (°C)	*granulin*		*egt*		*lef-8/lef-9*	
		Time (sec)	Cycles	Time (sec)	Cycles	Time (sec)	Cycles
1	95	180	1 x	300	1 x	300	1 x
2	95	30		30		30	
	*52/50/58/55	60	35 x	40	30 x	30	$45/30 x
	72	60		45		#147/120	
3	72	600	1 x	300	1 x	300	1 x

*****Annealing temperature of *granulin* gene (52°C), *egt* gene (50°C), *lef*-8 gene (58°C) and *lef*-9 gene (55°C). #147 seconds elongation time for *lef*-8 gene and 120 seconds for *lef*-9 gene. $45 cycles for *lef*-8 gene and 30 cycles for *lef*-9 gene.

2.2.5 Sequence alignments of the *granulin, egt, lef*-8 and *lef*-9 genes

The *granulin, egt* and *lef*-9 PCR amplicons were sequenced by Inqaba Biotechnical Industries (Pty) Ltd., South Africa. The *egt* and *lef*-9 genes were sequenced in the forward and reverse direction using gene specific primer pairs while M13 sequencing primers were used to sequence the *granulin* gene in both the forward and reverse direction (Table 2.1).

MEGA 5.2 (Tamura *et al.*, 2011) was used to align sequence data obtained for each of the genes. Ambiguous nucleotides were edited using the alignments and sequence chromatograms viewed in Finch TV (USA) version 1.4.0. A single consensus sequence was assembled for the *granulin, lef*-9 and *egt* sequence alignments. The PCR amplification of the *lef*-8 gene was unsuccessful. The sequence for this gene was acquired from the complete genome sequence obtained for the CpGV isolate in this study (see Chapter 3).

The sequences obtained for the *granulin*, *egt* and *lef*-9 genes were submitted to BLAST to identify similar sequences. MEGA 5.2 was used to align the *granulin*, *egt, lef*-8 and *lef*-9 sequence data obtained in this study against the respective gene sequences from the reference isolate genome, CpGV-M1, and the sequences of other geographically different CpGV isolates available on the GenBank database, shown in Table 2.3. The alignments were examined for synonymous or non-synonymous SNPs.

Table 2.3: Data available on GenBank for CpGV *granulin*, *egt*, *lef*-8 and *lef*-9 genes of geographically different isolates

Isolate	Origin	Source	*granulin*		*egt*		*lef-8*		*lef-9*	
			Size (nt)	Accession no.	Size (nt)	Accession no.	Size (nt)	Accession no.	Size (nt)	Accession no.
*SA	South Africa	This study	502		1382		2622		1500	
*M1	Mexico	Luque *et al.* (2001)	747	NC_002816	1455	NC_002816	2622	NC_002816	1500	NC_002816
*E2	England	Gebhardt *et al.* (2014)	747	KM217577	1455	KM217577	2622	KM217577	1500	KM217577
*I12	Iran	Gebhardt *et al.* (2014)	747	KM217576	1455	KM217576	2622	KM217576	1500	KM217576
*M	Mexico	Gebhardt *et al.* (2014)	747	KM217575	1455	KM217575	2622	KM217575	1500	KM217575
*I07	Iran	Gebhardt *et al.* (2014)	747	KM217574	1455	KM217574	2622	KM217574	1500	KM217574
*S	Canada	Gebhardt *et al.* (2014)	747	KM217573	1455	KM217573	2622	KM217573	1500	KM217573
A11-2	Russia	Jehle *et al.* (2006)	501	AY706670			757	AY706561	258	AY706626
M39-1	*Grapholitha molesta*	Jehle *et al.* (2006)	500	AY706668			757	AY706557	258	AY706622
A6-4	*Grapholitha funebrana*	Jehle *et al.* (2006)	499	AY706667			655	AY706556	260	AY706621
G01	Georgia	Eberle *et al.* (2009)	474	EU370250			641	EU370241	258	EU370259
G02	Georgia	Eberle *et al.* (2009)	510	EU370249			742	EU370242	258	EU370260
I66	Iran	Eberle *et al.* (2009)	467	EU370248			756	EU370239	258	EU370257
I68	Iran	Eberle *et al.* (2009)	510	EU370247			756	EU370240	258	EU370258
I12	Iran	Eberle *et al.* (2009)	510	EU370246			744	EU370238	258	EU370256
I07	Iran	Eberle *et al.* (2009)	510	EU370245			726	EU370237	258	EU370255
I08	Iran	Eberle *et al.* (2009)	465	EU370244			756	EU370236	258	EU370254
I01	Iran	Eberle *et al.* (2009)	467	EU370243			699	EU370235	258	EU370253
E2	England	Eberle *et al.* (2009)	496	EU428824			670	EU428825		
Col19	Argentina	Arnedo *et al.* (2015)	418	KF584298			609	KF584310		
P7	Argentina	Arnedo *et al.* (2015)	418	KF584297			609	KF584309		
P118	Argentina	Arnedo *et al.* (2015)	418	KF584296			609	KF584308		
2.17	Argentina	Arnedo *et al.* (2015)	418	KF584295			609	KF584307		
3.8	Argentina	Arnedo *et al.* (2015)	418	KF584294			609	KF584306		
6.16	Argentina	Arnedo *et al.* (2015)	418	KF584293			609	KF584305		
6.9	Argentina	Arnedo *et al.* (2015)	418	KF584292			609	KF584304		
C6	Argentina	Arnedo *et al.* (2015)	418	KF584291			609	KF584303		
C1	Argentina	Arnedo *et al.* (2015)	418	KF584290			609	KF584302		
M3	Argentina	Arnedo *et al.* (2015)	418	KF584289			609	KF584301		
M10	Argentina	Arnedo *et al.* (2015)	418	KF584288			609	KF584300		
M18	Argentina	Arnedo *et al.* (2015)	418	KF584287			609	KF584299		
KS1	China	Fan & Wu, unpublished	747	KJ184166						
ALE2	China	Fan & Wu, unpublished	747	KJ184165						
ALE1	China	Fan & Wu, unpublished	747	KJ184164						
WW1	China	Fan & Wu, unpublished	747	KJ184163						
ZY1	China	Fan & Wu, unpublished	747	KJ184162						
CJ01	China	Fan & Wu, unpublished	210	JQ003555						

*CpGV isolates used in the phylogenetic and genome type analysis

2.2.6 Phylogenetic analysis of the *granulin*, *lef*-8 and *lef*-9 genes

The nucleotide sequences for the *granulin, lef*-8 and *lef*-9 genes of the CpGV isolate in this study were re-aligned against the respective genes of 6 CpGV isolates (M1, E2, I12, M, I07 and S) (Table 2.3). These CpGV isolates have previously been allocated to genome types and were chosen due to the availability of full length sequences for each of the genes (Luque *et al.,* 2001; Gebhardt *et al.,* 2014). These alignments were then concatenated into a single nucleotide sequence for each of the isolates and maximum likelihood and maximum parsimony trees were inferred using MEGA 5.2. The baculovirus *Cryptophlebia leucotreta* GV (CrleGV) was used as an out group for each inference. The best model was identified in MEGA 5.2 for the maximum likelihood tree with each phylogenetic tree constructed with 1000 bootstrap replicates.

2.3 Results

2.3.1 Transmission electron microscopy of occlusion bodies

Purified virus extracted from a mixture of field collected *C. pomonella* larvae and supplied by River Bioscience (Pty) Ltd., South Africa, was detected using TEM and a high number of ovocylindrical OBs were observed (Figure 2.2). These OBs were of typical granulovirus morphology with single dense regions visible in the centre of a number of the virus particles (Figure 2.2C), which had a consistent morphology and size (Figure 2.2A-B). The average size of the OBs was 330.9 ± 24 nm (n = 50) in length and 171.7 ± 19 nm (n = 50) in width.

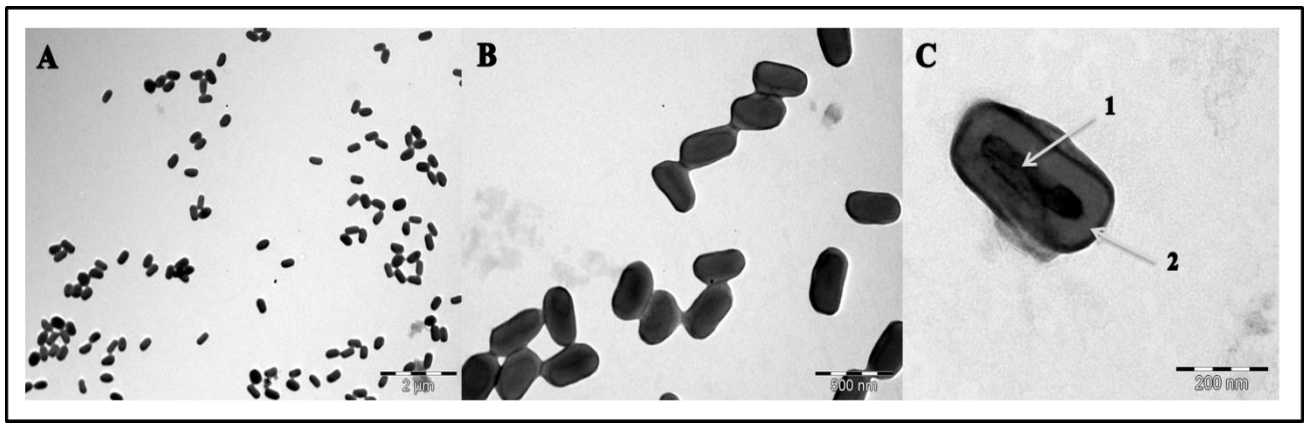


Figure 2.2: Transmission electron micrographs of particles extracted from *Cydia pomonella* larvae. A-B) Images of particles observed from separate virus samples. C) Particle with a dense central region (1) and a lighter outer region (2).

2.3.2 Genomic DNA extraction and DNA concentration

DNA was successfully extracted from aliquots of purified OBs using the CTAB method described in section 2.2.2. The typical appearance of DNA extracted from OBs is shown in Figure 2.3. A concentrated, bright band forms above the 10000 bp mark but the actual size of the DNA cannot be determined due to its large size and limitations in gel resolution. The average DNA concentration was 105.5 ± 34.8 ng/µl (n=5).

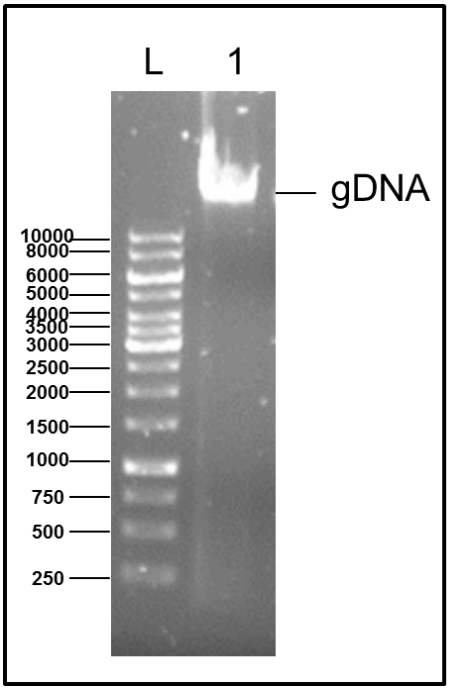


Figure 2.3: DNA extraction using the CTAB method run on a 0.6% agarose gel. L – GeneRuler™ 1 Kb DNA ladder (Thermo Scientific, USA), 1 – extracted DNA (2 µl).

2.3.3 Amplification of the *granulin, egt* and *lef*-9 gene sequences

The *granulin, egt* and *lef*-9 genes were successfully amplified using genomic DNA extracted from purified OBs (Figure 2.3). The amplified products were analysed by 0.7% agarose gel electrophoresis. The amplification of the *granulin* gene produced a dense, bright band of approximately 500 bp (Figure 2.4A) and the amplified product of the *egt* gene formed a band slightly below the 1500 bp mark (Figure 2.4B). The *lef-9* gene produced a band that was estimated to be about ±1500 bp (Figure 2.4C). The amplicon sizes were similar to the estimated sizes of the products based on where the oligonucleotides are positioned in the reference genome. The estimated sizes of the *granulin, egt, lef*-9 genes are 507-510, 1455 and 1679 nt respectively and are shown in Table 2.1.

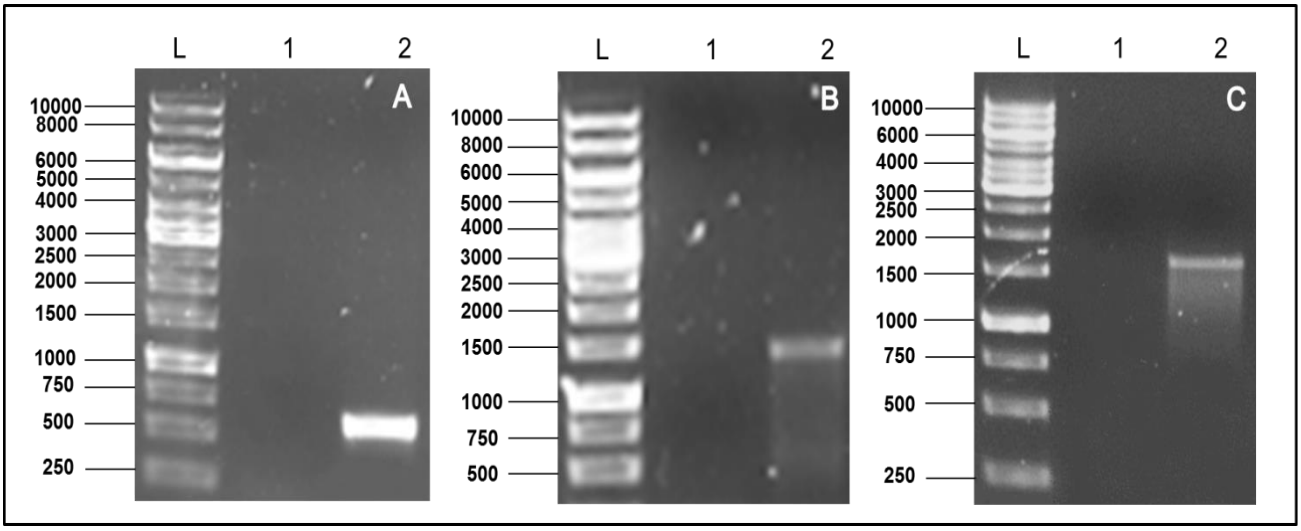


Figure 2.4: PCR amplification of the *granulin, egt* and *lef*-9 genes run on 0.7% agarose gels. L– GeneRuler™ 1 Kb DNA ladder, 1– No template control, 2– A) *granulin* amplicon, 2– B) *egt* amplicon, 2– C) *lef*-9 amplicon.

2.3.4 Analysis of the *granulin*, *lef*-8, *lef*-9 and *egt* gene sequences

Partial *granulin* and *egt* sequences were obtained as the oligonucleotides were designed within the gene sequences. The complete *lef*-9 sequence was obtained as the oligonucleotides were designed to flank the gene (Figure 2.1). An alignment of the partial *granulin,* partial *egt* and complete *lef*-9 genes from the isolated granulovirus against the complete *granulin, egt* and *lef*-9 genes from the reference isolate, CpGV-M1, showed a 99% identity between each sequence, with a few synonymous and one non-synonymous SNP detected.

The *granulin* sequence acquired was 502 nt with BLAST results showing an identity of 99% (E value of 0.0) when aligned against the reference isolate, CpGV-M1. Analysis of the *granulin* alignments revealed a total of three SNPs at nucleotides 246C→T, 249A→C and 564T→C with none resulting in amino acid changes (Table 2.4). The *egt* sequence acquired was 1382 nt with BLAST indicating an identity of 99% (E value of 0.0) when compared against the reference isolate. Analysis of the *egt* alignments revealed a SNP at nucleotide 121144A→G which does not result in an amino acid change. An additional SNP was detected at the start of the *egt* gene when the sequence data were aligned with the complete genome sequence of the virus isolate used in this study (see Chapter 3). This SNP is at nucleotide 122295A→G and results in an amino acid change (Table 2.4). The sequence acquired for *lef*-9 was 1500 nt with BLAST indicating an identity of 99% (E value of 0.0) against the reference isolate. Analysis of the *lef*-9 alignments revealed a SNP at

nucleotide 99629G→A which does not result in an amino acid change (Table 2.4). This high degree of identity reveals the recovered virus, obtained from River Bioscience Pty Ltd, is a novel member of the species CpGV and was thereafter named CpGV-SA.

PCR amplification of the *lef*-8 gene was unsuccessful (data not shown). The sequence for this gene was acquired from the complete genome sequence of CpGV-SA (see Chapter 3) and was 2622 nt in length. BLAST indicated an identity of 99% (E value of 0.0) against the reference isolate, CpGV-M1. Analysis of the *lef*-8 alignments revealed a SNP at nucleotide 113579C→G which resulted in an amino acid change shown in Table 2.4.

Multiple alignments of the *granulin* (Table 2.5A), *lef*-8 (Table 2.5B), *lef*-9 (Table 2.5C) and *egt* (Table 2.5D) sequence data obtained in this study against the respective gene sequences of other geographically different CpGV isolates available on the GenBank database (Table 2.3) revealed several sets of SNPs across each of the isolates in comparison to the reference isolate, CpGV-M1. Based on the SNPs observed, 16 isolates have previously been assigned to genome types A-E (Eberle *et al.,* 2009; Arneodo *et al.,* 2015). In Table 2.5A-D, the nucleotide positions are numbered according to positions in CpGV-M1, the nucleotides highlighted in green are an indication of nucleotides that would be observed for genome groups A-E and the nucleotides highlighted in blue are an indication of those that have not been classified as belonging to a genome type.

Table 2.4: Single nucleotide polymorphisms (SNPs) found in the *granulin*, *egt, lef*-8 and *lef*-9 genes after alignment with the reference isolate, CpGV-M1

Gene	Nucleotide position/codon	Isolate		Amino acid
		Mexico	South Africa	
granulin	246	CA**C**	CA**T**	Histidine
	249	CC**A**	CC**C**	Proline
	564	TC**T**	TC**C**	Serine
egt	#122295	**A**CT	**G**CT	Threonine » Alanine
	121144	GA**A**	GA**G**	Glutamic acid
* *lef-8*	113579	**C**AA	**G**AA	Glutamine » Glutamic acid
lef-9	99629	TT**G**	TT**A**	Leucine

Letters in red bold colour indicate single nucleotide polymorphisms. (») indicates amino acid changes. *Sequence obtained from complete genome sequence (Chapter 3). #SNP observed from complete sequence obtained from genome sequence of CpGV-SA

Table 2.5A: SNPs found in the *granulin* gene after multiple alignments with the respective gene from CpGV isolates of different geographical origins

Isolate	Origin	Genome Type	Size (nt)	Nucleotide Position											
				106	213	246	249	325	362	381	515	549	564	603	627
M1	Mexico	A	747	C	A	C	A	G	T	G	T	C	T	G	A
SA	South Africa	ua	502	C	A	T	C	G	T	G	T	C	C	G	A
E2	England	B	747	C	A	T	C	G	T	G	T	C	C	G	A
I12	Iran	D	747	C	A	C	A	G	T	G	T	C	T	G	A
M	Mexico	A	747	C	A	C	A	G	T	G	T	C	T	G	A
I07	Iran	C	747	T	A	C	A	G	T	A	T	C	C	G	A
S	Canada	E	747	C	A	T	C	G	T	G	T	C	C	A	A
A11-2	Russia	ua	501		A	T	A	G	T	G	T	C	Y	G	A
M39-1	*Grapholitha molesta*	ua	500		A	C	A	G	T	G	T	C	T	G	A
A6-4	*Grapholitha funebrana*	ua	499		A	C	A	G	T	G	T	C	T	G	A
G01	Georgia	C	474		A	C	A	G	T	A	T	C	C	G	A
G02	Georgia	A	510		A	C	A	G	T	G	T	C	T	G	A
I66	Iran	A	467		A	C	A	G	T	G	T	C	T	G	A
I68	Iran	C	510		A	C	A	G	T	A	T	C	C	G	A
I12	Iran	D	510		A	C	A	G	T	G	T	C	T	G	A
I07	Iran	C	510		A	C	A	G	T	A	T	C	C	G	A
I08	Iran	D	465		A	C	A	G	T	G	T	C	T	G	A
I01	Iran	D	467		A	C	A	G	T	G	T	C	T	G	A
E2	England	B	496		A	T	C	G	T	G	T	C	C	G	A
Col19	Argentina	ua	418			T	C	G	T	G	T	C	C	G	A
P7	Argentina	ua	418			T	C	A	T	G	T	C	C	G	A
P118	Argentina	A	418			C	A	G	T	G	T	C	T	G	A
2.17	Argentina	*A	418			C	A	G	T	G	T	C	T	A	A
3.8	Argentina	A	418			C	A	G	T	G	T	C	T	G	A
6.16	Argentina	A	418			C	A	G	T	G	T	C	T	G	A
6.9	Argentina	A	418			C	A	G	T	G	T	C	T	G	A
C6	Argentina	ua	418			T	C	G	T	G	T	C	C	G	A
C1	Argentina	ua	418			T	C	G	T	G	T	C	C	G	A
M3	Argentina	ua	418			T	C	G	T	G	T	C	C	G	A
M10	Argentina	ua	418			T	C	G	T	G	T	C	C	G	A
M18	Argentina	ua	418			T	C	G	T	G	T	C	C	G	A
KS1	China	ua	747	C	A	T	C	G	C	G	T	C	C	G	G
ALE2	China	ua	747	C	A	C	A	G	T	G	T	C	T	G	A
ALE1	China	ua	747	C	G	C	A	G	T	G	A	C	T	G	A
WW1	China	ua	747	C	A	T	C	G	T	G	T	C	C	G	G
ZY1	China	ua	747	C	A	T	C	G	T	G	T	T	T	G	A
CJ01	China	ua	210					G	T	G					

*Sequence does not match that of CpGV-M1 but isolate has been assigned to genome type A (Arneodo *et al.*, 2015). Genome type refers to the complete set of genes specific to a group of viruses (Rohrmann, 2013).

Table 2.5B: SNPs found in the *lef*-8 gene after multiple alignments with the respective gene from CpGV isolates of different geographical origins

Isolate	Origin	Genome Type	Size (nt)	Nucleotide Position														
				113190	113233	113579	113677	113760	114102	114213	114267	114274	114354	114654	114693	114822	114882	115705-115706
M1	Mexico	A	2622	G	T	G	C	G	G	G	G	G	C	T	G	G	G	
SA	South Africa	ua	2622	G	T	C	C	G	G	G	G	G	C	T	G	G	G	
E2	England	B	2622	G	C	C	C	G	A	G	A	G	C	A	G	G	G	
I12	Iran	D	2622	G	T	G	A	A	G	G	G	A	C	T	A	G	G	
M	Mexico	*A	2622	G	T	C	C	G	G	G	G	G	C	T	G	G	G	
I07	Iran	C	2622	A	T	C	C	G	G	G	A	G	C	T	A	G	A	TGT
S	Canada	E	2622	G	T	C	C	G	G	G	G	G	C	T	G	A	G	
A11-2	Russia	ua	757						G	G	G	G	C	T	G			
M39-1	*Grapholitha molesta*	ua	757						G	G	G	G	C	T	G			
A6-4	*Grapholitha funebrana*	ua	655						G	G	G	G	C	T	G			
G01	Georgia	C	641						G	G	A	G	C	T				
G02	Georgia	A	742						G	G	G	G	C	T	G			
I66	Iran	A	756						G	G	G	G	C	T	G			
I68	Iran	C	756						G	G	A	G	C	T	A			
I12	Iran	D	744						G	G	G	A	C	T	A			
I07	Iran	C	726						G	G	A	G	C	T	A			
I08	Iran	D	756						G	G	G	A	C	T	A			
I01	Iran	D	699						G	G	G	A	C	T	A			
E2	England	B	670						A	G	A	G	C	A	G			
Col19	Argentina	ua	609						G	G	G	G	C	T	G			
P7	Argentina	ua	609						G	A	A	G	T	T	A			
P118	Argentina	A	609						G	G	G	G	C	T	G			
2.17	Argentina	A	609						G	G	G	G	C	T	G			
3.8	Argentina	A	609						G	G	G	G	C	T	G			
6.16	Argentina	A	609						G	G	G	G	C	T	G			
6.9	Argentina	A	609						G	G	G	G	C	T	G			
C6	Argentina	ua	609						G	A	A	G	T	T	A			
C1	Argentina	ua	609						G	A	A	G	T	T	A			
M3	Argentina	ua	609						G	A	A	G	T	T	A			
M10	Argentina	ua	609						G	A	A	G	T	T	A			
M18	Argentina	ua	609						G	A	A	G	T	T	A			

***Sequence does not match that of CpGV-M1 but isolate has been assigned to type A (Gebhardt *et al.,* 2014).

Table 2.5C: SNPs found in the *lef*-9 gene after multiple alignments with the respective gene from CpGV isolates of different geographical origins

Isolate	Origin	Genome Type	Nucleotide position					
			Size (nt)	99354	99359	99381	99629	100247
M1	Mexico	A	1500	T	A	T	G	A
SA	South Africa	ua	1500	T	A	T	A	A
E2	England	B	1500	T	A	T	G	G
I12	Iran	D	1500	T	A	T	G	A
M	Mexico	A	1500	T	A	T	G	A
I07	Iran	C	1500	G	G	C	G	A
S	Canada	E	1500	T	A	T	A	A

Table 2.5D: SNPs found in the *egt* gene after multiple alignments with the respective gene from CpGV isolates of different geographical origins

Isolate	Origin	Genome Type	Nucleotide position											
			Size (nt)	121046	121048	121066	121075	121384	121454	121466	121675	122071	122149	122279-122280	122295
M1	Mexico	A	1455	T	T	T	G	G	T	G	A	C	A		T
SA	South Africa	ua	1455	T	C	T	G	G	T	G	A	C	A		C
E2	England	B	1455	T	T	T	G	G	T	G	A	C	A		T
I12	Iran	D	1455	T	T	T	G	G	T	G	A	C	A		T
M	Mexico	A	1455	T	T	T	G	G	T	G	A	C	A		T
I07	Iran	C	1437	C	C	C	G	R	C	G	G	T	G	C	T
S	Canada	E	1455	T	C	T	A	G	C	C	A	C	G		T

2.3.5 Phylogenetic analysis of the *granulin*, *lef*-8 and *lef*-9 genes

Phylogenetic analysis of concatenated nucleotide sequence alignments for *granulin, lef*-8 and *lef*-9 genes of CpGV isolates M1, E2, I12, M, I07, S and the isolate from this study were performed and are shown below in Figure 2.5. For each data set a maximum likelihood consensus tree and a maximum parsimony consensus tree were created each with 1000 bootstrap replications. The Kimura 2-parameter (K2) model was identified as the best model used with frequencies and discrete Gamma distribution for the maximum likelihood tree determined with the model select feature in Mega 5.2. The bold letters next to the isolate names are the genome groups.

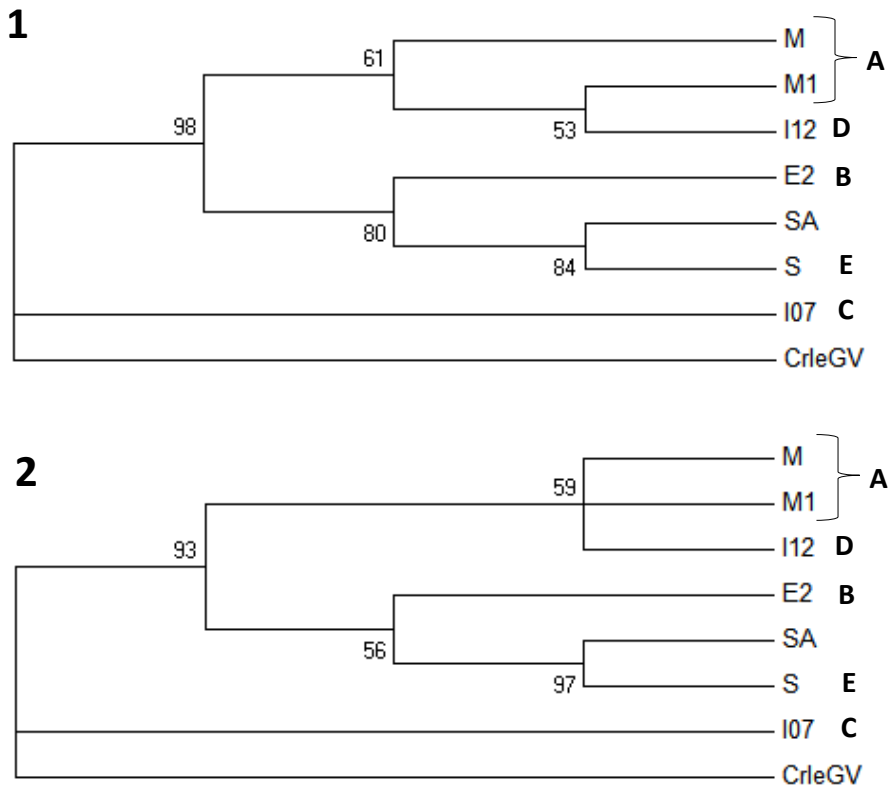


Figure 2.5: Phylogenetic reconstruction inferred using A) Maximum Likelihood and B) Maximum Parsimony of concatenated *granulin, lef*-8 and *lef*-9 nucleotide sequences. Bootstrap values based on 1000 replicates, CrleGV set as the out group.

2.4 Discussion

The use of baculoviruses to control insect pests has increased over the years as predicted by Moscardi (1999) and the frequent use of the formulated CpGV products

over a period of years has resulted in the occurrence of reduced susceptibility and resistance to the virus (Fritsch *et al.,* 2005). It is therefore important to continue the search for novel CpGV isolates and conduct further studies in order to develop resistance management strategies. Purified virus particles isolated from field collected larvae and supplied by River Bioscience (Pty) Ltd. South Africa were imaged and analysed with a transmission electron microscope. The particles observed were ovocylindrical and approximately 331 nm by 172 nm in size. The virus particles were similar in shape but slightly smaller than CpGV OBs analysed by Tanada (1964), which were approximately 360 nm by 190 nm in size. The TEM analysis indicated the presence of a purified baculovirus, specifically a GV, from *C. pomonella* larvae, which had morphological traits similar to viruses obtained from several populations of *C. pomonella* (Zimmermann *et al.,* 2013). However, the species identity of the GV could not be determined by examining the morphology alone and therefore some genetic analysis of the purified virus was performed.

