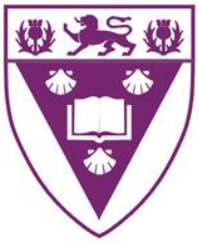


RHODES UNIVERSITY

Where leaders learn

Genetic characterisation of a range of geographically distinct *Helicoverpa armigera* nucleopolyhedrovirus (HearNPV) isolates and evaluation of biological activity against South African populations of the African bollworm, *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae)

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

DOCTOR OF PHILOSOPHY

By

KUDZAI TAPIWANASHE ESAU MTAMBANENGWE

Abstract

The African bollworm, *Helicoverpa armigera* Hübner (Lepidoptera: Noctuidae) is a pest of economic and agricultural importance globally. It is a polyphagous pest that feeds on a wide range of host plants including economically important crops. The impact it has on agricultural systems makes its control a priority. The most common method of control is using chemical pesticides; however, continuous application of the pesticides has resulted in the development of resistance. The use of biological control has been investigated and established as an effective method of control as a standalone or part of an integrated pest management (IPM) system. The use of the baculovirus Helicoverpa armigera nucleopolyhedrovirus (HearNPV), has shown promise in the control of *H. armigera*. Commercial formulations based on the virus are available in many global markets. However, the identification of novel HearNPV isolates will aid in the control of *H. armigera* as well as provide alternative isolates that may have better virulence.

Three new HearNPV isolates were purified and identified from three distinct geographical South African locations *H. armigera* cadavers and named HearNPV-Albany, HearNPV-KZN and HearNPV-Haygrove. The genomes of two of the HearNPV isolates, namely HearNPV-Albany and HearNPV-KZN were genetically characterised and compared to other geographically distinct HearNPV isolates. Virulence studies were performed comparing the new HearNPV isolates against established commercial HearNPV formulations, Helicovir™ and Helicovex® and other geographically distinct isolated HearNPV, HearNPV-G4 and HearNPV-SP1.

Two laboratory colonies were established using *H. armigera* collected from South African fields in the Belmont Valley near Grahamstown labelled as Albany colony and a colony provided from Haygrove Eden farm near George labelled as Haygrove colony. Biological studies were carried out using the Albany *H. armigera* colony comparing the rate of development, survival and fertility on bell green peppers, cabbage leaves and on artificial diet. From the biological studies, it was recorded that development and survivorship was best on artificial diet. Regular quality control was required for the maintenance of the colony and continuous generations of healthy larvae were eventually established.

Diseased cadavers with signs of baculovirus infection were collected after bioprospecting from the Kwa-Zulu Natal Province in South Africa and were labelled KZN isolate; Belmont Valley near Grahamstown and were labelled Albany isolate; and Haygrove Eden farm near

George and were labelled Haygrove isolate for the study. A fourth isolate made up of a crude extract of occlusion bodies (OBs) first described by Whitlock was also analysed and labelled Whitlock isolate. Occlusion bodies were extracted, purified and morphologically identified from the KZN, Albany, Haygrove and Whitlock isolates using TEM. Genomic DNA, which was extracted from the purified OBs. Using PCR, the identity of the OBs as HearNPV was confirmed.

Genomic analyses were performed on HearNPV-Albany and HearNPV-KZN through genetic characterisation and comparison with other geographically distinct HearNPV genomes to confirm novelty and establish potential genetic relationships between the isolates through evolutionary distances. Full genomic sequencing of the isolated HearNPV and comparison with other geographically distinct HearNPV isolates identified genomic differences that showed that the HearNPV isolates were novel. HearNPV-Albany and HearNPV-KZN were successfully sequenced and identified as novel isolates with unique fragment patterns and unique gene sequences through deletions or insertions when compared to other geographically distinct HearNPV. This raised the potential for differences in biological activity against *H. armigera* larvae when tested through biological assays. HearNPV-Whit genome assembly had low quality data which resulted in many gaps and failed assembly.

The biological activity of HearNPV isolates from Spain, China, South Africa and two commercial formulations were studied against the laboratory established *H. armigera* South African colony. The LC_{50} values of the different South African HearNPV isolates were established to be between 7.7×10^1 OBs.ml^{-1} for the most effective and 3.2×10^2 OBs.ml^{-1} for the least effective. The Spanish and Chinese HearNPV isolates resulted in LC_{50} values of 2.0×10^2 OBs.ml^{-1} and 1.2×10^1 OBs.ml^{-1} respectively. The commercial formulations resulted in the least virulence observed with an LC_{50} of 5.84×10^2 OBs.ml^{-1} and 9.0×10^2 OBs.ml^{-1} for Helicovex® and Helicovir™ respectively.

In this study, novel South African HearNPV isolates were isolated and identified. Through characterisation and bioassays against South African *H. armigera* populations the HearNPV isolates were shown to have different virulence in comparison to geographically distinct isolates. From this research, there is potential for development of new *H. armigera* biopesticides based on the novel isolates after field trial testing.

Table of contents

List of figures

List of tables

List of abbreviations

$ - dollar

22°14’21” E – 22 degrees 14 minutes 21 seconds east of GMT

33°25’22” S – 33 degrees 25 minutes 22 seconds south of the equator

A – adenine

AGE – agarose gel electrophoresis

alk-exo – alkaline exonuclease

ANOVA – analysis of variance

ATP – adenosine tri-phosphate

BLAST – basic local alignment search tool

bp – base pairs

bro – baculovirus repeat ORF

Bt – Bacillus thuringiensis

BV – budded virus

C – cytosine

CE – controlled environment

cm – centimetres

CTAB – cetrimonium bromide

dbp – DNA binding protein

ddH_2O – distilled deionised water

DDT – dichlorodiphenyltrichloroethane

df – degrees of freedom

DNA – deoxy-ribonucleic acid

DNApol – DNA polymerase

Dr. – doctor

egt – ecdysteroid UDP-glucosyl transferase

FAO – food and agricultural organisation

fgf – fibroblast growth factor

fp – fusion protein

G – guanine

gp – glycoprotein

gran - granulin

GV – granulovirus

HearNPV – *Helicoverpa armigera* nucleopolyhedrovirus

hrs – homologous repeat regions

iap – inhibitor of apoptosis

ie – immediate early

IPM – integrated pest management

kDa – kilodaltons

kg. ha^{-1} – kilograms per hectare

km – kilometres

KZN – Kwa-Zulu Natal

L × W × H – length × width × height

L:D – light: day

LC_{50} – lethal concentration that results in mortality of 50% of the population

LD_{50} – lethal dose that results in mortality of 50% of the population

lef – late expression factor

LT_{50} – lethal time needed for the mortality of 50% of the population

mg – milligrams

mm – millimetres

MNPV – multiple-nucleocapsid nucleopolyhedrovirus

NCBI – national centre for biotechnology information

ng – nanograms

nm – nanometres

NPV – nucleopolyhedrovirus

nt – nucleotides

OBs – occlusion bodies

ODV – occlusion derived virus

ODV-e – occlusion derived virus envelope

orf – open reading frame

PCR – polymerase chain reaction

PIFs – *per os* infectivity factors

polh -polyhedrin

RNA – ribonucleic acid

SDS – sodium dodecyl sulphate

SE – standard error

SNP – single nucleotide polymorphism

SNPV – single-nucleocapsid nucleopolyhedrovirus

T – thymine

TEM – transmission electron microscope

UK – United Kingdom

US – United States

USA – United States of America

USD – United States Dollar

UV – ultraviolet

V – volts

$^{v}/_{v}$ – volume per volume

vlf – very late expression factor

vp – virus protein

$^{w}/_{v}$ – weight per volume

μm - micrometres

Acknowledgements

My deepest gratitude goes to my supervisor professor C. Knox, for her unending patience and assistance in the research conducted for this study. Her wisdom, knowledge, exceptional supervision and belief in her students goes beyond the expected norm. Without her, this dissertation and the research would never have been possible. I would also like to extend my gratitude to my co-supervisor professor M. P. Hill. The broad knowledge in entomology, the exposure to entomological theories and practices was very rewarding. The expertise, patience and unending support during this period helped during periods where it seemed there was no light at the end of the tunnel. I would also like to thank my co-supervisor doctor S. Moore. He brought unimaginable industrial expertise to the research, assisting in bridging the gap between the laboratory research and its role in the industry. The role played by him cannot be expressed in words and for that I am deeply grateful.

Gratitude is also extended to Rhodes University, for providing the environment enough for continued university education, from undergraduate all the way up to postgraduate studies. The academic and postgraduate environment provided for by Rhodes University can never be underestimated. Sincere gratitude to Melissa Peyper who was with me during the bollworm scouting and collections, was involved in helping during the rearing process and always made sure I had enough feed for the bollworm. Gratitude to the manager, Mr. Ferguson, for allowing us access to the farm for live bollworm collections. Special thanks to Dr. Michael Jukes, B. Motsoeneng, Nicole, Melissa, Onosi and Marcel, for their continued assistance in the laboratory and being part of my research family.

I would like to thank Rhodes University and NRF for funding this project.

Funding for this work was provided by the South African Research Chairs Initiative of the Department of Science and Technology and the National Research Foundation of South Africa. Any opinion, finding, conclusion or recommendation expressed in this material is that of the authors and the NRF does not accept any liability in this regard.

To T. Mutsvairo, Mr N. Mtambanengwe, Prof. F. Mtambanengwe, K. Mtambanengwe and the rest of the Mtambanengwe family that were with me during the research journey, thank you for your continued support, assistance and understanding.

1 Literature review and project motivation

1.1 An introduction to *Helicoverpa armigera*

Helicoverpa armigera Hübner (Lepidoptera, Noctuidae) is a stout-bodied, brown, nocturnal moth (Karel and Autrique, 1989). It is indigenous to Africa and is a major agricultural problem (Cherry *et al*., 2003). It was previously also known as *Heliothis armigera* (Hübner), *Chloridea armigera* (Hübner), *Heliothis obsoleta* (Auctorum), and bears the common names, old world (African) bollworm, corn earworm and cotton bollworm (Lammers and McLeod, 2007). There are four heliothine species that are considered to be of economic importance in Africa, namely *H. armigera*, *H. assulta afra* (Hardwick), *H. fletcheri* (Hardwick) and *H. peltigera* (Schiffermüller) and of these four species, only *H. armigera* is considered of major economic importance (Cherry *et al*., 2003). It is a lepidopteran insect and the larvae have biting and chewing type mouthparts with well-developed mandibles, which facilitate their voracious feeding on the plant and crop foliage (Sree and Varma, 2015). It is important to understand *H. armigera* as a pest and the methods involved in controlling it. This is because although many pests have been known to cause damage to crop in Africa, *H. armigera* has a high propensity to cause sustainable damage and crop losses to a range of crops (Kranthi and Russell, 2009). This review describes the pest, its impact on agriculture and methods of control globally, with special focus on South Africa where it has a devastating effect on crop production (Bell and McGeoch, 1996).

Host range of Helicoverpa armigera

Helicoverpa armigera is a polyphagous pest that is often found on vegetable crops, cereals and fruit plants (Figure 1.1) (Vaissayre and Caiquil, 2000). It has been recorded on 182 plant species with most of the plants considered to be of agricultural and economic importance (Kranthi and Russell, 2009). A preference for host plant species has been observed and the preference is dependent on the temporal and spatial availability of the hosts at the feeding stage of development. Other host factors that influence *H. armigera* attack, besides plant species, include plant height and plant physiological state (Haase *et al*., 2015). A hierarchy in host choice has been observed and the most important damaged plants are grain sorghum, corn, cotton, tobacco, tomatoes, potatoes, flax, soya, and beans (Smith-Pardo, 2014).

Figure 1.1: The larvae of *H. armigera* feeding on different host plants. a: larva feeding on the flesh in the interior of green bell pepper. b: larva feeding on the surface of a kale leaf with leaf incision visible. c: first instar larva feeding on the surface of cabbage leaves. d: larva feeding on wheat (Reddy *et al*., 2015).

1.2 Global distribution of *Helicoverpa armigera*

The distribution of *H. armigera* is widespread, being found in tropical, subtropical and warm temperature regions (Cherry *et al.*, 2003). *Helicoverpa armigera* infestations have been reported in Pakistan, India, Central Asia, southeastern Asia, Africa, Middle East, southern Europe, eastern and northern Australia, New Zealand and many eastern Pacific islands (Figure 1.2) (Karim, 2000). In Europe it is widely present in Greece, Portugal and Spain and is less abundant in Austria, Czech Repubulic, France, Germany, Hungary, Italy and Lithuania (Lammers and McLeod, 2007).

The ability of *H. armigera* to move long distances, allows spatial redistribution by migration to different areas and different plant hosts (Karim, 2000). Its migration patterns have been reported to be up to 1,000 km (Lammers and McLeod, 2007). It was reported as a quarantine

pest in the Americas but recent studies have detected its presence in Brazil, Paraguay and Argentina, with two new *H. armigera* populations being confirmed in Brazil (Tay *et al*., 2013; Haase *et al.*, 2015). In India prior to 1980, *H. armigera* was not considered as a significant crop pest, but this changed in the mid to late 1980s when pesticide resistance to synthetic pyrethroids was reported (Kranthi and Russell, 2009). The intercropping of cotton and maize in many African farming settings exacerbates the transfer of *H. armigera* from maize to cotton (Russell, 2004). *Helicoverpa armigera* is distributed in West, Central, East and southern Africa because of the tropical climate (Cilas *et al.*, 2015). It is considered as the number one lepidopteran pest of agricultural crops in Africa (Cherry *et al*., 2003). *Helicoverpa armigera* has the highest pest status of the Lepidoptera in South Africa with a pest status value of 233 based on entries in South African journals, which is four times as much as the second highest pest, *Agrotis segetum* (Dennis and Schiffermüller) (Bell and McGeoch, 1996).

Figure 1.2: Global distribution of *H. armigera* with orange stars highlighting countries where it is considered an important economic agricultural pest (Kriticos *et al.*, 2016).

1.3 Global economic importance of *Helicoverpa armigera* and damage to host plants

Helicoverpa armigera has high reproduction rates, short generation times and a high mobility. These qualities result in continuous generations that can move to different host

plants (Cherry *et al*., 2003). The larvae destroy the terminal buds and flowers before attacking the fruiting body. The larvae are very aggressive and will outcompete other lepidopteran species that may be feeding on the host plant (Vaissayre and Cauquil, 2000). The incidence of host plant attacks is unpredictable due to the dispersive and migrational attributes of *H. armigera* and damage of host plants varies in severity between crops as well as regions on a temporal scale (Gouse, 2013). The larvae are mainly fruit feeders, although they cause major damage to the leaves of plants such as tobacco and potatoes (Karim, 2000). They ingest large amounts of food during the larval stage and sometimes even feed on smaller insects as well as being cannibalistic. The attack is mainly focused on the harvestable and fruiting parts of the host plants such as leaves, flowers, buds, developing pods, fruits and seed (Figure 1.3) (Cherry *et al*., 2003). The early instar larvae feed on the flowering parts and the young pods by making clean spherical holes. The main damage on crop plants is caused by older larvae, mainly from the third instars that burrow into the flowering parts and fruit (Karel and Autrique, 1989).

The damage to the crops results in high economic costs and losses as well as high socio-economic impact on subsistence farming (Gouse, 2013). If left uncontrolled, severe crop losses can occur and it is also important to note that in some instances the cost of control may exceed crop value (Cherry *et al*., 2003). The control of *H. armigera* accounts for approximately 30% of cotton production cost. Nearly 25% of the global use of chemical pesticides in farming is used in the control of agricultural pests on cotton and *H. armigera* is the main cotton pest (Gouse, 2013). Between 1990 and 2000, Indian farmers were using approximately 43% of their production costs in the control of pests, with 80% of the control dedicated to *H. armigera*. Losses of up to USD 100 million were reported in some Indian farming regions (Kranthi and Russell, 2009). If left uncontrolled *H. armigera* can damage 50–100% of the cotton yield (Reed and Pawar, 1982; Sehgal and Ujagir, 1990; Gouse, 2013). It is considered the most significant pest of cotton in Australia, China and India (Gordh and Headrick, 2011). Between 2011 and 2013, an estimated USD 10 billion in economic losses in Brazil were attributed to *Helicoverpa* species (Smith-Pardo, 2014). It has also been reported to account for yield losses of between 10 and 50% in beans in Ethiopia, Kenya and Tanzania (Oluoch-Kasura *et al*., 2013). In studies conducted in Ivory Coast between 1978 and 1983, 60% of cotton crop losses were attributed primarily to *H. armigera* damage. In Zimbabwe significant cotton losses were recorded and in Tanzania losses were estimated at about USD

20 million (Smith-Pardo, 2014). In South Africa it the most important plant and crop pest and is a major restriction to the production of food and crops (Moran, 1983).

Figure 1.3: Damage to different host crops through extensive feeding by *H. armigera* larvae. (images captured in the field)

1.4 Life cycle of *Helicoverpa armigera* and morphological identification

The life cycle of *H. armigera* has been observed to be linked with environmental and climatic conditions, with up to 10 generations recorded under optimal tropical conditions. When conditions are harsh, the caterpillar and the pupae enter diapause, which results in the lengthening of the life cycle (Cilas *et al*., 2015). The lifecycle involves the laying of eggs by adult moths singly or in groups on young leaves, flowers and shoots. The eggs hatch in 3-8 days and the larval lifecycle is between 15-30 days. Pupation occurs in the soil and requires 10-25 days (Amphofo, 1994).

The eggs of *H. armigera* are white, turning yellow to brownish just before hatching. The eggs are spherical in shape and 0.5 mm in diameter. The area surrounding the micropyle is smooth and the rest of the egg surface is covered in longitudinal ribs (Smith-Pardo, 2014). At optimum temperatures the eggs generally hatch within 4 days (Hill, 1987). Adult females lay on average 450 eggs and a female moth lays between 150-1500 eggs during its lifecycle (Karim, 2000; Gordh and Headrick, 2011).

The larval stage has six instars and lasts 14 to 30 days during summer temperatures of 25-37°C but can be up to 50 days at low temperatures (Karel and Autrique, 1989; Mironidis and Savapoulou-Soultani, 2008; Gordh and Headrick, 2011). First and second instar larvae are pale in colour usually yellowish-white to reddish brown and lack permanent defining marks and features. A colour pattern develops in subsequent instars in different shades including green, yellow, pink, red, brown and black (Smith-Pardo, 2014). Larger larvae have a pale white longitudinal band against darker bands on each side of the body (Figure 1.4), with fully grown larvae growing up to 4 cm long (Karel and Autrique, 1989). Cannibalism has been recorded in third instar larvae, with reports of older larvae travelling to different parts of the plant to attack other larvae (Karim, 2000).

Figure 1.4: Newly moulted fourth instar *H. armigera* captured under a dissecting microscope with longitudinal band visible (Own photograph).

The final instar usually burrows into the soil, where pupation commences, and this occurs close to the base of the plant. Pupae are brown, averaging 12-18 mm in length, with smooth surfaces rounded at both ends with tapering parallel spines at the posterior tips (Smith-Pardo, 2014). When temperatures are optimal and suitable (between 25 and 37°C), maturation of pupae takes approximately 16 days, and when temperatures are too cold (below 20°C), the pupae enter diapause and may spend winter as pupae (Hackett and Gatehouse, 1982; Lu and

Xu, 2010). The life cycle can be completed within 8 weeks if conditions are favourable (Figure 1.5), resulting in continuous generations throughout the year (Hill, 1987).

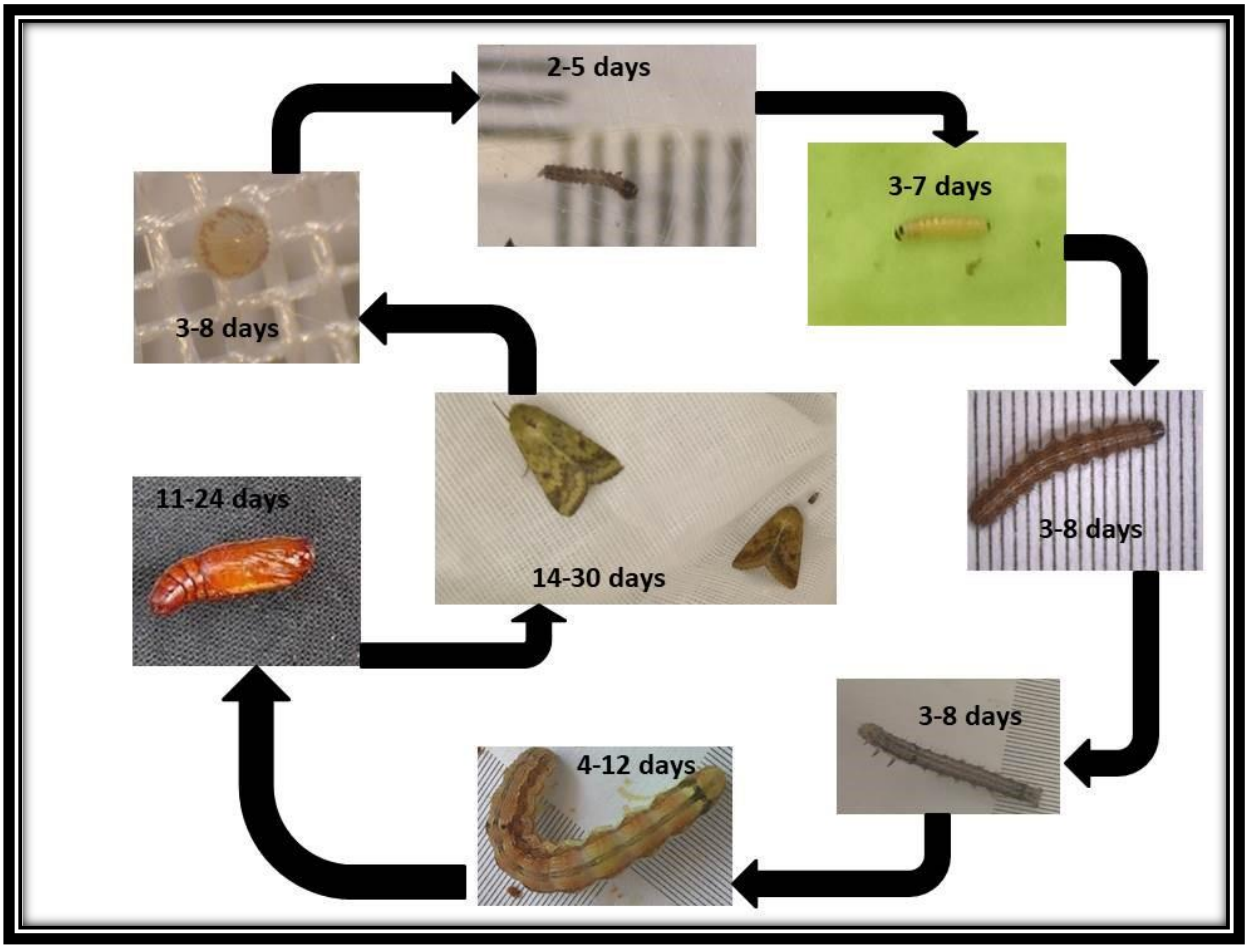


Figure 1.5: Different life stages of *H. armigera* captured during biological studies from egg to adult with directional arrows showing the progress expected to be observed (Own photograph).

The adult moths are mainly nocturnal, with a noctuid body shape, a wingspan of 30-45 mm and a body length that averages 12-20 mm. The females are brownish, and the males are lighter, appearing light greenish/yellowish (Figure 1.5) and a line of blackish dots is visible on the margins of the wings with a broad, irregular transverse band (Smith-Pardo, 2014). Eclosion of the adult moths from pupae occurs at night, and the adults dry their wings by crawling up vertical substrates. After four days, the females start releasing pheromones which trigger mating. The adult female moths are highly fecund, with fecundity being dependent on temperature, humidity and nutrition of the adult moths (Karim, 2000).

1.5 The control of *Helicoverpa armigera*

The control of *H. armigera* is of major importance, as without pest regulation, yield losses would be very high. To control *H. armigera*, natural control (abiotic and biotic) and applied control can be used. Applied control includes physical, cultural, biological, genetic, chemical and integrated pest management measures (Oluoch-Kasura *et al*., 2013). The control of *H. armigera* is required throughout the season, if the crop is susceptible, which often disrupts natural enemies and leads to outbreaks of secondary pests (Peshin *et al*., 2009).

1.5.1 Integrated pest management and its role in the control of *Helicoverpa armigera*

Integrated pest management (IPM) is a long-term, science-based decision-making process that identifies and reduces risks from pests and pest management strategies (Mitchell and Hutchison, 2009). IPM involves the consideration and utilisation of all suitable techniques and methods, including natural and chemical control to maintain pest populations at levels below those causing economic effects (Kogan, 1998). The strategy aims at the reduction or elimination of chemical pesticides which have been put under stricter control by governments because of their effects on the ecosystem and the general consumer green movement (Haase *et al*., 2015). IPM consists of a multicomponent approach to the management of pests, through the selection of required methods and decisions for the most effective application to defend crop plants from pests (Kranthi and Russell, 2009). IPM coordinates the use of pest biology, environmental information and available technology to prevent the damage of crops beyond an economic threshold level whilst not exposing people, the environment, property and resources to risk (Mitchell and Hutchison, 2009). For IPM to be effective an understanding of the landscape ecology, local biodiversity and habitat management is required and knowledge of the beneficial organisms within the ecosystem is also an important factor (Oluoch-Kasura *et al*., 2013). Biological and chemical control methods are both necessary for effective IPM (Kogan, 1998). Pesticides are recommended to be applied when monitoring indicates they are needed according to strategy established guidelines (Dreistadt, 2012).

IPM has been successful in the reduction of insecticide application, enhancing yields and ensuring sustainable environmentally friendly pest management. The introduction of new chemical insecticides coupled with biological control and good pest management saw a significant reduction in *H. armigera* infestation levels in India to below economic threshold levels (Kranthi and Russell, 2009). However, there have been barriers in the adoption of IPM, including lack of practical sampling tools, challenges involved in fully integrating biological

methods, the need to establish multiple pest-damage relationships, changing economic conditions and the multiple consumer base with different cultural and motivational farming techniques (Mitchell and Hutchison, 2009).

In the USA, evaluation of IPM programmes noted that the strategies resulted in lower pesticide use as well as saved hundreds of millions of dollars as a result of pesticide reduction. In Europe there was a proposed compulsory implementation of IPM to commence in 2014, as part of a bid to reduce or eliminate pesticides as well as protect the environment (European Union, 2009). Pesticide use in Europe has been reduced by up to 50% because of IPM strategies (Abrol and Shankar, 2012). In Australia, IPM has been implementated on a number of crops including cotton, wine grapes, pome and stone fruits, vegetable crops and citrus. The IPM strategies in Australia include the use of predatory species, mating disruption and selective insecticides. Pesticide use in Australia in IPM regions has been reduced by up to 30% and problem pests such as lettuce aphid have been managed through an overall IPM strategy (Abrol and Shankar, 2012). The availability of *Bt* cotton and selective control options for *H. armigera* such as spinosad, indoxacarb and emamectin pesticides have greatly encouraged the uptake of IPM, as growers facilitate the management of *H. armigera*. The ban of certain pesticides in India as well as government directives facilitated the uptake of IPM strategies (Abrol and Shankar, 2012). It can be noted that IPM strategies have been highly succesful in Asian rice agriculture systems as well as the highly innovative push-pull systems in East African maize (Fitt and Wilson, 2012). In Africa, IPM has been adopted with successful results mainly by subsistence farmers (Abrol and Shankar, 2012). Cotton companies and the government provide technical expertise, arrange for reliable input such as pesticides, fertilisers and seeds and supply them according to need. Farmer field schools for cotton farming on a smaller scale have been implemented through the Food and Agricultural Organisation (FAO) in several African countries (Kranthi and Russell, 2009).

1.5.2 Chemical control of *Helicoverpa armigera*

Farmers use pesticides which are fast acting, broad spectrum, less labour intensive, more energy efficient and relatively safe to non-target vertebrates if properly used to control *H. armigera* (Oluoch-Kasura *et al*., 2013). The egg and pupal stages are biologically resistant, and many pesticides are ineffective against them. Late instar caterpillars are also less susceptible to most chemicals (Hill, 2008). When using contact insecticides, the timing of the application is very important for control. Sprays should be applied at weekly intervals during the egg-laying period of *H. armigera* (Hill, 1987).

Several chemical broad-spectrum insecticides such as pyrethroids, organophosphates, dichlorodiphenyltrichloroethane (DDT), carbamates and endosulfan have been used to control *H. armigera* (Zhu, 2005; Quinn *et al*., 2011). Organophosphate compounds are contact and systemic insecticides with a high toxicity to mammals; carbamates are more broad-spectrum but less toxic to vertebrates (Hill, 2008). The use of broad-spectrum pesticides to control *H. armigera* results in the killing of non-target insects, i.e. natural predators and parasitoids (Zhu, 2005). Newer chemicals, such as Spinosad, indoxacarb and emamectin are effective against *H. armigera* as well as being less toxic to beneficial insects (Kranthi and Russell, 2009).

1.5.3 Natural and biological control of *Helicoverpa armiger*a

Biological control involves the use of living organisms or their products in the management and control of *H. armigera* (Sree and Varma, 2015). Biological control agents are relatively inexpensive, environmentally friendly and safe (Abate, 2012). They also offer effective solutions for pests that are difficult to manage such as *H. armigera* (Oluoch-Kasura *et al*., 2012). Microbial pesticides such as bacteria based, fungi based, and virus based are the most widely used and relatively the cheapest method of biological control (Szewczyk *et al*., 2006). Biological control of *H. armigera* includes the use of bacteria, viruses, plant extracts, fungi and nematodes (Hill, 2008).

Natural control involves the use of non-human intervention measures. These measures include climatic factors such as wind, sunshine, rain and temperature, topographic features such as rivers, lakes and mountains (Oluoch-Kasura *et al*., 2013). The timing of planting, harvesting, crop rotation and varied cropping systems help to reduce *H. armigera* build up in the farming ecosystem (Karim, 2000). Changing the planting cycles may also lead to considerable changes in pest damage. However, simplified agricultural landscapes which result from clearing of land for agricultural purposes, up-scaling and expansion of crop fields, have led to a reduction of natural habitats, which has resulted in a loss of biodiversity with respect to natural pest control (Sree and Varma, 2015).

1.5.3.1 Natural enemies

The natural enemies of *H. armigera* have huge variations in occurrence and the impact is dependent on the host plant association (Cherry *et al*., 2003). A total of 112 natural enemies have been reported for *H. armigera* in India, which are 83 parasitoids and 29 predators. Of the 83 parasitoids, nine are known to be egg parasitoids, four are egg-larval parasitoids, 61

are larval, seven are larval-pupal and two are pupal parasitoids (Lingappa and Hedge, 2001). Mud wasps such as *Delta* spp. have been recorded to carry off many large *H. armigera* larvae to feed their young. Predation by insectivorous birds and arthropods also accounts for substantial natural control of *H. armigera* (Rao and Babu, 2009).

1.5.3.1.1 Egg parasitoids and predators

Trichogramma pretiosum is the most commonly used egg parasitoid biocide in the control of *H. armigera* by Australian growers and pest managers on cotton (Mensah and Pyke, 2007). In India *T. chilonis* (Ishii) and *T. brasiliensis* (Ashmead) were recorded to be successful parasitoids of *H. armigera* eggs in field trials, although parasitism was recorded to be influenced by host plants (Rao and Babu, 2009). Parasitoids such as *Telenomus* spp. are more abundant during the middle to late cotton season in Australia (Mensah and Pyke, 2007). The larvae of *Chrysopa* spp. are active parasitoids of *H. armigera* eggs (Mensah and Pyke, 2007). The effectiveness of parasitoids from the family Trichogrammatidae as a control method were measured and not permanently established in South Africa (Cherry *et al*., 2003).

1.5.3.1.2 Larval and pupal parasitoids and predators

Microplitis demolitor (Wilkinson) is a parasitoid for the larval stage of *H. armigera* that is used in cotton and cereal fields (Mensah and Pyke, 2007). The adult stage of *M. demolitor* lays its eggs in young *H. armigera* larva, and the *M. demolitor* larva develops by feeding inside the host (Mensah and Pyke, 2007). *Campoletis chlorideae* (Uchida) is a parasitoid that is mainly active in the first to third instars whereas *Corcelia illiota* (Curron) is most common in the late instars (Rao and Babu, 2009). *Cermatulus nasalis* (Westwood), *Orius* spp., *Ochelia schellenbergii* (Guérin-Méneville) are predatory bugs that feed on *H. armigera* larvae (Mensah and Pyke, 2007). Other early instar larval parasitoids namely *Eucelatoria bryani* (Sabrosky) and *Hyposoter didymator* (Thumb) resulted in successful parasitism under laboratory conditions but unsuccessful in the field in India (Rao and Babu, 2009). *Hyposoter didymator* (Thunberg) has been reported to have a substantial impact on the control of *H. armigera* in Europe (Cherry *et al*., 2003). The predatory mite, *Phytoseiulus persimilis* (Athias-Henriot) found in Kenya, has been successfully used as a control measure (Abate, 2012). The use of *Rhynocoris fuscipes* (Fabricius) and *R. marginatus* (Fabricius) greatly suppressed *H. armigera* infestations in groundnuts and bhendi fields (Sahayaraj and Balasubramanian, 2016). However, in India, the use of egg-larval and larval parasitoids has been met with limited success in *H. armigera* control (Abate, 2012). In South Africa the use

of *Habrobracon braiicornis* (Wesnel) had no practical success (Cherry *et al*., 2003). A new parasitoid, *Chelomus texanus* (Cresson) was imported from the United States of America in the 1940s and over 1.4 million parasitoids were released in South Africa with no subsequent recoveries. The release of parasitoids against *H. armigera* has generally not been regarded as successful (Cherry *et al*., 2003). The use of ant species namely *Anoplolepis custodiens* (F. Smith) and *Pheidole megacephala* (Fabricius) to attack the pupal stage of *H. armigera* was demonstrated to affect the survival of the pest (Bownes *et al*., 2014).

1.5.3.2 Plant extracts used against *Helicoverpa armigera*

The use of plants or plant extracts to control *H. armigera* has not been adopted substantially. The plant extract neem oil, which is extracted from the seeds of *Azadirachta indica* (A. Juss) is a repellent, antifeedant and an insecticide (Siegwart *et al*., 2015). The use of *A. indica*, *Khaya senegalensis* (Desr.) and *Hyptis suaveolons* (L. poit) in isolation or as part of an IPM system has been reported to result in substantial control (Akhtar and Isman, 2013). Neem oil or neem seed extracts can be utilised in the control of *H. armigera* 2-3 months after planting (Kranthi and Russell, 2009). The crude extract from the seeds of the plant *Annona squamosa* (L.) has been documented to have considerable insecticidal properties against *H. armigera* (Gopalakrishnan *et al*., 2014). A hot water extract of *Ipomonea carnea* (Jacq.) and *Vitex negundo* (Linn.) resulted in the suppression of *H. armigera* infestations in field studies on groundnuts (Sahayaraj and Balasubramanian, 2016).

1.5.3.3 Bacterial control of *Helicoverpa armigera*

Bacillus thuringiensis is a gram-positive spore-forming bacterium that is soil dwelling with parasporal crystals that can be used for control of agricultural pests (Szewczyk *et al*., 2006). The majority of *B. thuringiensis* (*Bt*) products are for the suppression and management of lepidopteran species such as *H. armigera* (Lacey, 2017). During spurolation, many *Bt* strains produce proteinaceous inclusions known as delta-endotoxins, which have insecticidal action (Walter, 2014). *Bt* toxins are activated in the insect midgut through proteolytic cleavage. The toxins bind to the epithelial cells in the intestine, resulting in pore formation and eventual cell lysis (Madigan *et al.,* 2012). The spores form midgut pores and limit regeneration of cells through septicemia which results in insect death (Siegwart *et al*., 2015). Genetically engineered crops and pesticide sprays containing insecticidal proteins of *Bt* have been developed and applied in the field (Gunning *et al.*, 2005). The introduction of *Bt* cotton and Cry toxins in *H. armigera* control resulted in novel environmentally friendly control methods

that are effective (Kranthi and Russell, 2009). Several subspecies of *Bt* are effective for the control of *H. armigera* (Cherry *et al*., 2003).

Yield increases in *Bt* cotton compared to conventional cotton have been observed, with increases between 14 and 20% for large scale farming and increases of between 20 and 85% for small scale farming over a number of seasons (Gouse, 2013). Yield increases are however dependent on level of *H. armigera* infestation in a particular season as well as the effectiveness of chemical pest control (Gouse, 2013). In South Africa the biopesticide, DiPel (*B. thuringiensis*), is utilised in the control of *H. armigera* (Moore and Kirkman, 2010). Of the total national cotton area in South Africa in 2002, 40% was *Bt* cotton. Initially, the use of *Bt* cotton resulted in maintained or increased cotton yields while simultaneously reducing chemical pesticide use by up to 75%. This was particularly helpful to small scale farmers struggling with *H. armigera* (Russell, 2004).

1.5.3.4 Nematode control of *Helicoverpa armigera*

Entomopathogenic nematode juveniles that belong to the families of Steinernematidae and Heterorhabditidae are obligate parasites of insect species and are associated with symbiotic bacteria (Sree and Varma, 2015). They are motile predators, which are highly virulent therefore quickly kill their hosts, have a high reproductive potential and can be cultured *in vitro*. They have a broad host range and are safe to vertebrates and other non-target organisms. Research into the control of *H. armigera* has shown them to be opportunistic pathogens rather than active biocides. This has led to a limitation in their effective use as control measures (Karim, 2000). In Pakistan *Steinernerma pakistanense, S. asiaticum, S. feltiae* and *Heterorhabditis indica* were tested against *H. armigera* on okra and cotton plants in greenhouse studies and field studies. Significant control of *H. armigera* was recorded for both studies with S. *pakistanense* being the best performer in terms of mortality (Fayyaz *et al.,* 2017). The nematodes *Overmemis* and *Hexamermis* were recorded to attack *H. armigera* larvae on pigeon pea plants in India (Rao and Babu, 2009). Mermithid nematodes have been reported to be parasitic to *H. armigera* larvae, with up to 93% infection (Rao and Babu, 2009). In South Africa *H. bacteriophora* and *S. tophus* were applied against pre-pupal and pupal stages of *H. armigera* in a greenhouse study with initial high mortalities being recorded for both species and subsequent significant decline in mortalities observed for further generations (Hatting and Malan, 2017). Commercial applications have not been reported yet in Africa. Research for the development of nematode bioinsecticides from endemic strains in South Africa is still on-going (Kaya *et al.*, 2006).

1.5.3.5 Virus control of *Helicoverpa armigera*

Viruses that occur naturally in *H. armigera* can be utilised in the control of the pest. Helicoverpa armigera stunt virus (HaSV), a small RNA virus that belongs to the Tetraviridae family has been reported to be infectious to *H. armigera* and cause mortality in neonates (Hanzlik *et al*., 1995). However, the use of small RNA viruses as control agents is limited by the common instability of RNA viruses and their relatedness to vertebrate viruses (Hanzlik *et al*., 1993). Helicoverpa armigera cytoplasmic polyhedrosis virus HaCPV has been reported to cause mortality and be synergistic in mortality effects when combined with *Bt* (Marzban *et al*., 2009). Mortality of bollworm in the early immmature stages using HaCPV has been reported to vary between 7-14 days. The infection and activity of CPV is limited to some cells in the insect midgut (Marzban *et al*., 2009). The main group of entomopathogenic viruses, known as baculoviruses, have been investigated and succesfully formulated as *H. armigera* control agents on a commercial scale globally (Rohrmann 2013), including in South Africa (Moore and Kirkman, 2010).

Helicoverpa armigera resistance to control measures

Helicoverpa armigera has been reported to have a great capacity to develop resistance to insecticides (Maiti and Kumari, 2016). The intensive and continued application of pesticides as well as the indiscriminate application of these pesticides has resulted in the development of resistance in *H. armigera* (Armes *et al.*, 1996; McCaffery, 1998; Peshin *et al*., 2009; Windley *et al.*, 2012). In a study in India by Kranthi *et al*., (2001), high levels of pyrethroid resistance from field collected *H. armigera* populations were recorded, with most resistance being identified from regions where pyrethroid use was most frequent. In West Africa the development of significant resistance to pyrethroids in *H. armigera* populations was reported from 1996 (Russell, 2004). The failure of chemical insecticides due to resistance has led to the search for alternative control methods and adoption of bio-insecticide control (Xu *et al*., 2005).

The continued control of *H. armigera* with *Bt* and abamactin (natural derivative of *Streptomyces avermitilis*) has resulted in the development of resistance in some populations (Zhuang and Gill, 2003; Zhu, 2005; Gassman *et al.*, 2011). The use of *Bt* in the control of *H. armigera* has been unsatisfactory, even with repeated application with observations that the larvae feed sparingly on applied areas and move onto covered or internal foliage where they continue to cause damage (Kranthi and Russell, 2009). In Australia, China and India,

transgenic cotton expressing Cry1Ac reported development of *H. armigera* resistance to the toxin (Ackhurst *et al.*, 2003). In Pakistan some *H. armigera* populations have also developed resistance to the toxin Cry1Ac (Alvi *et al*., 2012). In India, geographically distinct *H. armigera* populations showed different reactions to the Cry1Ac toxin, with some populations showing a higher tolerance to the toxin (Gujar *et al*., 2008). The developed resistance to the toxin has resulted in low suppression of *H. armigera* populations by *Bt* (Lingappa and Hedge, 2001). Research by Gunning *et al.*, (2005) showed evidence that some *H. armigera* strains have developed resistance to *Bt* toxin Cry1Ac in South Africa resulting in the decreased efficacy of some transgenic crops (Huesing and English, 2004).

1.6 An introduction to baculoviruses

Baculoviruses are insect pathogens that have a high specificity, narrow host range, environmental persistence with the ability to act synergistically with other natural enemies and may result in epizootics of the host (Claus e*t al*., 2012). They have a high pathogenicity to target species and are less harmful to the environment than chemical pesticides (Huber, 1986; Groner, 1990; Cunningham, 1995). Baculovirus diseases affect mainly the larval stages of hosts (Federici, 1997). Insects that belong to the orders Lepidoptera, Diptera and Hymenoptera are the only insects known to be affected by baculoviruses (Lapointe *et al*., 2012).

1.6.1 Baculoviruses in biological control and commercial importance

Baculoviruses result in lethal infections of susceptible hosts upon uptake of environmentally stable OBs (Berretta *et al.*, 2013). For baculovirus epizootics to be possible, larvae must consume contaminated foliage (Eldered and Reilly, 2014). Baculoviruses are orally infectious; alkaline soluble OBs are ingested and dissolved in the insect midgut because of the high pH resulting in the release of virions (Miller, 1997; Claus *et al*., 2012; Haase *et al*., 2013).

Virus transfer from one host to the other can occur through horizontal or vertical transmission. Horizontal transfer is dependent on contact rate between susceptible hosts as well as the plant architecture (Georgievska, 2009). The infection of the insect host by baculovirus results in cessation of cell division and an enlargement of the cells (Berretta *et al*., 2013). The effectiveness of the baculovirus infection depends on several factors including larval age, nutritional status, dose and virulence of the strain as well as environmental factors. Signs of infection include the change of insect colour, altered feeding behaviour, lack of

appetite and the total interruption of feeding after a few days (Claus *et al*., 2012). Few signs of infection can be observed in the first three days post infection, but by the fourth day the infected larvae begin to respond much more slowly to external stimuli and feeding slows down and usually ceases by the seventh day. From the fourth day, larval swelling and glossy cuticles can be observed from the baculovirus progeny building up in the epidermal and fat body nuclei (Federici, 1997). The growth of infected larvae is significantly delayed and the length of the interval between infection and death varies from 3-21 days (Claus *et al.*, 2012). A failure in moulting is observed after infection and the death of the host is followed by the liquefaction of the cadaver. The liquefaction of the diseased larvae occurs during the process of the progeny virus being converted from insect biomass (Miller, 1997). Before liquefaction, the larvae will crawl to the top of the feeding host and then die. They hang by their proleg crotchets. The fragile cuticle ruptures in time releasing lots of OBs from the lysed cells (Federici, 1997).

1.6.2 Baculovirus morphology and life cycle

Baculoviruses produce a paracrystalline structure that envelopes OBs which surround the virions during late infections of host insect cells (Berretta *et al*., 2013). Baculoviruses have two different progenies during their lifecycle, known as budded virus (BV) and occlusion derived virus (ODV) (Figure 1.6) (Claus *et al*., 2012). The ODV envelope is adapted for interacting with polyhedron structures in the occlusion process and facilitating the infection of the midgut epithelium under the harsh conditions of the insect gut. The occluded viral form is a highly stable form even in the insect midgut (Miller, 1997). The nucleocapsids of the ODV and the BV have similar structures with the major difference being observed in the envelopes which have differing compositions (Funk *et al.*, 1997). The ODVs are made up of one or more nucleocapsids in a membrane that is synthesised within the nucleus of infected cells (Claus *et al.*, 2012). A crystalline single polypeptide protein matrix surrounds the virions resulting in occlusion bodies (OBs) (Claus *et al*., 2012).

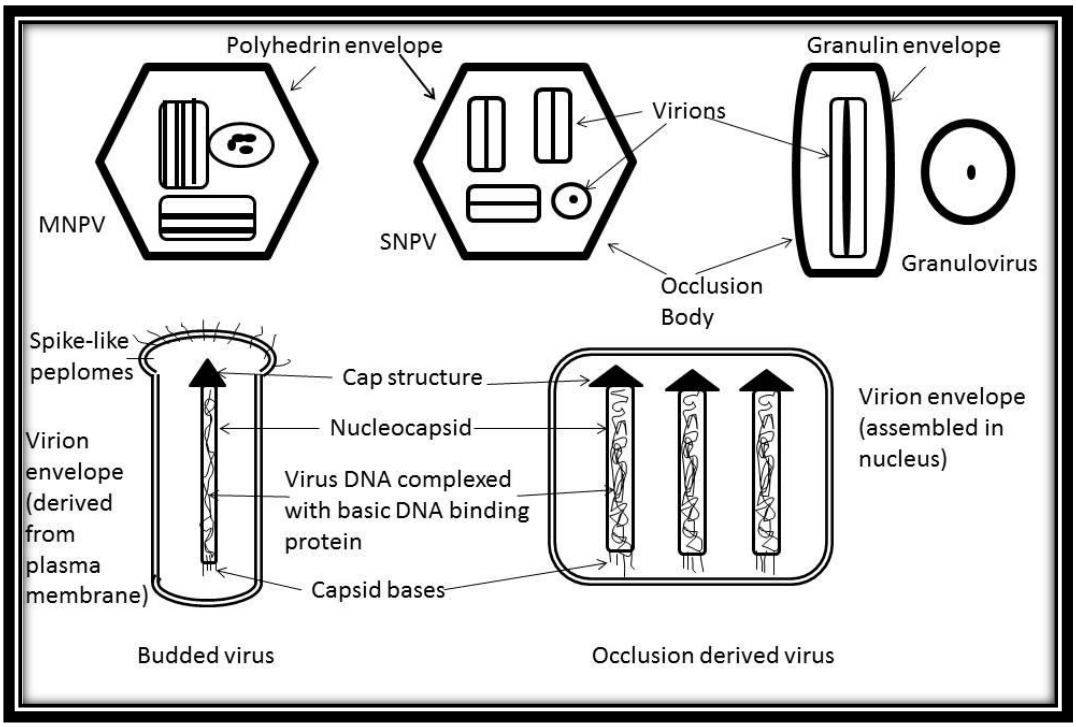


Figure 1.6: The general structure of the two different genera of baculoviruses namely nucleopolyhedrovirus (NPV) and granulovirus (GV). The two common states of NPV are represented, showing the cross-sectional structure of multiple NPV and single NPV. The different phenotypes that are expressed during insect host infection of budded virus and occluded derived virus are also shown as a cross-section (Fauquet *et al*., 2005).

The BV and ODV are responsible for different functions in the insect cell host infections (Haase *et al*., 2013). The BV envelope is adapted for efficient cellular movement and infection of tissues whereas the ODV envelope is adapted for interaction with polyhedrin and granulin structures in the occlusion process and facilitation of midgut infection (Funk *et al*., 1997; Claus *et al*., 2012).

The BVs are produced when nucleocapsids bud through the plasma membrane surface of an infected cell (Figure 1.7) (Fauquet *et al*., 2005). The BV spread the infection through the cells of the insect host through receptor mediated endocytosis (Berretta *et al*., 2013). The OBs protect the ODV and are responsible for the transmission of infection. The ODVs pass through the midgut peritrophic matrix and fuse the viral envelope with the brush border membrane of the columnar midgut epithelial cells and the viral nucleocapsids enter the cytoplasm (Claus *et al*., 2012). The ODV envelope is also believed to aid in virion stability in infected host cells and in the environment (Funk *et al*., 1997).

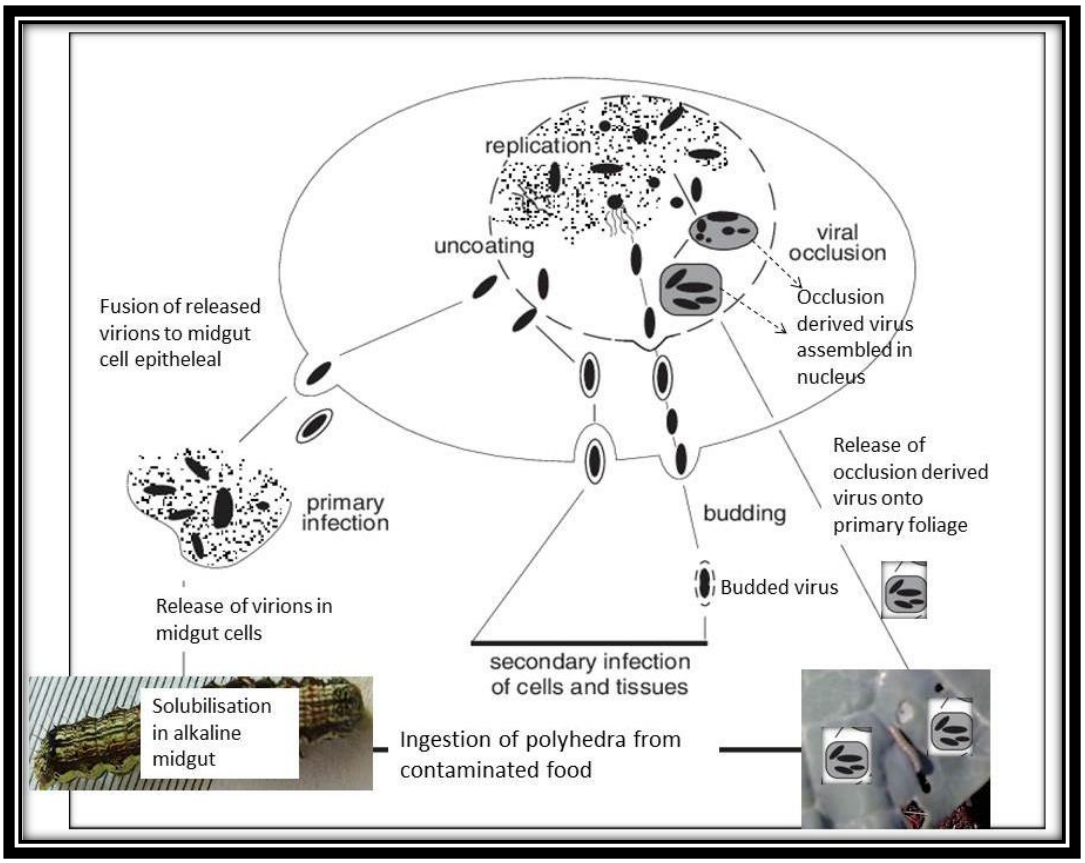


Figure 1.7: The baculovirus infection cycle of a susceptible insect host showing morphology stages and roles in infection (Murphy *et al*., 2004).

1.6.2.1 Primary infection of insect cells

After ingestion of OBs, viral entry is through the midgut epithelium where the OBs are dissolved in the alkaline environment of the midgut which ranges from pH 8-11 (Figure 1.7) (Federici, 1997). The delivery of nucleocapsids into the midgut epithelial cells during early infection is performed through a mechanism of membrane fusion, which is aided by per os infectivity proteins (PIFs) (Berretta *et al*., 2013). Cell entry is mediated by glycoprotein 64 (gp64) which interacts with cell receptors resulting in acidification of internalised vesicles therefore releasing the nucleocapsid to the cytoplasm (Pidre *et al*., 2013). In the early stages of infection, budded viruses are produced, and they spread infection within the insect (Miller, 1997). A cell cycle arrest is performed by baculoviruses prior to DNA replication and the host protein synthesis is also subsequently shut off. Virogenic stroma is assembled early in the infection cycle in the nucleus and used for the assembly and replication of viral progeny (Berretta *et al*., 2013).

1.6.2.2 Secondary infection of insect cells

After the midgut infection, the BV particles spread throughout the insect body in the haemolymph, infecting the haematocytes, fat body cells, trachea cells and hypodermis cells (Grzywacz *et al*., 2007). The spreading of viruses in the insect cells by BVs is part of the secondary infection of baculoviruses (Berretta *et al*., 2013). The viruses cross the peritrophic membrane whereby the virion envelope encounters the microvilli membrane of columnar midgut epithelial cells. After contact, the virion envelope fuses with the microvillar membrane, which allows the nucleocapsids to enter the microvilli (Federici, 1997). The nucleocapsids are transported to the nucleus through actin polymerisation. In the nucleus, using the naked virions, transcription is initiated, which leads to the assembly of new nucleocapsids (Claus *et al*., 2012). Newly replicated baculovirus genomes are packaged with viral proteins to form nucleocapsids. The newly formed nucleocapsids bud from infected cells in the process acquiring a virally modified plasma membrane which forms the budded virus envelope (Herniou *et al*., 2003).

1.6.2.3 Late infection of the insect cells

In late infection stages, the natural cellular transcription is mainly shut off, with 95% of the cell protein activity being directed to polyhedrin/granulin protein production (Berretta *et al*., 2013). In the later stages of infection, virions within the insect cell nuclei become encapsulated and occluded. The ODVs are encapsulated in either a polyhedrin protein or granulin protein depending on the genera of the baculovirus, whereas BVs are not encapsulated (Funk *et al*., 1997). The *gp64* gene encodes for the BV and aids in envelope fusion (Federici, 1997). Occluded virions are released when the diseased insects disintegrate which results in contamination of the foliage allowing for further infection. The baculovirus envelope is adapted for movement and infection of insect tissues (Miller, 1997).

1.6.3 Baculovirus genome

Baculovirus genes are expressed in three phases, early, late and very late gene expression. Early genes are transcribed by the host RNA polymerase II before virus DNA replication, the early gene products are then used for late and very late gene expression (Berretta *et al*., 2013). Baculovirus early genes can be categorised into immediate early (*ie*) and delayed early genes (Berretta *et al*., 2013). Some genes are conserved within genera or large groups of related baculoviruses, some genes are present in only a small group of baculoviruses and some genes are only unique to a single baculovirus genome (Harrison and Hoover, 2012). DNA genomes require primers for DNA polymerases to initiate replication on the ends,

which may result in mutation of primer binding sites, circulisation of the ends eliminates the need for primers and is done by baculoviruses (Ball, 2007). The core genes of the baculovirus genome and their function are summarised in Table 1.1.

Table 1.1: The core genes expressed during the baculovirus infection cycle in the host insect cells and their function (Funk *et al.,* 1997; Herniou *et al.,* 2003; Fauquet *et al.,* 2005; Slack and Arif, 2007; Lapointe *et al.,* 2012; Berretta *et al.,* 2013; Rohrmann, 2013)

Core genes	Cycle	Function
lef-2, DNA polymerase	Replication	Involved in DNA replication in late infection cycle
lef-1	Replication	Encodes for a putative DNA primase
Lef-7	Replication	Possibly codes for DNA binding protein
lef-3, dbp-1, me-53	Replication	DNA replication role in lepidopteran baculoviruses
ie-1	Replication	Dimerisation, nuclear import, binding, transactivation
p143 helicase	Replication	Unwinding DNA
pe38, ie-2	Replication	Regulation of expression of early genes
Lef-4, lef-8, lef-9, p47	Transcription	Products of the genes form complexes that are responsible for viral RNA polymerase
Lef-10, pp31/39, vlf-1	Transcription	Essential for late gene expression
Lef-5	Transcription	Definite function unclear
Polh/gran	Structural	Major structural components of occlusion bodies
p6.9	Structural	Condensation of DNA during packaging
vp39	Structural	Major capsid protein
vp1054, vp91	Structural	Associated with BV and ODV
fp	Structural	Involved in fusion of BV envelopes with host endosomal membranes
gp64	Structural	Modifies the BV envelope and serves as a fusion protein
pp78/83	Structural	Essential for NPV viability and is associated with both the BV and ODV
gp41	Structural	Required for the exit of nucleocapsids from infected cells
ptp	Structural	Protein tyrosinase phosphatase protein that is associated with the BV and ODV
pk1	Structural	Involved in DNA unpackaging
ODV-e66	Structural	Associated with nucleocapsids upon occlusion
ODV-e56	Structural	Associated exclusively with the envelope
ODV-e25	Structural	Major protein of polyhedrin or granulin
ODV-e18	Structural	Localised to the envelopes surrounding ODV
p74	Structural	Important for the oral infectivity of the ODV envelope by facilitating the binding of virions to the midgut cells
Pif-1, pif-2, pif-3, pif-4/19k/ODV-e28, pif-5	Infection	Infection of the midgut epithelial cells
p35, iap	Infection	Block apoptosis of infected cells
calyx	Auxiliary	Localises to the margins of p10
p10	Auxiliary	Responsible for electron-dense spacers and is expressed late in the infection cycle. Associated with polyhedrin protein and may be involved in cell lysis, virus stability and efficient dissemination of polyhedra
Lef-6, lef-12	Auxiliary	Role in late gene expression
Egt	Auxiliary	Blocks normal larval development allowing the larvae to continue feeding

1.6.4 Nucleopolyhedrovirus and granulovirus taxonomy

The size and the shape of the baculoviruses allows for relatively simple morphological identification of GV or NPV. In NPVs the polyhedra typically range from 1-5 μm in diameter and GV average 300-600 nm in length and 150 nm in diameter (Berretta *et al*., 2013). In GVs two or more virions are very rare, although NPVs and GVs are closely related bicochemically and structurally (Fauquet *et al*., 2005). Phylogenetic analysis has allowed further insight into the baculoviruses, which has in turn allowed classification into four genera based on phylogenetic relationships with species affected (Berretta *et al*., 2013). The four genera of baculoviruses are alphabaculovirus, betabaculovirus, gammabaculovirus and deltabaculovirus, with alphabaculoviruses being NPVs that infect Lepidoptera, betabaculoviruses being GVs that infect Lepidoptera, gammabaculoviruses being NPVs that infect Hymenoptera and deltabaculoviruses being NPVs that infect Diptera (Claus *et al*., 2012). The lepidopteran NPVs can be subdivided into group I NPVs such as *Autographa californica* MNPV and group II NPVs such as HearSNPV and *Spodoptera exigua* MNPV (Herniou *et al*., 2003). Group I NPVs contain the protein gp64, which is a membrane fusion protein required for viral entry into the cells, group II NPVs lack gp64 and homologues of gp64, making use of the f-protein instead (Haase *et al*., 2013).

For baculoviruses, the NPVs are the most common and widely distributed and they have been reported from more than 400 insect species, most commonly the Lepidoptera and have been reported to be responsible for major insect epizootics (Federici, 1997). The NPVs are the most studied and utilised virus for biocontrol and are named after the host they were first isolated from. Because of the large virus size, they can be identified by light microscopy allowing for relatively easy detection if available (Grzywacz *et al*., 2007).

1.7 *Helicoverpa armigera* nucleopolyhedrovirus: use in control and production

Helicoverpa armigera NPV (HearNPV) is a baculovirus that infects *H. armigera* and has been reported in regions where *H. armigera* is endemic (Grzywacz *et al*., 2007). HearNPV particles are spherical and sometimes irregular in shape, with sizes ranging from 0.6-2.3 μm (Rao *et al*., 2015). Of the reported potential *H. armigera* natural control agents, NPV has been shown to portray the most potential and has been reported to cause up to 90% mortality in the field (Lingappa and Hedge, 2001). After oral ingestion of the virus, the alkaline midgut of *H. armigera* activates the virus activity (Kalawate, 2014). A combination of HearNPV and

insecticides has been reported to exhibit symbiotic mortalities in *H. armigera* populations (Reddy and Manjunatha, 2000). A combination of HearNPV and azadirachtin (a derivative of the neem tree) demonstrated symbiotic efficacy compared to the individual components of the insecticides (Kumar *et al*., 2008). Wild-type HearNPV can be mass produced *in vivo* for commercialisation (Table 1.2) (Sun, 2015). Mass production of *H. armigera* is required for the mass production of HearNPV. Inoculation of the larval stage with virus is performed for virus propagation before harvesting the cadavers and extracting virus (Kalawate, 2014).

Table 1.2: Commercial *Helicoverpa* spp. NPV based biopesticides mainly used for the control of *H. armigera* globally (Kalawate, 2014; Ibarra and Rincón-Castro, 2009; Moore and Kirkman, 2010; Paul, 2004; Haase *et al*., 2015; Noune and Hauxwell, 2016, Plantwise knowledge bank, 2017)

Active product	Product	Country
HearNPV	Helicide (Pest Control India Ltd)	India
HearNPV	Biovirus –H (Biotech International Ltd)	India
HearNPV	Heligard (Margo Biocontrols Pvt. Ltd.)	India
HzNPV	Elcar (Sandoz AG)	USA
HearNPV	HaNPV (Biocontrol research laboratory)	India
HzNPV	HzNPV CCAB (AgBiTech (Pty) Ltd)	Australia
HearNPV	Diplomata (Koppert Biological Systems)	Brazil
HearNPV	Virin (HS)	Russia
HearNPV	Vivus and Vivus Max (AgBiTech (Pty) Ltd)	Australia
HzNPV	Gemstar (Certis USA L.L.C)	USA
HearNPV	DOA Bio V2 (Department of Agriculture)	Thailand
HearNPV	Helicovex (Andermatt Biocontrol AG)	Switzerland
HearNPV	Helicovir (River Bioscience (Pty) Ltd)	South Africa

HearNPV has become an integral component of IPM systems (Buerger *et al*., 2007). HearNPV is commercially mass produced and has been shown to be a biologically viable option for the control of *H. armigera* with efficacy demonstrated in a host of crops in India (Singh *et al*., 2004). The department of biotechnology in India has invested and supported research and development for HearNPV products for the control of *H. armigera* and large-scale field utilisation of the products has been demonstrated (Paul, 2004). In China in 1993, a HearNPV suspension was authorised and commercialised and by 2012 HearNPV was the most produced viral insecticide (Sun and Peng, 2007; Sun, 2015). By 1999, 50,000 L of HearNPV was being used for the control of *H. armigera* in Australia (Grzywacz, 2001). In combination with IPM systems and government regulations, HearNPV and Helicoverpa/Heliothis NPV based formulations (Table 1.2) have increased in their usage and adoption by consumers

(Reddy and Manjunatha, 2000; Buerger *et al.*, 2007; Mazid *et al*., 2011).