The species identity of the purified virus isolated from *C. pomonella* was achieved through the comparison of *granulin, egt, lef*-8 and *lef*-9 sequences to sequence data of other virus isolates obtained from the NCBI's GenBank database. The *granulin, egt* and *lef*-9 genes were successfully PCR amplified and sequenced, while the amplification of the *lef*-8 gene was not successful due to the long length of the sequence, internal primers were designed based on the initial sequences returned but no complete sequence was obtained. The sequence of the *lef*-8 gene was however acquired through the sequencing of the full genome of the isolated virus (see Chapter 3) and used in alignments to detect SNPs and in the phylogenetic analysis of the virus. The identification of the isolated granulovirus was achieved through the submission of the gene sequences to BLAST, which revealed high percentage identities to sequences from various CpGV isolates. The *granulin, egt, lef*-8 and *lef*-9 genes from the reference isolate appeared in the BLAST results for each of the respective genes, with identity scores of >99% and E values of 0.0. This high degree of identity suggested that the isolate was a member of the CpGV family.

The multiple alignments and SNPs detected in the highly conserved genes of the isolate in this study showed that it could not be assigned to either of the proposed genome groups A-E but could rather be of a mixed genome type (Eberle *et al.,* 2009). In comparison to the reference isolate, CpGV-M1, a few SNPs were detected

two were non-synonymous, one within the *egt* gene and other within the *lef*-8 gene. From the two SNPs detected in the *egt* gene the non-synonymous SNP was detected by analysing the complete gene from the full genome sequence of CpGV-SA (Chapter 3) as opposed to the partial sequence obtained through amplification and sequencing of the specific gene, which revealed only the synonymous SNP. The *egt* sequence could be compared to only the sequences of CpGV isolates that have been fully sequenced as there is no partial sequence data for the other isolates available for comparison. The *egt* gene sequence did not have a 100% identity to any of these other isolates, revealing the novelty of the CpGV isolate. Multiple alignments using the *egt* gene of CpGV isolates has not been conducted prior to this study. The SNPs observed in the *granulin* gene of the South African isolate are identical to the SNPs observed for the CpGV-E2 isolate which has been assigned to genome type B. Therefore, the isolate could be of genome type B. However, when assigning the isolates to a genome type single genes cannot be studied in isolation particularly for the CpGV family. This has been shown for several Argentinian isolates which have genome type B SNPs in the *granulin* gene but form a unique cluster when analysing the *lef*-8 gene sequence data and could therefore not be assigned to a genome type (Arneodo *et al.,* 2015). The *lef*-9 gene sequence of the South African isolate was 100% identical to CpGV-S, suggesting the isolate is of mixed genome types B and E. The SNP observed in the *lef*-8 gene of the isolate in this study, matches the SNP detected in CpGV-M. The CpGV-E2 isolate is assigned genome type B but has some features of the genome type A sequence therefore it is expected that if the South African isolate is comprised of genome type B and E attributes then there will be some minor similarities with genome type A (Eberle *et al.,* 2009).

The only possible limiting factor to this method of genotyping is that most of the sequences available on the GenBank database are not complete sequences and therefore alignments for other isolates have important regions missing. This was the case for the *lef-9* gene (data not shown). It would seem that most of the isolates listed in Table 2.3 as having this sequence data would have identical *lef*-9 gene sequences to CpGV-M1, but the regions that have been sequenced do not contain the nucleotide positions where SNPs may be detected, which is the reason for not including these data in Table 2.5C. Alternatively, quantitative real-time PCR (qPCR)

could be used to determine whether CpGV-SA is of mixed genome types B and E and determine the genome type ratios of which the isolate's genome is comprised of. A qPCR-based assay is an accurate and useful tool used to study population genetics, as it is able to detect genotype frequencies that are low in virus populations (Zwart *et al.*, 2008).

The phylogeny of this virus was determined by constructing phylogenetic trees using the aligned and concatenated nucleotide sequences for the highly conserved *granulin, lef*-8 and *lef*-9 genes as conducted by Eberle *et al.,* (2009). The construction of a maximum likelihood tree and a maximum parsimony tree for 7 CpGV isolates (including the isolate of this study) resulted in the grouping of South African (SA) isolate with CpGV-S and CpGV-E2. This clustering received a bootstrap support value of 80 and 56 in the maximum likelihood tree and maximum parsimony tree respectively indicating that this clustering was present in over 80% and 56% of 1000 replicate trees. CrleGV was included in the phylogenetic analysis as an out group as it is closely related to the CpGV species (Lange & Jehle, 2003; Jehle *et al.,* 2006b). The phylogenetic tree clustering is based on the similarity between the sequences and therefore the similarity of SNPs detected amongst the isolates. Although, the South African isolate groups with the English (CpGV-E2) and Canadian (CpGV-S) isolates, the genotyping is not an indication of where the virus isolate is from. The genotyping of CpGV isolates based on the SNPs detected in conserved genes does correspond with the geographical origin of the isolate (Eberle *et al.,* 2009).

In conclusion, the identification and morphological characterisation by transmission electron microscopy confirmed that a granulovirus was isolated from *C. pomonella* larvae, collected from a field in the Free State, South Africa. The granulovirus was identified as a CpGV isolate through amplification and sequence analysis of viral genes *granulin, egt, lef*-8 and *lef*-9, which also showed genetic variations, suggesting the novelty of the isolate obtained. The *granulin* and *lef*-9 gene were 100% identical to isolates E2 and S respectively, this shows that the South African isolate could be of mixed genome types B and E. This result was verified in the grouping of the isolate with the E2 and S isolates in the phylogenetic trees that were generated. An isolate of this mixed genotype has not previously been described. Therefore the novel isolate was named *Cydia pomonella* granulovirus-South Africa (CpGV-SA).

The next chapter describes further genetic characterisation of CpGV-SA by restriction endonuclease analysis of genomic DNA and full genome sequencing revealing more genetic variations between CpGV-SA and the reference isolate, CpGV-M1.

Chapter 3

Genomic analysis of *Cydia pomonella* granulovirus- South Africa

3.1 Introduction

The morphological and genetic identification of a novel South African *C. pomonella* granulovirus (CpGV-SA) was described in Chapter 2. The identification was achieved through the use of transmission electron microscopy as well as the amplification and alignments of highly conserved genes. It is important to further genetically characterise novel virus isolates in order to verify initial genotyping and to determine the evolution of genetic variations between isolates of the same species (Eberle *et al.,* 2009). Studies regarding the genomics of novel baculoviruses are essential as genetic variability may result in differences in virulence, which would have an effect on (i) the development of the virus as a biopesticide, (ii) the success of current baculovirus application programmes and (iii) resistance management strategies (Moscardi, 1999; Eberle *et al.,* 2008; Berling *et al.,* 2009b).

Cydia pomonella granulovirus- Mexico (CpGV-M) was the first CpGV isolate to be discovered, biologically characterised and developed into a biopesticide (Tanada, 1964; Huber, 1981; Falcon & Huber, 1991; Moscardi, 1999). Although, the CpGV-M based products are highly efficient, after several years of use reduced susceptibility and even resistance of field populations of *C. pomonella* to CpGV-M was reported first in Germany and France, and currently across several countries in Europe (Fritsch *et al.,* 2005; Sauphanor *et al.,* 2006; Schmitt *et al.,* 2013). Studies have shown that CpGV isolates of different geographical origins and genotypes are able to overcome resistance in laboratory reared resistant *C. pomonella* strains and in resistant field populations (Eberle *et al.,* 2008; Eberle *et al.,* 2009; Berling *et al.,* 2009a; Berling *et al.,* 2009b). The resistance in *C. pomonella* has been shown to be specific to CpGV-M and related isolates of the genome type A. Isolates from the other CpGV genome types B-E are able to infect and successfully replicate in resistant *C. pomonella* populations (Gebhardt *et al.,* 2014). New CpGV biopesticides are based on isolates genetically distinct from CpGV-M, emphasising the need and importance of bioprospecting (Zingg, 2008). However, it is also important to study the mechanism of resistance and therefore investigate the genetic differences between

isolates in order to prevent resistance in other geographical regions and prevent the resistance to the new CpGV biopesticides from occurring (Schmitt *et al.,* 2013).

Restriction endonuclease analysis (REN) is a molecular technique commonly used to genetically characterise and determine the novelty of baculovirus isolates (Abdulkadir *et al.*, 2013; Opoku-Debrah *et al.*, 2013; Graillot *et al.*, 2014; Moore *et al.*, 2015). DNA profiles are generated using a range of restriction enzymes, which cleave the DNA at specific sites resulting in fragments of various sizes. The DNA fragments produced are then separated by agarose gel electrophoresis and when analysed different profiles between genetically distinct isolates are revealed as a result of insertions, deletions or point mutations which have occurred within the genome sequence (Rohrmann, 2013). REN analysis has also been used to determine whether isolates belong to a single genotype or are comprised of a mixture of genome types (Eberle *et al.,* 2009). The genotypic variants are detected by observing the submolar bands generated by the restriction enzymes (Berling *et al.,* 2009b). The construction of physical maps of the genomic DNA of various CpGV isolates has been achieved through the use of REN analysis, proving that this molecular technique is highly useful in genomic studies (Crook *et al.*, 1985; Crook *et al.,* 1997).

Although, REN analysis is considered to be a convenient and an inexpensive technique, in order to conduct comprehensive accurate characterisation of virus isolates, complete genome sequencing is required. The first baculovirus genome to be sequenced was of *Autographa californica* nucleopolyhedrovirus (Ayres *et al.*, 1994). The number of completely sequenced baculovirus genomes has increased since then to over 78 sequences (Zhu *et al.*, 2014). Information regarding the diversity and evolution of baculoviruses can be obtained through the analysis of the available genome sequences. However, only a few complete genome sequences of baculoviruses of the same species are available, resulting in fewer studies being done on genotypic variations between isolates (Li *et al.*, 2005; Zhang *et al.*, 2005; Ogembo *et al.*, 2009; Xu *et al.*, 2013; Gebhardt *et al.,* 2014). The lack of complete genome sequences for genomic research is changing due to next generation sequencing (NGS) technologies increasing and sequencing prices decreasing (Metzker, 2010; Quail *et al.*, 2012). Recently, new NGS platforms such as 454 sequencing (Roche, Switzerland), Ion Torrent (Thermo Scientific, USA) and Illumina sequencing (Illumina, USA) have been used in baculovirus research (Rohrmann *et al*., 2013; Gebhardt *et al.,* 2014; Zhang *et al.*, 2014). The quality of the genomic sequencing is high, accurate and rapidly produced through the use of NGS technologies (Liu *et al.*, 2012). This sequencing

technology is based on the principles of polymerase chain reaction (PCR), millions of short reads are generated resulting in high coverage across the genome sequence. Access to complete genome sequences of baculoviruses allows further studies on the evolution, phylogeny, classification and the functional genetic variability between isolates to be conducted comprehensively (Herniou *et al.*, 2001; Jehle *et al.,* 2006b; Gebhardt *et al.,* 2014)

The Mexican CpGV isolate is used as the reference isolate in all CpGV studies (Eastwell *et al.*, 1999; Rezapanah *et al.*, 2008; Eberle *et al.,* 2009; Arneodo *et al.,* 2015). Restriction profiles and the full genome sequence of an *in vivo* cloned strain, CpGV-M1 (GenBank Accession number: NC_002816), are available for genomic comparative studies (Crook *et al.,* 1997; Luque *et al.,* 2001). Recently, the complete genome of CpGV-M (GenBank Accession number: KM217575), isolated from larvae, was sequenced and is also available on the NCBI's GenBank database (Gebhardt *et al.,* 2014). Several minor nucleotide and annotation differences have been observed between the two genomic sequences and therefore both sequences are used for comparison purposes in this study (Eberle, 2010).

The overall aim of this chapter was to conduct a comprehensive characterisation of the CpGV-SA genome by restriction endonuclease analysis and complete genome sequencing. As an internal reference and the first objective of this chapter, CpGV-M occlusion bodies (OBs) were purified from Carpovirusine® (Arysta Lifescience, France), a commercial formulation of the virus, available in South Africa. Secondly, the genomic DNA extracted from the CpGV-M (extracted from Carpovirusine®) and CpGV-SA OBs were digested using the restriction enzymes *Bam*HI, *Eco*RI, *Pst*I and *Xho*I and the profiles generated were compared. The third objective was to determine the genetic variations between the South African isolate and the reference isolate CpGV-M1 (Luque *et al.,* 2001) and CpGV-M (Gebhardt *et al.,* 2014), by analysing the complete genomes and determining the gene identities by nucleotide and amino acid alignments. Further comparative genomics was conducted in comparison to other CpGV isolates from different genome types and is discussed in Chapter 4.

3.2 Materials and methods

3.2.1 Occlusion body purification from a commercial biopesticide

In order to obtain the CpGV-M isolate for comparative purposes, a modified glycerol gradient purification method described by Opoku-Debrah *et al*. (2013) was performed on a CpGV-M based biopesticide, Carpovirusine® (Arysta Lifescience, France) of which details of the formulation are unknown as these are proprietary. The product was stored according to the manufacturer's instructions at -20°C and used within the products shelf life of 2 years.

Gradients were prepared by layering glycerol solutions in order of most dense to least dense, 80-70-60-50-40-30% (v/v), in 14 mm × 89 mm ultracentrifuge tubes. Two JA-20 centrifuge tubes were filled with the CpGV-M formulation and centrifuged at 7840 ×*g* for 30 min at 4°C in a Beckman Coulter Avanti® J-E centrifuge. The supernatants were discarded and the pellets were suspended by filling the tubes with 0.1% SDS, this was followed by centrifugation and repeated twice in order to remove any remaining formulation additives. The resultant pellets were suspended in 3 ml ddH_2O and carefully added to a glycerol gradient. The gradient was centrifuged at 27783 ×*g* for 15 min at 4°C in a Beckman Coulter Optima™ L-90 K ultracentrifuge.

The OBs forming a distinct white band across the centre of the ultracentrifugation tube were collected using a pipette and transferred into two JA-20 tubes. The tubes were filled with ddH_2O and centrifuged at 7840 ×*g* for 30 min at 4°C. Centrifugation was repeated three times to ensure the glycerol solution was removed. The virus pellet was suspended in 1 ml of ddH_2O in a 1.5 ml tube and stored at -20°C. Genomic DNA was extracted from the purified OBs using the protocol described in Section 2.2.2 and served as a reference isolate in the restriction endonuclease analysis, whilst the remainder of the virus was tested along with the South African isolate in the biological activity assays discussed in Chapter 5.

3.2.2 Restriction endonuclease analysis (REN)

Restriction enzyme digestions were performed on the genomic DNA extracted from both the CpGV-SA and CpGV-M isolates using the following FastDigest™ enzymes, *Bam*HI, *Eco*RI, *Pst*I and *Xho*I (Thermo Scientific, USA). The reactions contained approximately 200 ng of the template DNA, 1X FastDigest™ buffer (Thermo Scientific, USA), 1X of the respective restriction enzyme and ddH_20 to a total volume of 30 µl. The reaction tubes were

incubated for 15 min at 37°C and the fragmented DNA was separated by gel electrophoresis alongside GeneRuler High Range DNA Ladder (Thermo Scientific, USA) and GeneRuler 1 Kb DNA Ladder (Thermo Scientific, USA). The digests were visualised by 0.6% agarose gel electrophoresis run at 30 V for 16 hours in 1X TAE buffer (40 mM Tris-acetate, 20 mM acetic acid, 1 mM EDTA) stained with ethidium bromide and gel images captured with a UVIpro chemi (UVItec,UK) UV trans-illuminator. The DNA ladders combined provided a wider range of marker sizes to be used when estimating the band sizes of the fragmented DNA using the UviBand software (UVItec, UK).

3.2.3 Complete genome analysis of CpGV-SA

The full genome of CpGV-SA was sequenced by Inqaba Biotechnical Industries (Pty) Ltd., South Africa, using a MiSeq Desktop Sequencer (Illumina). Genomic DNA, approximately 100 ng, extracted from CpGV-SA OBs (Section 2.2.2) was sequenced producing a total of 1976868 paired reads. These reads were paired together and 35% of the data were used to produce contigs through the de novo assembly of the raw reads. A total of 10000 contigs were assembled and mapped to the complete genome sequence of the reference isolate, CpGV-M1 (Accession number: NC_002816), with medium sensitivity and a total of 5 iterations producing a single consensus sequence. A pairwise alignment using the progressive mauve algorithm was performed between the consensus sequence and the genome sequence of CpGV-M1. The consensus sequence was annotated using annotation data from both CpGV-M1 and CpGV-M (Accession number: KM217575).

Pairwise nucleotide and amino acid ClustalW (Thompson *et al*., 1994) alignments of the CpGV-SA predicted open reading frames (ORF) with the ORFs of CpGV-M1 isolate were conducted. Alignments were repeated for ORFs where differences between CpGV-M1 (Accession number: NC_002816) and CpGV-M (Accession number: KM217575) were observed. Alignments were then manually checked and plots were produced. The genome assembly, annotations and analysis were carried out in Geneious (New Zealand) version R7 (Kearse *et al*., 2012).

3.3 Results

3.3.1 Occlusion body purification

The purification of CpGV-M OBs from Carpovirusine® using a 30-80% glycerol gradient was successful (Figure 3.1). A white milky band was visible across the centre of the glycerol gradient and debris, comprising of the formulation additives, was observed at the bottom of the tube. The virus band was collected, suspended in 1 ml of ddH_2O and then viewed under a transmission electron microscope using the protocol in Section 2.2.1 of Chapter 2, before DNA was extracted and further studies could be conducted.

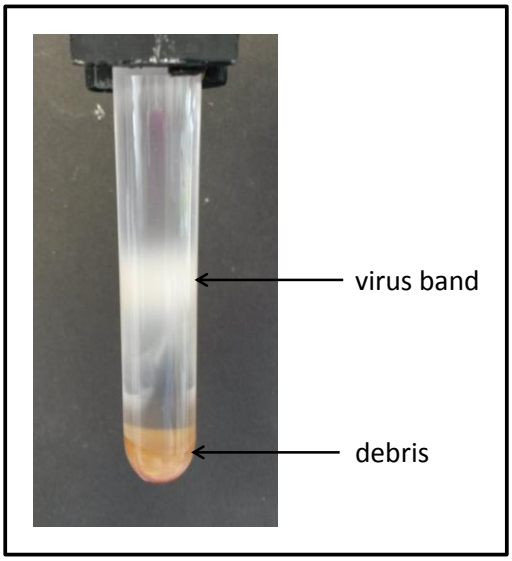


Figure 3.1: Purified CpGV-M occlusion bodies on a 30-80% glycerol gradient

3.3.2 Restriction endonuclease profiles

Restriction endonuclease analysis was carried out in order to obtain profiles of the CpGV-SA genomic DNA for comparison with CpGV-M purified from Carpovirusine®. The profiles generated through the use of restriction enzymes *Bam*HI, *Eco*RI, *Pst*I and *Xho*I revealed several bright bands for DNA fragments of a high molecular size and fainter bands for fragments of lower molecular sizes; fragments below 1000 bp were not detected (Figure 3.2). To estimate the sizes of the fragments GeneRuler High Range DNA Ladder (Thermo Scientific, USA) and GeneRuler 1 Kb DNA Ladder (Thermo Scientific, USA) were run alongside the restriction digests. Panel A of Figure 3.2 shows the DNA profiles of the CpGV-M and CpGV-SA by digestion with *Bam*HI and *Eco*RI. Restriction digestion with

*Bam*H1 revealed the presence of 10 and 9 bands for CpGV-M and CpGV-SA respectively (Figure 3.2, panel A, lanes 3 and 4). The *Eco*RI profile generated 10 bands for both isolates with high similarity in sizes (Figure 3.2, panel A, lanes 5 and 6). Panel B of Figure 3.2 shows the DNA profiles obtained of the CpGV-M and CpGV-SA by digestion with *Pst*1 and *Xho*1 enzymes. REN analysis with *Pst*1 produced 10 and 12 bands for CpGV-M and CpGV-SA respectively (Figure 3.2, panel B, lane 3 and 4). The *Xho*1 profile generated 10 and 9 bands for CpGV-M and CpGV-SA respectively (Figure 3.2, panel B, lanes 5 and 6). The following sections describe a detailed comparison of each of the CpGV-SA profiles with that of the reference isolate profiles, which were also generated *in vitro*.

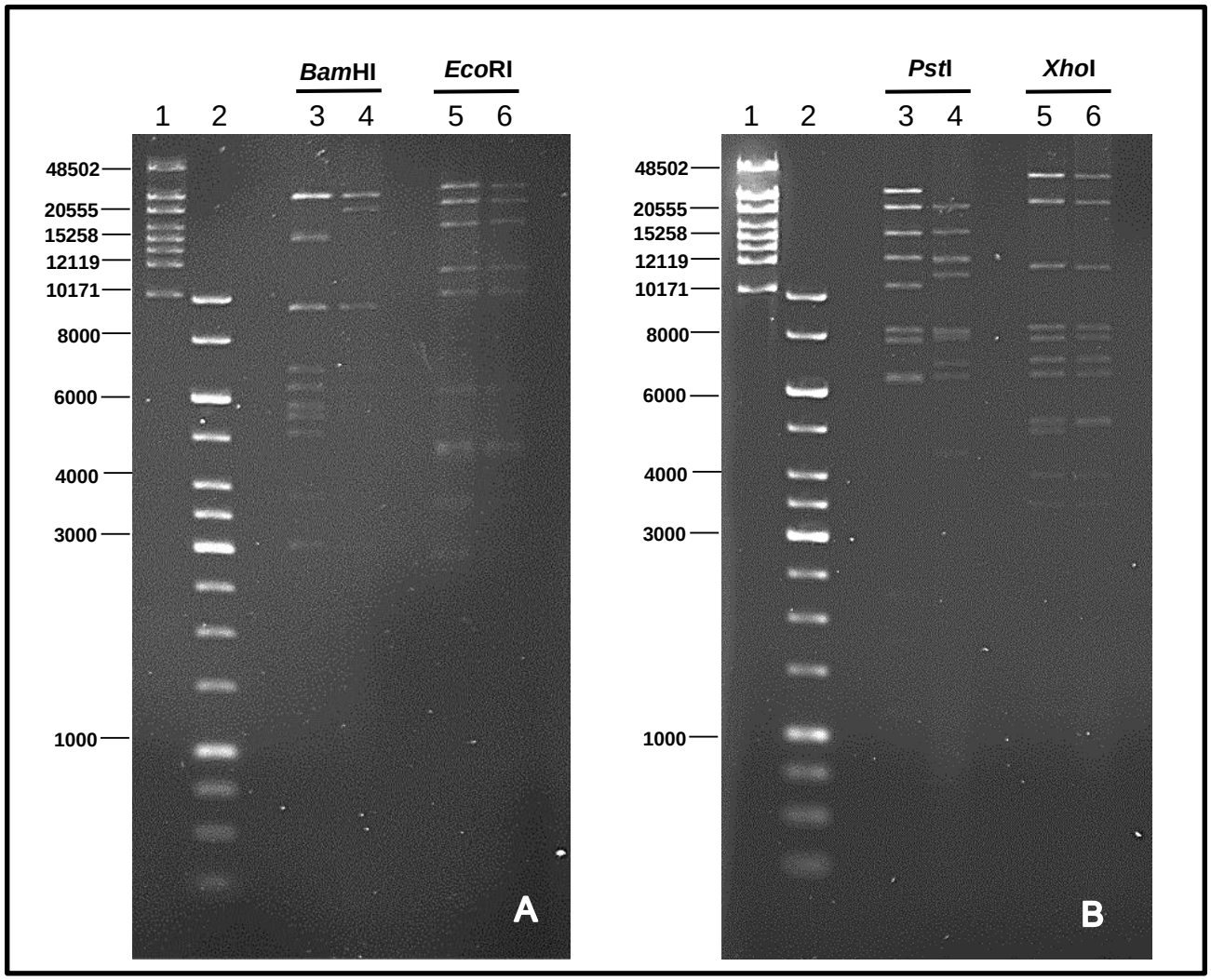


Figure 3.2: Restriction endonuclease profiles of CpGV-M and CpGV-SA genomic DNA run on 0.6% agarose electrophoresis at 30 V for 16 hours. 1- High Range DNA ladder, 2- 1 Kb DNA ladder. Panel A: *Bam*HI and *Eco*RI digests for CpGV-M and CpGV-SA respectively. Panel B: *Pst*I and *Xho*I digests for CpGV-M and CpGV-SA respectively.

3.3.2.1 *Bam*HI restriction endonuclease profile

The *Bam*HI restriction profile of the CpGV-SA genomic DNA is shown in Figure 3.3, alongside the profile generated for CpGV-M using the same enzyme. The digestion of CpGV-M DNA produced 10 fragments resulting in an approximate genome size of 88309 bp. The analysis between the two profiles was conducted by matching bands of similar sizes, fragments A2 (15285) and A8 (5156) were absent in the CpGV-SA profile, which had a fragment equivalent in size to these two fragments, B2 (20902). The B2 fragment of CpGV-SA was not present in the CpGV-M profile, therefore the CpGV-SA profile had 9 fragments. The fragments produced for both isolates ranged between 26477-3000 bp but the approximate CpGV-SA genome size was 88880 bp, which is slightly larger than the CpGV-M genome.

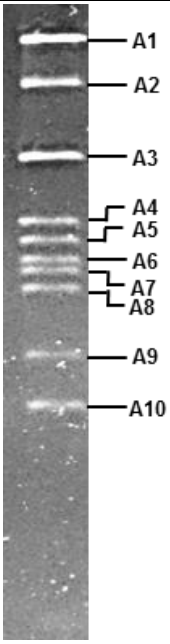
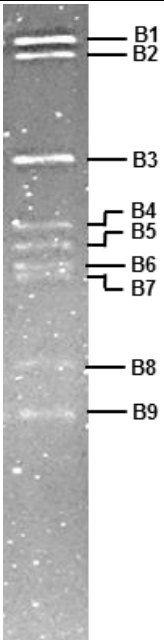
CpGV-M			CpGV-SA		
Size (bp)	**Fragment**		**Size (bp)**	**Fragment**	
26477	A1		26477	B1	
-	-		20902	B2	
15285	A2		-	-	
9679	A3		9743	B3	
7046	A4		7046	B4	
6426	A5		6426	B5	
5866	A6		5866	B6	
5552	A7		5620	B7	
5156	A8		-	-	
3822	A9		3800	B8	
3000	A10		3000	B9	
Number of fragments	10		**Number of fragments**	9	
Total	88309		**Total**	88880	

Figure 3.3: Comparison of CpGV-M and CpGV-SA *Bam*HI DNA restriction profiles

3.3.2.2 *Eco*RI restriction endonuclease profile

The digestion of CpGV-SA DNA with enzyme *Eco*RI generated a profile highly similar to that of CpGV-M. A total of 12 fragments ranging from 29278-1723 bp were observed for both isolates (Figure 3.4). Minor differences, in size, were observed between fragments A2-A7 of CpGV-M with the corresponding fragments B2-B7 of CpGV-SA. The approximate genome sizes were 120588 bp and 121050 bp for CpGV-M and CpGV-SA respectively. Although, the genome sizes determined using *Eco*RI were much larger than the sizes calculated using the *Bam*HI enzyme, the genome size of the South African isolate remained larger than the genome size of the reference isolate.

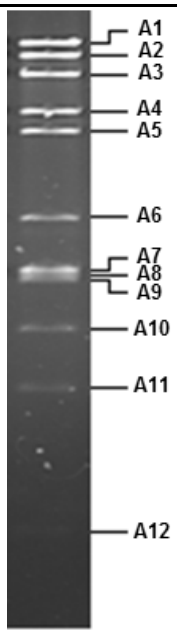
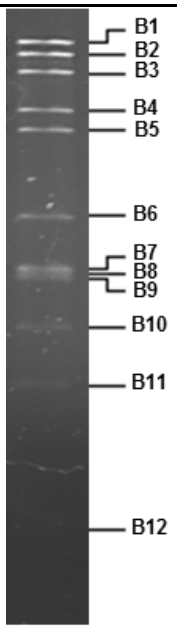
CpGV-M			CpGV-SA		
Size (bp)	**Fragment**		**Size (bp)**	**Fragment**	
29278	A1		29278	B1	
21903	A2		21644	B2	
17000	A3		17563	B3	
11811	A4		12004	B4	
10414	A5		10322	B5	
6448	A6		6448	B6	
5115	A7		5172	B7	
5029	A8		5029	B8	
4920	A9		4920	B9	
3861	A10		3861	B10	
3077	A11		3077	B11	
1732	A12		1732	B12	
Number of fragments	12		**Number of fragments**	12	
Total	120588		**Total**	121050	

Figure 3.4: Comparison of CpGV-M and CpGV-SA *Eco*RI DNA restriction profiles

3.3.2.3 *Pst*I restriction endonuclease profile

The *Pst*I restriction profiles of the CpGV-M and CpGV-SA genomic DNA are shown below in Figure 3.5. The digestion of CpGV-SA DNA produced a total of 12 fragments of which 9 could be matched to similar sized fragments in the reference DNA profile. Fragments B7 (6979), B9 (4524) and B11 (2013) were not present in the CpGV-M *Pst*I profile, comprised of 10 fragments and with an approximate size of 113523 bp. Fragment A1 (27113) was absent in the CpGV-SA profile and when fragments were analysed fragment B3 (12403) was observed to be of a greater intensity than the upper fragment and could therefore represent a doublet. If the doublet is included the approximate size of the CpGV-SA genome would be 113699 bp whereas if it excluded the approximate size is 101296 bp (Figure 3.5).