Although it shows much promise, variation has been shown for dose recommendation from 2.0×10^9 OB.ml^{-1} up to 2.25×10^{12} OB.ml^{-1} for field applications (Sun, 2015). An American isolate of HzNPV was mass produced in Australia by Ag Biotech *in vivo* in a single colony of Australian *H. armigera* to produce a baculovirus control product for *H. armigera* known as ViVUS (Buerger *et al*., 2007). Initial production of ViVUS was based on manual processes including transfer of live larvae onto diet using brushes, which was labour intensive and costly (Buerger *et al*., 2007). The American company Sandoz Inc. used HzNPV as a biopesticide under the product names Elcar™ and GemStar™ to control *H. armigera* (Szewczyk *et al*., 2006). In China HzNPV was substituted by HearNPV for the control of *H. armigera* with improved performance recorded (Szewcyk *et al*., 2006). As of 2014 in China there were 35 registered biopesticides and 14 were based on HearNPV (Reid *et al*., 2014). A substitution of HzNPV with HearNPV isolated from local populations resulted in increased virulence and a new commercial product known as ViVUS Gold (Buerger *et al*., 2007). ViVUS Gold has been registered in Australia since 2004 and application for control is recommended for smaller larvae (Buerger *et al*., 2007). A Swiss based company, Andermatt Biocontrol AG, developed a biopesticide product based on HearNPV and registered it as HELICOVEX® and was tested in field trials in Italy where the efficacy was reported to be up to 90%, being significantly higher than *Bt* and Indoxacarb (Kessler *et al.,* 2008). The control of *H. armigera* on tomato plants using HearNPV sprayed using a knapsack resulted in 100% reduction for a concentration of 1.15×10^7 OB.ml^{-1} after 7 days and 100% reduction after 16 days for a concentration of 1.15×10^6 OB.ml^{-1} (Moore *et al.,* 2004). In South Africa, a commercial HearNPV product known as Helicovir™ has been marketed and demonstrated to be able to control *H. armigera* populations in agriculture (Moore *et al*., 2004; Moore and Kirkman, 2010).

1.7.1 HearNPV commercial use: opportunities and barriers

Use of commercial baculovirus pesticides has faced limited acceptance because of reasons such as slow speed of kill, registration problems due to biological constraints, marketing of biological insecticides and difficulties experienced for mass production of the insecticides (Claus *et al*., 2012). For baculoviruses to be a successful alternative, there needs to be a highly pathogenic isolate, mass production at a cost that the market can bear, fields where baculoviruses can be readily applied, a crop and pest environment that supports horizontal transfer of the virus and systems where chemical pesticides are required to be reduced (Grzywacz, 2017). The current production of baculoviruses is dependent on baculovirus

infection and propagation in susceptible hosts as well as harvesting and purification. This usually results in high labour requirements, lack of expertise in standardisation and validation of rearing and propagation, difficulties in upscaling the rearing and production system to levels that are economically viable and profitable and difficulties in quality control of production and quality assurance of the final product (Claus *et al.*, 2012). Environmental factors, biotic and abiotic factors affect the persistence of baculoviruses in the environment and field after application. Without reapplication of baculovirus in the field, significant losses of more than 70% after a year of application in OB density have been recorded (Mohamed *et al.*, 1982). These losses in virus density from plant foliage and soil may have been a result of natural factors such as surface run-off from the rain and UV damage from the sun (Mohamed *et al*., 1982; Reid *et al.*, 2014). UV radiation from sunlight plays a major role in the degradation of baculovirus (Grzywacz, 2017). The addition of photo stimulants and UV protectants has been shown to increase the efficacy of HearNPV (Lingappa and Hedge, 2001). The mixture of HearNPV with polymers such as starch, gelatine or pectin has been utilised in the formulation of HearNPV biopesticides as protectants of the baculovirus (Ibarra and Rincón-Castro, 2009).

The market size for HearNPV has been increasing with the relative decrease being recorded for chemical pesticides (Siegwart *et al*., 2015). The success of HearNPV is dependent on being part of an overall IPM system (Zhu, 2005). HearNPV is not as fast acting as chemical pesticides, because of the disease cycle required for its progression which results in a slower speed of kill (Behle and Birthisel, 2014). With the availability of low cost synthetic chemical pesticides that are fast acting, some farmers opt against the use of HearNPV biopesticides (Buerger *et al*., 2007). HearNPV has a narrow specificity and its application may require coupling with other insecticides to control other non-target pests (Nawaz *et al.*, 2016).

However, the development of resistance in *H. armigera* populations to many chemical pesticides raises the opportunity for HearNPV to be adopted as a potential control agent (Lammers and McLeod, 2007). HearNPV is non-toxic and non-pathogenic to organisms that are not targeted; therefore, it does not directly affect beneficial fauna (Nawaz *et al*., 2016). It is also important to formulate the virus in a medium that optimises storage and application (Haase *et al*., 2015). With continued research into HearNPV, there is vast potential for its use in IPM systems and adoption as a viable pest control agent for *H. armigera* and related species (Sun, 2015).

1.7.2 Genetic diversity of HearNPV isolates

The NPV from *H. armigera* can be single or multiple NPV (single or multiple enveloped virions) known as HearSNPV or HearMNPV (Rohrmann, 2013). Many HearNPV isolates have been isolated and sequenced, highlighting viral diversity. Morphological studies of purified HearNPV OB through TEM analysis can aid in the identification of the HearNPV ODV variant namely MNPV or SNPV (Lua *et al*., 2002; Tang *et al*., 2012). Isolation of HearNPV from diseased larvae and extraction of genomic DNA facilitates analysis of genetic variation through restriction endonuclease analysis (REN) or genomic sequencing (Chen *et al*., 2000; Chen *et al*, 2001; Christian *et al*., 2001; Arrizubieta *et al.,* 2015). Isolates of HearNPV from Spain, Australia, Portugal, China, Kenya, Zimbabwe and South Africa have been reported to have different REN profiles (Christian *et al.,* 2001; Lua *et al.,* 2002; Ogembo *et al.,* 2007; Figueiredo *et al.,* 2009; Arrizubieta *et al.,* 2015, Noune and Hauxwell, 2016) which means that each of the strains are different. Genomic sequencing aids in the inference of genetic diversity of the viruses aiding in the identification of different viral strains. Some HearNPV isolates from similar and different geographical locations have been sequenced and published on genome databanks such as GenBank highlighting differences on a nucleotide level hence the novelty of the viruses (Table 1.3) (Chen *et al.,* 2001; Zhang *et al.,* 2005; Ogembo *et al.,* 2007).

Table 1.3: Table of HearNPV isolates that have been fully sequenced showing the country of origin and the GenBank accession number

Country of origin	Isolate name	Sequence length (bp)	GenBank accession number	Reference
Spain	HearNPV-SP1A	132,481	KJ701032.1	Arrizubieta *et al.,* 2015
Spain	HearNPV-SP1B	132,265	KJ701033.1	Arrizubieta *et al.,* 2015
China	HearNPV-C1	130,759	NC_003094.2	Chen *et al*., 2001
China	HearNPV-G4	131,405	AF271059.2	Zhang *et al.,* 2005
China	HearMNPV	154,196	NC_011615	Tang *et al*., 2012
India	HearNPV-L1	136,760	KT013224.1	Raghavendra *et al*., 2016
Australia	HearNPV-Au	130,992	JN584482.1	Zhang *et al.,* 2014
Australia	HearNPV-AC53	130,442	KJ909666.1	Noune and Hauxwell, 2016
Australia	HearNPV-H25EA1	130,440	KJ922128.1	Noune and Hauxwell, 2016
Kenya	HearNPV-NNg1	132,425	AP010907.1	Ogembo *et al.,* 2007

1.8 Virulence of HearNPV isolates

The virulence of HearNPV isolates has been reported to be different (Ogembo *et al.* 2007; Arrizubieta *et al.,* 2013). Bioassays have been conducted to test the virulence of different HearNPV isolates against various host populations from geographically distinct locations (Ogembo *et al*., 2007; Figueiredo *et al*., 2009). The virulence has been recorded to vary based on the HearNPV isolate as well as the geographical source of the host population (Georgievska *et al*., 2010). For example, the virulence of HearNPV-SP1 variants against second instar *H. armigera* were recorded to be similar whereas those of HearNPV isolates from Zimbabwe, South Africa, Kenya, Thailand and China that were tested against third instar *H. armigera* larvae from Japan differed significantly (Arrizubieta et al., 2013; Ogembo *et al*., 2007).

1.9 Problem statement

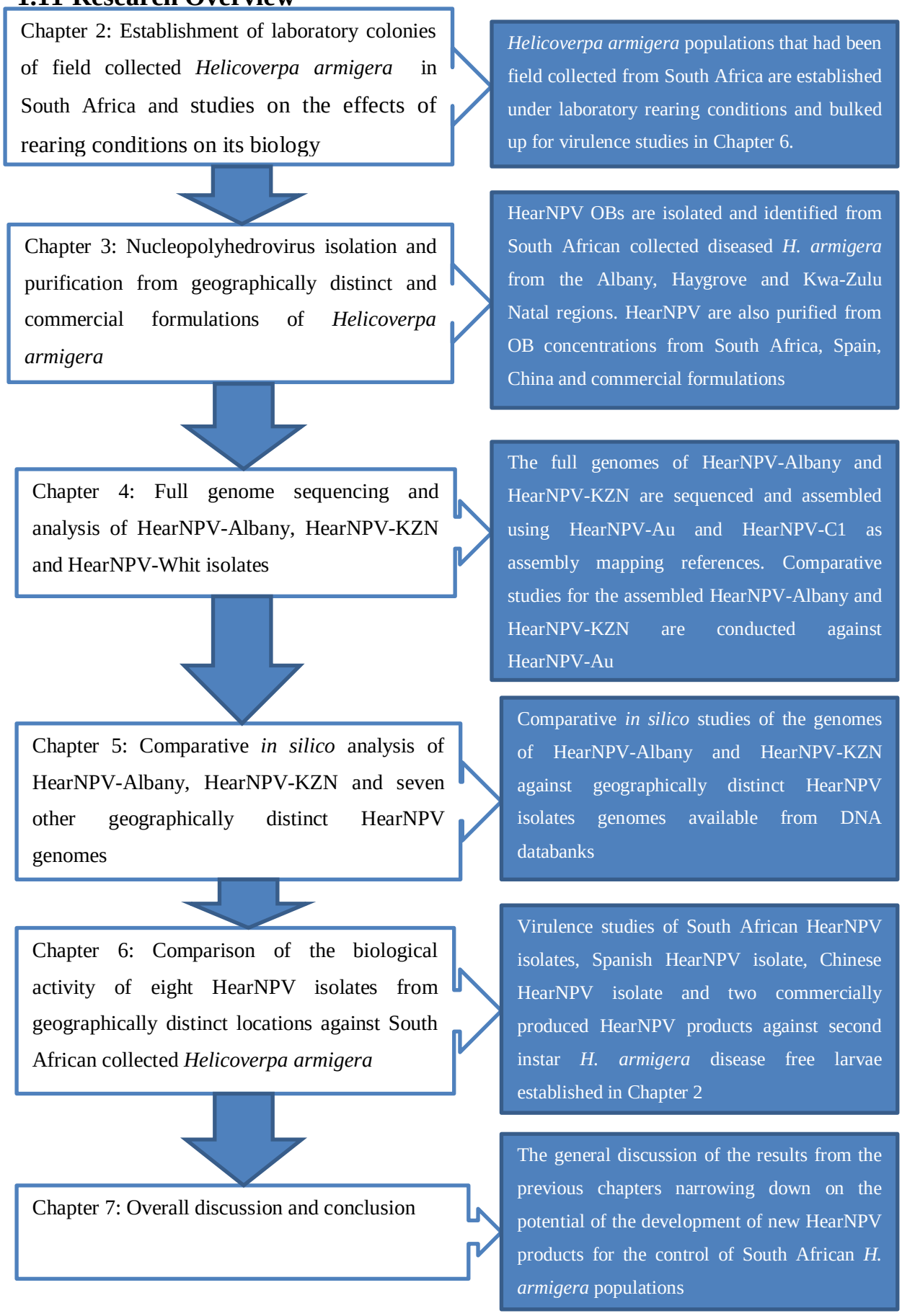
The African bollworm, *H. armigera* is regarded the most important crop pest in South Africa and a major crop pest globally, attacking crops such as maize, cotton, tomatoes, wheat, citrus and many others (Bell and McGeoch, 1996; Cherry et al., 2003; Kranthi and Russell, 2009; Gouse, 2013). Due to its polyphagous nature, *H. armigera* results in major crop losses and economic losses. Chemical and biological pesticides are available for its control but *H. armigera* has a history of pesticide resistance. It has been reported to have developed resistance against some chemical pesticides and some biopesticides, such as *Bt* toxin, which has resulted in the decline in the efficacy of these control methods. The baculovirus, *H. armigera* nucleopolyhedrovirus (HearNPV), offers a promising alternative control method and has been isolated and identified in local host populations globally including in South Africa. The specificity of HearNPV to *H. armigera* and related species makes it a potential candidate for control of *H. armigera*. In South Africa, a commercial HearNPV biopesticide known as Helicovir™ that is manufactured by River Bioscience ((Pty) Ltd) is used to control the pest in South Africa. However, to avoid a build up of resistance to the virus in South Africa, it is necessary to characterise different HearNPV isolates in terms of genotype and virulence against South African populations of *H. armigera*. The characterisation of HearNPV isolates aids in the identification of novel isolates which may be used to infer virulence differences in the control of *H. armigera.*

1.10 Overall aims of the study

The first aim of the study was to establish healthy, disease-free laboratory colonies of South African collected *H. armigera*. The second aim of the study was to isolate HearNPV from South African *H. armigera* and genetically characterise isolates to establish novelty on the genomic level. The final aim of the study was to compare South African isolated HearNPV to globally isolated HearNPV including Spanish isolated HearNPV and Chinese isolated HearNPV in terms of genetics and biological activity against South African collected disease-free *H. armigera* populations.

The specific objectives for the study were to initially collect *H. armigera* healthy and diseased larvae from geographically distinct locations in South Africa and establish healthy colonies in sufficient quantities under laboratory conditions. The second objective was the purification of HearNPV OBs from the collected diseased insect cadavers and from concentrated HearNPV OBs from Spain, China and commercial HearNPV formulations. The third objective of the study was to extract genomic DNA from South African isolated HearNPV. The fourth objective was to sequence and assemble genomic DNA from South African HearNPV isolates. The fifth objective was to perform comparative studies of genomic sequences from HearNPV isolated from South Africa against globally isolated HearNPV genomes available from online genomic databases such as GenBank. The final objective was to perform virulence studies with South African isolated HearNPV, Spanish isolated HearNPV, Chinese isolated HearNPV and commercial HearNPV products against South African collected *H. armigera* populations.

1.11 Research Overview



2 Establishment of laboratory colonies of field collected *Helicoverpa armigera* in South Africa and studies on the effects of rearing conditions on its biology

2.1 Introduction

Pest insect studies require a constant supply of healthy insects, which can be achieved through appropriate mass rearing protocols (Parker, 2005). An in-depth understanding of the biology of the insect to be reared is required because, for example, insects such as *Helicoverpa zea* and *H. armigera* are cannibalistic during their immature stages which makes it important to rear the larvae in separate incubators (Arrizubieta *et al*., 2016).

Having a disease-free colony allows for the analysis of viral effects as well as propagation of virus progeny with minimal probability of contamination of other disease factors. Further, it is essential to have a disease-free colony for bioassay studies to allow isolation and identification of disease-causing factors during studies (Grzywacz *et al*., 2007). The establishment of a disease-free colony can be achieved through controlled rearing conditions and strict hygiene (Jones, 2000). In most cases, establishing an *H. armigera* colony by obtaining one from a disease-free facility is not possible and it is common to establish a laboratory colony from field collected insects. However, direct use of field collected *H. armigera* for studies has the disadvantage of not guaranteeing enough quantities of insects required for the studies (Dent, 2000). Additionally, field collected *H. armigera* are rarely adapted to laboratory rearing conditions which usually results in high mortality rates for the early generations. Field collected larvae can be established in the laboratory by providing optimal conditions. An *H. armigera* colony from larvae collected from glasshouses and reared over a period of five years has been established for several laboratory studies (Arrizubieta *et al*., 2013). For example, Mironidis and Savapoulou-Soultani (2008) established a laboratory colony from *H. armigera* collected from the field in Greece and they reared the colony on artificial diet.

Helicoverpa armigera is difficult to propagate under laboratory conditions because of its high level of plasticity under natural conditions, making it highly vulnerable under controlled conditions (Singh *et al.,* 2004). This is because under natural conditions, *H. armigera* can move and adopt a new host and environment if the conditions are not suitable for it compared to a controlled setup which may induce stress factors in the larvae. Colony stabilisation is

often only observed on average after five generations (Parker, 2005). Field collected moths should be examined for microsporidia infection, particularly *Vairimorpha* spp., which commonly occur and infect *H. armigera* in laboratory reared colonies, weakening the insects and resulting in lower fecundity. Vertical transmission of the pathogen occurs, which leads to reduced fecundity for subsequent generations and may even result in sterile generations (Grzywacz *et al*., 2007).

Without an ideal propagation facility for *H. armigera,* difficulties in colony establishment and maintenance of a healthy colony could be experienced (Grzywacz and Moore, 2017). It is important to understand the lifecycle and the time taken to complete. This would be useful in devising methods for the control or utilisation of the insects (Parker, 2005). The oviposition medium is different under laboratory conditions to that in the field, in many cases being made of fabric or a membrane. Emergence cages often serve as both mating and oviposition cages (Calkins and Parker, 2005). After oviposition, the eggs can be surface sterilised using hypochlorite solution and have the solution rinsed with water until all the hypochlorite is removed (Kalawate, 2014).

Larval rearing can be carried out on artificial diet or natural host plants. Under laboratory conditions, it is recommended to rear *H. armigera* on artificial diet, which can be made according to requirements (Kalawate, 2014). Artificial diets are more convenient for handling and preparation and the larvae can be reared on many semi-synthetic variants (Grzywacz *et al*., 2007). The natural diet of the adult moth is nectar, and this can be substituted with 10% sucrose or honey solution in the laboratory (Grzywacz *et al.*, 2007). Most of the baculovirus propagation is performed through laboratory reared insects that are reared on artificial diet in controlled environments (Grzywacz and Moore, 2017). The culture for propagation and maintenance of the colony should be kept separate from the culture for virus testing and bioassays in the rearing facilities (Kalawate, 2014).

The aim of the study in this chapter was to establish a laboratory culture of *H. armigera* populations collected in the field in South Africa. The specific objectives of the study were to determine the life cycle of *H. armigera* collected from the Belmont Valley near Grahamstown (Eastern Cape), under optimal conditions, and further determine the effect of the different diets on the colony growth. The second objective was the rearing of two South African *H. armigera* populations from geographically distinct locations using the most

suitable conditions and the best possible diet, for conducting bioassays against these larvae with various H. armigera nucleopolyhedrovirus (HearNPV) isolates.

2.2 Method and materials

2.2.1 Laboratory populations of *Helicoverpa armigera* used in the study

A laboratory colony of *H. armigera* was established by collecting wild-type larvae from farms in the Belmont Valley, near Grahamstown in the Eastern Cape. This colony was referred to as the Albany colony. *Helicoverpa armigera* larvae were collected from green bell pepper fields and cabbage fields that had been reported to have *H. armigera* larvae during field scouting (Figure 2.1). Field collections were initially done at two-week intervals. Twenty to 50 larvae were collected during each field collection and mixed with the existing laboratory colony from the same field population that had been collected earlier, to minimise genotype thinning and increase culture strength.

Figure 2.1: The fields in which *H. armigera* larvae were collected in the Belmont valley in Grahamstown. a: field of green bell peppers. b: field of young cabbages. c: green bell pepper with spherical incision. d: the interior of the bell green pepper with a healthy *H. armigera* larva feeding on the flesh of the fruit. e: a healthy *H. armigera* larva feeding on the edge of an infected cabbage leaf with feeding holes on the leaf surface.

After the Albany colony was reduced to very few individuals, field collections were increased to weekly intervals and the colony was kept in three different controlled environment (see 2.2.3 for rearing conditions) rooms to minimise the risk of colony demise. After collection, larvae were monitored daily until pupation and any diseased or parasitised larvae were removed from the colonies.

The Albany colony was the only colony collected from field scouting for the study (33°19'20.2" S; 26°38'23.4" E and 33°19'35.0" S; 26°39'03.0" E) (Figure 2.2). Towards the end of the study, late instar *H. armigera* were also provided from Haygrove Eden Farm near George, in the Western Cape, South Africa (33°57'59.0" S; 22°22'54.9"') (Figure 2.2). The colony was referred to as Haygrove colony. From the Haygrove colony, 108 larvae were supplied. From the supplied larvae, 42 displayed signs of baculovirus infection and these were transferred to sterile tubes and stored at -20°C for downstream experiments. The healthy larvae were used to establish a laboratory colony.

Figure 2.2: Sites where *H. armigera* larvae were collected in the Belmont Valley near Grahamstown and Haygrove Farms near George.

2.2.2 Diets studied during biological studies of Albany *Helicoverpa armigera*

The response of *H. armigera* to different diets was studied, however, for these studies, only the Albany colony was used because of the late acquisition of the Haygrove colony.

Helicoverpa armigera under laboratory conditions has been reared on a range of diets including cabbage leaves, green bell peppers and artificial diets (Patankar *et al.*, 2001; Liu *et al*., 2004; Balogh *et al*., 2009). The Albany colony was initially reared on green bell pepper cubes as described by Balogh *et al.*, (2009). The green bell pepper cubes were considered because this was the host plant. Green bell pepper cubes (0.5 cm × 0.5 cm) were washed under tap water and rinsed in ddH_20. The green bell pepper cubes were changed every 48 hours due to fungal growth. Cabbage leaves were then substituted as feed because of abundance throughout the year in comparison to green bell pepper cubes as well as ease of handling. Cabbage leaves were also used because collections of the Albany colony were from cabbage fields. Cabbage has been used as a diet in *H. armigera* studies (Yankanchi and Patil, 2009). Cabbage heads were bought from local supermarkets and the leaves were rinsed in ddH_2O before being used as feed. Cabbage pieces were placed in vials with individual larvae and changed every 48 hours. An artificial diet was studied as another dietary option based on literature (Patankar *et al*., 2001; Grzywacz *et al*., 2007; Mironidis and Savapoulou-Soultani, 2008). The artificial diet was based on a protocol from River Bioscience (Pty) Ltd (2014). The artificial diet pre-mix was prepared by adding soy flour (412.8 g), wheat germ (288 g), yeast (240 g), sorbic acid (6.4 g), ascorbic acid (16 g) and nipagin (16 g) to give a total of 972.9 g for the premix. To prepare the diet, 235.27 g of the premix was sterilised in an oven at 180°C for 10 minutes. To the sterilised pre-mix, 1.527 ml of propionic acid and 0.145 ml of phosphoric acid were added. Agar (13g in 400 ml ddH_2O) that had been sterilised through autoclaving at 121°C for 15 minutes was added to the premix. The mixture was stirred, and 400 ml of boiled hot water was added to the mixture until a uniform paste was obtained (Figure 2.3a). The diet was then poured in individual glass/plastic vials (Figure 2.3b and Figure 2.3c) and left for 48 hours in a laminar flow cabinet to cool off before placing *H. armigera* larvae. Diet that was not used was stored at 4°C to prevent desiccation. The mixing bowl, glass vials, plastic vials and wooden mixing stick were sterilised by cleaning in 0.35% hypochlorite solution ($^{v}/_{v}$) and rinsed off with ddH_2O before drying off in a laminar flow cabinet.

Adult *H. armigera* (moths) were fed 10% ($^{w}/_{v}$) sucrose solution (Zhou *et al.*, 2000; Perkins *et al.*, 2009). The solution was prepared by adding 10 g biological grade sucrose (Merck (Pty) Ltd RSA) to 100 ml ddH_20 and autoclaved at 121°C for 15 minutes. Sucrose solution (10 ml) was placed in a petri dish and cotton wool was placed on top to avoid drowning of the moths.

The petri dish was then placed in the eclosion cage and the sucrose solution was changed every two days.

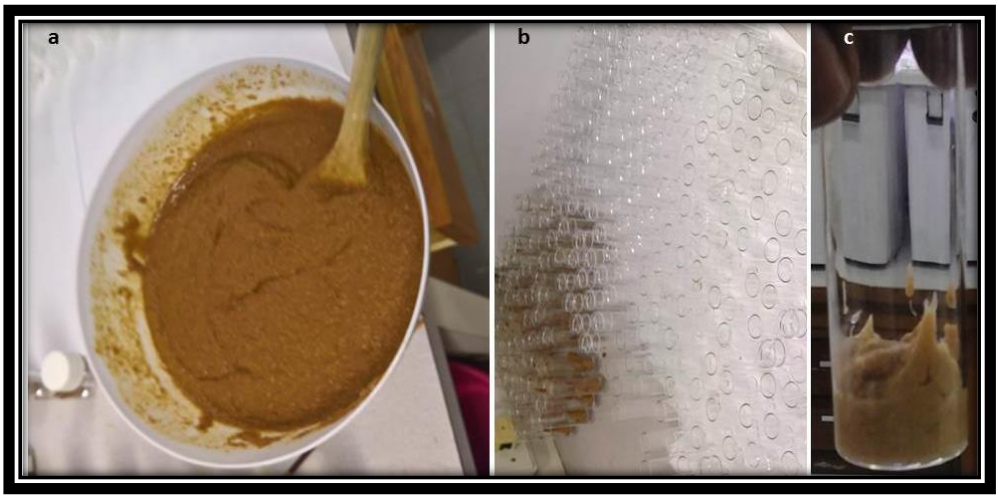


Figure 2.3: Artificial bollworm diet preparation under sterile conditions. A: mixing of the diet under sterile conditions. B: glass and plastic vials for larval rearing on diet. C: diet in a glass vial.

2.2.3 The insectary for *Helicoverpa armigera* maintenance

Rearing of the colony was done in controlled environment rooms (CE rooms) at 27°C (± 2.0°C) (Armes *et al*., 1992; Grzywacz *et al*., 2007). Humidity was low, ranging from 21% to 40% until a humidifier was introduced and humidity was maintained between 45% and 65%. Temperature and humidity in the CE rooms where the insects were being reared was measured daily using a thermometer with hygrometer (MajorTECH® model MT662). The humidity was maintained by using a humidifier (LAICA® Ultrasonic Humidifier, HI3006T) that had ddH_2O water. A bucket of water was also placed in the CE rooms to also maintain high humidity. The CE rooms had 14:10 hours light: dark periods. *Helicoverpa armigera* larvae were reared in groups of 20 on cabbage leaves (first and second instar) and individually until pupation in glass vials on artificial bollworm diet (from third instar) because of cannibalism (Armes *et al*., 1992). Pupae were transferred to eclosion containers in groups at a ratio of 1:1 male to female (differentiation is described in 2.2.4). If pupal numbers were more than 15, larger eclosion containers with dimensions of 55 cm × 35 cm × 40 cm (L × W × H) (Figure 2.4a) were used. If pupal numbers were below 15, eclosion containers with dimensions of 20 cm × 20 cm × 20 cm (L × W × H) (Figure 2.4b) were used. Not more than 50 pupae were placed in the eclosion containers for each generation. The eclosion containers had cut out lids with thin muslin cloth placed on top to allow the internal microenvironment

in the containers to be like the external environment. White muslin cloth was placed on the internal walls of the containers for moth oviposition.

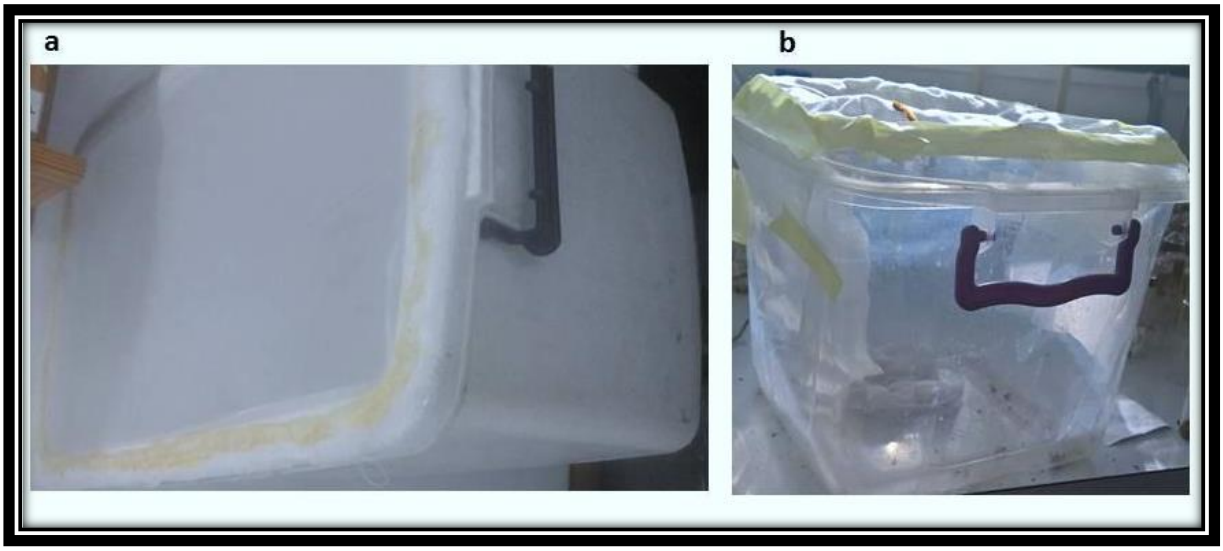


Figure 2.4: Eclosion containers used in the study where pupae were placed before adult moth eclosion. A: large eclosion container B: small eclosion container

2.2.4 *Helicoverpa armigera* biology

The host biology was only studied for the Albany colony because of the late acquisition of the Haygrove colony. The life cycle for the colony was monitored and recorded. The host fitness was recorded using the following parameters; fecundity, egg hatch, pupal weight and length, pupal survival, life cycle duration and adult eclosion.

2.2.4.1 Egg stage

Fecundity was studied using the modified protocol by Opoku-Debrah *et al.*, (2014) for false codling moth. Five newly eclosed virgin female moths and five newly eclosed virgin male moths were placed in eclosion containers. The process was repeated for 10 containers to give a total of 50 female moths and 50 male moths per each generation studied. On the third day of egg laying the adult moths were transferred to new containers (Figure 2.5). The process was repeated after three days and continued until the death of the moths. The total number of eggs laid was quantified and the mean per female moth was determined. From each container of eggs, 50 eggs were selected to measure the egg-hatch (Liu *et al*., 2004). During hatching, the eggs were monitored daily and the number of eclosed neonates was recorded. The process was repeated for eight moth generations. All experiments were performed in triplicate per generation.

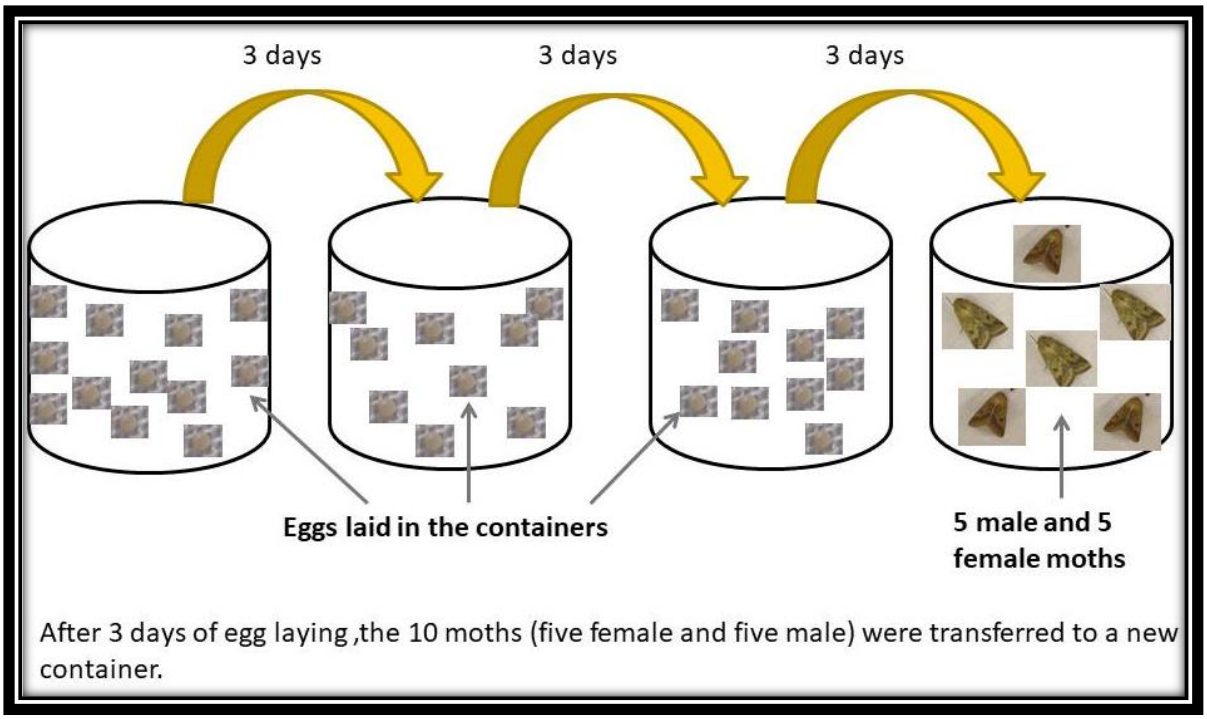


Figure 2.5: Egg quantification procedure, with active adult mating moths being transferred to a new container every 3 days and laid eggs quantified.

2.2.4.2 Larval stage

When the eggs had hatched, neonates were transferred to a solid diet medium and reared in groups of 20-50 according to a modified protocol from Mironidis and Savapoulou-Soultani (2008). From the hatched larvae, 750 were selected in groups of 250 and were set aside for larval studies. Larval development was monitored daily, and stages were determined through head capsule moulting (Liu *et al.*, 2004). Larval lengths and head capsule widths were measured daily by selecting 25 random larvae from each group, placing them on a graduated paper and taking measurements (Figure 2.6). The larvae were monitored daily until pupation.

Figure 2.6: The procedure of measuring head capsule width using a graduated paper. Each line is 2 cm long with the space between the parallel lines being 0.5 mm.

2.2.4.3 Pupal stage

Pupae were transferred to eclosion containers and pupal biology was studied using a modified protocol from Opoku-Debrah *et al*., (2014) for false codling moth. Pupal mass and length were determined by measuring and weighing 40 pupae from each generation (20 male and 20 female). Pupal sex was determined through the observation of the abdominal characteristics (Figure 2.7) to determine sex ratio. The pupal mass was measured by placing 20 female pupae and 20 male pupae in a petri dish and weighing using a laboratory balance scale (4 digital places), less the weight of the empty petri dish and averaged to find the mass of one pupa. Pupal length was measured for each pupa of the 20 pupae using a string and then aligning the string to a ruler. Pupae that were malformed or injured were excluded from the analyses. This was done for six generations of the colony. The pupal period was recorded,

and percentage pupation of larvae was determined for 40 pupae. All measurements were done in triplicate.

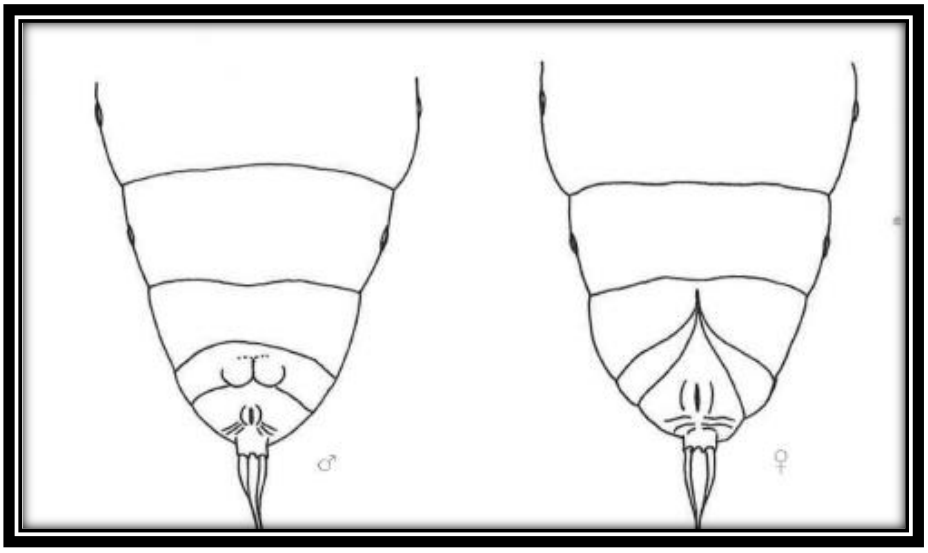


Figure 2.7: The abdominal ventral view of male pupa on the left and female pupa on the right showing sex differences for *H. armigera* (Armes *et al.*, 1992).

2.2.4.4 Adult stage

The sex of the adult moths was distinguished through the wing colour. The females are brownish in colour and the males are lighter, appearing light greenish/yellowish (Smith-Pardo, 2014). The adults were monitored daily to determine longevity, preoviposition and oviposition period as described by Mironidis and Savapoulou-Soultani (2008).

2.2.4.5 Development of the different stages

The collected eggs were quantified and monitored daily for neonate hatching. After 10 days the eggs that had not hatched were discarded. First instars were monitored daily on cabbage or artificial diet. Survival of the larvae was recorded daily through the measurement of larval mortality. Data recorded for the colony during rearing were duration of each larval instar, body length of the instars, head capsule width, pupal weight, duration of the pupal period, pupation percentage, percentage of moth eclosion, longevity of the adults, fertility of the adults and egg-hatch.

2.2.4.6 Survivorship of the larval stage

The survivorship of the different larval stages on cabbage leaf diet and artificial bollworm diet was studied using a modified protocol adapted from Liu *et al*., (2004). The initial colony sample for survivorship used was 250 larvae per diet and this was performed in triplicate. For the larvae on the artificial bollworm diet and on cabbage diet, the instar survivorship was

deduced as a percentage of the number of active larvae at the new instar compared to the number of active larvae for the previous instar. This was calculated using the following formula $\frac{number\ of\ larvae\ observed\ in\ new\ instar\ stage}{number\ of\ larvae\ observed\ in\ previous\ instar\ stage} \times 100\%$. For the sixth instar larvae, the pupae were used as the next instar stage.

2.2.5 Albany and Haygrove colony development

The larvae from the Albany colony that were not used for biological studies were used in the colony bulking up. Artificial bollworm diet and cabbage leaves diet were used for the colony development, with cabbage leaves used when artificial bollworm diet was unavailable. The established rearing protocol for Albany colony was then used to develop the Haygrove colony for downstream biological studies. The two colonies were reared in different insectaries to eliminate possibility of colony mixing. Adults that escaped from the eclosion containers were killed. The oviposition cloths with the eggs were placed next to a humidifier with the surrounding humidity between 60-70%. The Albany colony was maintained for several generations (more than 10 generations) until required for downstream studies and the Haygrove colony was reared until enough numbers for bioassay were achieved.

2.2.6 Quality control

For each generation the rearing containers and vials were cleaned using 5% sodium hypochlorite ($^{v}/_{v}$) before going through a dishwashing cycle at 60°C. The cotton wool used as vial plugs was either discarded or cleaned by soaking in 5% sodium hypochlorite ($^{v}/_{v}$) for 48 hours followed by autoclaving at 120°C for 15 minutes. All other material used was sterilised by using 5% sodium hypochlorite ($^{v}/_{v}$) solution.

2.2.7 Statistical analysis

Raw data recorded were analysed and graphically displayed in Microsoft Excel. Statistical analysis on the raw data of each stage that was quantified was conducted using Statistica version 13.2 (Dell). Descriptive analyses were carried out to determine the means and standard error. For raw data on different diet analyses, mean test graphs were performed in Statistica version 13.2 and one-way ANOVA tests were performed including the t-test for independence for the multiple comparisons of means. The level of significance was established at $p < 0.05$

2.3 Results

2.3.1 *Helicoverpa armigera* biology

During rearing, higher levels of moisture were observed in green bell peppers than in cabbage leaves through moisture precipitation on the rearing glass vials. This contributed to accelerated secondary infection from fungus and other microbes resulting in signs of stress from the larvae. Consequently, green bell pepper was discarded as a rearing diet and subsequent generations of larvae were reared on cabbage leaves or artificial bollworm diet.

2.3.1.1 Egg stage

The eggs were laid in clusters or singly on the oviposition cloth. The eggs were spherical in shape with longitudinal lines observed on the eggs (Figure 2.8). New eggs were white (Figure 2.8a) and older eggs that were closer to hatching were a shade of brown or cream (Figure 2.8b). The fully developed larvae were more visible prior to hatching (Figure 2.8c).

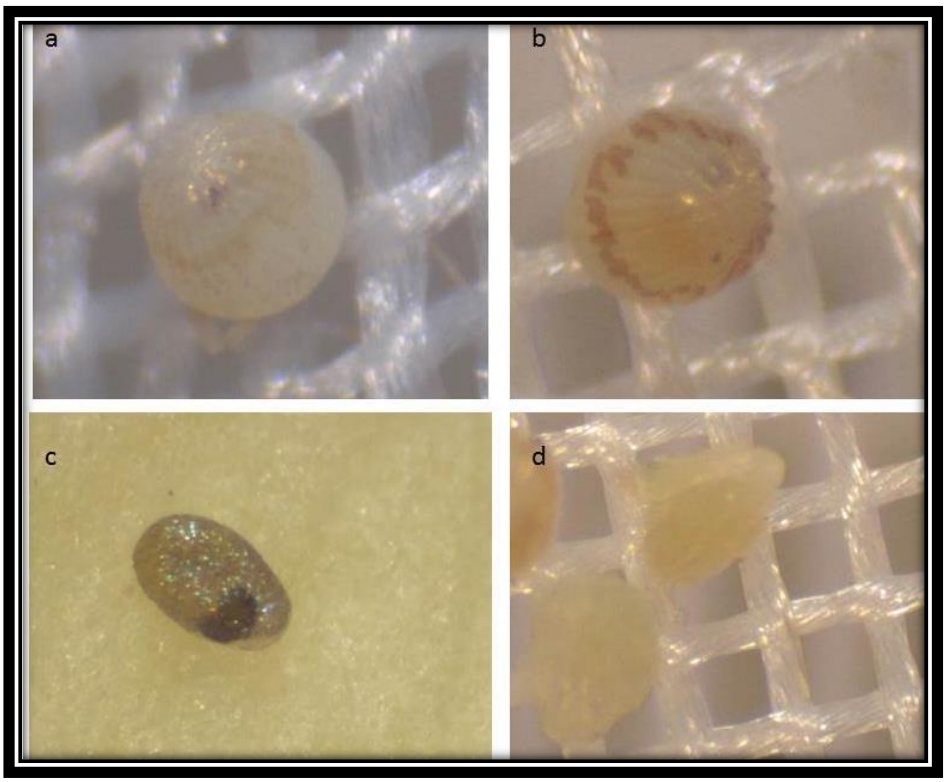


Figure 2.8: The different egg developmental stages viewed under stereo microscopy. a: a newly laid egg b: older egg darker in colour with developing larva. c: the eggs just before hatching showing the fully developed larva. d: the egg after hatching of the larva.

Some of the eggs that were laid desiccated and did not hatch (Figure 2.9). For some generations fungal infection on the eggs was also observed which affected egg survival.

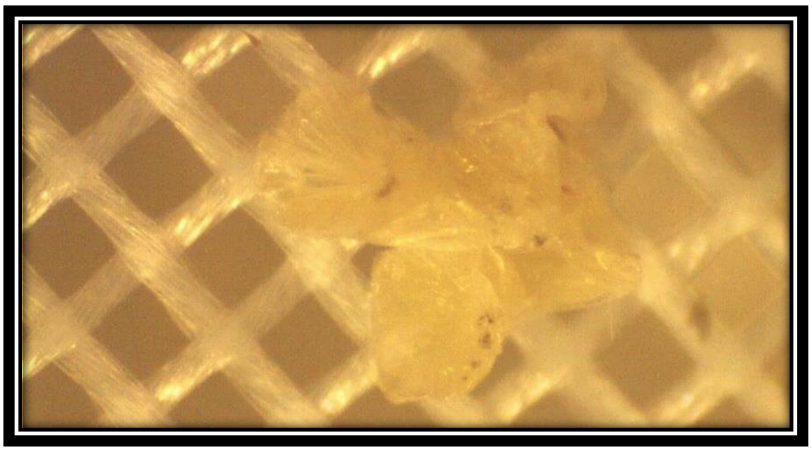


Figure 2.9: Desiccated eggs that show signs of dehydration and malformation.

The egg-hatch for the eight generations was similar with no significant difference recorded between the generations with mean percentage egg-hatch ranging from 61.7% to 77.7% ($p>0.05$) (Figure 2.10).

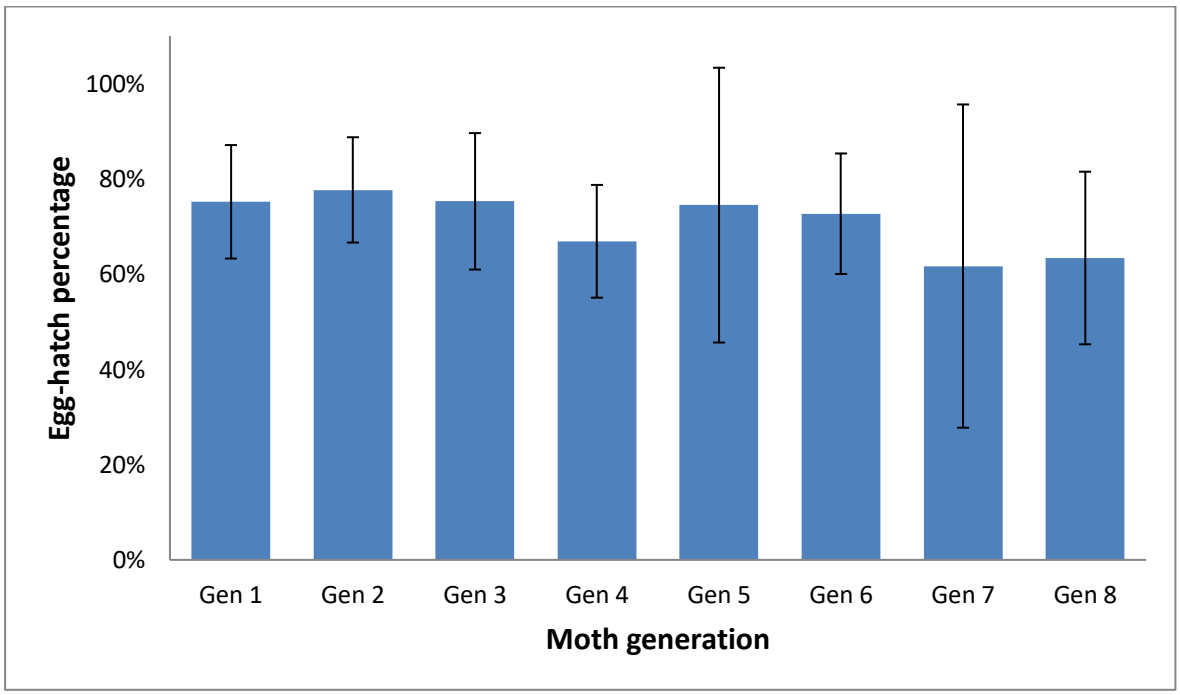


Figure 2.10: The mean percentage (± SE) of egg-hatch for the eight generations studied.

The earliest egg-hatch time was recorded to be 3 days and the latest egg hatch was recorded at day 10. The mean incubation time for the eight generations was between 4.03 (±0.57) days and 6.81 (±1.00) days with no specific pattern observed as the generations continued (Figure 2.11). This could have been a result of the temperature in the CE rooms ranging from 25°C to 28°C.

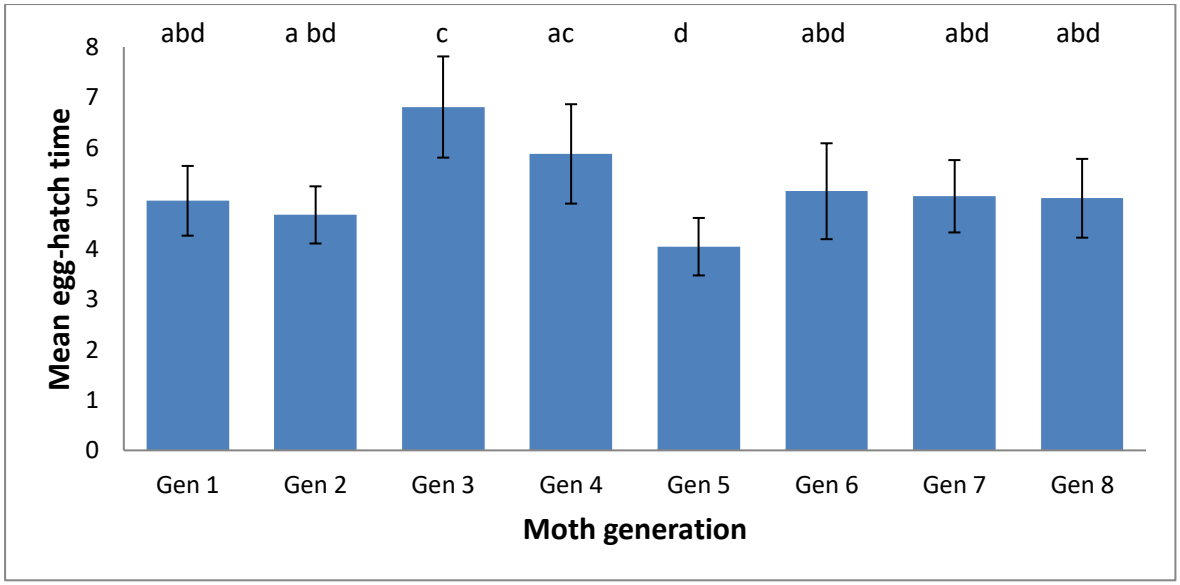


Figure 2.11: The mean duration of incubation of the eggs recorded for eight continuous generations (means and standard errors calculated from replicate experiments N=3). Columns with different letters are significantly different from each other (P<0.05).

2.3.1.2 Larval stage

Larval development and instars were determined through body length, head capsule moulting and head capsule width. The first and second instar were pale in colour with no visible body colour differentiation. The third, fourth and later instars had varying shades of brown, green or black. The first instar had a relatively large black head in comparison to the body with hairy spikes on the body (Figure 2.122a). The second instar was slightly larger than first instar, they were pale to cream with black head capsule, but a more defined colouration was recorded than in the first instar (Figure 2.122b). In Figure 2.12b, head capsule moulting is visible. The third instar colour had a darker colour, with most being brown or grey. Body markings were visible at this stage and the head capsule was black (Figure 2.122c). The fifth and sixth instar were the largest immature stages. Differences were recorded in the body colour and feeding behaviour, with fifth instar having a darker body colour (Figure 2.1212e). Whereas, the sixth instar had a pale body colour in comparison and was larger (Figure 2.1212f).

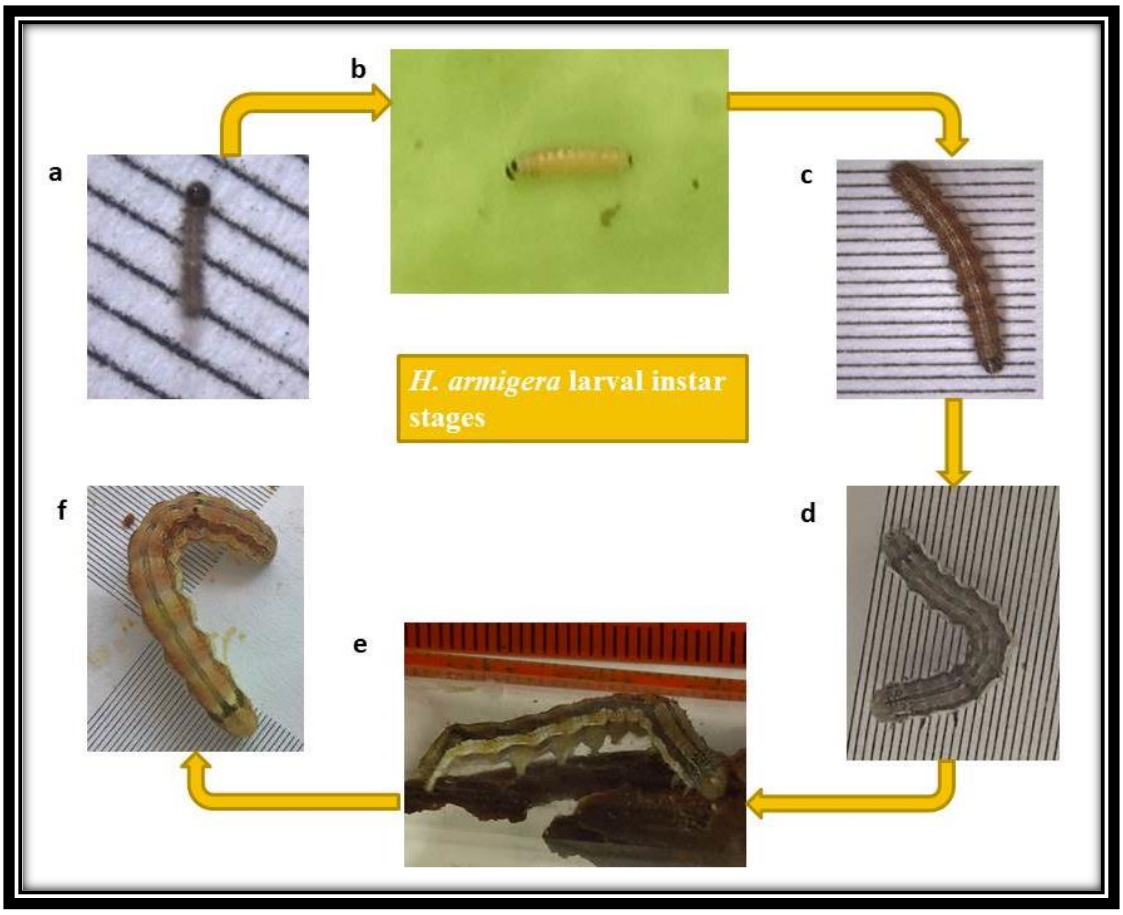


Figure 2.12: Different instars that were observed during development of *H. armigera*. The parallel lines in the pictures are 0.5 mm apart. a: first instar. b: second instar, black spot behind the head showing head capsule moulting. c: third instar. d: fourth instar. e: fifth instar. f: sixth instar.

The rate of development of larvae on artificial bollworm diet and cabbage leaves was determined. Larvae reared on artificial bollworm diet developed faster for each stage compared to larvae reared on cabbage leaves (Table 2.1). Six instars were observed on each of the diets. The first instar on artificial bollworm diet had a mean (±SE) developmental duration of 2.36 (±0.58) days compared to 5.90 (±0.15) days on cabbage leaves diet. For the second instar on artificial diet, development lasted 2.06 (±0.863) days compared to 5.86 (±0.07) on cabbage leaves diet. For the third and fourth instar, development lasted 3.24 (±0.84) days and 4.96 (±0.85) days respectively on artificial bollworm diet compared to 8.00 (±0.52) days and 9.17 (±0.36) days respectively on cabbage leaves. Development for the fifth and sixth instar on artificial diet was much longer than the early immature stage with a mean of 7.32 (±2.16) and 7.62 (±0.85) days respectively on artificial diet and 10.44 (±1.50) and 10.82 (±2.38) days respectively on cabbage leaves diet. The body lengths for each of the instars on the artificial bollworm diet and the cabbage leaves diet were similar.

Table 2.1: Developmental period in days and developmental length of *H. armigera* larvae reared on artificial bollworm diet and on cabbage leaves. Experiments were performed in triplicate. The mean (±SE) was recorded for the developmental period and the minimum and maximum body lengths was recorded. Values highlighted in red are significantly different from the preceding instar on the same diet ($p < 0.05$).

Stage	**Artificial bollworm diet**		**Cabbage leaves diet**	
	Days at stage	Body length (mm)	Days at stage	Body length (mm)
First instar	2.36 (± 0.583)	1.98-3.37 (±0.147)	5.90 (±0.153)	1.98 -4.16 (±0.147)
Second instar	2.06 (± 0.863)	3.37-9.56 (±0.275)	5.86 (±0.068)	4.16-11.0 (±0.495)
Third instar	3.24 (± 0.840)	9.56-19.3 (±1.08)	8.00 (±0.516)	11.0-21.0 (±0.941)
Fourth instar	4.96 (±0.849)	19.3-29.5 (±2.20)	9.17 (±0.363)	21.0-26.6 (±1.04)
Fifth instar	7.32 (± 2.18)	29.5-35.3 (±1.79)	10.4 (±1.60)	26.6-36.2 (±1.37)
Sixth instar	7.62 (± 0.852)	35.3 (±2.59)	10.8 (±2.89)	36.2 (±1.36)

Mean total developmental time of larvae on the artificial diet (27.1 ±5.7 days) was significantly faster than on a diet of cabbage (41.1 ±8.2 days) ($t=2.83$, $df=10$, $p=0.0179$) (Figure 2.133)

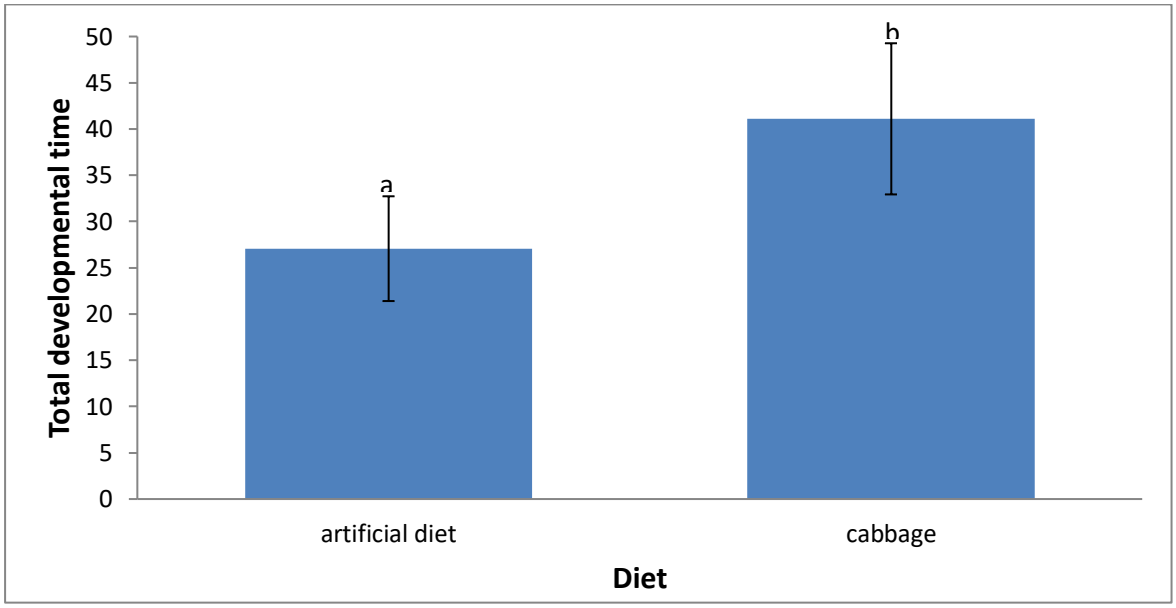


Figure 2.13: The total average developmental time of *H. armigera* from egg hatch to pupation on artificial diet and cabbage diet. N=25. (a and b represent means that are significantly different from each other; $p < 0.05$).

The head capsule measurement for each instar on artificial diet was not significantly different from those on cabbage diet. The head capsule widths on both diets follow a geometrical function expressed with the growth rate and from the assumption of Dyar's rule ($D_{i,\ i+1} = M_{i+1}/M_i$), the ratio is constant between the successive instars with a constant of 1.26 (Figure 2.144).

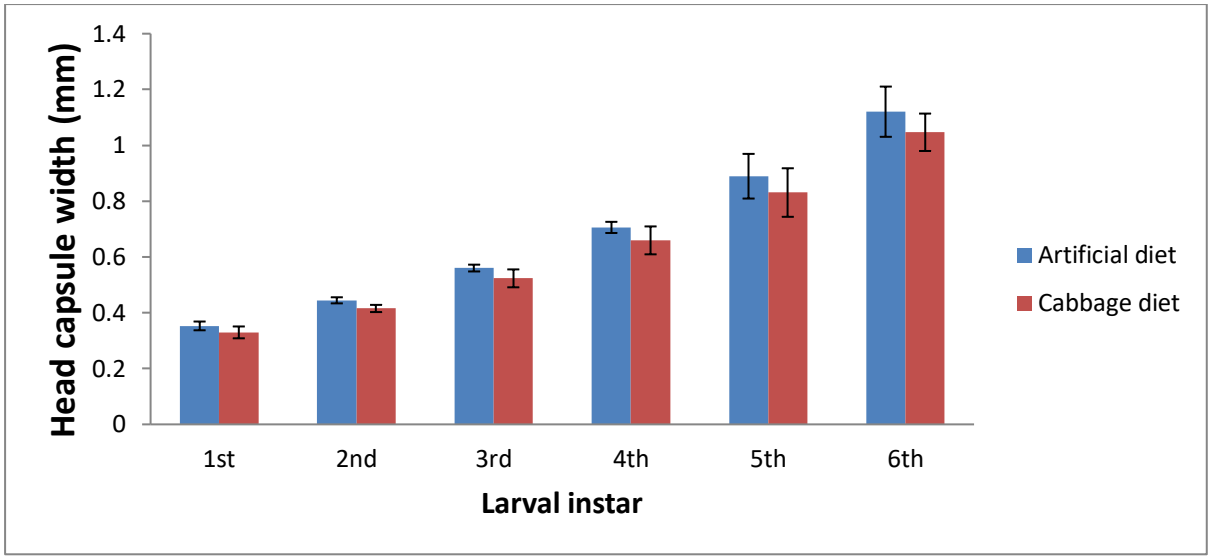


Figure 2.14: The head capsule width of the larval instar on artificial bollworm diet and on cabbage diet.

The mean survivorship of the larvae on artificial bollworm diet from egg hatch to pupation (69.3% ±7.8%) was not significantly different from mean survivorship on cabbage leaves diet (57.2% ±10.4%) ($p = 0.162$). The survivorship for each instar on both diets was significantly different (second instar $p = 0.0073$; third instar = 0.0034; fourth instar = 0.0024; fifth instar = 0.0076; sixth instar = 0.0015) (Figure 2.155).

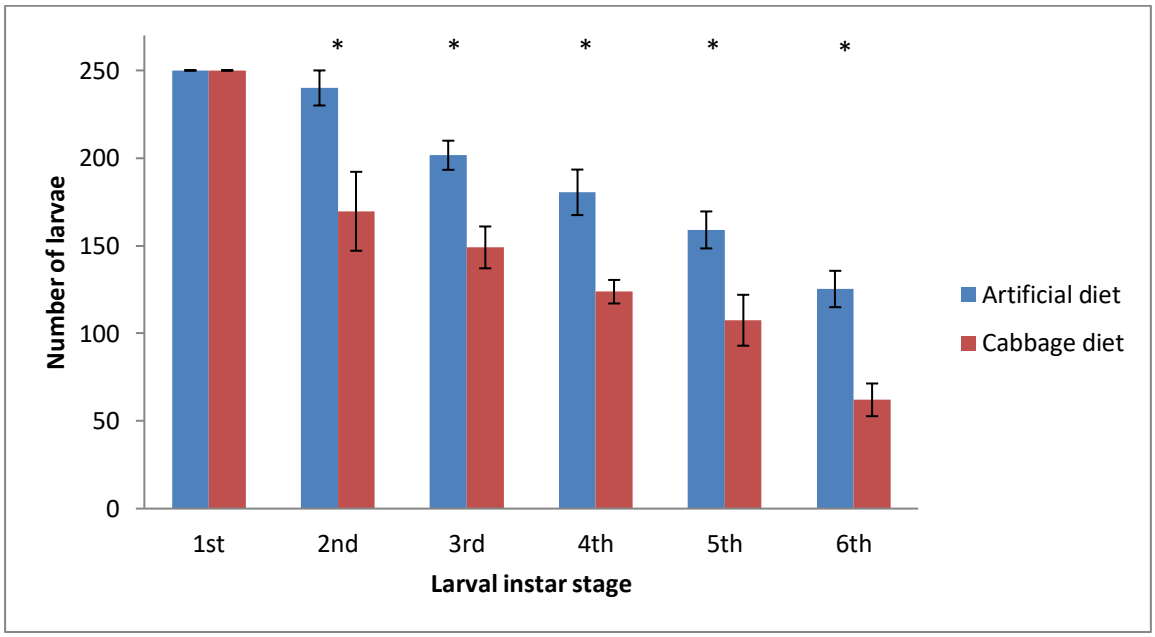


Figure 2.15: Survivorship of the immature stages from first instar to sixth instar on artificial bollworm diet and on cabbage diet. The initial colony had 250 larvae for the first instar and the experiment was conducted in triplicate. **(*** represent statistical significance for the different diets at each instar stage)

2.3.1.3 Pupal stage

For the pupal stage, only larvae reared on artificial diet were considered for analyses. Pupal eclosion was observed at the bottom of the vials in the diet. The pupal sex ratio was an average of 0.86:1.00 male: female. The mean pupal length for females was 18.6 (± 2.36) mm and for males was 20.3 (±1.72) mm. The mean pupal weight was 0.243 (±0.07) g for females and 0.229 (±0.01) g for males; no significant difference was recorded between the male and female weights (p = 0.17). The mean time for pupal eclosion was recorded to be 16.7 (±3.24) days for females and 19.1 (±4.12) days for males.