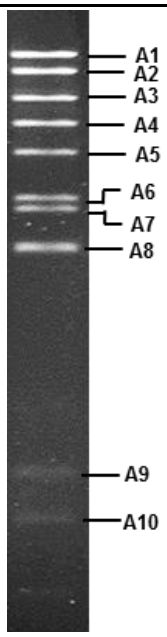
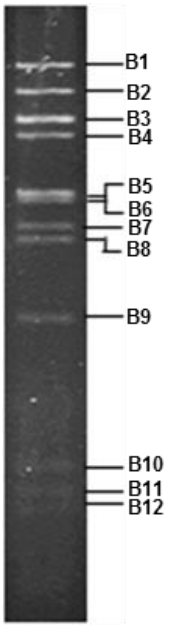
CpGV-M			CpGV-SA		
Size (bp)	**Fragment**		**Size (bp)**	**Fragment**	
27113	A1		-	-	
21188	A2		21510	B1	
15809	A3		16002	B2	
12551	A4		12403	B3	
10305	A5		11072	B4	
8344	A6		8344	B5	
7814	A7		7910	B6	
-	-		6979	B7	
6433	A8		6558	B8	
-	-		4524	B9	
2230	A9		2245	B10	
-	-		2013	B11	
1736	A10		1736	B12	
Number of fragments	10		**Number of fragments**	12	
Total	113523		**Total**	101296	

Figure 3.5: Comparison of CpGV-M and CpGV-SA *Pst*I DNA restriction profiles

3.3.2.4 *Xho*I restriction endonuclease profile

The digestion of CpGV-SA DNA with enzyme *Xho*I generated a profile with a high similarity to the CpGV-M *Xho*I restriction profile (Figure 3.6). The CpGV-SA profile had a total of 10 fragments and lacked the A9 (4960) fragment. However, the B8 (5180) fragment had a greater intensity than the upper fragments and possibly represents a doublet which would result in a genome size of 121909 bp. The CpGV-M profile had a total of 11 fragments and the total size of the CpGV-M genome was calculated to be approximately 120545 bp.

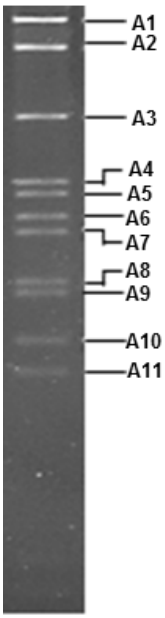
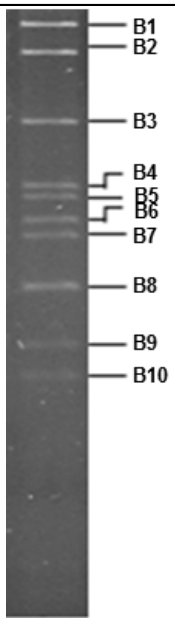
CpGV-M			CpGV-SA		
Size (bp)	**Fragment**		**Size (bp)**	**Fragment**	
38665	A1		39742	B1	
22734	A2		22734	B2	
11618	A3		11700	B3	
8330	A4		8330	B4	
7903	A5		7903	B5	
7048	A6		7048	B6	
6609	A7		6578	B7	
5180	A8		5180	B8	
4960	A9		-	-	
3969	A10		3985	B9	
3529	A11		3529	B10	
Number of fragments	11		**Number of fragments**	10	
Total	120545		**Total**	116729	

Figure 3.6: Comparison of CpGV-M and CpGV-SA *Xho*I DNA restriction profiles

3.3.3 Complete genome analysis of CpGV-SA

A complete genome sequence for CpGV-SA was produced using the Illumina MiSeq desktop sequencer. A total of 1.97 million paired reads were sequenced of which 691903 were used in the *de novo* assembly producing a total of 10001 contigs, which were rated with the highest quality in Geneious R7 and stored. The largest contig produced was comprised of 103423 reads and had a coverage of 545.4 ± 191.3. A total of 94 stored contigs were mapped to the reference isolate genome sequence CpGV-M1 producing a

single consensus sequence of 123597 bp with a GC percentage of 45.3%. Alignment of the consensus sequence for CpGV-SA against the full genome sequence of CpGV-M1 showed a pairwise identity of 97.2%.

3.3.3.1 Genome annotations

The CpGV-SA genome sequence was initially annotated using the gene data from the reference sequence CpGV-M1 (Luque *et al.,* 2001). However, a new complete sequence of the Mexican isolate (CpGV-M; accession number KM217575) is available on the NCBI's GenBank database (Gebhardt *et al.,* 2014). The open reading frames of CpGV-M that were determined as different from the CpGV-M1 were chosen as the annotations, due to having higher percentage identities when assessing the nucleotide and amino acid sequences. A total of 141 genes were identified out of 143 and 142 genes present in the complete sequences of CpGV-M1 and CpGV-M respectively (Table 3.1). Open reading frames 25 and 38 of CpGV-M1 are absent from the CpGV-SA genome. The gene sequence of ORF38 in CpGV-M1 was determined to be a mis-annotation in the sequence as the annotation appears to be within ORF37 when viewing the annotations of the CpGV-M sequence (Eberle, 2010). All annotations were manually checked and considered as complete coding regions if a start and stop codon were present and if the ORF would translate to a protein of 50 amino acids or more (Luque *et al.*, 2001; Lange & Jehle, 2003). The nucleotide sequence of ORF25 was a short length, 87 nt in the CpGV-SA, and therefore codes for only 29 amino acids. The nucleotide sequence of ORF38 was also short, 114 nt, which would code for only 38 amino acids. Both ORFs are included in Table 3.1 (shown in red) but were excluded from the rest of the genomic analysis. The ORFs in the CpGV-SA genome in comparison to the ORFs found in the reference isolate are recorded in Table 3.1 showing the ORF positions, lengths and percentage identity, determined by performing pairwise nucleotide ClustalW alignments.

In order to verify the assembly of the genome and assess the annotation accuracy, nucleotide sequences (obtained and discussed in Chapter 2) of the *granulin, lef-9* and *egt* genes were aligned with the complete genome and each of the genes were identical to the corresponding annotations in the CpGV-SA genome sequence.

Table 3.1: The positions, sizes and percentage identity of open reading frames (ORFs) detected in the CpGV-SA genome in comparison to the reference isolate ORFs. Numbering of ORFs in the left column is according to the numbering the CpGV-M1 genome. Annotations which differ in CpGV-M are shown in green and ORFs to be excluded, due to size, are shown in red.

ORF	Name	CpGV-M1/M			CpGV-SA			
		Position		Length (nt)	Position		Length (nt)	% Identity
		From	To		From	To		
1	*granulin*	1	747	747	1	747	747	99.6
2		749	1,273	525	749	1,273	525	100
3	*pk-1*	1,254	2,093	840	1,254	2,093	840	99.9
4		2,173	2,739	567	2,172	2,738	567	100
5		2,729	2,971	243	2,728	2,970	243	100
6		3,122	3,298	177	3,121	3,351	231	75.8
7	*ie-1*	3,391	4,857	1,467	3,444	4,910	1,467	99.9
8		4,963	5,541	579	5,016	5,594	579	100
9		5,581	5,886	306	5,634	5,939	306	100
10	*chitinase*	6,027	7,811	1,785	6,080	7,864	1,785	99.6
11	*cathespin*	7,934	8,935	1,002	7,987	8,988	1,002	99.9
12		9,015	9,248	234	9,068	9,301	234	100
13	*gp37*	9,318	10,073	756	9,371	10,126	756	100
14	*odv-e18*	10,145	10,399	255	10,199	10,453	255	100
15	*p49*	10,400	11,773	1,374	10,969	11,829	861	99.7
16		12,145	12,735	591	12,217	12,807	591	100
17	*iap-3*	12,865	13,692	828	12,937	13,764	828	99.6
18	*odv-e56*	13,730	14,797	1,068	14,139	14,870	732	99.6
19		15,171	15,398	228	15,244	15,471	228	100
20		15,458	16,165	708	15,531	16,229	699	98.2
21		16,434	16,613	180	16,498	16,677	180	100
22	*pep/p10*	16,836	17,879	1,044	16,899	17,942	1,044	99.7
23		17,970	18,428	459	18,033	18,491	459	100
24	*pe-38*	18,571	19,719	1,149	18,636	19,760	1,125	97.6
25		20,165	20,329	165	20,269	20,355	87	88.2
26		20,328	21,314	987	20,911	21,339	429	99.8
27		20,358	21,827	1,470	21,254	21,832	579	99.8
28/29		22,688	24,058	1,371	22,693	24,063	1,371	100
30		24,629	25,174	546	24,602	25,147	546	99.6
31	*F protein*	25,306	27,099	1,794	25,287	26,807	1,521	99.5
32		27,325	28,665	1,341	27,310	28,677	1,368	97.8
33		28,737	29,621	885	28,749	29,633	885	99.9
34		29,664	30,353	690	29,676	30,365	690	99.6
35		30,247	30,822	576	30,259	30,859	601	100
36a		30,998	31,162	165	31,010	31,174	165	100
36b		30,901	31,083	183	30,913	31,095	183	100
37	*odv-e66*	31,205	33,439	2,235	32,052	33,488	1,437	96.9
38		33,166	33,391	343	33,325	33,441	114	98.3
39		33,481	33,798	318	33,530	33,847	318	99.1
40		33,851	34,180	330	33,908	34,237	330	100
41	*lef-2*	34,321	34,836	516	34,378	34,893	516	99.6
42		34,928	35,176	249	34,985	35,233	249	100
43		35,223	35,567	345	35,280	35,624	345	100
44		35,632	36,255	624	35,689	36,312	624	99.8
45		36,314	36,778	465	36,371	36,835	465	100
46	*metalloproteinase*	36,836	38,473	1,638	36,893	38,530	1,638	99.8
47	*p13*	38,480	39,289	810	38,537	39,346	810	99.4
48		39,333	40,451	1,119	39,390	40,508	1,119	100
49		40,448	40,837	390	40,505	40,894	390	100
50/51		40,773	44,090	3,318	40,830	44,150	3,321	97.2
52a		44,099	44,839	741	44,159	44,899	741	100
52b		43,709	44,071	363	43,766	44,131	366	98.6
53		44,849	44,995	147	44,909	45,055	147	100
54	*v-ubi*	45,071	45,355	285	45,145	45,429	285	100
55		45,434	46,498	1,065	45,508	46,572	1,065	100
56		46,505	46,714	210	46,579	46,788	210	100
57	*39K*	46,791	47,516	726	46,865	47,632	768	98.9
58	*lef-11*	47,467	47,871	405	47,541	47,987	447	96.6
59	*sod*	47,826	48,224	399	47,942	48,340	399	100
60	*p74*	48,598	50,664	2,067	48,718	50,784	2,067	100
61		50,883	51,095	213	51,002	51,214	213	96.7
62		51,067	51,636	570	51,458	51,550	93	98.9
63	*bro*	51,774	51,941	168	51,688	51,930	243	97.6
64		52,654	53,346	693	52,792	53,484	693	99.9
65		53,444	53,683	240	53,582	53,821	240	100
66	*ptp-2*	53,838	54,059	222	53,976	54,192	217	97.3
67		54,134	54,397	264	54,267	54,530	264	100
68	*p47*	54,369	55,751	1,383	54,502	55,884	1,383	99.8
69		55,790	56,452	663	55,923	56,585	663	100
70		56,526	57,080	555	56,658	57,191	534	96.2

71	*p24capsid*	57,150	57,761	612	57,253	57,864	612	100
72		57,800	58,144	345	57,903	58,247	345	100
73	*38.1Kd*	58,338	58,934	597	58,441	59,037	597	99.5
74	*lef-1*	58,915	59,622	708	59,018	59,725	708	99.9
75	*pif*	59,748	61,364	1,617	59,851	61,467	1,617	99.8
76	*fgf*	61,523	62,203	681	61,626	62,306	681	99.9
77		62,271	62,582	312	62,374	62,679	306	97.8
78		62,591	62,818	228	62,688	62,915	228	100
79		62,842	63,312	471	62,939	63,409	471	99.6
80	*lef-6*	63,309	63,614	306	63,406	63,711	306	99.7
81	*dbp*	63,693	64,565	873	63,790	64,662	873	99.8
82a		64,773	65,465	693	64,851	65,543	693	99.9
82b		64,592	64,849	258	64,670	64,927	258	99.6
83	*p45*	65,355	66,674	1,320	65,433	66,752	1,320	99.2
84	*p12*	66,706	67,035	330	66,784	67,113	330	99.7
85	*p40*	67,094	68,236	1,143	67,172	68,314	1,143	99.7
86	*p6.9*	68,268	68,417	150	68,346	68,495	150	100
87	*lef-5*	68,518	69,246	729	68,596	69,324	729	99.6
88	*38K*	69,070	70,101	1,032	69,148	70,179	1,032	99.9
89		70,269	70,754	486	70,347	70,832	486	99.6
90	*helicase*	70,738	74,133	3,396	70,816	74,211	3,396	100
91	*odv-e25*	74,245	74,886	642	74,323	74,964	642	99.8
92		74,976	75,461	486	75,054	75,539	486	99.2
93		75,525	76,280	756	75,603	76,358	756	99.3
94	*iap*	76,324	77,055	732	76,402	77,133	732	100
95	*lef-4*	77,060	78,502	1,443	77,138	78,580	1,443	99.8
96	*vp39capsid*	78,574	79,431	858	78,652	79,509	858	99.9
97	*odv-ec27*	79,573	80,439	867	79,651	80,517	867	100
98	*ptp-2*	80,753	81,238	486	80,831	81,316	486	100
99		81,329	82,513	1,185	81,407	82,591	1,185	100
100		82,545	82,871	327	82,623	82,949	327	99.8
101	*vp91capsid*	82,906	84,903	1,998	82,984	84,981	1,998	99.8
102	*tlp20*	84,884	85,534	651	84,962	85,609	648	99.5
103		85,509	86,084	576	85,584	86,159	576	100
104	*gp41*	86,110	86,979	870	86,185	87,054	870	99.9
105		87,059	87,319	261	87,134	87,394	261	100
106	*vlf-1*	87,276	88,412	1,137	87,351	88,487	1,137	99.9
107		88,508	88,762	255	88,583	88,837	255	100
108		88,874	89,320	447	88,949	89,395	447	100
109		89,415	89,990	576	89,490	90,077	588	98.4
110		90,352	90,729	378	90,394	90,801	408	100
111	*dnapol*	90,849	94,004	3,156	90,921	94,076	3,156	99.9
112	*desmoplakin*	93,946	96,102	2,157	94,018	96,171	2,154	99.8
113	*lef-3*	96,313	97,374	1,062	96,383	97,444	1,062	99.6
114		97,343	97,723	381	97,413	97,793	381	100
115		97,845	98,351	507	97,934	98,467	534	99.4
116	*iap-5*	98,533	99,360	828	98,619	99,446	828	99.8
117	*lef-9*	99,335	100,834	1,500	99,421	100,920	1,500	99.9
118	*fp*	100,870	101,355	486	100,956	101,441	486	100
119		101,468	101,956	489	101,554	102,042	489	99.6
120	*dnaligase*	102,007	103,719	1,713	102,074	103,786	1,713	99.8
121		104,019	104,234	216	104,086	104,301	216	100
122		104,328	104,528	201	104,395	104,595	201	100
123	*fgf*	104,618	105,820	1,203	104,685	105,887	1,203	99.8
124		105,973	106,272	300	106,041	106,340	300	99.3
125	*alk-exo*	106,435	107,631	1,197	106,503	107,699	1,197	99.9
126	*helicase-2*	107,555	108,928	1,374	107,623	108,996	1,374	99.8
127	*rr1*	109,017	110,891	1,875	109,085	110,959	1,875	99.8
128	*rr2a*	111,019	112,107	1,089	111,086	112,174	1,089	99.4
129/130		112,140	113,120	981	112,207	113,187	981	100
131	*lef-8*	113,198	115,819	2,622	113,265	115,886	2,622	100
132		115,907	116,302	396	115,973	116,368	396	99.2
133		116,364	116,552	189	116,430	116,618	189	99.5
134		116,542	116,943	402	116,608	117,009	402	100
135		117,012	118,142	1,131	117,078	118,208	1,131	99.9
136		118,148	118,369	222	118,214	118,435	222	100
137	*lef-10*	118,320	118,589	270	118,386	118,655	270	100
138	*vp1054*	118,447	119,445	999	118,513	119,511	999	99.7
139		119,335	119,655	321	119,401	119,721	321	99.7
140	*fgf*	119,756	120,799	1,044	119,822	120,865	1,044	99.3
141	*egt*	120,882	122,336	1,455	120,949	122,403	1,455	99.9
142		122,354	122,557	204	122,421	122,624	204	100
143	*me53*	122,530	123,441	912	122,597	123,508	912	99.9

3.3.3.2 Open reading frame variability

A total of 57 ORFs within the CpGV-SA genome are 100% identical to the corresponding ORFs in the CpGV-M1/M isolates. Two out of the 57 have slightly longer nucleotide sequence lengths, ORF35 and ORF110. The 55 ORFs that are identical to the reference in terms of nucleotide sequence and length were excluded from the amino acid sequence alignment analysis as the result would be a 100% identical match in terms of the amino acid sequence as well. The percentage identity between the amino acid sequences of the remaining ORFs of CpGV-SA and CpGV-M1/M were determined and are shown in Figure 3.7. Sixty ORFs were 99.4-100% identical to the reference ORFs with 15 genes having greater than 96% identity. Eleven ORFs in CpGV-SA showed a high degree of variation (<95%) to the corresponding ORFs in the reference isolate.

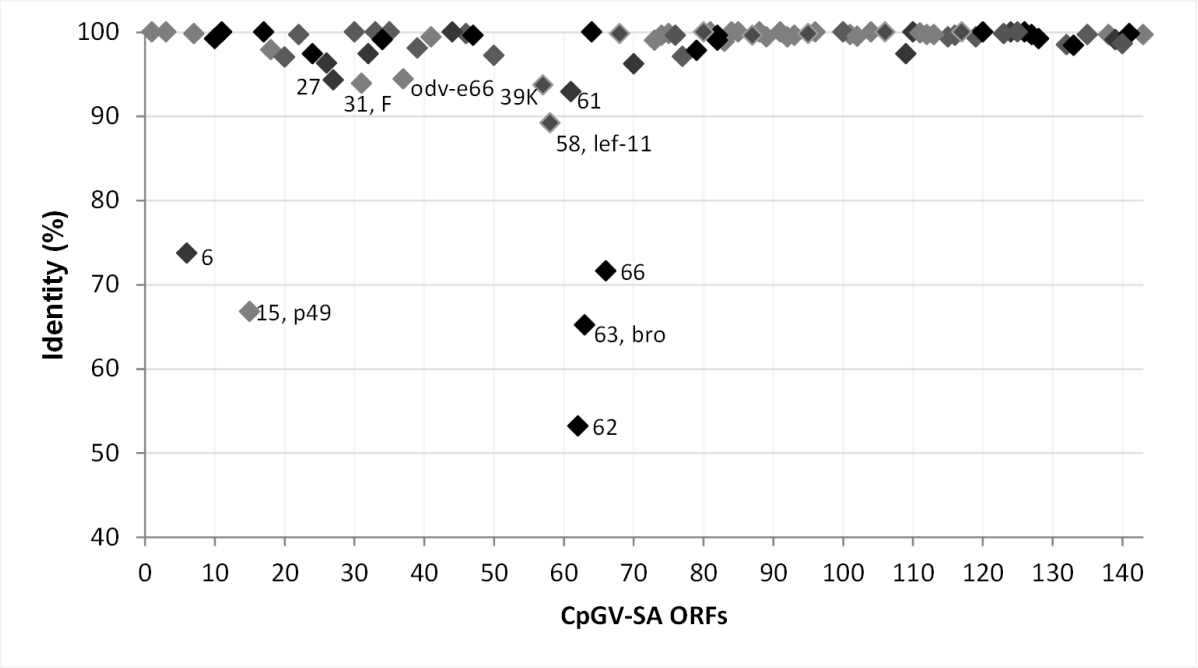


Figure 3.7: Scatter graph of the amino acid percentage identity for the ORFs in CpGV-SA to ORFs in CpGV-M1/M. Eleven outlying open reading frames are shown.

3.3.3.3 Genome organisation of CpGV-SA

The complete genome organisation of CpGV-SA is shown in Figure 3.8. The 30 core genes described by Herniou *et al*. (2003) were identified within the CpGV-SA genome, with replication genes highlighted in green, transcription genes highlighted in purple, structural genes highlighted in orange, auxiliary genes highlighted in brown and core genes with unknown functions highlighted in grey. The remaining ORFs were highlighted according to categories outlined by Luque *et al*. (2001); genes unique to the CpGV family were highlighted in red; genes specific to granuloviruses were highlighted in blue; black genes are those present in some NPVs and GVs.

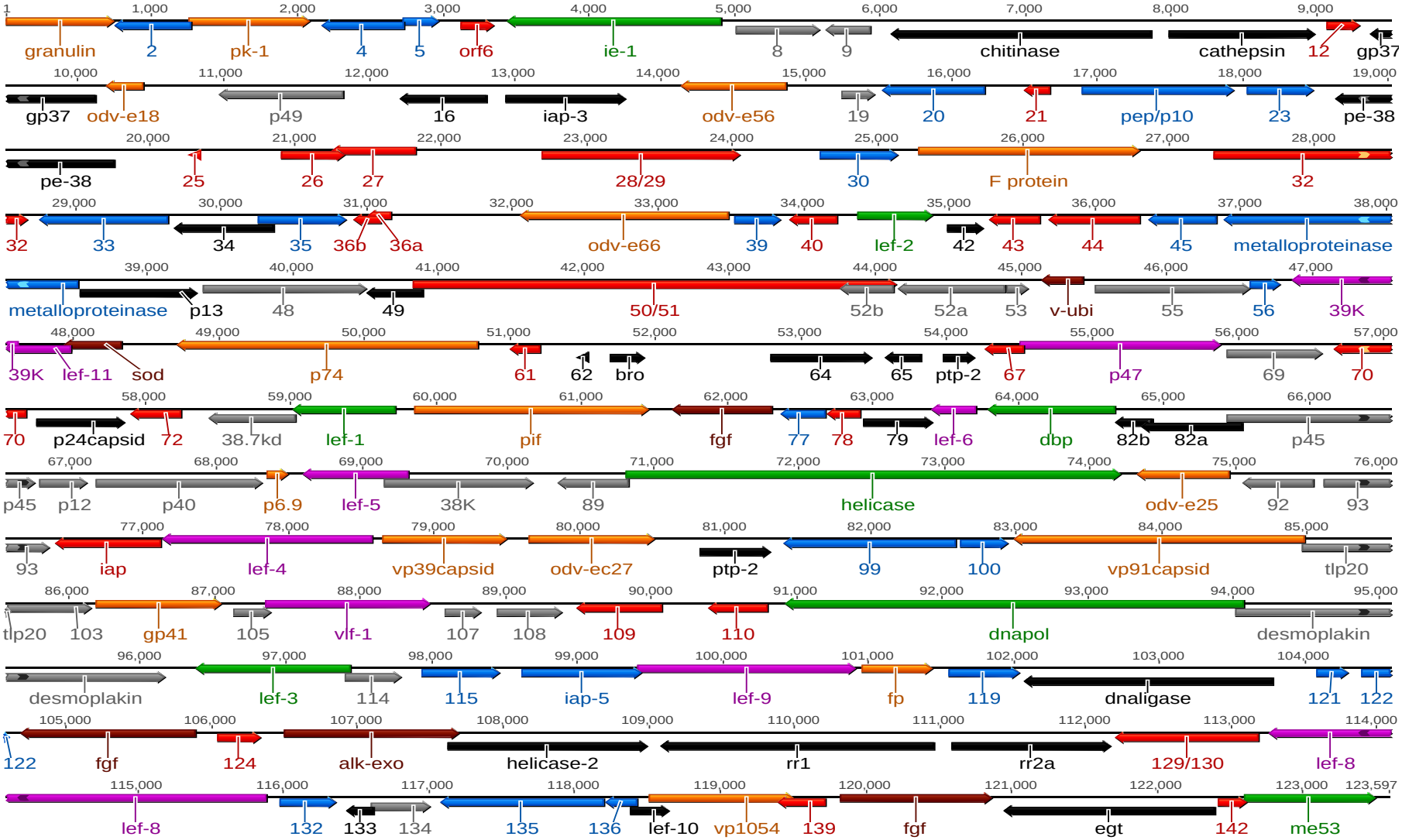


Figure 3.8: Genome organisation of the CpGV-SA isolate, with the *granulin* gene represented as the first ORF.

3.4 Discussion

The OBs were successfully purified from a commercially available formulation of CpGV-M using the glycerol gradient method. Genomic DNA was extracted from these OBs using the CTAB method which has been shown to be a reliable method for obtaining DNA from baculoviruses (Opoku-Debrah *et al.,* 2013, Abdulkadir *et al.,* 2013; Jukes *et al.,* 2014). The CpGV-M DNA served as an internal reference when conducting restriction endonuclease analysis (REN) to ensure full digestions were carried out through the use of the FastDigest™ (Thermo Scientific, USA) enzymes, as several restriction profiles for CpGV-M are available for comparison from literature (Crook *et al.,* 1985; Crook *et al.,* 1997; Eberle *et al.,* 2008; Berling *et al.,* 2009a; Eberle *et al.,* 2009). The CpGV-M isolate was also included as a reference when performing the REN analysis to verify that differences detected in the DNA profiles of the South African isolate were not due to factors such as partial digests or non-specific cleaving of the genomic DNA.

The purpose of performing REN analysis was to genetically characterise the CpGV-SA isolate and determine whether it is genetically distinct from the CpGV-M isolate. Previous studies have used the enzymes *Eco*RI, *Bam*HI, *Pst*I, *Sal*I and *Xho*I to generate profiles for the characterisation of novel CpGV isolates discovered in different geographical regions (Rezapanah *et al.,* 2008; Graillot *et al.,* 2014). Analysis of the profiles has been used as a tool to assign CpGV isolates into different genotypes or reveal isolates of mixed genotypes. The genotypic variation is a result of genome sequence alterations occurring at the different restriction sites resulting in different profiles (Eberle *et al.,* 2009; Graillot *et al.,* 2014). The CpGV-M profiles from this study were compared with those found in the literature to validate the use of the profile for comparison purposes. The CpGV-M profile matched the profiles generated by Crook *et al.,* (1997), of the *in vivo* cloned strain of the Mexican isolate (CpGV-M1) (data not shown). The only difference is that possibly due to the concentration of DNA the fragments of lower molecular weights were not visible and this also explains the smaller genome sizes calculated for both isolates across the different enzymes used. The *Bam*HI profiles of CpGV-SA and CpGV-M were different, the CpGV-M profile had two fragments that were not observed in the CpGV-SA profile and the CpGV-SA profile had a fragment that was not present in the CpGV-M profile. Another digestion that revealed that the isolate differs genetically was the *Pst*I profiles generated which showed the CpGV-SA isolate to have three fragments that are absent in the reference isolate but one fragment in

the reference absent in the CpGV-SA profile. The profiles generated with the enzymes *Eco*RI and *Xho*I were very similar to the profiles generated for CpGV-M. There was only one difference in the XhoI profile and that was the presence of an additional fragment in the CpGV-M DNA profile.

Although the comparison between restriction profiles revealed that the CpGV-SA isolate is genetically different from the CpGV-M isolate, it is still closely related to the isolate due to the observations of highly similar profiles. A more accurate method for comparing DNA profiles is required as several factors affect the quality of *in vitro* restriction digests such as the difficulty to detect low molecular weight fragments and distinguishing between fragments of similar sizes (doublet or triplet bands) as well as low gel resolution. To improve the detection of low molecular weight fragments a higher concentration of gel could be used and the amount of DNA loaded onto the gel could be increased. It is also difficult to detect the presence of mixed genotypes using *in vitro* restriction digests, if the frequency of a genotype is low resulting in restriction fragments that are below detection levels. *In silico* digestion offers an efficient and reliable method for performing restriction digests but is also dependent on the availability of complete genome sequences. The comparison of *in silico* digests of various CpGV isolates representing the different genome types is investigated in Chapter 4 in order to determine if the CpGV-SA isolate is of mixed genome types as initially proposed in Chapter 2, that the isolate in this study may potentially be of mixed genome types B and E based on single nucleotide polymorphism (SNP) detection.

To determine the genetic variations between the South African isolate and the reference isolate CpGV-M1/M, analysis of the complete genomes by determining the gene identities through nucleotide and amino acid alignments was conducted. Through the nucleotide alignments it was determined that 55 ORFs were 100% identical in nucleotide sequence and size to the respective ORFs in the reference isolate. Amino acid alignments were conducted for these ORFs but the data were not shown, as the result was an identical match to the reference isolate. The remaining ORFs were further analysed at the amino acid level and the ORFs that resulted in a 100% match showed that the variation observed in terms of nucleotide variation were silent SNPs, if the gene size was the same. All the ORFs with percentage identities of 99.4% and above contained non-silent SNPs, if the gene size was the same. Overall a total of 26 ORFs showed greater differences in the predicted amino acid sequences and of these 11 were classified as outliers. The

differences observed are due to multiple additions or deletions of nucleotides in various positions, leading to frameshifts in the sequence and amino acid changes.

In conclusion, the novel CpGV-SA isolate was characterised by REN analysis of the genomic DNA and assembly of the complete genome sequence. A comparison of the genes present in CpGV-SA and CpGV-M1/M was completed, resulting in the identification of 11 genes with high variation. Further comparative genomics were conducted in comparison to other CpGV isolates from different genome types and are discussed in Chapter 4.

Chapter 4

Comparative genomic analysis of the CpGV-SA genome with CpGV isolates of different genome types

4.1 Introduction

A comprehensive characterisation and comparison of the CpGV-SA genome by restriction endonuclease (REN) analysis and complete genome sequencing to CpGV-M1 (Luque *et al.,* 2001) and CpGV-M (Gebhardt *et al.,* 2014) was conducted in Chapter 3. It was determined that the South African isolate is genetically distinct from the CpGV-M1 and CpGV-M isolates. This observation is important as the resistance in *C. pomonella* has been shown to be highly specific to CpGV-M, which belongs to genome type A. It is possible that other genome type A isolates, G02, I66, P118, 2.17, 3.8, 6.16 and 6.9 (Table 2.5A section 2.3.4 of Chapter 2), may be associated with resistance although this has not yet been investigated and complete genomes are not available. Naturally occurring isolates E2, I12, I07 and S belonging to CpGV genome types B-E (Table 2.5A section 2.3.4 of Chapter 2), are able to infect and successfully replicate in resistant *C. pomonella* populations and are referred to as resistance-overcoming strains (Eberle *et al.,* 2008; Eberle *et al.,* 2009; Berling *et al.,* 2009b; Zichová *et al.*, 2013; Gebhardt *et al.,* 2014). The availability of complete genome sequences and comparative genomic analysis of CpGV isolates is limited to isolates M, E2, I12, I07 and S of genome types A, B, C, D and E respectively (GenBank Accession numbers: KM217573-KM217577) (Gebhardt *et al.,* 2014). Recently, the functional difference between the resistance overcoming CpGV isolates and the CpGV-M isolate was discovered and it was revealed that a single common difference in the *pe38* gene of isolates able to infect resistant *C. pomonella* could be used as a genetic marker when bioprospecting for resistance overcoming isolates (Gebhardt *et al.,* 2014). Further comparative genomics analysis is therefore necessary (i) to verify that the CpGV-SA isolate is novel and therefore genetically distinct from all the CpGV isolates available for comparison and (ii) to determine if CpGV-SA shares the common difference found in the *pe38* gene of resistance overcoming isolates.

In silico genomic DNA profiles can be performed using various software programmes, to determine genome differences and identify novel isolates, if complete genome sequences of various isolates are available. This would result in an accurate analysis based on the

detection of sequence specific cleavage sites of the restriction enzymes across the sequence data. *In silico* digestions are usually used to support physical restriction maps of extracted genomic DNA from the sequenced isolate but could be used for comparison with a reference isolate or across various isolates (Mochizuki *et al.*, 2011; Abdulkadir *et al.*, 2015; Craveiro *et al.*, 2015). A wider variety of enzymes could be selected for the DNA profiling when performing *in silico* digests which would be less costly and time-consuming to conduct in comparison with physical REN analysis. Additionally, the sizes of fragments that appear as doublets or triplet bands on agarose gels can be determined, further improving the accuracy of the analysis carried out.

Research regarding the mechanism of action in which the resistance in *C. pomonella* occurs or the functional difference between various CpGV isolates is limited. Gebhardt *et al.* (2014), through whole genome sequencing and phylogenetic analysis, discovered that the Mexican isolate, CpGV-M, has a mutation within the *pe38* gene that is not present in isolates that are considered to be infectious in *C. pomonella* resistant strains. The mutation is a repeat insertion of 24 nucleotides that results in two additional repeats of an amino acid motif. The sequence difference in this gene was further investigated and determined to be a functional difference through the use of baculovirus expression technology also known as the Bacmid system (Rohrman, 2013). When the *pe38* gene from CpGV-M was knocked out and replaced with the *pe38* gene from CpGV-S, the resistant *C. pomonella* strain was susceptible to the virus. It was therefore concluded that the resistance observed in *C. pomonella* was virus isolate-dependent and the *pe38* gene is a factor in overcoming the resistance (Gebhardt *et al.,* 2014).

The overall aim for this chapter was to perform a comparative genomic analysis of the CpGV-SA genome with complete genomes available for CpGV isolates representing the genome types A to E. The first objective was to perform and compare *in silico Bam*HI, *Eco*RI, *Pst*I and *Xho*I restriction digests of the CpGV-SA with M1, M, E2, I12, I07 and S genome sequences. Secondly, each of the complete genome sequences was aligned with the genome of CpGV-SA to determine the percentage identity across coding and non-coding regions. Lastly, a multiple alignment of the *pe38* gene of all the CpGV isolates was conducted to determine whether the South African isolate could be considered a resistance overcoming strain.

4.2 Material and Methods

4.2.1 *In silico* restriction enzyme digests

The CpGV (-M1, -M, -E2, -I12, -107, -S) complete genome sequences were downloaded from GenBank (NC_002816; KM217573-KM217577) and imported into Geneious software (New Zealand) version R7 (Kearse *et al*., 2012). *In silico* digests using the complete genome sequence of CpGV-SA and each of the other CpGV isolates' genome sequences were performed for each of the enzymes used in the REN analysis in Chapter 3, *Bam*HI, *Eco*RI, *Pst*I and *Xho*I. Virtual gels displaying the genomic DNA profiles for each of the isolates and enzymes selected were generated. The fragment size data was downloaded and analysed using Microsoft Excel® 2010. The standard deviation between the fragments found in CpGV-SA with the corresponding fragments in the other CpGV isolates was determined and the average standard deviation across all the fragments for each of the enzymes was used to determine the degree of similarity across the profiles.

4.2.2 Alignments of the CpGV genome and *pe38* gene sequences

A pairwise alignment using the progressive mauve algorithm was performed between the CpGV-SA genome sequence and the genome sequences of each of the genetically distinct CpGV isolates using Geneious software (New Zealand) version R7 (Kearse *et al*., 2012).

The *pe38* nucleotide and amino acid sequences were extracted from each of the CpGV complete genome sequences and Geneious software (New Zealand) version R7 (Kearse *et al*., 2012) was used to conduct a multiple alignment of the sequence data.

4.3 Results

4.3.1 *In silico* restriction enzyme digests

The *in silico* digests of the seven CpGV genome sequences were analysed by observing profiles generated and displayed on a virtual gel and comparing the DNA fragment sizes. Differences and similarities in the profiles generated for the CpGV-SA isolate and the profiles of the other CpGV isolates, after digestion with *Bam*HI, *Eco*RI, *Pst*I and *Xho*I, were detected. The *in silico* DNA profiles of each of the enzymes generated for CpGV (-M1, -M and -SA) matched the physical REN profiles obtained in Chapter 3, section 3.3.2.1, with the

exception of the lower sized fragments not being visible on the agarose gels due to poor resolution. The CpGV-M profiles generated *in silico* revealed few differences in comparison to the CpGV-M1 profiles although both represent the genome type A.

The virtual gel fragments (Figure 4.1 - 4.4) are labelled to the left of each isolates' profile and corresponding fragments of similar sizes shown in Tables 4.1 - 4.4 are given the same label. Regarding the fragment sizes, the similarity in fragments was displayed between the South African isolate and each of the other CpGV isolates including CpGV-M1. The calculated genomes sizes 123597, 123500, 123529, 123858, 124269,120816 and 123193 bp (Tables 4.1 - 4.4) for each of the CpGV isolates (SA, M1, M, E2, 1I2, I07 and S) respectively remained constant with the selection of different enzymes as the complete genome sequences were computationally digested.

4.3.1.1 Comparison of *in silico* CpGV *Bam*HI profiles

In Figure 4.1, the profiles displayed were based on the *Bam*HI restriction sites found in the genome sequences of genetically different CpGV isolates. The *in silico* CpGV-SA *Bam*HI profile mostly matched the CpGV-S (genome type E) profile where the only difference detected was the absence of fragment K in the virtual gel image of CpGV-SA. More differences between these two profiles were revealed by further analysing the sizes of the fragments that appeared similar in the virtual gel (Table 4.1). The remaining five CpGV isolates did not have a fragment that would match up in size to fragment C present in the CpGV-SA and -S profiles and the CpGV-I07 profile did not have a corresponding fragment for fragment A2 present in all the *Bam*HI profiles. Fragments A1, A2; E1, E2; G1, G2 and I1, I2 were doublet bands of which the sizes were determined and accounted for when calculating the genomes sizes (Table 4.1). Fragments that are produced in CpGV (-M1, -M, -E2, -I12, -I07 and -S) that did not have corresponding fragments within the CpGV-SA *Bam*HI profile are shown in red (Figure 4.1).

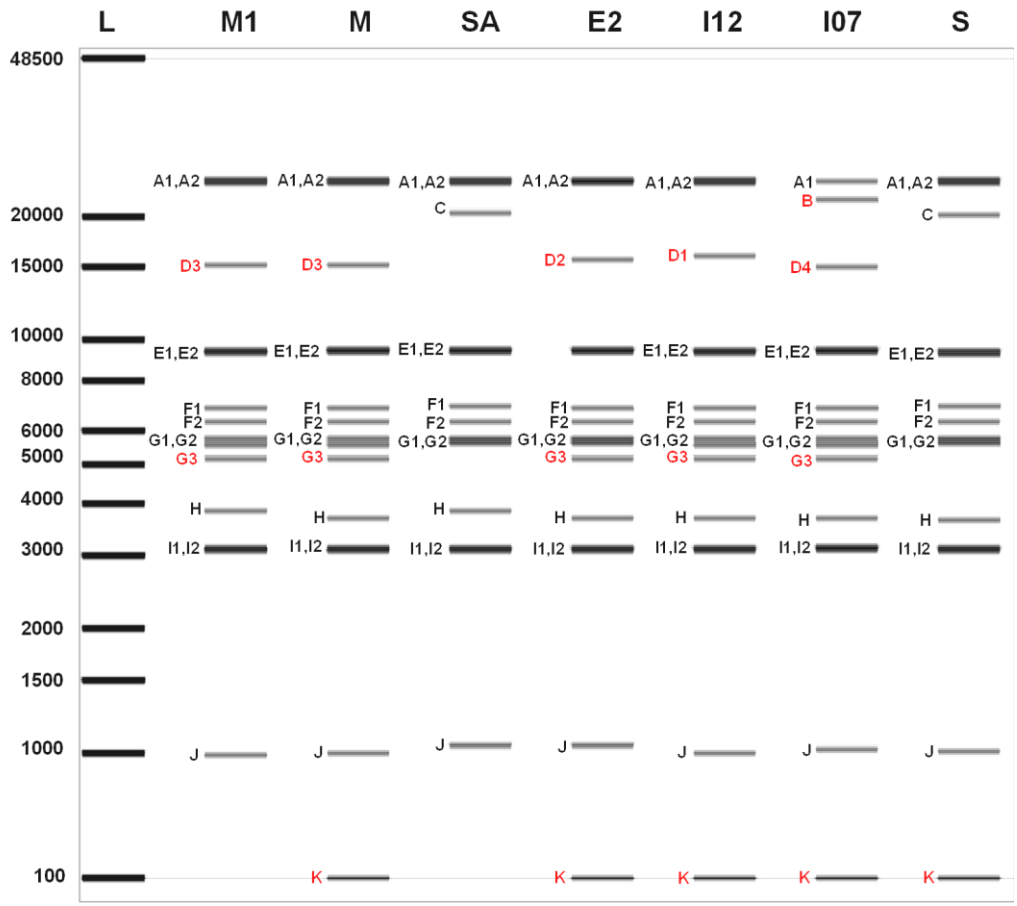


Figure 4.1: Virtue gel of *in silico* CpGV *Bam*HI profiles. L- DNA ladder; CpGV isolate is indicated above the respective genomic DNA profile. Red labels show fragments that are not present in the CpGV-SA genome.

Further comparisons were conducted regarding the fragments sizes generated by the *Bam*HI enzyme (Table 4.1). This revealed that although the virtual gel profiles of CpGV-SA and CpGV-S are highly similar, fragments A1, C, E2, F1 and H differ in size by more than 30 bp, which is the average difference in base pairs across all the *Bam*HI generated fragments of each of the CpGV isolates. Fragments that differed by less than 30 bp have a blue bar to the right of the fragment and fragments identical in size to fragments produced for the CpGV-SA isolate have a green bar to the right of the fragment. The CpGV-SA fragment G1 (5808) was identical in size with the G1 fragments of CpGV (-M1, -M, and -I12) and fragment I2 (3101) was identical in size with the I2 fragment of CpGV (-M, -E2, I12 and -S). The *in silico Bam*HI profile of CpGV-SA generated 13 fragments whereas the enzyme generated 14 fragments in isolates M1 and S and 15 fragments in isolates E2, I12 and 107.

Table 4.1: Comparison of *in silico* CpGV *Bam*HI profiles

Fragment	CpGV Isolates						
	SA	M1	M	E2	I12	I07	S
A1	24466	24501	24508	24324	24546	24202	24522
A2	24155	24162	24161	24165	24159		24157
B						21861	
C	20342						20126
D1					15976		
D2				15643			
D3		15261	15258				
D4						15098	
E1	9456	9425	9425	9451	9426	9461	9414
E2	9425	9374	9393	9427	9378	9386	9249
F1	6865	6847	6846	6840	6847	6851	6920
F2	6343	6344	6346	6331	6345	6362	6342
G1	5808	5808	5808	5787	5808	5794	5811
G2	5611	5559	5558	5597	5558	5535	5587
G3		5143	5143	5167	5143	5145	
H	3868	3866	3686	3686	3686	3688	3660
I1	3118	3121	3121	3121	3121	3121	3121
I2	3101	3100	3101	3101	3101	3113	3101
J	1039	989	995	1038	995	1019	1003
K			180	180	180	180	180
Bands	13	14	15	15	15	15	14
Size	123597	123500	123529	123858	124269	120816	123193

Fragments identical to SA (bp) | Fragments similar to SA (bp)

4.3.1.2 Comparison of *in silico* CpGV *Eco*RI profiles

The *in silico Eco*RI profiles of the CpGV isolates are shown in Figure 4.2 and the profiles were different across the genome types. The CpGV-SA profile was identical to the CpGV-M1 (genome type A) profile when observing the virtual gel image. The differences between these two profiles are shown by further analysing the sizes of the fragments that appeared similar in the gel (Table 4.2). The remaining five CpGV isolates did not have a fragment that would match up in size to fragment K3 present in the CpGV-SA and -M1 profiles. The *Eco*RI profiles of CpGV-E2 and –I12 both lacked fragment A whilst, the CpGV-I07 and –S profiles did not contain fragment D. The *in silico Eco*RI digestion of the CpGV-I07 isolate differed more from the CpGV-SA profile by the absence of fragments G1 and HI. Fragments B1, B2; H1, H2; K1, K2; K3, K4 and P1, P2 were doublet bands, whilst P1, P2, P3 was a triplet band, which were accounted for when calculating the genomes sizes (Table 4.2). Fragments that are produced in CpGV-M1, -M, -E2, -I12, -I07 and -S that did not have corresponding fragments within the *in silico* CpGV-SA *Eco*RI profile are shown in red.

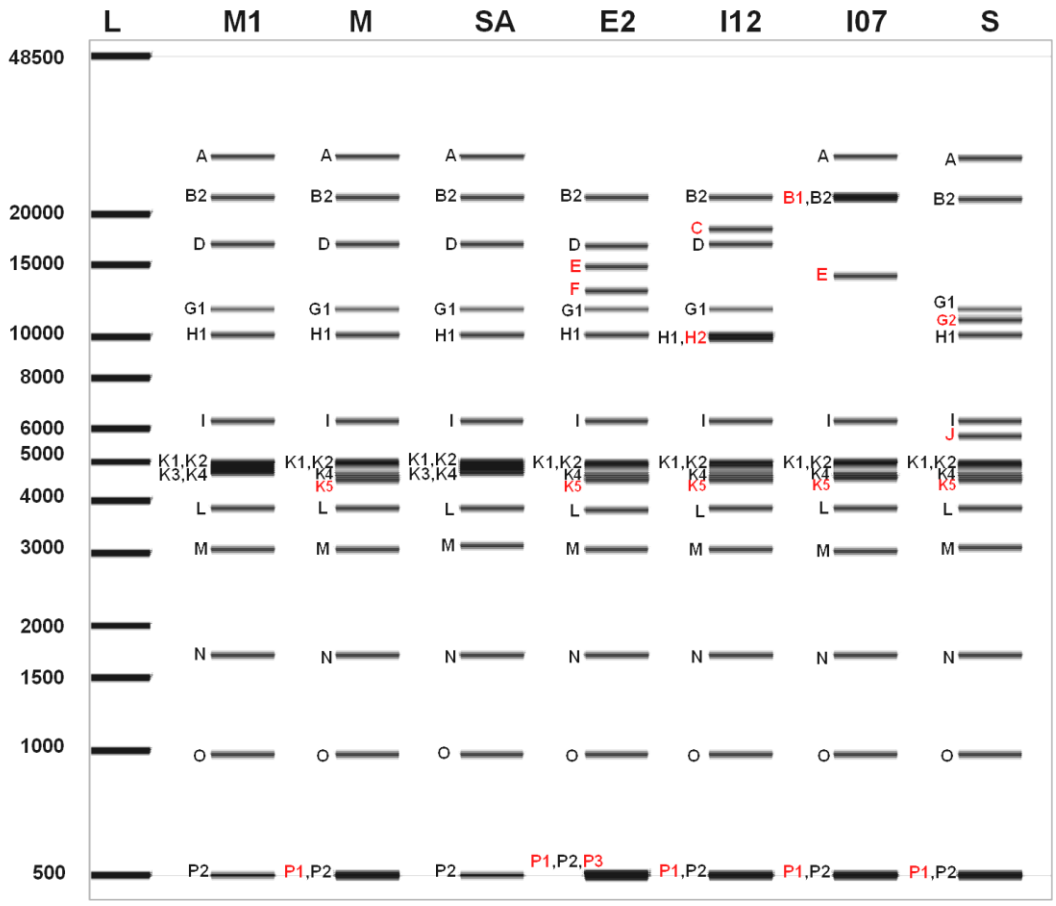


Figure 4.2: Virtue gel of *in silico* CpGV *Eco*RI profiles. L- DNA ladder; CpGV isolate is indicated above the respective genomic DNA profile. Red labels show fragments that are not present in the CpGV-SA genome.

The fragment sizes generated by the *Eco*RI enzyme (Table 4.2) revealed that the CpGV-SA profile was highly similar to CpGV-M1 with only two fragments having a difference of more than 14 bp, which is the average difference in base pairs across all the *Eco*RI generated fragments of each of the CpGV isolates. Fragments identical in size to fragments produced for the CpGV-SA profile, such as fragment P2 (786) present in all the profiles, have a green bar to the right of the fragment. Fragment O (970) was observed in all the profiles but the CpGV (-M1 and -I07) O fragment was highly similar and not identical to the rest. Fragments that differed by less than 14 bp have a blue bar to the right of the fragment. Fragments K4 (4711) and L (3843) of CpGV (-M1, -M and -I12) profiles and fragment N (1698) of the same isolates including CpGV-S were identical to the CpGV-SA fragment. Although, in Figure 4.2 the profiles of CpGV-SA and CpGV-M1 appear to be identical, fragments B2 and D differed in size by more than 14 bp (Table 4.2) but the CpGV-SA profile still remained mostly similar to the CpGV-M1 profile. The *Eco*RI profile of CpGV-SA, M1 and I07

generated 15 fragments whereas the enzyme generated 16 fragments in isolate CpGV-M, 17 fragments in isolates I12 and S, and 18 fragments in isolate E2.

Table 4.2: Comparison of *in silico* CpGV *Eco*RI profiles

Fragment	CpGV Isolates						
	SA	M1	M	E2	I12	I07	S
A	27557	27551	27553			27435	27361
B1						21932	
B2	21800	21821	21822	21829	21823	21857	21646
C					18311		
D	16859	16830	16836	16651	16874		
E				14936		14191	
F				13081			
G1	11745	11748	11747	11748	11748		11747
G2							11069
H1	10123	10124	10125	10124	10122		10124
H2					9961		
I	6262	6264	6264	6268	6264	6257	6261
J							5802
K1	4972	4973	4975	4967	4974	4991	4971
K2	4948	4913	4931	4960	4916	4921	4931
K3	4819	4817					
K4	4711	4711	4711	4708	4711	4690	4714
K5			4531	4531	4531	4533	4505
L	3843	3843	3843	3825	3843	3840	3842
M	3113	3061	3060	3060	3060	3040	3089
N	1698	1698	1698	1697	1698	1694	1698
O	970	969	970	970	970	972	970
P1			286	286	286	286	286
P2	177	177	177	177	177	177	177
P3				40			
Bands	15	15	16	18	17	15	17
Size	123597	123500	123529	123858	124269	120816	123193

Fragments identical to SA (bp) | Fragments similar to SA (bp)

4.3.1.3 Comparison of *in silico* CpGV *Pst*I profiles

The profiles shown in Figure 4.3 are based on the *Pst*I restriction sites found in the genome sequences of the genetically different CpGV isolates. The *in silico* CpGV-SA *Pst*I profile contained fragments which each corresponded to fragments in the other CpGV isolates. The CpGV-M, -E2, -I12, -I07 and -S isolates did not have a fragment that would match up in size to fragment B which was present in the CpGV-SA and -M1 profiles. Fragments D2 and K1 corresponded with fragments in the all CpGV profiles with the exception of the CpGV-S profile. Fragments F2, I2, L and P1 were present in both CpGV-SA and -S profiles but were not generated in the other profiles. The profiles for the CpGV-M1, -M, -E2, -I12 and -S

isolates lacked fragment G and isolates -M1, -M, -I12 and –I07 further lacked fragments J and S which were only produced in the profiles for CpGV-SA, -E2 and -S. Fragment I3 corresponded with fragments in the all CpGV profiles with the exception of the CpGV-E2 profile. Lastly, fragment N2 was absent in the profiles of CpGV-E2 and -S but was present in the other profiles. Although, a unique CpGV-SA *Pst*I profile was obtained and several differences were observed in comparison to each of the other profiles, analysis of the fragment sizes revealed, by the number of identical and highly similar fragments, that the *Pst*I profile of CpGV-SA is mostly similar to the CpGV-E2 (genome type B) and CpGV-S (genome type E) profiles (Table 4.3). Fragments F1, F2; I1, I2; K1, K2 and N1, N2 were doublet bands found in several of the profiles of which the sizes were accounted for when calculating genomes sizes (Table 4.3). Fragments that were produced in CpGV-M1, -M, -E2, -I12, -I07 and -S that did not have corresponding fragments within the CpGV-SA *Pst*I profile are shown in red.

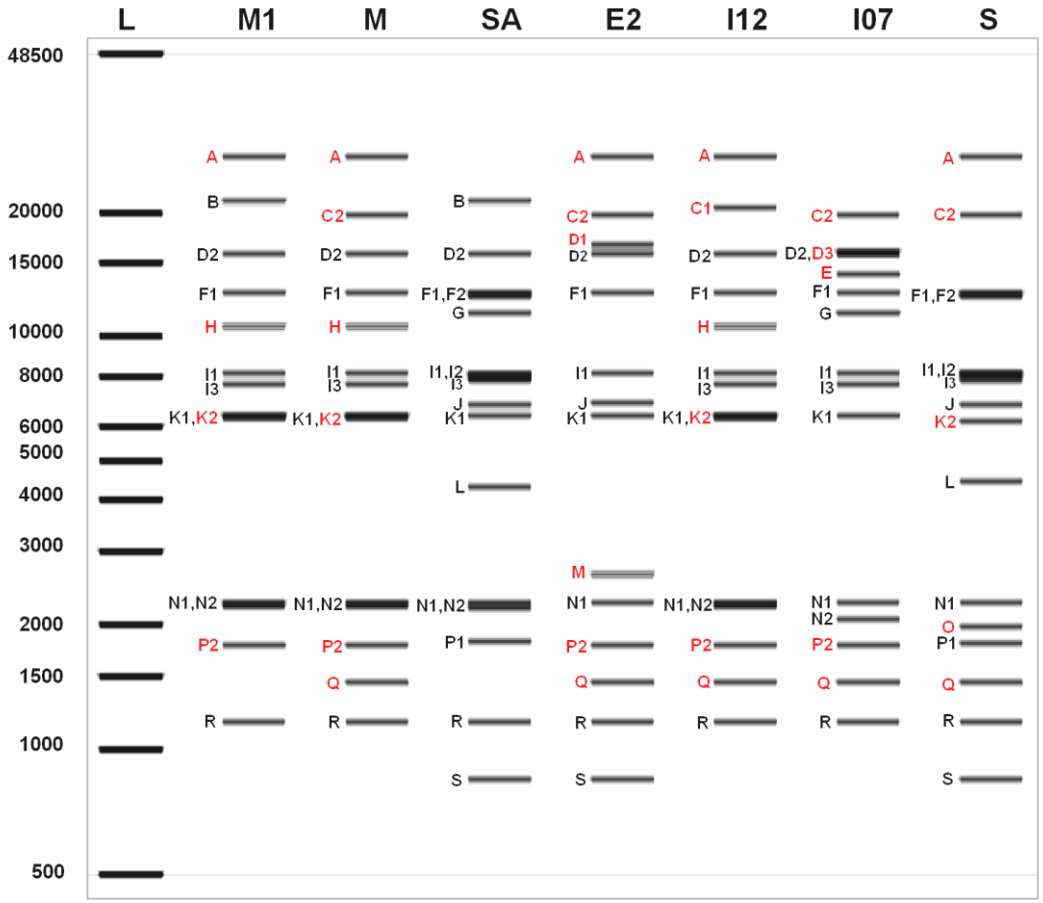


Figure 4.3: Virtue gel of *in silico* CpGV *Pst*I profiles. L- DNA ladder; CpGV isolate is indicated above the respective genomic DNA profile. Red labels show fragments that are not present in the CpGV-SA genome.

Further analysis regarding the fragment sizes generated by the *Pst*I enzyme are shown in Table 4.3. Fragments identical in size to fragments produced for the CpGV-SA profile have a green bar to the right of the fragment. Fragment S was the smallest fragment produced and was present in the CpGV-SA, E2 and S profiles at a size of 837 bp. Fragment R (1160) was observed in all the profiles but the CpGV-I07 R (1167) fragment was highly similar and not identical to the rest. Fragments that differed by less than 25 bp, which is the average difference in base pairs across all the *Pst*I generated fragments of each of the CpGV isolates, have a blue bar to the right of the fragment. Fragment N1 (2264) in all the CpGV profiles, with exception of the CpGV-S fragment N (2267), were identical to the CpGV-SA fragment. Fragment K1 (6423) was only identical between the profiles of CpGV-SA, -M1 and -M. The *Pst*I profiles of CpGV-SA and -S generated 16 fragments, whereas the enzyme generated 14 fragments for CpGV-M, -E2, -I12 and -I07 and the CpGV-M1 profile contained 13 fragments.

Table 4.3: Comparison of *in silico* CpGV *Pst*I profiles

Fragment	CpGV Isolates						
	SA	M1	M	E2	I12	I07	S
A		27231	27231	27238	27229		27204
B	21244	21205					
C1					20479		
C2			19760	19789		19816	19826
D1				16672			
D2	15914	15841	15843	15825	15842	15976	
D3						15888	
E						14215	
F1	12793	12741	12741	12784	12741	12695	12760
F2	12537						12431
G	11315					11311	
H		10518	10524		10526		
I1	8100	8118	8119	8124	8119	8118	8119
I2	8040						8039
I3	7794	7662	7680		7665	7649	7797
J	6860			6876			6843
K1	6423	6423	6423	6425	6424	6427	
K2		6333	6333		6369		6247
L	4301						4443
M				2639			
N1	2264	2264	2264	2264	2264	2264	2267
N2	2190	2222	2226		2226	2064	
O							1977
P1	1825						1801
P2		1782	1783	1783	1783	1782	
Q			1442	1442	1442	1444	1442
R	1160	1160	1160	1160	1160	1167	1160
S	837			837			837
Bands	16	13	14	14	14	14	16
Size	123597	123500	123529	123858	124269	120816	123193

Fragments identical to SA (bp) | Fragments similar to SA (bp)

4.3.1.4 Comparison of *in silico* CpGV *XhoI* profiles

In Figure 4.4, the profiles obtained from the *in silico Xho*I restriction digests of genetically different CpGV isolates are shown. The CpGV-SA profile was identical to the CpGV-M1 (genome type A) profile when observing the virtual gel image. The differences between these two profiles are revealed by further analysing the sizes of the fragments that appeared similar in the gel (Table 4.4). The remaining five CpGV isolates did not have a fragment that would match up in size to fragment A present in the CpGV-SA and -M1 profiles. The *Xho*I profiles of CpGV-I07 and -S both lacked the B2 fragment and the CpGV-I07 profile also did not contain an H2 fragment, which was also not generated in the CpGV-E2 profile. The *in silico Xho*I digestion of the CpGV-E2 isolate differs more from the CpGV-SA profile by the absence of fragments G1 and I. Fragment I present in the CpGV-SA profile was also absent in the CpGV-S profile. Fragment H1, H2 was a doublet band found in five of the profiles of which the sizes were accounted for when calculating genomes sizes (Table 4.4). Fragments that were produced in CpGV-M1, -M, -E2, -I12, -I07 and -S that did not have corresponding fragments within the CpGV-SA *Xho*I profile are shown in red.

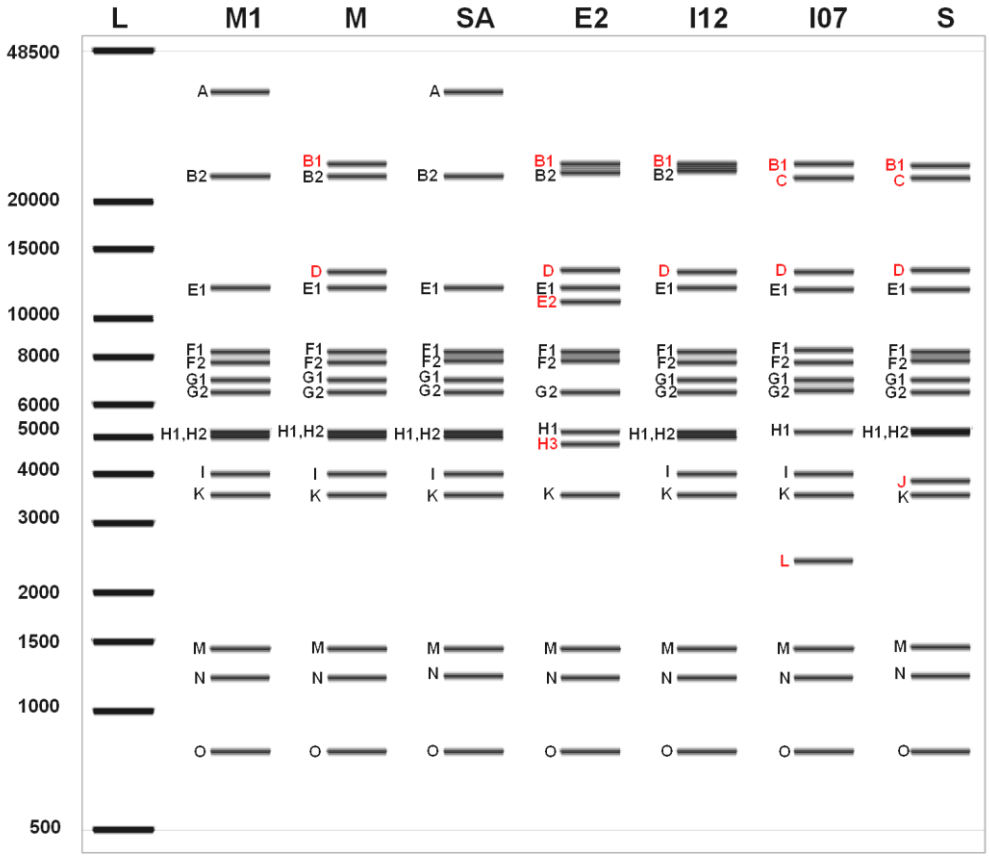


Figure 4.4: Virtue gel of *in silico* CpGV *Xho*I profiles. L- DNA ladder; CpGV isolate is indicated above the respective genomic DNA profile. Red labels show fragments that are not present in the CpGV-SA genome.