2.3.1.4 Adult stage

Mating of the moths was observed 3-5 days after female moth eclosion. Adult moth survival for males (20.8 ± 0.7 days) was not significantly different than for females (17.5 ±2.5 days) (p=0.109).

Oviposition of the eggs by adult moths was observed three days after mating, with low initial oviposition and an increased oviposition after the first 24 hours. The highest number of eggs per female was recorded for the first four generations and the lowest number of eggs per female was recorded for the next four generations with significant difference ($p<0.05$) between the first four and the next four generations (Figure 2.166).

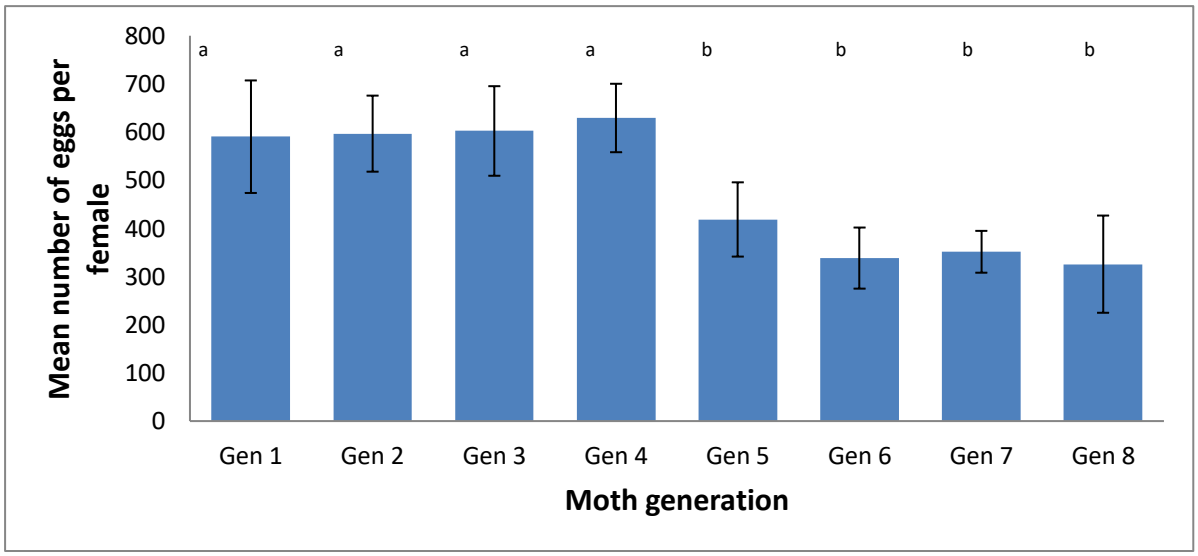


Figure 2.16: Mean (±SE) number of eggs per female for the eight generations studied. Columns with different letters represent significantly different means ($p<0.05$).

2.4 Albany and Haygrove colony development of disease-free insects

Colony build-up of disease-free insects was achieved through the separation of diseased larvae from healthy larvae in the insectary. Cannibalism for both colonies was recorded from

the third instar, which prompted separation of third instar from the colonies for later generations. From the fourth generation a minimum of 1,500 larvae for each colony were kept and maintained with minimal signs of disease recorded. Survivorship of the larvae was recorded for each generation as a ratio of pupae to initial neonates and for the Albany colony the mean was 72.1% (±8.6%) and for the Haygrove colony the mean was 68.3% (±12.5%).

2.5 Diseased larvae

Larvae that had signs of infection during field collections were placed in Eppendorf tubes. For the Albany colony, a total of 37 larvae that exhibited disease factors were collected from the field after bioprospecting and from the Haygrove colony 42 larvae collected had signs of infection. The diseased larvae were stored at -20°C. Characteristics of virus infection were glossy cuticle, discolouration and liquefaction (Figure 2.17).

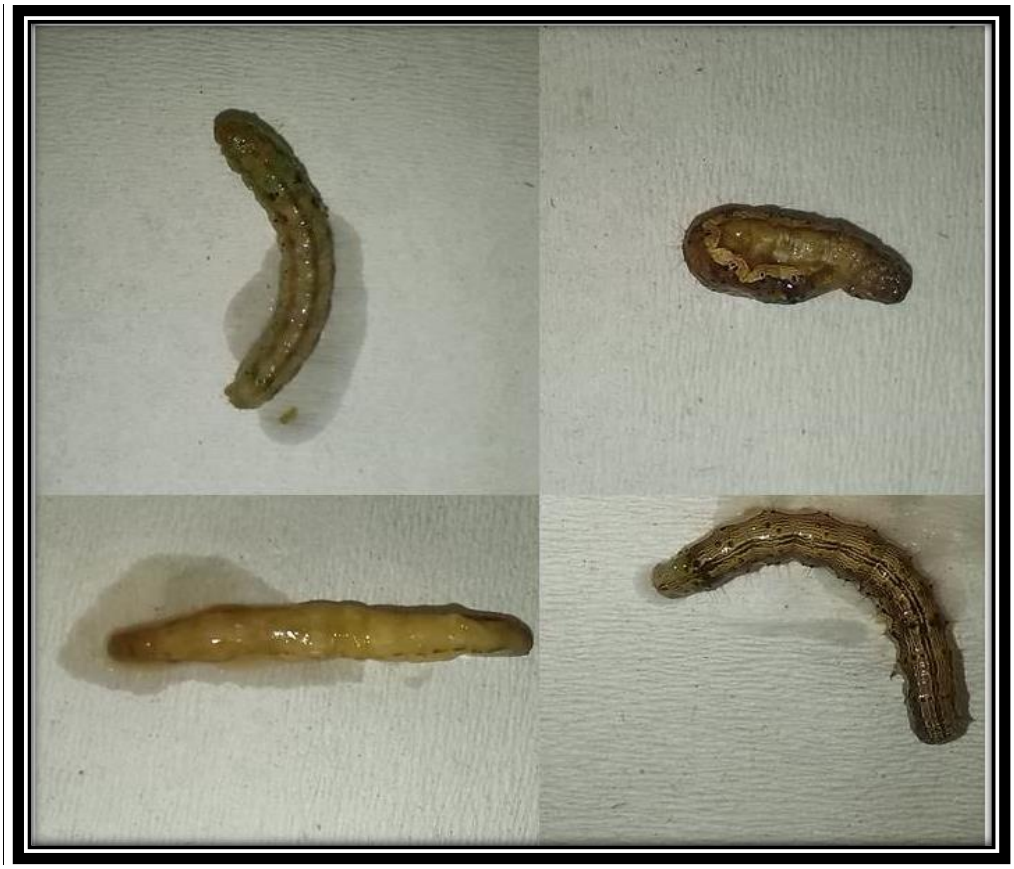


Figure 2.17: Larval cadavers with signs of baculovirus infection including discolouration and liquefaction. The glossy characteristic of the cuticle can be observed in the diseased cadavers.

2.6 Discussion

The aim of this chapter was to establish colonies of *H. armigera* populations collected from South Africa under laboratory rearing conditions. *Helicoverpa armigera* populations were

collected from two geographically distinct locations in South Africa. One colony was established from insects collected from Grahamstown in the Eastern Cape and the second colony was established from insects collected from George in the Western Cape.

Colony demise was recorded during the initial establishment of the Albany colony. Establishment of a laboratory colony is challenging when the rearing conditions are not optimal (Grzywacz *et al.*, 2007). The temperature was maintained at the optimal conditions during the rearing period in the CE rooms; however, the relative humidity that is reported in the literature is high and a lack of a humidifier and therefore constant humidity in the insectary resulted in a lower relative humidity of between 40% and 65%. Figueiredo *et al.*, (2009), reared their *H. armigera* colony at 25°C ±2°C, at 70-80% relative humidity with a 16:8 h day-night photoperiod. Gupta *et al.*, (2007), reared their colony at 26.0°C ±2.0°C and a 70%±10% relative humidity, 10:14 hour day: light cycle. Two colonies of *H. armigera* collected from South Africa were successfully established.

For the host biology, the egg stage was studied through morphology and time to hatch. The egg morphology was like reports in literature by Karel and Autrique (1989), and Hill (1987). The mean percentage egg-hatch ranged from 61.7% to 77.7% between the generations with no significant difference. The egg-hatch recorded in this study was not too different from the egg-hatch recorded by Mironidis and Savapoulou-Soultani (2008) of between 51.7% and 67.7% for temperature ranging from 20.0°C to 27.5°C. Under the laboratory rearing conditions, the mean egg-hatch time ranged from 4.03 days to 6.81 days between the generations with the mean egg-hatch for the generations of 5.19 (±0.83) days. The incubation period of the eggs was longer than that reported by Jallow and Matsumura (2001), in their developmental days versus temperature studies. They recorded that egg hatch at 27.9°C took 2.8 days (± 0.29) and at 20°C took 5.5 days (± 0.37). This could have been because of the low humidity in the insectary which would have resulted in the conditions not being optimal for accelerated egg-hatch.

Upon hatching of neonates from eggs, rearing of first instars was performed in groups on cabbage leaves. This process was also reported in literature by Mironidis and Savapoulou-Soultani (2008) whereby they reared newly hatched larvae in groups of 50-100 in plastic cups on artificial diet. The first instar larvae were also observed to be fragile, with mortality during transferring.

Developmental times of the different instars on the different diets were significantly different. Larval development and maturation were accelerated on artificial bollworm diet in comparison to cabbage diet. The average larval stage took 27 days on artificial diet and 41 days on cabbage diet. The larvae reared on cabbage leaves had to be moved every 48 hours because of fungal and microbial secondary infection that grew on the cabbage diet. This may have had a stressing effect on the larvae, therefore affecting their development. A hierarchy in the desired diet has also been observed for *H. armigera* (Smith-Pardo, 2014). The choice in diet could be based on the suitability of the food, which would result in development on a less suitable host being negatively affected. Liu *et al.*, (2004) recorded significant differences in the developmental time on different diets. Suitable diet for the insects would ensure faster developmental time and short generations, resulting in population increases. Developmental stages were determined through cuticle colouration, head capsule colour, body length and head capsule measurements. Larval stages can last between 16 and 27 days on artificial diet at temperatures of 20-25°C (Jallow and Matsumura, 2001). Developmental time of the larvae is also dependent on the diet, with observations of up to 42 days on plant-based diets (Armes *et al*, 1992; Liu *et al*., 2004).

The body lengths of each of the instars were not significantly different on the different diets. The lengths are also like those reported in the literature (Karim, 2000; Christian *et al*., 2001). Using Dyar's law, the head capsule width ratio was constant between successive instars resulting in a logarithmic straight line on both artificial diet and cabbage diet. Head capsule measurement therefore allowed for differentiation of each of the instars. The head capsule sizes reported were like those reported in literature correlating with the instar stages (Armes *et al*., 1992).

The mean survivorship from first instar to pupation on artificial diet was 50% and on cabbage leaves diet was 25%. There was no significant difference in larval survivorship from each stage. From the third instar on both diets, the survivorship had a significant fall for third instar larvae. Mortality of the larvae from the third instar could be attributed to cannibalism as well as escape from rearing vials due to enhanced activity and movement. Cannibalism has been reported to be enhanced from the third instar, with reports of older larvae travelling to different parts of the plant to attack other larvae (Karim, 2000). An increase in the population densities of the larvae has been shown to also result in increased cannibalism (Arrizubieta *et al*., 2016). Secondary mould on the diets was sometimes observed on the faecal pellets especially on the cabbage leaves. Phorid flies were attracted to the artificial diet, with phorid

flies laying eggs on the diet and maggots hatching in the rearing vials. This aided in stressing the larvae and in turn interfering with larval growth and survival. Fungi and phorid flies on the diet are important factors that may stress the larvae (Grzywacz *et al*., 2007). Survivorship for the sixth instar was mainly attributed to malformation of pupae, possibly due to stress factors, fungal factors and phorid fly infestations.

Pupation was observed at the bottom of the converted diet in burrows. The pupation of *H. armigera* in nature is observed in the soil at the base of the plant and they have been recorded to burrow into the diet to form a pupation cell (Armes *et al.,* 1992; Smith-Pardo, 2014). There was no significant difference in the sex of the pupae with an almost 1:1 ratio of male to female being observed for the pupae. The male to female ratio of the pupae was like literature reports (Armes *et al*., 1992). Pupation in the female had a shorter residence time than in the males with no significant difference being recorded in the pupation time. The pupation period was like the period recorded at 22.5°C from a study by Jallow and Matsumara (2001).

The longevity of adult moths was a mean of 19.8 (± 2.30) days for the males and 17.9 (± 1.44) for the females. These results were like reports by Mironidis and Savapoulou-Soultani (2008), where they recorded shorter lifespans for the females. They recorded the male lifespans to be 24 days at 20°C and 21 at 27.5°C and female lifespans to be 20 days at 20°C and 17 days at 27.5°C. In this study the fertility of the earlier female generation was significantly higher than for the latter generations. Fertility has been reported to be dependent on the larval diet as well as colony stability (Liu *et al.,* 2004). This could mean that as the generations progressed, the colony lost vigour.

Colony bulking up and maintenance of the Albany and Haygrove colonies was managed successfully. Infected larvae from initial collections were removed from the colonies and stored to facilitate identification of the pathogen before bioassay studies.

2.7 Conclusions

The biology of *H. armigera* collected in South Africa was studied and recorded under laboratory conditions. A study on the life cycle of *H. armigera* on cabbage and artificial diet was successfully carried out and an in depth understanding of *H. armigera* population dynamics resulted. A laboratory reared colony was successfully reared for 10 generations, which resulted in the successful establishment of Albany and Haygrove colonies for downstream experiments. Initial colony establishment resulted in the demise of the Albany colony. Conditions for rearing were improved and optimised through literature studies,

resulting in a stable colony rearing protocol which was applied for further Albany and Haygrove rearing. Conditions for rearing had to be sterile, with quality in the rearing facilities being controlled to improve colony stability and minimise host infection and colony crashing.

The analysis of the Albany *H. armigera* lifecycle through biological life studies allowed for an establishment of a suitable rearing protocol that enabled the propagation and maintenance of the colony. The rearing protocol adapted and modified for the Albany colony was utilised for the rearing of the Haygrove colony. This process ensured that healthy colonies were always available in enough quantities that would allow for further studies. The next chapter describes the protocol for baculovirus isolation and purification from diseased larvae, OB suspensions and commercial formulations.

3 Nucleopolyhedrovirus isolation and purification from geographically distinct and commercial formulations of *Helicoverpa armigera*

3.1 Introduction

Helicoverpa armigera is a major lepidopteran pest that affects crop production and is considered as extremely important in South Africa (Moran, 1983; Bell and McGeoch, 1996). Chemical pesticides are mostly used to control *H. armigera* but development of resistance by the insect has resulted in a decrease in control efficiency (Osibanjo *et al.*, 2002; Maiti and Kumari, 2016). Andermatt-Biocontrol has a commercial product for the European market from HearNPV that is registered as Helicovex®; it has been reported to be an effective control measure of *H. armigera* during field trials, resulting in up to 90% mortality (Kessler *et al*., 2008). In South Africa, over a dozen field trials have been conducted to control *H. armigera* on citrus plants using HearNPV resulting in the development of Helicovir™, a HearNPV based biopesticide manufactured by River Bioscience (South Africa) (Moore and Kirkman, 2010; Hatting *et al*., 2018). Other biopesticides based on HearNPV are available as commercial products in South Africa under the registered names Bolldex® and Graboll® (Andermatt-Biocontrol AG) (Knox *et al*., 2015; Hatting *et al*., 2018). Variants of the same baculovirus isolated from different geographical locations have been reported to have differing genetic makeup and differences in virulence (Cory *et al*., 2005). The differences in the virulence and pathogenicity of the viruses would be a useful tool in the development of biopesticides in the control of *H. armigera* and in turn limit the effects of resistance development (Moscardi, 1999). It is therefore important to isolate viruses from geographically different host populations and identify genetically different viruses for effective pest control. From Chapter 2, diseased larvae that exhibited symptoms of baculovirus infection were collected and an attempt to purify OBs and identify potential baculoviruses was made.

Characterisation is performed by extracting and purifying baculoviruses from the infected larvae. The diseased larvae are macerated and homogenised in suspension, often in waster or SDS so as to release the virus (Christian *et al.,* 2001; Cory *et al.,* 2005; Figueiredo *et al.,* 2009). Large cellular debris is removed from the homogenate through filtering or low speed centrifugation before concentration and purification of the virus (Christian *et al.,* 2001; Arrizubieta *et al.,* 2015). Virus concentration and purification is mainly performed through

gradient centrifugation or differential centrifugation (Ogembo *et al.*, 2007, Raghavendra *et al.*, 2017). For gradient centrifugation, the use of inert solutes such as glycerol or sucrose in water allow for the production of different density gradient layers (Gryzwacz *et al*., 2007). Most commercial NPV solutions contain glycerol as part of the formulation, which aids in limiting microbial contaminants growth (Behle and Birthisel, 2014). The use fo glycerol for gradient purification would be preferrable in this case as it does not introduce another factor that would need to be rinsed off. The virus size for HearNPV is relatively large in comparison to GV, which results in a significant loss of the virus in the pellet during gradient centrifugation (Grzywacz *et al*., 2007). After purification of the virus, occlusion bodies (OB) identification can be perfomed through the use of an electron or light microscope (Berreta *et al*., 2013). The size and the shape of the occlusion bodies (OB) aid in morphological identification of the baculovirus (Eroglu *et al*., 2018). Although visual studies aid in determination of the virus, molecular and genetic work allows for accurate identification (Jones, 2000).

Polymerase chain reaction (PCR) amplification is a useful tool in the accurate identification of organisms including baculoviruses (Sun *et al*., 2002). Baculovirus genomes have highly conserved genes which makes it possible to design universal primer sequences that target specific genes such as polyhedrin or granulin in different baculoviruses (Lange *et al*., 2004; Rowley *et al.*, 2011; King *et al.*, 2012). Through the examination of genome sequences of lepidopteran-specific GVs and NPVs, degenerate primers that target the highly conserved genes namely *polyhedrin/granulin* (*polh*/*gran*), *late expression factor 8* (*lef-8*) and *lef-9* were developed. The degenerate primers were designed using the conserved regions of lepidopteran alpha- and betabaculoviruses including HearNPV (Lange *et al*., 2004). A HearGV has been previously isolated from South African bollworm populations (Whitlock, 1974). Therefore, in identifying the baculovirus from the wild type *H. armigera*, it was important to use degenerate primers that could amplify both HearGV and HearNPV.

In this chapter, OBs were purified from diseased cadavers collected from South African host populations after bioprospecting in the field, commercial HearNPV formulations namely Helicovir™ and Helicovex® and concentrated OB suspensions supplied from Spain and China. The concentrated OB suspensions were HearNPV-SP1 and HearNPV-G4.

The aim of this chapter was to isolate and identify HearNPV from field collected South African *H. armigera* cadavers, semi-purified suspensions and commercial formulations as

well as HearNPV OBs first isolated in South Africa by Whitlock, (1974). The first objective of the study was to isolate and concentrate OBs from the different samples through centrifugation. The second objective of the study was to identify the nature of the baculovirus through morphological studies of the purified OBs by TEM. The third objective was to extract genomic DNA from the purified OBs and then PCR amplify the *polh/gran* region using degenerate baculovirus primers (Lange *et al*., 2004).

3.2 Methods and materials

3.2.1 Host populations used for occlusion body purification

The samples used for the analysis are described in Table 3.1. The Whitlock isolate and Helicovir™ were received from Sean Moore (Citrus Research International) and HearNPV-SP1, HearNPV-G4 and Helicovex® were provided by Primitivo Caballero (Instituto de Agrobiotecnologia, Public University of Navarra in Pamplona, Spain).

Table 3.1: Helicoverpa armigera nucleopolyhedrovirus isolate sources and dates when received, that were used in this study

Isolate name	Isolate source	Source date
Albany	Belmont Valley near Grahamstown (33°19’20.2” S; 26°38’23.4” E)	Collected from May – August 2014
KZN	Cappeny Estates in KZN (29°28'40.6"S 31°10'19.9"E)	Diseased cadavers received in September 2014
Haygrove	Haygrove Eden Farm, near George (33°57’59.0” S; 22°22’54.9”)	Diseased cadavers received in December 2016
Whitlock	Citrus Research International (1997)	Homogenate of NPV-infected larvae received in September 1997
SP1	Toledo, Spain	Concentrated OBs received in November 2014
G4	Hubei Province, China	Concentrated OBs received in November 2014
Helicovir™	River Bioscience ((Pty) ltd)	Commercial biopesticide received August, 2014
Helicovex®	Andermatt-Biocontrol AG	Commercial biopesticide received November, 2014

3.2.2 Occlusion body purification from diseased cadavers

For the OB isolation and purification from the different samples, a 50-60% glycerol gradient and differential centrifugation were used.

3.2.2.1 50-60% glycerol extraction

The 50-60% glycerol gradient extraction was performed using a modified protocol from Grzywacz *et al*., (2007). A single diseased cadaver (Albany, KZN or Haygrove) was added to 1.0 ml of 0.1% SDS ($^{w}/_{v}$) and macerated and the resulting homogenate was spun down 100 ×g for a minute. The resulting supernatant was transferred to a new tube and the loose pellet was suspended in 1.0 ml of 0.1% SDS and centrifuged at 100 ×g for 1 minute. For SP1, G4, Whitlock, Helicovir™ and Helicovex®; 500 µl of the suspensions were added to 500 µl of 0.1% SDS ($^{w}/_{v}$) in a 1.5 ml tube. The suspensions were centrifuged at 100 ×g for 1 minute and the supernatant was transferred to a new tube. To the pellet, 1.0 ml of 0.1% SDS was added and centrifuged at 100 ×g for 1 minute. The resulting supernatant was combined with the initial supernatant.

The combined supernatants were centrifuged at 2,500 ×g for 5 minutes. The supernatant was discarded, and the pellet was suspended in 1.0 ml ddH_2O and centrifuged at 3,500 ×g for 5 minutes and the final pellet was suspended in 500 µl of 20% glycerol with the supernatant being discarded.

Sterile solutions of 50% glycerol and 60% glycerol were prepared. To a fresh tube, 500 µl of 60% ($^{v}/_{v}$) glycerol was added and a similar volume of 50% ($^{v}/_{v}$) glycerol was overlaid onto this. From the suspended pellet in 20% glycerol, 100 µl was overlaid onto the 50-60% glycerol gradient and centrifuged using a benchtop microcentrifuge (STURDY SA-300 VF, Sturdy Industrial Co. Ltd.) at 12,100 ×g for 20 minutes. A white band was observed at the interface and the band was siphoned using a micropipette, transferred to a new sterile tube and diluted with 1.0 ml ddH_2O. The new suspension was centrifuged at 12,100 ×g for 30 minutes and the supernatant was discarded. The pellet was suspended in ddH_2O and centrifuged at 12,100 ×g for 30 minutes. The final pellet was suspended in 200 µl of ddH_2O.

3.2.2.2 Differential centrifugation

Differential centrifugation was performed using a modified protocol adapted from Grzywacz *et al*. (2007). A single diseased cadaver was added to 1.0 ml of 0.1% SDS ($^{w}/_{v}$) and macerated. The resulting homogenate was centrifuged at 100 × g for a minute. The resulting supernatant from the suspensions was transferred to a new tube and the loose pellet was suspended in 1.0 ml of 0.1% SDS and centrifuged at 100 × g for 1 minute. The supernatants were combined and centrifuged at 2,500 ×g for 5 minutes. The supernatant was discarded, and the pellet was suspended in 1.0 ml ddH_2O and centrifuged at 3,500 × g for 5 minutes.

For the samples provided from Spain (SP1) and China (G4), as well as the Whitlock sample 500 µl of the concentrated OBs were added to 500 µl of 0.1% SDS ($^{w}/_{v}$) in a 1.5 ml tube. For the commercial formulations Helicovex® and Helicovir™ 500 µl of the commercial formulations were added to 500 µl of 0.1% SDS ($^{w}/_{v}$) in a 1.5 ml tube.

The prepared suspensions were then centrifuged at 3,500 × g for five minutes using a benchtop microcentrifuge (STURDY SA-300 VF, Sturdy Industrial Co. Ltd.) and the supernatant was discarded. The pellet was dissolved in 1.0 ml 0.1% SDS ($^{w}/_{v}$) and centrifuged at 3,500 × g for five minutes. The supernatant was disposed, and the pellet dissolved in 1 ml ddH_2O and centrifuged at 2,500 × g for five minutes. The supernatant was discarded, and the pellet was dissolved in 500 µl ddH_2O. The dissolved pellets were pooled and mixed before being placed back in fresh 1.5 ml tubes.

3.2.3 Morphological identification of purified baculovirus

Sample preparation was performed by placing 5.0 µl of purified OBs on to formvar, carbon coated grids (Wirsam Scientific, South Africa) for 30 seconds. Excess liquid was removed by gently dabbing the grid edges with filter paper. The grids were stored overnight at room temperature and viewed the following day using a transmission electron- microscope (Libra® 120 Plus (Zeiss, Germany)). Images were captured using an Olympus Megaview^{G2} CCD camera and viewed using iTEM software (Olympus SZX16) that was part of the camera setup. The iTEM software was used to measure the size and scale of the captured images.

3.2.4 Genomic DNA extraction from purified OBs

Total genomic DNA extraction from all the isolates described in Table 3.1 was performed using a modified CTAB DNA extraction protocol adapted from Opoku-Deborah *et al.*, (2013). To 200 µl of purified OBs, 90 µl of Na_2CO_3 was added and the solution was gently mixed before incubating at 37°C in a hot water bath for 30 minutes. After the incubation period, 120 µl of 1M Tris-HCl at pH 6.8 was added sequentially followed by 50 µl of SDS (10% $^{w}/_{v}$) and 50 µl of proteinase-K at a concentration of 25 mg.ml^{-1} and the mixture was incubated for a further 30 minutes at 37 °C. After the incubation 10 µl of RNaseA at a concentration of 10 mg.ml^{-1} was added and a further incubation performed at 37°C for 30 minutes. After the incubation the suspension was centrifuged at 12,100 ×g for 3.0 minutes using a benchtop microcentrifuge. The supernatant was transferred to a fresh tube and 400 µl of CTAB buffer which had been pre-heated to 70°C was added. The CTAB buffer was prepared by adding 54 mM CTAB with 0.1 M Tris-HCl at pH 8.0, 20 mM Na_2EDTA and 1.4

M NaCl. The resulting solution was incubated at 70°C for 60 minutes and the tube was inverted to gently mix the solution every 10 minutes. After 60 minutes of incubation, 400 µl of cold chloroform (stored at 4°C) was added to the mixture and centrifuged at 6,700 × g for 10 minutes. The upper organic layer was transferred to a fresh Eppendorf tube and 400 µl of ice cold iso-propanol that had been stored at -20°C was added. The resultant solution was stored overnight at -20 °C. The overnight solution was centrifuged at 12,100 ×g for 20 minutes. The supernatant was disposed and to the resulting pellet 1.0 ml of ice cold 70% ethanol ($^{v}/_{v}$) that had been stored at -20°C was added. The resulting solution was centrifuged at 12,100 × g for 5 minutes and the supernatant was disposed. The pellet was left to air dry overnight and the final dry pellet was re-suspended in 20 µl 10 mM Tris-HCl at pH 8.0. The resulting DNA solution was stored at -20°C until required for downstream experiments.

3.2.5 DNA concentration and quality

DNA concentration was determined using a NanoDrop 2000 spectrophotometer (Thermo-Scientific, USA). The spectrophotometer was blanked using 1.0 µl of 10 mM Tris-HCl at pH 8.0. When the machine had been fully calibrated 1.0 µl of the DNA sample was placed on the optical surface of the NanoDrop 2000 and this was performed in triplicate.

Agarose gel electrophoresis was used to determine the DNA quality. To 8.0 µl of DNA, 2.0 µl of loading dye was added and the mixture was loaded on a 0.7% ($^{w}/_{v}$) agarose gel was stained with ethidium bromide. The gel was run at 90 V for 30 minutes in 1× TAE buffer (40 mM Tris-acetate, 20 mM acetic acid and 1 mM EDTA). The gel was viewed, and images captured using a UVIpro chemidoc (UVItec, UK) UV trans-illuminator.

3.2.6 PCR amplification of the *polh* gene

The *polh* region of the extracted genomic DNA from Albany, KZN, Whitlock, Haygrove, SP1, G4, Helicovir™ and Helicovex® samples was amplified using universal degenerate primers that specifically target the *gran/polh* region as described by Lange *et al.*, (2004) (Table 3.2). For the amplification Taq ReadyMix PCR kit (Kapa Biosystems, USA) was used. From the ReadyMix, 12.5 µl was mixed with 10 µM of the forward and 10 µM of the reverse primers. From the calculated DNA concentration, approximately 10 ng of DNA template was used, and the final mixture was brought to 25 µl by adding ddH_2O.

Table 3.2: Degenerate *gran/polh* primers used for PCR amplification resulting in amplicon size of between 507-510 bp (Lange *et al.*, 2004)

Gene target	Primer name	AcMNPV genome position	Tm (°C)	Sequence
Gran/polh	prPH-1	42,075-42,088	38-54	**TGTAAAACGACGGCCAGT** NRCNGARGAYCCNTT
Gran/polh	prPH-2	41,373-41,389	38-52	**CAGGAAACAGCATATGACC** DGGNGCRAAYTCYTT

Nucleotides in bold are the standard sequencing primers (-21) M13 forward and (-29) M13 reverse. N = C, A, T or G; Y = C or T, R = A or G and D = A, G or T.

3.2.7 Analysis of amplified region

The partial *polh* amplicon was sequenced by Inqaba Biotechnical Industries (Pty) Ltd (South Africa) for both the forward and reverse sequences. Ambiguous nucleotides were corrected using sequence chromatograms in Chromas lite 2.1. The PCR sequences were aligned in Mega 6.0 against each other using the *polh* gene sequence available online on GenBank (National Centre for Biotechnology Information) from HearNPV-Au (accession number JN584482.1) and HearNPV-G4 (accession number AF271059.2). For ambiguous nucleotides, the nucleotide with the higher base quality was used. The combined forward and reverse PCR sequences for each of the samples were analysed through BLAST (Basic Local Alignment Search Tool) for identification on the NCBI database.

3.3 Results

3.3.1 Occlusion body purification and morphological identification by transmission electron microscopy

Larval homogenates were subjected to either gradient or differential centrifugation during the purification of OBs (Figure 3.1). For the 50-60% zonal-rate glycerol centrifugation, a milky white band was recorded between the 50% and 60% layer and a pellet (Figure 3.1a). For the differential centrifugation a pellet with no other layer was recorded (Figure 3.1b).

Figure 3.1: Centrifugation of OBs from macerated larval cadavers. a: Pellet and milky white band formed between the 50 and 60% glycerol solutions. b: Pellet with OBs from differential centrifugation.

Isolation of the white band from the 50%-60% glycerol gradient purification resulted in baculovirus-like particles being recorded in low concentrations or being absent for different purifications in the white band (Figure 3.1a). For the 50-60% glycerol gradient, more concentrated OBs were recorded from the pellet (Figure 3.1b).

The differential centrifugation was recorded to result in the highest concentration of OBs for the purification steps (Figure 3.2c and 3.2d). The baculovirus like particles identified were irregular and polyhedral in shape, with sizes that ranged from 0.7-2.2 μm with most between 0.8-1.1. As a result, all further purification protocols used were based on differential centrifugation.

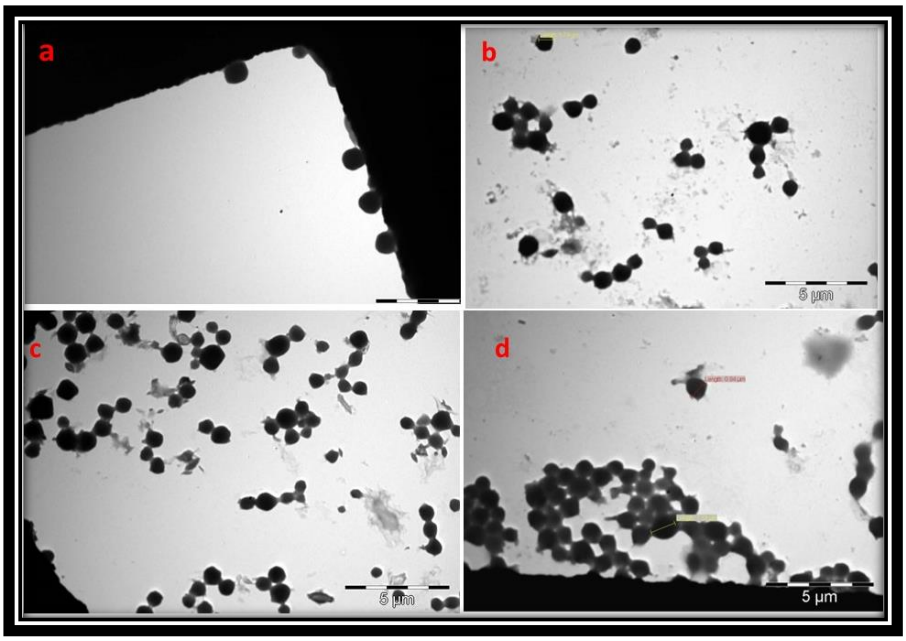


Figure 3.2: The extraction and isolation of baculovirus-like particles and morphological identification under TEM. a: OBs after 50-60% glycerol gradient purification from Albany isolate. b: OBs from the 50-60% glycerol gradient pellet from KZN isolate. c: OBs from differential centrifugation from Albany cadavers. d: OBs from KZN from differential centrifugation

3.3.2 Genomic DNA extraction from purified baculovirus like structures

Genomic DNA was extracted from the OBs purified using differential centrifugation for the isolates (Table 3.1). From the A260/A280 ratio calculated using the NanoDrop 2000, DNA with a ratio of between 1.8 and 2.0 was viewed using AGE (Figure 3.3).

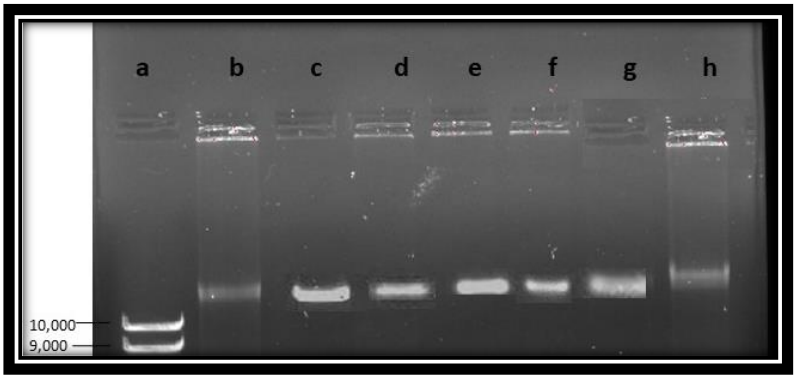


Figure 3.3: Genomic DNA extracted from purified Albany and KZN baculovirus-like particles. a: 1 kb GeneRuler™ DNA ladder with the top band at 10,000 bp. b: Helicovir™ DNA band. c: Albany DNA band. d: KZN DNA band. e: Whitlock DNA band. f: Haygrove DNA band. g: SP1 DNA band. h: G4 DNA band

3.3.3 PCR amplification using degenerate *gran/polh* primers

Using degenerate *gran/polh* primers as described by Lange *et al*., (2004), PCR amplification was performed on the extracted genomic DNA from isolates described in Table 3.1. The *polh* region of the isolates was successful amplified (Figure 3.4), however, degradation was recorded for the Helicovir™ isolate (Figure 3.4j).

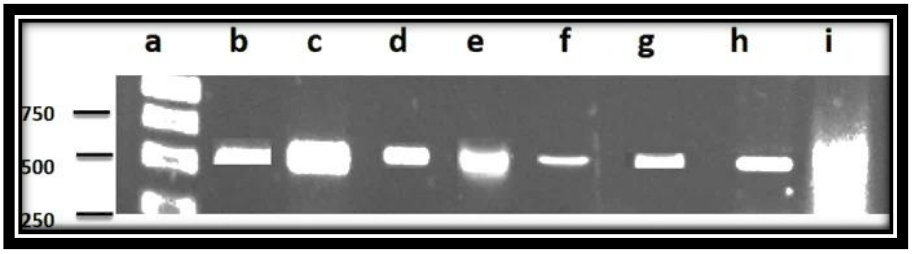


Figure 3.4: PCR amplification of the *polh/gran* region. a: GeneRuler 1 kb DNA ladder. b: Albany isolate. c: KZN isolate. d: Whitlock isolate. e: Haygrove isolate. f: SP1 isolate. g: G4 isolate. h: Helicovex isolate. i: Helicovir™ isolate

The PCR products for the KZN extracts and Albany extracts genomic DNA were sequenced. The resulting sequences after trimming and replacement of ambiguous nucleotides had sizes ranging from 494 bp to 505 bp (consensus sequence in Appendix A). Analysis of the PCR sequences through the BLAST tool confirmed the identity as the *polh* region belonging to Helicoverpa zea NPV or HearNPV with an identity of 99% and an E value of 0.0 were observed for a query cover of 93% (Table 3.3).

Table 3.3: BLAST results of PCR amplified *polh/gran* region of Albany and KZN extracted genomic DNA using degenerate primers

PCR sample	BLAST description	BLAST total score	BLAST query cover	BLAST E value	BLAST identity	Accession number
KZN	HzSNPV	931	93%	0.0	100%	KM596835.1
	HearSNPV	931	93%	0.0	100%	KJ922128.1
	HearNPV	915	93%	0.0	100%	KJ707030.1
Albany	HzSNPV	934	93%	0.0	100%	KJ922128.1
	HearSNPV	926	93%	0.0	100%	KU738904.1
	HearNPV	920	93%	0.0	100%	JN584421.1

3.4 Discussion

In this chapter, OBs were successfully extracted and purified from diseased *H. armigera* larval cadavers, concentrated OB suspensions and commercial formulations. Morphological

identification of the extracted OBs was performed through TEM. The morphology and size of the suspected NPV OBs from Albany, KZN, Whitlock and Haygrove were like that of HearNPV-SP1 and HearNPV-G4 as well as OBs from Helicovir™ and Helicovex®. The sizes of the OBs ranged from 0.7-2.2 μm. The shape and sizes correlated to HearNPV extracted OBs (Tang *et al*., 2012). The new South African isolates were referred to as HearNPV and were referred to as HearNPV-Albany, HearNPV-Haygrove, HearNPV-KZN and HearNPV-Whit.

Total genomic DNA was successfully extracted from all the HearNPV samples studied in this chapter and genomic DNA bands were observed on agarose gels. Using degenerate primers, the *polh* region of HearNPV-Albany, HearNPV-Haygrove, HearNPV-Whit and HearNPV-KZN was amplified. The PCR products for HearNPV-Albany and HearNPV-KZN were sequenced and, after assembly, the sequences had sizes of between 494 and 505 bp. This was like the sequenced PCR products using the degenerate primers which averaged 507-510 bp by Lange *et al*., (2004). The PCR product sequences were analysed using BLAST, and the identified as Helicoverpa zea or Helicoverpa armigera SNPV. HzSNPV and HearSNPV are closely related, with high amino acid identity observed in some genes including *polh* (Chen *et al*., 2002; Rowley *et al*., 2011). These two NPVs share the same heliothine species host range, although biological activity differs against individual heliothine species (Chen *et al*., 2002).

3.5 Conclusions

The aims of this Chapter were to purify, isolate and identify HearNPV OBs from field collected South African *H. armigera* cadavers that were collected after bioprospecting and compare them against HearNPV OBs from Spain, China and commercial formulations through morphological studies. Using purification protocols such as 50-60% glycerol gradient and differential centrifugation it was observed that OBs were purified in significant quantities using the latter method in comparison to the former. The morphology identified through TEM was typical of NPV. The use of molecular and bioinformatics tools such as genomic DNA extraction, PCR amplification and BLAST confirmed the Albany and KZN isolates as HearNPV.

The next chapter describes further molecular characterisation of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit by full genomic sequencing as well as REN analysis of genomic DNA to determine if the South African isolates are novel.

4 Full genome sequencing and analysis of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit isolates

4.1 Introduction

Potentially novel HearNPV isolates were identified as described in Chapter 3 from various South African sources. The identification of novel isolates is important as it offers potential candidates for effective control of *H. armigera* globally, including in South Africa. Characterisation is important in identifying novel viral isolates that may be used in developing new biopesticides for target pests. It has been reported that susceptibility of local host pests may vary against geographically different isolates and it is therefore necessary to identify new viral strains from local geographical populations as methods for biocontrol (Figueiredo *et al*., 2009; Opoku-Debrah *et al*., 2016).

For accurate identification of the virus, molecular analysis is required (Jones, 2000). Complete genome analysis and comparative genome studies of baculoviruses can aid in the inference of their phylogeny and possible inseciticidal activity (Herniou *et al*., 2001). Determining the structure of the virus genome at the nucleotide sequence level, enables isolation of mutated nucleotide variants from each gene which in turn directly allows the analysis of mutational effects on a phenotypic level (Condit, 2007). It is important to understand the role of genome diversity on virulence as well as insect-pathogen dynamics and evolution (Cory *et al*., 2005). The replication of baculovirus in host cells results in an interaction of the virus and the host resulting in long-term co-evolution (Theze *et al.*, 2018). Comparative genome analysis may reveal factors that are adaptive (Koonin *et al.*, 2000), therefore giving insight into insect-pathogen evolution. Genetic variation of HearNPV may result in different phenotypic traits that include pathogenicity and speed of kill (Baillie and Bouwer, 2013). HearNPV is a well characterised and commercially available virus in many countries (Baillie and Bouwer, 2013) and genotypic variation has been reported for HearNPV from the same location as well as geographically distinct locations with strains showing varying levels of virulence (Cory *et al*., 2005; Figueiredo *et al.*, 2009; Baillie and Bouwer, 2013).

For genomics to be carried out, sequencing, mapping and assembly of the genomes to be studied is required. The process of sequencing, mapping and assembly requires reference genome strains for alignment and qualitative nucleotide selection and this can be performed

by extracting sequenced genomes from online databases such as NCBI's GenBank, EMBL Nucleotide Sequence Database and DNA Data Bank of Japan (Dubey, 2014). The use of more than one closely related genome for assembly results in reads that have a higher quality (Ekblom and Wolf, 2014). There are 22 fully published genome sequences for HearNPV currently available on the NCBI database, GenBank from Spain, China, Australia, Kenya and India. The use of HearNPV genomes as references during assembly would result in higher quality reads. Genome annotation and prediction for the raw genomic data can be performed by aligning single genes or fully sequenced genomes to available gene and genome sequences on GenBank. Gene sequences of conserved HearNPV genes such as *polh, lef-8, lef-9, egt* and *p10* as well as full genomic sequences of isolates such as HearNPV-G4 and HearNPV-Au are available on GenBank and can be used for HearNPV *de novo* assembly and quality assessment (Chen *et al*., 2001; Ogembo *et al*., 2009; Raghavendra *et al.*, 2009; Zhang *et al*., 2014).

There are various methods that can be used for DNA sequencing, including classical Sanger sequencing and next generation sequencing (Men *et al*., 2008). Genomic sequencing using the Illumina sequencer allows for multiple fragments of DNA to be sequenced simultaneously, resulting in high quality bases (Lakdawalla and VanSteenhouse, 2008). Next generation sequencing allows for rapid sample preparation and an enhanced throughput (Du and Egholm, 2008). This sequencing technology is relatively fast, has a read length of 300 bp which reduces the data analysis time and is suitable for small genome analysis such as for HearNPV (El-Metwally *et al*., 2014).

There are several methods that can be used in the analysis of full genomes such as generation of restriction endonuclease (REN) profiles, which is commonly used for comparative purposes (Lua *et al*., 2002; Figueiredo *et al*., 2009; Arrizubieta *et al.*, 2013). The use of REN also allows identification of polymorphisms between genomes that are divergent and may aid in evolutionary studies and analysis (Cory *et al*., 2005). Studies involving REN in genomics allow for considerable resolution in genome diagnostics. It is a very useful method for distinguishing distinct HearNPV viral DNA, as each enzyme creates a characteristic pattern that is unique to the genome (Rohrmann, 2013; Rao *et al.*, 2015). Enzymes can have many restriction sites on a genome creating more than one fragment after digestion, resulting in genetic maps being created. This allows for the comparison of genetic maps that have been created and an in-depth analysis of the genomes through in silico comparative studies

(Camoretti-Mercado, 2008). REN analysis to create genomic profiles is commonly used as a method for identification of different HearNPV genomes (Jones, 2000).

The aim of this chapter was to sequence and assemble the genomes of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit using two reference strains from geographically distinct locations, namely HearNPV-Au and HearNPV-G4. The second aim of the chapter was to identify the novelty of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit. The first objective of the study was to assemble the sequenced genomes of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit through mapping against HearNPV-Au and HearNPV-G4. The second objective was to annotate the newly sequenced and assembled genome ORFs through the use of HearNPV-Au as the reference. The final objective was to perform comparative studies of the assembled genomes to HearNPV-Au through ORF alignment and REN to identify novelty of the viruses.

4.2 Methods and materials

4.2.1 Complete genomic sequencing and analysis of South African HearNPV isolates: HearNPV-Albany, HearNPV-KZN and HearNPV-Whit

Samples of genomic DNA for HearNPV-Albany, HearNPV-KZN and HearNPV-Whit extracted as described in Chapter 3 were sent for sequencing. The sequencing was performed by Inqaba Biotechnical Industries (Pty) Ltd (South Africa) using next generation DNA sequencing (NGS) technology, by using an Illumina® MiSeq desktop sequencer. Approximately 200 ng of the genomic DNA of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit were used for the sequencing. Using Geneious 7.1.7 (New Zealand) the reads produced were paired and assembled *de novo*. HearNPV-Au (GenBank Accession number JN 584482) and HearNPV-G4 (GenBank Accession number AF271059) genomes were used as the reference sequences for assembly and HearNPV-Albany, HearNPV-KZN and HearNPV-Whit were mapped against them. Medium sensitivity was used for the assembly and from the contigs a single consensus sequence was generated for HearNPV-Albany, HearNPV-KZN and HearNPV-Whit. Ambiguous nucleotides from the samples were analysed and corrected manually, on individual nucleotides using the highest quality compared to the two reference sequences.

Pairwise multiple alignments were performed on the generated consensus sequences and open reading frames (ORFs) were predicted when mapped against HearNPV-Au. The pairwise nucleotide and amino acid alignments were performed using ClustalW. Annotation

and prediction of the ORFs from the sequenced genomes was performed in Geneious 7.1.7 at a setting identifying 25% similarity using HearNPV-Au as the annotation and alignment reference. Predicted ORFs were exported to MS Excel 2010 edition in the csv format with ORF sizes, percentage similarity and position highlighted.

4.2.2 ORF variability and analysis of HearNPV-Albany and HearNPV-KZN against HearNPV-Au

The ORFs for HearNPV-Albany genome and HearNPV-KZN genome that had been annotated and predicted from using HearNPV-Au genome as the reference strain were organised and grouped according to size and analyses were performed in MS Excel 2010. ORFs were labelled as small (150 to 999 bp), medium (1,000 to 1,999 bp) and large (>2,000 bp). The average and standard deviation of the percentage similarity from each of the groups was calculated, and ORFs that were significantly different were reported.

4.2.3 *In silico* restriction endonuclease analysis (REN) of HearNPV-Albany, HearNPV-KZN and HearNPV-Au

The process of *in silico* digestion of full genomic DNA is explained by Camoretti-Mercado, (2008) and this method was adapted for the REN studies. The complete genomic sequences of HearNPV-Albany, HearNPV-KZN and HearNPV-Au were analysed through *in silico* REN using A plasmid Editor (ApE) version 2.0.47. The enzymes selected have been reported in the literature, namely *Eco*RI (Arrizubieta *et al*., 2013), BglII and PstI (Zhang and Wu, 2001; Figueiredo *et al*., 2009) and BamHI, HindIII, KpnI and XbaI (Zhang and Wu, 2001). For each of the enzymes selected, the genomes were digested, and the virtual profiles were aligned to highlight similarities and differences in fragment sizes.

4.3 Results

4.3.1 Complete genomic sequencing of the South African HearNPV isolates, HearNPV-Albany and HearNPV-KZN

Genomics were conducted for HearNPV-Albany and HearNPV-KZN and HearNPV-Whit. For effective genome sequencing, mapping and assembly against reference strains is required (Dubey, 2014). Using NGS technology, the genomic DNA for HearNPV-Albany, HearNPV-KZN and HearNPV-Whit was sequenced, mapped and assembled using HearNPV-Au and HearNPV-G4 as the reference strains.

4.3.1.1 HearNPV-Albany genomic assembly and genome organisation

Sequencing of HearNPV-Albany resulted in the generation of 1,742,702 paired reads and 10,002 contigs. The largest contig had a sequence length of 10,566, which had been assembled from 2,435 sequences. The HearNPV-Albany genome was assembled into a contiguous sequence with a length of 130,360 bp, a GC content of 39.1% and contained 135 ORFs that could encode for proteins with more than 50 amino acids. The final consensus of HearNPV-Albany had a 95.6% pairwise identity with HearNPV-Au and a 94.9% pairwise identity with HearNPV-G4 genome.

4.3.1.2 HearNPV-KZN genomic assembly and genome organisation

Sequencing of HearNPV-KZN resulted in the generation of 580,110 paired reads and 1,168 contigs were produced. The largest contig had a sequence length of 72,218 nucleotides and had been assembled from 209,089 sequences. The HearNPV-KZN genome was assembled into a contiguous sequence with a length of 130,636 bp, a GC content of 38.9% and contained 132 ORFs that could encode for proteins with more than 50 amino acids. The final consensus had a 96.1% sequence identity with HearNPV-Au genome and a 95.9% identity with the HearNPV-G4 genome.

4.3.1.3 HearNPV-Whit genomic assembly and genome organisation

The sequencing of HearNPV-Whit resulted in the generation of 619,004 paired reads and 10,002 contigs. The largest contig had a sequence length of 58,817, which was assembled from 150,642 sequences. The HearNPV-Whit assembly identified many gaps with only 51.2% of the genome successfully assembled and 64 partial and complete ORFs identified.

4.3.2 HearNPV-Albany, HearNPV-KZN and HearNPV-Whit genome organisation and ORF annotations

Annotation of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit was performed using HearNPV-Au genome as a reference strain (Accession number JN 584482) (Zhang *et al*., 2014) that is available on the NCBI's GenBank database. The sequence alignments between the South African HearNPV genomes against HearNPV-Au genome identified potential putative functions of the genes from the corresponding homologues. After annotation and alignment, the gene sizes, gene direction and percentage identity of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit against HearNPV-Au reference strain were deduced, and results were tabulated (Appendix B).

4.3.2.1 HearNPV-Albany genome assembly after alignment to HearNPV-Au

For HearNPV-Albany, a total of 135 ORFs were identified compared to the 133 ORFs in HearNPV-Au genome. In the HearNPV-Au genome, ORF 58 (*bro-a*) is available, but the ORF is absent in HearNPV-Albany genome (Appendix B). ORFs 26, 42 and 121 (*hypothetical proteins*) are observed in HearNPV-Albany genome but absent in HearNPV-Au genome. Of the 135 ORFs in the HearNPV-Albany genome, 70 were forward ORFs and 65 were reverse ORFs (Figure 4.1). In the HearNPV-Albany genome, four *hr* regions *hr1, hr2, hr4* and *hr5* were identified with varying identity with HearNPV-Au genome.

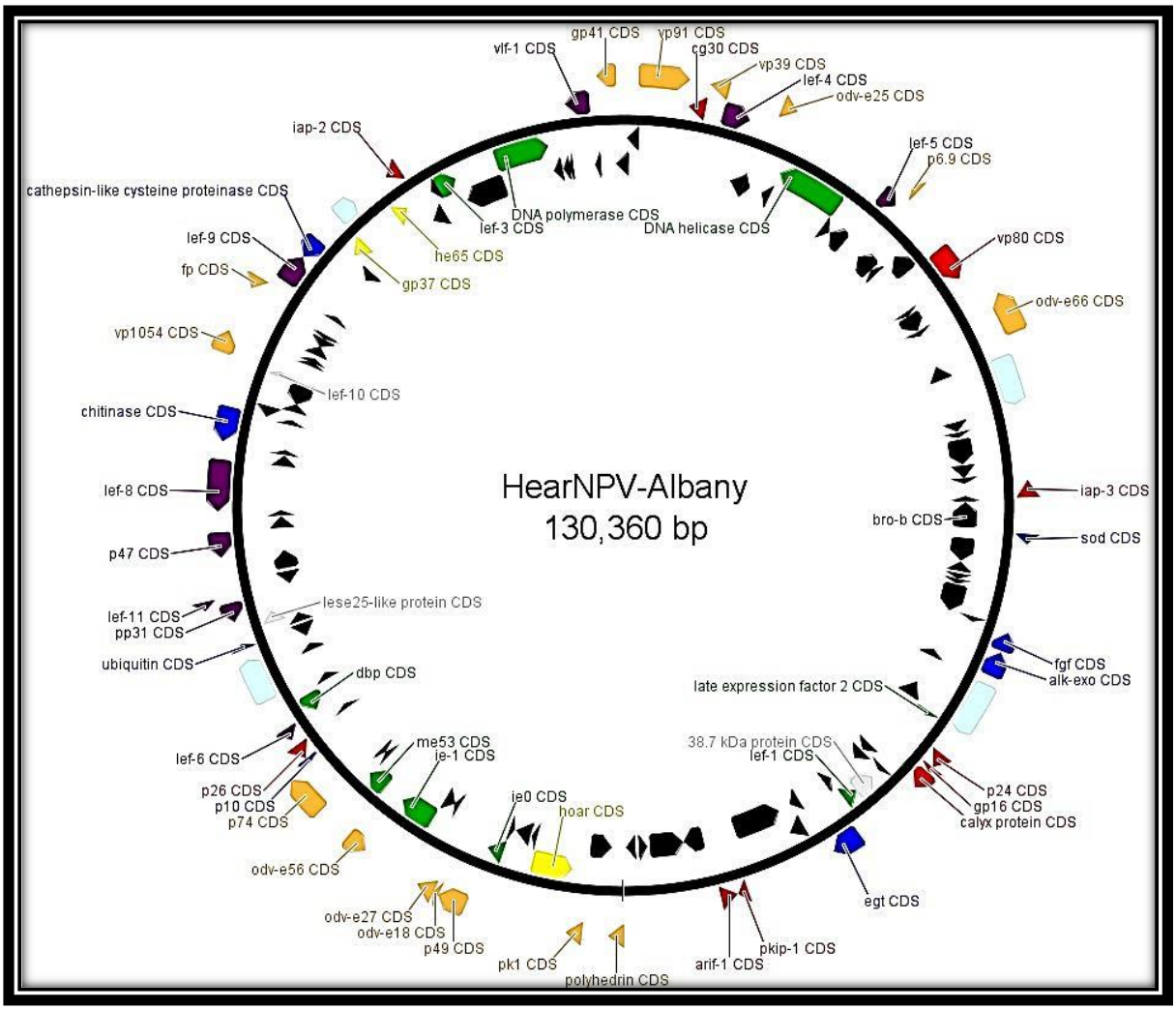


Figure 4.1: HearNPV-Albany genome after annotating using HearNPV-Au as the reference strain. Green: replication genes; Purple: transcription genes; Orange: structural genes; Dark blue: auxiliary genes; Grey: unknown function; Red: unique to lepidopteran NPVs; Yellow: specific to some NPVs including HearNPV; Black: *bro* genes and hypothetical proteins with no name label; Light blue: *hr* repeats. The map was generated in Geneious 7.1.

4.3.2.2 HearNPV-KZN genome assembly after alignment to HearNPV-Au

For HearNPV-KZN, a total of 132 ORFs were identified compared to the 133 ORFs in HearNPV-Au genome. In the HearNPV-Au genome, ORF 56 and 57 (*hypothetical protein* and *gp37*) are available but are absent in the HearNPV-KZN genome (Appendix B). ORF 118 (*hypothetical protein*) is present in HearNPV-KZN but absent in the HearNPV-Au genome. Of the 132 ORFs in the HearNPV-KZN genome, 69 were forward ORFs and 63 were reverse ORFs (Figure 4.2). In the HearNPV-KZN genome, five *hr* repeats were identified with varying identity in comparison to HearNPV-Au *hr* repeats.

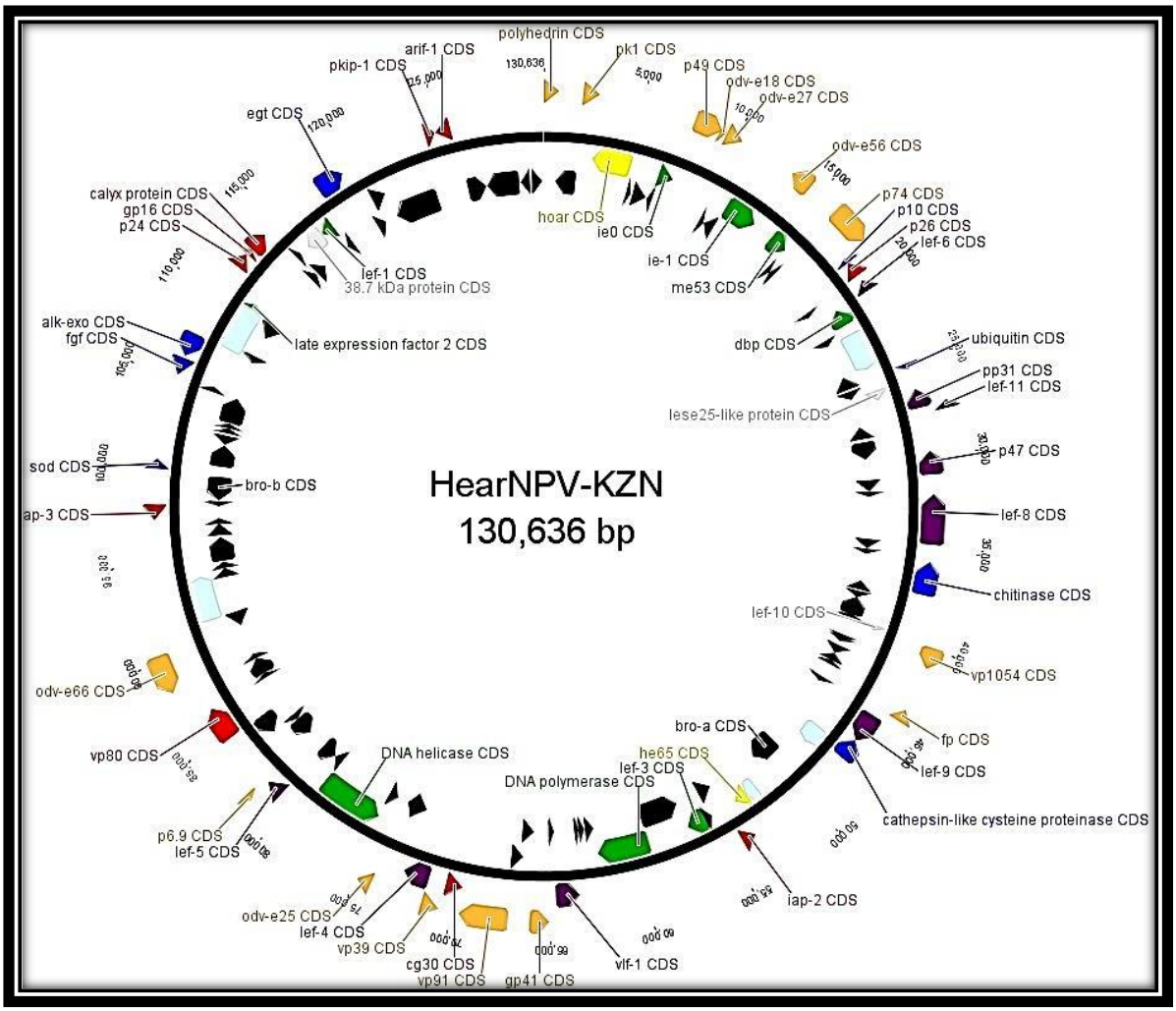


Figure 4.2: HearNPV-KZN genome after annotating using HearNPV-Au as the reference strain. Green: replication genes; Purple: transcription genes; Orange: structural genes; Dark blue: auxiliary genes; Grey: unknown function; Red: unique to lepidopteran NPVs; Yellow: specific to some NPVs including HearNPV; Black: *bro* genes and hypothetical proteins with no name label; Light blue: *hr* repeats. HearNPV-Whit genome assembly and putative gene function in comparison to Australian HearNPV. The map was generated in Geneious 7.1.

4.3.2.3 HearNPV-Whit genome assembly after alignment to HearNPV-Au

For the HearNPV-Whit genome, the sequence quality was low with many gaps in the sequenced data. Data that were recovered successfully from the genome were aligned against HearNPV-Au genome and potential ORFs were predicted. For the HearNPV-Whit genome, a potential 70 ORFs were identified from referencing to the HearNPV-Au genome. However, from the identified ORFs, some ORFs were not included in the analyses because of incomplete sequencing (Appendix B). From the HearNPV-Whit genome, one full *hr2* repeat was identified, with a length of 1,163 and percentage similarity of 45.19% when compared to the Australian genome. Full genomic analysis of the HearNPV-Whit genome was discontinued, and any further analysis was performed on completely sequenced ORFs.

4.3.3 HearNPV-Albany and HearNPV-KZN ORF variability and significant nucleotide difference in comparison to HearNPV-Au ORFs

For detailed comparative analyses of the genomes, a comparison of the ORFs from HearNPV-Albany and HearNPV-KZN against HearNPV-Au was carried out. The ORF studies were meant to identify the presence or absence of ORFs or nucleotide variability in the coding sequences between the HearNPV genomes.

4.3.3.1 HearNPV-Albany ORF analysis and comparison to HearNPV-Au

For the 135 identified putative ORFs in the HearNPV-Albany genome, 91 had sizes of between 150 bp and 999 bp; 28 ORFs were between 1,000 bp and 1,999 bp and the remaining 16 ORFs had sizes above 2,000 bp. For the small ORFs, the mean ORF identity was 97.25%, with a standard error of 7.45%. From the small ORFs, seven had a similarity lower than 89.80% compared to the corresponding HearNPV-Au ORFs. Of the ORFs, five were for hypothetical proteins and two with a putative function were for *pp31* and *ubiquitin* (Figure 4.3a). For the medium sized ORFs, the average similarity was 98.88% with a standard error of 1.26%. Only three ORFs were observed to have a similarity lower than 98.47%, with one being a *hypothetical protein* and two having a known putative function. The two ORFs were *p47,* with a similarity of 94.56%, and *38.7 kDa protein,* with a similarity of 97.28% (Figure 4.3b). For the large ORFs the average compared similarity was 98.71% with a standard error of 1.84% and only the *hoar* ORF with a similarity of 93.52% had a low similarity that 98.5% (Figure 4.3c).