Further analysis regarding the fragment sizes generated by the *Xho*I enzyme are shown in Table 4.4. Fragments identical in size to fragments produced for the CpGV-SA profile have a green bar to the right of the fragment. Fragment O was the smallest fragment produced and was present in all the profiles at a size of 786 bp. Similarly, fragment M (1438) was observed in all the profiles but the CpGV-S M (1441) fragment was highly similar and not identical to the rest. Fragments that differed by less than 24 bp, which is the average difference in base pairs across all the *Xho*I generated fragments of each of the CpGV isolates, have a blue bar to the right of the fragment. Fragment I (4036) of CpGV-M1, M, I12 and I07 profiles were identical to the CpGV-SA fragment. Fragment G1 (6967) was only identical across the profiles of CpGV-SA, -M1, -M and -S. Although, in Figure 4.4 the profiles of CpGV-SA and CpGV-M1 appear to be identical, fragments A, E1 and F2 differ in size by more than 24 bp (Table 4.4), but the CpGV-SA profile still remained mostly similar to the CpGV-M1 profile. The *Xho*I profiles of CpGV-SA, -M1 and -E2 generated 14 fragments whereas the enzyme generated 15 fragments for the remainder of the isolates.

Table 4.4: Comparison of *in silico* CpGV *Xho*I profiles

Fragment	CpGV Isolates						
	SA	M1	M	E2	I12	I07	S
A	37931	37879					
B1			24684	24684	24682	24744	24638
B2	23013	23025	23028	23480	23746		
C						22900	22760
D			13193	13226	13194	13176	13297
E1	11935	11983	11990	11990	11992	11898	11848
E2				11010			
F1	8234	8236	8237	8237	8237	8281	8237
F2	7828	7740	7759	7799	7744	7757	7783
G1	6967	6967	6967		6968	6970	6967
G2	6510	6512	6513	6498	6512	6517	6509
H1	5172	5178	5178	5175	5178	5152	5175
H2	4989	4976	4976		5012		5132
H3				4787			
I	4036	4036	4036		4036	4036	
J							3860
K	3533	3535	3535	3539	3535	3528	3532
L						2422	
M	1438	1438	1438	1438	1438	1438	1441
N	1225	1209	1209	1209	1209	1211	1228
O	786	786	786	786	786	786	786
Bands	14	14	15	14	15	15	15
Size	123597	123500	123529	123858	124269	120816	123193

Fragments identical to SA (bp) | Fragments similar to SA (bp)

4.3.2 Alignments of the CpGV genome sequences and the *pe38* gene

4.3.2.1 Complete genome sequence alignments

Alignments of the complete genome sequence, which included both coding and non-coding regions, for CpGV-SA against the full genome sequences of CpGV-M1, -M, -E2, -I12, I07 and S showed a pairwise identity of 97.2, 97.5, 96.5, 96.9, 94.2 and 97.9% respectively with each of the isolates.

4.3.2.2 Single nucleotide polymorphisms in the *pe38* gene

A multiple alignment of the *pe38* sequence data extracted from the complete genome sequences of genetically different CpGV isolates available on the GenBank database (NC_002816; KM217573-KM217577) and that of CpGV-SA revealed several sets of single nucleotide polymorphisms (SNPs) across the isolates (Table 4.5). The SNPs were observed in comparison to the reference isolate genome, CpGV-M1, and the nucleotide positions are therefore numbered according to positions in this isolate (Luque *et al.,* 2001). The nucleotides highlighted in red are an indication of the SNP and where an amino acid change was observed, the amino acid is shown in red letters. The *pe38* gene of CpGV-M (Gebhardt *et al.,* 2014) was 100% identical to the respective gene in the CpGV-M1 (Luque *et al.,* 2001) genome and therefore no SNPs were detected. The *pe38* genes of CpGV-SA and CpGV-S were 100% identical, but both varied from the Mexican isolates, as four SNPs were detected, two of which were synonymous and two of which were non-synonymous, resulting in amino acid changes. The SNPs detected in the respective gene of the CpGV-E2 isolate were identical to those found in CpGV-SA and -S. However an additional non-synonymous SNP was also detected. Three synonymous and two non-synonymous SNPs were detected in the *pe38* gene of CpGV-I12. The *pe38* gene of CpGV-I07 contained five SNPs in comparison to the reference isolate, CpGV-M1; two of these were synonymous and three were non-synonymous.

Table 4.5: Single nucleotide polymorphisms (SNPs) found in the *pe38* gene after multiple alignments with the genetically distinct CpGV isolates in comparison to the reference isolate, CpGV-M1

Isolate	Nucleotide position/codon										
	19543	**19347**	**19137**	**19062**	**18901**	**18708**	**18703**	**18686**	**18677**	**18660**	**18639**
M1	TGT	GAG	TGT	AGT	TCT	CCG	TCA	ATT	AGG	AAG	CAG
Amino acid	Cysteine	Glutamic acid	Cysteine	Serine	Alanine	Proline	Serine	Isoleucine	Arginine	Lysine	Glutamine
M	TGT	GAG	TGT	AGT	TCT	CCG	TCA	ATT	AGG	AAG	CAG
Amino acid	Cysteine	Glutamic acid	Cysteine	Serine	Alanine	Proline	Serine	Isoleucine	Arginine	Lysine	Glutamine
SA	TGT	GAG	TGT	AGT	TC**C**	CCG	TC**C**	ATT	AGG	**C**AG	**A**AG
Amino acid	Cysteine	Glutamic acid	Cysteine	Serine	Alanine	Proline	Serine	Isoleucine	Arginine	Glutamine	Lysine
E2	TGT	**C**AG	TGT	AGT	TC**C**	CCG	TC**C**	ATT	AGG	**C**AG	**A**AG
Amino acid	Cysteine	Glutamine	Cysteine	Serine	Alanine	Proline	Serine	Isoleucine	Arginine	Glutamine	Lysine
I12	TG**C**	GAG	TGT	AGT	TC**C**	CCG	TC**C**	A**A**T	A**A**G	AAG	CAG
Amino acid	Cysteine	Glutamic acid	Cysteine	Serine	Alanine	Proline	Serine	Asparagine	Lysine	Lysine	Glutamine
I07	TGT	GAG	**C**GT	**G**GT	TC**C**	**T**CG	TC**C**	ATT	AGG	AAG	CAG
Amino acid	Cysteine	Glutamic acid	Arginine	Glycine	Alanine	Serine	Serine	Isoleucine	Arginine	Lysine	Glutamine
S	TGT	GAG	TGT	AGT	TC**C**	CCG	TC**C**	ATT	AGG	**C**AG	**A**AG
Amino acid	Cysteine	Glutamic acid	Cysteine	Serine	Alanine	Proline	Serine	Isoleucine	Arginine	Glutamine	Lysine

4.3.2.3 Mutation in the *pe38* gene

In addition to the SNPs detected, according to Gebhardt *et al.* (2014), there is a repeat of 24 nucleotides within the *pe38* gene of the Mexican isolate that is absent in the other isolates and that have been shown to overcome resistance, revealing a single common difference for the screening of resistance overcoming isolates. Table 4.6 shows the alignment results of the mutated region of the *pe38* gene of the Mexican isolates in comparison to the other CpGV isolates and the isolate from this study, CpGV-SA. The *pe38* gene can be found in the reverse orientation for all the CpGV isolates and the length of the gene in CpGV-M1 and -M is 1149 bp but is 24 bp shorter for all the other CpGV isolates (1125 bp). Both Mexican isolates contained the 24 nt insertion between positions 18789 and 18759, which coded for the amino acid motif D- Aspartic acid, T- Threonine, V- Valine and D- Aspartic acid repeated three times in the sequence, as described by Gebhardt *et al.* (2014) (Table 4.6). This region of the *pe38* gene of CpGV-SA was found to match the other isolates, as the gene sequence did not contain the 24 nucleotide repeats present in CpGV-M1/M and therefore the amino acid motif only appeared once in the sequence.

Table 4.6: Multiple alignment of the single common difference detected in the *pe38* gene of CpGV resistance overcoming isolates. The roman numbers shown in red, highlight the amino acid motif repeated three times in the Mexican isolates and once in the other isolates.

Isolate	Codon position													
	I				II				III					
	18789	**18786**	**18783**	**18780**	**18777**	**18774**	**18771**	**18768**	**18765**	**18762**	**18759**	**18756**	**18753**	**18750**
M1	GAC	ACA	GTG	GAT	GAC	ACA	GTG	GAT	GAC	ACA	GTG	GAT	GAC	ACA
Amino acid	Aspartic acid	Threonine	Valine	Aspartic acid	Aspartic acid	Threonine	Valine	Aspartic acid	Aspartic acid	Threonine	Valine	Aspartic acid	Aspartic acid	Threonine
M	GAC	ACA	GTG	GAT	GAC	ACA	GTG	GAT	GAC	ACA	GTG	GAT	GAC	ACA
Amino acid	Aspartic acid	Threonine	Valine	Aspartic acid	Aspartic acid	Threonine	Valine	Aspartic acid	Aspartic acid	Threonine	Valine	Aspartic acid	Aspartic acid	Threonine
SA	GAC	ACA	GTG	GAT	-	-	-	-	-	-	-	-	GAC	ACA
Amino acid	Aspartic acid	Threonine	Valine	Aspartic acid									Aspartic acid	Threonine
E2	GAC	ACA	GTG	GAT	-	-	-	-	-	-	-	-	GAC	ACA
Amino acid	Aspartic acid	Threonine	Valine	Aspartic acid									Aspartic acid	Threonine
I12	GAC	ACA	GTG	GAT	-	-	-	-	-	-	-	-	GAC	ACA
Amino acid	Aspartic acid	Threonine	Valine	Aspartic acid									Aspartic acid	Threonine
I07	GAC	ACA	GTG	GAT	-	-	-	-	-	-	-	-	GAC	ACA
Amino acid	Aspartic acid	Threonine	Valine	Aspartic acid									Aspartic acid	Threonine
S	GAC	ACA	GTG	GAT	-	-	-	-	-	-	-	-	GAC	ACA
Amino acid	Aspartic acid	Threonine	Valine	Aspartic acid									Aspartic acid	Threonine

4.4 Discussion

To verify that the CpGV-SA isolate is novel, the genomes of CpGV isolates representing each of the genome types A-E were subjected to *in silico* restriction enzyme digestion for comparison with the *in silico* profiles of the CpGV-SA genome sequence. The *Bam*HI, *Eco*RI, *Pst*I and *Xho*I restriction enzymes were chosen for the *in silico* analysis as the physical restriction endonuclease (REN) analysis of CpGV-SA and CpGV-M (extracted from Carpovirusine®) was performed using these enzymes (Chapter 3). The *in silico* digestions proved to be a more accurate analysis, as the sizes of the fragments in the doublet or triplet bands and lower sized fragments could be determined. However, the physical REN profiles could also be validated using the *in silico* profiles therefore supporting the use of both methods. The CpGV-SA strain was isolated from a field population of *C. pomonella* where no biopesticides including CpGV-M (Madex® or Carpovirusine®) or CpGV-S (Virosoft®) based biopesticides had ever been applied. The *in silico* digestion profiles of CpGV-SA revealed that the isolate is novel and is comprised of a mixture of three genotypes A (CpGV-M1/M), B (CpGV-E2) and E (CpGV-S) displayed through the profiles generated by enzymes *Eco*RI/*Xho*I, *Pst*I and *Bam*HI respectively. The profiles of CpGV-M1 (Luque *et al.,* 2001) and CpGV-M (Gebhardt *et al.*, 2014) varied due to the additional 29 bp in the CpGV-M isolate and possibly the minor sequence variations (Chapter 3) however both are classified as belonging to genome type A. It is not rare for CpGV isolates to be constituted of mixed genotypes and it has been suggested that it could be a result of various genetically different CpGV isolates interacting in a host (Rezapanah *et al.,* 2008; Berling *et al.*, 2009b). Through the classification of isolates from different geographical locations into the genome types A-D, it has also been shown that the distinct genotypes are not limited to one geographical location and within one geographical location various genotypes can exist (Eberle *et al.*, 2009).

Pairwise alignments of the available complete genomes of genetically different CpGV isolates in comparison to CpGV-SA genome sequence were conducted and it was found that the genome of CpGV-S had the highest percentage identity. This result may be an indication as to which genome type the CpGV-SA strain predominantly belongs to, although factors such as the genome lengths and whether the variations are in the coding or non-coding regions would need to be taken into consideration for further investigations. However, it has been shown in Chapter 2 that the *lef-9* gene of CpGV-S is 100% identical

to that of CpGV-SA and it was determined in this chapter that the *pe38* gene is also 100% identical to that of CpGV-SA. The phylogenetic analysis of the CpGV isolates showed that the CpGV-SA isolate groups closely with CpGV-E2 and CpGV-S isolates. This grouping was based on the SNPs detected in the highly conserved genes *granulin, lef-8* and *lef-9*. The SNPs detected in the *pe38* gene shows a grouping of the CpGV-SA isolate with the genome types B and E, although the pairwise genome alignment of CpGV-E2 with CpGV-SA showed the second lowest percentage identity. It is evident from this that the CpGV-SA isolate is predominantly of genome type E but contains features of the genome types B. The CpGV-E2 isolate is an *in vivo* cloned strain of an isolate (CpGV-E) originating from England. The CpGV-E strain existed in a mixture of two genotypes, one of which (CpGV-E1) was identical to CpGV-M1 and the other (CpGV-E2) had minor differences to the CpGV-M1 (Crook *et al.,* 1985). CpGV-E2 was therefore assigned to genome type B but the isolate contains features of genome type A (Eberle *et al.,* 2009; Gebhardt *et al.,* 2014). It was shown in Chapter 3 that the CpGV-SA isolate is genetically different from the Mexican isolates, although there were similarities in the physical REN profiles.

Multiple alignment of the *pe38* gene of all the CpGV isolates was conducted to determine whether the South African isolate could be considered a resistance overcoming strain. The *pe38* sequence of CpGV-SA matched the sequences of the resistance overcoming isolates in the region where a single common difference to the Mexican isolates was detected. Gebhardt *et al*. (2014) discovered that the resistance observed in *C. pomonella* to the granulovirus is exclusive to the CpGV-M isolate due to the mutation in the *pe38* gene of this isolate. The *pe38* gene was also shown to be required for the infection cycle to take place, as it codes for a 38-kDa nuclear protein, which is one of the proteins that has a role in trans-activating early gene expression (Krappa & Knebel-Mörsdorf, 1991).

In conclusion the CpGV-SA isolate was shown to be predominantly of genome type E but displayed a genome with a genotype mixture of genome types A, B and E. It was also determined to be an isolate that would be able to overcome the resistance observed in *C. pomonella* populations and could therefore be developed into a biopesticide for use in South Africa, in the case that resistance to the currently used CpGV-M based biopesticides occurs. The biological activity of the CpGV-SA isolate in comparison to the activity of the CpGV-M isolate was investigated and is discussed in the following chapter (Chapter 5).

Chapter 5

Determining the biological activity of CpGV-SA against *C. pomonella* neonate larvae

5.1 Introduction

In the previous chapter, the complete genome sequence of CpGV-SA was compared by *in silico* restriction digests to the available complete genome sequences of CpGV isolates, each of which represented one of the genome types, A-E. Furthermore, the *pe38* genes of these isolates were aligned and it was revealed that the CpGV-SA isolate did not contain the mutation found in CpGV-M, which has been investigated as a cause in the occurrence of resistance in *C. pomonella* populations (Gebhardt *et al.,* 2014). Overall, the genetic analysis conducted in Chapters 2-4 confirmed that CpGV-SA is a novel CpGV isolate and that it would be able to overcome *C. pomonella* resistance to CpGV-M based biopesticides. This chapter describes the evaluation of the biological activity and therefore virulence of CpGV-SA against *C. pomonella* neonate larvae in comparison to CpGV-M, extracted from Carpovirusine® (Arysta Lifescience, France).

Biological assays are conducted in order to obtain information regarding the virulence of a pathogen against its host (Shapiro-Ilan *et al.*, 2005). Knowledge of the biological activity would aid in developing formulations and provide information about the application of a virus if it were to be developed into a biopesticide. Bioassays can be conducted using various inoculation methods such as (i) droplet feeding where a droplet of virus suspension is mixed with food colouring, (ii) surface contamination (used in this study), which involves adding a known volume of virus suspension to the surface of the diet, (iii) the diet plug method where one small piece of artificial diet containing a virus dose is ingested by each larva and (iv) the diet incorporation method which involves mixing a known concentration of the virus suspension with artificial diet (Lacey, 2012). Factors to consider when deciding which inoculation method is the most suitable for the host population tested include the instar of the larvae, the virus-host interaction, feeding habits of the larvae and the availability of the virus stock (Lacey, 2012). The virulence of virus isolates is measured by determining the concentration-mortality and time-mortality response relationships (Shapiro-Ilan *et al.,* 2005). The concentration-mortality response relationship provides information on the lethal concentrations of the virus with the ability to kill 50% and 90% of the host

population, referred to as LC_{50} and LC_{90} values respectively. The time-mortality response relationship provides a LT_{50} value, which is the median lethal time revealing the speed of kill of the virus (Hughes *et al.*, 1986; Hughes and Shapiro, 1997; Sporleder *et al.*, 2005). An important aspect of conducting bioassays is that the virulence between different isolates can be compared.

CpGV and in particular the Mexican isolate used for the production of most CpGV-based biopesticides, is regarded as a highly efficient control agent for *C. pomonella* due to its high virulence and host specificity (Lacey *et al.,* 2008). However, due to reported resistance cases, several geographically different isolates have been characterised and investigated. It has been shown that genetically distinct CpGV isolates have different degrees of virulence when tested against specific *C. pomonella* populations. The virulence of these isolates not only differs between the isolates but also across various *C. pomonella* populations (Eberle *et al.,* 2009, Gund *et al.*, 2012; Gebhardt *et al.,* 2014). The reason for the difference in virulence recorded is not well understood and therefore novel isolates need to be tested and compared with existing isolates against the same host population in order to determine the relative biological activity of the virus.

In South Africa, Madex® (Andermatt Biocontrol, Switzerland) and Carpovirusine® (Arysta Lifescience, France), which are both formulated with CpGV-M, are currently the only granulovirus based biopesticides used to control *C. pomonella* populations. No resistance in insect populations against these biopesticides in South African orchards has been reported to date. However, it is important to fully characterise novel isolates both genetically and in terms of the biological activity in order to manage resistance should it occur. Investigations regarding the biological activity of a novel isolate would also aid in developing the virus into a biopesticide. In this study, CpGV-SA was tested against an insect population originating from the Western Cape Province of South Africa, as this was the only established *C. pomonella* culture available. An insect population from the Free State Province, from which CpGV-SA was isolated, was not available. Therefore, the biological assays were conducted alongside CpGV-M, extracted from Carpovirusine, for comparative purposes.

The aim of this chapter was therefore to determine the biological activity of CpGV-SA and compare it to that of the CpGV-M isolate from Carpovirusine®. The specific objectives included determining CpGV stock concentrations by light microscopy enumeration and then

performing a surface concentration-mortality response bioassay in order to determine the LC_{50} and LC_{90} concentrations of CpGV-SA against neonate larvae. Another objective was to then use the LC_{90} concentration to perform surface time-mortality response bioassays on neonate larvae. Lastly, PCR amplification of the *egt* gene from DNA extracted from dead larvae collected from the concentration-mortality response bioassays was conducted to confirm that mortality was a result of CpGV-SA and CpGV-M infection.

5.2 Material and methods

5.2.1 Occlusion body enumeration

The concentrations of CpGV-SA and CpGV-M occlusion body (OB) extracts supplied by River Bioscience (Pty) Ltd., South Africa (Section 2.3.1 of Chapter 2) and extracted from Carpovirusine® (Section 3.3.1 of Chapter 3) respectively, were determined using light microscopy. The virus stock samples were vortexed prior to preparing 1:5 dilutions with ddH_2O, which were used to make successive 1:5 dilutions with 0.07% SDS (w/v) in 1.5 ml tubes. The resultant 1:25 virus dilutions were mixed then sonicated for 60 seconds at 60 Hz with a Vibra Cell (Sonics and Materials, USA). Next, 1:20 dilutions were prepared by mixing the sonicated solution with ddH_2O in new 2 ml tubes.

The OB enumeration was performed using a standard method described by Hunter-Fujita *et al.* (1998) and Jones (2000). A Thoma bacterial counting chamber (Hawksley®, UK) with a depth of 0.02 ml was cleaned using 70% ethanol (v/v) lens cleaning tissue and inspected under a light microscope set to dark field illumination. A coverslip was partially placed over the chamber and 5 µl of the CpGV-SA 1:20 diluted virus solution was loaded into the chamber via capillary action. The coverslip was cautiously slid to fully cover the chamber and the slide was left to stand for 5 min to allow non-virus particles to settle. Moving particles viewed at 400 x magnification were counted in the large squares in top left, top right, bottom left, bottom right and one random square from the centre using a hand held tally counter. The procedure was repeated using the same sample two more times and repeated using the CpGV-M 1:20 diluted virus suspension, with cleaning and inspection of the slide in between sample counts. The mean number of virus particles was used to calculate the concentration of the virus using the following formula:

$$\text{OBs/ml} = \frac{(\text{D x X})}{(\text{N x V})}$$

D = total dilution factor, X = average number of OBs counted, N = number of small squares and V = volume capacity of small square. The number of small squares counted was 80 (16 per large square) and the volume of each small square was 0.00005 µl.

5.2.2 Preparation of dilutions for the neonate bioassays

In two separate 1.5 ml tubes 1:100 dilutions of both virus stocks (described in Section 5.2.1) were prepared with ddH_2O. These dilutions were used to make successive 1:70 and 1:50 dilutions of CpGV-SA and CpGV-M suspensions respectively. The resultant 7000 and 5000 fold dilutions were made to lower the concentrations of the virus stocks to similar concentrations of 1.18 x 10^6 and 1.20 x 10^6 OBs/ml, labelled SA^A and M^A. These virus suspensions were used to make six seven-fold serial dilutions to a total volume of 4900 µl, labelled SA $^{(-D1,-D2,\ -D3,\ -D4,\ -D5,\ -D6)}$ and M $^{(-D1,-D2,\ -D3,\ -D4,\ -D5,\ -D6)}$ (Figure 5.1). These dilutions were prepared by adding 4200 µl of ddH_2O into sterile bottles, then vortexing virus suspension A prior to pipetting 700 µl into bottle D1. The mixture was vortexed to achieve homogeneity and 700 µl was transferred into bottle D2. This procedure was repeated subsequently for D3, D4, D5 and D6. The dilutions for the CpGV-SA and CpGV-M virus suspensions were prepared separately to avoid cross contamination and were made in preparation for neonate concentration-mortality response bioassays.

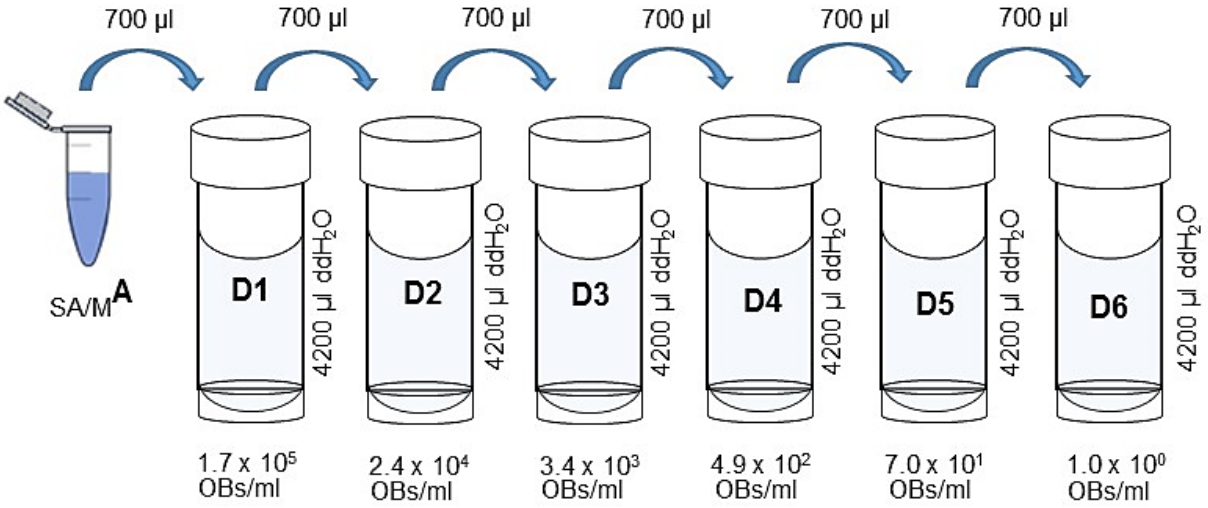


Figure 5.1: Diagram of CpGV-SA and CpGV-M six seven-fold dilutions prepared from 1.18 x 10^6 (SA^A) and 1.20 x 10^6 (M^A) OBs/ml virus stock suspensions. D1 - dilution 1, D2 - dilution 2, D3 - dilution 3, D4 - dilution 4, D5 - dilution 5, D6 - dilution 6.

5.2.3 Surface concentration-mortality response bioassay with neonate larvae

Surface-treated bioassays for CpGV-SA were conducted alongside CpGV-M (extracted from Carpovirusine® stored at -20°C and used within the products shelf life) in 24-well plates, using modified methods described by Chambers (2014). Pre-mixed artificial diet, supplied by River Bioscience (Pty) Ltd., South Africa, was prepared by first sterilising 235 g of the diet in an oven at 180°C for 10 min. 0.1 M propionic and 0.1 M phosphoric acid were added to the diet prior to the addition of 400 ml of an agar solution consisting of 13 g of agar. Water was added to the mixture for consistency before pouring the mixture into 24-well trays and allowing the diet to cool (Figure 5.2). *Cydia pomonella* egg sheets were obtained from ENTOMON Technologies Pty Ltd, Stellenbosch, South Africa, which rears *C. pomonella* for a sterile insect release programme, and the neonate larvae were used for the bioassays. When the *C. pomonella* larvae hatched, 50 µl of the virus dilutions prepared in section 5.2.2 was added to cover the entire surface of the diet in each well and allowed to dry for 30 min. Twenty four larvae were treated per dilution and the control plate was treated with 50 µl of ddH_20. One neonate larva was placed into each well using a 000 paint brush before the trays were sealed, wrapped in layers of paper towel and incubated in a controlled environment room set at 27°C with a photoperiod of 16:8 hours (L:D). After eight days the trays were inspected and the number of live larvae was recorded. The concentration-mortality response bioassays were replicated independently three times.

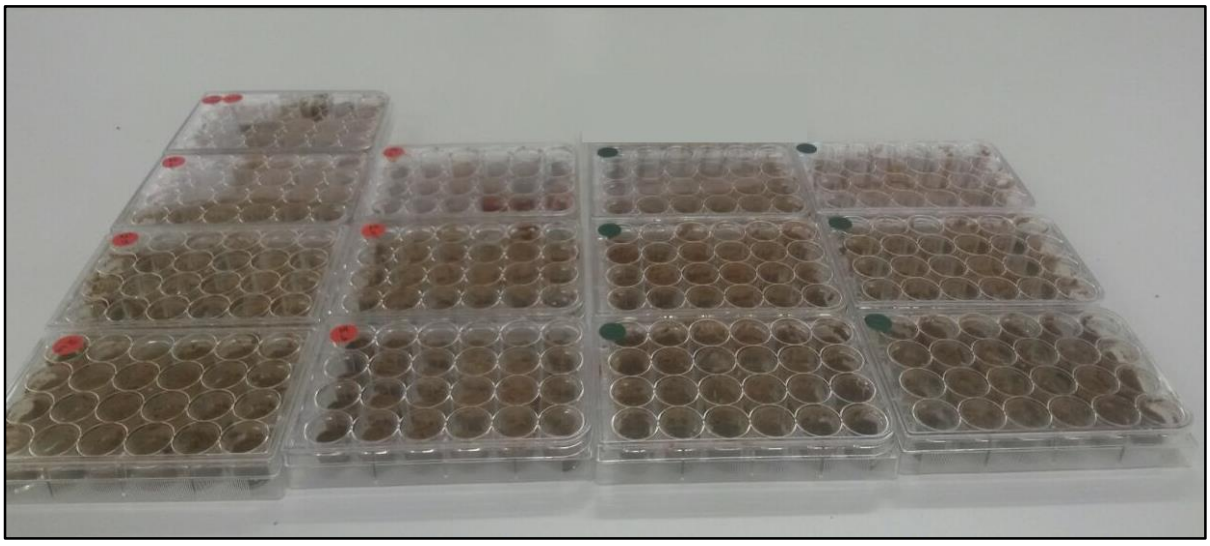


Figure 5.2: 24-well bioassay trays with artificial diet poured into each well. Six trays for CpGV-SA (red) and six trays for CpGV-M (green) were prepared. One control (2 red stickers) inoculated with ddH_2O was included with each replicate.

5.2.4 Time-mortality response bioassay with neonate larvae

Time-mortality response bioassays were conducted in 22 ml glass vials instead of the 24-well trays, in order to inspect the larvae showing symptoms of viral infection with minimal disturbances and contamination (Figure 5.3). The diet was prepared as described in section 5.2.3 and the control assay was set up using 24 glass vials inoculated with 100 µl of ddH_2O, upon the hatching of the larvae. A set of 24 vials were inoculated with 100 µl of the LC_{90} concentration of virus, calculated from the concentration-mortality response bioassays, for each of the CpGV isolates. The virus suspensions were left to dry for 30 min prior to 1 neonate larvae been placed into each vial. The vials were then sealed with cotton wool stoppers and incubated at 22°C and at a photoperiod of 16:8 hours (L:D). After 16 hours, the vials were inspected one by one for any dead larvae, which were then removed from the bioassay. Thereafter, the vials were inspected every eight hours until all the test treatment larvae had been removed. The time-mortality response bioassays were replicated independently three times.

Figure 5.3: Time-mortality response bioassay set up: each replicate included 24 vials for the CpGV-SA treatment, 24 vials for the CpGV-M treatment and 24 vials as a control (inoculated with ddH_2O).