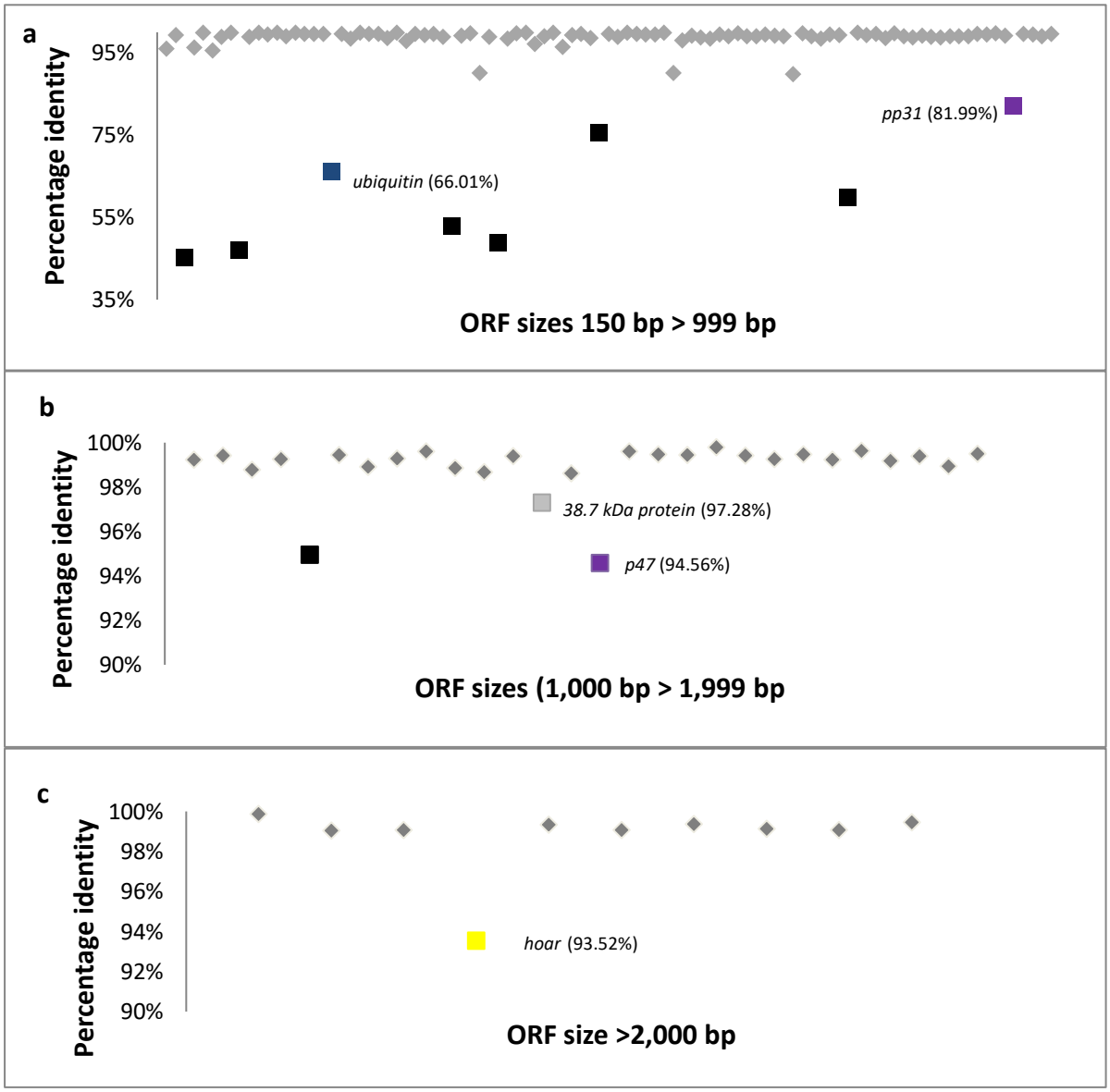


Figure 4.3: Percentage identity of the putative ORFs in the HearNPV-Albany genome compared to HearNPV-Au genome. a: ORFs with sizes below 1,000 bp. b: ORFs observed between 1,000 to 1,999 bp. c: ORFs with sizes above 2,000 bp. ORFs that were different had different data points and for ORFs with a putative function the data points were colour coded according to function as described in Figure 4.1

4.3.3.2 HearNPV-KZN ORF analysis and comparison to HearNPV-Au

For the 132 identified putative ORFs in the HearNPV-KZN genome, 86 had sizes of between 150 bp and 999 bp. Of the 86 ORFs, the mean ORF similarity to the HearNPV-Au ORFs was 98.90%, with a standard error of 5.67%. Only one ORF for a hypothetical protein with a similarity of 45.22% was significantly different (Figure 4.4a). For medium sized ORFs (1,000 bp to 1,999 bp), 29 ORFs were identified with the remaining ORFs being larger than 2,000 bp. For the medium sized ORFs, the average similarity was 96.26% with a standard deviation

of 10.32%. From the ORFs, three were observed to have a percentage similarity lower than 98.5% (Figure 4.4b). The three ORFs had a known function, being for *cathepsin-like protein*, with a similarity of 50.45%, the *bro-a*, with a similarity of 78.69%, and the *ie-1*, with a similarity of 78.37%. For the large ORFs, the average identity was 99.03%, with a standard error of 1.13%. Only the *hoar* ORF with an identity of 95.89% in comparison to HearNPV-Au *hoar* gene had a similarity lower than 98.5% (Figure 4.4c).

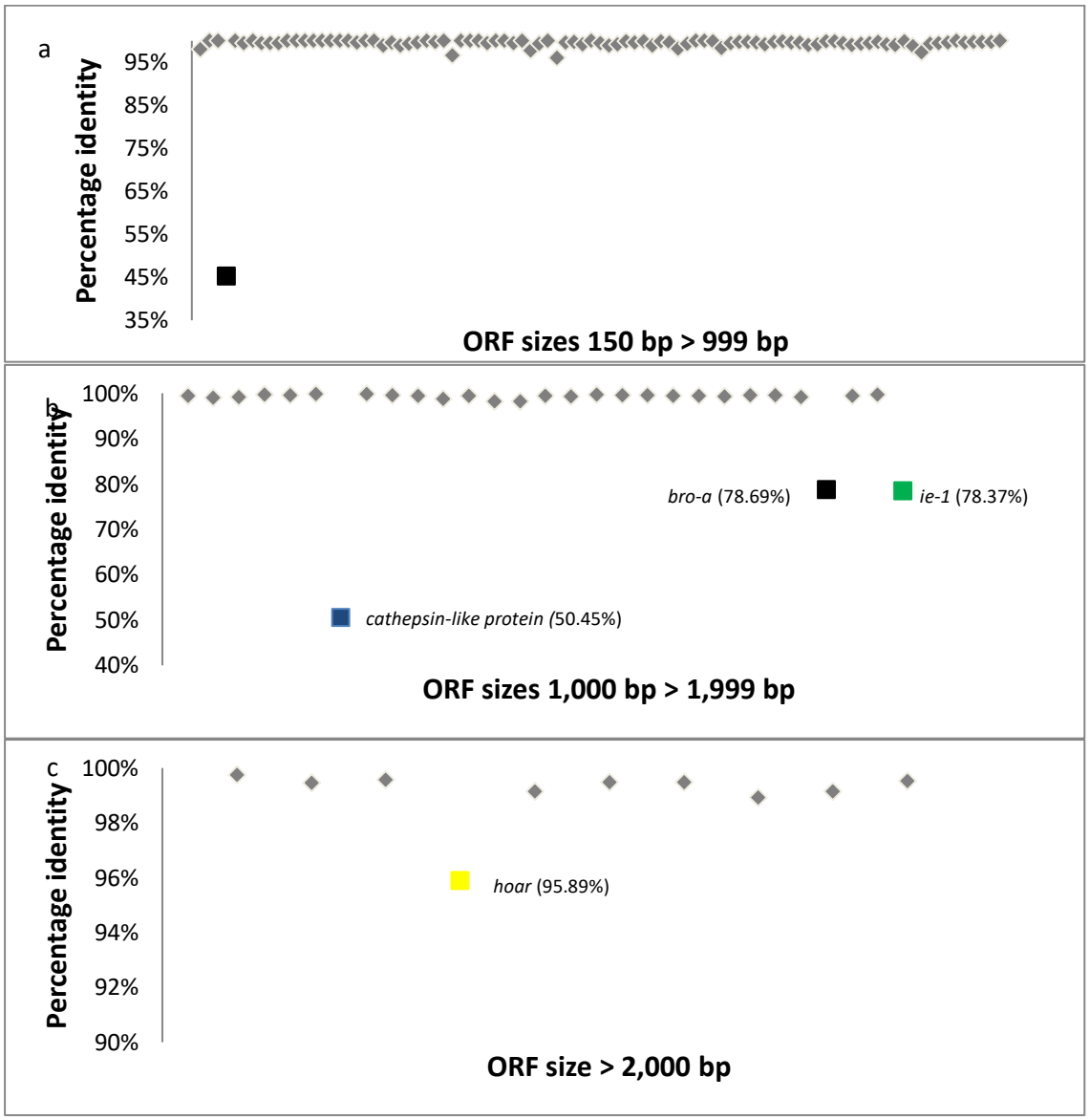


Figure 4.4: Percentage identity of the putative ORFs in the HearNPV-KZN genome compared to the HearNPV-Au genome. a: ORFs with sizes below 1,000 bp with variable ORFs with different data points. b: Variable ORFs observed between 1,000 to 1,999 bp. c: Variable ORFs with sizes above 2,000 bp. Data points highlighted in different colour represent function as described in Figure 4.2

4.3.4 Restriction endonuclease (REN) analyses for HearNPV-Albany, HearNPV-KZN and HearNPV-Au

4.3.4.1 REN analyses using EcoRI of HearNPV-Albany, HearNPV-KZN and HearNPV-Au *in silico*

In silico EcoRI profiles for HearNPV-Albany, HearNPV-KZN and HearNPV-AU genomes were generated (Table 4.1). The HearNPV-Albany genome had 31 fragments, HearNPV-KZN had 30 fragments and HearNPV-Au had 29 fragments. The fragments that had similar sizes in all three genomes are highlighted in red and there were 22 fragments recorded.

The genome of HearNPV-Albany and that of HearNPV-KZN shared four fragments that were absent in the HearNPV-Au genome. The bands had sizes of 13,166 and 13,110; 4,700 and 4,687; 3,995 and 4,175; 414 and 414 respectively.

The HearNPV-Au genome had one fragment with a size of 10,154 that was unique to the genome. The HearNPV-Albany genome had two fragments unique to the genome, with fragment sizes of 3,345 and 2,830 base pairs. The HearNPV-KZN genome had a fragment with a size of 9,195 which was the only unique fragment to the genome.

The HearNPV-Au and the HearNPV-KZN shared two fragment pairs that were absent in the HearNPV-Albany. The fragment sizes were 9,056 in HearNPV-Au and 9,050 in HearNPV-KZN and at 6,685 in HearNPV-Au and 6,593 in the HearNPV-KZN genome.

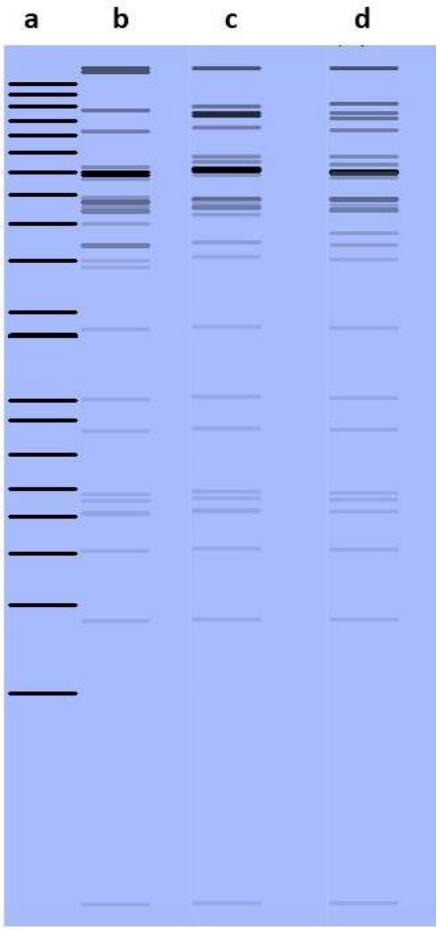
For the HearNPV-Albany and HearNPV-Au, two fragments with sizes of 4,919 and 13,575 for HearNPV-Albany and 4,833 and 13,441 for Australian HearNPV were identified to be similar.

There were double bands for HearNPV-Albany with sizes of 414 and 410, whereas the same double bands in HearNPV-KZN had sizes of 414 and 390. HearNPV-Au only had one of the bands, and this had a size of 410.

Table 4.1: EcoRI *in silico* digestion of HearNPV-Albany, HearNPV-KZN and HearNPV-Au. Red: fragments similar in all three genomes. Green: fragments similar in only HearNPV-Albany and HearNPV-KZN. Black: fragments unique to each genome. Purple: fragments similar in HearNPV-Au and HearNPV-KZN. Grey: fragments similar in HearNPV-Au and HearNPV-Albany

HearNPV-Albany 130,360 bp	HearNPV-KZN 130,636 bp	HearNPV-Au 130,992 bp	DNA Ladder
13,575		13,441	
13,166	13,110		12,000
		10,154	11,000
9,740	9,756	9,484	10,000
	9,195		9,000
	9,050	9,056	8,000
8,228	8,233	8,226	7,000
	6,593	6,685	6,000
6,230	6,294	6,285	5,000
5,987	5,997	5,992	4,000
5,958	5,946	5,944	3,000
5,843	5,844	5,843	2,000
5,840	5,837	5,838	1,650
5,689	5,687	5,688	1,000
4,919		4,833	850
4,753	4,733	4,751	
4,700	4,687		650
4,570	4,567	4,567	500
4,425	4,419	4,411	400
4,404	4,396	4,396	300
3,995	4,175		
		3,680	250
3,374	3,352	3,346	
3,345			
3,001	2,998	2,998	100
2,830			
1,743	1,743	1,742	
1,007	1,007	1,007	
784	784	784	
476	475	477	
453	453	453	
414	414		
410		410	
	390		
306	306	306	
176	176	176	
19	19	19	
130,360	130,636	130,992	

a: DNA ladder. b: HearNPV-Albany. c: HearNPV-KZN. d: HearNPV-Au



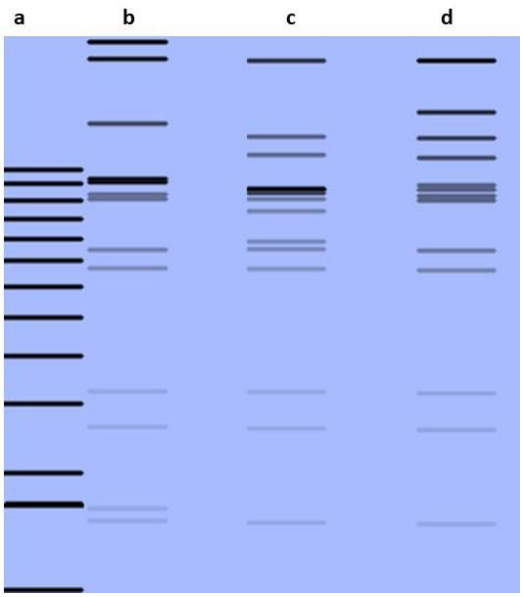
4.3.4.2 REN analyses using HindIII of HearNPV-Albany, HearNPV-KZN and HearNPV-Au genomes *in silico*

The *in silico* HindIII profiles for HearNPV-Albany, HearNPV-KZN and HearNPV-Au were generated (Table 4.2). In the three genomes, 11 fragment bands were recorded to be similar

for all three genomes. The HearNPV-Albany genome had 13 fragments, HearNPV-KZN had 14 fragments and HearNPV-Au had 13 fragments.

HearNPV-Albany genome had one fragment unique to the genome with a fragment size of 1,626 bp. HearNPV-Au genome had one unique fragment with a size of 16,865. HearNPV-KZN had two fragments with sizes of 9,354 bp and 7,840 bp that were only found in this genome. The HearNPV-KZN and HearNPV-Au genome shared one similar fragment at 12,965 bp in HearNPV-Au and 12,974 bp in HearNPV-KZN.

Table 4.2: HindIII *in silico* digestion of HearNPV-Au, HearNPV-KZN and HearNPV-Albany. Red: fragments similar in all three genomes. Black: fragments unique to each genome. Purple: fragments similar in HearNPV-Au and HearNPV-KZN.

HearNPV-Albany 130,360 bp	HearNPV-KZN 130,636 bp	HearNPV-Au 130,992 bp	DNA Ladder	a: DNA ladder. b: HearNPV-Albany. c: HearNPV-KZN. d: HearNPV-Au
25,495				
23,027	22,603	22,897		
		16,865		
15,721				
	14,493	14,582	12,000	
	12,974	12,965	11,000	
11,350			10,000	
11,111			9,000	
	10,720	10,982	8,000	
	10,632	10,717	7,000	
10,399	10,407	10,376	6,000	
10,062	10,050	10,050	5,000	
	9,354		4,000	
	7,840			
7,495	7,500	7,483	3,000	
6,696	6,684	6,690		
3,243	3,247	3,254	2,000	
2,626	2,623	2,622	1,650	
1,626				
1,509	1,509	1,509		
130,360	130,636	130,992	1,000	

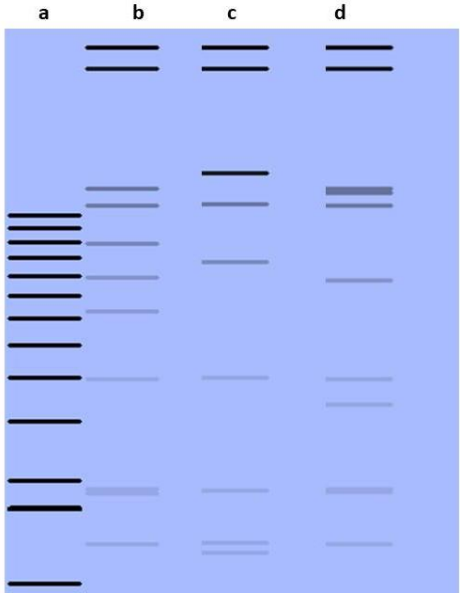
4.3.4.3 REN analyses using BamHI of HearNPV-Albany, HearNPV-KZN and HearNPV-Au genomes *in silico*

The *in silico* BamHI profiles for HearNPV-Albany genome, HearNPV-KZN genome and HearNPV-Au genome were generated (Table 4.3). The HearNPV-Albany gen and HearNPV Au had 11 digested fragments each and HearNPV-KZN had 10 digested fragments. The genomes shared six fragments that were similar in size.

The HearNPV-Albany genome and HearNPV-Au genome had three fragments, which were similar in size, of 14,392 bp, 7,719 bp and 1,892 bp for HearNPV-Au genome and 14,397 bp, 7,901 bp and 1,892 bp for HearNPV-Albany genome that were absent in HearNPV-KZN.

The HearNPV-Albany genome had two fragments unique to the genome, with fragment sizes of 9,934 bp and 6,287 bp. For the HearNPV-KZN genome, four fragments out of the 10 were unique with sizes of 15,847 bp, 15,801 bp, 8,725 bp and 1,223 bp. The HearNPV-Au genome had two fragments with sizes of 13,992 bp and 3,357 bp that were unique to the genome.

Table 4.3: BamHI *in silico* digestion of HearNPV-Albany, HearNPV-KZN and HearNPV-Au. Red: fragments similar in all three genomes. Black: fragments unique to each genome. Grey: fragments similar in HearNPV-Au and HearNPV-Albany

HearNPV-Albany 130,360 bp	HearNPV-KZN 130,636 bp	HearNPV-Au 130,992 bp	DNA Ladder	a: DNA ladder. b: HearNPV-Albany. c: HearNPV-KZN. d: HearNPV-Au
37,456	36,984	37,294		
32,523	32,083	32,385		
	15,847			
	15,801			
14,397		14,392		
		13,992		
12,839	12,825	12,831	12,000	
9,934			11,000	
	8,725		10,000	
7,901		7,719	9,000	
6,287			8,000	
3,985	3,985	3,985	7,000	
		3,357	6,000	
1,892		1,892	5,000	
1,843	1,860	1,842		
1,303	1,303	1,303	4,000	
	1,223			
			3,000	
			2,000	
130,360	130,636	130,992	1,000	

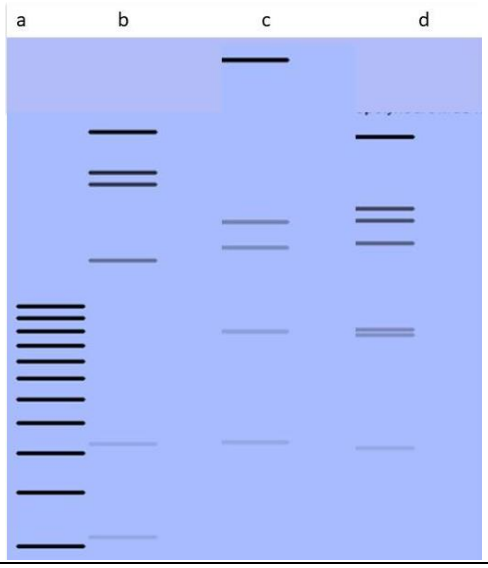
4.3.4.4 REN analyses using SacI of HearNPV-Albany, HearNPV-KZN and HearNPV-Au genomes *in silico*

The *in silico* SacI profiles for HearNPV-Albany, HearNPV-KZN and HearNPV-Au were generated (Table 4.4). Only one fragment size was similar for all three genomes with a size of 4,296 bp, 4,297 bp and 4,282 bp for HearNPV-Albany, HearNPV-KZN and HearNPV-Au respectively. The HearNPV-Albany genome had six fragments, HearNPV-KZN had five fragments and HearNPV-Au had seven fragments.

HearNPV-Albany genome had five fragments unique to the genome, with sizes of 44,394, 32,685, 29,963, 16,893 and 2,129 bp. The HearNPV-Au genome had three fragments with sizes of 41,191, 23,346 and 10,132 bp, that were unique to the genome. HearNPV-KZN genome had a fragment with a size of 75,551 bp that was unique to the genome.

Three fragment pairs were identified that were similar in the HearNPV-Au and HearNPV-KZN genomes with fragment sizes of 22,406 bp, 18,892 bp and 9,743 bp in HearNPV-Au and 22,407 bp, 18,538 bp and 9,843 bp in the HearNPV-KZN genome.

Table 4.4: SacI *in silico* digestion of HearNPV-Albany, HearNPV-KZN and HearNPV-Au. Red: fragments similar in all three genomes. Black: fragments unique to each genome. Purple: fragments similar in HearNPV-Au and HearNPV-KZN.

HearNPV-Albany 130,360 bp	HearNPV-KZN 130,636 bp	HearNPV-Au 130,992 bp	DNA Ladder	a: DNA ladder. b: HearNPV-Albany. c: HearNPV-KZN. d: HearNPV-Au
	75,551			
44,394				
		41,191		
32,685				
29,963				
		24,346		
	22,407	22,406		
	18,538	18,892		
16,893			12,000	
		10,132	11,000	
	9,843	9,743	10,000	
4,296	4,297	4,282	9,000	
2,129			8,000	
			7,000	
			6,000	
			5,000	
			4,000	
			3,000	
130,360	130,636	130,992	2,000	

4.3.4.5 REN analyses using PstI of HearNPV-Albany, HearNPV-KZN and HearNPV-Au genomes *in silico*

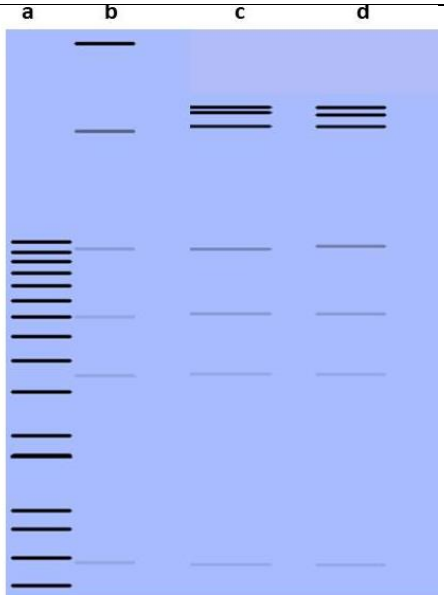
The *in silico* PstI profiles for HearNPV-Albany genome, HearNPV-KZN genome and HearNPV-Au genome were generated (Table 4.5). The HearNPV-Albany genome had six digested fragments and HearNPV-KZN and HearNPV-Au had seven digested fragments that were similar in size for both genomes. Of the six HearNPV-Albany fragments, five were similar in size to those generated for HearNPV-KZN and HearNPV-Au genomes.

For HearNPV-Albany, one band with a size of 75,221 bp was unique to the genome. The HearNPV-Au and HearNPV-KZN shared two bands that were absent in the HearNPV-Albany, with sizes of 39,365 bp and 37,351 bp in the HearNPV-KZN genome and 39,418 bp and 37,015 in HearNPV-Au.

Table 4.5: PstI *in silico* digestion of HearNPV-Au, HearNPV-KZN and HearNPV-Albany. Red: fragments similar in all three genomes. Purple: fragments similar in HearNPV-Au and HearNPV-KZN. Black: fragments unique to each genome

HearNPV-Albany 130,360 bp	HearNPV-KZN 130,636 bp	HearNPV-Au 130,992 bp	DNA Ladder
75,521			
	39,365	39,418	
	37,351	37,015	12,000
33,393	32,957	33,251	11,000
11,315	10,836	11,186	10,000
6,033	6,038	6,027	9,000
3,479	3,470	3,476	8,000
619	619	619	7,000
			6,000
			5,000
			4,000
			3,000
			2,000
			1,650
			1,000
			850
			650
130,360	130,636	130,992	500

a: DNA ladder. b: HearNPV-Au. c: HearNPV-KZN. d: HearNPV Albany



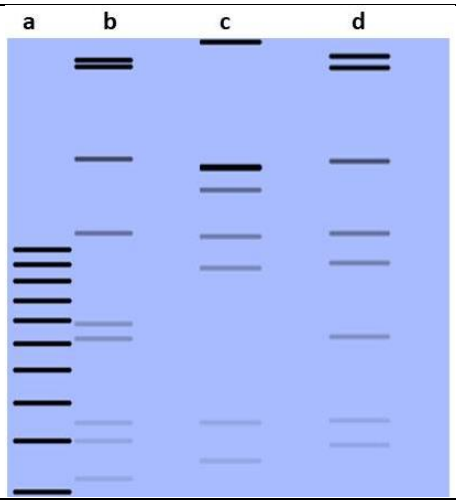
4.3.4.6 REN analyses using XhoI of HearNPV-Albany, HearNPV-KZN and HearNPV-Au genomes *in silico*

The *in silico* XhoI profiles for the HearNPV-Albany genome, HearNPV-KZN genome and HearNPV-Au genome were generated (Table 4.6). The three genomes shared three fragments with similar sizes and these had sizes of 20,058 bp, 19,999 bp and 20,273 bp; 13,210 bp, 13,211 bp and 13,209 bp; 4,459 bp, 4,466 bp and 4,460 bp for HearNPV-Albany, HearNPV-KZN and HearNPV-Au respectively. HearNPV-Albany had eight digested fragments and HearNPV-KZN and HearNPV-Au had eight bands each.

In all three genomes, unique bands were recorded, HearNPV-Albany having three bands at 35,673 bp, 7,847 bp and 3,226 bp, HearNPV-KZN having four bands at 41,188 bp, 19,666 bp, 17,458 bp and 3,587 bp and HearNPV-Au having two bands with sizes of 36,796 bp and 3,871 bp. In HearNPV-Au and HearNPV-KZN, one band was recorded to be similar with a size of 11,061 bp and the band was absent in HearNPV-Albany. The HearNPV-Albany and HearNPV-Au shared two bands that were absent in HearNPV-KZN at 34,447 bp and 34,313 bp; 7,335 bp and 7,208 bp.

Table 4.6: XhoI *in silico* digestion of HearNPV-Albany, HearNPV-KZN and HearNPV-Au. Red: fragments similar in all three genomes. Green: fragments similar in only HearNPV-Albany and HearNPV-KZN. Black: fragments unique to each genome. Purple: fragments similar in HearNPV-KZN and HearNPV-Au. Grey: fragments similar in HearNPV-Albany and HearNPV-Au

HearNPV-Albany 130,360 bp	HearNPV-KZN 130,636 bp	HearNPV-Au 130,992 bp	DNA Ladder	a: DNA ladder. b: HearNPV-Albany. c: HearNPV-KZN. d: HearNPV-Au
	41,188			
		36,796		
35,673				
34,447		34,314		
20,258	19,999	20,073		
	19,666			
	17,458			
13,210	13,211	13,209	12,000	
	11,061	11,061	11,000	
7,847			10,000	
7,235		7,208	9,000	
4,459	4,466	4,460	8,000	
4,005			7,000	
		3,871	6,000	
	3,587		5,000	
3,226			4,000	
130,360	130,636	130,992	3,000	



4.4 Discussion

In this chapter genomics were performed for HearNPV-Albany, HearNPV-KZN and HearNPV-Whit. DNA from the three samples was sequenced, mapped and assembled using HearNPV-Au and HearNPV-G4 as references. HearNPV-G4 was used as one of the assembly references as it was one of the first fully assembled HearNPV genome (Chen *et al.*, 2001) HearNPV-G4 has been used as a genome assembly reference (Zhang *et al.*, 2005; Ogembo *et al.*, 2009). HearNPV-Au was used as the assembly reference and the alignment reference because it was used as the HearNPV source in Helicovir™ (Moore and Kirkman, 2010), and it was therefore proposed that it may have been introduced to the local population. Assembled genomes of HearNPV-Albany, HearNPV-KZN and HearNPV-Whit were aligned against HearNPV-Au for ORF annotation and prediction. The use of HearNPV-Au for ORF annotation, prediction and analyses was performed because of the potential of HearNPV isolates from South Africa to be identical to HearNPV-Au which is formulated in Helicovir™. Therefore, these initial analyses helped in identifying if HearNPV-Albany and HearNPV-KZN were novel isolates and not isolates of the applied biopesticide. With the predicted putative ORF functions, comparative studies were performed on HearNPV-Albany and HearNPV-KZN against HearNPV-Au. HearNPV-Whit had assembly gaps that arose

from sequencing and no further analyses were performed on the genome. The sequencing of HearNPV-Whit was not successful, with gaps identified in the genome data resulting in corrupted sequencing and assembly. HearNPV-Haygrove was not sequenced as the isolate was only identified towards the end of the research. After complete assembly of the genomes annotation of the open reading frames (ORFs) and comparative studies of the HearNPV-Albany, HearNPV-KZN and HearNPV-Whit were conducted using HearNPV-Au as the reference.

After mapping and assembly against reference strains, HearNPV-Albany was found to have a size of 130,360 bp and HearNPV-KZN had a size of 130,636 bp. These sizes were like those of HearNPV sequenced genomes from geographically distinct locations in Australia, Asia, Europe and Africa with size that ranged from 130,759 bp to 132,481 bp (Chen *et al*., 2001; Zhang *et al*., 2005, Ogembo *et al*., 2007; Zhang *et al.,* 2014 Arrizubieta *et al*., 2015; Noune and Hauxwell, 2016; Eroglu *et al*., 2018).

Nucleotide variability was observed for both the HearNPV-Albany and HearNPV-KZN isolates during genome assembly and, for ambiguous nucleotides, the nucleotide with the higher base score was used for substitution. During assembly the identity of nucleotides at specific positions may not clearly be identifiable, which results in ambiguous bases (Dubey, 2014). The first adenine of the *polh* gene was initiated as position 1 of the genome (Chen *et al.,* 2001; Summers, 2006). Using this analysis, the physical gene maps of the complete sequenced HearNPV-Albany and HearNPV-KZN genomes were constructed. ORFs with potential to code for proteins of over 50 amino acids are considered as potential ORFs. This has been adopted as the conventional method for identifying potential baculovirus genes (Harrison and Hoover, 2012).

From the physical gene map, HearNPV-Albany comprised of 135 ORFs and HearNPV-KZN comprised of 132 ORFs, coding for proteins with 50 amino acids or more. The number of ORFs identified is proportional to the number of ORFs in relation to genome size that has been observed in other baculovirus genomes (Chen *et al.,* 2001). Characterisation of the ORFs was performed by comparing the HearNPV-Albany or HearNPV-KZN individual ORFs to the Australian HearNPV-Au ORFs. The analysis was separated into three distinct groups in order of size of the ORFs, from small, medium and large. This allowed for a more representative analysis of ORF variability throughout the genomes of HearNPV-Albany and HearNPV-KZN in comparison to the Australian HearNPV-Au ORFs. Based on this analysis,

it was observed that for the small ORFs, the average standard error was higher than for the medium and large ORFs even in the absence of outright outliers. In HearNPV-Albany, *ubiquitin* was found to be an outlier ORF in terms of nucleotide sequence identity. In baculoviruses *ubiquitin* has been reported to have a high degree of phylogenetic diversity (Rohrmann, 2013). For HearMNPV compared to HearNPV-G4, *ubiquitin* had an amino acid identity of 76.6% (Tang *et al*., 2012). *Ubiquitin* has been reported in some viruses to show degrees of divergence between genetically different samples (Ogembo *et al*., 2009; Tang *et al*., 2012). The *ubiquitin* is an auxiliary gene and has no direct involvement in the expression of viral genes, viral genome replication and viral progeny formation (Herniou *et al*., 2003). High sequence identity of *p47* and *38.7kDa protein* genes were observed in HearNPV-Albany being above 97% but below the standard error lower limit of 98.5%. Although it is highly conserved, low variability has been observed in the gene, with HearNPV-NNg1 having a nucleotide similarity of 79.9% with HearNPV-G4 (Ogembo *et al.,* 2009). The *38.7kDa protein* gene has an unknown function and no phenotypic effects were observed when it was deleted (Herniou *et al*., 2003; Rohrmann, 2013). The *38.7kDa protein* gene has been reported to show considerable variability in reported HearNPV genomes, with nucleotide similarity being between 94% and 96% when HearNPV-NNg1 was compared to HearNPV-C1 and HearNPV-G4 (Ogembo *et al.,* 2009). For the *hoar* ORF, variability was observed in both HearNPV-Albany and HearNPV-KZN, with similarities of 93.52% and 95.89% respectively when compared to HearNPV-Au *hoar*. The *hoar* gene is found in some alphabaculoviruses, including HearNPV, with its function being unknown (Rohrmann, 2013). It is variable with percentage identity for the ORF amongst HearNPV ranging between 90% and 94% (Ogembo *et al*., 2009). For HearNPV-KZN, variability was observed for *cathepsin-like protein*, *ie-1* and *bro-a* in comparison to the ORFs in HearNPV-Au. The c*athepsin-like protein* gene product is a proteinase that is found in most alphabaculoviruses that is involved in host cuticle disintegration and liquefaction and *ie-1* is involved in the transactivation of early infection promoters (Rohrmann, 2013; Haase *et al*., 2015). Both are highly conserved in currently sequenced HearNPV genomes (Ogembo *et al*., 2009; Zhang *et al*., 2014). *Bro* genes have low sequence similarities in HearNPV and show diverse variability probably due to recombination that occurs with the viral genomes (Bideshi *et al*., 2003; Noune and Hauxwell, 2016). Bro proteins have DNA binding activity and the genes are expressed as early genes (Rohrmann, 2013). The number of *bro* genes in each baculovirus genome varies from one virus to the next with up to 16 copies having been reported (Kuzio *et al*., 1999; Kang *et al*., 1999). In *bro* gene products the N-terminus that consists of the DNA binding domain is

conserved but the C-terminus is highly variable and intra-specific as well as interspecific polymorphisms are common in the *bro* genes for double stranded DNA insect viruses (Bideshi *et al.*, 2003).

Through *in silico* digests, using 6 different enzymes, similarities and differences of HearNPV-Albany, HearNPV-KZN and HearNPV-Au were recorded. Previous studies have used enzymes including EcoRI, HindIII, BamHI, SacI, PstI, XbaII, BglII and XhoI (Lua *et al.*, 2002; Chen *et al.*, 2000; Christian *et al*., 2001; Ogembo *et al*., 2007; Eroglu *et al*., 2018). Enzymes such as EcoRI have been used in the digestion of HearNPV genomes as they result in polymorphic fragments that allow a better comparison and distinction between genome isolates (Arrizubieta *et al.*, 2013). Genetic variation of similar genomes can be identified through the genomic maps generated after REN (Chen *et al*, 2000; Christian *et al*., 2001). Isolates of HearNPV from different geographical locations have been reported to generate different profiles therefore highlighting the genetic differences of the viruses on a genomic level (Ogembo *et al.*, 2007; Figueiredo *et al*., 2009; Noune and Hauxwell, 2016). The fragments generated after the *in-silico* digestion of HearNPV-Albany, HearNPV-KZN and HearNPV-Au showed that the three genomes had genetic differences that could be deduced from enzyme profiles. From REN analysis of genomes for the South African isolates it was confirmed that HearNPV-Albany and HearNPV-KZN are novel.

4.5 Conclusions

The aim of this chapter was to sequence, assemble and annotate three South African HearNPV isolates namely HearNPV-Albany, HearNPV-KZN and HearNPV-Whit. The sequencing and assembly of HearNPV-Whit resulted in low quality data and was discontinued. HearNPV-Albany and HearNPV-KZN were successfully sequenced, assembled and annotated. Through comparative analysis with HearNPV-Au it was confirmed that HearNPV-Albany and HearNPV-KZN are novel virus isolates from South African host populations.

In Chapter 5, a broader genetic analysis of the South African HearNPV genomes against full genomic sequences isolated from geographically distinct populations will be carried out to get an in depth understanding of genetic variations and identity and how they may relate to virulence.

5 Comparative *in silico* analyses of HearNPV-Albany, HearNPV-KZN and seven other geographically distinct HearNPV genomes

5.1 Introduction

In this chapter the genomes of HearNPV-Albany and HearNPV-KZN recovered from South African *H. armigera* host populations were compared to those of seven HearNPV isolates from different geographic regions namely, HearNPV-NNg1 from Kenya, HearNPV-Au, HearNPV-AC53 and HearNPV-H25EA1 from Australia, HearNPV-C1 and HearNPV-G4 from China and HearNPV-SP1A from Spain. The partial sequencing of HearNPV-Whit because of low quality DNA and the lack of a sequence for HearNPV-Haygrove resulted in the exclusion of the two isolates in the comparative analyses. The comparative studies of HearNPV-C1 and HearNPV-G4 isolated from the same province in China reported a high nucleotide similarity of 98.1% with some highly variable regions such as the *hr* and *bro* sequences in the genomes as well as significantly different insecticidal activity (Zhang *et al*., 2005). Comparative genome analysis can be used to infer evolutionary relationships with host populations and how it may affect virulence.

Genetic characterisation of HearNPV isolates aids in the understanding of the virus and in the establishment of a robust virus-based biological control system (Ardisson-Araújo *et al*., 2015). A high degree of genetic variation exists in baculovirus genomes from the same population and from geographically distant host populations (Harrison and Hoover, 2012). It is suggested that co-evolution of baculoviruses and their hosts has resulted in the genetic divergence of baculoviruses over time (Lua *et al*., 2002; Possee *et al*., 2010). As a result, genomes from the same host population may exhibit genetic variation (Erlandson, 2009). In depth comparative analysis of HearNPV genomes aids in understanding genetic variation and this may be linked to speed of kill and virulence (Baillie and Bouwer, 2013).

Genetic variation of HearNPV genomes can be identified from a general overview of ORF and genome sizes. Potentially, viruses isolated from different geographical locations or different host origins have different genomic sequences (Opoku-Debrah *et al*., 2013). Comparative genomic analyses of HearNPV genomes from distinct geographical locations have shown that the gene content and arrangements of HearNPV genomes is highly conserved (Ogembo *et al*., 2009). The presence and absence of some ORFs has been reported

to affect virulence in HearNPV genomes (Ogembo *et al.,* 2009; Raghavendra *et al*., 2017). Similarities of the genomes from distinct locations range between 94-100%, except for HearNPV-L1 from India, which has a low sequence similarity of below 85% compared to currently sequenced HearNPV genomes, because of genomic rearrangements (Noune and Hauxwell, 2016).

HearNPV genomes also contain unique sequences in their genomes, known as baculovirus repeat ORFs (*bro*) that comprise four families (*bro-a, bro-b, bro-c* and *bro-d*) and homologous repeat regions (*hrs*) of which there are five groups (*hr1, hr2, hr3, hr4* and *hr5*). These unique sequences are considered to represent viral diversity (Raghavendra *et al*., 2017). Divergence and differences of HearNPV strains is often observed in the *bro* and *hrs* regions which may be associated with virulence (Ogembo *et al*., 2009). Viral isolates have variants of *bro* genes, *hrs* and other repeated sequences which suggest that rearrangements are introduced through intramolecular and intermolecular homologous recombination (Harrison and Hoover, 2012). The Chinese isolates of HearNPV-C1 and HearNPV-G4 have high sequence identity with different virulence. The different virulence was attributed to the variation in the *hrs* or the *bro* region (Zhang *et al*., 2005). Virulence of strains isolated from Spain and Portugal were also different, even though the genomes shared a high sequence similarity within the putative genes and the principal difference in the genomes being observed for the *hrs* (Arrizubieta *et al.* 2015). The genome differences and identities can be used to infer evolutionary divergence of HearNPV strains.

Phylogenetic analysis of HearNPV genomes allows for the understanding of evolutionary relationships of HearNPV isolates from different populations. The construction of evolutionary trees from whole genomes compared to single genes provides an overall phylogenetic analysis (Herniou *et al.,* 2003). This is because whole genome changes compared to specific gene changes are less common making them more useful as phylogenetic markers (Goodman *et al*., 2007). From the construction of phylogenetic trees, it is commonly observed that HearNPV isolates from similar or close geographical locations are grouped together, showing recent common ancestral roots (Noune and Hauxwell, 2016).

Restriction endonuclease digestion of genomic DNA can be used to identify genetic variation between isolates (Erlandson, 2009). The use of REN for genome analysis is useful in identifying genotype differences through the fragmentation pattern generated after enzymatic digestions (Ogembo *et al.*, 2009; Harrison and Hoover, 2012). Similar sized band fragments

identify close relationships in comparative studies, with different band fragments potentially identifying genome additions or deletions (Cory *et al*., 2005). The generation of different fragment patters after REN is indicative of genetic variance between genomes (Figueiredo *et al.*, 2009).

Although significant conservation is observed in the genes of HearNPV, the specific gene sequences have been reported to have some variation between populations. Genomic studies have shown that specific genes play a role in virulence (Ogembo *et al*., 2007; Zhang *et al.*, 2005). Less conserved genes and repeat regions may be responsible for phenotypic factors that are specific to the virus for the host population representing co-evolution as well as targeted virulence (Arrizubieta *et al*., 2015). It is important to understand the role of genetic variation of HearNPV strains and the effect on biological activity for the improved control of *H. armigera* colonies globally (Rowley *et al.*, 2011). An analysis of the conserved regions aids in identifying evolutionary patterns. From the GenBank database, the *polh*, *lef-8* and *lef-9* genes are the mose sequenced genes for HearNPV. Single nucleotide polymorphisms (SNPs) are a useful aid in phylogenetic analysis. Conserved genes such as *polh*, *lef-8* and *lef-9* have been studied for SNPs (Khan *et al*., 2004; Arrizubieta *et al*., 2013). As a result of previous work on these genes, and the in-depth molecular work done on *polh* and *lef*-9, these two genes were chosen for SNP studies.

Several geographically distinct HearNPV strains have been sequenced and their full data have been published on GenBank. The published sequences include Iberian HearNPV strains isolated from Portugal and Spain namely HearNPV-SP and HearNPV-LB (Arrizubieta *et al*., 2015), Chinese HearNPV strains from similar geographical locations, HearNPV-G4 and HearNPV-C1 (Chen *et al*., 2001; Zhang *et al*., 2005), Australian HearNPV strains, HearNPV-Au, HearNPV-AC53 and HearNPV-H25EA1 (Zhang *et al*, 2014; Noune and Hauxwell, 2015) and a Kenyan HearNPV isolate HearNPV-NNg1 (Ogembo *et al*., 2007).

The aim of this chapter was to compare the genomes of two South African HearNPV isolates, namely HearNPV-Albany and HearNPV-KZN with those of geographically distinct isolates. The first objective was characterisation of the genomes including a general overview of ORFs and genome sizes to highlight differences in the genomes as well as genomic identities. The second objective was a comparison of the ORFs of HearNPV-Albany and HearNPV-KZN genomes against geographically isolated HearNPV genomes to identify ORFs present or absent in the genomes as well as their percentage identities. The third objective was the

identification of synonymous and non-synonymous SNPs in the *polh* and *lef-9* genes of HearNPV-Albany and HearNPV-KZN in comparison to geographically distinct HearNPV genomes. The fourth objective was the construction of phylogenetic trees to understand evolutionary paths of the different HearNPV strains. The final objective was to conduct *in silico* REN analysis of HearNPV-Albany and HearNPV-KZN against geographically distinct HearNPV genomes.

5.2 Methods and materials

5.2.1 Genome size and ORF comparison of HearNPV-Albany, HearNPV-KZN and seven geographically distinct HearNPV genomes

Geographically distinct full genomes of HearNPV were downloaded from the NCBI GenBank for the genomic analysis and comparative studies with HearNPV-Albany and HearNPV-KZN. The genomes used for comparative analysis (Table 5.1) and alignments were performed in Geneious R7.1.7. Data were exported to Microsoft Excel in the .csv format for further analysis. The size of each genome was confirmed in MEGA 7 and the number of ORFs for each genome were confirmed in Geneious R7.1.7 through the selection of coding sequences only option.

Table 5.1: HearNPV genomes that were used in the comparative studies against HearNPV-Albany and HearNPV-KZN and their accession numbers from NCBI's GenBank

Country	Isolate name	Accession number	Reference
Spain	HearNPV-SP1A	KJ701032	Arrizubieta *et al.*, (2015)
China	HearNPV-G4	AF271059	Chen *et al*., (2001)
China	HearNPV-C1	AF303045	Zhang *et al.*, (2005)
Australia	HearNPV-Au	JN584482	Zhang *et al.*, (2014)
Australia	HearNPV-AC53	KJ909666	Noune and Hauxwell (2015)
Australia	HearNPV-H25EA1	KJ922128	Noune and Hauxwell (2015)
Kenya	HearNPV-NNg1	AP010907	Ogembo *et al*., (2007)

5.2.2 Comparative ORF studies of HearNPV-Albany and HearNPV-KZN against seven geographically distinct HearNPV genomes

The ORFs for HearNPV-Albany and HearNPV-KZN were aligned against each of the seven HearNPV genomes (Table 5.1) in Geneious R7.1.7 at a 25% sequence similarity identity option and the nucleotide ORF percentage identity was calculated. The data with the ORF percentage identity and the lengths of the ORFs was exported to Ms Excel 2010 in the .csv format and comparative ORF studies were performed.

5.2.3 Sequence alignments of the *polh* and *lef-9* genes to identify single nucleotide polymorphisms

The complete genomes of HearNPV-Albany and HearNPV-KZN were aligned against the seven genomes in Geneious 7.1.7 using a complete nucleotide alignment. For the *polh* and *lef-9* coding sequences, nucleotide variability was studied, and the presence of single base substitutions was noted. The total number of SNPs observed for HearNPV-Albany and HearNPV-KZN were recorded and the nucleotide sequences were translated into amino acids to observe the effect of the SNPs on the protein sequence.

5.2.4 Phylogenetic analysis of HearNPV isolates

Complete genome phylogenetic analyses of the HearNPV-Albany, HearNPV-KZN and geographically distinct HearNPV genomes that are described in Table 5.1 were performed in Mega7. The evolutionary history was inferred using the maximum likelihood method based on the Tamura-Nei model and the neighbour joining method (Saitou and Nei, 1987; Tamura and Nei, 1993). These two methods were used for the analyses because they have been reported in HearNPV genome evolutionary tree analyses (Arrizubieta *et al.*, 2013; Ardisson-Araújo *et al.*, 2015 Noune and Hauxwell, 2016). The bootstrap consensus of the tree was deduced from 1,000 replicates which represent the evolutionary history of the taxa analysed. A discrete Gamma distribution was used to model the evolutionary rate differences among sites. For the analysis all positions that contained gaps and missing data were eliminated. The full genome of HearGV (accession number NC_010240.1) (Harrison and Popham, 2008) was used as an outgroup. Initial trees for the heuristic search were obtained automatically by applying the neighbour joining method and BioNJ (Bio neighbour joining) algorithms to a matrix of pairwise distances, estimated using the maximum composite likelihood (MCL) approach and then taking the topology with the superior log likelihood. The results from the two analyses were compared and the best model was used for the results.

5.2.5 *In silico* restriction endonuclease analysis

The complete genomes of HearNPV-Albany, HearNPV-KZN, HearNPV-Au, HearNPV-AC53, HearNPV-H25EA1, HearNPV-C1, HearNPV-G4, HearNPV-NNg1 and HearNPV-SP1 were digested *in silico* using A plasmid Editor (ApE) version 2.0.47. The enzymes selected were EcoRI BamHI and HindIII as reported by Arrizubieta *et al.*, (2013) and Zhang and Wu, (2001). For enzyme digestion, the DNA was selected in the circular conformation and virtual gels as well as band fragment size generated were saved and tabulated for each genome.

5.3 Results

5.3.1 Genome sequence analysis of HearNPV-Albany and HearNPV-KZN from South Africa against HearNPV genomes from different geographical locations

Genome analysis and comparison was performed for HearNPV-Albany, HearNPV-KZN and geographically distinct HearNPV genomes (Table 5.1). The genomes were similar in size with HearNPV-Albany having the smallest genome with a size of 130,360 bp and HearNPV-SP1A having the largest genome with a size of 132,481 bp (Table 5.2). The mean genome size for the nine genomes was 131,104 (±830) bp. HearNPV-KZN genome had the lowest ORFs with 132 and HearNPV-NNg1 had the most ORFs with 143 ORFs.

Table 5.2: A comparison of HearNPV genomes from geographically distinct regions showing the genome size and number of ORFs for each genome

Country	Isolate name	Genome size	ORFs
Spain	HearNPV-SP1A	132,481	136
China	HearNPV-G4	131,405	135
China	HearNPV-C1	130,759	137
Australia	HearNPV-Au	130,992	133
Australia	HearNPV-AC53	130,442	139
Australia	HearNPV-H25EA1	130,440	139
Kenya	HearNPV-NNg1	132,425	143
South Africa	HearNPV-Albany	130,360	135
South Africa	HearNPV-KZN	130,636	132

5.3.2 Nucleotide alignment of geographically distinct HearNPV genomes against South African HearNPV-Albany and HearNPV-KZN genomes

An alignment of HearNPV-Albany and HearNPV-KZN against seven HearNPV genomes resulted in relatively high nucleotide similarities of the genomes of between 93.9% and 95.3% (Table 5.3). The full genome alignments showed a mean nucleotide identity of 95.0% (±0.3%) for HearNPV-Albany genome and 94.6% (±0.4%) for HearNPV-KZN genome in comparison to the geographically distinct HearNPV genomes. The HearNPV-Albany genome had the least similarity in nucleotides when compared to HearNPV-C1 at 94.6% identity. The highest identical nucleotides after alignment were recorded for HearNPV-Albany genome and HearNPV-NNg1 with an identity of 95.3%. HearNPV-KZN shared the least nucleotide similarity with HearNPV-NNg1 with an identity of 93.9%. For HearNPV-KZN the highest nucleotide similarity was recorded when aligned against HearNPV-G4 for a 94.9% nucleotide similarity (Table 5.3).

Table 5.3: Comparison of full genomic identity of HearNPV full genomes to the South African HearNPV genomes after nucleotide alignment analyses

Genome Identity	HearNPV-Albany		HearNPV-KZN	
	Identical sites	Percentage identity	Identical sites	Percentage identity
HearNPV-G4	125,574	94.7%	125,513	94.9%
HearNPV NNg1	127,147	95.3%	124,900	93.9%
HearNPV-Au	125,162	95.2%	125,075	95.0%
HearNPV-SP1A	126,896	95.1%	125,139	94.1%
HearNPV-AC53	125,886	95.2%	124,916	94.8%
HearNPV-H25EA1	125,850	95.1%	124,911	94.8%
HearNPV-C1	125,162	94.6%	125,075	94.7%

5.3.3 Comparison of HearNPV-Albany and HearNPV-KZN ORFs against geographically distinct HearNPV genomes

A comparison of the HearNPV-Albany and HearNPV-KZN ORFs against geographically distinct HearNPV genomes (Table 5.1) after gene alignment in Geneious 7.1.7 to identify percentage similarity of the ORFs with a predicted putative function against each other was performed. The mean percentage similarity of the ORFs was calculated for each comparison as well as the standard error (Table 5.4)

For HearNPV-Albany, differences in ORF identity compared to HearNPV genomes were mainly recorded for hypothetical proteins. For ORFs with a predicted putative function in HearNPV-Albany, *ubiquitin, pp31, hoar*, *lef-11* and *lese-25 like protein* were recorded to be significantly different in nucleotide percentage identity compared to HearNPV genomes (Table 5.4).

For HearNPV-KZN ORFs when compared to the HearNPV genomes *bro-a, cathepsin, ie-1, hoar* and *bro-b* were recorded to be different for the nucleotide percentage identity (Table 5.4). The HearNPV-KZN genome also lacked the *gp37* gene. For both HearNPV-Albany and HearNPV-KZN it was recorded that *bro-c* gene is absent.

For the repeat regions, HearNPV-Albany had one *bro* ORF for *bro-b* and four *hr* repeats for *hr1, hr2, hr4* and *hr5* and HearNPV-KZN had three *bro* ORFs for a *bro-a* and two *bro-b* as well as five *hr* repeats for *hr1, hr2, hr3, hr4* and *hr5* (Table 5.4). The nucleotide sequences of the *hr* repeats were significantly different between the genomes compared to HearNPV-Albany and HearNPV-KZN. However, in HearNPV-KZN, *hr3* had a nucleotide similarity of

91.93% with the other HearNPV genomes used in the analysis (Table 5.1). In HearNPV-Albany, hr3 was absent.

Table 5.4: Comparative analysis of the nucleotide sequences of ORFs with a known putative function for HearNPV-Albany and HearNPV-KZN against HearNPV genomes. Identical results were recorded for HearNPV-AC53 and HearNPV-H25EA1 and only one of the genomes is reported. The similarity in only HearNPV-AC53 is marked with [a] and the similarity in only HearNPV-C1 is marked with [b].

HearNPV isolate	ORF identity	Nucleotide percentage similarity in HearNPV-Albany	Nucleotide percentage similarity in HearNPV-KZN
All in table 5.1	*ubiquitin*	66.01%	100.00%
All in table 5.1	*pp31*	81.99%	99.79%
All in table 5.1 (HearNPV-AC53)[a] (HearNPV-C1)[b]	*hoar*	93.52% (93.61%)[a]	95.89% (94.53%)[b]
All in table 5.1 (HearNPV-AC53)[a]	*lef-11*	90.10% (88.34%)[a]	100.00%
All in table 5.1	*lese-25 like protein*	90.10%	98.95%
All in table 5.1	*ie-1*	99.49%	78.37%
All in table 5.1	*cathepsin-like cysteine protease*	98.91%	50.45%
All in table 5.1	*gp37*	98.81%	ORF absent
All in table 5.1	*ie-1*	99.49%	78.37%
All in table 5.1	*bro-a*	ORF absent	78.69%
HearNPV-AC53	*bro-b*	ORF absent	53.78%
All in table 5.1	*bro-b*	99.40%	99.40%
HearNPV-SP1A HearNPV-NNg1	*bro-c*	ORF absent	ORF absent
All in table 5.1	*hr1*	46.16%	71.40%
All in table 5.1	*hr2*	67.98%	45.38%
All in table 5.1	*hr3*	Absent	91.93%
All in table 5.1	*hr4*	72.87%	72.41%
All in table 5.1	*hr5*	60.57%	67.10%

The genomes for Hear-NPV-Albany and HearNPV-KZN were aligned against geographically distinct HearNPV genomes. An ORF comparison for the presence or absence of ORFs predicted to be for hypothetical proteins was performed. Most of the *hypothetical protein* ORFs were shared between at least two HearNPV genomes; ORF 121 in HearNPV-Albany and ORF-118 in HearNPV-KZN were unique to the respective genomes (Table 5.5). HearNPV-Albany had a unique ORF at position 42 with a length of 192 bp that was absent in all the other HearNPV genomes used in the study. For HearNPV-KZN, an ORF with a size of 588 bp that was found at positions 55-58 in the other HearNPV genomes was absent.

Table 5.5: Comparative analysis of HearNPV ORFs of hypothetical proteins for HearNPV-Albany and HearNPV-KZN against geographically distinct HearNPV genomes. The table reports the position and sizes of the hypothetical protein ORFs if they are present in a HearNPV genome. Identical results were recorded for HearNPV-AC53 and HearNPV-H25EA1 and only one of the genomes is reported

HearNPV-Albany	HearNPV-KZN	HearNPV genomes corresponding ORF and size	HearNPV-absent
ORF-5 162 bp	ORF-5 180 bp	HearNPV-C1 ORF-5 192 bp HearNPV-(H25EA1, AC53, G4 and Au) ORF-5 180 bp HearNPV-NNg1 ORF-5 162bp	HearNPV-SP1A
ORF-26 393 bp	Absent	HearNPV-(AC53, NNg1) ORF26 153 bp	HearNPV-(C1, G4, SP1A and Au)
ORF-42 192 bp	Absent		HearNPV-(AC53, NNg1, Au, SP1A, G4, C1)
ORF-58 588 bp	Absent	HearNPV-(C1, G4, Au) ORF-56 588 bp HearNPV-AC53 ORF-57 606 bp HearNPV-NNg1 ORF-57 588 bp HearNPV-SP1A ORF-55 588 bp	
Absent	Absent	HearNPV-AC53 ORF-61 180 bp HearNPV-NNg1 ORF-62 180 bp	HearNPV-(Au, C1, G4, SP1A)
Absent	Absent	HearNPV-C1 ORF-97 168 bp HearNPV-NNg1 ORF-100 168 bp	HearNPV-(G4, Au, AC53, SP1A)
Absent	Absent	HearNPV-C1 ORF-119 408 bp HearNPV-NNg1 ORF 124 360 bp HearNPV-SP1A ORF-118 396 bp	HearNPV-(G4, AC53, Au)
ORF-121 157 bp	ORF-118 157 bp		HearNPV-(AC53, NNg1, Au, SP1A, G4, C1)

5.3.4 Phylogenetic studies of HearNPV genomes

Phylogenetic analyses were performed on the HearNPV-KZN, HearNPV-Albany, HearNPV-Au, HearNPV-C1, HearNPV-NNg1, HearNPV-AC53, HearNPV-H25EA1 and HearNPV-SP1A. The neighbour joining tree method and the maximum likelihood estimation resulted in similar trees.

Positions containing gaps and missing data were excluded from the analysis. A total of 125,922 positions were used in the final dataset. The analysis produced identical trees and the tree from the maximum likelihood method was reported for the analysis (Figure 5.1). The

analysis revealed that there were three distinct clusters. Hear-NPV-Albany was clustered with the other African isolate, HearNPV-NNg1 and the European isolate HearNPV-SP1A. HearNPV-KZN was clustered with the Chinese HearNPV isolates and HearNPV-Au. The two other Australian isolates, HearNPV-AC53 and HearNPV-H25EA1 were clustered by themselves. The phylogenetic analysis suggested that HearNPV-Albany was more closely related to the Kenyan and Spanish strain that to HearNPV-KZN. The pairwise nucleotide distance for the HearNPV isolates ranged from 0.001 to 0.019.

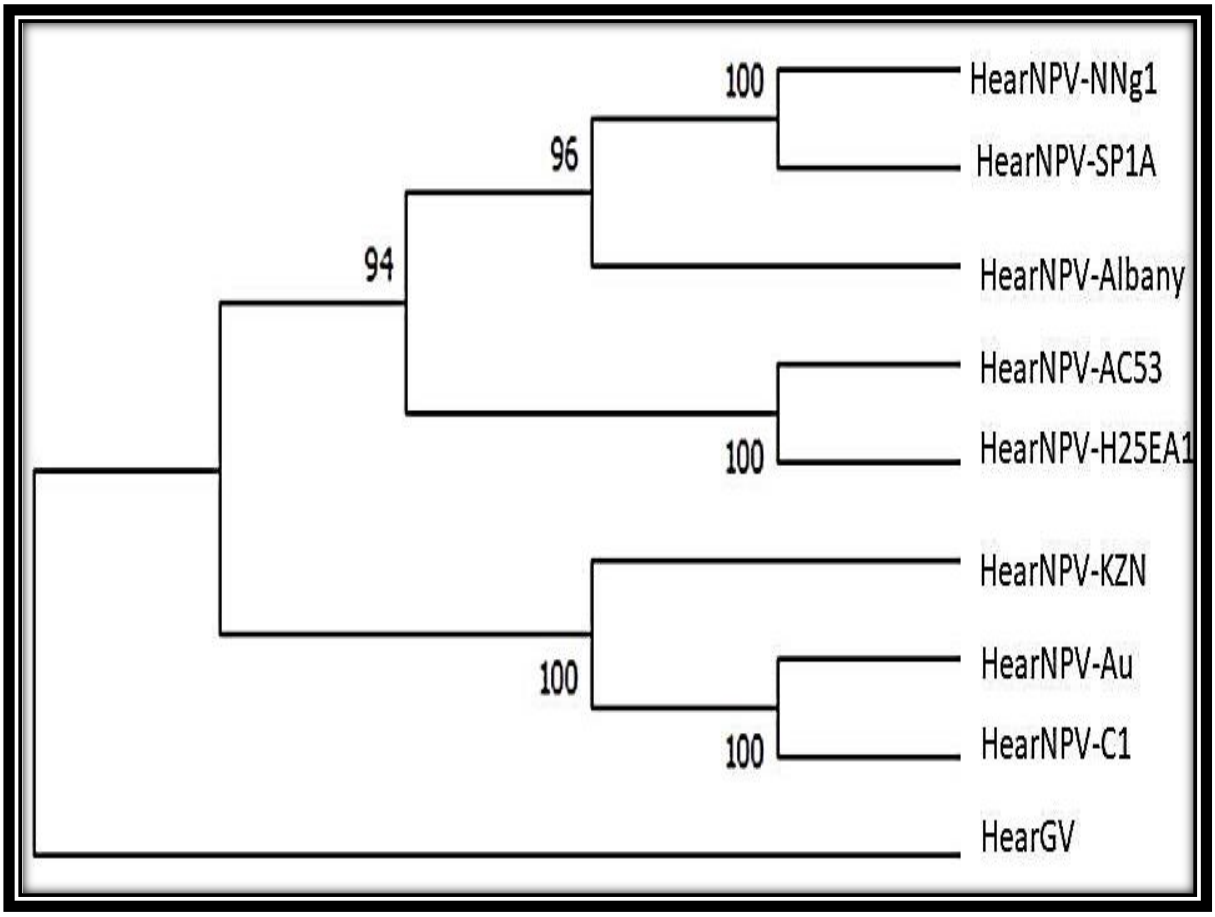


Figure 5.1: Phylogenetic tree generated through the maximum likelihood method of HearNPV-Albany, HearNPV-KZN, HearNPV-Au, HearNPV-SP1A, HearNPV-C1, HearNPV-G4 HearNPV-NNg1, HearNPV-AC53, HearNPV-H25EA1 and HearGV.