5.2.5 Statistical data analysis

The data obtained from the concentration-mortality response assays were subjected to probit analysis using Proban statistical software (Van Ark, 1995). The concentrations of each of the dilutions tested were transformed to $\log_{10}$ and the percentage mortality to empirical probits. Percentage mortality was also adjusted according to the control mortality using Abbott's formula (Abbott, 1925):

$$Corrected\ mortality\ (\%) = \frac{Treatment\ mortality\ (\%) - Control\ mortality\ (\%)}{100 - Contol\ mortality\ (\%)} \times 100$$

The regression lines of the three replicates for each CpGV assay were determined and for each replicate a Chi-squared (X^2) test was conducted to test the fit of the line. Furthermore, the regression lines across the three replicates for each CpGV isolate were compared by obtaining X^2 values for Bartlett's test for homogeneity of residual variances and for testing the parallelism of the slopes. The comprehensive comparison of the regression lines was performed to determine whether the data obtained from the replicates could be compared before pooling the replicate values for the calculation of LC_{50} and LC_{90} values.

The time-mortality response relationship was determined using a logistic version (logit) of a probit analysis performed using Statistica version V12 software (StatSoft, 2013) (Bliss & Stevens, 2008). The data obtained from the three replicates on a total of 72 larvae for each replicate were used to determine the median lethal time (LT_{50}) taking into account the control mortality.

5.2.6 PCR amplification of the *egt* gene

To confirm that mortality observed from the concentration-mortality response bioassays was a result of CpGV-SA and CpGV-M infections, OBs were extracted from dead larvae collected from each treatment in all three replicates for each of the CpGV assays. The larvae were macerated in 500 µl 0.1% SDS (w/v) in a 2 ml tube before centrifugation at 400 ×*g* for 30 seconds. The supernatant was collected in a new 2 ml tube and the pellet suspended in 500 µl 0.1% SDS (w/v) and centrifuged at 400 ×*g* for 30 seconds. The supernatant was collected and combined with the previously collected supernatant. The pellet was suspended and centrifuged a further two times with the supernatant collected and combined each time. The combined supernatant was centrifuged at 10000 ×*g* for 30 min. The resultant supernatant was discarded and the pellet suspended in 100 µl ddH_2O, producing a final OB extract for each sample treated.

Genomic DNA was extracted using the CTAB method described in Chapter 2, section 2.2.2, and stored at -25°C. PCR amplification of the *egt* gene was performed as described in Chapter 2, section 2.2.4 and Table 2.2. The PCR amplicons were visualised by viewing a 0.7% agarose gel, run at 90 V for 45 min in 1X TAE, stained with ethidium bromide. Gel

images were captured on a ChemiDoc™ XRS+ system (Bio-Rad Laboratories, Inc. USA) with the Image Lab (v5.1) software (Bio-Rad Laboratories, Inc. USA).

5.3 Results

5.3.1 Occlusion body enumeration

The concentration of the CpGV-SA and CpGV-M (extracted from Carpovirusine®) OB extracts were determined by light microscopy through counting the number of OBs in five large blocks of a counting chamber. The counting of the OBs was conducted three times for each CpGV stock and the average number of CpGV-SA OBs for the three counts was 66 OBs. The average number of CpGV-M OBs was slightly lower at 47.6 OBs for the three counts. The total average number of OBs counted was used to calculate the OB concentrations using the formula given in section 5.2.1. The OB concentrations of CpGV-SA and CpGV-M stocks were determined to be 8.25×10^9 and 5.95×10^9 OBs/ml respectively.

5.3.2 Surface concentration-mortality bioassay with neonate larvae

The surface-treated bioassays were used to estimate the LC_{50} and LC_{90} values of the South African isolate in comparison to the corresponding values of the CpGV-M (extracted from Carpovirusine®) isolate against *C. pomonella* neonate larvae. It was observed from preliminary bioassays (data not shown) that for both isolates, virus concentrations of 4×10^5 OBs/ml resulted in 100% mortality of the neonate larvae. Therefore lower concentrations were used for subsequent bioassays as the mortality for the bioassays must range from 10 – 90%, in order to determine the concentration-mortality response relationship (Jones, 2000). The mortality of the larvae was observed to increase with increasing virus concentration in all three replicates for each of the CpGV isolates (Figure 5.4 and 5.5).

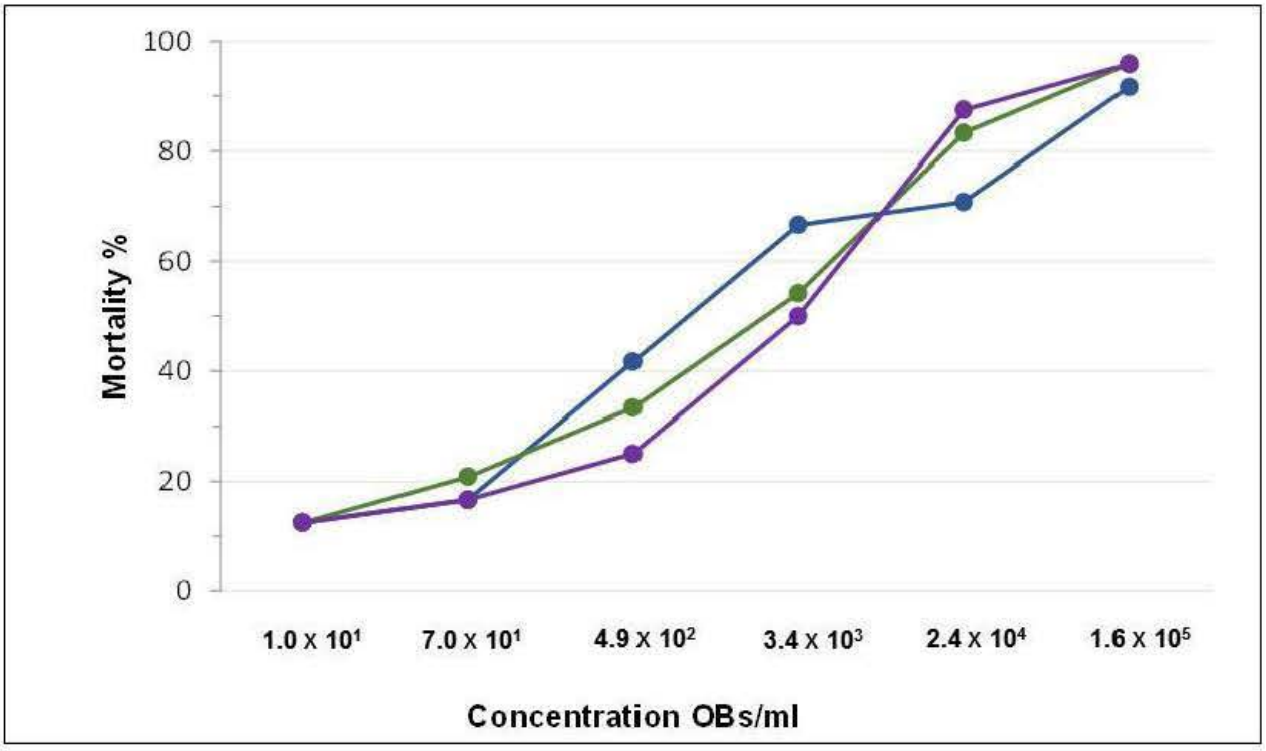


Figure 5.4 Mortality from the three replicates of CpGV-SA neonate dose-response bioassays (—) replicate 1, (—) replicate 2, (—) replicate 3

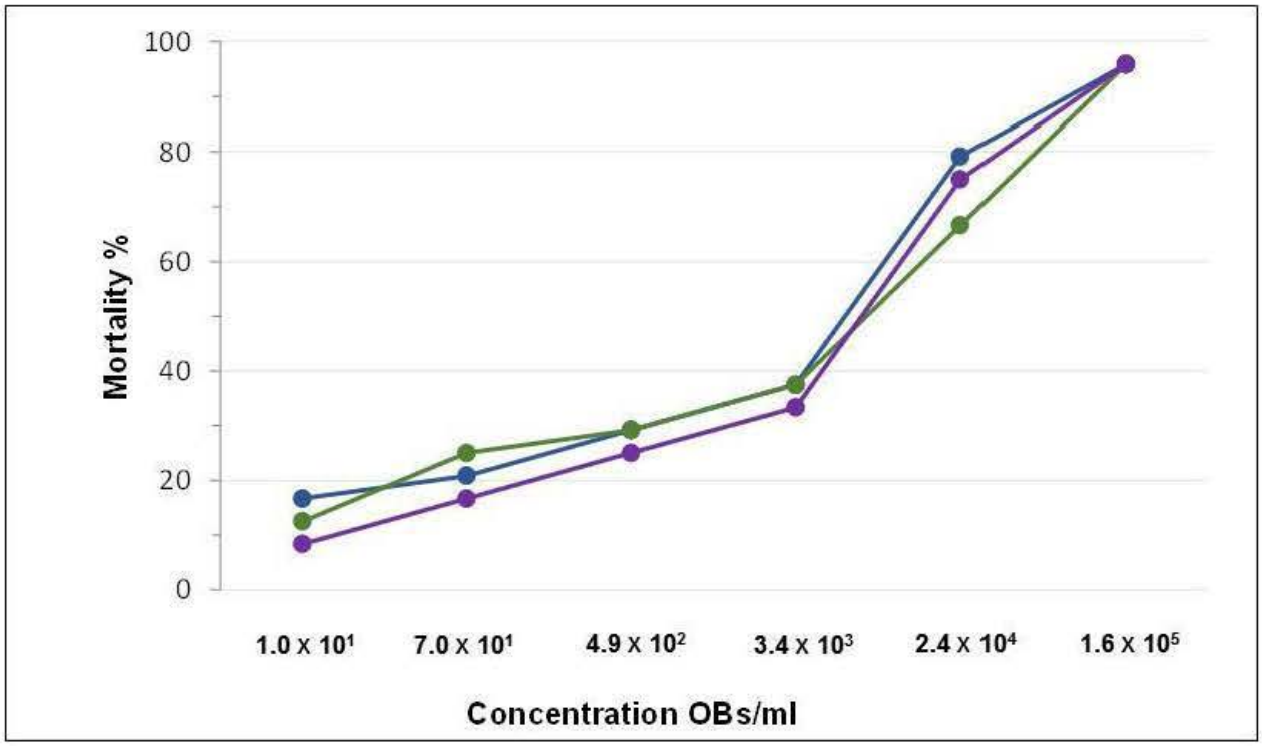


Figure 5.5 Mortality from the three replicates of CpGV-M neonate dose-response bioassays (—) replicate 1, (—) replicate 2, (—) replicate 3

The CpGV-SA bioassays were conducted alongside CpGV-M bioassays and the results obtained are presented for each of the isolates. The control mortality for the first replicate was 4% and the corrected treatment mortality ranged from 8.85 – 91.32% and 13.19 – 95.66% for CpGV-SA and CpGV-M respectively (Table 5.1). The G for fiducial limits was 0.1011 and 0.1053. According to Van Ark (1995), the experimental procedure or the value of the probit line is questionable if the G value exceeds 0.25, while a G value greater than 0.025 indicates a large difference in mortality. The deviations from the regression lines were homogeneous for CpGV-SA ($X^2 = 1.757$, df = 4 and $p < 0.05$) and CpGV-M ($X^2 = 6.801$, df = 4 and $p < 0.05$). The equations for the regression lines were y = 2.9152 + 0.6457x (SE of slope = 0.1048) and y = 2.7015 + 0.6728x (SE of slope = 0.1114).

Table 5.1: Mortality of neonate larvae from the first replicate of CpGV-SA and CpGV-M concentration-mortality response bioassays

Dose OBs/ml	Number exposed	CpGV-SA			CpGV-M		
		Larval mortality	Mortality %	Corrected mortality	Larval mortality	Mortality %	Corrected mortality
Control	24	1	4.00	-	1	4.00	-
1.7×10^5	24	22	91.67	91.32	23	95.83	95.66
2.4×10^4	24	17	70.83	69.62	19	79.17	78.30
3.4×10^3	24	16	66.67	65.28	9	37.50	34.90
4.9×10^2	24	10	41.67	39.24	7	29.17	26.22
7.0×10^1	24	4	16.67	13.19	5	20.83	17.53
1.0×10^1	24	3	12.50	8.85	4	16.67	13.19

The control mortality for the second replicate was also 4% and the corrected treatment mortality ranged from 8.85 – 95.66% for both CpGV-SA and CpGV-M (Table 5.2). The G for fiducial limits was 0.0952 and 0.1123 for CpGV-SA and CpGV-M respectively and the deviations from the regression line were homogeneous CpGV-SA ($X^2 = 1.486$, df = 4 and $p < 0.05$) and CpGV-M ($X^2 = 5.404$, df = 4 and $p < 0.05$). The equations for the regression lines were y = 2.6126 + 0.7423x (SE of slope = 0.1168) and y = 2.7915 + 0.6218x (SE of slope = 0.1063).

Table 5.2 Mortality of neonate larvae from the second replicate of CpGV-SA and CpGV-M concentration-mortality response bioassays

Dose OBs/ml	Number exposed	CpGV-SA			CpGV-M		
		Larval mortality	Mortality %	Corrected mortality	Larval mortality	Mortality %	Corrected mortality
Control	24	1	4.00	-	1	4.00	-
1.7×10^5	24	23	95.83	95.66	23	95.83	95.66
2.4×10^4	24	20	83.33	82.64	16	66.67	65.28
3.4×10^3	24	13	54.17	52.26	9	37.50	34.90
4.9×10^2	24	8	33.33	30.56	7	29.17	26.22
7.0×10^1	24	5	20.83	17.53	6	25.00	21.88
1.0×10^1	24	3	12.50	8.85	3	12.50	8.85

The control mortality for the third replicate was 0% and the corrected treatment mortality ranged from 12.50 – 95.83% and 8.33 – 95.83% for CpGV-SA and CpGV-M respectively (Table 5.3). The G for fiducial limits was 0.0768 and 0.0812 and the deviations from the regression line were homogeneous for CpGV-SA (X^2 = 5.411, df = 4 and $p < 0.05$) and CpGV-M (X^2 = 5.883, df = 4 and $p < 0.05$). The equations for the regression lines were y = 2.7231 + 0.7131x (SE of slope = 0.1008) and y = 2.5897 + 0.689 x (SE of slope = 0.1002).

Table 5.3 Mortality of neonate larvae from the third replicate of CpGV-SA and CpGV-M concentration-mortality response bioassays

Dose OBs/ml	Number exposed	CpGV-SA			CpGV-M		
		Larval mortality	Mortality %	Corrected mortality	Larval mortality	Mortality %	Corrected mortality
Control	24	0	0	-	0	0	-
1.7×10^5	24	23	95.83	95.83	23	95.83	95.83
2.4×10^4	24	21	87.50	87.50	18	75.00	75.00
3.4×10^3	24	12	50.00	50.00	8	33.33	33.33
4.9×10^2	24	6	25.00	25.00	6	25.00	25.00
7.0×10^1	24	4	16.67	16.67	4	16.67	16.67
1.0×10^1	24	3	12.50	12.50	2	8.33	8.33

The regression lines from the three bioassay replicates with each virus isolate were compared. The Bartlett's test for homogeneity of residual variances was conducted and revealed that the residual variances were homogeneous for both CpGV-SA (X^2 = 1.272, df = 2 and $p < 0.01$) and CpGV-M (X^2 = 0.033, df = 2 and $p < 0.01$). The slopes calculated were comparable and parallel for the CpGV-SA (X^2 = 0.225, df = 2, and $p < 0.05$) and

CpGV-M (X^2 = 0.33, df = 2, and $p < 0.05$) replicates. The average results from the three replicates were then used to calculate the LC_{50} and LC_{90} values for each of the virus isolates. The LC_{50} and LC_{90} values were determined to be 1.6×10^3 and 1.2×10^5 OBs/ml respectively for CpGV-SA (Table 5.4). Also shown in Table 5.4 were the LC_{50} and LC_{90} values of CpGV-M, which were 3.1×10^3 and 2.8×10^5 OBs/ml respectively.

Table 5.4 CpGV-SA and CpGV-M LC_{50} and LC_{90} of neonate larvae

Lethal concentration	CpGV-SA			CpGV-M		
	OBs/ml	Fudicial limits		OBs/ml	Fudicial limits	
		Upper	Lower		Upper	Lower
LC_{50}	1.6×10^3	3.9×10^3	6.7×10^2	3.1×10^3	6.9×10^3	1.3×10^3
LC_{90}	1.2×10^5	2.8×10^6	3.4×10^4	2.8×10^5	3.5×10^6	7.1×10^4

5.3.3 Time-mortality response bioassay with neonate larvae

The time-mortality response bioassays were performed as described in section 5.2.4. Every eight hours the larvae were inspected and for all the replicates of the CpGV-SA and CpGV-M bioassays, no larvae were found dead after 48 hours of exposure to the LC_{90} virus concentrations. The mortality observed thereafter was considered to be due to viral infection by the CpGV isolates. For both isolates the first mortality was observed after 54 hours (Table 5.5 and Table 5.6). The time-mortality response bioassays were terminated once 100% mortality was achieved for the virus treatments for each replicate. The control mortality for each of the replicates was 0% and therefore no corrected mortality was calculated.

Table 5.5: Mortality of *Cydia pomonella* neonate larvae in time-mortality response bioassays with the LC_{90} concentration of CpGV-SA (1.2×10^5 OBs/ml)

Virus exposure time		Cumulative mortality % (n=24)		
Day	Hours	Replicate 1	Replicate 2	Replicate 3
3	56	0	8.33	0
	64	4.16	8.33	8.33
	72	4.16	8.33	8.33
4	80	12.5	12.5	12.5
	88	12.5	16.67	16.67
	96	16.67	16.67	16.67
5	104	25	20.83	25
	112	29.17	29.17	29.17
	120	33.33	33.33	33.33
6	128	37.5	45.83	37.5
	136	41.67	50	45.83
	144	45.83	58.33	58.33
7	152	54.17	62.5	62.5
	160	58.33	66.67	66.67
	168	79.17	83.33	79.17
8	176	83.33	87.5	83.33
	184	91.67	91.67	91.67
	192	91.67	100	95.83
9	200	100		100

Table 5.6: Mortality of *Cydia pomonella* neonate larvae in time-mortality response bioassays with the LC_{90} concentration of CpGV-M (2.8×10^5 OBs/ml)

Virus exposure time		Cumulative mortality % (n=24)		
Day	Hours	Replicate 1	Replicate 2	Replicate 3
3	56	0	4.16	4.16
	64	4.16	8.33	8.33
	72	8.33	8.33	12.5
4	80	12.5	8.33	12.5
	88	12.5	12.5	12.5
	96	12.5	16.67	16.67
5	104	12.5	25	20.83
	112	16.67	29.17	25
	120	16.67	33.33	37.5
6	128	20.83	33.33	45.83
	136	29.17	37.5	54.17
	144	41.67	54.17	58.33
7	152	62.5	58.33	75
	160	66.67	70.83	83.33
	168	75	75	87.5
8	176	91.67	79.17	91.67
	184	95.83	87.5	95.83
	192	100	91.67	95.83
9	200		100	100

The time-mortality response relationship for the CpGV isolates against *C. pomonella* neonate larvae were analysed using a logit regression (Figure 5.6-5.7 and Table 5.7-5.8). From the statistical analysis, the LT_{50} value for each replicate was calculated. The mean LT_{50} value for CpGV-SA was 135 hours (SE = 0.003) and the LT_{50} value of CpGV-M was 136 hours (SE = 0.003). The calculated Chi-squared values revealed that the data obtained fit the logit regression model defined by the line equation y=exp(β+(x-intercept)×x)/(1+exp(β+x-intercept) (Table 5.7-5.8).

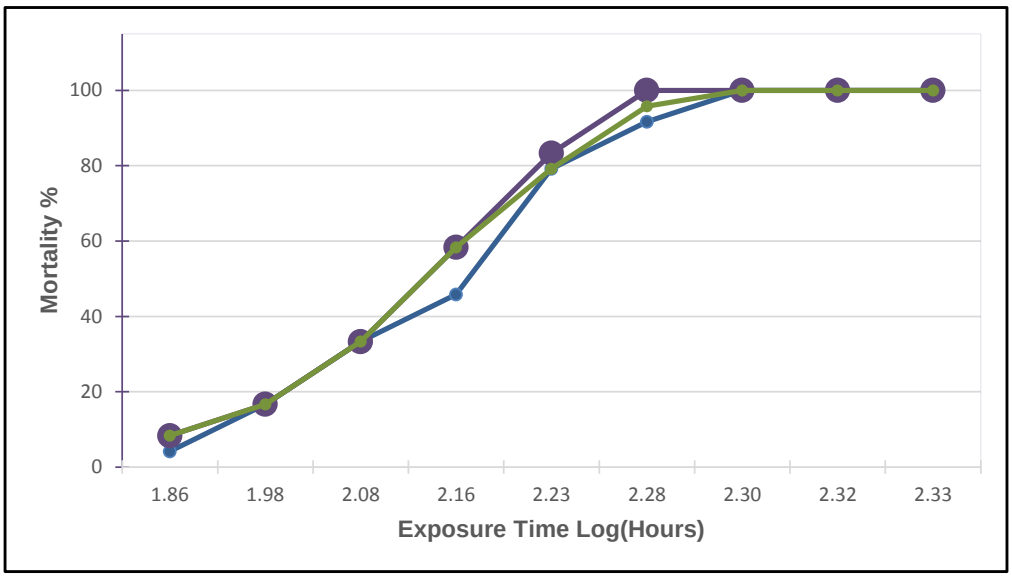


Figure 5.6: Time-mortality response relationship of LC_{90} concentration of CpGV-SA (1.2×10^5 OBs/ml) against *Cydia pomonella* neonate larvae

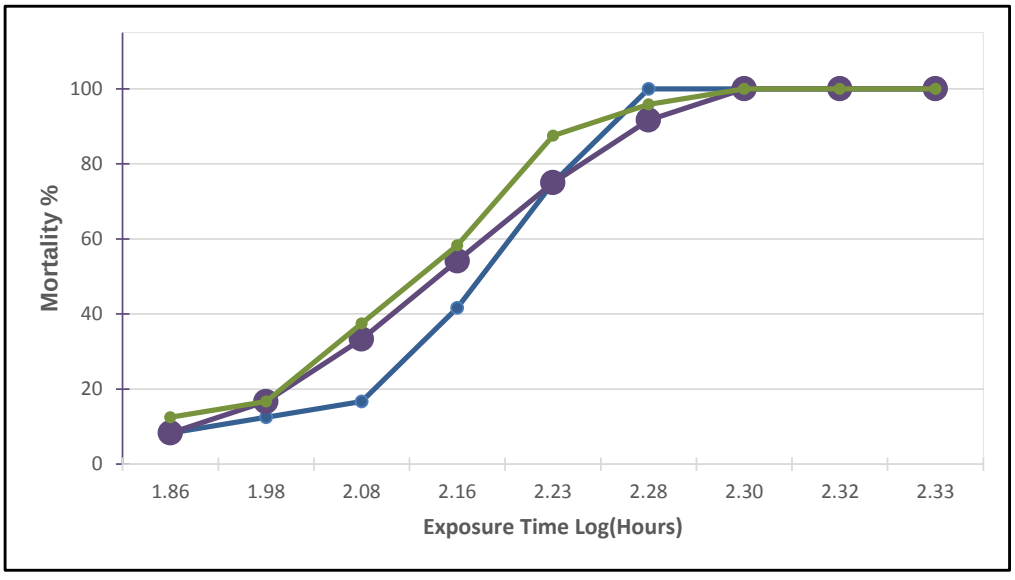


Figure 5.7: Time-mortality response relationship of LC_{90} concentration of CpGV-M (2.8×10^5 OBs/ml) against *Cydia pomonella* neonate larvae

Table 5.7: Logistic regression data for the mortality of neonate larvae subjected to the LC_{90} concentration of CpGV-SA (1.2×10^5 OBs/ml)

Replicates	Chi-square (D.F = 1; $p < 0.001$)	Const. B0 (β)	S.E.	X-intercept (S.E. 0.003)	LT_{50} (days)
1	540.02	6.45	0.44	-0.046	5.8
2	437.51	6.56	0.47	-0.049	5.5
3	548.92	6.29	0.43	-0.047	5.6

Table 5.8: Logistic regression data for the mortality of neonate larvae subjected to the LC_{90} concentration of CpGV-M (2.8×10^5 OBs/ml)

Replicates	Chi-square (D.F = 1; $p < 0.001$)	Const. B0 (β)	S.E.	X-intercept (S.E. 0.003)	LT_{50} (days)
1	411.57	7.16	0.52	-0.049	5.9
2	510.60	5.95	0.40	-0.043	5.7
3	548.89	6.09	0.42	-0.047	5.4

5.3.4 PCR amplification of the *egt* gene

5.3.4.1 Surface concentration-mortality response bioassay observations

The surface-treated bioassays were evaluated by recording the number of larvae which survived exposure to the different concentrations of virus, as opposed to counting dead larvae because dead larvae were not always visible on the surface of the diet. When dead larvae were visible on the surface of the diet a difference in the external appearance and response to external stimuli was observed between the larvae which had been exposed to virus and those which had not been exposed, in the control plate (Figure 5.8). *Cydia pomonella* granulovirus infected larvae had a distinct milky white appearance (Figure 5.8A) and upon attempting to collect these larvae most would liquefy, as shown in Figure 5.8B. The collection of dead larvae was therefore difficult; a sterile pick and bud were used for the collection of all dead larvae from the CpGV-SA and CpGV-M bioassays, of which the larvae were pooled from the different dilutions for all the replicates. The head capsules of some dead larvae could be found on the surface on the diet, shown in Figure 5.8C and no collection could be made from this surface as components of the diet could affect downstream applications. *Cydia pomonella* larvae which survived were easily found burrowed into the diet, under a layer of silk and frass (Figure 5.8D). Uninfected larvae in the control plates were pinkish white and would respond to external stimuli by burrowing deeper into the diet.

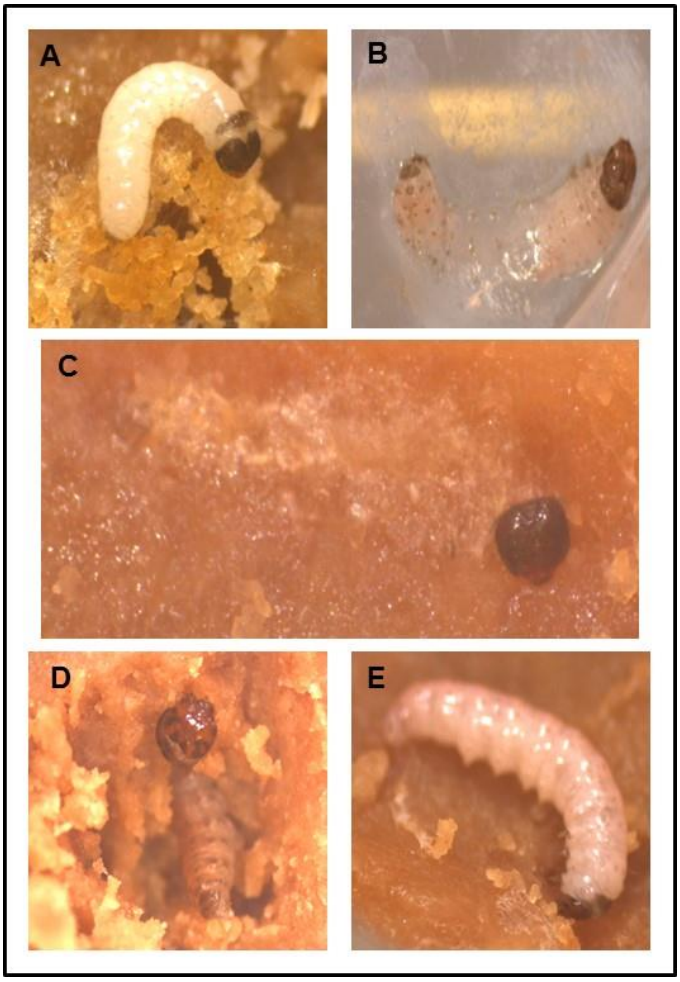


Figure 5.8: General observations of *Cydia pomonella* larvae across the surface concentration-mortality response bioassays. A) Virus infected larva, B) Liquidation of larva upon applying external stimuli, C) Head capsule of dead larva, D) Larva found alive burrowed into the diet, E) Appearance of uninfected larva on control plate.

5.3.4.2 Analysis of the *egt* gene sequences

OBs were recovered from pooled samples of dead larvae from the CpGV-SA bioassays and the CpGV-M bioassays. DNA was then extracted from the OBs and used as a template to amplify the CpGV *egt* gene. This was done to ensure that the virus used in the bioassays was the cause of the mortality observed. The *egt* gene was successfully amplified using genomic DNA extracted from the purified OBs of both samples (Figure 5.9). The amplified products were analysed by 0.7% agarose gel electrophoresis and bright bands slightly below the 1500 bp mark were observed.

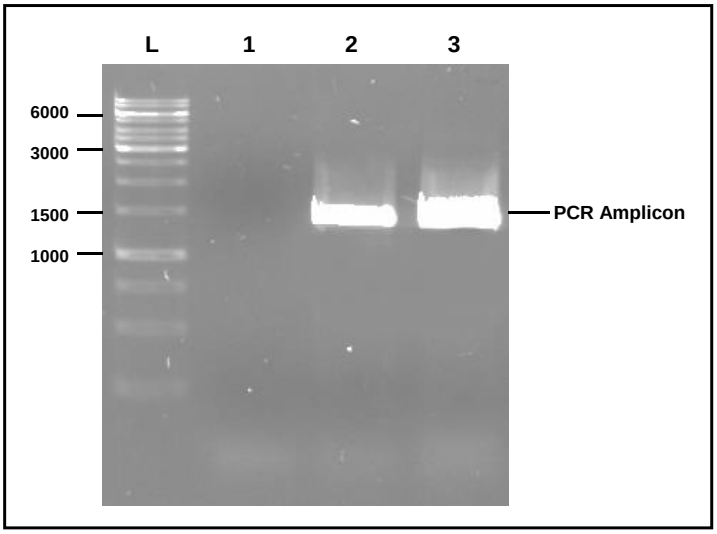


Figure 5.9: PCR amplification of the CpGV *egt* gene from dead larvae collected from the surface concentration-mortality response bioassays. L- GeneRuler™ 1Kb DNA ladder; 1- no template control; 2- CpGV-SA sample; 3- CpGV-M sample

The partial *egt* sequences were obtained from Inqaba Biotechnical Industries (Pty) Ltd., South Africa. Ambiguous nucleotides were edited using the alignments and sequence chromatograms viewed in Finch TV (USA) version 1.4.0. Single consensus sequences were assembled for the *egt* sequences using MEGA 5.2 (Tamura *et al*., 2011) by aligning the sequence data with the respective sequences from CpGV-SA and CpGV-M. The *egt* sequence obtained from the CpGV-SA bioassay sample was 1359 nt with an identity of 100% when aligned against the *egt* gene of CpGV-SA and the *egt* sequence obtained from the CpGV-M bioassay sample was 1373 nt with an identity of 100% when aligned against the *egt* gene of CpGV-M1 (Luque *et al.,* 2001), as it is a representative for all CpGV-M isolates. In Chapter 2, the *egt* sequence of CpGV-SA acquired was 1382 nt with BLAST, indicating an identity of 99% (E value of 0.0) when compared against the reference isolate, CpGV-M1. The two SNPs at nucleotide positions 121144A→G and 122295A→G, which were respectively synonymous and non-synonymous (Table 2.4 of section 2.3.4 in Chapter 2), were detected again in the *egt* sequence obtained from the CpGV-SA bioassay sample. This revealed that the nucleotide substitutions were retained in the *egt* gene thereby confirming the presence of CpGV-SA as the cause of mortality.