5.3.5 Analysis for single nucleotide polymorphisms in *polh* and *lef-9*

For HearNPV genomes, the *polh* and *lef-9* are among the most sequenced and published genes on the NCBI's GenBank database with over 15 gene sequences available for each. The *polh* and *lef-9* genes were chosen for further analysis from the genomes using the full genome sequences that have been studied in this chapter. Alignment of the full genomes using ClustalW alignment was performed in Geneious 7. The comparison of the aligned genes consisted of full gene sequences.

5.3.5.1 Comparative analysis of the *polh* gene for SNPs

For the *polh* gene eight SNPs in total were recorded for the comparison and all were synonymous (Table 5.6). HearNPV-Albany had at least two SNPs when compared to each individual HearNPV genome. The least SNPs were recorded for HearNPV-AC53 and HearNPV-NNg1 in comparison to HearNPV-Albany with only two synonymous SNPs.

HearNPV-KZN had one SNP when compared to HearNPV-G4 and two SNPs when compared to HearNPV-Au and HearNPV-AC53.

5.3.5.2 Comparative analysis of the *lef-9* gene for SNPs

For the *lef-9* gene eight SNPs were recorded across the HearNPV genomes with six being synonymous and two being non-synonymous (Table 5.7). The non-synonymous SNPs resulted in nucleotide code changes from GAA to AAA which resulted in an amino acid change of Glutamic Acid to Lysine and TTG to ATG which resulted in an amino acid change of Leucine to Methionine.

For HearNPV-Albany, no SNPs were recorded compared to HearNPV-SP1A and synonymous SNPs were recorded in comparison to HearNPV-NNg1, HearNPV-Au and HearNPV-AC53.

No SNPs were recorded between HearNPV-KZN and HearNPV-C1. When HearNPV-KZN was compared to all the other HearNPV genomes, both synonymous and non-synonymous SNPs were recorded, except for HearNPV-G4. For HearNPV-KZN, one non-synonymous SNP was recorded compared to HearNPV-G4, which changed the codon ATG:TTG, resulting in an amino acid change of Methionine to Leucine.

Table 5.6: Single nucleotide polymorphisms (SNPs) found in the *polh* gene for geographically distinct genomes compared to HearNPV-Albany and HearNPV-KZN. Red; identical codons between the genomes. Green; synonymous SNPs.

HearNPV genome	nt	nt	**HearNPV-KZN** Codons:	Amino acids	nt	**HearNPV-Albany** Codons	Amino acids
HearNPV -G4	192	192	CCT:CCT	Proline	192	CCC:CCT	Proline
	225	225	GAG:GAG	Glutamic acid	225	GAG:GAG	Glutamic acid
	279	279	CGC:CGC	Threonine	279	CGC:CGC	Threonine
	303	303	ACG:ACG	Threonine	303	ACT:ACG	Threonine
	549	549	AAT:AAC	Asparigine	549	AAC:AAC	Asparigine
	582	582	GAT:GAT	Aspartic Acid	582	GAC:GAT	Aspartic acid
	612	612	TTC:TTT	Phenylalanine	612	TTC:TTT	Phenylalanine
	672	672	GAG:GAA	Glutamic acid	672	GAG:GAA	Glutamic acid
HearNPV -C1	192	192	CCT:CCT	Proline	192	CCC:CCT	Proline
	225	225	GAG:GAA	Glutamic acid	225	GAG:GAG	Glutamic acid
	279	279	CGC:CGC	Threonine	279	CGC:CGC	Threonine
	303	303	ACG:ACG	Threonine	303	ACT:ACG	Threonine
	549	549	AAT:AAC	Asparagine	549	AAC:AAC	Asparagine
	582	582	GAT:GAT	Aspartic acid	582	GAC:GAT	Aspartic acid
	612	612	TTT:TTC	Phenylalanine	612	TTT:TTC	Phenylalanine
	672	672	GAG:GAA	Glutamic acid	672	GAG:GAA	Glutamic acid
HearNPV -NNg1	192	192	CCT:CCC	Proline	192	CCC:CCC	Proline
	225	225	GAG:GAA	Glutamic acid	225	GAG:GAA	Glutamic acid
	279	279	CGC:CGT	Threonine	279	CGC:CGT	Threonine
	303	303	ACG:ACT	Threonine	303	ACT:ACT	Threonine
	549	549	AAT:AAC	Asparagine	549	AAC:AAC	Asparagine
	582	582	GAT:GAC	Aspartic acid	582	GAC:GAC	Aspartic acid
	612	612	TTT:TTT	Phenylalanine	612	TTT:TTT	Phenylalanine
	672	672	GAG:GAG	Glutamic acid	672	GAG:GAG	Glutamic acid
HearNPV -Au	192	192	CCT:CCT	Proline	192	CCC:CCT	Proline
	225	225	GAG:GAG	Glutamic acid	225	GAG:GAG	Glutamic acid
	279	279	CGC:CGC	Threonine	279	CGC:CGC	Threonine
	303	303	ACG:ACG	Threonine	303	ACT:ACG	Threonine
	549	549	AAT:AAC	Asparagine	549	AAC:AAC	Asparagine
	582	582	GAT:GAT	Aspartic acid	582	GAC:GAT	Aspartic acid
	612	612	TTT:TTC	Phenylalanine	612	TTT:TTC	Phenylalanine
	672	672	GAG:GAA	Glutamic acid	672	GAG:GAA	Glutamic acid
HearNPV -SP1A	192	192	CCT:CCT	Proline	192	CCC:CCT	Proline
	225	225	GAG:GAA	Glutamic acid	225	GAG:GAA	Glutamic acid
	279	279	CGC:CGT	Threonine	279	CGC:CGT	Threonine
	303	303	ACG:ACT	Threonine	303	ACT:ACT	Threonine
	549	549	AAT:AAC	Asparagine	549	AAC:AAC	Asparagine
	582	582	GAT:GAC	Aspartic acid	582	GAC:GAC	Aspartic acid
	612	612	TTT:TTT	Phenylalanine	612	TTT:TTT	Phenylalanine
	672	672	GAG:GAG	Glutamic acid	672	GAG:GAG	Glutamic acid
HearNPV -AC53	192	192	CCT:CCT	Proline	192	CCC:CCT	Proline
	225	225	GAG:GAG	Glutamic acid	225	GAG:GAG	Glutamic acid
	279	279	CGC:CGC	Threonine	279	CGC:CGC	Threonine
	303	303	ACG:ACT	Threonine	303	ACT:ACT	Threonine
	549	549	AAT:AAT	Asparagine	549	AAC:AAT	Asparagine
	582	582	GAT:GAC	Aspartic acid	582	GAC:GAC	Aspartic acid
	612	612	TTT:TTT	Phenylalanine	612	TTT:TTT	Phenylalanine
	672	672	GAG:GAG	Glutamic acid	672	GAG:GAG	Glutamic acid
HearNPV -KZN	192	192	-	-	192	CCC:CCT	Proline
	225	225	-	-	225	GAG:GAG	Glutamic acid
	279	279	-	-	279	CGC:CGC	Threonine
	303	303	-	-	303	ACT:ACG	Threonine
	549	549	-	-	549	AAC:AAT	Asparagine
	582	582	-	-	582	GAC:GAT	Aspartic acid
	612	612	-	-	612	TTT:TTT	Phenylalanine
	672	672	-	-	672	GAG:GAG	Glutamic acid

Table 5.7: The alignment of the *lef-9* gene and SNPs observed in HearNPV-Albany and HearNPV-KZN compared to geographically distinct HearNPV genomes. Red; identical codons between the genomes. Green; synonymous SNPs. Blue; non-synonymous SNPs.

HearNPV genome	nt	nt	**HearNPV-KZN** Codons:	Amino acids	nt	**HearNPV-Albany** Codons	Amino acids
HearNPV -G4	44,515	44,471	AAA:AAA	Lysine	44,765	GAA:AAA	Glutamic acid: Lysine
	44,582	44,539	GTG:GTG	Valine	44,833	GTG:GTG	Valine
	44,604	44,560	GTA:GTA	Valine	44,854	GTT:GTA	Valine
	44,689	44,645	ATG:TTG	Methionine: Leucine	44,939	TTG:TTG	Leucine
	44,841	44,797	GTA:GTA	Valine	45,091	GTT:GTA	Valine
	45,366	45,322	GAC:GAC	Aspartic acid	45,617	GAT:GAC	Aspartic acid
	45,639	45,595	CTC:CTC	Leucine	45,889	CTA:CTC	Leucine
	45,786	45742	GCT:GCT	Alanine	46,037	GCC:GCT	Alanine
HearNPV -C1	44,426	44,471	AAA:AAA	Lysine	44,765	GAA:AAA	Glutamic acid: Lysine
	44,494	44,539	GTG:GTG	Valine	44,833	GTG:GTG	Valine
	44,515	44,560	GTA:GTA	Valine	44,854	GTT:GTA	Valine
	44,600	44,645	ATG:ATG	Methionine	44,939	TTG:ATG	Leucine: Methionine
	44,752	44,797	GTA:GTA	Valine	45,091	GTT:GTA	Valine
	45,278	45,322	GAC:GAC	Aspartic acid	45,617	GAT:GAC	Aspartic acid
	45,550	45,595	CTC:CTC	Leucine	45,889	CTA:CTC	Leucine
	45,698	45742	GCT:GCT	Alanine	46,037	GCC:GCT	Alanine
HearNPV -NNg1	44,503	44,471	AAA:GAA	Lysine: Glutamic acid	44,765	GAA:GAA	Glutamic acid
	44,571	44,539	GTG:GTA	Valine	44,833	GTG:GTA	Valine
	44,592	44,560	GTA:GTA	Valine	44,854	GTT:GTA	Valine
	44,677	44,645	ATG:TTG	Methionine: Leucine	44,939	TTG:TTG	Leucine
	44,829	44,797	GTA:GTT	Valine	45,091	GTT:GTT	Valine
	45,355	45,322	GAC:GAT	Aspartic acid	45,617	GAT:GAT	Aspartic acid
	45,627	45,595	CTC:CTA	Leucine	45,889	CTA:CTA	Leucine
	45,775	45742	GCT:GCC	Alanine	46,037	GCC:GCC	Alanine
HearNPV -Au	44,551	44,471	AAA:GAA	Lysine: Glutamic acid	44,765	GAA:GAA	Glutamic acid
	44,619	44,539	GTG:GTA	Valine	44,833	GTG:GTA	Valine
	44,640	44,560	GTA:GTA	Valine	44,854	GTT:GTA	Valine
	44,725	44,645	ATG:TTG	Methionine: Leucine	44,939	TTG:TTG	Leucine
	44,877	44,797	GTA:GTA	Valine	45,091	GTT:GTA	Valine
	45,403	45,322	GAC:GAC	Aspartic acid	45,617	GAT:GAC	Aspartic acid
	45,675	45,595	CTC:CTC	Leucine	45,889	CTA:CTC	Leucine
	45,823	45742	GCT:GCT	Alanine	46,037	GCC:GCT	Alanine
HearNPV -SP1A	44,353	44,471	AAA:GAA	Lysine: Glutamic acid	44,765	GAA:GAA	Glutamic acid
	44,421	44,539	GTG:GTG	Valine	44,833	GTG:GTG	Valine
	44,617	44,560	GTA:GTT	Valine	44,854	GTT:GTT	Valine
	44,702	44,645	ATG:TTG	Methionine: Leucine	44,939	TTG:TTG	Leucine
	44,854	44,797	GTA:GTT	Valine	45,091	GTT:GTT	Valine
	45,380	45,322	GAC:GAT	Aspartic acid	45,617	GAT:GAT	Aspartic acid
	45,652	45,595	CTC:CTA	Leucine	45,889	CTA:CTA	Leucine
	45,800	45742	GCT:GCC	Alanine	46,037	GCC:GCC	Alanine
HearNPV -AC53	44,249	44,471	AAA:GAA	Lysine: Glutamic acid	44,765	GAA:GAA	Glutamic acid
	44,317	44,539	GTG:GTA	Valine	44,833	GTG:GTA	Valine
	44,338	44,560	GTA:GTA	Valine	44,854	GTT:GTA	Valine
	44,423	44,645	ATG:TTG	Methionine: Leucine	44,939	TTG:TTG	Leucine
	44,575	44,797	GTA:GTA	Valine	45,091	GTT:GTA	Valine
	45,101	45,322	GAC:GAT	Aspartic acid	45,617	GAT:GAT	Aspartic acid
	45,373	45,595	CTC:CTA	Leucine	45,889	CTA:CTA	Leucine
	45,521	45742	GCT:GCC	Alanine	46,037	GCC:GCC	Alanine
HearNPV -Albany	44,471	44,471	AAA:GAA	Lysine: Glutamic acid	44,765	-	-
	44,833	44,539	GTG:GTG	Valine	44,833		
	44,854	44,560	GTA:GTT	Valine	44,854	-	-
	44,939	44,645	ATG:TTG	Methionine: Leucine	44,939	-	-
	45,091	44,797	GTA:GTT	Valine	45,091	-	-
	45,617	45,322	GAC:GAT	Aspartic acid	45,617	-	-
	45,889	45,595	CTC:CTA	Leucine	45,889	-	-
	46,037	45742	GCT:GCC	Alanine	46,037	-	-

5.3.6 Restriction endonuclease profiles of HearNPV-Albany and HearNPV-KZN against geographically distinct HearNPV genomes

Restriction endonuclease profiles were generated for the HearNPV-KZN and HearNPV-Albany genomes and were compared to the profiles of other geographically distinct HearNPV genomes. For the REN analyses, three enzymes were used, EcoRI, BamHI and HindIII (Lua *et al.*, 2002; Chen *et al*, 2000; Christian *et al*., 2001 and Ogembo *et al*., 2007). The digests were performed *in silico* using ApE v2.0.

5.3.6.1 *In silico* digestion of HearNPV genomes using EcoRI

After digestion with the enzyme EcoR1, the fragment profiles generated resulted in 29 band fragments for HearNPV-Au and HearNPV-AC53; 30 band fragments for HearNPV-C1, HearNPV-G4, HearNPV-KZN, HearNPV-NNg1 and HearNPV-SP1A; and 31 band fragments for HearNPV-Albany (Table 5.8).

From the digestions, 9 band fragments had similar sizes in all the genomes. From the similarities and differences, a total of five unique bands were identified from the digests.

Of these unique bands, HearNPV-SP1A had two with sizes of 5,256 bp and 974 bp and HearNPV-AC53, HearNPV-G4 and HearNPV-Albany had one each. The unique bands for HearNPV-AC53, HearNPV-G4 and HearNPV-Albany had sizes of 9,030 bp, 1,482 bp and 3,374 bp, respectively.

The unique band in HearNPV-Albany with a size of 3,345 bp was part of a doublet band fragment, with the other band having a size of 3,374. Compared to the other genomes, only HearNPV-Albany had the doublet fragments at that size.

The two South African genomes had 2 band fragments with sizes of 13,110 bp and 13,166 bp; 4,175 bp and 3,995 bp that were similar in the Chinese isolate HearNPV-C1.

Table 5.8: The fragment sizes of the EcoRI profile for HearNPV-KZN and HearNPV-Albany against other geographically distinct HearNPV genomes. Red: band fragments that are similar in all the genomes. Purple: band fragments that are similar in five or more of the genomes. Green: band fragments found in four genomes, including either HearNPV-KZN or HearNPV-Albany. Black: band fragments found in HearNPV-Albany or HearNPV-KZN and not found in more than two additional genomes or shared with four HearNPV genomes but absent in HearNPV-Albany and HearNPV-KZN.

HearNPV -Au 130,992	HearNPV -KZN 130,636	HearNPV -Albany 130,360	HearNPV -AC53 130,442	HearNPV -C1 130,759	HearNPV -G4 131,405	HearNPV -NNg1 132,425	HearNPV -SP1A 132,481
				14,135	14,131		
13,441		13,575			13,446	13,510	13,547
	13,110	13,166	13,153	12,845			
10,154			10,140		10,154	10,188	10,180
9,484	9,756	9,740		9,752		9,733	9,732
	9,195					9,201	9,198
9,056	9,050		9,056	9,053	9,050		
			9,030				
8,226	8,233	8,228	8,170			8,231	8,234
				6,906			7,156
6,685	6,593			6,534	6,641	6,601	
			6,383	6,300	6,356		6,300
6,285	6,294		6,289		6,290	6,297	
		6,230	6,246			6,230	
5,992	5,997	5,987	5,975	5,994	5,991	5,989	5,984
5,944	5,946	5,958	5,955			5,986	5,926
5,843	5,844	5,843	5,843	5,844	5,843	5,846	5,846
5,838	5,837	5,840	5,837	5,838	5,838	5,843	5,843
5,688	5,687	5,689	5,686	5,672	5,672	5,683	5,685
							5,256
4,833		4,919					
4,751	4,733	4,753	4,745	4,743	4,749	4,748	4,733
	4,687	4,700		4,650			
4,567	4,567	4,570	4,567	4,567	4,567	4,573	4,570
4,411	4,419	4,425	4,419	4,411	4,417	4,412	4,415
4,396	4,396	4,404	4,396	4,396	4,396	4,399	4,402
	4,175	3,995		4,138	4,137		
3,680					3,680		
3,346	3,352	3,374	3,355	3,361	3,359	3,343	3,345
		3,345					
2,998	2,998	3,001	2,998	2,998	2,998	2,998	2,992
		2,830	2,827	2,834	2,834	2,830	2,830
1,742	1,743	1,743	1,742	1,743	1,743	1,743	1,741
	1,397				1,482		
1,007	1,007	1,007	1,007	1,007	1,007	1,007	1,007
							974
784	784	784	784	784	784	781	784
477	475	476	475	476	476	475	476
453	453	453	453	453	453	453	
	414	414		414		414	414
410		410	410	410	410	410	410
306	306	306	306	306	306	306	306
176	176	176	176	176	176	176	176
19	19	19	19	19	19	19	19
130,992	130,636	130,360	130,442	130,759	131,405	132,425	132,481

5.3.6.2 *In silico* digestion of HearNPV genomes using HindIII

After digestion with HindIII, *in silico* using ApE v2.0, HearNPV-Albany, HearNPV-Au, HearNPV-C1 and HearNPV-G4 have 13 digestion sites and HearNPV-AC53, HearNPV-

KZN, HearNPV-NNg1 and HearNPV-SP1A have 14 digestion sites (Table 5.9). In the assayed genomes, eight bands of similar sizes were recorded in all the genomes. The bands were approximately at 22,600 bp, 10,350 bp, 10,050 bp, 7,500 bp, 6,700 bp, 3,250 bp, 2,600 bp and 1,509 bp. There were three bands of similar sizes that are found in all the genomes, except HearNPV-Albany, with sizes of approximately 10,700, 12,900 and 14,400. HearNPV-KZN had two band fragments, at 10,632 bp and 7,840 bp that were shared with HearNPV-AC53, HearNPV-SP1A and HearNPV-NNg1. HearNPV-Albany had three unique bands at 25,495 bp, 15,721 bp and 1,626 bp, whereas HearNPV-KZN has no band unique to the genome,

Table 5.9: The fragment sizes of the HindIII digest for HearNPV-KZN and HearNPV-Albany against other geographically distinct HearNPV genomes. Red: band fragments that are similar in all the genomes. Purple: band fragments that are similar in five or more of the genomes. Green: band fragments found in four genomes including either HearNPV-KZN or HearNPV-Albany. Black: band fragments found in HearNPV-Albany or HearNPV-KZN and not found in more than two additional genomes or shared with four HearNPV genomes but absent in HearNPV-Albany and HearNPV-KZN.

HearNPV-Au 130,992	HearNPV-KZN 130,636	HearNPV-Albany 130,360	HearNPV-AC53 130,442	HearNPV-C1 130,759	HearNPV-G4 131,405	HearNPV-NNg1 132,425	HearNPV-SP1A 132,481
		25,495					
22,897	22,603	23,027	22,642	23,155	22,604	22,605	22,603
16,865				17,125	17,642		
		15,721					
14,582	14,493		14,273	14,440	14,538	14,496	14,601
12,965	12,974		12,911	12,921	12,917	12,972	12,974
		11,350					
10,982		11,111			10,980	11,088	11,085
10,717	10,720		10,712	10,706	10,705	10,790	10,769
	10,632		10,681			10,718	10,721
10,376	10,407	10,399	10,392	10,409	10,398	10,420	10,358
				10,375			
10,050	10,050	10,062	10,067	10,050	10,050	10,060	10,060
	9,354		9,451				
	7,840		7,743			7,745	7,743
7,483	7,500	7,495	7,492	7,496	7,495	7,447	7,489
6,690	6,684	6,696	6,690	6,687	6,684	6,702	6,702
3,254	3,247	3,243	3,256	3,256	3,254	3,250	3,250
2,622	2,623	2,626	2,623	2,629	2,629	2,623	2,617
		1,626					
1,509	1,509	1,509	1,509	1,509	1,509	1,509	1,509
130,992	130,636	130,360	130,442	130,759	131,405	132,425	132,481

5.3.6.3 *In silico* digestion of HearNPV genomes using BamHI

Digestion of the genomes *in silico* using ApE v2.0 and using BamHI as the enzyme was performed. From the digestion, HearNPV-KZN had 10 bands, whereas HearNPV-AC53, HearNPV-Albany, HearNPV-Au, HearNPV-C1, HearNPV-G4, HearNPV-NNg1 and HearNPV-SP1A each had 11 bands. There were six fragments of similar size that were

recorded in all the genomes. There were three bands that are absent in HearNPV-KZN and present in the other genomes. The bands that were in the other genomes and absent in HearNPV-KZN genome were at approximately 14,300 bp, 7,700 bp and 1,900 bp. There were four bands at 15,847 bp, 15,801 bp, 8,725 bp and at 1,223 bp that were unique to HearNPV-KZN. HearNPV-Albany genome had two fragments at 9,934 bp and 6,287 bp that are unique to the genome.

Table 5.10: The fragment sizes of the BamHI digest for HearNPV-KZN and HearNPV-Albany against geographically distinct HearNPV genomes. Red: band fragments that were similar in all the genomes. Purple: band fragments that are similar in five or more of the genomes. Orange: band fragments found in all the genomes excluding either HearNPV-KZN or HearNPV-Albany.

HearNPV-Au 130,992	HearNPV-KZN 130,636	HearNPV-Albany 130,360	HearNPV-AC53 130,442	HearNPV-C1 130,759	HearNPV-G4 131,405	HearNPV-NNg1 132,425	HearNPV-SP1A 132,481
37,294	36,984	37,456	36,976	36,677	37,278	37,410	37,392
32,385	32,083	32,523	31,940	32,645	32,091	32,103	32,099
	15,847						
	15,801						
14,392		14,397	14,382	14,361	14,367	14,382	14,364
13,992			13,694	13,855	13,952	13,918	14,025
12,831	12,825	12,839	12,848	12,828	12,825	12,835	12,835
		9,934					
	8,725						
7,719		7,901	7,691	7,717	7,719	7,702	7,700
		6,287					
					4,149	5,057	5,051
3,985	3,985	3,985	3,974	3,985	3,985	3,979	3,979
3,357			3,901	3,640			
1,892		1,892	1,891	1,893	1,893	1,894	1,892
1,842	1,860	1,843	1,842	1,861	1,843	1,842	1,841
1,303	1,303	1,303	1,303	1,297	1,303	1,303	1,303
	1,223						
130,992	130,636	130,360	130,442	130,759	131,405	132,425	132,481

5.4 Discussion

In this chapter, comprehensive comparative studies of HearNPV-Albany and HearNPV-KZN genomes with seven geographically distinct HearNPV published genomes, namely HearNPV-Au, HearNPV-AC53, HearNPV-H25EA1, HearNPV-C1, HearNPV-G4, HearNPV-SP1A and HearNPV-NNg1 were conducted. The studies were performed to analyse genomic variance of HearNPV isolates from different locations in comparison to HearNPV-Albany and HearNPV-KZN. Relationships of HearNPV-Albany and HearNPV-KZN with the geographically distinct HearNPV-genomes was then analysed through various analytical tools including ORF comparison, phylogenetic trees, SNP studies and REN analysis.

Comparison of the genome sizes showed slight variation across the genomes with HearNPV-Albany being the smallest studied, with 130,360 bp and HearNPV-SP1A being as the largest genome, with 132,485 bp. There was no direct correlation observed for the genome size and number of ORFs. The genome size comparison and number of ORFs confirmed that the size and number of ORFs observed for HearNPV-Albany and HearNPV-KZN were like literature reports (Chen *et al.*, 2001; Zhang *et al.*, 2005; Ogembo *et al.*, 2007; Zhang *et al.*, 2014; Noune and Hauxwell, 2016). The nucleotide identity of HearNPV-Albany and HearNPV-KZN compared to the seven other HearNPV genomes was high, ranging between 93.9% and 95.3%. This showed that the viral genomes are highly conserved even if they are from different geographical locations (Herniou *et al*., 2003; Cory *et al*., 2005).

A more in-depth analysis of the genomes was performed by alignment of HearNPV-Albany and HearNPV-KZN to the seven HearNPV genomes and by comparing the ORF percentage nucleotide identity. ORF identity for HearNPV-Albany and HearNPV-KZN against the seven geographical HearNPV genomes was highly conserved with percentage identities for most of the ORFs being above 99%. A few ORFs were variable with a lower percentage identity than 97% in HearNPV-Albany and HearNPV-KZN. Variability was observed mainly for *hrs* and *bro* regions, and this has been reported in literature as the main contributing factor to genome variations in HearNPV isolates from different locations (Zhang *et al*., 2005; Ogembo *et al.*, 2009). Other ORFs with a predicted putative function in both HearNPV-Albany and HearNPV-KZN exhibited variability. For the HearNPV-Albany genome, *hoar, lese-25 like protein, lef-11, pp31* and *ubiquitin* exhibited a more variable nucleotide sequence when compared to the other HearNPV genomes. The *hoar* gene has been reported to have a less conserved nucleotide sequence in comparison to other conserved HearNPV genes (Noune and Hauxwell, 2016), and the results of the ORF alignment confirmed this to be the case for HearNPV-Albany. The role and function of *hoar* is not fully understood. It is hypothesised that it is potentially a ubiquitin ligase and has been reported to be unstable (Rohrmann, 2013). The *hoar* gene for HearNPV-KZN also exhibited nucleotide variability compared to the other HearNPV genomes. These results confirmed that in both HearNPV-Albany and HearNPV-KZN, the *hoar* gene had a higher variability than the more conserved ORFs. In HearNPV-Albany, *ubiquitin* had a 66.01% nucleotide sequence identity compared to the other seven HearNPV genomes in the analysis. For the HearNPV-KZN genome, a low gene identity of 78.37% for the *ie-1* gene was observed when compared to HearNPV-G4, HearNPV-Au, HearNPV-NNg1, HearNPV-AC53, HearNPV-H25EA1 and HearNPV-SP1A. When

compared to HearNPV-C1 a similarity of 99.45% was observed for the gene. The *ie-1* gene is a replication gene responsible for the immediate early protein, with domains associated with dimerisation, nuclear import, DNA binding, transactivation and replication and has been reported in all sequenced lepidopteran baculoviruses (Berretta *et al.*, 2013). The low sequence identity compared to the other genomes, as well as the lack of *gp37* in HearNPV-KZN genome suggests a different primary infection potential. A high nucleotide sequence identity of the *ie-1* gene from HearNPV-KZN with HearNPV-C1 of 99.45% compared to HearNPV-KZN and the other HearNPV genomes of 78.37%, leads to the possibility that the evolution of the gene is not isolated to HearNPV-KZN and may play an active role in viral diversity and virulence. The *gp37* gene is associated with the ODV and may be involved in oral infection (Liu *et al.*, 2006). The *gp37* gene is believed to be involved in the disruption of the peritrophic membrane along with the FP proteins (Herniou *et al*., 2003), and in HearNPV-KZN the *fp* gene had a high nucleotide identity to the other HearNPV strains. Because of the lack of *gp37* gene in HearNPV-KZN it can be inferred that unlike the other HearNPV strains, HearNPV-KZN may only use FP for targeting the chitin matrix for virus access. A low sequence similarity for *cathepsin-like protein* in the HearNPV-KZN genome of 50.45% was observed compared to the HearNPV genomes. *Cathepsin-like protein* is an auxiliary gene that is probably associated with host cuticle disruption and host liquefaction (O'Reilly, 1997). The genes that are absent or significantly variable in the HearNPV-KZN genome are directly associated with the viral interaction with host during the infectivity cycle (Herniou *et al.*, 2003; Liu *et al.*, 2006; Rohrmann, 2013).

For the *bro* genes and *hrs* repeats, low conservation was observed in both HearNPV-Albany and HearNPV-KZN genomes compared to the other studied HearNPV genomes. It has been reported that *hrs* do not share significant sequence identity between baculoviruses, showing significant difference in structure and sequence (Hu *et al*., 2003). *Bro* genes have also been reported to be highly divergent in baculovirus genomes (Harrison and Hoover, 2012). The HearNPV-Albany genome only encoded for one *bro* gene namely *bro-b*; HearNPV-KZN encoded for three *bro* genes namely a *bro-a* and two *bro-b*. The sequence identities of *hrs* in HearNPV-Albany and HearNPV-KZN compared to HearNPV genomes studied in this chapter were very low, most of them being below 72% identity. HearNPV-Albany had four *hrs* identified, lacking *hr3* and HearNPV-KZN had five *hrs* identified. Arrizubieta *et al*. (2015) identified two *bro* genes and five *hrs* for HearNPV-SP1A and HearNPV-SP1B. Five *hrs* 1-5 are located on the HearNPV-G4 genome and when compared to HearNPV-C1 that

had been isolated from the same region in China, it was reported that the *hrs* were highly variable with a low sequence similarity in three of them (Chen *et al*., 2001; Zhang *et al*., 2005). The *hrs* are common in NPVs and serve as enhancers as well as putative origins of DNA replication, being genetically diverse in the different isolated HearNPV genomes as they may be sites for intragenomic recombination and variation (Erlandson, 2009; Rohrmann, 2013).

From the phylogenetic studies using the neighbour-joining method and maximum-likelihood method, similar trees were produced. The pairwise nucleotide distance of the HearNPV isolates ranged from 0.001 to 0.019, which was consistent to what was reported by Rowley *et al.,* (2011). Three distinctive clusters of HearNPV isolates could be discerned, with one group having African and European isolates excluding HearNPV-KZN, the other group having Australian isolates and the third cluster having HearNPV-KZN, HearNPV-Au and the Chinese isolates. Kaur *et al*., (2014) also reported the clustering of HearNPV-C1, HearNPV-G4 and HearNPV-Au and Rowley *et al.,* (2011) reported the grouping of HearNPV-NNg1 with a South African HearNPV isolate. The phylogenetic analysis of the HearNPV isolates seemed to group them according to geographical location with the exception being HearNPV-KZN. For phylogenetic trees the grouping of African and European HearNPV genomes has been observed (Noune and Hauxwell, 2016). The HearNPV-KZN genome showed the greatest genetic variance and evolutionary difference from the second branch. The lack of *gp37* and a very low nucleotide sequence similarity for *cathepsin* and *ie-1* that was observed when analysing HearNPV-KZN, could account for this evolutionary divergence of the genome.

Synonymous SNPs were observed in the *polh* gene and synonymous and non-synonymous SNPs were observed in *lef-9* gene. The nucleotide variation observed for the HearNPV genomes, did not translate into amino acid changes for the *polh* gene therefore maintaining the protein structure and sequence across the genomes. The Polyhedrin protein is one of the most conserved baculovirus proteins (Rohrmann, 2013). For the *lef-9* gene, eight SNPs in total were observed with two of the SNPs being non-synonymous. The non-synonymous SNPs resulted in amino acid substitutions of Glutamic acid and Lysine as well as Leucine and Methionine. Glutamic Acid is an acidic amino acid, whereas Lysine is a basic amino acid. The hydration potential of Glutamic acid is -10.19 and that of Lysine is -9.52 (Strickberger, 2000). The hydration potential of the two amino acids is similar, therefore substituting one for the other should not affect the structure and polarity of the protein. Leucine and

Methionine are both non-polar amino acids and the substitutions of the amino acids has no overall effect on the protein polarity with the major difference being potentially contributed by side chain bonds of the amino acids (Wilson and Walker, 2005). The observations of the trends of the SNPs confirm the phylogenetic tree, with HearNPV-Albany being closely related to HearNPV-AC53, HearNPV-NNg1 and HearNPV-SP1A whereas HearNPV-KZN shared more similarities with the Chinese HearNPV and HearNPV-Au.

In silico REN profiles of the HearNPV genomes using BamHI, EcoRI and HindIII resulted in unique genetic fingerprints for each of the genomes studied. This result confirmed the novelty of HearNPV-Albany and HearNPV-KZN compared to geographically distinct genomes. The study confirmed the diversity and novelty of each of the HearNPV genomes, supporting the theory that the viruses evolve and mutate along with host genes (Summers, 2006)

5.5 Conclusions

The HearNPV-Albany and HearNPV-KZN genomes were compared to geographically distinct HearNPV genomes from Australia, Asia, Europe and Africa. Genetic variation between the genomes was observed confirming novelty of HearNPV-Albany and HearNPV-KZN. An alignment of the genomes and the ORFs showed some variances in the genomes confirming the differences of the HearNPV isolates. Phylogenetic analysis of the HearNPV isolates identified potential co-evolution and divergent evolution of HearNPV-Albany and HearNPV-KZN in comparison to other geographically distinct HearNPV isolates. This raises the potential for identification of migration patterns of HearNPV and their relationships globally. The difference in the genetic makeup of HearNPV-Albany and HearNPV-KZN against other HearNPV genomes was confirmed through REN and analyses of SNPs. REN profiles generated, and SNPs observed confirmed relationships of the HearNPV genomes and identified unique genetic fingerprints for HearNPV-Albany and HearNPV-KZN.

The genetic differences observed in HearNPV-Albany and HearNPV-KZN against HearNPV isolates from different geographical locations represents viral diversity at the genome level and potential differences in virulence factors. In the next chapter, the biological activity against Albany *H. armigera* of HearNPV-Albany, HearNPV-KZN, HearNPV-Whit, HearNPV-Haygrove will be compared to geographically distinct HearNPV isolates from Spain and China namely HearNPV-SP1 and HearNPV-G4 as well as two commercial HearNPV products against Albany *H. armigera*. Further biological activity against Haygrove

H. armigera will be tested for HearNPV-Albany, HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove.

6 Comparison of the biological activity of eight HearNPV isolates from geographically distinct locations against South African collected *Helicoverpa armigera*

6.1 Introduction

In this chapter, eight HearNPV isolates that had been isolated and purified in Chapter 3 were assayed against laboratory established South African *Helicoverpa armigera* larvae that were reported in Chapter 2. Genomic characterisation of four of the viruses in Chapter 5 namely HearNPV-Albany, HearNPV-KZN, HearNPV-G4 and HearNPV-SP1, identified genotypic differences. The virulence of the different HearNPV isolates was studied to determine biological differences through dose-response (concentrations required for mortality) and time response (speed of kill) bioassays.

Baculovirus-infected larvae can live for days or even weeks post-infection, during which period they continue feeding, resulting in continued crop damage (Harrison and Hoover, 2012); this raises the need for fast killing baculoviruses (Sun, 2015). Early instar *H. armigera* larvae cause less feeding damage and are more susceptible to HearNPV and this makes the targeting of early instar larvae economical and the most effective control method (Siegwart *et al*., 2015). The feeding damage can be minimised or kept below an economic threshold by timing the application of HearNPV so that early instar larvae are affected. As *H. armigera* larvae mature, they become more resistant to HearNPV infection (Siegwart *et al*., 2015). There are two types of viral infections namely acute infection that results in rapid mortality and latent infection where the virus is present at a low level, resulting in it being non-infective (without mortality) (Villareal, 1999). Acute virus infections are common in *H. armigera* because of the high population structure of having various life stages of the pest at any moment, as well as high population density (Villareal, 1999).

The amount of HearNPV biopesticide and the speed of kill of the biopesticide are used to determine the activity of the biopesticide (Behle and Birthisel, 2014). Effective control of *H. armigera* populations can be aided using large baculovirus concentrations, which result in shorter larval survival time (Harrison and Hoover, 2012). The virulence of the virus can be tested *in vitro* using cell cultures or *in vivo* using whole live insects. Cell cultures eliminate the potential of latent virus however and therefore are not a true reflection of virus efficacy as they do not consider all conditions required for virus infection in an insect host (Jones, 2000).

Indigenous HearNPV isolates have been reported to be more effective against local *H. armigera* populations than exotic isolates (Cherry *et al*., 2003). High dosages of up to 1×10^{13} OB per hectare are required to match standard chemical insecticide results in the field (Cherry *et al*., 2003). Different strains of HearNPV isolated from different geographical locations result in different potencies of the viruses against host populations (Jones, 2000). In trials in Botswana, HearNPV showed promising results in the control of *H. armigera* in sorghum, cowpeas and maize fields. The use of H. zea NPV or *Bt* in the control of *H. armigera* was shown not to be as effective as the use of local HearNPV (Cherry *et al*., 2003). Seasonal variations and changes did not affect the symptom induction on *H. armigera* by HearNPV in South Africa (Whitlock, 1974). In South Africa, *H. armigera* NPV was first reported by Whitlock (1974), from *H. armigera* larvae that had succumbed to baculovirus infection. Application of HearNPV against early immature stages of *H. armigera* in the field and laboratory resulted in mortality within the first 9 days (Whitlock, 1974). In 2004 the virus isolated by Whitlock (1974) was tested for virulence against *H. armigera* field populations in South Africa and found to be effective (Moore *et al*., 2004). In South Africa Helicovir™ (HearNPV) is currently registered as a bio-insecticide for the control of *H. armigera* on citrus and other crops including tomatoes, peppers, berries and vegetables. Helicovir is formulated using a HearNPV isolate from Australia (Moore and Kirkman, 2010). *Helicoverpa armigera* on citrus plants was susceptible to Helicovir™ and the HearNPV formulation resulted in the lowest infestation of *H. armigera* over a prolonged period compared to a chemical pesticide (Mevinphos) and a *Bacillus thuringiensis* bacterial product (DiPel, *Bt*) (Moore and Kirkman, 2010). *Helicoverpa armigera* neonate larvae collected from tobacco and cotton plants in fields from Limpopo Province in South Africa had a dose-dependent mortality resulting from HearNPV applications on all the populations studied (Grant and Bouwer, 2009).

For a good bioassay of HearNPV against *H. armigera*, variation of the testing conditions should be minimised, and the assays should accurately reflect the conditions being tested. The quantity of virus to be used for the assays must be accurately known and this can be done for HearNPV through the enumeration of OBs (Jones, 2000). Mass dosing assays allow for the determination of a dose-response relationship (including estimation of a lethal concentration that kills 50% of individuals in a sample of the population) and a time-response relationship (including an estimation of the time to death for 50% of the individuals in a sample of the population) between pathogen and host (Kalawate, 2014). For assays against *H. armigera*, droplet feeding of the larvae with the virus, incorporation of the virus into the

larval diet or surface dosing of the virus onto the diet can be performed (Kalawate, 2014). The handling of larvae during bioassays may result in the death of the larvae. Newly hatched neonates are more susceptible to handling mortality whereas older larvae may have diverging growth patterns and from the third instar of *H. armigera*, also be cannibalistic (Grzywacz *et al.*, 2007). It is necessary to minimise variation of the larvae in the bioassay setup (Jones, 2000). In Chapter 2, instar was recorded through head capsule moulting and head capsule measurement. The identification of newly moulted second instar larvae was established and confirmed consistently, with minimum variation recorded for the newly moulted second instar larvae as this was the life stage to be used for bioassays.

The aim of this chapter was to conduct virulence studies with eight HearNPV isolates, namely HearNPV-Albany, HearNPV-KZN, HearNPV-Whit, HearNPV-Haygrove, HearNPV-G4, HearNPV-SP1, Helicovir™ and Helicovex®, against South African collected second instar *H. armigera* larvae. The first objective of the study was the establishment of the virulence of the different virus isolates through determining doses required for minimum (LC_{20},) median (LC_{50}) and 90% (LC_{90}) mortality. The second objective was the establishment of the time taken to achieve 50% (LT_{50}) and 90% (LT_{90}) mortality of the *H. armigera* larvae in a sample. The third objective was to compare the virulence of South African HearNPV isolates against different geographically distinct South African populations of *H. armigera*.

6.2 Methodology

6.2.1 Source of insects

For the bioassays, two *H. armigera* strains that were reared and maintained in the insectary at Rhodes University at 26.0°C (± 2.0°C) on artificial bollworm diet (Chapter 2) were used. One colony was established from larvae collected from the Belmont Valley near Grahamstown, South Africa and this was referred to as the Albany colony. The second colony was established from larvae collected from Haygrove Farms near George, South Africa and this was referred to as the Haygrove colony. The Albany colony had been maintained in the insectary for more than 15 generations and the Haygrove colony had been maintained for two generations.

6.2.2 Geographically distinct HearNPV isolates

HearNPV-Albany, HearNPV-KZN, HearNPV-Haygrove and HearNPV-Whit from South Africa that had been isolated and purified, as described in Chapter 3, were used in the assays. A commercially produced HearNPV product from South Africa that uses an Australian

isolated HearNPV, produced by River Bioscience, called Helicovir™, was one of two commercial products used in the assays. There were three other isolates that were included for the assays, provided by Primitivo Caballero (Instituto de Agrobiotecnologia, Public University of Navarra in Pamplona, Spain). The HearNPV isolates provided were HearNPV-SP1 isolated from Spain (Arrizubieta *et al*., 2013), HearNPV-G4 isolated from China (Chen *et al*., 2001) and a commercial formulation Helicovex® made by Andermatt Biocontrol AG from Switzerland.

6.2.3 Occlusion body enumeration

The concentrations of occlusion bodies (OBs) for HearNPV-Albany, HearNPV-KZN, HearNPV-Whit, HearNPV-Haygrove, HearNPV-G4, HearNPV-SP1, Helicovir™ and Helicovex® isolates were enumerated by counting under a light microscope using a method adapted from Grzywacz *et al*. (2007). A Reichert improved Neubauer haemocytometer (Bright-line®, Hausser Scientific, USA), with a depth of 0.1 mm was used. The virus stock suspensions were pooled and mixed prior to virus enumeration. Before virus enumeration, 10-fold serial dilutions were performed on the stocks, up to 10^3 dilutions. The counting of OBs was performed for each isolate in five large squares (with 16 small squares making up one large square) and this process was repeated five times (with a freshly prepared sample before each bioassay. The five large squares counted were four from each corner and the middle square. The concentration of the HearNPV suspension was calculated using the formula $\frac{total\ number\ of\ OBs \times dilution\ factor}{number\ of\ squares \times volume\ of\ one\ square}$. OB counts were performed using phase-contrast light microscopy.

Virus enumeration was performed by adding 10 µl of HearNPV isolates suspensions to a Neubauer improved cell counting chamber on the middle of the counting grids with the cover slip already on using the V-grooves on the haemocytometer and allowing the suspension to spread and fixed through capillary attraction. The counting chamber was left for five minutes to allow for non-viral particles to sediment. The enumeration was performed on a light microscope (Olympus BX50) at 400 × magnification. The OBs were visualised as bright vibrating particles and were counted and recorded if the total OBs for the five squares was between 200 and 400. The counting chamber was rinsed with 70% ethanol and ddH_2O.

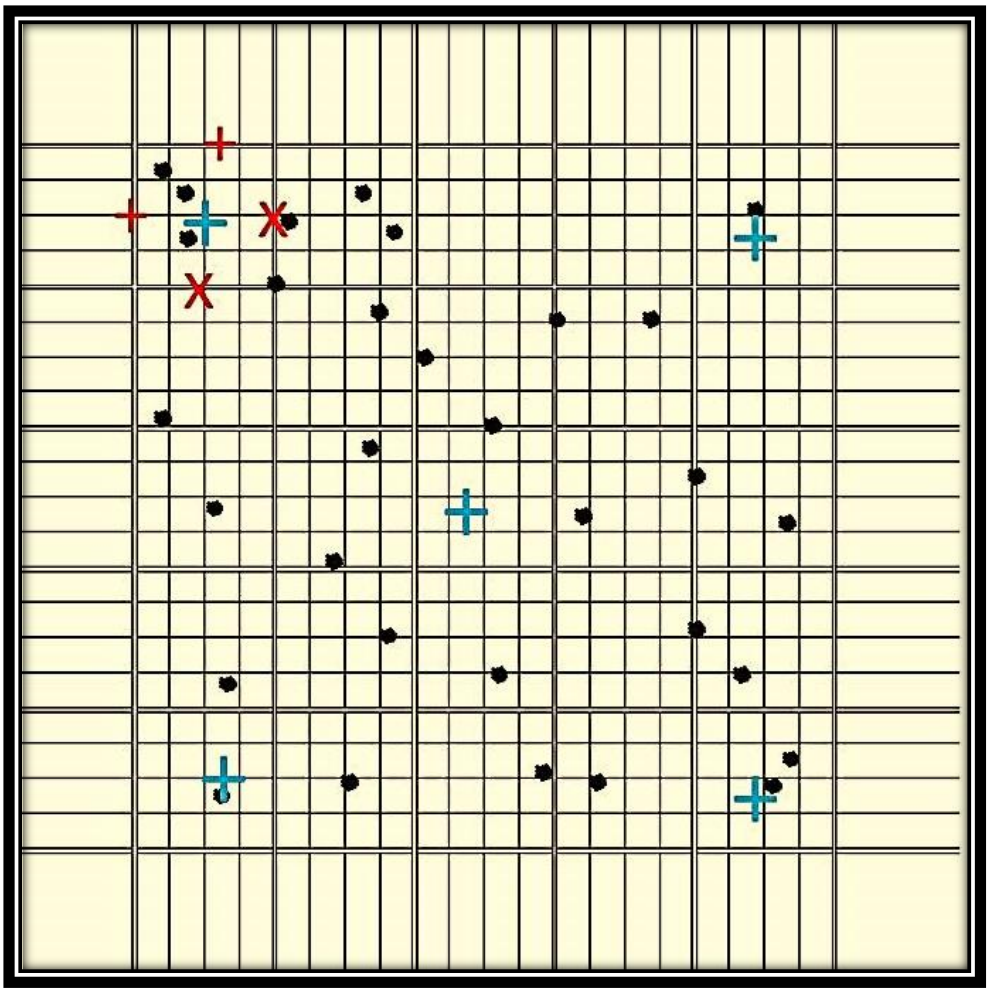


Figure 6.1: The design of the counting chamber showing grids where counts are conducted. Total counts are conducted for OBs in 16 small squares that make up one large grid and a total of 5 large grids are counted. +: grids that were used for OB enumeration. +: The top and the left edges of the grid that are included in the count if OBs are on them. **x**: The bottom and right edges that are excluded from OB enumeration.

6.2.4 Dose preparation for dose-response bioassays

In 20 ml sterile tubes, serial dilutions from the stock virus concentration were performed before bioassays were conducted. 10-fold serial dilutions were performed initially for HearNPV-Albany, HearNPV-KZN, HearNPV-Whit, HearNPV-Haygrove, HearNPV-G4, HearNPV-SP1, Helicovir™ and Helicovex® and preliminary bioassays were performed on the Albany collected *H. armigera* second instar larvae to establish concentrations required for between 10% and 91-99% mortality. The bioassays were performed by treating the diet with 200 µl of the virus suspension. The artificial bollworm diet (Chapter 2) had a volume of 0.78 cm^3 in the individual glass vials (one cm height of diet) (Grzywacz *et al.*, 2007; Bouwer *et al.*, 2009; Pudijanto *et al.*, 2016). Control experiments that substituted HearNPV-Albany for ddH_20 were setup. The total number of larvae in each test assay experiment was 25. Doses

that resulted in 10% to 91-99% mortality were then identified. Mortality was determined through the following formula:

$$[\frac{25-\text{number of dead larvae in test experiment}}{25-number\ of\ dead\ larvae\ in\ control\ experiment}]*100$$

From the initial concentrated stock suspension, 10-fold serial dilutions were performed until a concentration that resulted in between 90-99% mortality was reached, and this was referred to as the working suspension (D1). From the dose stock suspension, 5-fold serial dilutions were performed for 5 concentrations (Figure 6.2). This process was repeated for all the other HearNPV isolates that were being assayed. All experiments were conducted in triplicate.

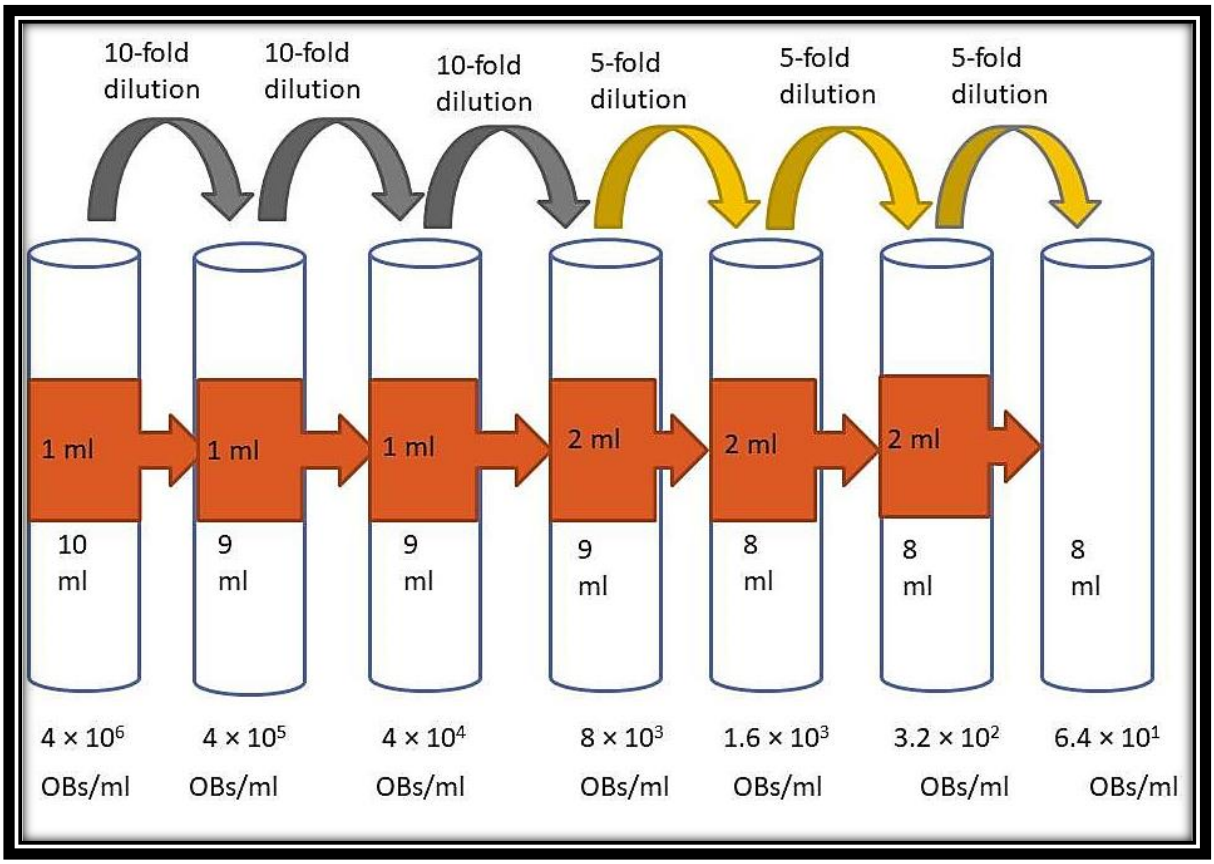


Figure 6.2: Process of 10-fold and 5-fold serial dilutions of HearNPV concentrations for bioassays, explaining the calculations.

6.2.5 Surface dose bioassays for dose-response studies

Surface dose bioassays for the HearNPV isolates were performed on newly moulted second instar *H. armigera* Albany larvae or Haygrove larvae (Figueiredo *et al.*, 2009). Bioassays were conducted in individual glass vials (Chapter 2). The surface dose method was performed using a modified protocol from Christian *et al.*, (2001) and Grzywacz *et al.*, (2007). A volume of 200 µl of the treatment virus was placed on the diet in the glass rearing vials. The

virus was left for 24 hours before adding larvae. A range of five virus doses were used for the assays that resulted in mortalities between 10% and 91-99% (Table 6.1). For each of the experimental setups 25 larvae were used. Control experiments using ddH_2O instead of virus were setup and run concurrently. The bioassays were carried out for 10 days with initial observations being carried out every 12 hours for the first 48 hours. Larvae that had died within the first 48 hours were excluded from the assay because virus induced mortality would only occur more than 48 hours post infection (Claus *et al*., 2012). Final mortality was recorded after 10 days for the different concentrations. The bioassays were separated over time and/or space (i.e. at different times or in different controlled environment (CE) rooms) and were performed in triplicate. The CE rooms used were at 27°C ± 2.0°C with the same day: light photoperiods of 14:10 hours.

Table 6.1: The concentration of the HearNPV isolates ($OB.ml^{-1}$) that were used for dose-response studies after 5-fold dilutions.

Isolate	Concentration ($OB.ml^{-1}$)				
HearNPV-Albany	4.52×10^3	9.04×10^2	1.81×10^2	3.62×10^1	7.23
HearNPV-KZN	5.19×10^3	1.04×10^3	2.08×10^2	4.15×10^1	8.31
HearNPV-Whit	3.96×10^3	7.92×10^2	1.58×10^2	3.17×10^1	6.33
HearNPV-Haygrove	2.14×10^3	4.27×10^2	8.54×10^1	1.71×10^1	3.41
HearNPV-SP1	3.00×10^3	6.00×10^2	1.20×10^2	2.40×10^1	4.80
HearNPV-G4	3.00×10^3	6.00×10^2	1.20×10^2	2.40×10^1	4.80
Helicovir™	2.00×10^4	4.00×10^3	8.00×10^2	1.60×10^2	3.20×10^1
Helicovex®	2.00×10^4	4.00×10^3	8.00×10^2	1.60×10^2	3.20×10^1

6.2.6 Surface dose bioassays for time-response determination

Surface dose bioassays for the HearNPV isolates were performed using the calculated concentrations after probit analysis that resulted in LC_{90} mortalities in the dose-response bioassays. Surface dose bioassays were set up as described above using LC_{90} concentrations. The virus was left to dry for 24 hours and 25 recently moulted second instar larvae were placed in the individual vials. Mortality was recorded every 12 hours until 100% mortality was recorded. Mortality recorded within the first 48 hours was excluded from the analyses because virus induced mortality would not occur sooner than 48 hours post infection (Claus *et al*., 2012).

6.2.7 Statistical data analyses

The data recorded from the dose-response assays and time-response studies were analysed in ToxRat Professional v3.2.1. Raw data were entered for the different treatment concentrations to get a quantal mortality response. A probit linear maximum likelihood regression analysis was performed on the data for the dose-response and a logit analysis for time response. Abbott's correction was used to compensate for the control response (Abbott, 1925). The lethal concentration confidence limits were based on Fieller's theorem and concentrations of LC_{20}, LC_{50} and LC_{90} were chosen as point estimates. LT_{20} LT_{50} and LT_{90} were chosen as end points for time response estimates. Statistical analyses including one-way ANOVA, student t-test, Levene's test on variance homogeneity and Shapiro-Wilk's test on normal distribution were also performed in ToxRat Professional v3.2.1. A statistical comparison of the samples was performed by using HearNPV-Albany as the reference sample and the other HearNPV isolates as the test samples. A summary t-test was performed on the reference isolate against the test isolates for each of the HearNPV samples in Statistica 13.2, using $p < 0.05$ for statistical significance. Relative potency was calculated using the probit values, which are deduced from logarithmic calculations, resulting in maintenance of parallel samples (Bursa, 2016). For the relative potency, HearNPV-Albany was used as the reference strain (Ref) and the other strains as the test strains (tes) and the relative potency was calculated using the following formula:

$$\text{Relative potency} = \frac{Ref\ probit\ value}{tes\ probit\ value}$$

If< 1 it means less potent, =1 means similar potency and > 1 means more potent. The relative mean time course which represent the time required to achieve mortality of 50% or 90% as a ratio of the reference strain was calculated using HearNPV-Albany as the reference. This was calculated as:

$$\text{Relative mean time course} = \frac{tes\ logit\ value}{Ref\ logit\ value}$$

If < 1 it means faster speed of kill, = 1 means similar speed of kill and > 1 means slower speed of kill.

6.2.8 PCR amplification of *polh*

To confirm that the mortality recorded in the dose-response bioassays was a result of baculovirus infection and not some other factor, OBs were extracted from larvae that had died in each of the bioassays with each of the HearNPV isolates. Genomic DNA was

extracted from the OBs and PCR amplification of *polh* was performed to confirm the presence of NPVs in the cadavers. PCR amplification was performed as described in section 3.2.6 in Chapter 3 using polh/gran degenerate primers (Lange *et al*., 2004). The PCR products were analysed on 0.7% agarose gel, using agarose gel electrophoresis. Agarose gel electrophoresis was performed to determine the DNA quality. To 8 µl of DNA, 2 µl of loading dye was added and the mixture was loaded on a 0.7% AGE that had ethidium bromide (2 µl for every 100 ml of AGE). The gel was run at 90 V for 30 minutes using 1× TAE buffer (40 mM Tris-acetate, 20 mM acetic acid and 1 mM EDTA) as the running buffer. The gel was viewed, and images captured using a UVIpro chemidoc (UVItec, UK) UV trans-illuminator.

6.3 Results

6.3.1 Enumeration of occlusion bodies in HearNPV isolates

The OB concentrations of HearNPV-Albany, HearNPV-KZN, HearNPV-Whit, HearNPV-Haygrove, HearNPV-SP1, HearNPV-G4, Helicovir™ and Helicovex® isolates were determined by enumeration under phase-contrast light microscopy. The OBs were identified as vibrating molecules that refracted light (Figure 6.3).

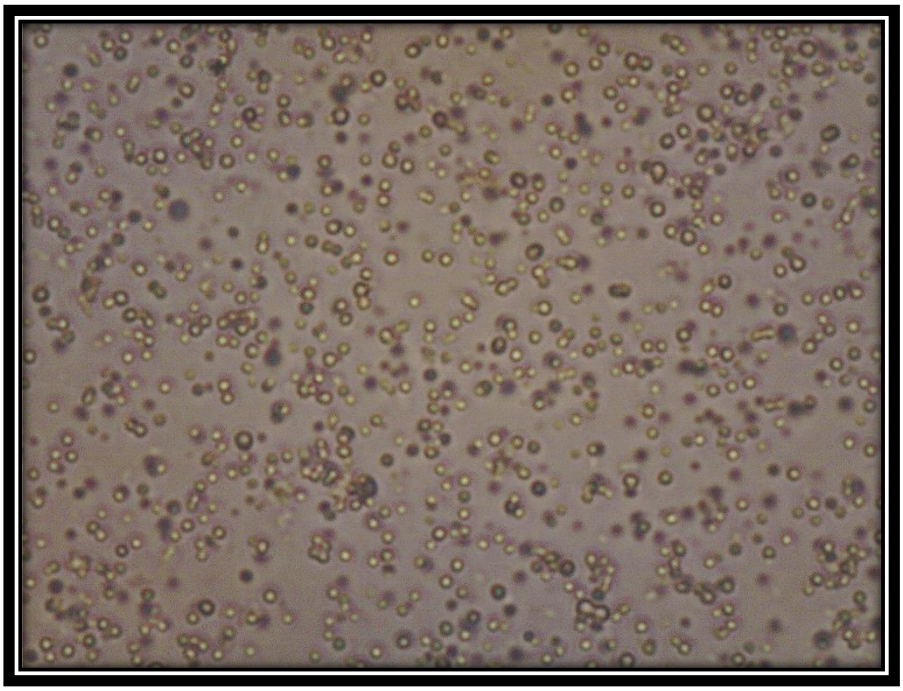


Figure 6.3: Image captured during visualisation of HearNPV-Albany OBs under phase-contrast light microscopy. Viewed under objective and eyepiece at ×400

6.3.2 Surface dose bioassays against second instar Albany *H. armigera* larvae

Surface dose assays were performed to estimate the LC_{50} and LC_{90} values of the HearNPV isolates against Albany second instar *H. armigera* larvae. Preliminary bioassay studies were performed using all the HearNPV isolates to establish the concentrations required for a mortality below 100% of the population that would allow for the calculation of LC_{50} and LC_{90} values. Larvae that succumbed to HearNPV were identified through symptoms including hanging upside down (Figure 6.4), a glossy cuticle or complete liquefaction of the larva on the interior sides of the rearing vial (Federici, 1997; Murphy *et al.*, 2004). For all bioassays, final mortality was recorded after 10 days.

Figure 6.4: Baculovirus diseased *H. armigera* cadaver in a larval rearing vial after surface dose assaying showing symptoms of virus infection; hanging upside down and liquefaction.

6.3.2.1 HearNPV-Albany bioassays

The mortality of second instar Albany *H. armigera* larvae after treatment with HearNPV-Albany was studied. The number of larvae that had succumbed to the HearNPV-Albany treatment for the different concentrations was recorded with the control mortality recorded as the first treatment (Figure 6.5). The concentrations used in the study were between 7.2 $OB.ml^{-1}$ and 4.5×10^3 $OB.ml^{-1}$.

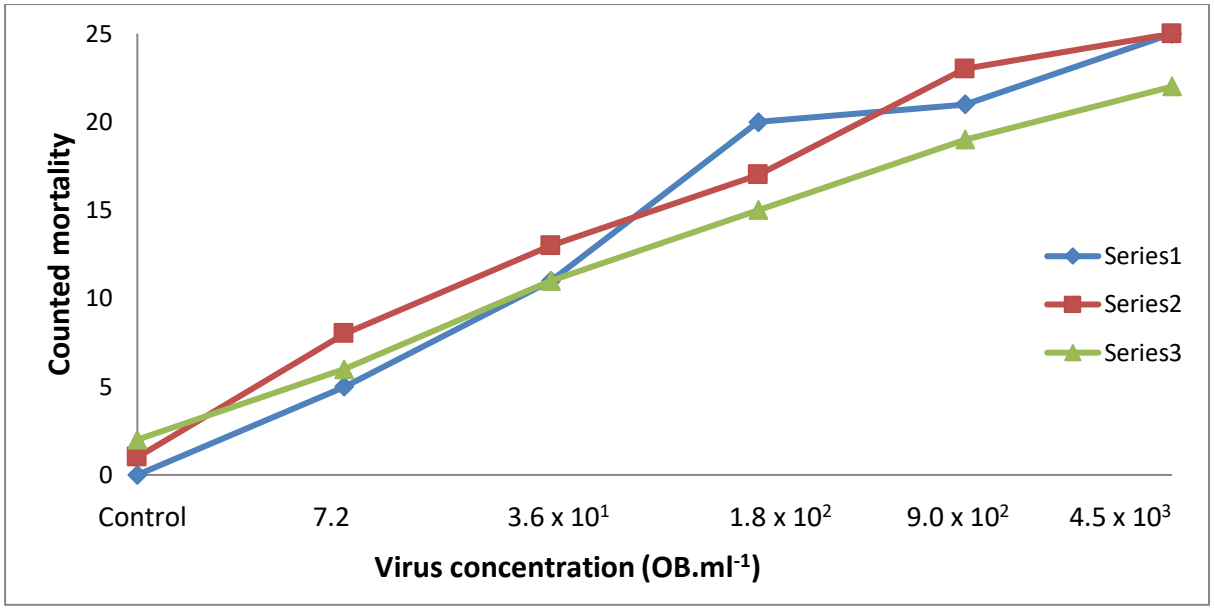


Figure 6.5: The counted mortality of second instar *H. armigera* larvae against different concentrations of HearNPV-Albany for three replicates namely Series 1, 2 and 3. N = 25

6.3.2.2 HearNPV-KZN bioassays

The mortality of second instar Albany *H. armigera* larvae when treated with different concentrations of HearNPV-KZN was recorded and the control mortality was used as the first data point (Figure 6.6). The concentrations studied were between 8.3 and 5.2×10^3 OB.ml^{-1}.

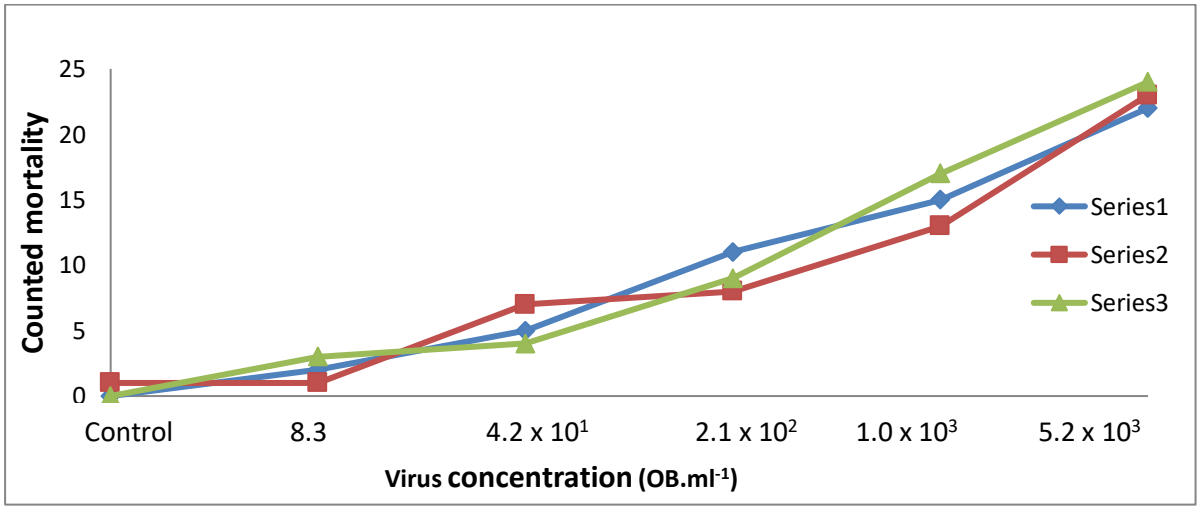


Figure 6.6: The counted mortality of second instar Albany *H. armigera* when treated with different concentrations of HearNPV-KZN for three replicates namely Series 1, 2 and 3. N = 25

6.3.2.3 HearNPV-Whit bioassays

The mortality of second instar *H. armigera* after being treated with HearNPV-Whit at different concentrations was recorded and the control mortality recorded as the first data point (Figure 6.7). The concentrations studied were between 4.2×10^1 and 2.6×10^4 OB.ml^{-1}.

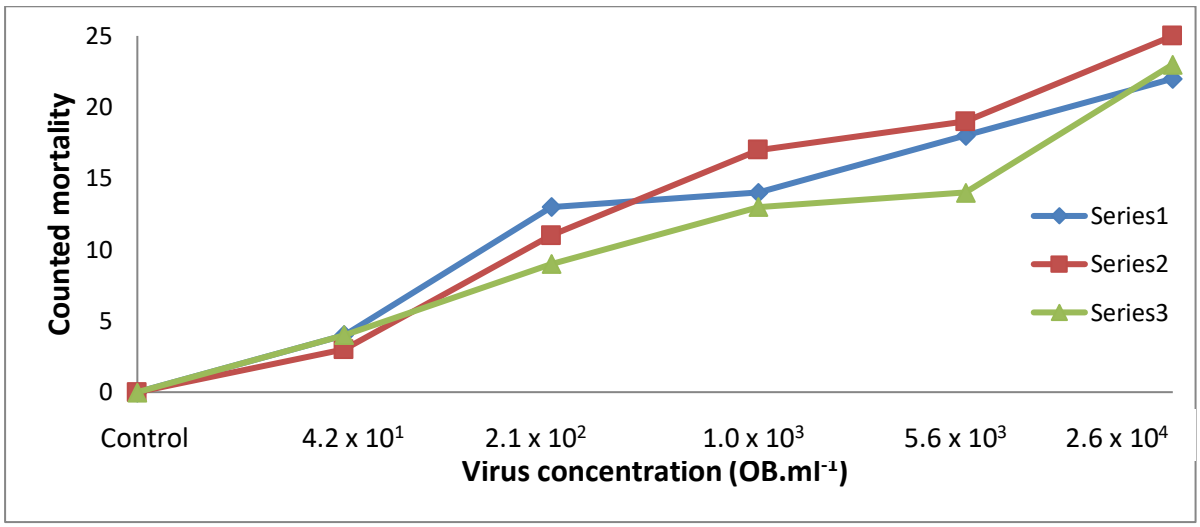


Figure 6.7: HearNPV-Whit bioassays against second instar Albany *H. armigera* showing the counted mortality versus virus concentration for three replicates namely Series 1, 2 and 3. N = 25

6.3.2.4 HearNPV-Haygrove bioassays

Second instar Albany *H. armigera* were treated with different concentrations of HearNPV-Haygrove and the control mortality was recorded as the first data point (Figure 6.8). The concentrations studied were between 3.4 and 2.1×10^3 OB.ml^{-1}.

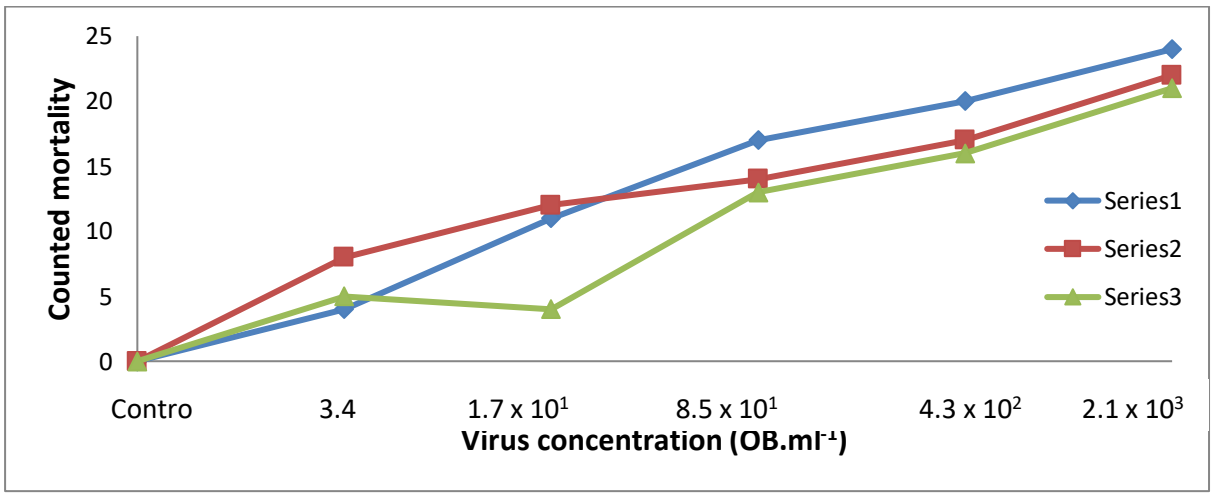


Figure 6.8: The mortality of second instar Albany *H. armigera* when treated with different concentrations of HearNPV-Haygrove over a period of 10 days for three replicates namely Series 1, 2 and 3. N = 25

6.3.2.5 HearNPV-G4 bioassays

Second instar *H. armigera* larvae were treated with different concentrations of HearNPV-G4 the mortality was recorded, with the control mortality set as the first data point (Figure 6.9). Concentrations studied were between 4.8 and 3.0×10^3 OB.ml^{-1}.

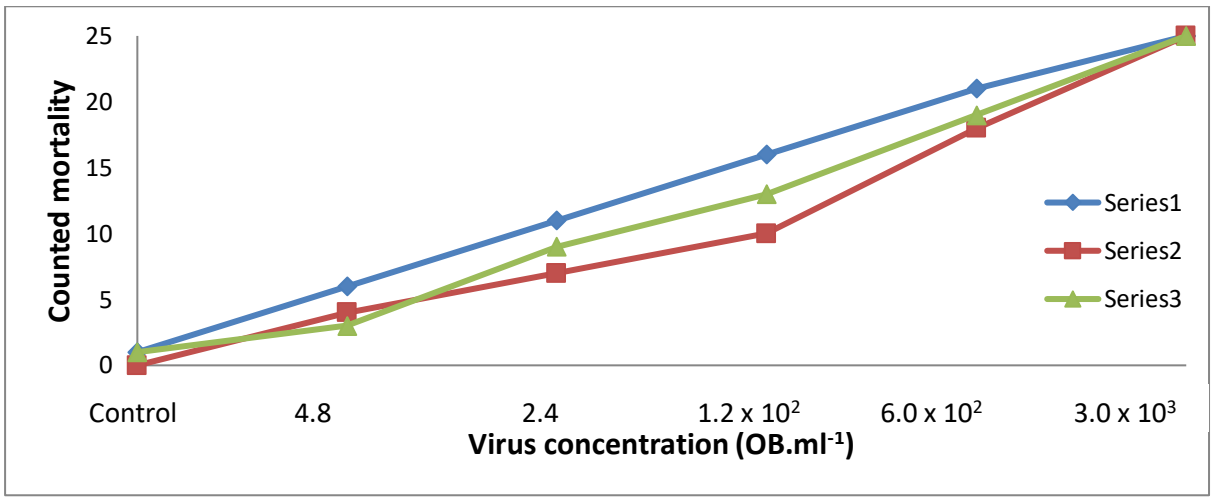


Figure 6.9: The dose responses of second instar Albany *H. armigera* to different concentrations of HearNPV-G4 over a period of 10 days for three replicates namely Series 1, 2 and 3. N = 25

6.3.2.6 HearNPV-SP1 bioassays

Bioassays were carried out on second instar *H. armigera* larvae using different concentrations of HearNPV-SP1 and the control mortality was recorded as the first data point (Figure 6.10). The concentrations used were between 4.8 and 3.0×10^3 OB.ml^{-1}.

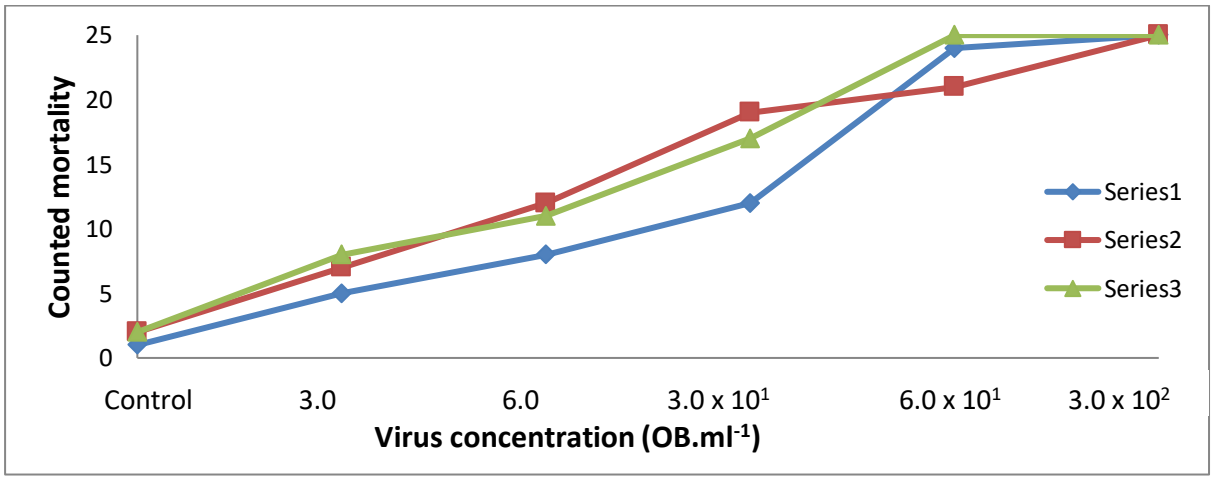


Figure 6.10: The dose responses of second instar Albany *H. armigera* to different concentrations of HearNPV-SP1 over a period of 10 days for three replicates namely Series 1, 2 and 3. N = 25

6.3.2.7 Helicovir™ HearNPV bioassays

Helicovir™ bioassays were performed against second instar Albany *H. armigera* and the control mortality was recorded as the first data point (Figure 6.11). The concentrations used were between 3.2×10^1 and 2.0×10^4 OB.ml^{-1}.

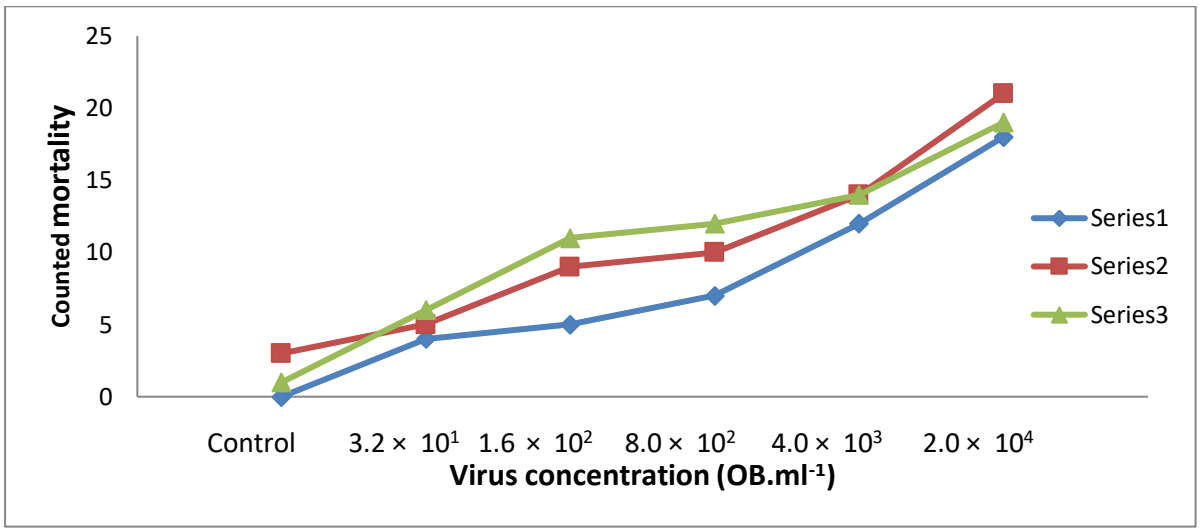


Figure 6.11: The dose response of Albany second instar *H. armigera* to Helicovir™ HearNPV showing percentage mortality versus concentration for three replicates namely Series 1, 2 and 3.

6.3.2.8 Helicovex® HearNPV bioassays on 2nd instar *H. armigera* larvae

Helicovex® virulence studies were performed against second instar Albany *H. armigera* larvae and the control mortality was recorded as the first data point (Figure 6.12). The concentrations used were between 3.2×10^1 and 2.0×10^4 OB.ml^{-1}.

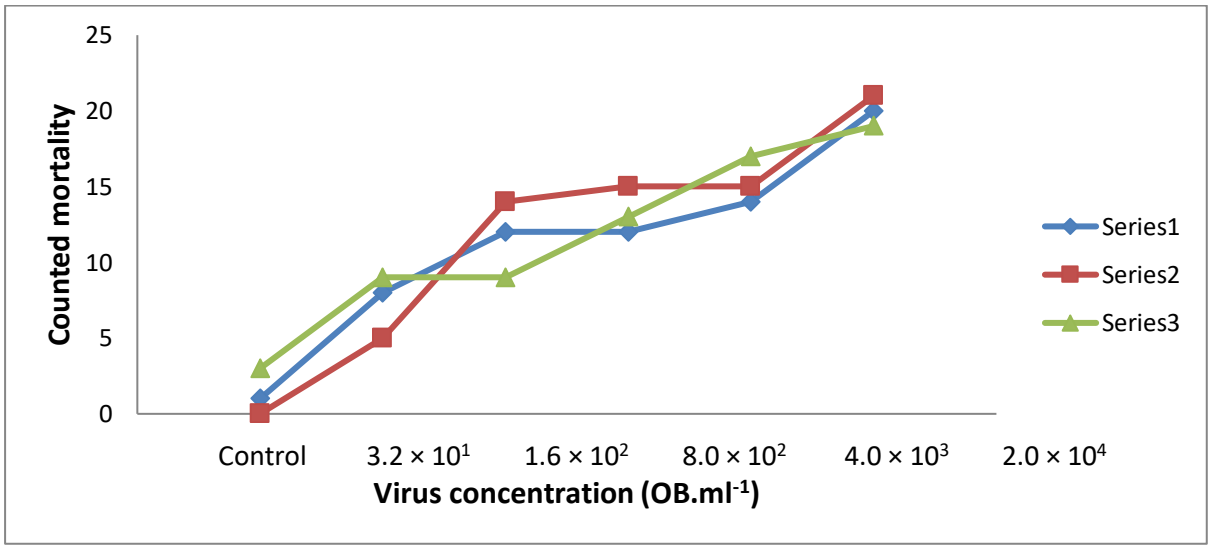


Figure 6.12: The counted mortality of Albany second instar *H. armigera* after treatment with different concentrations of Helicovex® for three replicates namely Series 1, 2 and 3. N = 25

6.3.3 Dose dependency analysis in the quantal response of second instar *H. armigera* to different HearNPV treatments

Table 6.2: Statistical evaluations of the quantal response of second instar Albany *H. armigera* to different HearNPV isolates using the probit analysis. Values in parenthesis represent the lower and upper 95% confidence limits according to Fieller's theorem. The relative potency is given as the dose-response potency at LC_{50} (LC_{50RP}) and LC_{90} (LC_{90RP}) in reference to HearNPV-Albany. The values highlighted in red are significantly different from the dose-response recorded for HearNPV-Albany after using the independent *t*-test, $p < 0.05$.

Data analysed	HearNPV-Albany	HearNPV-KZN	HearNPV-Whit	HearNPV-Haygrove	HearNPV-G4	HearNPV-SP1	Helicovir™	Helicovex®
LC_{20} ($OB.ml^{-1}$)	1.5×10^1 ($7.58 - 2.3 \times 10^1$)	7.1×10^1 ($4.7 \times 10^1 - 9.7 \times 10^1$)	2.0×10^1 ($8.67 – 3.6 \times 10^1$)	1.2×10^1 ($5.56 - 2.1 \times 10^1$)	1.3×10^1 ($6.6 – 2.1 \times 10^1$)	5.5×10^1 ($3.8 \times 10^1 – 7.5 \times 10^1$)	5.3×10^1 ($2.1 \times 10^1 – 9.9 \times 10^1$)	4.1×10^1 ($2.0 \times 10^1 - 7.0 \times 10^1$)
LC_{50} ($OB.ml^{-1}$)	7.7×10^1 ($5.5 \times 10^1 - 1.0 \times 10^2$)	3.2×10^2 ($2.5 \times 10^2 - 4.2 \times 10^2$)	2.8×10^2 ($1.8 \times 10^2 – 4.4 \times 10^2$)	9.9×10^1 ($6.8 \times 10^1 - 1.4 \times 10^2$)	1.2×10^2 ($8.4 \times 10^1 – 1.8 \times 10^2$)	2.0×10^2 ($1.5 \times 10^2 - 2.6 \times 10^2$)	9.0×10^2 ($6.0 \times 10^2 - 1.5 \times 10^3$)	5.84×10^2 ($3.7 \times 10^2 – 9.5 \times 10^2$)
LC_{50RP}	1.0	0.24	0.27	0.77	0.62	0.23	0.09	0.13
LC_{90} ($OB.ml^{-1}$)	9.7×10^2 ($6.3 \times 10^2 – 1.8 \times 10^3$)	3.1×10^3 ($2.0 \times 10^3 – 5.9 \times 10^3$)	1.5×10^4 ($7.1 \times 10^3 - 4.5 \times 10^4$)	2.3×10^3 ($1.3 \times 10^3 - 5.9 \times 10^3$)	3.9×10^3 ($2.1 \times 10^3 - 9.8 \times 10^3$)	1.4×10^3 ($9.7 \times 10^2 - 2.3 \times 10^3$)	-	-
LC_{90RP}	1.0	0.31	0.06	0.40	0.24	0.66	0.01	0.03
Slope	1.161	1.289	0.735	0.923	0.865	1.505	0.682	0.727
SE log LC_{50}	0.069	0.059	0.099	0.077	0.086	0.060	0.103	0.103
$p(Chi^2)$	0.644	0.164	0.108	0.641	0.457	0.810	0.877	0.131
p(F)	0.001	0.004	0.006	0.001	0.002	<0.001	<0.001	0.005

The quantal response of second instar Albany *H. armigera* larvae against HearNPV-Albany, HearNPV-KZN, HearNPV-Haygrove, HearNPV-Whit, HearNPV-SP1, HearNPV-G4, Helicovir™ and Helicovex® was assayed. The effective concentration needed for a mortality of 20%, 50% and 90% of the population over a period of 10 days was derived using the probit analysis in ToxRat professional v3.2.1 and the LC_{20}, LC_{50} and LC_{90} values were in turn calculated (Table 6.3).

The raw data for the mortality below 100% for each replicate was entered in the ToxRat application. For accuracy, data between 01% to 99.9% for intrapolation of effective levels through linear regression (ToxRat, 2015). For raw data that did not reach 90% mortality, the extrapolation limiter was removed from the ToxRat software allowing for extrapolation of data. This is a less accurate method for analyses but because of time limitations some of the experiments were not repeated and therefore on LC_{20} and LC_{50} were considered (Helicovir™ and Helicovex®). The probit analysis data were derived using linear maximum likelihood regression for the counted mortalities. From the counted mortality for the 3 replicates, the log function of the concentration was used, and this resulted in the deduction of the empirical and regular probit values.

6.3.4 Time response bioassays of second instar *H. armigera* larvae

Time response bioassays were performed for the HearNPV samples using LC_{90} concentrations and mortality was monitored every 12 hours. Mortalities that were recorded within 48hours of exposure to the doses were excluded from the results. The time response bioassays were run for a period that resulted in 100% mortality.

6.3.4.1 HearNPV-Albany

For the HearNPV-Albany time response bioassays a concentration of 1.5×10^3 $OBs.ml^{-1}$ was used. For all three replicates, 100% mortality of the test population was recorded after 156 hours (Figure 6.13)

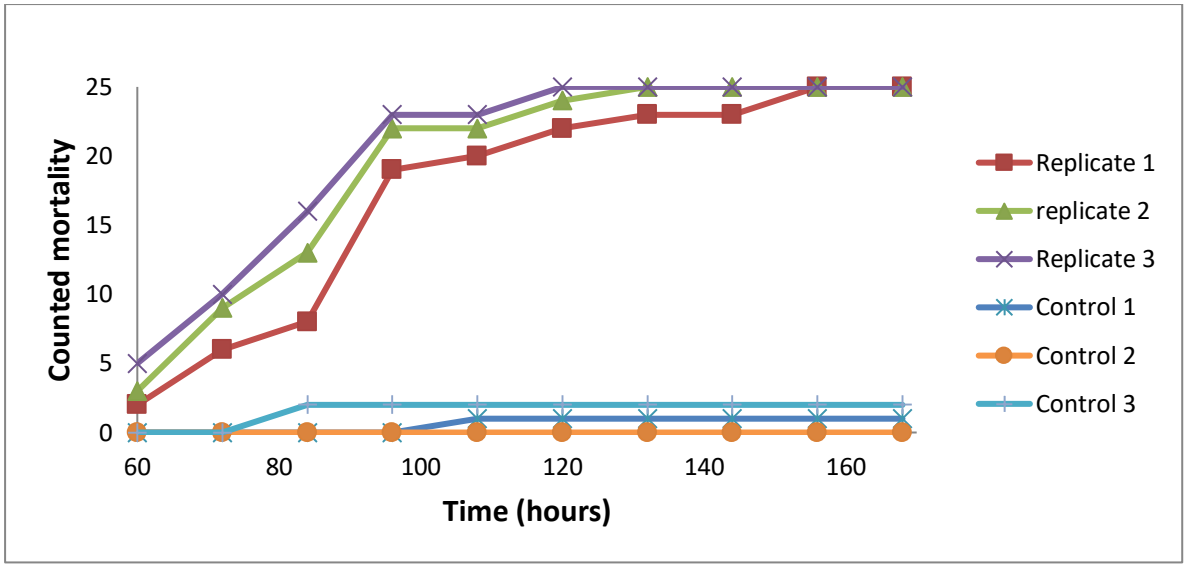


Figure 6.13: HearNPV-Albany mortality versus time bioassays measured every 12 hours for an LC_{90} concentration for three replicates and three controls. N=25

6.3.4.2 HearNPV-KZN

HearNPV-KZN at a concentration of 2.5×10^3 OBs.ml^{-1} was used to determine the lethal time studies. For all three replicates, 100% mortality of the test population was recorded after 180 hours (Figure 6.14).

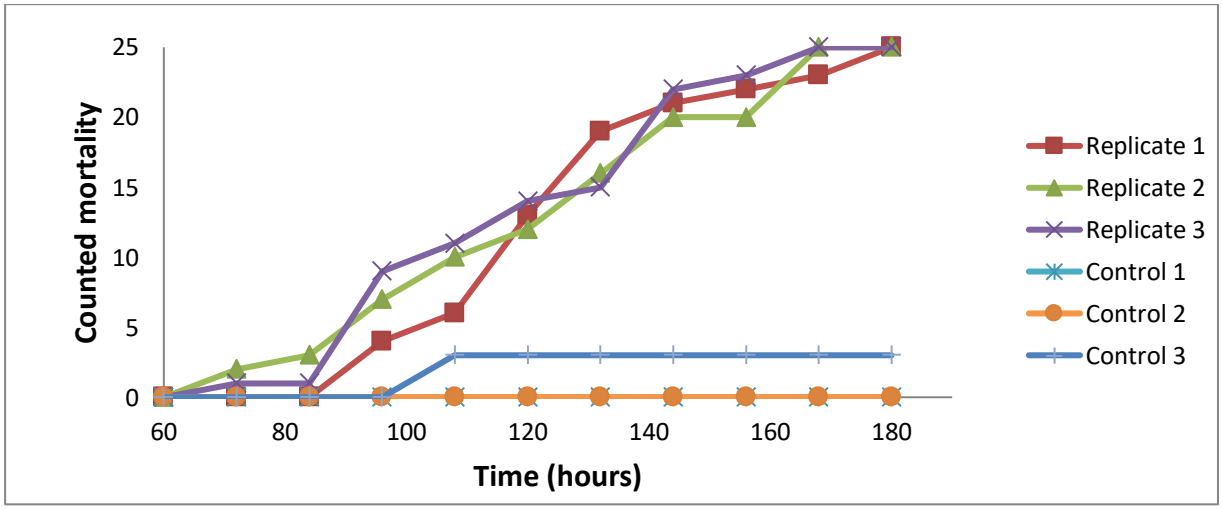


Figure 6.14: HearNPV-KZN mortality versus time studies performed on second instar Albany *H. armigera* for three replicates and three controls. N = 25

6.3.4.3 HearNPV-Haygrove

HearNPV-Haygrove at a concentration of 4.4×10^3 OBs.ml^{-1} was used to determine the lethal time studies. For all three replicates, 100% mortality of the test population was recorded after 216 hours (Figure 6.15).

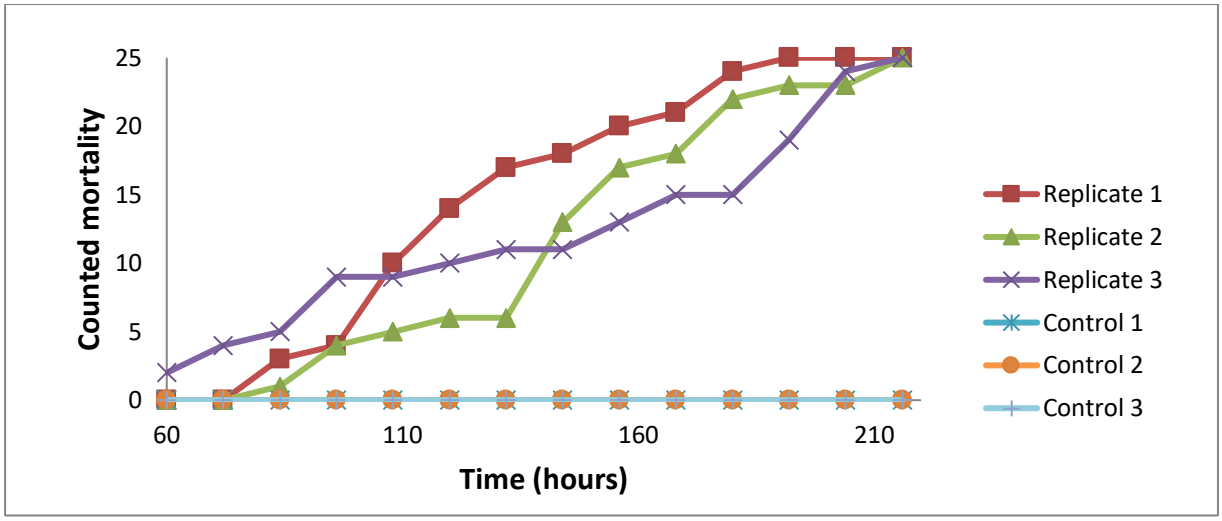


Figure 6.15: HearNPV-Haygrove mortality versus time studies performed on second instar Albany *H. armigera* for three replicates and three controls. N = 25

6.3.4.4 HearNPV-Whit

Time response bioassays were performed on second instar Albany *H. armigera* larvae using HearNPV-Whit at a concentration of 1.5×10^4 OBs.ml^{-1}. For all three replicates, 100% mortality of the test population was recorded after 216 hours (Figure 6.16).

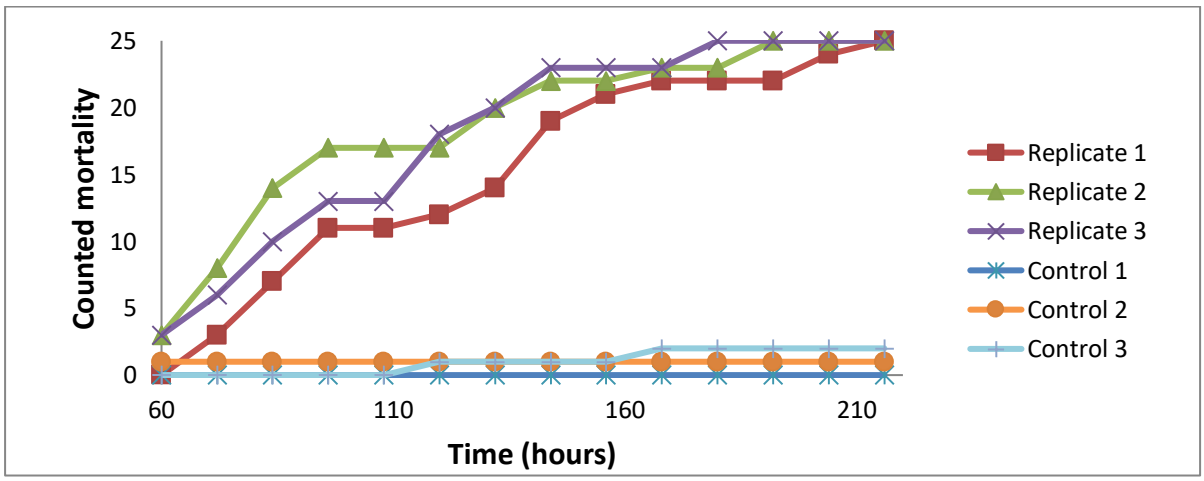


Figure 6.16: HearNPV-Whit mortality versus time-based studies on second instar *H. armigera* for three replicates and three controls. N = 25

6.3.4.5 HearNPV-G4

HearNPV-G4 at a concentration of 4.0×10^3 OBs.ml^{-1} were assayed against second instar Albany *H. armigera* and mortality was recorded periodically until 100% mortality was reached. For all three replicates, 100% mortality of the test population was recorded after 168 hours (Figure 6.17).

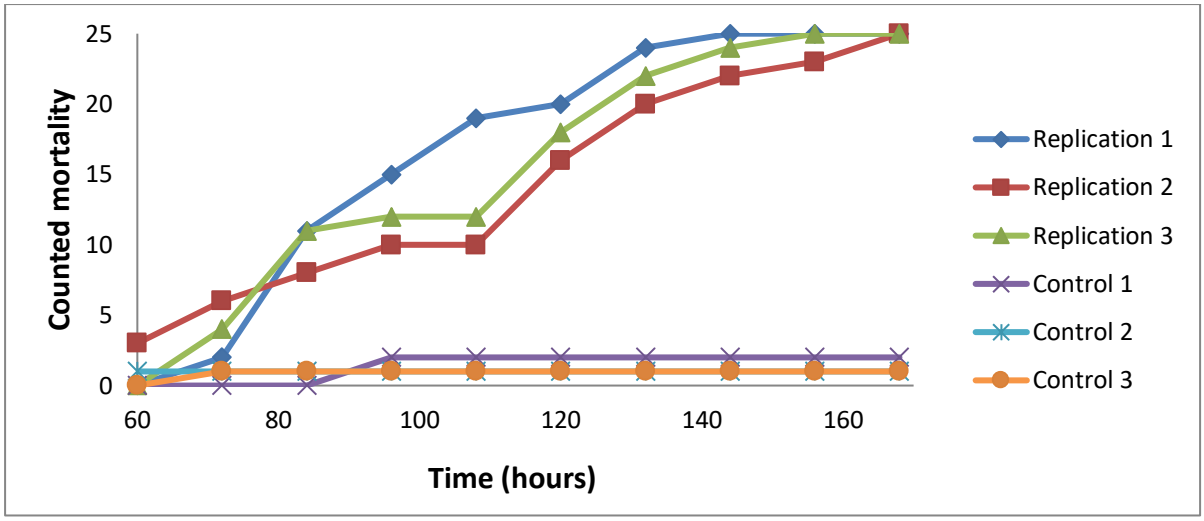


Figure 6.17: HearNPV-G4 for the three replicates for mortality versus time against second instar *H. armigera*. N = 25

6.3.4.6 HearNPV-SP1

The time response bioassay for HearNPV-SP1 against second instar Albany *H. armigera* were assayed using a concentration of 1.5×10^3 OBs.ml^{-1}. For all three replicates, 100% mortality of the test population was recorded after 192 hours (Figure 6.18).

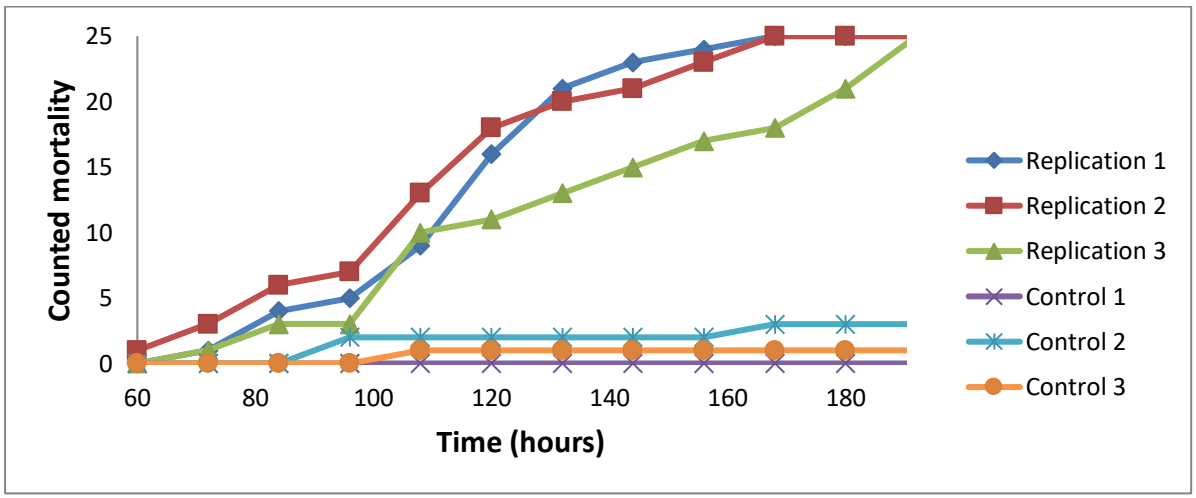


Figure 6.18: The mortality versus time response for SP1 HearNPV against second instar *H. armigera* for three replicates. N=25

6.3.4.7 Helicovir™

Time response mortality assays were performed against second instar *H. armigera* larvae using Helicovir™ at a concentration of 6.0×10^4 OBs.ml^{-1} and mortality was measured every 12 hours. For all three replicates, 100% mortality of the test population was recorded after 308 hours (Figure 6.19).

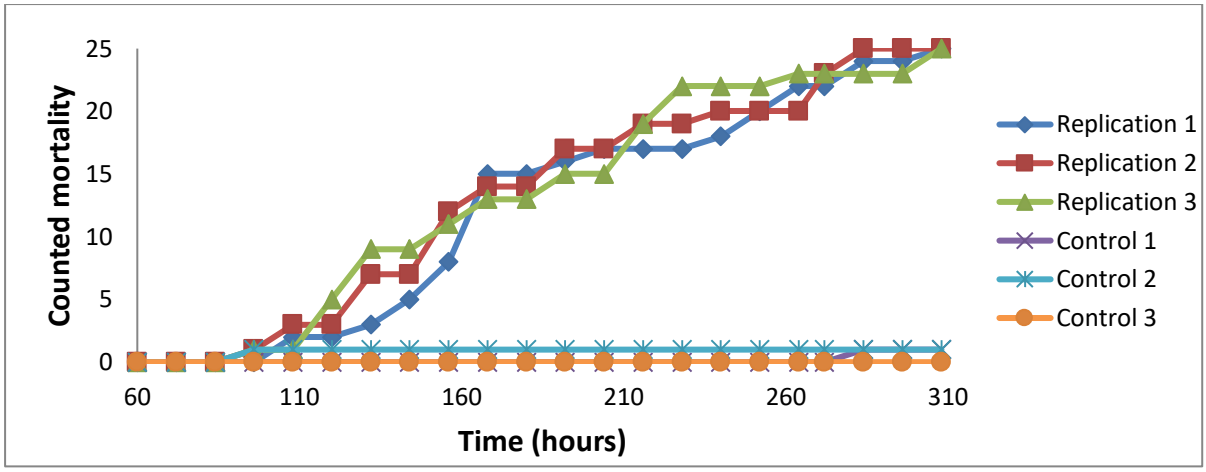


Figure 6.19: Helicovir™ mortality versus time studies against second instar Albany *H. armigera* for three replicates and three controls. N = 25

6.3.4.8 Helicovex®

Time response bioassays using Helicovex® at a concentration of 6.0×10^4 OBs.ml^{-1} were studied against second instar Albany *H. armigera*. For all three replicates, 100% mortality of the test population was recorded after 276 hours (Figure 6.20).

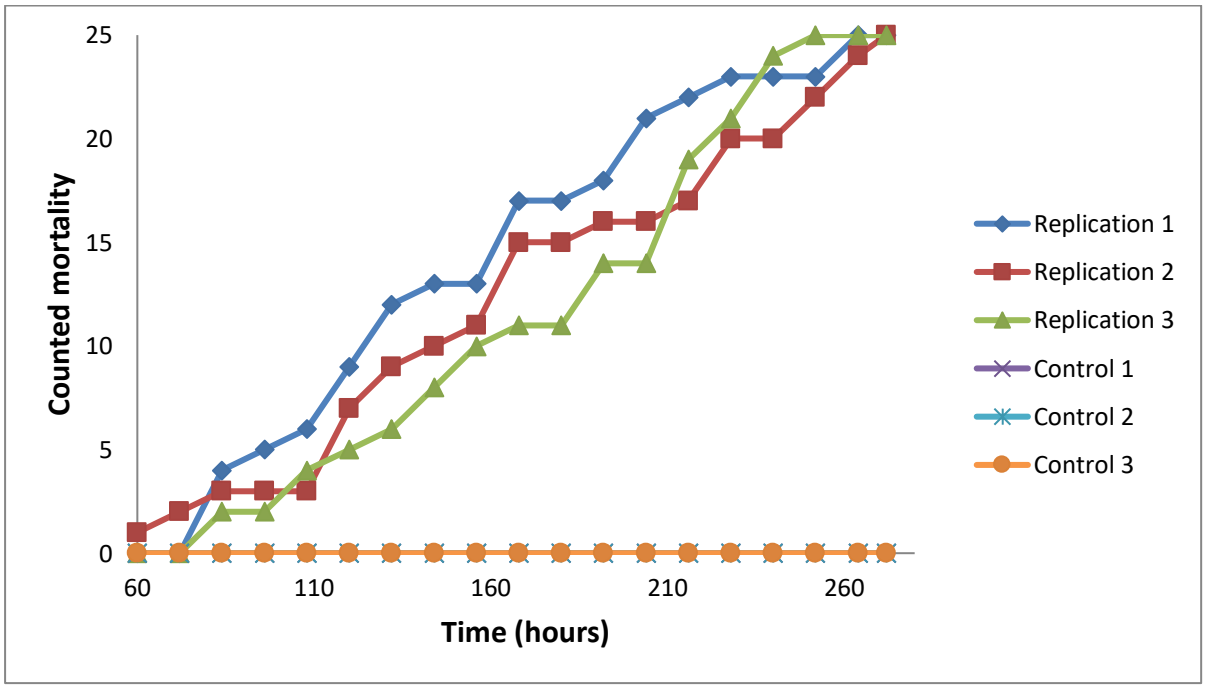


Figure 6.20: Time response bioassays showing percentage mortality over time using Helicovex HearNPV for three replicates. N = 25

6.3.5 Time response analyses for HearNPV isolates against second instar Albany *H. armigera*

Linear regression analyses were performed on the time-based study results using the logit analysis as a means to calculate LT_{50} and LT_{90}. The analyses were performed in ToxRat v3.2.1 and using the counted mortality over time and the Abbott's correction, data for the LT_{50} and LT_{90} was calculated and recorded (Table 6.4). The relative mean time course (time required to achieve 50% or 90% mortality of the assayed population in comparison to a reference) was calculated using HearNPV-Albany as the reference.

Table 6.3: The LT_{50} and LT_{90} values deduced using the logit analysis for different HearNPV isolates against second instar Albany *H. armigera*. The relative mean time course represents the time course required to achieve an LT_{50} or LT_{90} in reference to HearNPV-Albany. The values highlighted in red are significantly different from the LT_{50} or LT_{90} of HearNPV-Albany using the independent *t*-test ($p < 0.05$). N = 25.

Isolate	LT_{50} (95% cl) (Hours)	Relative mean time course LT_{50}	LT_{90} (95% cl) (Hours)	Relative mean time course LT_{90}
HearNPV-Albany	81.82 (79.09 –84.43)	1.00	108.44 (104.05 – 114.18)	1.00
HearNPV-KZN	119.65 (116.43 – 122.98)	1.46	158.22 (151.24 – 167.73)	1.46
HearNPV-Whit	99.15 (94.79 – 103.34)	1.21	168.53 (157.76 – 183.35)	1.55
HearNPV-Haygrove	130.89 (126.45 – 135.40)	1.60	208.76 (196.81 - 224.58)	1.93
HearNPV-SP1	117.35 (114.09 – 120.60)	1.43	160.00 (153.99 – 167.51)	1.47
HearNPV-G4	99.21 (94.11 – 104.04)	1.21	136.71 (127.72 – 147.21)	1.25
Helicovir™	176.78 (171.90 – 181.72)	2.16	271.65 (259.27 – 287.18)	2.51
Helicovex®	155.06 (148.22 – 161.87)	1.89	256.62 (239.75 – 279.53)	2.37

6.3.6 HearNPV-Albany, HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove dose-response assays

Further bioassays were performed on Haygrove *H. armigera* using South African HearNPV isolates. The study was performed to record the effects of geographically distinct local HearNPV isolates on different local *H. armigera* populations. HearNPV-Albany, HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove were assayed using the surface dose method to determine dose-response and time-response. Appropriate mortality responses against Haygrove *H. armigera* was recorded for concentrations of between 3.0 and 2.0×10^3 OB.ml^{-1} for the assayed HearNPV. These concentrations were then used for all the dose-response assays.

6.3.6.1 HearNPV-Albany dose-response

The dose-response of second instar *H. armigera* larvae treated with HearNPV-Albany is shown as counted mortality against virus concentration (Figure 6.21). The counted mortality was finalised after 10 days for the three replicates. The control mortality is represented as the first data point.

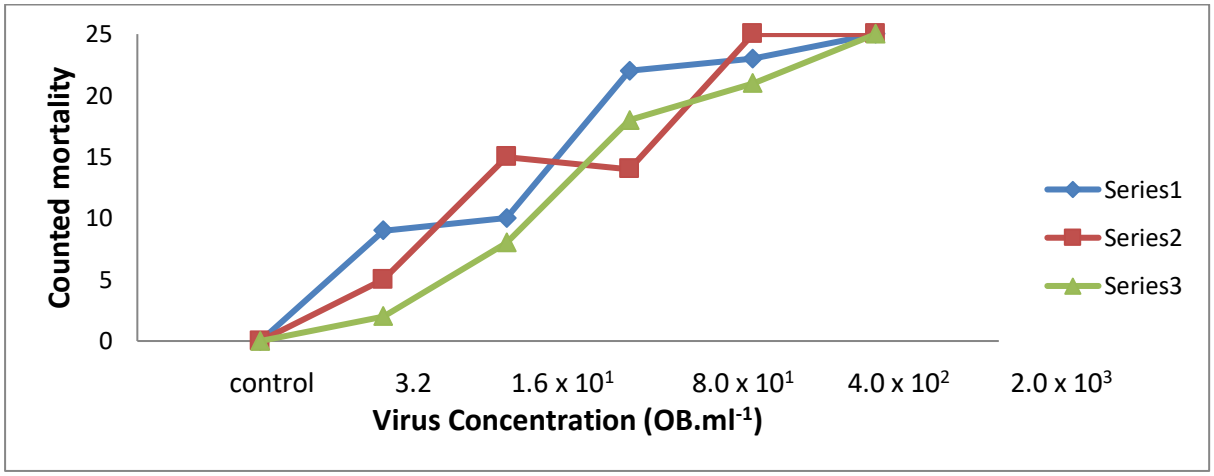


Figure 6.21: The dose-response of second instar *H. armigera* larvae when treated with HearNPV-Albany for three replicates namely Series 1, 2 and 3. N = 25

6.3.6.2 HearNPV-KZN dose-response

The dose-response of second instar *H. armigera* larvae treated with HearNPV-KZN is shown as counted mortality against virus concentration (Figure 6.22). The counted mortality was finalised after 10 days for the three replicates. The control mortality is represented as the first data point.

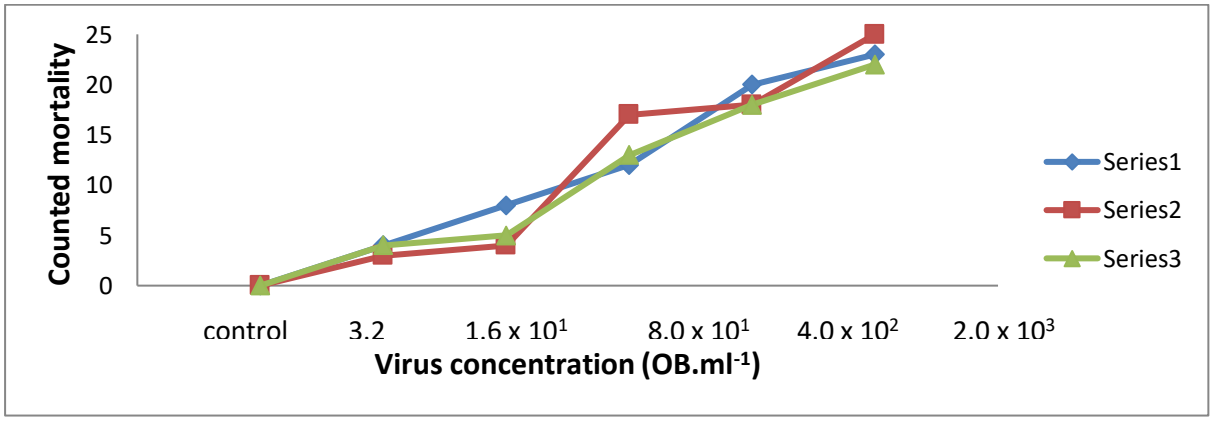


Figure 6.22: The dose-response of second instar *H. armigera* larvae when treated with HearNPV-KZN for three replicates namely Series 1, 2 and 3. N = 25

6.3.6.3 HearNPV-Whit dose-response

The dose-response of second instar *H. armigera* larvae treated with HearNPV-Whit is shown as counted mortality against virus concentration and the control mortality is represented as the first (Figure 6.23). The counted mortality was finalised after 10 days for the 3 replicates. The control mortality is represented as the first data point.

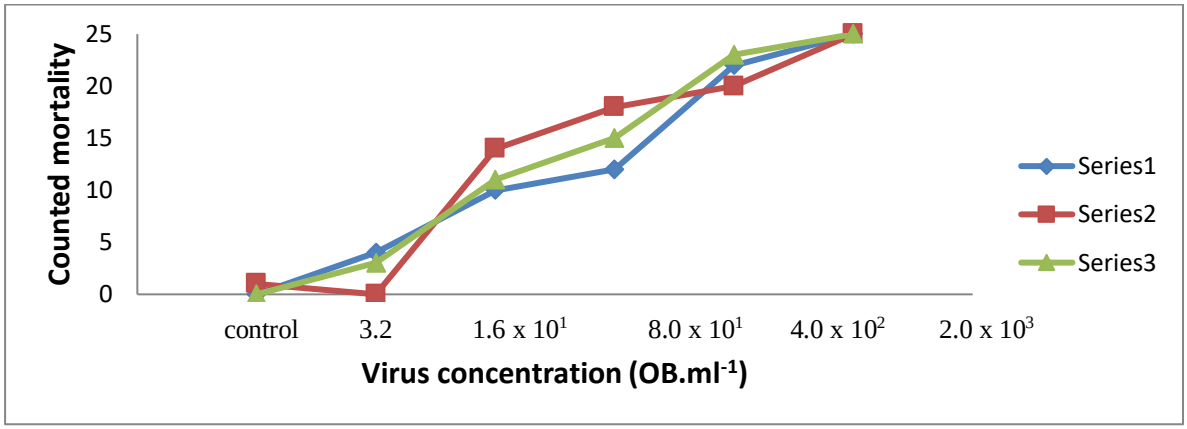


Figure 6.23: The dose-response of second instar *H. armigera* larvae when treated with HearNPV-Whit for three replicates namely Series 1, 2 and 3. N = 25

6.3.6.4 HearNPV-Haygrove dose-response

The dose-response of second instar *H. armigera* larvae treated with HearNPV-Haygrove is shown as counted mortality against virus concentration (Figure 6.24). The counted mortality was finalised after 10 days for the three replicates. The control mortality is represented as the first data point.

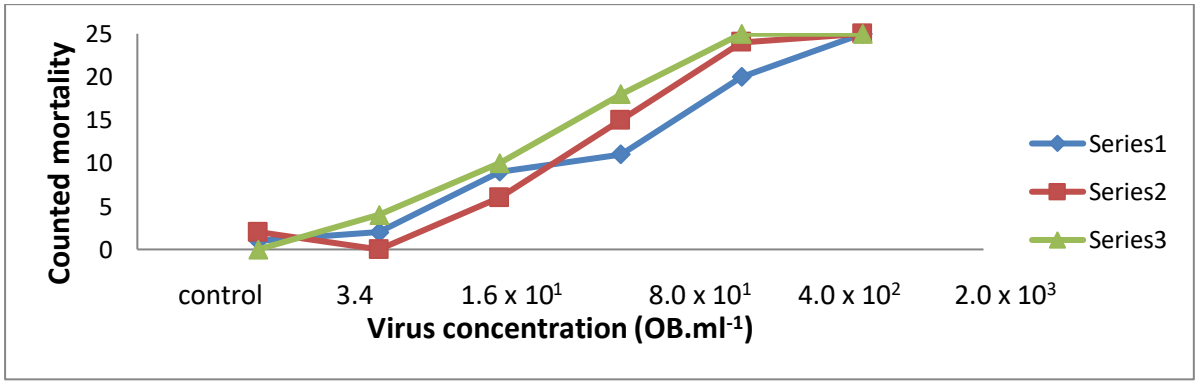


Figure 6.24: The dose-response of second instar *H. armigera* larvae when treated with HearNPV-Haygrove for three replicates namely Series 1, 2 and 3. N = 25

The probit analysis was carried out to determine the dose-response required for a mortality of 20%, 50% and 90% of the Haygrove *H. armigera* population when treated with HearNPV-Albany, HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove (Table 6.5). The data used

for the probit analysis were the logarithm of the concentration and the number of replicates, which resulted in the generation of the empirical probit and regular probit which were used to generate probit values. For all the HearNPV isolates tested, the data were scattering around the computed dose response as can be observed by the low p(Chi^2) value. Values that resulted in 100% mortality were automatically rejected from the analyses.

Table 6.4: Statistical evaluations of the quantal response of second instar Haygrove *H. armigera* to South African HearNPV isolates using the probit analysis. Values in parenthesis represent the lower and upper 95% confidence limits according to Fieller's theorem. The relative potency is given as the dose-response potency at LC_{50} and LC_{90} in reference to HearNPV-Albany. The values highlighted in red are considered to be significantly different from the dose-response recorded for HearNPV-Albany, using the independent *t*-test ($p < 0.05$).

Isolates	HearNPV-Albany	HearNPV-KZN	HearNPV-Whit	HearNPV-Haygrove
LC_{20} (OB.ml^{-1})	3.08 (1.68 - 4.83)	8.43 (4.59 – 1.3×10^1)	6.81 (1.0 – 1.8×10^1)	1.2×10^1 ($8.3 - 1.7 \times 10^1$)
LC_{50} (OB.ml^{-1})	1.6×10^1 (1.1×10^1 - 2.1×10^1)	6.6×10^1 (4.6×10^1 - 9.4×10^1)	3.5×10^1 (1.2×10^1 -9.2×10^1)	4.7×10^1 (3.5×10^1 - 6.1×10^1)
LC_{50RP}	1.00	0.24	0.43	0.33
LC_{90} (OB.ml^{-1})	1.8×10^2 (1.2×10^2 - 3.2×10^2)	1.5×10^3 (8.5×10^2 - 3.3×10^3)	4.3×10^2 (1.5×10^2 - 5.3×10^3)	3.5×10^2 (2.4×10^2 - 5.7×10^2)
LC_{90RP}	1.00	0.12	0.41	0.52
Slope	1.200	0.943	1.177	1.460
SE log LC_{50}	0.0733	0.0792	0.0690	0.0597
p(Chi^2)	0.125	0.490	0.026	0.326
p(F)	0.005	0.001	0.007	0.002

6.3.7 HearNPV-Albany, HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove time-response bioassays

6.3.7.1 HearNPV-Albany

The mortality of *H. armigera* larvae when treated with HearNPV-Albany at a concentration of 9.5×10^2 OB.ml^{-1} was recorded after 60 hours until 100% mortality was recorded (Figure 6.25).

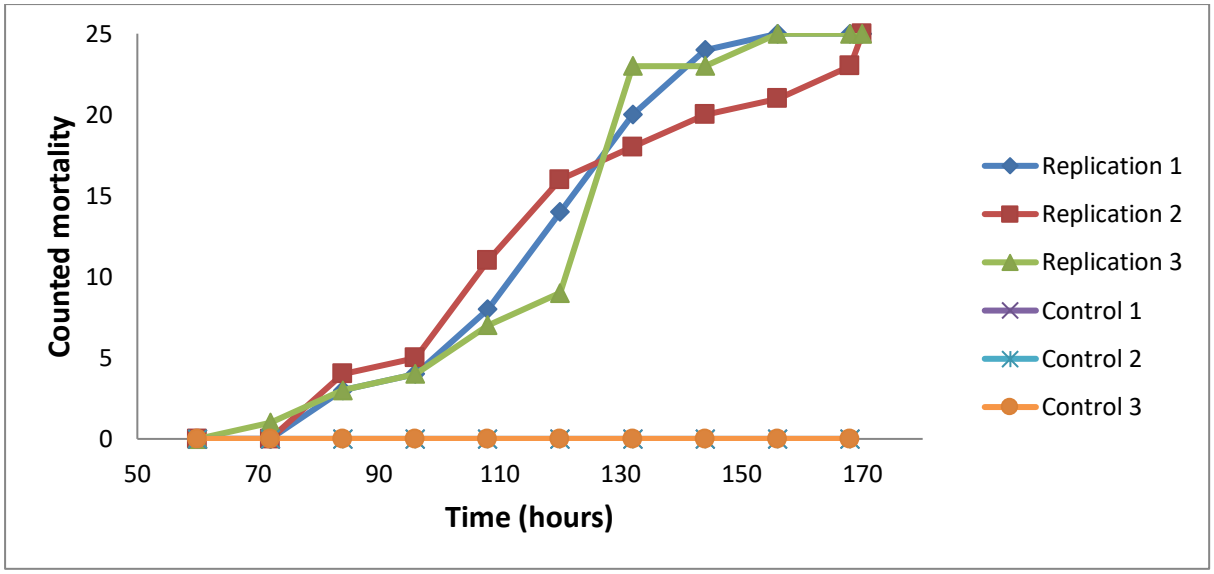


Figure 6.25: The time-response of *H. armigera* larvae after treatment with HearNPV-Albany.

6.3.7.2 HearNPV-KZN

The mortality of *H. armigera* larvae when treated with HearNPV-KZN at a concentration of 1.5×10^3 OB.ml^{-1} was recorded after a period of 60 hours post-infection at 12-hour intervals until 100% mortality was reached (Figure 6.26).

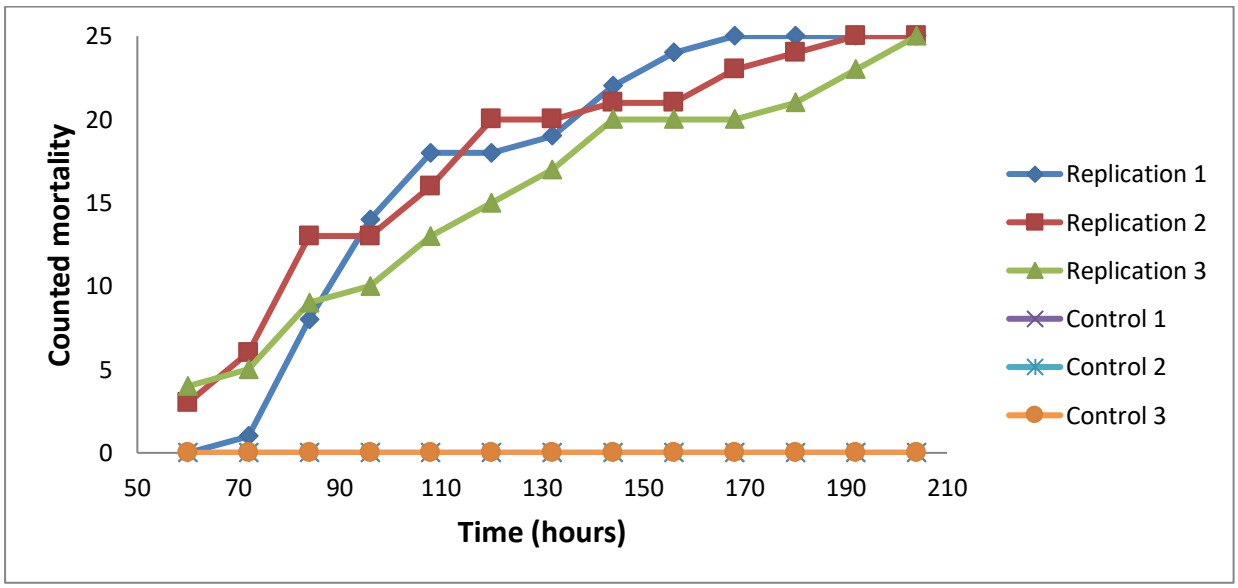


Figure 6.26: The time-response of *H. armigera* larvae after treatment with HearNPV-KZN.

6.3.7.3 HearNPV-Whit

The mortality of *H. armigera* larvae when treated with HearNPV-Whit at a concentration of 1.0×10^3 OB.ml^{-1} was recorded after a period of 60 hours post-infection at 12-hour intervals until 100% mortality was reached (Figure 6.27).

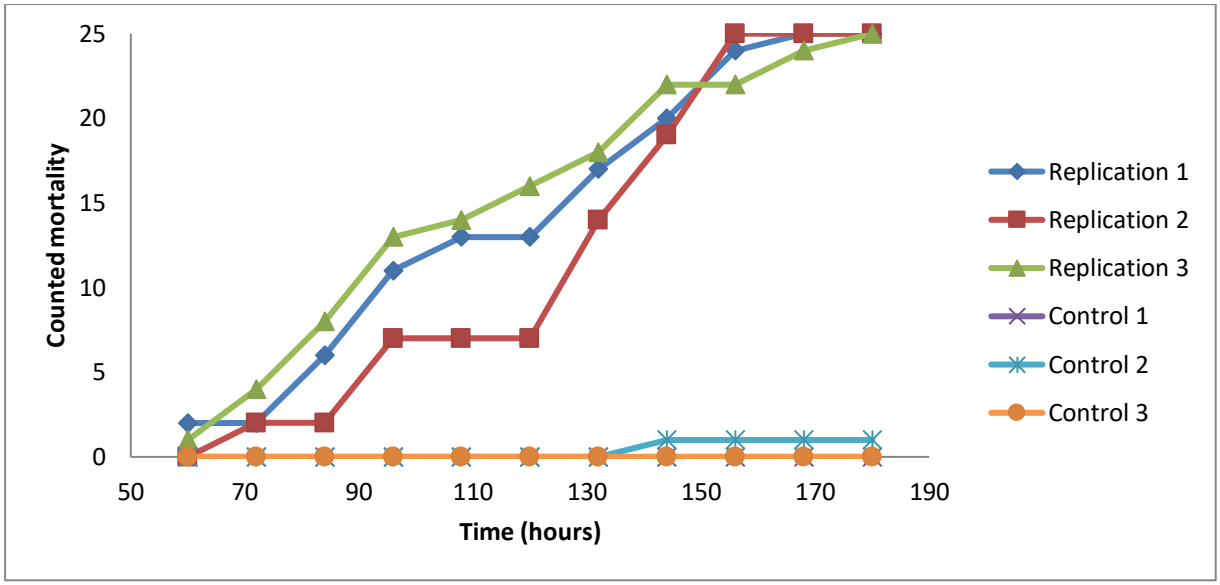


Figure 6.27: The time-response of *H. armigera* larvae after treatment with HearNPV-Whit.

6.3.7.4 HearNPV-Haygrove

The mortality of *H. armigera* larvae when treated with HearNPV-Haygrove at a concentration of 1.2×10^3 OB.ml^{-1} was recorded after a period of 60 hours post-infection at 12-hour intervals until 100% mortality was reached (Figure 6.28).

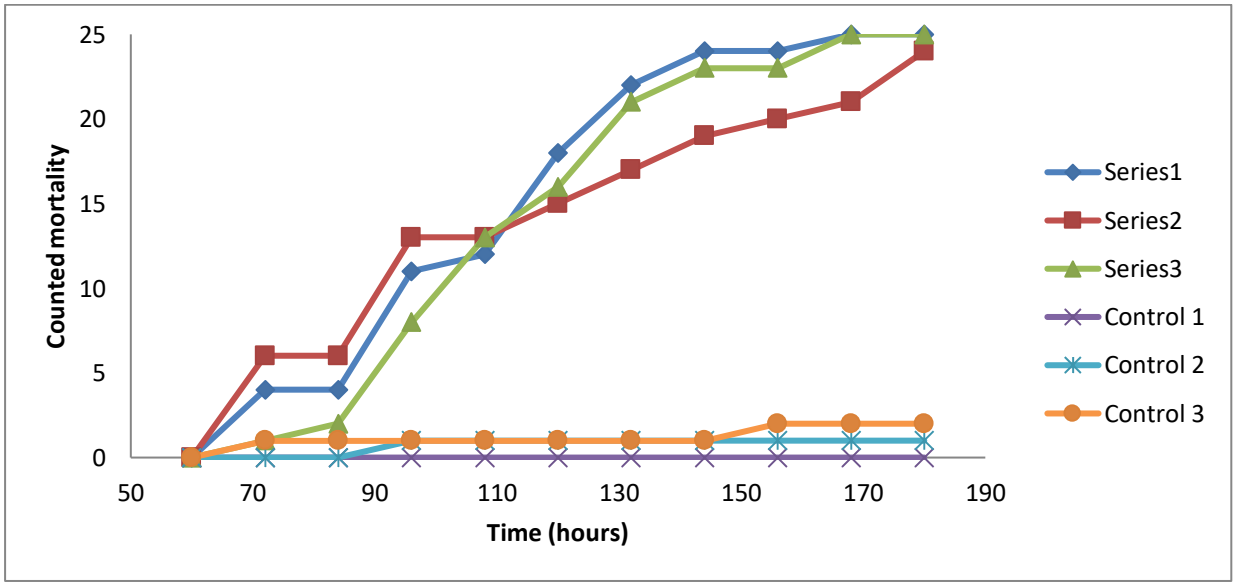


Figure 6.28: The time-response of *H. armigera* larvae when treated with HearNPV-Haygrove.