5.4 Discussion

In order to determine the biological activity of CpGV-SA against neonate larvae, the virus stock was diluted to different concentrations for the surface concentration-mortality

response bioassays and, upon hatching, the *C. pomonella* larvae were subjected to the different dilutions. Once the LC_{90} concentration was determined, time-mortality response bioassays were performed over a period of 9 days. The same procedures were conducted for CpGV-M, extracted from Carpovirusine®, alongside the CpGV-SA bioassays for comparative purposes. In this study, the *C. pomonella* egg sheets were obtained from an insect culture originating from the Western Cape Province of South Africa and maintained by ENTOMON Technologies Pty Ltd, Stellenbosch, South Africa. The biological activity of CpGV-SA could not be tested against the insect population from which it was isolated because this was not available. However, the LC and LT values obtained in this study are considered an indication of how virulent the isolate would be against a population of *C. pomonella* which would usually be controlled with a CpGV-M based biopesticide such as Carpovirusine®. An added advantage of using the Western Cape culture is that the largest pome fruit production sites are situated in this Province and therefore more control agents are applied in this area (Addison, 2005; Hortgro, 2014).

The LC and LT values obtained in this study revealed that the biological activity of CpGV-SA is similar to that of CpGV-M (Carpovirusine®) against the Western Cape *C. pomonella* population tested. However, the LC_{50} and LC_{90} values for CpGV-SA (1.632×10^3 and 1.163×10^5 OBs/ml) were lower than the LC_{50} and LC_{90} for CpGV-M (Carpovirusine®) (3.107×10^3 and 2.828×10^5 OBs/ml) indicating that the CpGV-SA isolate may be slightly more virulent. The LT_{50} values of 135 and 136 hours (5.6 and 5.7 days) for CpGV-SA and CpGV-M were similar against the *C. pomonella* neonate larvae. A possible reason for the difference in biological activity observed could relate to the *C. pomonella* population tested. Geographically distinct *C. pomonella* populations have been shown to be genetically different and this may affect the observed efficacy of the virus tested (Gund *et al.,* 2012; Opoku-Debrah *et al.,* 2014).

It is not possible to compare the LC and LT values obtained in this study with the respective values found in literature (section 1.6 of Chapter 1). Bioassays of several CpGV isolates and mostly of CpGV-M against *C. pomonella* neonate larvae have been conducted (Lacey *et al.,* 2008). The lethal concentrations and time to kill 50% of the neonate larvae may vary depending on the inoculation methods used for the bioassays. In this study, surface inoculation was used due to the limited availability of the virus stocks. However, due to the feeding habits of *C. pomonella* larvae, the diet incorporation method is considered preferential for laboratory bioassays on artificial diet, as the larvae burrow into the diet while

feeding. The diet incorporation method does however require larger volumes of virus suspensions (Lacey, 2012). The surface treatment method could be used to determine the concentration-mortality response relationships, as larvae ingested the virus particles with the diet prepared and this was revealed through the observation of infected and dead larvae. The concentration-mortality response bioassays revealed that mortality of neonate larvae increased with increasing virus concentration, an observation that is supported by similar studies using susceptible insects (Jacques *et al.,* 1987; Fritsch *et al.,* 2007; Zichová *et al.,* 2011). However, where there is resistance to the virus in the host population tested, the biological activity is decreased and the trend is only visible with resistance overcoming isolates as was shown with CpGV-I12 (Eberle *et al.,* 2008).

The preferred inoculation method for time-mortality response bioassays is droplet feeding (Lacey, 2012). However, in this study, it was observed that, when the virus suspension was not completely dry, the neonate larvae died in any wet areas due their size (data not shown). Therefore the surface inoculation method was chosen for the time-mortality response bioassays. The LT_{50} values of 5.6 and 5.7 days were calculated for CpGV-SA and CpGV-M respectively. Sheppard & Stairs (1977), calculated LT_{50} values of 9.7 days for 3 OBs per larva to 3.7 days for 280 OBs per larva for CpGV against *C. pomonella* neonate larvae. The biological assay setup for their studies differed from the setting of the bioassays in this study and therefore the values cannot be compared. Furthermore, LC and LT values may differ based on the CpGV susceptibility of the population of *C. pomonella* tested. This has been shown for genetically different CpGV isolates tested against susceptible and resistant *C. pomonella* populations where the biological activity of the CpGV isolates varied for each of the populations tested and across the different isolates (Gebhardt *et al.,* 2014).

After conducting the surface concentration-mortality response bioassays, virus was recovered from the samples of pooled dead larvae for each of the CpGV bioassays. The PCR amplification and sequencing of *egt* revealed that there was no cross contamination between the two isolates, CpGV-SA and CpGV-M (Carpovirusine®), when performing the bioassays for all three replicates. The results obtained also showed that the SNPs detected in the *egt* gene of CpGV-SA in Chapter 2 were maintained in the virus recovered. The sequences analysed confirmed that the mortality observed from the concentration-mortality response bioassays was a result of CpGV-SA and CpGV-M infections.

In conclusion the biological activity of CpGV-SA was observed to be similar to Carpovirusine® indicating that it has potential for development and application as a biopesticide for control of *C. pomonella* in South Africa. Moreover, it could be used as an alternative option interchangeably with CpGV-M based products in managing resistance in insect populations should this occur in South Africa but to also delay or prevent the occurrence of resistance.

Chapter 6

6.1 General Discussion

The codling moth, *Cydia pomonella* (L.) (Lepidoptera: Tortricidae), is a globally distributed, serious pest of pome fruit such as apples, pears and quinces. It is responsible for not only causing extensive damage to pome fruit but for also feeding on other crops such as cherries, plums, nectarines, peaches, apricots and walnuts, resulting in significant economic losses in fruit industries worldwide (Ciglar, 1998; Wearing *et al.,* 2001). In the past, the control of this insect pest was dependent mostly on the use of broad spectrum insecticides (Rield *et al.,* 1998; Pajač *et al.*, 2011). However, over the years increased cases of *C. pomonella* resistance to several classes of these insecticides has been reported and chemical insecticides have been shown to cause various types of environmental damage (Dunley & Welter, 2000; Franck *et al.,* 2007; Reyes *et al.,* 2007; Rodriguez *et al.,* 2012). In an attempt to reduce the use of insecticides and the risks to human health and ecosystems due to the presence of chemical residues in the environment and foods, integrated pest management (IPM) is being applied worldwide, including in South Africa (de Waal *et al.,* 2010; Rodriguez *et al.,* 2012). An important part of IPM is the use of biological control measures, specifically the use of microbial pest control agents such as baculoviruses, which are insect viruses that have been successfully used as biopesticides (Szewczyk *et al.,* 2006; Lapointe *et al.,* 2012). *Cydia pomonella* granulovirus is a baculovirus used to suppress *C. pomonella* populations. Although, highly specific and virulent the occurrence of *C. pomonella* resistance to CpGV, in particular to the Mexican isolate (CpGV-M), has led to the increased need to search for and investigate the use of alternative CpGV isolates (Lacey *et al.,* 2008).

The overall aim of this study was to genetically and biologically characterise a *C. pomonella* granulovirus isolate recovered from dead and diseased *C. pomonella* larvae collected from a farm in the Free State Province of South Africa. The study was conducted in order to firstly determine whether the isolate is novel and genetically different from CpGV-M, which is the active ingredient of the granulovirus based biopesticides used in South Africa. Secondly, and most importantly, the study was conducted to evaluate the isolate's potential for use in resistance management strategies, which involves both preventing the onset of resistance and managing resistance, should it occur in South Africa. The specific objectives which were to morphologically and genetically identify the virus prior to characterising the

complete genome and conducting comparisons with other CpGV isolates, as well as determining the biological activity against neonate *C. pomonella* larvae were achieved.

The transmission electron microscopy analysis confirmed that the virus recovered from the field collected *C. pomonella* larvae was a granulovirus, due to the morphological traits observed. The genetic identification was achieved through the PCR amplification of the *granulin*, *late expression factor* 8 (*lef*-8) and *late expression factor* 9 (*lef*-9) gene sequences, as described by Jehle *et al.* (2006) and Lange *et al.* (2004). In addition to the three highly conserved genes, the *ecdysteroid UDP-glucosyltransferase* (*egt*) gene sequence of the isolate characterised in this study was also PCR amplified and analysed. Comparisons of the respective genes with the reference isolate, CpGV-M1 (Luque *et al.,* 2001), were conducted and this confirmed that the virus isolate was a member of the CpGV species. Furthermore, through single nucleotide polymorphism (SNP) detection in the *granulin*, *lef*-8 and *lef*-9 genes compared with all the respective sequence data from geographically different CpGV isolates available on the NCBI's GenBank database, it was revealed that an isolate of mixed genome types B and E was recovered. The *egt* sequence data did not show an identical match with any of the CpGV isolates of the different genome types showing that the isolate was novel and as a result was referred to CpGV-SA throughout the remainder of the study. The SNP analysis was performed in accordance with the proposed genotyping suggested by Eberle *et al.* (2009). The phylogenetic analysis was based on the SNPs detected and supported the finding that the CpGV-SA was of mixed genome types B and E. As stated by Eberle *et al.* (2009), the genome types are not an indication of the origin of the isolate, therefore no conclusions as to where the CpGV-SA isolate may have been introduced from can be made.

Further genetic characterisation of CpGV-SA was then conducted to obtain more information about the complete genome which was achieved by firstly performing restriction endonuclease (REN) analysis and then complete genome sequencing of the genomic DNA. For this study, it was important to determine if CpGV-SA is genetically different from CpGV-M, as CpGV-M has been used for most CpGV based biopesticides and the resistance observed in *C. pomonella* to CpGV has been shown to be specific to mostly CpGV-M (genome type A) and would potentially be specific to all genome type A isolates (Gebhardt *et al.,* 2014). Therefore, alternative control agents would need to be genetically different from CpGV-M in order to be considered for biopesticide production. The REN analysis was conducted alongside CpGV-M extracted from Carpovirusine®, a CpGV-M based

biopesticide used in South Africa and the profiles generated revealed that the South African isolate is distinctly different from the Mexican isolate. The complete genome sequence was used to validate the SNPs detected previously in the *granulin*, *egt*, *lef-8* and *lef-9* genes and the previous SNP analysis was shown to be accurate. The complete genome was then compared to the reference isolate, CpGV-M1 (Luque *et al.,* 2001) and a recently added genome sequence of a CpGV-M isolate, also representing the genome type A (Gebhardt *et al.,* 2014) and 11 open reading frames with high genetic variation in the nucleotide and amino acid sequences between the South African isolate and the Mexican isolates were detected. Genome variation in terms of the sizes of the genomes and mutations in non-coding regions, although not shown, were also observed. The complete genome analysis revealed that the CpGV-SA isolate is genetically different from CpGV-M and based on this comprehensive genetic analysis the novel isolate could be considered as an alternative control agent to CpGV-M.

The REN analysis was successful in revealing genome variations between isolates of the same species. However, it would not be possible to obtain DNA from all the geographically and genetically distinct CpGV isolates. Therefore, *in silico* restriction digests were conducted using the complete genome sequences available for genetically different CpGV isolates, which was limited to isolates M, E2, I12, I07 and S of genome types A, B, C, D and E respectively (GenBank Accession numbers: KM217573-KM217577) (Gebhardt *et al.,* 2014). The reference isolate, CpGV-M1 of genome type A, was also included in the comparative genomic analysis. The *in silico* digest profiles showed high similarities with the profiles of isolates of genome type A, B and mostly E, further supporting the suggestion that CpGV-SA is comprised of mixed genotypes. Further investigations using the available complete genome sequences were performed and the genome sequences of naturally occurring isolates E2, I12, I07 and S belonging to CpGV genome types B-E, that are also considered to be resistance overcoming isolates, were each aligned with the CpGV-SA genome (Eberle *et al.,* 2008; Eberle *et al.,* 2009; Berling *et al.,* 2009b; Zichová *et al.,* 2013; Gebhardt *et al.,* 2014). The highest percentage identity obtained for the genome alignments was between CpGV-SA and CpGV-S (genome type E), suggesting that although the CpGV-SA isolate is of mixed genotypes, it may be predominantly of genome type E. Recently, the functional difference between the resistance overcoming CpGV isolates and the CpGV-M isolate was discovered to be a mutation in the *pe38* gene of CpGV-M isolate that is not present in the other isolates able to infect resistant *C. pomonella* (Gebhardt *et al.,* 2014).

When the *pe38* gene of CpGV-SA was aligned with the *pe38* gene of the Mexican isolates and resistance overcoming isolates, the mutation observed in the Mexican isolate was not observed and therefore it was concluded that the South African isolate could be considered a resistance overcoming isolate. In addition to this observation, SNPs in this gene were detected and it was revealed that the *pe38* gene of CpGV-SA is identical to the respective gene in CpGV-S, supporting the previous information suggesting that although CpGV-SA is of mixed genome types A, B and E, and is predominantly of genome type E. The comparative genomic analysis results showed that the CpGV-SA isolate is novel, comprised of mixed genotypes but predominantly of genome type E and can be considered a resistance overcoming isolate. This would have significant implications for resistance management strategies, particularly if resistance were to occur in South Africa. However, even before any onset of resistance, CpGV-M and CpGV-SA can be used interchangeably, in a resistance management approach aimed at averting the development of resistance by *C. pomonella* to CpGV.

The initial step to developing a new biopesticide would be to first evaluate the biological activity of the virus isolate against a host population under laboratory conditions. This is necessary in order to determine the lethal concentrations and speed of kill of the virus isolate. Surface concentration-mortality and time-mortality response bioassays were conducted for CpGV-SA alongside CpGV-M, extracted from Carpovirusine® for comparative purposes, against a Western Cape culture of *C. pomonella*. Although, the data obtained could not be compared with that found in the literature, due to differences in bioassay methods, the surface inoculation method was successfully used to determine LC and LT values, which were found to be similar between the isolates. This is advantageous as CpGV-M is highly efficient at controlling *C. pomonella*. It has been shown that single applications of CpGV-M based products can result in a 50% decrease in overwintering *C. pomonella* populations and two to three applications could reduce damage to fruit by 85%, which was shown to be comparable with reductions in injury to fruit achieved with chemical insecticides (Jaques *et al.*, 1981; Kienzle *et al.*, 2002). The biological activity supported the genetic analysis in that the novel CpGV-SA isolate can be considered for the development of a new biopesticide.

6.2 Conclusions and future work

In this study, a novel CpGV-SA isolate was characterised, it was determined to be genetically different from CpGV-M and shown to match the resistance overcoming isolates genetically with regard to the functional difference observed in comparison to CpGV-M. The biological activity of this isolate was also evaluated under laboratory conditions. It has been suggested that the application of a single isolate of a single genome type should be avoided in the commercial applications of baculoviruses in order to prevent the development of resistance to a product that is applied frequently over an extended period (Eberle, 2010; Opoku-Debrah *et al*., 2013). Therefore, the novel CpGV-SA isolate could be used alongside CpGV-M based products. The CpGV-SA isolate is also a naturally occurring isolate of mixed genotypes and it may potentially have an advantage over isolates of a single genome type, when it comes to resistance management (Eberle, 2010). Although, genotyping was achieved in this study confirmation of the mixed genotypes and determining the genotype frequencies of the virus isolate could be determined using quantitative real-time PCR (qPCR) and would be an ideal initial step to furthering this study (Zwart *et al.*, 2008).

There are several advantages and disadvantages of using baculoviruses found naturally in pest populations as biological control agents (Rodriguez *et al.,* 2012). The insect viruses are considered safe and have narrow host ranges and can therefore not cause harm to non-target organisms in the environment (Szewczyk *et al.,* 2006). *Cydia pomonella* granulovirus is one of those baculoviruses that is highly specific for *C. pomonella* although it is also able to infect (albeit at a much higher concentration) and replicate in *T. leucotreta*, which is a closely related host species and is also an insect pest (Chambers, 2014). It is important to characterise novel baculovirus isolates as alternative control agents, as the viruses are safe for human health and do not pose a threat to the environment (Rodriguez *et al.,* 2012). Another important advantage of using baculoviruses is that within an IPM programme the suppression of host populations can be significantly increased, due to the viruses targeting the larvae before significant damage is caused and due to the fact that the activity of the virus is not affected by the use of chemical insecticides (Copping *et al.*, 1992). One of the disadvantages with the use of baculoviruses is the costs involved. Also, the narrow host range means that alternative control methods will still be needed for other insect pests present in the fruit production systems (Szewczyk *et al*., 2011). The costs

involved in the production of biological control agents is usually high (Ignoffo *et al.*, 1977). This cost will directly affect the cost of the products and will make it difficult for private companies to invest in the production and sales of the biopesticides (Canningham, 1995). This makes it difficult for biopesticides to be competitive with chemical insecticides (Black *et al*. 1997). Another disadvantage of baculoviruses is that the late larval stages are more resistant to infections and the neonate larvae needs to be targeted or higher concentrations of virus will need to be applied to achieve control over the pest (Washburn *et al*., 2003). This is of course not an issue with *C. pomonella*, as it would in any case only be the neonate larva that could be exposed to virus, as the instars thereafter are cryptic within the fruit. The persistence of the virus in the fields is affected by some cultural techniques and solar radiation, which is known to degrade the virus particles (Moscardi, 1999; Rodriguez *et al.,* 2012). This would mean that more applications of the virus would be needed than would be the case for chemical insecticides and would increase the cost of the control method. However, the biggest disadvantages of baculoviruses are that insects must ingest the virus in order to get infected and the infection has a slow speed of kill. Infected larvae continue feeding throughout the period of infection causing damage to crops due to the slow speed of kill (Ignoffo, 1992). Although there are several disadvantages to the use of baculoviruses in controlling insect pests, CpGV used to control *C. pomonella* has been developed into a successful control agent. Field trial results have shown that the reduction in the damage caused to apples is the same as with chemical insecticide use (Huber & Dickler, 2009).

Several other studies would need to be conducted before a virus product can be developed and made available for commercial use. The focus of these studies would involve the establishment of an efficient and cost- effective method for mass rearing of the host and mass production of the virus, field trials and the evaluation of the product's feasibility in integrated management programmes (Moore, 2002). A study has been conducted, which focused on developing a method to reliably produce CpGV in *T. leucotreta,* as it was shown that mass rearing *T. leucotreta* is more affordable than rearing *C. pomonella,* due to diet expenses and insect yields and therefore CpGV would be produced at a lower cost (Chambers, 2014). Future studies would involve expansions on such studies. In the same study, mass production trials were conducted and virus yields produced in *T. leucotreta* were satisfactory. The only problem was the quality of the product: high levels of CrleGV contamination in both the fourth and fifth instar production samples were detected and therefore methods would need improvements. Future work would also need to include

determining the biological activity of more isolates across a wider range of insect populations as differences have been shown to exist between the genetics of geographically different insect host populations and because it is not known for certain that resistance to the newly characterised isolates will not occur (Opoku-Debrah *et al.,* 2013).

Before resistance cases were reported CpGV-M products were shown to be the most efficient biological control agents for suppressing *C. pomonella* populations (Falcon *et al.*, 1968; Glen & Payne, 1984; Jaques *et al.*, 1981; Jaques *et al.*, 2012). CpGV products containing resistance overcoming isolates as the active ingredient have proven to be reliable alternatives to CpGV-M-based products (Schmitt *et al.*, 2013). It is the diversity regarding the genetics of the CpGV isolates used in new products that will lower the risk of resistance occurring in the future (Jehle, 2008). Therefore the continuation of bioprospecting for novel CpGV isolates is still necessary and the characterisation of CpGV-SA was important as it has the potential to be developed in a biopesticide for the management of resistance and use in controlling *C. pomonella* in South Africa.

References

Abbott, W. S. (1925). A method of computing the effectiveness of an insecticide. *Journal of the American Mosquito Control Association* **3**, 302–303.

Abdulkadir, A. F., Marsberg, T., Knox, C. M., Hill, M. P. & Moore, S. D. (2013). Morphological and Genetic Characterization of a South African Plutella xylostella Granulovirus (Plxy GV) Isolate. *African Entomology* **21**, 168–171.

Abdulkadir, F., Knox, C., Marsberg, T., Hill, M. P. & Moore, S. D. (2015). Genetic and biological characterisation of a novel Plutella xylostella granulovirus, PlxyGV-SA. *BioControl* 1–9.

Ackermann, H.-W. & Smirnoff, W. A. (1983). A morphological investigation of 23 baculoviruses. *Journal of Invertebrate Pathology* **41**, 269–280.

Addison, M. F. (2005). Suppression of Codling Moth Cydia pomonella L . (Lepidoptera : Tortricidea) Populations in South African Apple and Pear Orchards Using Sterile Insect Release. *IX International Pear Symposium* 555–557.

Aghdam, H. R., Fathipour, Y., Radjabi, G. & Rezapanah, M. (2009). Temperature-Dependent Development and Temperature Thresholds of Codling Moth (Lepidoptera: Tortricidae) in Iran. *Environmental Entomology* **38**, 885–895.

Agnello, A. & Kain, D. (1996). Codling Moth Fact Sheet.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology* **215**, 403–10.

Andermatt, M., Mani, E., Wildbolz, T. & Lüthy, P. (1988). Susceptibility of Cydia pomonella to Bacillus thuringiensis under laboratory and field conditions. *Entomologia Experimentalis et Applicata* **49**, 291–295.

Arneodo, J. D., De Anna, J., Salvador, R., Farinon, M., Quintana, G. & Sciocco-Cap, a. (2015). Prospection and molecular analysis of CpGV isolates infecting\n Cydia pomonella\n at different geographical locations in Argentina. *Annals of Applied Biology* **166**, 67–74.

Arthurs, S. P. & Lacey, L. A. (2004). Field evaluation of commercial formulations of the codling moth granulovirus: persistence of activity and success of seasonal applications against natural infestations of codling moth in Pacific Northwest apple orchards. *Biological Control* **31**, 388–397.

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

Audemard, H., (1991). Population dynamics of the codling moth. In *World crop pests. Tortricid pests. Their biology, natural enemies and control*. pp 329-338 Edited by L.P.S. van der Geest & H.H. Evenhuis, Amsterdam, Elsevier.

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.

Ballard, J., Ellis, D. J. & Payne, C. C. (2000). Uptake of Granulovirus from the Surface of Apples and Leaves by First Instar Larvae of the Codling Moth Cydia pomonella L. (Lepidoptera: Olethreutidae). *Biocontrol Science and Technology* **10**, 617–625.

Barnes, B. N. (1990). FFTRI monitoring manual for orchard research institute, Stellenbosch, pp. 21.

Barnes, M. M. (1991). Codling moth occurrence, host race formation and damage. In *Tortricid pests, their biology, natural enemies and control.* Edited by Edited by L.P.S. van der Geest & H.H. Evenhuis, Amsterdam, Elsevier.

Barnes, B. N. & Blomefield, T. L. (1997). Confusion amongst codling moth fellows continues: A commercial perspective on the implementation of codling moth mating disruption in North America. In *Technology transfer in mating disruption*. pp. 57-63. Edited by P. Witzgall & H. Arn, Montepellier. IOBC wprs Bulletin.

Barnes, B. N., Hofmeyr, J. H., Groenewald, S., Conlong, D. E. & Wohlfarter, M. (2015). Review article The sterile insect technique in agricultural crops in South Africa : a metamorphosis .... but will it fly ? *African Entomology* **23**, 1–18.

Bell, J. & McGeoch, M. (1996). Evaluation of the pest status and research conducted on phytophagous Lepidoptera on cultivated plants in South Africa RID F-8353-2011. *African Entomology* **4**, 161–170.

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Sayers, E. W. (2013). GenBank. *Nucleic acids research* **41**, D36–42.

Benz, G. A. (1986). Introduction: Historical perspectives. In *The biology of baculoviruses.* pp. 1-35. Edited by R. R. Granad & B. A. Federici. CRC Press.

Bergold, G. H. (1948). Uber die Kapselvirus-Krankheit*. Z. Naturforsch* **3**, 338 – 342.

Berling, M., Rey, J.-B., Ondet, S.-J., Tallot, Y., Soubabère, O., Bonhomme, A., Sauphanor, B. & Lopez-Ferber, M. (2009a). Field Trials of CpGV Virus Isolates Overcoming Resistance to CpGV-M. *Virologica Sinica* **24**, 470–477.

Berling, M., Blachere-Lopez, C., Soubabere, O., Lery, X., Bonhomme, A., Sauphanor, B. & Lopez-Ferber, M. (2009b). Cydia pomonella granulovirus Genotypes Overcome Virus Resistance in the Codling Moth and Improve Virus Efficiency by Selection against Resistant Hosts. *Applied and Environmental Microbiology* **75**, 925–930.

Black, B.C., Brennan, L.A., Dierks, P.M. & Gardm, I.E. (1997). Commercialization of baculoviral insecticides. In: *The Baculoviruses*, Edited by Miller L.K. pp. 341-387. Plenum Press.

Bliss, C. I. & Stevens, W. L. (2008). The calculation of the time-mortality curve. *Annals of Applied Biology* **24**, 815–852.

Bloem, K. A., Bloem, S. & Tan, K.-H. (2000). SIT for codling moth eradication in British Columbia, Canada. In *Area-wide control of fruit flies and other insect pests Joint proceedings of the international conference on area-wide control of insect pests, 28 May-2 June, 1998 and the Fifth International Symposium on Fruit Flies of Economic Importance, Penang, Malay*, pp. 207–214. Penerbit Universiti Sains Malaysia.

Blomefield, T. L. & Giliomee, J. H. (2014). Validation of the phenology model for the codling moth, Cydia pomonella (Lepidoptera: Tortricidae), in South African pome fruit orchards. *African Entomology* **22**, 30–48.

Blomefield, T. L. & Giliomee, J. H. (2012). Fecundity and mortality of codling moth, Cydia pomonella (L.) (Lepidoptera: Tortricidae), under field conditions in South Africa. *African Entomology* **20**, 316–324.

Brown, R. C. (1979). The valid generic and tribal names for the codling moth, Cydia pomonella (Olethreutinae: Tortricidae). Annuals of the Entomological Society of America **72**, 565-567.

Calkins, C.O. (1998). Review of the codling moth area wide suppression program in the Western United States. *Journal of Agricultural Entomology* **15**, 327-333.

Capinera, J. L. (2008). *Encyclopedia of Entomology, Volume 4*. Springer Science & Business Media.

Carde, R. T. & Minks, A. K. (1995). Control of Moth Pests by Mating Disruption: Successes and Constraints. *Annual Review of Entomology* **40**, 559–585.

Carpio, C., Dangles, O., Dupas, S., Léry, X., López-Ferber, M., Orbe, K., Páez, D., Rebaudo, F., Santillán, A. & other authors. **(2013).** Development of a viral biopesticide for the control of the Guatemala potato tuber moth Tecia solanivora. *Journal of invertebrate pathology* **112**, 184–91.

Chambers, C. B. (2014). *Production of Cydia pomonella granulovirus (CpGV) in a heteralogous host, Thaumatotibia leucotreta (Meyrick) (False Codling Moth)*. Rhodes University.

Charleston, D., Kfir, R., van Rensburg, N., Barnes, B., Hattingh, V., Conlong, D., Visser, D. & Prinsloo, G. (2003). Intergrated pest management in South Africa. In *Integrated Pest Management in the Global Arena*, pp. 169–196. Edited by K. M. Maredia, D. Dakouo & D. Mota-Sanchez. CABI.

Ciglar, I. (1998). Integrirana zaŝtita voćaka I vinove loze. Zrinski, Ĉakovec. Claus,

Copping, L. G., Green, M. B. & Rees, R. T. (1992). *Pest Management in Soybean* (L. G. Copping, M. B. Green & R. T. Rees, Eds.). Dordrecht: Springer Netherlands.

Craveiro, S. R., Inglis, P. W., Togawa, R. C., Grynberg, P., Melo, F. L., Ribeiro, Z., Ribeiro, B. M., Báo, S. N. & Castro, M. (2015). The genome sequence of Pseudoplusia includens single nucleopolyhedrovirus and an analysis of p26 gene evolution in the baculoviruses. *BMC Genomics* **16**, 127.

Crook, N. E., James, J. D., Smith, I. R. & Winstanley, D. (1997). Comprehensive physical map of the Cydia pomonella granulovirus genome and sequence analysis of the granulin gene region. *The Journal of general virology* **78**, 965–74.

Crook, N. E., Spencer, R. A., Payne, C. C. & Leisy, D. J. (1985). Variation in Cydia pomonella Granulosis Virus Isolates and Physical Maps of the DNA from Three Variants. *Journal of General Virology* **66**, 2423–2430. Microbiology Society.

Cross, J. V., Solomon, M. G., Chandler, D., Jarrett, P., Richardson, P. N., Winstanley, D., Bathon, H., Huber, J., Keller, B. & other authors. **(1999).** Biocontrol of Pests of Apples and Pears in Northern and Central Europe: 1. Microbial Agents and Nematodes. *Biocontrol Science and Technology* **9**, 125–149.