Using the counted mortality over time and the Abbott's correction, data for the LT_{50} and LT_{90} was calculated and recorded (Table 6.6). The relative mean time course (time required to

achieve 50% or 90% mortality of the assayed population in comparison to a reference) was calculated using HearNPV-Albany as the reference.

Table 6.5: The quantal response the time-response bioassays of South African HearNPV isolates on Haygrove *H. armigera* larvae. The relative mean time course represents the time course required to achieve an LT_{50} or LT_{90} in reference to HearNPV-Albany. The values highlighted in red are significantly different from the LT_{50} or LT_{90} of HearNPV-Albany using the independent *t*-test ($p < 0.05$). N = 25

Isolate	LT_{50} (95% cl) (Hours)	Relative mean time course LT_{50}	LT_{90} (95% cl) (Hours)	Relative mean time course LT_{90}
HearNPV-Albany	113.7 (110.1 - 117.2)	1.00	147.4 (141.6 - 154.9)	1.00
HearNPV-KZN	98.2 (94.8 - 101.5)	0.86	159.1 (152.6 - 167.1)	1.08
HearNPV-Whit	108.1 (101.2 - 114.8)	0.95	157.7 (146.5 – 176.6)	1.07
HearNPV-Haygrove	104.3 (100.0 – 108.5)	0.91	145.4 (138.1 – 155.2)	1.01

6.3.8 PCR amplification of *polh* from bioassay disease cadavers

From the larval cadavers that had succumbed to HearNPV treatments, genomic DNA was extracted and the *polh* was amplified to establish presence of HearNPV in the bioassay cadavers (Figure 6.29).

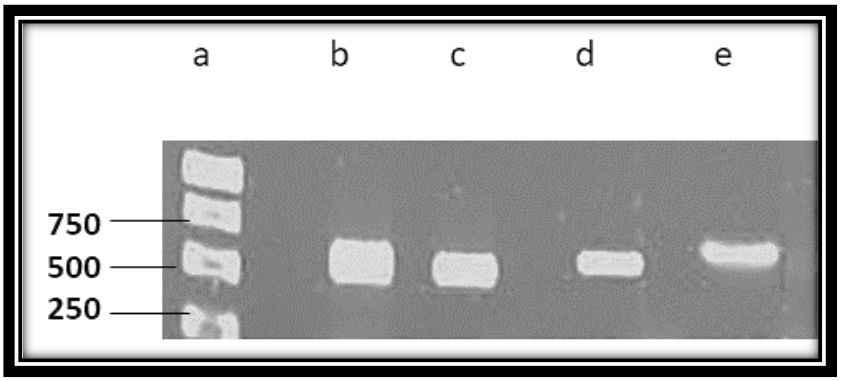


Figure 6.29: PCR amplification of the *polh* of HearNPV purified from bioassay larvae. A: 1kb GeneRuler DNA ladder. B: HearNPV-Albany PCR product. C: HearNPV-KZN PCR product. D: HearNPV-Whit PCR product. E: HearNPV-Haygrove PCR product.

6.4 Discussion

In this chapter, the virulence of geographically distinct HearNPV isolates was assayed against two South African collected *H. armigera* larval populations. Quantification of the HearNPV isolates was performed through OB enumeration in a counting chamber. This method was utilised because of the paracrystalline structure of the OBs, which makes them appear as bright refractive particles under phase-contrast light microscopy (Grzywacz *et al*., 2007).

After successful quantification of the viruses, dose-response and time-response studies were performed to determine the concentrations required for mortality as well as the speed of kill.

Surface dose bioassays to analyse dose-response of eight HearNPV isolates namely HearNPV-Albany, HearNPV-KZN, HearNPV-Whit, HearNPV-Haygrove, HearNPV-SP1, HearNPV-G4, Helicovir™ and Helicovex® were performed against second instar Albany *H. armigera*. The data points for the assays had a good fit to the regression line used for analysis, with a high coefficient of determination for all the assays. The statistical analyses performed confirmed the dose-response to be significantly different from the control mortality. For the calculation of dose-response, intrapolation using the mortalities of the studied population is required, hence the need for more than 90% mortality for the deduction of LC_{90}. However, Helicovir™ and Helicovex® dose response studies did not result in mortalities above 90% and using ToxRat, the extrapolation factor was used and the LC_{90} values of Helicovir™ and Helicovex® were calculated (ToxRat, 2015). The 95% Fiducial confidence limits for Helicovir™ and Helicovex® deviated greatly from the LC_{90} value, highlighting the minimal efficiency of calculations through extrapolation. It would be important to get more accurate results by repeating the experiments with doses that result in more than a 90% mortality of the studied population.

For comparison of the dose-response of the HearNPV isolates against each other, HearNPV-Albany was chosen as the reference isolate with the other seven being compared as the test isolates. The reason for comparing the other isolates to HearNPV-Albany was based on the premise that the virus isolate had been identified and purified from the local host population, Albany *H. armigera*. In Chapter 5 it was reported that HearNPV-Albany, HearNPV-KZN, HearNPV-SP1 and HearNPV-G4 were genotypic variants, and it was also hypothesised that HearNPV-Whit and HearNPV-Haygrove were also variants. Genotypic variants of HearNPV have been reported to have different biological activity against host populations isolated from the same and different geographical regions (Ogembo *et al.*, 2007; Rowley *et al.*, 2011; Arrizubieta *et al.*, 2013). It was therefore hypothesised that testing the different HearNPV isolates against local host populations of *H. armigera* should result in different biological activity. A comparison of the LC_{50} and LC_{90} values against those of HearNPV-Albany resulted in determination of relative potency. The relative potency is a representation of how potent the test preparation is, in respect to the reference (Council of Europe, 2013). The dose-response assays showed that the activity differed significantly between some of the isolates and HearNPV-Albany. HearNPV-Albany expressed the highest activity against Albany *H.*

armigera for both LC_{50} and LC_{90}. In comparison to the LC_{50}, only HearNPV-Haygrove and HearNPV-G4 had a relative potency that was above 60% of the activity recorded for HearNPV-Albany; whereas the rest of the isolates had the activity ranging from 9%-27%. For the LC_{90} relative potency in comparison to HearNPV-Albany, only HearNPV-SP1 and HearNPV-Haygrove were not significantly different ($p > 0.05$), with the rest of the isolates having their potency ranging from 1%-31%. The difference in the potency could be potentially explained on the genomic level. In Chapter 5 it was reported that the isolates had highly variable *hr* regions. The *hr* regions function as regions for putative origin for DNA replication (Erlandson, 2009). The *hr* regions have been reported to be potentially responsible for the difference in virulence for HearNPV-C1 and HearNPV-G4 which were isolated from the same region, with high nucleotide sequence similarity but differing activity (Zhang *et al*., 2005). The two commercial formulations, Helicovir™ and Helicovex® had the least potency. This result is similar to what was reported by Rowley *et al.*, (2011) where they stated that Gemstar (a commercial formulation) had a significantly higher LC_{50}. HearNPV-SP1 against second instar *H. armigera* collected from UK, using the droplet feeding method, had an LC_{50} of 3.6×10^4 OB.ml^{-1} and HearNPV-G4 had an LC_{50} value of 1.5×10^4 OBs.ml^{-1} (Arrizubieta et al., 2013; Arrizubieta *et al*., 2015), which is lower than the 2.0×10^2 OB.ml^{-1} and 1.24 OB.ml^{-1} recorded against second instar Albany *H. armigera* respectively. The ingredients used in the formulations of commercial biopesticides may affect the potency of the virus (Behle and Birthisel, 2014). The lower activity recorded for Helicovir™ and Helicovex® may be a result of the formulation. Another possibility for the low virulence recorded for the commercial biopesticides may be a result of the length of the biopesticides having been stored since the manufacture date. The effectiveness of the biopesticides has been recorded to be affected by storage factors and the age of the isolated virus (Claus *et al*., 2012). The low LC_{50} values recorded against Albany *H. armigera* for HearNPV-SP1 and HearNPV-G4 may represent the different insecticidal activity exhibited by the same isolates against host populations from geographically distinct populations. The other potential reason for the low values may be that on artificial diet lower concentrations have been reported to be required for mortality (Richards and Christian, 1999; Gupta *et al*., 2007). This may be a result of the low OB ingestion during the droplet method, where it has been recorded that third instar larvae ingest on average 1 µl of the HearNPV compared to surface dose whereby ingestion of the OB will be magnitudes higher because of quantities used to dose the diet surface and the continued feeding on infested diet surface (Georgievska *et al*., 2010). The control of *H. armigera in* the field will be dependent on the use of virulent strains of HearNPV, with

HearNPV-Albany being shown to be more virulent than the commercial formulations, and HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove proving to be comparably virulent to other geographically distinct strains and better than Helicovir™ and Helicovex®.

For the time-response studies, HearNPV-Albany was used as the reference for comparison with the other seven isolates to observe relative mean time course. For both the LT_{50} and LT_{90} times a significant difference was recorded for all the HearNPV isolates in comparison to HearNPV-Albany. HearNPV-KZN, HearNPV-Haygrove and HearNPV-SP1 the mean time was 1.5 times more than that of HearNPV-Albany. The *ie-1* gene and the gp37 gene which was absent in HearNPV-KZN, are involved in the primary infection cycle (Herniou *et al.,* 2003). The *ie-1* gene encodes for a protein that is associated with DNA replication and the *gp37* gene is responsible for a protein that may be involved in oral infection (Liu *et al.,* 2006; Berretta *et al.,* 2013). The *cathepsin-like protein* gene from HearNPV-KZN had a 50.45% similarity to the same gene in the other genome comparison in Chapter 5. This gene is involved in the late-infection cycle, responsible for the protein that is responsible for host cuticle disruption (O'Reilly, 1997). These genes are directly involved in the infection cycle and therefore may affect speed of kill or lethal dose. The least speed of kill was recorded for the two commercial products, Helicovir™ and Helicovex® with Helicovir™ having a speed of kill of up to 2.2 times slower than that of HearNPV-Albany. Similar results were recorded for the mean time course at LT_{90} with Helicovir™ having a speed of kill up to 2.5 times slower. The LT_{50} values in this study for HearNPV-SP1 and HearNPV-G4 were like literature reports of 102.8 hours compared to 117.4 hours for this study (Arrizubieta *et al.,* 2015). Against second instar *H. armigera* larvae, LT_{50} values of between 108 to 138 hours were recorded for HearNPV-SP isolates, and 126 hours for HearNPV-G4 (Figueiredo *et al*., 2009).

The higher speed of kill exhibited by the HearNPV isolates especially HearNPV-Albany in comparison to the commercial formulations, coupled with the dose-response studies, makes the use of these isolates in field control more appealing.

Bioassays of South African HearNPV isolates against second instar Haygrove *H. armigera* were studied. For HearNPV-Albany, HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove, virulence was recorded to be greater against the Haygrove colony compared to the Albany colony. Using HearNPV-Albany as the reference isolate, at LC_{50} the lowest relative potency was recorded for HearNPV-KZN which had a relative potency of 24% which was similar to what was recorded against Albany *H. armigera* and HearNPV-Haygrove had a

relative potency of 33%, which was lower to what was recorded against Albany *H. armigera*. At LC_{90} the relative potency of HearNPV-KZN was the lowest at 12%, with the other isolates having between 40% and 52%. For the LT_{90} no significant difference was recorded. The South African HearNPV isolates had similar lethal times against Haygrove *H. armigera*. Ogembo *et al.*, (2007) assayed African isolated HearNPV against new third instar larvae that had been starved for 16 hours. HearNPV-NNg1 resulted in an LC_{50} of 1.0×10^5 OB.ml^{-1}, HearNPV-Whit had an LC_{50} of 3.5×10^4 OB.ml^{-1} and HearNPV-NZ3 from Zimbabwe had an LC_{50} of 1.46×10^6 OB.ml^{-1}. For the time-response studies an LT_{50} for HearNPV-NNg1 was 4.0 days, for HearNPV-Whit was 5.1 days and for HearNPV-NZ3 was 4.8 days. The LT_{50} recorded for HearNPV-Whit is similar to what was recorded and reported in this chapter. The time-response assays show that for geographically similar isolates, although the concentrations required for mortality may differ, against the Haygrove colony, the time to death is not significantly different ($p > 0.05$). This makes the use of either isolate at the correct concentrations appealing for *H. armigera* control in the field, in an effort to minimise continued feeding damage post infection. The difference in virulence recorded for the South African HearNPV isolates against the two *H. armigera* population raise the possibility of different formulations and combinations that can be used in manufacturing biopesticides. A combination of HearNPV isolates that cover different geographical populations are important if HearNPV products are to be used in different regions. A mixture of the HearNPV may result in improved virulence against more geographically distributed host populations and also potentially decrease the potential of adaptation from the host. Mixtures of different genotype SeNPV against *Spodoptera exigua* exhibited different virulence in comparison to the individual NPV, possibly representing synergism (Virto *et al.*, 2017). This result supports the potential of using mixtures of the HearNPV and increasing the potential of the biopesticide in formulations when they are used in different geographical locations.

The PCR amplification of *polh* identified the presence of NPVs in the diseased larval cadavers from the bioassays. This confirmed the presence of NPVs in the larval cadavers collected from the bioassays.

6.5 Conclusions

In this chapter, virulence studies of eight geographically distinct HearNPV isolates were successfully tested against South African collected and laboratory reared second instar *H. armigera* larvae. The efficacy of the different viruses was deduced from analysing the concentration-responses and time evaluations. The geographically distinct HearNPV isolates

resulted in different dose-responses and time-responses against local South African *H. armigera* populations. The lowest virulence was recorded for the commercial formulations and the best performing isolates were HearNPV-Albany, HearNPV-Haygrove and HearNPV-SP1. For the speed of kill it was recorded that the South African isolated HearNPV had comparable times with HearNPV-Albany being the most effective against the Albany collected *H. armigera*. Against Haygrove *H. armigera*, it was recorded that the South African HearNPV-isolates were less virulent as well in comparison to Albany *H. armigera*. The results supported the hypothesis that the homologous virus is the most virulent.

7 General discussion and conclusion

7.1 Introduction

The overall aim of this research was to fulfil several objectives leading to a better understanding of the biology, genetics and virulence of novel Helicoverpa armigera NPV isolates for potential application as biopesticides against *H. armigera* in agricultural ecosystems. To achieve the aim of the research, several objectives were set as described in Chapters 2-6 of the thesis. The first objective, described in Chapter 2, was to establish colonies of two South African *H. armigera* populations from Grahamstown (Albany colony) and George (Haygrove colony) under laboratory conditions. The second objective, described in Chapter 3, was to isolate HearNPV OBs from South African collected diseased cadavers, crude extracts from other geographically distinct locations and commercial HearNPV products. The third objective, described in Chapter 4, was to perform molecular studies on the purified HearNPV isolates from South Africa through full genome sequencing and annotation of the genomes. The fourth objective was to identify differences between South African HearNPV isolates and geographically separate HearNPV isolates on a molecular level and this was described in Chapter 5. The final objective was to perform bioassays on South African collected and mass reared *H. armigera* colonies. Bioassays were performed against Albany colony using HearNPV-Albany, HearNPV-KZN, HearNPV-Whit, HearNPV-Haygrove, HearNPV-G4, HearNPV-SP1, Helicovir™ and Helicovex®. Bioassays against the Haygrove colony were performed using HearNPV-Albany, HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove. The bioassays enabled the comparison of geographically distinct HearNPV isolates against South African HearNPV isolates as well as the response of the two different South African collected *H. armigera* populations against different South African HearNPV isolates. The discussion below is broadly structured around the aims as set out in Chapter 1.

Helicoverpa armigera control is important because of the pest status as a phytophagous insect of global significance, which impacts on agricultural industries of many countries resulting in significant crop and economic losses (Vaissayre and Caiquil, 2000; Kranthi and Russell, 2009). The pest has been traditionally controlled using broad-spectrum chemical pesticides (Oluoch-Kasura *et al.*, 2013). However, *H. armigera* has a great capacity to develop resistance and high levels of resistance have been reported in the field (Zhuang and Gill, 2005; Xu *et al.*, 2005). To continue effective control of *H. armigera*, IPM was adopted

as an approach; the management system includes the use of biological control measures amongst other methods (Mitchell and Hutchison, 2009). Insect viruses of *H. armigera* have been reported to be a control measure, as standalone application or as part of an IPM strategy, and NPV has been successfully commercialised in global markets (Kalawate, 2014; Ibarra and Rincón-Castro, 2009; Moore and Kirkman, 2010; Paul, 2004; Haase *et al*., 2015; Noune and Hauxwell, 2016). The NPV for *H. armigera* has been studied as a control method for the pest with various commercial biopesticides having been formulated from the virus including Helicovir™ and Helicovex® (Moore *et al*., 2004; Kessler *et al*., 2008). The research carried out in this thesis aimed to isolate novel HearNPV isolates that could be utilised in the control of *H. armigera*. The introduction of novel HearNPV isolates would be useful in minimising potential pesticide resistance from the target pest and would be part of an IPM strategy. Different baculovirus isolates could show differences in virulence and infectivity against geographically different host populations (Georgievska *et al.,* 2010; Arrizubieta *et al.,* 2015; Opoku-Debrah *et al.,* 2016). Genetic characterisation aids in the identification of variation between isolates, and this may translate into phenotypic traits that result in differences in virulence against host populations in some cases (Chen *et al*., 2000; Baillie and Bouwer, 2013). Biological activity and characterisation of novel HearNPV isolates makes it possible to infer phenotypic differences including infectivity and virulence (Cory *et al*., 2005). Evaluation of biological activity against geographically different insect colonies is useful in understanding and comparing the virulence of novel isolates, which may have implications for their use and efficacy as biopesticides in the field.

7.1.1 Insect rearing and colony establishment

To be able to study the biological activity of HearNPV isolates on *H. armigera*, it was important to have a constant supply of healthy insects. Laboratory reared insects allow for enough quantities to be available in comparison to wild-type insects which are not guaranteed to be readily accessible when needed (Dent, 2000). It was necessary to study the host biology, including fertility, stage development, rearing environmental conditions, disease and pest management to establish perimeters that would suffice in increasing the population density of *H. armigera*. The optimum rearing conditions for *H. armigera* have been documented (Armes *et al*., 1992; Zhou *et al*., 2000; Bouwer *et al.,* 2009; Perkins *et al*., 2009). It is recommended to rear the insects in a controlled environment. The rearing temperatures for the larvae in the study were carried out at 26°C (±2°C). The humidity for moths and eggs is required to be between 70-90% with the humidity of larvae being optimal between 35% and

60% (Armes *et al*., 1992; Grzywacz *et al*., 2007). In the CE rooms used for rearing, there were no installed humidifiers; this meant the humidity for moth and egg stages was not optimal. The use of cabbage leaves or green pepper cubes as diet was met with shortfalls which included short lifespan, wilting and secondary infection from microorganisms being observed after 48 hours. The secondary infection was also potentially triggered by the faecal pellets from the larvae, which are known to often trigger mould infection (Grzywacz *et al.*, 2007). The plant-based diet also resulted in excessive moisture in the rearing vials from transpiration, which exacerbates microbial contamination (Armes *et al*., 1992). However, the use of whole cabbage leaves was preferred by neonates over artificial diet and as no cannibalism was recorded for the first and second instar, allowing for group rearing on the leaves. The diet for rearing was changed to artificial bollworm diet that had antimicrobial agents in the mixture for the third instar onwards, as this ensured better quality control and consistent diet for the insect generations (Cohen, 2015). The use of plant-based diet initially, as opposed to artificial diet, was based on the lack of the material required for artificial diet preparation. As a result, the diet used was based on the plants the *H. armigera* larvae had been collected from in the field. The need to change the diet constantly and the secondary mould infection resulted in the larvae being stressed and initial high mortalities being recorded. The change of diet from a plant-based diet to an artificial diet resulted in better rearing control and increased survivorship of the insects, with time spent on rearing being reduced. Comparisons of plant-based diets and artificial diets have been reported and artificial diet has been shown to be the better option for mass rearing (Patankar *et al*., 2001, Perkins *et al*., 2008). It is possible to mass rear first and second instar in groups as these stages are not cannibalistic (Mironidis and Savapoulou-Soultani, 2008). Mass rearing requires less labour as there is no requirement to transfer each larva to a single vial, which in turn minimises the time spent on rearing (Liu et al., 2004; Mironidis and Savapoulou-Soultani, 2008). From the third instar the larvae had to be transferred to individual rearing vials because of cannibalism. The rearing protocol adapted was successful in the establishment of continuous generations of healthy larvae.

The pathogens of concern during *H. armigera* rearing are viruses, microsporidia protozoa, fungi and bacteria (Armes *et al.*, 1992). The virus infection for the colonies was due to the introduction of virus from the field and separation of healthy larvae ensured that no horizontal transfer that resulted in mortality occurred. Baculoviruses can cause non-lethal covert infections that can be passed from generation to generation, which may be triggered to

lethal overt infections when the larvae are stressed through overcrowding, poor diet or presence of other pathogens (Graham *et al.,* 2015). Fungal infection during rearing was observed on the diet used and on the eggs in the insectary and was reduced by improving quality control and rearing protocol through improved hygiene and minimising contamination. This was done by cleaning the insectary with a dilute hypochlorite solution and spraying the solution into the room. Other experiments that were being conducted including one that involved entomopathogenic fungi were removed from the rearing insectary.

The importance of having a healthy colony of *H. armigera* and establishing a suitable rearing protocol is not limited to having insects for biological assays. The current technology that is utilised for virus production is through the infection of susceptible insects (Claus *et al.*, 2012). The management of the rearing protocol and improvement of colony stabilisation ensures that enough healthy larvae are available for biological assays of different virus isolates as well as *in vivo* virus propagation for suitable isolates and for commercialisation. The production and the quality of the virus are dependent on the health of the larvae among other factors (Buerger *et al.,* 2007). It is important to consider that the larvae are cannibalistic in nature and this can result in the reduction in individuals (Arrizubieta *et al.*, 2016). However, rearing the larvae individually also results in increased labour and resources including time spent on the rearing therefore directly affecting efficiency and cost (Buerger *et al*, 2007; Arrizubieta *et al*., 2016).

7.1.2 Characterisation of HearNPV isolates

The bioprospecting of new virus isolates to aid in the control of *H. armigera* is of importance. Different HearNPV isolates that vary in virulence and control of *H. armigera* have been isolated from different host populations (Harrison *et al.,* 2008; Arrizubieta *et al.,* 2013). The molecular analysis of the genomes from different baculovirus isolates has identified genotype variants (Zhang *et al.,* 2005; Ogembo *et al.,* 2009, Opoku-Debrah *et al.,* 2013). The sequencing of HearNPV genomes allows for the analysis of molecular mechanisms that may be responsible for the variation in virulence exhibited by the isolates (Zhang *et al*., 2005). In the scope of this study, it was important to identify whether the South African HearNPV isolates were genetically different from other HearNPV isolates and if this translated into differences in virulence. It can be hypothesised that genetically different isolates perform differently against geographically similar as well distinct host populations. In a report by Figueiredo *et al.* (2009), the potency of genetically different HearNPV isolates against second

instar *H. armigera* varied by up to 3-fold in comparison to each other. Therefore, there is an importance to identifying genetic variation of HearNPV isolates as this can be linked to the virulence. The full genomes of HearNPV-Albany and HearNPV-KZN were sequenced and assembled using HearNPV-Au and HearNPV-G4 as the reference strains. The genome of HearNPV-Whit was unsuccessfully assembled because of poor sequence quality. The reference strains were chosen because they have been extensively characterised (Zhang *et al*., 2005; Lange *et al*., 2004; Ogembo *et al*., 2009; Zhang *et al*., 2014).

The genome sequence of HearNPV-Whit had low quality data, with over 40% of the genome not successfully sequenced. This led to the exclusion of HearNPV-Whit from the analysis. HearNPV-Whit used in the study had been stored for a long period (1997 extract) under frozen conditions. It is not documented how many times this sample was thawed and refrozen again prior to the analyses in 2014, which may have affected the quality of the suspension. The defrosting period during prior studies is not known, which may have affected OB quality and structure. The prolonged storage of baculovirus OBs has been reported to affect its stability and viability as well as result in the degradation of viral DNA (Behle and Birthisel, 2014). Another factor that could have affected the quality of HearNPV-Whit is contamination by other microbes as the microbes may consume resources that are required by the baculovirus to be stable (Behle and Birthisel, 2014) during the unfrozen periods. The genomes of HearNPV-Albany and HearNPV-KZN were successfully sequenced, assembled and annotated. The sequencing and annotation of HearNPV-Albany and HearNPV-KZN genomes enabled the identification of conserved and unique ORFs confirming that they were novel isolates.

The isolation and sequencing of HearNPV from different geographical locations has identified that the genetic make-up of the isolates can be different (Lua *et al*., 2002; Zhang *et al*., 2005; Ogembo *et al*., 2007; Noune and Hauxwell, 2016). On this basis, comparative analysis of HearNPV-Albany and HearNPV-KZN was performed to confirm that the novel South African isolates were indeed genetically different from the currently sequenced isolates. The isolates used in the comparison were from Australia, China, Spain and Kenya and in the eight isolates used for comparison, genetic difference was recorded. The genetic difference may be a result of co-evolution of the host and the virus which results in genetic divergence (Lua *et al*., 2002; Possee *et al*., 2010). Different selection pressures associated with geographical distribution may also be responsible for different genotype variation of the host virus (Hitchman *et al*., 2007). Further analysis including REN, identification of SNPs in

conserved genes, nucleotide alignment to identify ORF variance as well as construction of phylogenetic trees confirmed that HearNPV-Albany and HearNPV-KZN were genetically distinct in comparison to other isolates. The nucleotide alignment of HearNPV-Albany and HearNPV-KZN against reference HearNPV strains immediately identified differences at the nucleotide level, with percentage identities of the genome sequences being between 93.5% and 96.0%. Most of the differences in the genomes were observed in the intergenic spaces of the genome. Deletions, insertion and substitutions are reported to be more common in the intergenic spaces of the genome (Chen *et al.*, 2002). The REN analyses demonstrated genetic differences for HearNPV-Albany and HearNPV-KZN in comparison to the reference isolates as different band fragments were generated for each of the isolates. Analyses of highly conserved genes through SNPs identified synonymous and non-synonymous mutations. Most of the mutations identified were synonymous. For the non-synonymous mutations, amino acid substitutions were of amino acids with side chains that shared similar polarity. The non-synonymous mutations for the important proteins had side chains with similar polarity therefore not affecting overall protein function (Patthy, 2008). This meant that the overall structure and functionality of the conserved proteins are potentially not altered (Wilson and Walker, 2009). Phylogenetic trees constructed using HearNPV gene sequences, showed that the two South African isolates, HearNPV-Albany and HearNPV-KZN were genetically distant, with evolutionary patterns grouping HearNPV-Albany with an African, European and Australian isolate whereas HearNPV-KZN was grouped with the Chinese and a different Australian isolate. A maximum-likelihood comparison of whole-genome sequences identified clusters of HearNPV genomes based on geographic origin of the isolates (Noune and Hauxwell, 2016). Analyses from phylogenetic trees using different loci on geographically different HearNPV isolates also identified clustering based on geographic origins (Rowley *et al.*, 2011). The differences observed for the genomes could be due to the co-evolution of the virus with its host population (Arrizubieta *et al.*, 2015). It has been reported that genotype phylogeny is not always associated with the geographical location of the isolate with some isolates from distinct geographical locations being more closely related on the genome level (Figueiredo *et al.*, 2009). The relations of the HearNPV genomes from geographically distinct locations that were reported through the clusters may be a result of the migratory patterns of *H. armigera* and the distribution of the host populations (Karim, 2000, Lammers and McLeod, 2007). Other factors that may explain the HearNPV genome relationship identified through clusters such as of HearNPV-Albany, HearNPV-SP1, HearNPV-NNg1 and HearNPV-AC53 and that of HearNPV-KZN, HearNPV-G4, HearNPV-C1 and HearNPV-Au

may also be a potential result of the migratory transport of HearNPV isolates through larval parasitoids and insectivorous birds (Figueiredo *et al*., 2009). In regions that are not conserved such as the *hrs* and *bro* genes, intragenomic recombination has been reported (Erlandson, 2009). Homologous recombination of variants of a baculovirus or closely related baculoviruses that co-infect the same host has also been reported (Erlandson, 2009). These factors may be responsible for the genetic differences that are observed in HearNPV isolates from geographically distinct locations.

7.1.3 Biological assays

Full nucleotide comparison and alignment identified differences in the HearNPV-Albany and HearNPV-KZN genomes in comparison to other genomes. The genetic differences of HearNPV-Albany and HearNPV-KZN in comparison to other isolates including HearNPV-Au, HearNPV-SP1 and HearNPV-G4 raised the potential for different phenotypic traits that include virulence. For the biological studies, four South African isolates namely HearNPV-Albany, HearNPV-KZN, HearNPV-Whit and HearNPV-Haygrove were studied against two geographically distinct host populations that were collected in South Africa. Other non-South African HearNPV isolates namely HearNPV-SP1 and HearNPV-G4 as well as the commercial products Helicovir™ and Helicovex® were tested against one South African host. It was important to monitor the movement of the larvae so that they were limited to the individual vials with the surface dosed diet. Considerable movement and escaping from the rearing vials were recorded for some of the second instar larvae. Movement of larvae upwards is a common observation in *H. armigera* larvae (Perkins *et al*., 2008), and because of the cotton plug used, this allowed escape of the larvae in some cases. These larvae had to be discarded from the biological analyses.

The results from the dose-response mortality studies showed that the virulence of the isolates differed and, in some instances, significantly, with HearNPV-Albany being the most potent and the commercial isolates the least potent. Ogembo *et al.* (2007) in comparative studies of six HearNPV isolates from Kenya, an isolate from South Africa, Zimbabwe, China and Thailand respectively, recorded significant differences in the LD_{50} between the isolates. HearNPV-G4 had significantly lower virulence in comparison to the HearNPV African isolates including HearNPV-Whit (Ogembo *et al*., 2007). For the bioassays conducted in Chapter 6, HearNPV-G4 had significantly lower virulence in comparison to HearNPV-Albany only for the South African isolates. Compared to the other isolates, the time-response for HearNPV-Albany was significantly different with between 1.2 times to 1.9 times delay in

mortality recorded in comparison and up to 2.5 times delay in mortality recorded for the commercial formulations. The exact causes for the difference in virulence recorded is not known but a hypothesis can be raised based on the genotype differences reported during the comparative studies in Chapter 5 that may have been well suited for the co-evolution of the insect host and host-pathogen (HearNPV-Albany and Albany *H. armigera*). Looking at the dose-response and time-response assays, compared to well characterised isolates such as HearNPV-G4, the South African isolates had comparable or improved virulence. This lays the foundation for the novel South African HearNPV isolates to be further analysed and characterised with the end goal of commercial product development. Field trial testing of the South African HearNPV isolates would confirm the virulence of the strains. Successful field testing against *H. armigera* host populations would be a required step in the development and registration of virulent strains of HearNPV isolated from South African native populations. The different responses of the two host populations to the South African HearNPV isolates confirmed that different isolates had different biological properties. The genome diversity plays a role in virulence because of insect-virus dynamics and evolution (Cory *et al*., 2005). Reduced susceptibility to isolates of the same virus by local host populations have been reported because of specific genes that have been identified in the hosts. Resistance was also reported after continued generations of infection by the same virus, with later generations been less susceptible to the virus strains (Siegwart *et al*., 2015). The resistance mechanism has been attributed to the insect midgut cells inhibiting transfer of virus to the rest of the insect cells or the blockage of early replication in the insect cells (Siegwart *et al*., 2015). The biological studies identified South African HearNPV isolates that have the potential to be utilised in the effective control of *H. armigera*. The genetic differences recorded for HearNPV-Albany, HearNPV-KZN, HearNPV-SP1 and HearNPV-G4 during the comparative studies were shown to translate to differences in virulence. The HearNPV-Au strain is also used as the strain to produce the South African commercial HearNPV formulation, Helicovir™ (Moore and Kirkman, 2010). The use of Helicovir™ in the control of South African *H. armigera* may lead to the covert expression and dispersion of the virus amongst local populations. Insects that survive biopesticide application have been reported to migrate and carry the virus covertly to new areas (Moscardi, 1999).

7.2 Conclusions and future work

The main aims of the study were to establish healthy, disease-free colonies of South African collected *H. armigera*, to isolate and characterise HearNPV from South African host

populations and to study the biological activity against South African *H. armigera* host populations and compare against geographically distinct isolates and commercial products. The aims of the project were achieved through the sequential fulfilment of the objectives.

This study has allowed further steps into the production of commercial HearNPV products for use in the control of *H. armigera* and understanding how genetic variation may affect biological activity. The perception of growers towards biological pesticides has been changing because of chemical pesticide resistance, continued education of detrimental effects of some chemical pesticides as well as government and trading bloc pressures such as the EU (Buerger *et al.*, 2007). The objectives undertaken in the study set a foundation on how to continue the studies of *H. armigera* control using HearNPV isolates and confirmed that geographically different isolates have considerable genetic differences. The biological activities of the HearNPV isolates also demonstrated significantly improved performance in comparison to commercially produced HearNPV products.

Laboratory based bioassays recorded with two host populations demonstrated significantly improved virulence when using HearNPV-Albany, HearNPV-KZN, HearNPV-Haygrove and HearNPV-Whit in comparison to the commercial formulations of Helicovir™ and Helicovex®. The performance of HearNPV-Whit against *H. armigera* shows that the virus was still viable, therefore, the hypothesis for the genome not being sequenced successfully due to viral degradation did not hold. This makes the possibility of contamination by other microbes therefore affecting the extraction of the DNA quality the more plausible explanation. The mass production of the South African HearNPV isolates leading up to field trials should be considered for future work on the foundation prepared. The bioprospecting conducted and the successful isolation of novel HearNPV isolates represents continuous potential for possible future identification of new isolates from distinct host populations. It would also be important to have full genomic sequences for HearNPV-Haygrove and HearNPV-Whit. The presence of full genomic sequence functions as references for future HearNPV studies. There is considerable potential for the South African isolates to be implemented in control programs as standalone biopesticides or as part of an IPM system.

8 References

Abate T., (2012), *Chrysoperla* spp. an important predator of African bollworm. Maize stalk borers of Ethiopia: quantitative data on ecology and management in meeting the challenges of global climate change and food security through innovative maize Research, proceeding of the 3rd national maize workshop of Ethiopia April 18-20, 2011; Addis Ababa, Ethiopia. (ed. Worku M., Twumasi-Afriyie S., Wade L., Tadesse B., Demisie G., Bagale G., Wegary D and Prasanna B.M.), CIMMYT and EIAR, pp 174-185

Abrol D.P. and Shankar U., (2012), History, overview and principles of ecologically-based pest management. Integrated pest management: principles and practice. (ed. Abrol D.P. and Shankar U.), CABI, UK, pp 1-26

Ackhurst R.J., James W., Bird L.J. and Beard C., (2003), Insecticide resistance and resistance management, Resistance to the Cry1Ac δ-endotoxin of *Bacillus thuringiensis* in the cotton bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae). Journal of economical entomology, volume 96, issue 4, pp 1290-1299

Akhtar Y. and Isman M.B., (2013), Plant natural products for natural pest management: The magic mixtures. Advanced technologies for managing insect pests. (ed. Ishaaya I., Palli S.R. and Horowitz R.). Springer, Dordrecht, pp 231-247

Alvi A.H.K., Sayyed A.H., Naeem M. and Ali M., (2012), Field evolved resistance in *Helicoverpa armigera* (Lepidoptera: Noctuidae) to *Bacillus thuringiensis* Toxin Cry1Ac in Pakistan. PLoS ONE volume 7, issue 10: e47309. doi: 10.1371/journal.pone.0047309

Ahmed K. and Khalique F., (2012), Oviposition and larval development of *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae) in relation with chickpea, *Cicer arietinum* L. (Fabaceae) crop phenology. Pakistan journal of zoology, volume 44, issue 4, pp 1081-1089

Amphofo K., (1994), Pod and seed feeders. Field pests of bean in Africa. (ed. Llanos C. and Walter A.), CIAT, Colombia, pp 26-32

Annecke D.P. and Morgan V.C., (1982), Insects and mites of cultivated plants in South Africa. Durban: Butterworth, South Africa.

Ardisson-Araujo D.M.P., Sosa-Gomez D.R., Melo F.L., Bao S.N. and Ribeiro B.M., (2015), Characterisation of *Helicoverpa zea* single nucleopolyhedrovirus isolated in Brazil during the

first old bollworm (Noctuidae: *Helicoverpa armigera*) nationwide outbreak. Virus reviews and research, volume 20, issue 1, pp 1-4

Armes N.J., Jadhav D.R. and DeSouza K.R., (1996), A survey of insecticide resistance in *Helicoverpa armigera* in the Indian subcontinent. Bulletin of entomological research, volume 86, issue 5, pp 499-514

Arrizubieta M., Williams T., Caballero P. and Simon O., (2013), Selection of a nucleopolyhedrovirus from *Helicoverpa armigera* as the basis for a biological insecticide. Pest management scientific, volume 70, issue 6, pp 967-976

Arrizubieta M., Simon O., Williams T. and Caballero P., (2015), Genomic sequences of five *Helicoverpa armigera* nucleopolyhedrovirus genotypes from Spain that differ in their insecticidal properties. Genome announcement, volume 3, issue 3: e00548-15. doi:10.1128/genomeA.00548-15.

Arrizubieta M., Simon O., Williams T. and Caballero P., (2016), Determinant factors in the production of a co-occluded binary mixture of *Helicoverpa armigera* alphabaculovirus (HearNPV) genotypes with desirable insecticidal characteristics, PLoS ONE, volume 11, issue 10: e0164486.doi: 10.1371/journal.pone.0164486

Asser-Kaiser S., Fritsch E., Undorf-Spahn K., Kienzle J., Eberle K.E., Gund N.A and Jehle J.A., (2007), Rapid emergence of resistance in codling moth due to dominant sex-linked inheritance. Science, volume 318, pp 1916-1917

Badii K.B. and Asante S.K., (2012), Efficacy of some synthetic insecticides for control of cotton bollworms in northern Ghana. Africa crop science journal, volume 20, issue 1, pp 59-66

Ball L.A., (2007), Virus replication strategies. Fields' virology, volume 1, 5th edition. (ed. Knipe D.M. and Howley P.M.). Lippincott Williams & Wilkins, USA, pp 119-140

Bell J.C. and McGeoch M.A., (1996), An evaluation of the pest status and research conducted on phytophagous Lepidoptera on cultivated plants in South Africa. African entomology, volume 4, issue 2, pp 161-170

Behle R and Birthisel T., (2014), Formulation of entomopathogens as bioinsecticides. Mass production of beneficial organisms: Invertebrates and entomopathogens. (ed. Morales-Ramos J., Rojas M. G. and Shapiro-Ilan D), 1st edition, Academic press, United States, pp 483-517

Berretta M.F., Ferrelli M.L., Salvador R., Sciocco A. and Romanowksi V., (2013), Baculovirus gene expression. Biochemistry, Genetics and Molecular Biology, "Current Issues in Molecular Virology – Viral Genetics and Biotechnological Applications, (ed. Romanowski V.), InTech, UK, pp 57-78

Bideshi D.K., Renault S., Statsiak K., Federici B.A. and Bigot Y., (2003), Phylogenetic analysis and possible function of bro-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria. Journal of general virology, volume 84, issue 9, pp 2531-2544

Blissard G., Black B., Crook N., Keddie B.A., Possee R., Rohrmann G., and Volkman L., (2000), Family Baculoviridae. Virus taxonomy: Seventh report of the international committee of taxonomy of viruses. (ed. van Regenmortel), Academic press, San Diego, USA, pp 195-202

Bouwer G., Nardini L. and Duncan F.D., (2009), *Helicoverpa armigera* (Lepidoptera: Noctuidae) larvae that survive sublethal doses of nucleopolyhedrovirus exhibit high metabolic rates. Journal of insect pathology, volume 55, issue 4, pp 369-374

Bownes A., Moore S.D. and Villet M.H., (2014), My enemy's enemies: recruiting Hemipteran-tending generalist ants for biological control in citrus orchards by spatial partitioning of foraging webs. African entomology, volume 22 issue 3, pp 519-529

Brouwer A., Longnecker M.P., Birnbaum L.S., Cogliano J., Kostyniak P. and Moore J., (1999), Characterisation of potential endocrine related health effects at low dose exposure to PCBs. Environmental health perspectives, volume 107, pp 639-649

Brown D.P., Wilkinson H.S. and Gatehouse J.A., (1997), Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. Insect biochemistry and molecular biology, volume 27, issue 7, pp 625-636

Buerger P., Hauxwell C. and Murray D., (2007), Nucleopolyhedrovirus introduction in Australia. Virologica Sinica, volume 22, issue 2, pp 173-179

Butler J.M., (2010), DNA extraction. Fundamentals of forensic DNA typing. (ed. Butler J.M.). Academic Press Elsevier, USA, pp 99-110

Calkins C.O. and Parker A.G., (2005), Sterile insect quality. Sterile insect technique: Principles and practice in area-wide integrated pest management, (ed. Dyck V.A., Hendrichs J. and Robinson A.S), Springer, Netherlands, pp 269-298

Calow P., (1981), Introduction. Invertebrate biology: a functional approach, first edition. Croom Helm London, London, pp 11-19

Camoretti-Mercado B., (2008), Two common DNA analysis tools. Bioinformatics: a practical approach. (ed. Ye S.C.), CRC press, USA, pp 55-79

Chen X., Hu Z., Jehle J.A., Zhang Y. and Vlak J.M., (1997), Analysis of the ecdysteroid UDP-glucosyltransferase gene of *Heliothis armigera* Single-Nucleocapsid baculovirus. Virus genes, volume 15, issue 3, pp 219-225

Chen X., Li M., Sun X., Arif B.M., Hu Z. and Vlak J.M., (2000), Genomic organisation of *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus, Archives of virology, volume 145, issue 12, pp 2539-2555

Chen X., Ijkel W.F.J., Tarchini R., Sun X., Sandbrink H., Wang H., Peters S., Zuidema D., Lankhorst R.K., Vlak J.M. and Hu Z., (2001), The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus, Journal of general virology, volume 82, issue 1, pp 241-257

Chen X., Zhang W.J., Wong J., Chun G., Lu A., McCutchen B.F., Presnail J.K., Herrmann R., Dolan M., Tingey S., Hu Z.H. and Vlak J.M., (2002), Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus. Journal of general virology, volume 83, issue 3, pp 673-684

Cherry A., Cock M., van der Berg H. and Kfir R., (2003), Biological control of *Helicoverpa armigera* in Africa. Biological control in an integrated pest management system in Africa. (ed. Neunshwande P., Borgemeister C. and Langewald J.), CABI publishing, UK, pp 329-346

Christian P.D., Gibb N., Kasprzak A.B. and Richards A., (2001), A rapid method for the identification and differentiation of *Helicoverpa* nucleopolyhedrovirus (NPV Baculoviridae) isolated from the environment. Journal of virological methods, volume 96, issue 1, pp 51-65

Cilas C., Goebel F-R, Babin R. and Avelino J., (2015), Tropical crop pests and diseases in a climate change setting – a few examples. Climate change and agriculture worldwide. (ed. Torquebiau E.), Springer, France, pp 73-82

Claus J.D., Gioria V.V., Micheloud G.A. and Visnovsky G., (2012), Production of insecticidal baculoviruses in insect cell cultures: potential and limitations. Insecticides – basic and other applications, (ed. Soloneski S. and Larramendy M.), InTech, UK, pp 127-152

Cohen A.C., (2015), Introduction. Insect diets: Science and technology, 2nd edition (ed. Cohen A.C.), CRC press, USA, pp 1-20

Condit R.C., (2007), Principles of virology. Fields' virology, volume 1, 5th edition. (ed. Knipe D.M. and Howley P.M.). Lippincott Williams & Wilkins, USA, pp 25-58

Cory J.S., Green B.M., Paul R.K. and Hunter-Fujita F., (2005), Genotypic and phenotypic diversity of baculovirus population within an individual insect host. Journal of invertebrate pathology, volume 89, issue 2, pp 101-111

Council of Europe. 2013. Statistical analysis of results of biological assays and tests. Pages 551-579 in European Pharmacopoeia, 8th ed. Council of Europe.

Cunningham J.C., (1995), Baculoviruses as microbial insecticides. Novel approaches in integrated pest management. (ed. Reuveni R.), Boca Ranton, Lewis, Florida, pp 261-292

Deng F., Wang R., Fang M., Jiang Y., Xu X., Wang H., Chen X., Arif B.M., Guo L., Wang H. and Hu Z., (2007), Proteomics analysis of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus identified two new occlusion derived virus associated proteins, HA44 and HA100. Journal of virology. volume 81, issue 17, pp 9377-9385

Dent D., (2000), Yield loss. Insect pest management, 2nd edition, Cabi publishing, UK, pp 48-80

Dreistadt S.H., (2012), Managing pests in citrus. Integrated pest management for citrus. 3rd edition (ed. Dreistadt S.H.). UC peer reviewed, Canada, pp 13-38

Du L. and Egholm M., (2008), The next-generation genome sequencing: 454/Roche GS FLX. Next generation genome sequencing: Towards personalised medicine. (ed. Janitz M.), Wiley-VCH Verlag GmbH & Co. KGaA, pp 43-56

Dubey R.C., (2014), Bioinformatics. Advanced biotechnology. (ed. Dubey R.C.), S Chand & Company PVT. LTD, New Delhi, pp 524-569

Eberle K.E., Asser-Kaiser S., Sayed S.M., Nguyen H.T. and Jehle J.A., (2008), Overcoming the resistance of codling moth against conventional *Cydia pomonella* granulovirus (CpGV-M) by a new isolate CpGV-112. Journal of invertebrate pathology, volume 98, pp 293-298

Eberle K.E., Jehle J.A. and Hüber J., (2012), Microbial control using insect viruses. Integrated pest management: principles and practice. (ed. Abrol D.P. and Shankar U.), CABI, UK, pp 281-290

Ekblom R. and Wolf J.B.W, (2014), A field guide to whole-genome sequencing, assembly and annotation. Evolutionary applications, doi.org/10.1111/eva.12178, pp 1026-1042

El-Metwally S., Ouda O.M. and Helmy M., (2014) Next-generation sequencing platforms. Next generation sequencing technologies and challenges in sequence assembly. (ed. El-Metwally S., Ouda O.M. and Helmy M.), Springer, New York, pp 37-44

Elderd B.D. and Reilly J.R., (2014), Warmer temperatures increase disease transmission and outbreak intensity in a host-pathogen system. Journal of animal ecology, pp 1-12

Erlandson M.A., (2009), Genetic variation of field populations of baculoviruses: Mechanisms for genetic variation and its potential role in baculovirus epizootiology, Virologica sinica, volume 24 issue 5, pp 458-469

Eroglu G.B., Demir I. and Demirbag Z., (2018), A novel alphabaculovirus isolated from the cotton bollworm, *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae): Characterisation and pathogenicity. Biologia, volume 78, issue 5, pp 545-551

European Union, (2009), Directive 2009/128/EC of the European parliament and of the council of 21 October 2009: establishing a framework for community action to achieve the sustainable use of pesticides. Official journal of the European Union. L309/71

Fauquet C.M., Mayo M.A., Moniloff J, Desselberger U. and Ball L.A., (2005), Baculoviridae. Virus taxonomy eighth report of the international committee on taxonomy of viruses, 1st edition. (ed. Fauquet C.M., Mayo M.A., Moniloff J, Desselberger U. and Ball L.A), Elsevier inc, San Diego, California, pp 177-186

Fayyaz S., Kazi F. and Khanum T.A., (2017), Status of entomopathogenic nematodes in integrated pest management strategies in Pakistan. Biocontrol agents: Entomopathogenic and

slug parasitic nematodes. (ed. Mahfouz M.M., Abd-Elgawad, Askary T.H. and Coupland J.) CABI, UK, pp 383-408

Federici B.A., (1997), Baculovirus infection. The baculoviruses. (ed. Miller L.K.), Plenum press, New York, pp 33-59

Figueiredo E., Munoz D., Murillo R., Mexia A. and Caballero P., (2009), Diversity of Iberian nucleopolyhedrovirus wild-type isolates infecting *Helicoverpa armigera* (Lepidoptera: Noctuidae). Biological control, volume 50, issue 1, pp 43-49

Fitt G.P., (1989), The ecology of Heliothis species in relation to agroecosystems. Annual review of entomology, volume 34, pp 17-53

Fitt G.P., (1991), Host selection in the Heliothine. Reproductive behaviour in insects – individuals and populations. (ed. Bailey W.J. and Ridsdill-Smith T.J.), Chapman and Hall, London, pp 172-201

Fitt G. and Wilson L., (2012), Integrated pest management for sustainable agriculture. Integrated pest management: principles and practice. (ed Abrol D.P. and Shankar U.), CABI, UK, pp 27-40

Fritsch E., Undorf-Spahn K., Kienzle J., Zebitz C.P. and Huber J., (2005), Codling moth granulovirus: first indications of variations in susceptibility of local codling moth populations. Nachrichtenbl. Deut. Pflanzenschutzd, volume 57, pp 29-34

Funk C.J., Braunagel S.C. and Rohrmann G.F., (1997), Baculovirus structure. The baculoviruses. (ed. Miller L.K.), Plenum press, New York, pp 7-32

Gassmann A.J., Petzold-Maxwell J.L., Keweshan R.S. and Dunbar M.W., (2011), Field evolved resistance to *Bt* maize by western corn rootworm. PLoS One, volume 6, issue 7, e22629. doi:DOI:10.1371/journal.pone.0022629

Georgievska L., Hoover K., van der Werf W., Munoz D., Caballero P., Cory J.S. and Vlak J.M., (2010), Dose dependency of time to death in single and mixed infections with a wildtype and egt deletion strain of *Helicoverpa armigera* nucleopolyhedrovirus. Journal of invertebrate pathology, volume 104, issue 1, pp 44-50

Gettig R.R., Burand J.P., Flore P.H., McCarthy W.J. and Wood H.A., (1987), Replications of Heliothis zea nuclear polyhedrosis virus in insect culture: virus-induced and structural proteins. Journal of general virology, volume 68, pp 1617-1625

Goodman D., Ollikainen N. and Sholley C., (2007), Baculovirus phylogeny based on genome rearrangements. Comparative genomics: RECOMB 2007, international workshop, RECOMB-CG 2007 San Diego, CA, USA, September 16-18, 2007 proceedings. (ed, Tesler G. and Durand D.), Springer, USA, pp 69-83

Goodwin W., Linacre A. and Hadi S., (2007), DNA extraction and quantification. An introduction to forensic genetics. (ed. Goodwin W., Linacre A. and Hadi S.,), John Wiley and Sons Ltd., England, pp 27-38

Gordh G. and Headrick D., (2011), African bollworm. A dictionary of entomology, 2nd edition. (Gordh G. and Headrick D.). CABI, UK, pp 32

Goulson D., (2003), Can host susceptibility to baculovirus infection be predicted from host taxonomy or life history? Population ecology, volume 32, issue 1, pp 61-70

Gouse M., (2013), Socioeconomic and farm-level effects of genetically modified crops: the case of *Bt* crops in South Africa. Genetically modified crops in Africa, economic and policy lessons from countries south of the Sahara. (ed. Falck-Zepeda J., Gruere G. and Sithole-Niang I.), International food policy residence institute, USA, pp 25-42

Graham R.I., Tummala Y., Rhodes G., Cory J.S., Shirras A., Grzywacz D. and Wilson K., (2015), Development of real-time qPCR assay for quantification of covert baculovirus infections in a major African crop pest. Insect, volume 6, issue 3, pp 746-759

Granados R.R. and Federici B.A., (1986), The biology of baculoviruses volume I & II, CRC press, Boca Raton, Florida.

Grant M. and Bouwer G., (2009), Variation in the time-mortality responses of laboratory and field populations of *Helicoverpa armigera* infected with nucleopolyhedrovirus. Journal of applied entomology, volume 133, issue 8, pp 647-650

Groner A., (1990), Safety to non-target invertebrates of baculoviruses. Safety of microbial insecticides. (ed. Laird M., Lacey L.A. and Davidson E.W.), CRC press, Boca Ranton, Florida, pp 135 – 147

Grzywacz D., (2001), Nucleopolyhedrovirus: Potential in the control of the pod-borer (Helicoverpa armigera) on chickpea in Nepal. On-farm IPM of chickpea in Nepal, proceedings of the international workshop on planning and implementation of on-farm chickpea IPM in Nepal, 6-7 September 2000, Kathmandu, Nepal. Patancheru 502 324,

Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics; and Chatham, Kent ME4 4TB, UK: Natural Resources Institute. 132 pp. ISBN 92-9066-438-X. Order code: CPE 134, pp 94-98

Grzywacz D., Rabindra R.J., Brown M., Jones K.A. and Parnell M., (2007), The *Helicoverpa armigera* NPV production manual, pp 1-104

Grzywacz D., (2017), Basic and applied research: Baculovirus. Microbial control of insect and mite pests: From theory to practice. (ed Lacey L.A.), Academic press, USA. Pp 27-46

Grzywacz D. and Moore S., (2017), Production, formulation, and bioassay of baculovirus for pest control. Microbial control of insect and mite pests: From theory to practice. (ed Lacey L.A.), Academic press, USA. Pp 109-124

Gujar G.T., Nair R., Singh B.P., Kumari A. and Kalia V., (2008), Toxicity to the cotton bollworm, *Helicoverpa armigera*, of some Cry1Ac toxins expressed in cotton in India. Crop protection, volume 27, issue 3-5, pp 537-544

Gunning R.V., Moores G.D. and Devonshire A.L., (1999), Esterase inhibitors synergise the toxicity to pyrethroids in Australian *Helicoverpa armigera* (Hubner) (Lepidoptera: Noctuidae). Pesticide biochemistry and physiology, volume 63, pp 50-62

Gunning R.V., Dang H.T., Kemp F.C., Nicholson I. C. and Moores G.D., (2005), New resistance mechanism in *Helicoverpa armigera* threatens transgenic crops expressing *Bacillus thuringiensis* Cry1Ac toxin. Applied and environmental microbiology, volume 75 issue 5, pp 2,558-2,563

Gupta R.K., Raina J.C., Arora R.K. and Bali K., (2007), Selection and field effectiveness of nucleopolyhedrovirus isolates against *Helicoverpa armigera* (Hübner). International journal of virology, volume 3 issue 2, pp45-59

Haase S., Ferrelli L., Pidre M.L. and Romanowski V., (2013), Genetic engineering of baculoviruses. Biochemistry, Genetics and Molecular Biology, “Current Issues in Molecular Virology – Viral Genetics and Biotechnological Applications, (ed. Romanowski V.), InTech, UK, pp 79-111

Haase S., Sciocco-Cap A. and Romanowski V., (2015), Baculovirus insecticides in Latin America: Historical overview, current status and future perspective. Virus, volume 7, issue 5, pp 2230-2267

HackettD.S. and Gatehouse A.G., (1982), Diapause in *Heliothis armigera* (Hübner) and *H. fletcheri* (Hardwick) (Lepidoptera: Noctuidae) in the Sudan Gezira. Bulletin of entomological research, volume 72, issue 3, pp 409-422

Hanzlik T.N., Dorrian S.J., Gordon K.H.J. and Christian P.D., (1993), A Novel small RNA virus isolated from the cotton bollworm, Helicoverpa armigera, Journal of general virology, volume 74, pp 1805-1810

Hanzlik T.N., Dorrian S.J., Johnson K.N., Brooks E.M. and Gordon K.H., (1995), Sequence of RNA2 of the Helicoverpa armigera stunt virus (Tetraviridae) and bacterial expression of its genes, Journal of general virology, volume 76, pp 799-811

Harrison R. and Hoover K., (2012), Baculoviruses and other occluded insect viruses. Insect pathology, 2nd edition. (ed. Vega F.E. and Kaya H.K.), Academic press, UK, pp 73-132

Hatting J.L. and Malan A.P., (2017), Status of Entomopathogenic nematodes in integrated pest management strategies in South Africa. Biocontrol agents: Entomopathogenic and slug parasitic nematodes. (ed. Mahfouz M.M., Abd-Elgawad, Askary T.H. and Coupland J.) CABI, UK, pp 409-428

Hatting J.L., Moore S.D. and Malan A.P., (2018), Microbial control of phytophagous invertebrate pests in South Africa: Current status and future prospects. Journal of invertebrate pathology, doi.org/10.1016/j.jip.2018.02.004, pp 1-13

Herniou E.A., Luque T., Chen X., Vlak J.M., Winstanley D., Cory J.S. and O'Reilly D.R., (2001), Use of whole genome sequence data to infer baculovirus phylogeny. Journal of virology, volume 75 issue 17, pp 8117-8126

Herniou E.A., Olszewski J.A., Cory J.S. and O'Reilly D.R., (2003), The genome sequence and evolution of baculoviruses. Annual review of entomology, volume 48, issue 1, pp 211-234

Hill D.S., (1987), Major tropical crop pests. Agricultural insect pests of the tropics and their control, 2nd edition. (ed. Hill D.S.), press syndicate of the University of Cambridge, pp 326

Hill D.S., (2008), Pesticides in current use. Pests of crops in warmer climates and their control. (ed. Hill D.S.), Springer, UK, pp 107-114

Hu Z., Chen X. and Sun X., (2003), Molecular biology of insect viruses. Advances in microbial control of insect pests. (ed. Upadhyay R.K.), pp 83-108

Huber J., (1986), Use of baculoviruses in pest management programs. The biology of baculoviruses: practical applications for insect control, volume 2. (ed. Granados R.R. and Federici B.A.), CRC press, Boca Ranton, Florida, pp 181-202

Huesing J. and English L., (2004), The impact of *Bt* crops on the developing world. Agricultural bioforum, volume 7, issue 1-2, pp 84-95: http://www.agbioforum.org/

Hunter E., (2007), Virus assembly. Fields' virology, volume 1, 5th edition. (ed. Knipe D.M. and Howley P.M.). Lippincott Williams & Wilkins, USA, pp141-168

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

Hurley P.M., Hill R.N. and Whiting R.J., (1998), Mode of carcinogenic action of pesticides inducing thyroid follicular cell tumours in rodents. Environmental health perspectives, volume 106, issue 437.

Ibarra J.E. and Rincón-Castro M.C.D., (2009), Insect viruses: Diversity, biology and use as bioinsecticides. Tropical biology and conservation management – volume 7: phytopathology and entomology: Encyclopedia of life systems. (ed. Claro K.D., Oliviera P.S. and Rico-Gray V), EOLSS publishers, Oxford, UK, pp 1-30

Jallow M.F.A. and Matsumara M., (2001), Influence of temperature on the rate of development of *Helicoverpa armigera* (Hubner) (Lepidoptera: Noctuidae). Applied entomology and zoology, volume 36, issue 4, pp 427-430

Jehle J.A., Blissard G.W., Bonning B.C., Cory J.S., Herniou E.A., Rohrmann G.F., Theilmann D.A., Thiem S.M. and Vlak J.M., (2006), On the classification and nomenclature of baculoviruses: A proposal for revision. Archives of virology, volume 151, issue 7, pp 1257-1266

Jones K.A., (2000), Bioassays of entomopathogenic viruses. Bioassays of entomopathogenic microbes and nematodes. (ed. Navon A. and Ascher K.R.S.), CABI publishing, UK, pp 95-140

Kalawate A.S., (2014), Microbial viral insecticides. Basic and applied aspects of biopesticides. (ed. Sahayaraj K.), Springer, India, pp 47-68

Kang W., Suzuki M., Zemskov E., Okano K. and Maeda S., (1999), Characterisation of baculovirus repeated open reading frames (bro) in *Bombyx mori* nucleopolyhedrovirus. Journal of virology, volume 73, issue 9, pp 10339-10345

Karel A.K. and Autrique A., (1989), Insects and other pests in Africa. Bean production problems in the tropics, 2nd edition. (ed. Schwartz H.F. and Pastor-Corrales M.A.), Centro Internacional de Agricultura Tropical (CIAT), pp 455-504

Karim S., (2000), Management of *Helicoverpa armigera*: A review and prospectus for Pakistan. Pakistan journal of biological sciences, volume 3, issue 8, pp 1213-1222

Kaur B., Gupta V.K. and Jindal V., (2014), Molecular phylogenetic analysis of Indian strains of Helicoverpa armigera nucleopolyhedrovirus (HearNPV), Indian journal of biotechnology, volume 13, pp 186-194

Kaya H.K., Aguillera M.M., Alumai A., Choo H.Y., Torre M.d.l., Fodor A., Ganguly S., Hazir S., Lakatos T., Pye A., Wilson M., Yamanaka S., Yang H. and Ehlers R.-U., (2006), Status of entomopathogenic nematodes and their symbiotic bacteria from selected countries or regions of the world. Biological control, volume 38, issue 1, pp 134-155

Kessler P., Benuzzi M. and Mayoral F., (2008), Developing new baculoviruses products or "How to walk a tightrope", pp 44-49

Khan S., Sneddon K., Fielding B., Ward V. and Davison S., (2003), Functional characterisation of the ecdysteroid UDP-glucosyl transferase gene of *Helicoverpa armigera* single-enveloped nucleopolyhedrovirus isolated in South Africa. Virus genes, volume 27, issue 1, pp 17-27

Khan S., Fielding B. and Davison S., (2004), Identification of three genes of *Helicoverpa armigera* single-enveloped nucleopolyhedrovirus and their application in determining HaSNPV-SA phylogeny. South African journal of science, volume 100, issue 5, pp 301-305

Kidwell K.K. and Osborn T.C., (1992), Simple plant DNA isolation procedures. Plant genomes: methods for genetic and physical mapping. (ed. Beckmann J.S. and Osborn T.C.), Kluwer academic publishers, USA, pp 1-14

King A. B., (1994)., *Heliothis/Helicoverpa* (Lepidoptera: Noctuidae). Insect Pests of Cotton. (ed. Mathews G.A. and Turnstall J.P.), CAB International, Wallingford, UK pp. 39-106

King A. M., Adams M. J., Carstens E. B. and Lefkowitz E. J. (2012*). Virus Taxonomy: Classification and Nomenclature of Viruses, Ninth Report of the International Committee on Taxonomy of Viruses.* Elsevier Academic Press, Boston

Knipling E.F., (1966), Introduction. Insect colonisation and mass production. (ed. Smith C.), Academic press, USA, pp 2-14

Knox C.K., Moore S.D., Luke G.A. and Hill M.A., (2015), Baculovirus-based strategies for management of insect pests: a focus on development and application in South Africa. Biocontrol science and technology, volume 25, issue 1, pp 1-20

Kogan M., (1998), Integrated pest management: Historical perspectives and contemporary developments, Annual review of entomology, volume 43, pp 243-270

Koonin E.V., Aravind L. and Kondrashov A.S., (2000), The impact of comparative genomics on our understanding of evolution. Cell, volume 101, issue 6, pp 573-576

Kranthi K.R., Jadhav D., Wanjari R., Kranthi S. and Russell D., (2001), Pyrethroid resistance and mechanisms of resistance in field strains of *Helicoverpa armigera* (Lepidoptera: Noctuidae). Insecticide resistance and resistance management, volume 94, issue 1, pp 253-263

Kranthi K.R. and Russell D., (2009), Changing trends in cotton pest management. Integrated pest management: Volume 1: Innovation-development process, (ed. Peshin R. and Dhawan A.K.), Springer, USA. Pp 499-541

Kriticos D.J., Ota N., Hutchison W.D., Beddow J., Walsh T., Tay W.T., Borchert D.M., Paula-Moreas S.V., Czepak C. and Zalucki M.P., (2015), The potential of distribution of invading *Helicoverpa armigera* in North America: Is it just a matter of time? PLOS ONE, DOI: 10.1371/journal.pone.0119618

Kumar N.S., Murugan K. and Zhang W., (2008), Additive interaction of Nucleopolyhedrovirus and azadirachtin, BioControl, volume 53, pp 869-880

Kumar M., Gupta G.P. and Rajam M.V., (2009), Silencing of acetylcholinesterase gene of *Helicoverpa armigera* by siRNA affects larval growth and its life cycle. Journal of insect pathology, volume 55, issue 3, pp 273-278

Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T., Slavicek J.M. and Rohrmann G.F., (1999), Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantia dispar*. Virology, volume 253, issue 1, pp 17-34

Lacey L.A., Frutos R., Kaya H.K. and Vail P., (2001), Insect pathogens as biological control agents: do they have a future? Biological control, volume 21, pp 230-248

Lacey L.A., (2017), Entomopathogens used as microbial control agents. Microbial control of insect and mite pests: From theory to practice. (ed Lacey L.A.), Academic press, USA. Pp 3-12

Lakdawalla A. and Vansteenhouse H., (2008), Illumina genome analyser II system. Next generation genome sequencing: Towards personalised medicine. (ed. Janitz M.), Wiley-VCH Verlag GmbH & Co. KGaA, pp 13-28

Lammers J.W. and MacLeod A., (2007), Report of a pest risk analysis Helicoverpa armigera (Hübner, 1808). Plant protection service (NL) and central science laboratory (UK) joint pest risk analysis for *Helicoverpa armigera*, pp 1-18

Lange M., Wang H., Zhihong H. and Jehle J.A., (2004), Towards molecular identification and classification system of lepidopteran-specific baculoviruses. Virology, volume 325, issue 1, pp 36-47

Lapointe R., Thumbi D. and Lucarotti C.J., (2012), Recent advances in our knowledge of baculovirus molecular biology and its relevance for the registration of baculovirus-based products for insect population control. Integrated pest management and pest control – current and future tactics, (ed. Soleneski S.,), InTech, UK, pp 481-522

Lingappa S. and Hedge R., (2001), Exploitation of biocontrol potential in the management of insect pests of pulse crops. Biocontrol potential and its exploitation in sustainable in agriculture volume 2, insect pests (ed. Ipadhyay R.K., Mukerji K.G. and Chamola B.P.), Springer science, New York, pp 321-344

Liu S., Li H., Sivakumar S. and Bonning B.C., (2006), Virus-derived genes for insect resistant transgenic plants. Advances in virus research: Insect viruses: Biotechnological applications volume 68. (ed. Bonning B.C, Maramorosch K. and Shatkin A.J.), Academic press, USA, pp 428-459

Liu Z., Li D., Gong P. and Wu K., (2004), Life table studies of the cotton bollworm, *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae), on different host plants. Population ecology, volume 33 issue 6, pp 1570-1576

Lu Y-X and Wie-Hua X., (2010), Proteomic and phosphoproteomic analysis at diapause initiation in the cotton bollworm, *Helicoverpa armigera*. Journal of proteome research, volume 9, issue 10, pp 5053-5064

Lua L.H.L., Pedrini M.R.S., Reid S., Robertson A. and Tribe D.E., (2002), Phenotypic and genotypic analysis of *Helicoverpa armigera* nucleopolyhedrovirus serially passaged in cell culture. Journal of general virology, volume 83, issue 4, pp 945-955

Madigan M.T., Martinko J.M., Stahl D.A. and Clark D.P., (2012), Brock biology of microorganisms, 13th edition, Benjamin Cummings, Boston

Maiti R. and Kumari A., (2016), Biotic stress: Insects, Bioresource and stress management, (ed. Maiti R., Kumari A., Thakur A.K. and Sarkar N.C.), Springer nature, Singapore, pp 101-130

Marzban R., He Q., Liu X. and Zhang Q., (2009), Effects of Bacillus thuringiensis toxin Cry1Ac and cytoplasmic polyhedrosis virus of Helicoverpa armigera (Hübner) (HaCPV) on cotton bollworm (Lepidoptera: Noctuidae), Journal of Invertebrate pathology, volume 101, issue 1, pp 71-76

Mazid M., Khan T.A. and Mohammad F., (2011), Role of secondary metabolites in defence of mechanisms of plants. Biology and medicine, volume 3, pp 232-249

McCaffrey, (1998), Resistance to insecticides in Heliothine Lepidoptera: a global view. Philosophical transactions of the royal society B: biological sciences, volume 353, issue 1376, pp 1735-1750

Men A.E., Wilson P., Siemering K. and Forrest S., (2008), Sanger DNA sequencing. Next generation genome sequencing: Towards personalised medicine. (ed. Janitz M.), Wiley-VCH Verlag GmbH & Co. KGaA, pp 1-11

Mensah R.K and Pyke B.A., (2007) Cotton: Beneficial's. Pests of field crops and pastures: identification and control. (ed. Bailey P.T.), CSIRO publishing, Australia, pp 102-121

Miele S.A.B., Guravaglia M.J., Belaich M.N. and Ghiringhelli P.D. (2011), Baculoviruses: molecular insight on their diversity and conservation (review). International journal of evolutionary biology: http://dx.doi.org/10.4061/2011/379424

Miller L.K., (1997), Introduction to baculoviruses. The baculoviruses. (ed. Miller L.K.), Plenum Press, New York, pp 1-7

Mironidis G.K. and Savopoulou-Soultani M., (2008), Development, survivorship, and reproduction of *Helicoverpa armigera* (Lepidoptera: Noctuidae) under constant and alternating temperatures. Environmental entomology, volume 37, issue 1, pp 16-28

Mitchell P.D. and Hutchison W.D., (2009), Decision making and economic risk in IPM. Integrated pest management: Concepts, techniques, strategies and case studies. (ed. Radcliffe E.B., Hutchison W.D. and Cancelado R.E.). Cambridge University press, UK, pp 33-51

Mohamed M.A., Coppel H.C and Podgwaite J.D., (1982), Persistence in soil and on foliage of nucleopolyhedrosis virus of the European pine sawfly, *Neodiprion sertifer* (Hymenoptera: Diprionidae). Environmental entomology, volume 11, issue 5, pp 1116-1118

Moore S.D., Pittaway T., Bouwer G and Fourie J.G., (2004), Evaluation of *Helicoverpa armigera* (Lepidoptera: Noctuidae) on citrus. Biocontrol science and technology, volume 14, issue 3, pp 239-250

Moore S.D. and Kirkman W., (2010), Helicovir™: a virus for the biological control of bollworm. SA fruit journal, volume 9, issue 4, pp 63-67

Moran V.C., (1983), The phytophagous insects and mites of cultivated plants in South Africa: patterns and pest status. Journal of applied ecology, volume 20, pp 439-450

Murphy C.I., Piwnica-Worms H., Grünwald S., Romanow W.G., Francis N. and Fan H-Y., (2004), Overview of the baculovirus expression system. Current protocols in molecular biology, volume 65, issue 1, pp 1-10, doi.org/10.1002/0471142727.mb1609s65

Nawaz M., Mabubu J.I. and Hua H. (2016), Current status and advancement of biopesticides: Microbial and botanical pesticides, Journal of Entomology and Zoology studies, volume 4, issue 2, pp 241-246

Noune C. and Hauxwell C., (2016), Comparative analysis of HaSNPV-AC53 and derived strains. Viruses, volume 31, issue 8, pp 1-18, doi:10.3390/v811028

Ogembo J.V., Chaeychomsri S., Kamiya K., Ishikawa H., Katou Y., Ikeda M. and Kobayashi M., (2007), Cloning and comparative characterisation of nucleopolyhedrovirus isolated from African bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae) in different geographic regions. Journal of insect biotechnology and sericology, volume 76, issue 1, pp 39-49

Ogembo J.V., Caoili B.L., Shikata M., Chaeychomsri S., Kobayashi M. and 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, volume 39, issue 2, pp 261-272

Oluoch-Kasura W.A., Muriuki A.W., Ulubaya F.M. and Kilala D., (2013), Why pests and disease regulation should concern mankind. Ecosystems service and carbon sequestration in the biosphere. (ed. Lal R., Lorenz K., Hüttl R.F., Schneider B.U. and Braun J.V.), Springer, New York, pp 209-218

O'Reilly D. R. (1989). A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyltransferase. Science, volume 245, pp 1110-1112.

O'Reilly D. R. and Miller L. K. (1990). Regulation of expression of a baculovirus ecdysteroid UDP-glucosyltransferase gene. Journal of Virology, volume 64, pp 1321-1328.

O'Reilly D. R. and Miller L. K. (1991). Improvement of a baculovirus pesticide by deletion of the egt gene. Bio-Technology, volume 9, pp 1086-1089.

O'Reilly D., (1997), Auxiliary genes of baculoviruses. The baculoviruses, 1st edition. (ed. Miller L.K.), Springer science+business media, LLC, New York, pp 267-300

Opoku-Debrah J.K., Hill M.P., Knox C. and 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, volume 112, issue 3, pp 219-228

Opoku-Debrah J.K., Hill M.P., Knox C. and Moore S.D., (2014), Comparison of the biology of geographically distinct populations of the citrus pest, *Thaumatotibia leucotreta* (Meyrick) (Lepidoptera: Tortricidae), in South Africa. African entomology, volume 22, issue 3, pp 530-537

Opoku-Debrah J.K., Hill M.P., Knox C. and Moore S.D., (2016), Heterogeneity in virulence relationships between *Cryptophlebia leucotreta* granulovirus isolates and geographically

distinct host populations: lessons from codling moth resistance to CpGV-M. BioControl, volume 61, issue 4, pp 449-459

Osibanjo O., Bouwman H., Bashir N. H., Okond'Ahoka J., Choong Kwet Yve R. and Onyoyo, H. A. (2002). Regionally based assessment of persistent toxic substances. Sub-Saharan Africa regional report. Geneva: UNEP Chemicals.

Parker A.G., (2005), Mass-rearing for sterile insect release. Sterile insect technique: Principles and practice in area-wide integrated pest management, (ed. Dyck V.A., Hendrichs J. and Robinson A.S), Springer, Netherlands, pp 209-232

Patankar A.G., Giri A.P., Harsulkar A.M., Sainani M.N., Deshpande V.V., Ranjekar P.K. and Gupta V.S., (2001), Complexity in specificities and expression of *Helicoverpa armigera* gut proteinases explains polyphagous nature of the insect pest. Insect biochemistry and molecular biology, volume 31, issue 4-5, pp 453-464

Patthy L., (2008), Mutations. Protein evolution, 2nd edition. (ed. Patthy L.), Blackwell publishing, Australia, pp 45-62

Paul A.V.N., (2004), Use of biological control for insect pest management. Biopesticides for sustainable agriculture: Prospects and constraints. (ed. Kaushik N.), The energy and resource institute press, India, pp 143-172

Perkins L.E., Cribb B.W., Hanan J., Glaze E., Beveridge C. and Zalucki M.P., (2008), Where to from here? The mechanisms enabling the movement of first instar caterpillars on whole plants using *Helicoverpa armigera* (Hübner). Arthropod-plant interactions, volume 2, pp 197-207

Perkins L.E., Cribb B.W., Hanan J. and Zalucki M.P., (2009), The role of two plant-derived volatiles in the foraging movement of 1st instar *Helicoverpa armigera* (Hubner): time to stop and smell the flowers. Arthropod-plant interactions, volume 3, pp 173-179

Peshin R., Bandral R.S., Zhang W., Wilson L. and Dhawan A.K., (2009), Integrated pest management: A global overview of history, programs and adoption. Integrated pest management: Volume 1: Innovation-development process, (ed. Peshin R. and Dhawan A.K.), Springer, USA, pp 1-47

Pidre M.L., Ferrelli M.L., Haase S. and Romanowski V., (2013), Baculovirus display: a novel tool for vaccination. Biochemistry, Genetics and Molecular Biology, "Current Issues in

Molecular Virology – Viral Genetics and Biotechnological Applications, (ed. Romanowski V.), InTech, UK, pp 137-164

Plantwise knowledge bank, (2017), Nucleopolyhedrosis virus. Plantwise technical factsheet. http://www.plantwise.org/KnowledgeBank/Datasheet.aspx?dsid=46667

Possee R.D., Griffiths C.M., Hitchman R.B., Chambers A., Murguia-Meca F., Danguah J., Jeshtadi A. and King L.A., (2010), Baculoviruses: Biology, replication and exploitation. Insect virology. (ed. Asgari S. and Johnson K.), Caister academic press, pp 35-58

Prasad V. and Srivastava S., (2016), Insect viruses. Eco-friendly pest management for food security. (ed. Omkar), Academic Press, UK, pp 411-442

Pudijanto, Shepard B.M., Shapiro M., Jackson D.M. and Carner G.R., (2016), Comparative analysis of homologous and heterologous nucleopolyhedrovirus of beet armyworm larvae. Journal of agricultural and urban entomology, volume 31, issue 1, pp 13-24

Quinn L.P., de Vos B.J., Fernandes-Whaley M.C.R., Bouwman H., Kylin H. and van der Berg J., (2011), Pesticide use in South Africa: one of the largest importers of pesticide in Africa. Pesticides in the modern world, pesticide use and management. (ed. Stoytcheva M.), InTech, Pretoria, doi:DOI:10.5772/16995

Rachel M., (2008), Culture of natural enemies on factitious foods and artificial diets. Encyclopedia of entomology, second edition. (ed. Capinera J.L.), Springer science and business media, Germany, pp 1133-1136

Raghavendra A.T., Jalali S.K., Ojha R., Shivalingaswamy T.M. and Bhatnagar R., (2017), Whole genome sequence and comparative genomic sequence analysis of *Helicoverpa armigera* nucleopolyhedrovirus (HearNPV-L1) isolated from India. Virus diseases, pp 1-8: doi:10.1007/s13337-016-0352-6

Rao A.G. and Babu T.R., (2009), Biological control of arthropod pest of pulse crops. Integrated pest and disease management. (ed. Upadhyay R.K., Mukerji K.G., Chamola B.P. and Dubey O.P.), APH publishing, India, pp 157-168

Rao G.V.R., Kumar C.S., Sireesha K. and Kumar P.L., (2015), Role of nucleopolyhedrovirus (NPVs) in the management of lepidopteran pests in Asia. Biocontrol of lepidopteran pests: Use of soil microbes and their metabolites. (ed. Sree K.S. and Varma A.), Springer, Switzerland. pp 11-52

Reddy G.V.P. and Manjunatha M., (2000), Laboratory and field studies on the integrated pest management of Helicoverpa armigera (Hübner) in cotton, based on pheromone trap catch threshold level. Journal of applied entomology, vol 124, issue 5-6, pp 213-221

Reed W. and Pawar C.S., (1982), *Heliothis*: A global problem. In: Proceedings of the international workshop on *Heliothis* management 15-20 November, pp 9-14

Reid S., Chan L. and Oers M.M.v., (2014), Production of entomopathogenic viruses. Mass production of beneficial organisms: Invertebrates and entomopathogens. 1st edition, (ed. Morales-Ramos J.A., Rojas M.G. and Shapiro-Ilan D.I.), Academic press, USA, pp 437-482

Richards A.R. and Christian P.D., (1999), A rapid bioassay screen for quantifying Nucleopolyhedroviruses (Baculoviridae) in the environment. Journal of virological methods, volume 82, issue 1, pp 63-75

Rohrmann G., (2013), Selected baculovirus genes without orthologs in the AcMNPV genome: Conservation and function. Baculovirus molecular biology, 3rd edition. (ed. Rohrmann G.), National Library of Medicine (US), National Center for Biotechnology Information; http://www.ncbi.nlm.nih.gov/books/NBK114593/ pp47-64

Rowley D. L., Popham H. J. and Harrison R. L. (2011). Genetic variation and virulence of nucleopolyhedroviruses isolated worldwide from the heliothine pests *Helicoverpa armigera*, *Helicoverpa zea*, and *Heliothis virescens*. Journal of Invertebrate Pathology, volume 107 issue 2, pp 112-126

Russell D., (2004), Integrated pest management for insect pests of cotton in less developed countries. Insect pest management: Field and protected crops. (ed. Horowitz A.R. and Ishaaya I.), Springer-Verlag Berlin Heidelberg, New York, pp 141-180

Sahayaraj K. and Balasabramanian R., (2016), Field evaluation. Artificial rearing of Reduviid predators for pest management. (ed. Sahayaraj K. and Balasabramanian R), Springer nature, Singapore, pp 147-166

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

Sehgal, V. K., & Ujagir, R. (1990). Effect of synthetic pyrethroids, neem extracts and other insecticides for the control of pod damage by Helicoverpa armigera (Hubner) on chickpea

and pod damage-yield relationship at Pantnagar in northern India. Crop protection, volume 9, pp 29-32.

Sekulic, R., Tatjana, K., Masirevic, S., Vajgand, D., Gordana, F., & Radojcic, S. (2004). Incidence and damage of cotton bollworm (*Helicoverpa armigera* Hbn.) in Vojvodina Province in 2003. Biljni Lekar Plant Doctor, volume 32, issue 2, pp 113-124.

Shanower, T. G., & Romeis, J. (1999). Insect pests of pigionpea and their management. Annual Review of Entomology,volume 44, pp 77-96.

Sharma, H. C. (2005). *Heliothis/Helicoverpa* Management: Emerging Trends and Strategies for Future Research. New Delhi, India: Oxford and IBH Publishers.

Sharma H.C., (2005), *Helicoverpa* the global problem. *Heliothis/Helicoverpa* management: emerging trends and strategies for future research. (ed. Sharma H.C.), Oxford and IBH publishers, India, pp 1-6

Siegwart M., Graillot B., Lopez C.B., Besse S., Bardin M., Nicot P.C. and Lopez-Ferber M., (2015), Resistance to bio-insecticides or how to enhance their sustainability: a review. Frontiers in plant science. Volume 6, issue 381. Pp 1-20

Singh A., Bambawale O.M. and Sharma O.P., (2004), Entomopathogen biopesticides in integrated pest management. Biopesticides for sustainable agriculture: Prospects and constraints. (ed. Kaushik N.), The energy and research institute press, India, pp 41-66

Slack J. and Arif B.M., (2007), The baculovirus occlusion-derived virus: Virion structure and function. Advances in virus research, volume 69. (ed. Maramorosch K. and Shatkin A.J.), pp 100-167

Slavicek, J. M., Popham, H. J., & Riege, C. I. (1999). Deletion of the Lymantria dispar multicapsid nucleopolyhedrovirus ecdysteroid UDP-glucosyltransferase gene enhances viral killing speed in the last instar of the gypsy moth. Biological Control, volume 16, pp 91-103.

Smith-Pardo A.H., (2014), The old-world bollworm *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae: Heliothine) its biology, economic importance and its recent introduction into the western hemisphere. Volume 6, issue 1, pp 18-28

Sparks A.N. and Harrell E.A., (1976), Introduction. Corn earworm rearing and mechanisation. Technical bulletin number 1544. Agricultural research service: United States Department of Agriculture, pp 1-11

Sree K.S. and Varma A., (2015), An introduction to entomopathogenic microorganisms. Biocontrol of lepidopteran pests: Use of soil microbes and their metabolites. (ed. Sree K.S. and Varma A.), Springer, Switzerland. pp 1-10

Strickberger M.W., (2000), Proteins and the genetic code. Evolution, 3rd edition (Strickberger M.W.), Jones and Bartlett, UK, pp 139 - 157

Summers M.D., (2006), Milestones leading to the genetic engineering of baculoviruses as expression vector systems and viral pesticides. Advances in virus research: Insect viruses: Biotechnological applications volume 68. (ed. Bonning B.C.), academic press, USA, pp 4-75

Sun (2010), Nucleic extraction and amplification. Molecular diagnostics: techniques and applications for the clinical laboratory, first edition. (ed. Gordy W.W., Nakamura R.M., Kiechle F.L. and Strom C.), Academic Press Elsevier, USA, pp 35-48

Sun X., (2015), History and current status of development and use of viral insecticides in China, Viruses, volume 7 issue 1, pp 306-319

Sun X., Chen X., Zhang Z., Wang H., Bianchi F.J.J.A., Peng H., Vlak J.M. and Hu Z., (2002), Bollworm responses to release of genetically modified *Helicoverpa armigera* Nucleopolyhedroviruses in cotton. Journal of invertebrate pathology. Volume 81, issue 2, pp 63-39

Sun X., Wang H., Sun X., Chen X., Peng C., Pan D., Jehle J.A., van der Werf W., Vlak J.M. and Hu Z., (2004), Biological activity and field efficacy of a genetically modified *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus expressing an insect-selective toxin from a chimeric promoter. Biological control, volume 29, issue 1, pp 124-137

Sun X. and Peng H., (2007), Recent advances in biological control of pest insects by using viruses in China. Virologica sinica, volume 22, issue 2, pp 158-162

Szewczyk B., Hoyos-Carvajal L., Paluszek M., Skrzecz I. and de Souza M.L., (2006), Baculoviruses – re-emerging biopesticides. Biotechnology advances, volume 24, issue 2, pp 143-160

Tang P., Zhang H., Li Y., Han B., Wang G., Qin Q. and Zhang Z., (2012), Genomic sequencing and analyses of HearMNPV-a new Multinucleocapsid nucleopolyhedrovirus isolated from *Helicoverpa armigera*. Virology Journal. Volume 9, issue 168, pp 1-18

Tay W.T., Soria M.F., Walsh T., Thomazoni D., Silvie P., Behere G.T., Anderson C. and Downes S., (2013), A brave new world for an old-world pest: *Helicoverpa armigera* (Lepidoptera: Noctuidae) in Brazil. PLoS One, volume 8, issue 11, pp 1-71

Thézé J., Lopez-Vaamonde C., Cory J.S. and Herniou E.A., (2018), Biodiversity, evolution and ecological specialisation of baculoviruses: a treasure trove for future applied research. Viruses, volume 10, issue 7, doi.org/10.3390/v10070366

Tisdall J.D., (2001), Biology and computer science. Beginning Perl for bioinformatics, 1st edition. (ed. LeJuene L.), O'Reilly media, CA, USA, pp 1-11

ToxRat® (2015), The evaluation. Toxicity response analysis and testing user manual. Toxrat solutions GmbH, Alsdorf, pp 22-76

van Lenteren J.C., (2000), Success in biological control of arthropods by augmentation of natural enemies. Biological control: measures of success. (ed. Gurr G., Wratten S.D. and Wratten S), Kluwer academic publishers, Netherlands pp 77-104

Vaissayre M. and Cauquil J., (2000), Bollworms. Main pests and diseases of cotton in Sub-Saharan Africa. (ed. Vaissayre M. and Cauquil J.). CIRAD-CTA, Netherlands, pp 22-31

Villareal L.P., (1999), DNA virus contribution to host evolution. Origin and evolution of viruses. (ed. Domingo E., Webster R. and Holland J.), Academic press, UK, pp 391-420

Virto C., Williams T., Navarro D., Tellez N.M., Murillo R. and Caballero P., (2017), Can mixtures of horizontally and vertically transmitted nucleopolyhedrovirus genotypes be effective for biological control of *Spodoptera exigua*? Journal of pest science, volume 90 issue 1, pp 331-343

Walter G.H., (2014), Pest management as an applied science. Insect pest management and ecological research, 1st edition, Content technologies, pp 1-13

Whitlock V.H., (1974), Symptomatology of two virus's infection *Heliothis armigera*. Journal of invertebrate pathology, volume 23, issue 1, pp 70-75

Whitlock V.H., (1978), Dosage-mortality studies of a granulosis and nuclear polyhedrosis virus of a laboratory strain of *Heliothis armigera*. Journal of invertebrate pathology, volume 14, issue 1, pp 386-387

Wilson K. and Walker J., (2005), Protein structure, purification, characterisation and function analysis. Principles and techniques of biochemistry and molecular biology, 6th edition. (Wilson K. and Walker J.), Cambridge university press, USA, pp 349 - 404

Windley M.J., Herzig V., Dziemborowicz S.A., Hardy M.C., King G.F. and Nicholson G.M., (2012), Spider-venom peptides as bio-insecticides. Toxins, volume 4, pp 191-227

Xu X., Yu L. and Wu Y., (2005), Disruption of the cadherin gene associated with resistance to Cry1Ac δ-endotoxin of *Bacillus thuringiensis* in *Helicoverpa armigera*. Applied and environmental microbiology, volume 71, issue 2, pp 948-954

Zhang C-X., Ma X-C. and Guo Z-J., (2005), Comparison of the complete genome sequence between C1 and G4 isolates of the *Helicoverpa armigera* single nucleopolyhedrovirus. Virology, volume 333, issue 1, pp 190-199

Zhu Y.J., (2005), Introduction. Efficiency of *Bacillus thuringiensis* with abamectin on different agricultural pests and their natural enemies. (ed. Sengoca C.), Cuvillier Verlogg Gőttingeng, Germany, pp 1-3

Zhuang M. and Gill S.S., (2003), Mode of action of *Bacillus thuringiensis* toxins. Chemistry of crop protection. (ed. Voss G. and Ramos G), Wiley VCH, Weinheim, Germany, pp 213-234

9 Appendix

9.1 Appendix A

The PCR sequence data for HearNPV- Albany and HearNPV-KZN *polyhedrin* region after assembly from the raw data.

9.1.1 PCR sequence for HearNPV-Albany

GAAGATCCTTTTTTGGGACCCGGCAAAAATCAAAAATTAACTTTGTTTAAAGAAA
TTCGCAGCGTTAAGCCCGACACAATGAAGCTTGTAGTTAACTGGAGCGGTCGCG
AATTTCTTCGCGAAACTTGGACGCGTTTCATGGAAGACAGTTTTCCCATTGTAAA
CGACCAAGAAATTATGGACGTGTTTCTGTCTGTTAATATGCGACCAACCAAACCG
AACCGTTGTTACCGATTCTTAGCGCAACACGCTCTGCGTTGTGATCCCGACTATA
TTCCTCACGAAGTCATTCGTATTGTAGAACCTTCCTATGTAGGCAGTAACAACGA
GTACAGAATTAGTTTAGCCAAAAAATACGGCGGTTGTCCCGTTATGAACTTGCAC
GCTGAATACACTAATTCCTTTGAAGATTTCATTACCAACGTAATTTGGGAGAACT
TTTACAAACCAATTGTTTACGTAGGCACTGATTCTGCCGAAGAAGAGGAAATACT
CCTAGAAGTTTCTTTGATATTTAAGATCAAAGAATTTGCACCTGACGCTCCGCTA
TACACTGGTCCTGCATATTA

9.1.2 PCR sequence for HearNPV-KZN

GAAGGCCCTTTTTTGGGACCTGGCAAAAATCAAAAACTATCTTTGTTTAAAGAGA
TTCGCAGCGTTAACCCCGACACAAGGAAGCTTGTAGTTAACGGGAGGGGTCGCG
AATTTCTTCGCGAAACTTGGACCCGTTTCATGGAAGACAGTTTTCCCATTGTAAA
CGACCAAGAAATTATGGACGTGTTTCTGTCTGTTAATATGCGACCAACCAAACCG
AACCGTTGTTACCGATTCTTAGCGCAACACGCTCTGCGTTGTGATCCCGACTATA
TTCCTCAAGAAGTCATTCGTTTTGAAAACCCTTCCTTGGGGGGCAGTAACAACGA
GTACAAAATTATTTTACCCAAAAAATCGGGGGGTGGCCCCGTTATAAATTGGCCC
GCCAAAAACCCTATTTCCTTTAAAAATTTYTTTCCCAGATGATTTGGGGAAAACT
TCTACAACCCATTTGTTTACGTAGGCCCTGATTCTGCCAAAAAAAAGAAAATCCC
CCTAGGGGTTTTTTTGATTTTAAAAATCAAAAAATTCCCCCCGGGCGATACGCTT
TACGCGG

9.2 Appendix B

9.2.1 HearNPV-Albany ORF alignment against HearNPV-Au

ORF comparison of the full HearNPV-Albany genome against the HearNPV-Au genome, showing percentage identity and ORF size and positions (Table 9.1). Replication genes were highlighted in green, transcription genes were highlighted in purple, structural genes were highlighted in orange, auxiliary genes were highlighted in blue and genes with unknown function were highlighted in grey. Genes that are unique to lepidopteran NPVs were labelled red, genes specific to some NPVs including HearNPV were labelled in yellow. Hypothetical proteins were labelled in black with no name label.

The comparison of HearNPV-KZN ORFs and HearNPV-Au ORFs showing all the percentage identities and sizes of the ORFs (Table 9.2). Replication genes were highlighted in green, transcription genes were highlighted in purple, structural genes were highlighted in orange, auxiliary genes were highlighted in blue and genes with unknown function were highlighted in grey. Genes that are unique to lepidopteran NPVs were labelled red, genes specific to some NPVs including HearNPV were labelled in yellow. Hypothetical proteins were labelled in black with no name label.

Table 9.1: From nucleotide alignment, potentially expressed ORFs and their positions are shown with the percentage nucleotide identity for HearNPV-Albany genome versus HearNPV-Au for the ORFs. ORFs with numbers and no names represent potential hypothetical proteins.

HearNPV-Albany NPV					**HearNPV-Au NPV**					
ORF position and Name	**Start**	**Stop**	**Size**	**Direction**	**ORF position and Name**	**Start**	**Stop**	**Size**	**Direction**	**% identity**
1.polyhedrin	1	740	740	forward	1.polyhedrin	1	741	741	forward	98.52
2.	737	1,972	1,236	reverse	2.	738	1,979	1,242	reverse	98.63
3.pk1	1,987	2,790	804	forward	3.pk1	1,994	2,797	804	forward	99.75
4.hoar	2,915	5,194	2,280	reverse	4.hoar	2,920	5,181	2,262	reverse	93.52
5.	5,377	5,551	175	forward	5.	5,377	5,556	180	forward	95.56
6.	5,723	6,580	858	forward	6.	5,724	6,575	852	forward	99.07
7.	6,793	6,954	162	reverse	7.	6,787	6,942	156	reverse	96.30
8.ie0	6,942	7,799	858	forward	8.ie0	6,930	7,787	858	forward	99.65
9.p49	7,816	9,222	1,407	forward	9.p49	7,804	9,210	1,407	forward	99.43
10.odv-e18	9,233	9,478	246	forward	10.odv-e18	9,221	9,466	246	forward	99.59
11.odv-e27	9,493	10,347	855	forward	11.odv-e27	9,481	10,335	855	forward	99.06
12.	10,392	10,670	279	forward	12.	10,381	10,659	279	forward	99.64
13.	10,697	11,308	612	reverse	13.	10,686	11,297	612	reverse	99.84

14.ie-1	11,350	13,317	1,968	forward	14.ie-1	11,339	13,306	1,968	forward	99.49
15.odv-e56	13,371	14,435	1,065	reverse	15.odv-e56	13,359	14,423	1,065	reverse	98.78
16.me53	14,596	15,675	1,080	forward	16.me53	14,584	15,663	1,080	forward	99.26
17.	15,678	15,845	168	forward	17.	15,666	15,833	168	forward	100.00
18.	15,898	16,179	282	reverse	18.	15,886	16,167	282	reverse	98.58
19.p74	16,200	18,266	2,067	forward	19.p74	16,188	18,254	2,067	forward	99.08
20.p10	18,319	18,582	264	reverse	20.p10	18,308	18,571	264	reverse	98.48
21.p26	18,665	19,468	804	reverse	21.p26	18,654	19,457	804	reverse	99.00
22.	19,581	19,784	204	forward	22.	19,570	19,773	204	forward	100.00
23.lef-6	19,860	20,423	564	reverse	23.lef-6	19,849	20,412	564	reverse	99.47
24.dbp	20,437	21,408	972	reverse	24.dbp	20,426	21,397	972	reverse	99.59
25.	21,627	22,028	402	forward	25.	21,617	22,018	402	forward	98.51
26.	23,611	24,003	393	forward						N/A
27.	24,521	25,274	754	reverse	26.	24,340	25,107	768	reverse	82.49
28.ubiquitin	25,052	25,303	252	forward	27.ubiquitin	24,947	25,198	252	forward	66.01
29.	25,368	25,873	506	forward	28.	25,262	25,768	507	forward	75.49
30.lese25-like protein	25,892	26,471	580	forward	29.lese25-like protein	25,788	26,360	573	forward	90.10

31.pp31	26,531	27,469	939	reverse	30.pp31	26,419	27,354	936	reverse	81.99
32.lef-11	27,432	27,818	387	reverse	31.lef-11	27,320	27,703	384	reverse	90.10
33.	27,787	28,499	713	reverse	32.	27,672	28,388	717	reverse	89.81
34.	28,734	29,813	1,080	forward	33.	28,620	29,699	1,080	forward	94.94
35.p47	29,887	31,130	1,244	reverse	34.p47	29,767	31,005	1,239	reverse	94.56
36.	31,203	31,874	672	forward	35.	31,078	31,749	672	forward	99.11
37.	31,960	32,202	243	forward	36.	31,835	32,077	243	forward	99.59
38.lef-8	32,199	34,904	2,706	reverse	37.lef-8	32,074	34,779	2,706	reverse	99.37
39.	34,960	35,540	581	forward	38.	34,832	35,410	579	forward	99.14
40.	35,681	35,833	153	forward	39.	35,551	35,703	153	forward	96.08
41.chitinase	35,841	37,613	1,773	reverse	40.chitinase	35,711	37,480	1,770	reverse	99.38
42.	37,528	37,719	192	forward						N/A
43.	37,650	38,192	543	reverse	41.	37,524	38,066	543	reverse	99.63
44.	38,309	38,719	411	forward	42.	38,184	38,594	411	forward	100.00
45.	38,726	39,865	1,140	reverse	43.	38,601	39,737	1,137	reverse	98.86
46.	39,873	40,100	228	reverse	44.	39,745	39,972	228	reverse	100.00
47.lef-10	40,060	40,275	216	forward	45.lef-10	39,932	40,147	216	forward	100.00

48.vp1054	40,148	41,203	1,056	forward	46.vp1054	40,020	41,075	1,056	forward	99.43
49.	41,323	41,529	207	forward	47.	41,195	41,401	207	forward	99.52
50.	41,530	41,724	195	forward	48.	41,402	41,596	195	forward	98.97
51.	42,004	42,523	520	forward	49.	41,875	42,390	516	forward	98.85
52.	42,573	43,049	477	reverse	50.	42,441	42,923	483	reverse	96.48
53.	43,061	43,327	267	reverse	51.	42,935	43,201	267	reverse	100.00
54.fp	43,539	44,192	654	reverse	52.fp	43,413	44,066	654	reverse	99.08
55.	44,364	44,549	186	forward	53.	44,238	44,423	186	forward	100.00
56.lef-9	44,675	46,234	1,560	forward	54.lef-9	44,545	46,104	1,560	forward	99.62
57.cathepsin-like cysteine proteinase	46,318	47,421	1,104	reverse	55.cathepsin-like cysteine proteinase	46,188	47,291	1,104	reverse	98.91
58.	47,462	48,049	588	reverse	56.	47,332	47,919	588	reverse	98.81
59.gp37	48,120	48,959	840	reverse	57.gp37	47,990	48,829	840	reverse	98.81
					58. bro-a	49,980	51,629	1,650	forward	N/A
60.he65	51,117	51,827	711	forward	59.he65	52,385	53,095	711	forward	99.02
61.iap-2	51,904	52,656	753	reverse	60iap-2	53,172	53,924	753	reverse	99.34
62.	52,704	53,528	825	reverse	61.	53,972	54,796	825	reverse	99.27

63.	53,497	53,898	402	reverse	62.	54,765	55,166	402	reverse	99.75
64.lef-3	53,918	55,057	1,140	forward	63.lef-3	55,186	56,325	1,140	forward	98.68
65.	55,164	57,521	2,358	reverse	64.	56,433	58,790	2,358	reverse	99.36
66.DNA polymerase	57,555	60,617	3,063	forward	65.DNA polymerase	58,821	61,883	3,063	forward	99.09
67.	60,694	61,152	459	reverse	66.	61,960	62,418	459	reverse	99.13
68.	61,218	61,601	384	reverse	67.	62,484	62,867	384	reverse	99.74
69.	61,607	61,864	258	reverse	68.	62,873	63,130	258	reverse	99.61
70.vlf-1	61,905	63,152	1,248	reverse	69.vlf-1	63,171	64,415	1,245	reverse	99.60
71.	63,165	63,497	333	reverse	70.	64,428	64,760	333	reverse	99.40
72.gp41	63,566	64,534	969	reverse	71.gp41	64,829	65,797	969	reverse	99.07
73.	64,464	65,189	726	reverse	72.	65,727	66,452	726	reverse	99.72
74.	65,062	65,739	678	reverse	73.	66,325	67,002	678	reverse	99.56
75.vp91	65,669	68,119	2,451	forward	74.vp91	66,932	69,382	2,451	forward	99.10
76.cg30	68,248	69,099	852	reverse	75.cg30	69,510	70,361	852	reverse	99.06
77.vp39	69,188	70,073	886	reverse	76.vp39	70,450	71,331	882	reverse	99.44
78.lef-4	70,072	71,457	1,386	forward	77.lef-4	71,330	72,715	1,386	forward	99.78
79.	71,510	72,275	766	reverse	78.	72,768	73,532	765	reverse	99.35

80.	72,277	72,765	489	forward	79.	73,534	74,022	489	forward	99.59
81.odv-e25	72,811	73,506	696	forward	80.odv-e25	74,068	74,760	693	forward	99.14
82.	73,538	74,038	501	reverse	81.	74,792	75,289	498	reverse	98.60
83.DNA helicase	74,057	77,818	3,762	reverse	82.DNA helicase	75,308	79,069	3,762	reverse	99.47
84.	77,775	78,296	522	forward	83.	79,026	79,547	522	forward	100.00
85.	78,355	79,320	966	reverse	84.	79,606	80,571	966	reverse	99.48
86.lef-5	79,216	80,163	948	forward	85.lef-5	80,467	81,414	948	forward	99.58
87.p6.9	80,157	80,486	330	reverse	86.p6.9	81,408	81,737	330	reverse	99.70
88.	80,551	81,660	1,110	reverse	87.	81,802	82,911	1,110	reverse	99.28
89.	81,706	82,074	369	reverse	88.	82,957	83,325	369	reverse	99.19
90.	82,074	83,207	1,134	reverse	89.	83,325	84,458	1,134	reverse	99.60
91.vp80	83,302	85,119	1,818	forward	90.vp80	84,553	86,370	1,818	forward	98.95
92.	85,116	85,292	177	forward	91.	86,367	86,543	177	forward	98.87
93.	85,307	86,392	1,086	forward	92.	86,558	87,643	1,086	forward	99.45
94.	86,438	86,722	285	forward	93.	87,688	87,972	285	forward	100.00
95.odv-e66	86,789	88,807	2,019	reverse	94.odv-e66	88,039	90,057	2,019	reverse	99.90
96.	88,828	89,658	831	reverse	95.	90,078	90,908	831	reverse	98.92

97.	91,916	92,515	600	forward	96.	93,508	94,107	600	forward	99.00
98.	92,305	92,664	360	forward	97.	94,111	94,467	357	forward	52.78
99.	92,759	94,291	1,533	forward	98.	94,563	96,095	1,533	forward	99.48
100.	94,370	95,131	762	forward	99.	96,174	96,935	762	forward	99.87
101.	95,146	95,478	333	forward	100.	96,950	97,282	333	forward	99.70
102.iap-3	95,537	96,343	807	reverse	101.iap-3	97,340	98,146	807	reverse	98.76
103.	96,340	96,495	156	reverse	102.	98,143	98,298	156	reverse	99.36
104.bro-b	96,600	98,105	1,506	reverse	103.bro-b	98,409	99,914	1,506	reverse	99.27
105.sod	98,273	98,752	480	forward	104.sod	100,082	100,561	480	forward	99.38
106.	98,759	100,132	1,374	forward	105.	100,568	101,941	1,374	forward	99.45
107.	100,185	100,763	579	reverse	106.	101,994	102,572	579	reverse	100.00
108.	100,932	101,288	357	forward	107.	102,742	103,098	357	forward	98.88
109.	101,299	101,565	267	forward	108.	103,109	103,375	267	forward	99.63
110.	101,635	103,221	1,587	forward	109.	103,443	105,029	1,587	forward	99.18
111.	103,218	103,454	237	forward	110.	105,026	105,262	237	forward	99.58
112.fgf	103,477	104,382	906	reverse	111.fgf	105,285	106,190	906	reverse	99.78
113.alk-exo	104,509	105,796	1,288	reverse	112.alk-exo	106,318	107,604	1,287	reverse	99.46

114.	105,816	106,206	391	reverse	113.	107,624	108,013	390	reverse	98.98
115.	108,013	108,939	927	reverse	114.	109,693	110,619	927	reverse	99.14
116.lef-2	109,140	109,355	216	forward	115.lef-2	110,820	111,035	216	forward	99.07
117.p24	110,561	111,307	747	forward	116.p24	112,229	112,975	747	forward	99.46
118.gp16	111,369	111,659	291	forward	117.gp16	113,037	113,327	291	forward	97.94
119.calyx protein	111,705	112,727	1,023	forward	118.calyx protein	113,379	114,401	1,023	forward	99.22
120.	112,806	113,270	465	forward	119.	114,480	114,944	465	forward	100.00
121.	113,239	113,395	157	forward						N/A
122.	113,401	113,991	591	forward	120.	115,075	115,665	591	forward	99.49
123. 38.7 kDa protein	114,035	115,204	1,170	reverse	121. 38.7 kDa protein	115,709	116,878	1,170	reverse	97.28
124.lef-1	115,206	115,943	738	reverse	122.lef-1	116,880	117,617	738	reverse	99.05
125.	115,918	116,352	435	reverse	123.	117,592	118,020	429	reverse	97.24
126.egt	116,497	118,044	1,548	forward	124.egt	118,165	119,712	1,548	forward	99.22
127.	118,251	118,830	580	forward	125.	119,912	120,490	579	forward	98.10
128.	118,781	119,581	801	forward	126.	120,441	121,241	801	forward	98.63
129.	119,662	122,505	2,844	reverse	127.	121,322	124,165	2,844	reverse	99.16

130.pkip-1	122,898	123,407	510	forward	128.pkip-1	124,571	125,080	510	forward	99.61
131.arif-1	123,474	124,271	798	reverse	129.arif-1	125,147	125,944	798	reverse	99.62
132.	124,534	125,685	1,152	forward	130.	126,205	127,356	1,152	forward	99.39
133.	125,726	127,759	2,034	reverse	131.	127,397	129,430	2,034	reverse	99.07
134.	127,903	128,448	546	reverse	132.	129,572	130,117	546	reverse	99.45
135.	128,642	129,229	588	forward	133.	130,299	130,886	588	forward	98.47

Table 9.2: From nucleotide alignment, potentially expressed ORFs and their positions are shown with the percentage nucleotide identity for HearNPV-KZN genome versus HearNPV-Au for the ORFs. ORFs with numbers and no names represent potential hypothetical proteins.

HearNPV-KZN					**HearNPV-Au**					
ORF position and Name	**Start**	**Stop**	**Size**	**Direction**	**ORF position and Name**	**Start**	**Stop**	**Size**	**Direction**	**% identity**
1.polyhedrin	1	741	741	forward	1.polyhedrin	1	741	741	forward	99.60
2.	738	1,970	1,233	reverse	2.	738	1,979	1,242	reverse	98.31
3.protein kinase	1,985	2,788	804	forward	3.protein kinase	1,994	2,797	804	forward	99.25
4.haor	2,913	5,183	2,271	reverse	haor	2,920	5,181	2,262	reverse	95.89
5.	5,379	5,558	180	forward	5.	5,377	5,556	180	forward	100.00
6.	5,726	6,583	858	forward	6.	5,724	6,575	852	forward	97.32

7.	6,795	6,950	156	reverse	7.	6,787	6,942	156	reverse	100.00
8.ie-0	6,938	7,795	858	forward	8.ie-0	6,930	7,787	858	forward	99.30
9.p49	7,812	9,218	1,407	forward	9.p49	7,804	9,210	1,407	forward	99.43
10.odv-e18	9,229	9,474	246	forward	10.odv-e18	9,221	9,466	246	forward	100.00
11.odv-ec27	9,489	10,343	855	forward	11.odv-ec27	9,481	10,335	855	forward	98.71
12.	10,387	10,665	279	forward	12.	10,381	10,659	279	forward	99.64
13.	10,692	11,303	612	reverse	13.	10,686	11,297	612	reverse	99.84
14.ie-1	11,369	13,330	1,962	forward	14.ie-1	11,339	13,306	1,968	forward	78.37
15.odv-e56	13,383	14,447	1,065	reverse	15.odv-e56	13,359	14,423	1,065	reverse	99.15
16.me53	14,608	15,687	1,080	forward	16.me53	14,584	15,663	1,080	forward	99.72
17.	15,690	15,857	168	forward	17.	15,666	15,833	168	forward	100.00
18.	15,910	16,191	282	reverse	18.	15,886	16,167	282	reverse	98.94
19.p74	16,212	18,278	2,067	forward	19.p74	16,188	18,254	2,067	forward	99.56
20.p10	18,332	18,595	264	reverse	20.p10	18,308	18,571	264	reverse	100.00
21.p26	18,678	19,481	804	reverse	21.p26	18,654	19,457	804	reverse	99.38
22.	19,594	19,797	204	forward	22.	19,570	19,773	204	forward	99.51
23.lef-6	19,873	20,436	564	reverse	23.lef-6	19,849	20,412	564	reverse	99.82

24.dbp	20,450	21,421	972	reverse	24.dbp	20,426	21,397	972	reverse	100.00
25.	21,640	22,041	402	forward	25.	21,617	22,018	402	forward	100.00
26.	24,272	25,039	768	reverse	26.	24,340	25,107	768	reverse	99.87
27.ubiquitin	24,879	25,130	252	forward	27.ubiquitin	24,947	25,198	252	forward	100.00
28.	25,194	25,700	507	forward	28.	25,262	25,768	507	forward	100.00
29.lese25-like protein	25,720	26,292	573	forward	29.lese25-like protein	25,788	26,360	573	forward	98.95
30.pp31	26,351	27,289	939	reverse	30.pp31	26,419	27,354	936	reverse	99.57
31.lef-11	27,255	27,638	384	reverse	31.lef-11	27,320	27,703	384	reverse	100.00
32.	27,607	28,323	717	reverse	32.	27,672	28,388	717	reverse	99.16
33.	28,554	29,633	1,080	forward	33.	28,620	29,699	1,080	forward	99.54
34.p47	29,708	30,946	1,239	reverse	34.p47	29,767	31,005	1,239	reverse	99.52
35.	31,019	31,690	672	forward	35.	31,078	31,749	672	forward	99.40
36.	31,776	32,018	243	forward	36.	31,835	32,077	243	forward	100.00
37.lef-8	32,015	34,720	2,706	reverse	37.lef-8	32,074	34,779	2,706	reverse	99.48
38.	34,773	35,357	585	forward	38.	34,832	35,410	579	forward	98.12
39.	35,498	35,650	153	forward	39.	35,551	35,703	153	forward	98.04
40.chitinase	35,658	37,427	1,770	reverse	40.chitinase	35,711	37,480	1,770	reverse	99.49

41.	37,464	38,006	543	reverse	41.	37,524	38,066	543	reverse	99.82
42.	38,122	38,532	411	forward	42.	38,184	38,594	411	forward	100.00
43.	38,539	39,675	1,137	reverse	43.	38,601	39,737	1,137	reverse	99.47
44.	39,683	39,910	228	reverse	44.	39,745	39,972	228	reverse	100.00
45.lef-10	39,870	40,085	216	forward	45.lef-10	39,932	40,147	216	forward	100.00
46.vp1054	39,958	41,013	1,056	forward	46.vp1054	40,020	41,075	1,056	forward	99.05
47.	41,133	41,339	207	forward	47.	41,195	41,401	207	forward	100.00
48.	41,340	41,534	195	forward	48.	41,402	41,596	195	forward	99.49
49.	41,814	42,333	520	forward	49.	41,875	42,390	516	forward	98.85
50.	42,384	42,851	468	reverse	50.	42,441	42,923	483	reverse	96.07
51.	42,863	43,129	267	reverse	51.	42,935	43,201	267	reverse	100.00
52.fp	43,341	43,994	654	reverse	52.fp	43,413	44,066	654	reverse	98.32
53.	44,166	44,351	186	forward	53.	44,238	44,423	186	forward	99.46
54.lef-9	44,465	46,024	1,560	forward	54.lef-9	44,545	46,104	1,560	forward	99.62
55.cathepsin-like cysteine proteinase	46,108	47,213	1,106	reverse	55.cathepsin-like cysteine proteinase	46,188	47,291	1,104	reverse	50.45
					56.	47,332	47,919	588	reverse	N/A

					57.gp37	47,990	48,829	840	reverse	N/A
56.bro-a	49,240	50,891	1,652	forward	58.bro-a	49,980	51,629	1,650	forward	78.69
57.he65	52,655	53,365	711	forward	59.he65	52,385	53,095	711	forward	99.58
58.iap-2	53,441	54,193	753	reverse	60.iap-2	53,172	53,924	753	reverse	99.07
59.	54,241	55,065	825	reverse	61.	53,972	54,796	825	reverse	99.15
60.	55,034	55,435	402	reverse	62.	54,765	55,166	402	reverse	99.50
61.lef-3	55,455	56,594	1,140	forward	63.lef-3	55,186	56,325	1,140	forward	98.86
62.	56,701	59,058	2,358	reverse	64.	56,433	58,790	2,358	reverse	99.15
63.DNA polymerase	59,089	62,151	3,063	forward	65.DNA polymerase	58,821	61,883	3,063	forward	99.15
64.	62,228	62,686	459	reverse	66.	61,960	62,418	459	reverse	99.35
65.	62,752	63,135	384	reverse	67.	62,484	62,867	384	reverse	99.48
66.	63,141	63,398	258	reverse	68.	62,873	63,130	258	reverse	99.61
67.vlf-1	63,439	64,677	1,239	reverse	69.vlf-1	63,171	64,415	1,245	reverse	99.28
68.	64,690	65,022	333	reverse	70.	64,428	64,760	333	reverse	99.70
69.gp41	65,091	66,059	969	reverse	71.gp41	64,829	65,797	969	reverse	99.69
70.	65,989	66,714	726	reverse	72.	65,727	66,452	726	reverse	99.72
71.	66,587	67,264	678	reverse	73.	66,325	67,002	678	reverse	99.71

72.vp91	67,194	69,644	2,451	forward	74.vp91	66,932	69,382	2,451	forward	99.47
73.cg30	69,772	70,623	852	reverse	75.cg30	69,510	70,361	852	reverse	99.88
74.vp39	70,712	71,593	882	reverse	76.vp39	70,450	71,331	882	reverse	99.43
75.lef-4	71,592	72,977	1,386	forward	77.lef-4	71,330	72,715	1,386	forward	99.64
76.	73,030	73,794	765	reverse	78.	72,768	73,532	765	reverse	99.87
77.	73,796	74,284	489	forward	79.	73,534	74,022	489	forward	99.80
78.odv-e25	74,330	75,022	693	forward	80.odv-e25	74,068	74,760	693	forward	99.71
79.	75,054	75,551	498	reverse	81.	74,792	75,289	498	reverse	99.20
80.DNA helicase	75,570	79,331	3,762	reverse	82.DNA helicase	75,308	79,069	3,762	reverse	99.52
81.	79,288	79,809	522	forward	83.	79,026	79,547	522	forward	99.23
82.	79,868	80,833	966	reverse	84.	79,606	80,571	966	reverse	99.79
83.lef-5	80,729	81,676	948	forward	85.lef-5	80,467	81,414	948	forward	99.79
84.p6.9	81,670	81,999	330	reverse	86.p6.9	81,408	81,737	330	reverse	100.00
85.	82,064	83,173	1,110	reverse	87.	81,802	82,911	1,110	reverse	99.91
86.	83,219	83,587	369	reverse	88.	82,957	83,325	369	reverse	100.00
87.	83,587	84,720	1,134	reverse	89.	83,325	84,458	1,134	reverse	99.65
88.vp80	84,816	86,633	1,818	forward	90.vp80	84,553	86,370	1,818	forward	99.72

89.	86,630	86,806	177	forward	91.	86,367	86,543	177	forward	99.44
90.	86,821	87,906	1,086	forward	92.	86,558	87,643	1,086	forward	99.82
91.	87,952	88,236	285	forward	93.	87,688	87,972	285	forward	99.30
92.odv-e66	88,303	90,321	2,019	reverse	94.odv-e66	88,039	90,057	2,019	reverse	99.75
93.	90,342	91,172	831	reverse	95.	90,078	90,908	831	reverse	99.04
94.	93,489	94,088	600	forward	96.	93,508	94,107	600	forward	100.00
95.	94,092	94,448	357	forward	97.	94,111	94,467	357	forward	100.00
96.	94,543	96,069	1,527	forward	98.	94,563	96,095	1,533	forward	99.35
97.	96,148	96,909	762	forward	99.	96,174	96,935	762	forward	99.21
98.	96,924	97,256	333	forward	100.	96,950	97,282	333	forward	100.00
99.iap-3	97,315	98,121	807	reverse	101.iap-3	97,340	98,146	807	reverse	99.75
100.	98,118	98,273	156	reverse	102.	98,143	98,298	156	reverse	100.00
101.bro-b	98,377	99,882	1,506	reverse	103.bro-b	98,409	99,914	1,506	reverse	99.40
102.sod	100,050	100,529	480	forward	104.sod	100,082	100,561	480	forward	99.58
103.	100,536	101,909	1,374	forward	105.	100,568	101,941	1,374	forward	99.56
104.	101,962	102,540	579	reverse	106.	101,994	102,572	579	reverse	99.83
105.	102,711	103,058	348	forward	107.	102,742	103,098	357	forward	96.64

106.	103,069	103,335	267	forward	108.	103,109	103,375	267	forward	98.88
107.	103,403	104,989	1,587	forward	109.	103,443	105,029	1,587	forward	99.18
108.	104,986	105,222	237	forward	110.	105,026	105,262	237	forward	100.00
109.fgf	105,245	106,150	906	reverse	111.fgf	105,285	106,190	906	reverse	99.56
110.alk-exo	106,277	107,563	1,287	reverse	112.alk-exo CDS	106,318	107,604	1,287	reverse	99.77
111.	107,583	107,972	390	reverse	113.	107,624	108,013	390	reverse	100.00
112.	109,311	110,237	927	reverse	114.	109,693	110,619	927	reverse	100.00
113.Lef-2	110,438	110,653	216	forward	115.Lef-2	110,820	111,035	216	forward	100.00
114.p24	111,849	112,595	747	forward	116.p24	112,229	112,975	747	forward	99.60
115.gp16	112,657	112,947	291	forward	117.gp16	113,037	113,327	291	forward	99.66
116.calyx protein	112,999	114,021	1,023	forward	118.calyx protein	113,379	114,401	1,023	forward	99.41
117.	114,100	114,564	465	forward	119.	114,480	114,944	465	forward	100.00
118.	114,533	114,689	157	forward						N/A
119.	114,695	115,285	591	forward	120.	115,075	115,665	591	forward	100.00
120. 38.7 kDa protein	115,329	116,507	1,179	reverse	121.38.7 kDa protein	115,709	116,878	1,170	reverse	98.22
121.lef-1	116,509	117,246	738	reverse	122.lef-1	116,880	117,617	738	reverse	99.86
122.	117,221	117,655	435	reverse	123.	117,592	118,020	429	reverse	97.70

123.egt	117,800	119,347	1,548	forward	124.egt	118,165	119,712	1,548	forward	99.61
124.	119,547	120,125	579	forward	125.	119,912	120,490	579	forward	99.65
125.	120,076	120,876	801	forward	126.	120,441	121,241	801	forward	99.00
126.	120,959	123,802	2,844	reverse	127.	121,322	124,165	2,844	reverse	98.91
127.pkip-1	124,208	124,717	510	forward	128.pkip-1	124,571	125,080	510	forward	99.41
128.arif-1	124,784	125,581	798	reverse	129.arif-1	125,147	125,944	798	reverse	99.50
129.	125,838	126,989	1,152	forward	130.	126,205	127,356	1,152	forward	99.48
130.	127,030	129,063	2,034	reverse	131.	127,397	129,430	2,034	reverse	99.46
131.	129,205	129,750	546	reverse	132.	129,572	130,117	546	reverse	99.63
132.	129,943	130,530	588	forward	133.	130,299	130,886	588	forward	99.32

The incomplete ORFs from the HearNPV-Whit genome compared to the HearNPV-Au ORFs

Table 9.3: From nucleotide alignment, potentially expressed ORFs and their positions are shown with the percentage nucleotide identity for HearNPV-Whit genome versus HearNPV-Au. ORFs with numbers and no names represent potential hypothetical proteins. Only ORFs with full bases included, ORFs with incomplete sequences deleted.

HearNPV-Whit					**HearNPV-Au**					
ORF position and Name	Start	Stop	Size	Direction	ORF position and Name	Start	Stop	Size	Direction	% identity

3.pk1	1,990	2,793	804	forward	3.pk1	1,994	2,797	804	forward	99.75%
4.hoar	2,981	5,251	2,271	reverse	4.hoar	2,920	5,181	2,262	reverse	69.13%
5.	106,735	106,897	175	forward	5.	5,377	5,556	180	forward	95.56%
6.	107,111	107,970	860	reverse	6.	5,724	6,575	852	forward	98.60%
7.	108,142	108,316	163	reverse	7.	6,787	6,942	156	reverse	93.87%
8.ie0	105,890	106,747	858	reverse	8.ie0	6,930	7,787	858	forward	99.30%
9.p49	104,467	105,873	1,407	reverse	9.p49	7,804	9,210	1,407	forward	99.43%
10.odv-e18	104,211	104,456	246	reverse	10.odv-e18	9,221	9,466	246	forward	99.59%
11.odv-e27	103,342	104,196	855	reverse	11.odv-e27	9,481	10,335	855	forward	99.06%
12.	103,019	103,297	279	reverse	12.	10,381	10,659	279	forward	99.64%
13.	102,381	102,992	612	forward	13.	10,686	11,297	612	reverse	99.84%
14.ie-1	100,365	102,339	1,975	reverse	14.ie-1	11,339	13,306	1,968	forward	99.14%
15.odv-e56	99,247	100,311	1,065	forward	15.odv-e56	13,359	14,423	1,065	reverse	98.59%
16.me53	98,007	99,086	1,080	reverse	16.me53	14,584	15,663	1,080	forward	98.98%
17.	97,837	98,004	168	reverse	17.	15,666	15,833	168	forward	100.00 %
18.	97,503	97,784	282	forward	18.	15,886	16,167	282	reverse	98.58%
19.p74	95,422	97,482	2,061	reverse	19.p74	16,188	18,254	2,067	forward	98.45%

20.p10	95,105	95,368	264	forward	20.p10	18,308	18,571	264	reverse	98.48%
21.p26	94,219	95,022	804	forward	21.p26	18,654	19,457	804	reverse	99.13%
22.	93,903	94,106	204	reverse	22.	19,570	19,773	204	forward	100.00 %
23.lef-6	93,264	93,827	564	forward	23.lef-6	19,849	20,412	564	reverse	99.47%
24.dbp	92,243	93,215	973	forward	24.dbp	20,426	21,397	972	reverse	71.46%
25.	91,623	92,024	402	reverse	25.	21,617	22,018	402	forward	98.26%
26.he65	52,596	53,306	711	forward	59.he65	52,385	53,095	711	forward	99.30%
27.iap-2	53,383	54,135	753	reverse	60.iap-2	53,172	53,924	753	reverse	99.60%
28.	54,183	55,007	825	reverse	61.	53,972	54,796	825	reverse	99.27%
29.	54,976	55,377	402	reverse	62.	54,765	55,166	402	reverse	99.75%
30.lef-3	55,397	56,536	1,140	forward	63.lef-3	55,186	56,325	1,140	forward	98.95%
31.	56,643	59,000	2,358	reverse	64.	56,433	58,790	2,358	reverse	99.32%
32.DNA polymerase	59,034	62,096	3,063	forward	65.DNA polymerase	58,821	61,883	3,063	forward	99.09%
33.	62,173	62,631	459	reverse	66.	61,960	62,418	459	reverse	99.13%
34.	62,697	63,080	384	reverse	67.	62,484	62,867	384	reverse	99.74%
35.	63,086	63,343	258	reverse	68.	62,873	63,130	258	reverse	99.61%

36.vlf-1	63,384	64,631	1,248	reverse	69.vlf-1	63,171	64,415	1,245	reverse	99.60%
37.	64,644	64,976	333	reverse	70.	64,428	64,760	333	reverse	99.40%
38.gp41	65,045	66,013	969	reverse	71.gp41	64,829	65,797	969	reverse	99.07%
39.	65,943	66,668	726	reverse	72.	65,727	66,452	726	reverse	99.72%
40.	66,541	67,218	678	reverse	73.	66,325	67,002	678	reverse	99.56%
41.vp91	67,148	69,598	2,451	forward	74.vp91	66,932	69,382	2,451	forward	99.10%
42.cg30	69,727	70,578	852	reverse	75.cg30	69,510	70,361	852	reverse	99.06%
43.vp39	70,667	71,548	882	reverse	76.vp39	70,450	71,331	882	reverse	99.89%
44.lef-4	71,547	72,932	1,386	forward	77.lef-4	71,330	72,715	1,386	forward	99.71%
45.	72,985	73,749	765	reverse	78.	72,768	73,532	765	reverse	99.48%
46.	73,751	74,239	489	forward	79.	73,534	74,022	489	forward	99.59%
47.odv-e25	74,285	74,980	696	forward	80.odv-e25	74,068	74,760	693	forward	99.14%
48.	75,012	75,512	501	reverse	81.	74,792	75,289	498	reverse	99.00%
49.DNA helicase	75,531	79,292	3,762	reverse	82.DNA helicase	75,308	79,069	3,762	reverse	99.49%
50.	79,249	79,770	522	forward	83.	79,026	79,547	522	forward	100.00%
51.	79,829	80,794	966	reverse	84.	79,606	80,571	966	reverse	99.48%

52.lef-5	80,690	81,637	948	forward	85.lef-5	80,467	81,414	948	forward	99.47%
53.p6.9	81,631	81,960	330	reverse	86.p6.9	81,408	81,737	330	reverse	99.70%
54.	82,025	83,134	1,110	reverse	87.	81,802	82,911	1,110	reverse	99.28%
55.	83,180	83,548	369	reverse	88.	82,957	83,325	369	reverse	99.19%
56.	83,548	84,681	1,134	reverse	89.	83,325	84,458	1,134	reverse	99.65%
57.vp80 CDS	84,776	86,593	1,818	forward	90.vp80	84,553	86,370	1,818	forward	98.95%
58.	86,590	86,766	177	forward	91.	86,367	86,543	177	forward	98.87%
59.	86,781	87,866	1,086	forward	92.	86,558	87,643	1,086	forward	99.45%
60.	87,912	88,196	285	forward	93.	87,688	87,972	285	forward	100.00 %
61.odv-e66	88,263	90,281	2,019	reverse	94.odv-e66	88,039	90,057	2,019	reverse	99.85%
62.	90,302	91,132	831	reverse	95.	90,078	90,908	831	reverse	98.92%