Cunningham, J.C. (1995). Baculoviruses as microbial pesticides. In Novel approaches to integrated pest management. Edited by R. Reuveni. pp261-292. Boca Raton, FL7 Lewis

Dunley, J. E. & Welter, S. C. (2000). Correlated insecticide cross-resistance in azinphosmethyl resistant codling moth (Lepidoptera: Tortricidae). *Journal of economic entomology* **93**, 955–62.

Eastwell, K. C., Cossentine, J. E. & Bernady, M. G. (1999). Characterisation of Cydia pomonella granulovirus from codling moths in a laboratory colony and in orchards of British Columbia. *Annals of Applied Biology* **134**, 285–291.

Eberle, K. E. & Jehle, J. A. **(2006).** Field resistance of codling moth against Cydia pomonella granulovirus (CpGV) is autosomal and incompletely dominant inherited. *Journal of Invertebrate Pathology* **93**, 201–206.

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. *Journal of Invertebrate Pathology* **98**, 293–298.

Eberle, K. E., Sayed, S., Rezapanah, M., Shojai-Estabragh, S. & Jehle, J. A. (2009). Diversity and evolution of the Cydia pomonella granulovirus. *Journal of General Virology* **90**, 662–671.

Eberle, K. E. (2010). *Novel isolates of Cydia pomonella granulovirus (CpGV): deciphering the molecular mechanism for overcoming CpGV resistance in codling moth (Cydia pomonella)*. University of Mainz.

European Commission. **(2009).** Regulation 1107/2009 concerning the placing of plant protection products on the market. *Official*.

Etzel, L. K. & Falcon, L. A. (1976). Studies of transovum and transstadial transmission of a granulosis virus of the codling moth. *Journal of Invertebrate Pathology* **27**, 13–26.

Falcon, L. A., Kane, W. R. & Bethell, R. S. (1968). Preliminary Evaluation of a Granulosis Virus for Control of the Codling Moth. *Journal of Economic Entomology* **61**, 1208–1213.

Falcon, L. A. & Huber, J. (1991). Biological control of the codling moth. In: *Tortricid Pests, their biology, natural enemies and control.* Edited by L.P.S. van der Geest & H.H. Evenhuis. pp. 355-369. Amsterdam, Elsevier.

Fan, J.B. and Wu, Z. W. (unpublished). *Virulence determinnation for three native Cydia pomonella granulovirus strains and their control effect in field*.

Fauquet, C. M., Mayo, M. a, Maniloff, J., Desselberger, U. & Ball, L. a. **(2005).** *Virus Taxonomy: classification and nomenclature of viruses. Acad Press*. Springer Science & Business Media.

Federici, B.A. (1997). Baculovirus Pathogenesis. In *The Baculoviruses.* Edited by L.K. Miller. pp. 33-59. Plenum Press.

Ferrelli, M.L., Berretta, M.F., Baleich, M.N., Ghiringhelli, P.D., Sciocco-Cap, A. & Romanowski, V. (2012). The Baculoviral genome. In *Viral Genomes - Molecular structure, diversity, gene expression mechanisms and host-virus interactions.* Edited by M.L. Garcia & V. Romanowski. pp. 1-32. InTech

Flint, M. L. & Bosch, R. van den. **(2012).** *Introduction to Integrated Pest Management*. Springer Science & Business Media.

Franck, P., Reyes, M., Olivares, J. & Sauphanor, B. (2007). Genetic architecture in codling moth populations: comparison between microsatellite and insecticide resistance markers. *Molecular Ecology* **16**, 3554–3564.

Fritsch, E., Undorf-spahn, K., Kienzle, J. & Zebitz, C. P. W. (2005). Apfelwickler-Granulovirus : Erste Hinweise auf Unterschiede in der Empfindlichkeit lokaler Apfelwickler-Populationen. *Journal fur Kulturpflanzen* **57**, 2–7.

Fuchs, L. Y., Woods, M. S. & Weaver, R. F. (1983). Viral Transcription During Autographa californica Nuclear Polyhedrosis Virus Infection: a Novel RNA Polymerase Induced in Infected Spodoptera frugiperda Cells. *Journal of virology* **48**, 641–6.

Garcia-Gutierrez, C., Gonzales-Maldonado, M. B., Medrano-Roldan, H. & Chairez-Hernandez, I. (2004). Evaluación de la cepa BbP1 de Beauveria bassiana, Mycotrol®, Meta-Sin® y azinfos-metilico contra *Cydia pomonella* L. (Lepidoptera: Tortricidae) en laboratorio y campo. *Folia Entomol. Mex*., **43**, 1-7.

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. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 15711–6.

Glen, D. M. & Payne, C. C. (1984). Production and field evaluation of codling moth granulosis virus for control of Cydia pomonella in the United Kingdom. *Annals of Applied Biology* **104**, 87–98.

Goto, C., Minobe, Y. & Iizuka, T. (1992). Restriction endonuclease analysis and mapping of the genomes of granulosis viruses isolated from Xestia c-nigrum and five other noctuid species. *The Journal of general virology* **73 (Pt 6)**, 1491–7.

Graillot, B., Berling, M., Blachere-López, C., Siegwart, M., Besse, S. & López-Ferber, M. (2014). Progressive Adaptation of a CpGV Isolate to Codling Moth Populations Resistant to CpGV-M. *Viruses* **6**, 5135–44.

Guarino, L. A., Xu, B., Jin, J. & Dong, W. (1998). A Virus-Encoded RNA Polymerase Purified from Baculovirus-Infected Cells. *J Virol* **72**, 7985–7991.

Gund, N. A., Wagner, A., Timm, A. E., Schulze-Bopp, S., Jehle, J. A., Johannesen, J. & Reineke, A. (2012). Genetic analysis of Cydia pomonella (Lepidoptera: Tortricidae) populations with different levels of sensitivity towards the Cydia pomonella granulovirus (CpGV). *Genetica* **140**, 235–247.

Haas-Stapleton, E. J., Washburn, J. O. & Volkman, L. E. (2004). P74 Mediates Specific Binding of Autographa californica M Nucleopolyhedrovirus Occlusion-Derived Virus to Primary Cellular Targets in the Midgut Epithelia of Heliothis virescens Larvae. *Journal of Virology* **78**, 6786–6791.

Hansen, L. D., Heidt, M. L. & Anderson, P. a. (2006). Bin sterilization to prevent reintroduction of colding moth. *Journal of Agricultural and Urban Entomology* **23**, 17–26.

Harvey, J. P. & Volkman, L. E. (1983). Biochemical and biological variation of Cydia pomonella (codling moth) granulosis virus. *Virology* **124**, 21–34.

Hawtin, R. E., Zarkowska, T., Arnold, K., Thomas, C. J., Gooday, G. W., King, L. A., Kuzio, J. A. & Possee, R. D. (1997). Liquefaction of Autographa californica nucleopolyhedrovirus-infected insects is dependent on the integrity of virus-encoded chitinase and cathepsin genes. *Virology* **238**, 243–53.

Herniou, E. A., Luque, T., Chen, X., Vlak, J. M., Winstanley, D., Cory, J. S. & O'Reilly, D. R. (2001). Use of whole genome sequence data to infer baculovirus phylogeny. *Journal of virology* **75**, 8117–26.

Herniou, E. A., Olszewski, J. A., Cory, J. S. & O'Reilly, D. R. (2003). The genome sequence and evolution of baculoviruses. *Annual Review of Entomology* **48**, 211–234.

Hess, R. T. & Falcon, L. a. (1987). Temporal events in the invasion of the codling moth, Cydia pomonella, by a granulosis virus: An electron microscope study. *Journal of Invertebrate Pathology* **50**, 85–105.

Hoddle, M. S. & Van Driesche, R. G. (2009). Biological control of insect pests. In *Encyclopedia of Insects*, pp. 505–511. Edited by V. H. Resh & R. T. Carde. Academic Press.

Hortgro. **(2014).** Key deciduous fruit statistics 2014.

Hu, Y.-C., Tsai, C.-T., Chang, Y.-J. & Huang, J.-H. (2003). Enhancement and prolongation of baculovirus-mediated expression in mammalian cells: focuses on strategic infection and feeding. *Biotechnology progress* **19**, 373–9.

Huber, J. (1981). Apfelwickler-Granulosevirus: Produktion und Biotests. *Mitteilungen der Deutschen Gesellschaft für Allgemeine und Angewandte Entomologie* **2**.

Huber, J. & Dickler, E. (2009). Das Granulosevirus des Apfelwicklers: Seine Erprobung für die biologische Schädlingsbekämpfung1,2. *Zeitschrift für Angewandte Entomologie* **82**, 143–147.

Hughes, P. R., van Beek, N. A. M. & Wood, H. A. (1986). A modified droplet feeding method for rapid assay of Bacillus thuringiensis and baculoviruses in noctuid larvae. *Journal of Invertebrate Pathology* **48**, 187–192.

Hughes, E., & Shapiro, M. (1997). Viruses. In *Manual of techniques in insect pathology.* Edited by L. Lacey. pp. 40-48. Academic Press.

Hunter-Fujita, F. R., Entwistle, P. F., Evans, H. F. & Crook, N. E. (1998). Insect viruses and pest management. John Wiley & Sons Ltd.

Ignoffo, C. M. (1992). Environmental factors affecting persistence of entomopathogens. *Florida entomologist (USA)* **75**, 516–525.

Ignoffo, C., Hostetter, D., Sikorowski, P., Sutter, G. & Brooks, W. (1977). Inactivation of representative species of entomopathogenic viruses, a bacterium, fungus and protozoan by an ultraviolet light source. *Environmental Entomology* **6**, 411–441.

Jaques, R. P., Laing, J. E., MacLellan, C. R., Proverbs, M. D., Sanford, K. H. & Trottier, R. (1981). Apple orchard tests on the efficacy of the granulosis virus of the codling moth,Laspeyresia pomonella [Lep.: Olethreutidae]. *Entomophaga* **26**, 111–117.

Jaques, R. P., Laing, J. E., Laing, D. R. & Yu, D. S. K. (2012). Effectiveness and persistence of the granulosis virus of the codling moth cydia pomonella (l.) (lepidoptera: olethreutidae) on apple. *The Canadian Entomologist* **119**, 1063–1067.

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. *Archives of Virology* **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. (2008). The future of Cydia pomonella granulovirus in biological control of codling moth. *Problems in Organic Fruit-Growing: Proceedings to the* 265–270.

Jones, K. S. (2000). Bioassays of Entomopathogenic Viruses. In *Bioassays of Entomopathogenic Microbes and Nematodes*, pp. 95–140. Edited by A. Navon & K. R. S. Asche. CABI.

Judd, G. J. R., Gardiner, M. G. T. & Thomson, D. R. (1997). Control of codling moth in organically-managed apple orchards by combining pheromone-mediated mating disruption, post-harvest fruit removal and tree banding. *Entomologia Experimentalis et Applicata* **83**, 137–146.

Jukes, M. D., Knox, C. M., Hill, M. P. & Moore, S. D. (2014). The isolation and genetic characterisation of a South African strain of Phthorimaea operculella granulovirus, PhopGV-SA. *Virus research* **183**, 85–8.

Kalmakoff, J. & Ward. V.K. (2003). Baculoviruses. University of Otago, Dunedin, New Zealand.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S. & other authors. **(2012).** Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)* **28**, 1647–9.

Kienzle, J., Schulz, C., Zebitz, C. P. W. & Huber, J. (2002, October 30). Persistence of the biological effect of codling moth granulovirus in the orchard - preliminary field trials. *10th International Conference on Cultivation Technique and Phytopathological Problems in Organic Fruit-Growing and Viticulture Proceedings to the Conference from 4th to 7th Feb 2002 at Weinsberg/Germany*.

Knight, A. (1994). Insect pest and natural enemy populations in paired organic and conventional apple orchards in the Yakima Valley, Washington. *Journal of the Entomological Society of British Columbia* **91**, 27–36.

Kost, T. A., Condreay, J. P. & Jarvis, D. L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature Biotechnology* **23**, 567–575.

Krappa, R. & Knebel-Mörsdorf, D. (1991). Identification of the very early transcribed baculovirus gene PE-38. *Journal of virology* **65**, 805–12.

Kumar, C. S., Ranga Rao, G. V., Sireesha, K. & Kumar, P. L. (2011). Isolation and characterization of baculoviruses from three major lepidopteran pests in the semi-arid tropics of India. *Indian Journal of Virology* **22**, 29–36.

Lacey, L. & Unruh, T. (2005). Biological control of codling moth (Cydia pomonella, Lepidoptera: Tortricidae) and its role in integrated pest management, with emphasis on entomopathogens. *Vedalia* **12**, 33–60.

Lacey, L. A. (2012). *Manual of Techniques in Invertebrate Pathology*. *Manual of Techniques in Invertebrate Pathology* (L. A. Lacey, Ed.). Academic Press.

Lacey, L. a & Shapiro-Ilan, D. I. (2008). Microbial control of insect pests in temperate orchard systems: potential for incorporation into IPM. *Annual review of entomology* **53**, 121–144.

Lacey, L. A., Unruh, T. R. & Headrick, H. L. (2003). Interactions of two idiobiont parasitoids (Hymenoptera: Ichneumonidae) of codling moth (Lepidoptera: Tortricidae) with the entomopathogenic nematode Steinernema carpocapsae (Rhabditida: Steinernematidae). *Journal of Invertebrate Pathology* **83**, 230–239.

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., Arthurs, S. P., Knight, A. & Huber, J. (2007). Microbial control of Lepidopteran pests in apple orchards. In *Field Manual of Techniques in Invertebrate Pathology: Application and Evaluation of pathogens for Control of Insects and other Invertebrate Pests.* Edited by L.A. Lacey & H.K. pp.528-546. Springer.

Lacey, L. A., Thomson, D., Vincent, C. & Arthurs, S. P. (2008). Codling moth granulovirus: a comprehensive review. *Biocontrol Science and Technology* **18**, 639–663.

Laing, D. R. & Jaques, R. P. (1980). Codling Moth: Techniques for Rearing Larvae and Bioassaying Granulosis Virus. *Journal of Economic Entomology* **73**, 851–853.

Lange, M. & Jehle, J. a. **(2003).** The genome of the Cryptophlebia leucotreta granulovirus. *Virology* **317**, 220–236.

Lange, M., Wang, H., Zhihong, H. & Jehle, J. a. (2004). Towards a molecular identification and classification system of lepidopteran-specific baculoviruses. *Virology* **325**, 36–47.

Lapointe, R., Thumbi, D. & Lucarotti, C. (2012). Recent advances in our knowledge of baculovirus molecular biology and its relevance for the registration of baculovirus-based products for insect pest population control. *Integrated Pest Management and Pest Control - Current and Future Tactics*.

Lazarovits, G., Goettel, M. S. & Vincent, C. (2007). Adventures in Biocontrol. In: *Biological control: a global perspective.* Edited by C. Vincent, M. S. Goettel & G. Lazarovits, G. pp. 1-6 CAB International.

Li, L., Li, Q., Willis, L. G., Erlandson, M., Theilmann, D. A. & Donly, C. (2005). Complete comparative genomic analysis of two field isolates of Mamestra configurata nucleopolyhedrovirus-A. *The Journal of general virology* **86**, 91–105.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L. & Law, M. (2012). Comparison of next-generation sequencing systems. *Journal of biomedicine & biotechnology* **2012**, 251364.

Lounsbury, C. P. (1898). Codling moth. *Agricultural Journal* 13, 597-616.

Luque, T., Finch, R., Crook, N., O'Reilly, D. R. & Winstanley, D. (2001). The complete sequence of the Cydia pomonella granulovirus genome. *Journal of General Virology* **82**, 2531–2547.

Madsen, H. F., Myburgh, a. C., Bosman, J. P. & Rust, D. J. (1974). Codling Moth (Lepidoptera: Olethreutidae): Studies on the Placement of Sex Attractant Traps in South African Orchards. *Phytophylitica* **6**, 189–194.

McWilliam, A. (2007). Environmental impact of baculoviruses. FAO.

Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nature reviews Genetics* **11**, 31–46.

Miele, S. A. B., Garavaglia, M. J., Belaich, M. N. & Ghiringhelli, P. D. (2011). Baculovirus: molecular insights on their diversity and conservation. *International journal of evolutionary biology* **2011**, 379–424.

Mikhaĭlov, V. S. (2003). Replication of the baculovirus genome. *Molecular Biology* **37**, 250–259.

Miller, L. K. (1996). Insect viruses. In: *Fundamental Virology*. Edited by Fields, B.N., Knipe, D. M., Howley, P.M., Chanock, R. M., Melnick, J. L., Monath, T. P., Roizman, B. & Straus, S. E. pp. 401-422. Lippincott - Raven Publishers.

Miller, L. K. (2013). *The Baculoviruses*. Springer Science & Business Media.

Mochizuki, T., Sako, Y. & Prangishvili, D. (2011). Provirus induction in hyperthermophilic archaea: characterization of Aeropyrum pernix spindle-shaped virus 1 and Aeropyrum pernix ovoid virus 1. *Journal of Bacteriology* **193**, 5412–5419.

Moore, S. D. (2002). The development and evaluation of *Cryptophlebia leucotreta* granulovirus (CrleGV) as a biological control agent for the management of false codling moth, *Cryptophlebia leucotret*a on citrus. Rhodes University.

Moore, S. D., Kirkman, W., Richards, G. I. & Stephen, P. R. (2015). The Cryptophlebia leucotreta granulovirus-10 years of commercial field use. *Viruses* **7**, 1284–312.

Moscardi, F. (1999). Assessment of the application of baculoviruses for control of Lepidoptera. *Annual review of entomology* **44**, 257–289.

Myburgh, a. C., Mad'Sen(1), H. F., Bosman, J. P. & Rust, D. J. (n.d.). Codling Moth (Lepidoptera: Olethreutidae): Studies on the Placement of Sex Attractant Traps in South African Orchards.

Myers, J. H., Savoie, A. & van Randen, E. (1998). Eradication and pest management. *Annual review of entomology* **43**, 471–491.

Myers, J. H., Simberloff, D., Kuris, A. M. & Carey, J. R. (2000). Eradication revisited: dealing with **O'Reilly, D. R., Crawford, A. M. & Miller, L. K. (1989).** Viral proliferating cell nuclear antigen. *Nature* **337**, 606.

van Oers, M. & Vlak, J. (2007). Baculovirus Genomics. *Current Drug Targets* **8**, 1051–1068.

Ogembo, J. G., Caoili, B. L., Shikata, M., Chaeychomsri, S., Kobayashi, M. & Ikeda, M. (2009). Comparative genomic sequence analysis of novel Helicoverpa armigera nucleopolyhedrovirus (NPV) isolated from Kenya and three other previously sequenced Helicoverpa spp. NPVs. *Virus genes* **39**, 261–72.

Opoku-Debrah, J. K., Hill, M. P., Knox, C. & Moore, S. D. (2013). Overcrowding of false codling moth, Thaumatotibia leucotreta (Meyrick) leads to the isolation of five new Cryptophlebia leucotreta granulovirus (CrleGV-SA) isolates. *Journal of Invertebrate Pathology* **112**, 219–228.

Pajač, I., Pejić, I. & Barić, B. (2011). Codling Moth, Cydia pomonella (Lepidoptera: Tortricidae)–Major Pest in Apple Production: an Overview of its Biology, Resistance, Genetic Structure and Control Strategies. *Agriculturae Conspectus Scientificus (ACS)* **76**, 87–92.

Possee, R.D. & Rohrmann, G.F. (1997). Baculovirus genome organization and evolution. In *The Baculoviruses.* Edited by Miller LK. Plenum Press.

Pringle, K. L. (2006). The use of economic thresholds in pest management: apples in South Africa. *South African Journal of Science* **102**, 201–204.

Pringle, K. L., Eyles, D. K. & Brown, L. (2003). Trends in codling moth activity in apple orchards under mating disruption using pheromones in the Elgin area, Western Cape Province, South Africa. *African Entomology* **11**, 65–75.

Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., Bertoni, A., Swerdlow, H. P. & Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics* **13**, 341.

Quinn, L. P., Vos, B. J. de, Fernandes-Whaley, M., Roos, C., Bouwman, H., Kylin, H., Pieters, R. & Berg, J. van den. **(2011).** Pesticides in the Modern World - Pesticides Use and Management. In *Pesticides in the Modern World - Pesticides Use and Management*, pp. 49–87. Edited by M. Stoytcheva. InTech.

Reuveny, H. & Cohen, E. (2004). Resistance of the codling moth Cydia pomonella (L.) (Lep., Tortricidae) to pesticides in Israel. *Journal of Applied Entomology* **128**, 645–651.

Reyes, M., Franck, P., Charmillot, P.-J., Ioriatti, C., Olivares, J., Pasqualini, E. & Sauphanor, B. (2007). Diversity of insecticide resistance mechanisms and spectrum in European populations of the codling moth, Cydia pomonella. *Pest management science* **63**, 890–902.

Rezapanah, M., Shojai-Estabragh, S., Huber, J. & Jehle, J. a. (2008). Molecular and biological characterization of new isolates of Cydia pomonella granulovirus from Iran. *Journal of Pest Science* **81**, 187–191.

Rezapanah, M. R., Kharazi Pakdel, A., Kamali, K. & Huber, J. (2002). Survey on Natural Occurrence of Cydia Pomonella Granulovirus in Apple Orchards of Iran. *Applied entomology and phytopathology* **69**, 49–56.

Riedl, H. (1983). Analysis of codling moth phenology in relation to latitude, climate and food availability. In *Diapause and Life Cycle Strategies in Insects*. Edited by Brown, V. K. & Hodek, I. pp. 233-252. The Hague, Dr W Junk Publishers.

Riedl, H., Blomefield, T. L. & Giliomee, J. H. (1998). A century of codling moth control in South Africa: II. Current and future status of codling moth management. *Journal of the Southern African Society of Horticultural Sciences* **8**, 32–54.

Road, M. (1991). Comparative testing of disinfectant and antiseptic products using proposed European suspension testing methods. *Letters in Applied Microbiology* **13**, 233–237.

Rodriguez, V., Belaich, M. & Ghiringhelli, P. (2012). Baculoviruses : Members of Integrated Pest Management strategies. *Integrated Pest Management and Pest Control - Current and Future Tactics* 463–480.

Rohrmann, G. F. (2013). *Baculovirus Molecular Biology*. National Library of Medicine (US), National Center for Biotechnology Information.

Rubinstein, R. & Poison, A. (1983). Midgut and Viral Associated Proteases of *Heliothis armigera*. *Intervirology* **19**, 16–25. Karger Publishers.

Sauphanor, B., Berling, M., Toubon, J. F., Reyes, M., Delnatte, J. & Allemoz, P. (2006). Carpocapse des pommes. Cas de résistance au virus de la granulose en vergers biologique. *Phytoma - La Défense des Végétaux* **590**, 24 – 27.

Schmitt, a., Bisutti, I. L., Ladurner, E., Benuzzi, M., Sauphanor, B., Kienzle, J., Zingg, D., Undorf-Spahn, K., Fritsch, E. & other authors. **(2013).** The occurrence and distribution of resistance of codling moth to Cydia pomonella granulovirus in Europe. *Journal of Applied Entomology* **137**, 641–649.

Setyobudi, L. (1989). *Seasonality of Codling Moth, Cydia Pomonella L. (Lepidoptera: Olethreutidae) in the Willamette Valley of Oregon: Role of Photoperiod and Temperature*. Oregon State University.

Shapiro-Ilan, D. I., Fuxa, J. R., Lacey, L. A., Onstad, D. W. & Kaya, H. K. (2005). Definitions of pathogenicity and virulence in invertebrate pathology. *Journal of invertebrate pathology* **88**, 1–7.

Sporleder, M., Kroschel, J., Huber, J. & Lagnaoui, A. (2005). An improved method to determine the biological activity (LC50) of the granulovirus PoGV in its host Phthorimaea operculella. *Entomologia Experimentalis et Applicata* **116**, 191–197.

Steineke, S. B. & Jehle, J. A. (2004). Investigating the horizontal transmission of the Cydia pomonella granulovirus (CpGV) in a model system. *Biological Control* **30**, 538–545.

Stern, V. M. (1973). Economic thresholds. Annual review of Entomology. **18,** 259-280.

Swezey, S. L. (2000). *Organic Apple Production Manual*. UCANR Publications.

Szewczyk, B., Hoyos-Carvajal, L., Paluszek, M., Skrzecz, I. & Lobo de Souza, M. (2006). Baculoviruses — re-emerging biopesticides. *Biotechnology Advances* **24**, 143–160.

Szewczyk, B., Lobo de Souza, M., Batista de Castro, M.E., Moscardi, M.L. & Moscardi, F. (2011). Baculovirus Biopesticides. In: *Pesticides - Formulations, Effects, Fate.* Edited by Stoytcheva M. pp. 21-36. InTech.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* **28**, 2731–9.

Tanada, Y. (1964). A granulosis virus of the codling moth, *Carpocapsae pomonella*. *Journal of Insect Pathology* **6**, 378 – 380.

Tanada, Y. & Leutenegger, R. (1968). Histopathology of a granulosis-virus disease of the codling moth, Carpocapsa pomonella. *Journal of Invertebrate Pathology* **10**, 39–47.

Tanada, Y. & Hess, R. T. (1991). Baculoviridae Granulosis viruses. In Atlas of invertebrate viruses. Edited by J.R. Adams & J.R. Bonami. pp. 227-257. CRC Press

Tanada, Y. & Kaya, H. (1993). *Insect Pathology*, Academic press: 666

Thaler, R., Brandstätter, A., Meraner, A., Chabicovski, M., Parson, W., Zelger, R., Dalla Via, J. & Dallinger, R. (2008). Molecular phylogeny and population structure of the codling moth (Cydia pomonella) in Central Europe: II. AFLP analysis reflects human-aided local adaptation of a global pest species. *Molecular phylogenetics and evolution* **48**, 838–49.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**, 4673–80.

Van Ark, H. (1995). *Introduction to the analysis of quantal response*. Agricultural Research Council Agrimetrics Institute: Pretoria. 78 pp.

Vincent, C., Goettel, M. S. & Lazarovits, G. (2007). Adventures in biocontrol. In *Biological control - A global perspective: Case studies from around the world*, pp. 1–6. Edited by C. Vincent, M. Geottel & G. Lazarovits. CABI.

Voudouris, C. C., Sauphanor, B., Franck, P., Reyes, M., Mamuris, Z., Tsitsipis, J. A., Vontas, J. & Margaritopoulos, J. T. (2011). Insecticide resistance status of the codling moth Cydia pomonella (Lepidoptera: Tortricidae) from Greece. *Pesticide Biochemistry and Physiology* **100**, 229–238.

de Waal, J. Y., Malan, a P., Levings, J. & Addison, M. F. (2010). Key elements in the successful control of diapausing codling moth, Cydia pomonella (Lepidoptera: Tortricidae) in wooden fruit bins with a South African isolate of Heterorhabditis zealandica (Rhabditida: Heterorhabditidae). *Biocontrol Science and Technology* **20**, 489–502.

Wahner, N. (2008, December 1). Initial investigation of Trichogrammatoidea lutea (Hymenoptera: Trichogrammatidae) as biological control agent of codling moth, Cydia pomonella (Lepidoptera: Tortricidae), in apple and pear orchards, under sterile insect technique (SIT). Stellenbosch : Stellenbosch University.

Washburn, J. O., Chan, E. Y., Volkman, L. E., Aumiller, J. J. & Jarvis, D. L. (2003). Early Synthesis of Multicapsid Nucleopolyhedrovirus Virulence in Orally Infected Heliothis virescens. *Journal of Virology* **77**, 280–290.

Wearing, C. ., Hansen, J. ., Whyte, C., Miller, C. . & Brown, J. (2001). The potential for spread of codling moth (Lepidoptera: Tortricidae) via commercial sweet cherry fruit: a critical review and risk assessment. *Crop Protection* **20**, 465–488.

Wearing, C. (1979). Integrated control of apple pests in New Zealand 10. Population dynamics of codling moth in Nelson. *New Zealand Journal of Zoology* **6**, 165–199.

Welter, S. (2009). Codling moth. In *Encyclopedia of Insects*, pp. 174–175. Edited by V. H. Resh & R. T. Carde. Academic Press.

Winston, M. L. (1999). *Nature Wars: People Vs. Pests*. Harvard University Press.

Witzgall, P., Stelinski, L., Gut, L. & Thomson, D. (2008). Codling moth management and chemical ecology. *Annual review of entomology* **53**, 503–22. Annual Reviews.

Xu, Y.-P., Cheng, R.-L., Xi, Y. & Zhang, C.-X. (2013). Genomic diversity of Bombyx mori nucleopolyhedrovirus strains. *Genomics* **102**, 63–71.

Yan, F., Bengtsson, M. & Witzgall, P. (n.d.). Behavioral Response of Female Codling Moths, Cydia pomonella, to Apple Volatiles. *Journal of Chemical Ecology* **25**, 1343–1351.

Zhang, C.-X., Ma, X.-C. & Guo, Z.-J. (2005). Comparison of the complete genome sequence between C1 and G4 isolates of the Helicoverpa armigera single nucleocapsid nucleopolyhedrovirus. *Virology* **333**, 190–9.

Zhang, X., Liang, Z., Yin, X., Wang, J. & Shao, X. (2014). Complete genome sequence of Agrotis segetum granulovirus Shanghai strain. *Archives of virology* **159**, 1869–72.

Zhu, Z., Yin, F., Liu, X., Hou, D., Wang, J., Zhang, L., Arif, B., Wang, H., Deng, F. & Hu, Z. (2014). Genome Sequence and Analysis of Buzura suppressaria Nucleopolyhedrovirus: A Group II Alphabaculovirus. *PloS one* **9**, e86450.

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*.

Zimmermann, G., Huger, A. & Kleespies, R. (2013). Occurrence and Prevalence of Insect Pathogens in Populations of the Codling Moth, Cydia pomonella L.: A Long-Term Diagnostic Survey. *Insects* **4**, 425–446.

Zingg, D. (2008, July 24). Madex Plus and Madex I12 overcome Virus Resistance of Codling Moth. *Ecofruit - 13th International Conference on Cultivation Technique and Phytopathological Problems in Organic Fruit-Growing: Proceedings to the Conference from 18thFebruary to 20th February 2008 at Weinsberg/Germany*.

Zwart, M. P., van Oers, M. M., Cory, J. S., van Lent, J. W. M., van der Werf, W. & Vlak, J. M. (2008). Development of a quantitative real-time PCR for determination of genotype frequencies for studies in baculovirus population biology. *J Virol Methods* **148**, 146–54