

Article

Comparative Analysis of HaSNPV-AC53 and Derived Strains

Christopher Nouné and Caroline Hauxwell *

Queensland University of Technology, Brisbane 4000, Australia; chris.noune@connect.qut.edu.au

* Correspondence: caroline.hauxwell@qut.edu.au; Tel.: +61-07-3138-8062

Academic Editor: Karyn Johnson

Received: 12 September 2016; Accepted: 21 October 2016; Published: 31 October 2016

Abstract: Complete genome sequences of two Australian isolates of *H. armigera* single nucleopolyhedrovirus (HaSNPV) and nine strains isolated by plaque selection in tissue culture identified multiple polymorphisms in tissue culture-derived strains compared to the consensus sequence of the parent isolate. Nine open reading frames (ORFs) in all tissue culture-derived strains contained changes in nucleotide sequences that resulted in changes in predicted amino acid sequence compared to the parent isolate. Of these, changes in predicted amino acid sequence of six ORFs were identical in all nine derived strains. Comparison of sequences and maximum likelihood estimation (MLE) of specific ORFs and whole genome sequences were used to compare the isolates and derived strains to published sequence data from other HaSNPV isolates. The Australian isolates and derived strains had greater sequence similarity to New World SNPV isolates from *H. zea* than to Old World isolates from *H. armigera*, but with characteristics associated with both. Three distinct geographic clusters within HaSNPV genome sequences were identified: Australia/Americas, Europe/Africa/India, and China. Comparison of sequences and fragmentation of ORFs suggest that geographic movement and passage in vitro result in distinct patterns of baculovirus strain selection and evolution.

Keywords: baculovirus; SNPV; *Helicoverpa*; Next Generation Sequencing; strain selection; virus evolution

1. Introduction

Baculoviruses (family *Baculoviridae*) are double-stranded DNA (dsDNA) viruses with a genome of between 80,000 and 180,000 base pairs [1,2]. The genetic diversity of the Group II nucleopolyhedroviruses (genus *Alphabaculoviruses*) from Lepidoptera of the genus *Helicoverpa* are of importance due to their worldwide distribution and widespread use as biopesticides against these significant polyphagous pests [3]. Group II singly-enveloped nucleopolyhedroviruses from species of the genus *Helicoverpa* (Lepidoptera: Noctuidae) were originally classified into two species; Old World *H. armigera* single nucleopolyhedrovirus (HaSNPV), isolated from *H. armigera* (Hübner) and New World *H. zea* single nucleopolyhedrovirus (HzSNPV) isolated from *H. zea* (Boddie) [3–13]. This has been recently revised to classify both types as a single species, HaSNPV, with similarities in DNA sequence and biological activity [12].

Old world isolates of HaSNPV, and New World isolates from *H. zea* are widely used in Australia as biopesticides against both *H. armigera* and *H. punctigera* (Wallengren) in a range of crops including sorghum, chickpea and cotton [14] and are also registered in South Africa and the USA. Two Australian HaSNPV isolates, H25EA1 and AC53, are of international interest as biopesticides. HaSNPV isolate HaSNPV-AC53 (AC53) is manufactured in Australia and included in the commercial biopesticides “Vivus” and “Vivus Max” (AgBiTech Pty Ltd., Brisbane, Queensland, Australia). H25EA1 was selected

by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) from a wild type isolate, and was used by the University of Queensland for in vitro baculovirus production [10,15–17].

Significant genotypic and phenotypic diversity exists within nucleopolyhedroviruses (NPV) isolates, which can be identified by cloning in vivo or in vitro [11,18–22]. For example, 25 of the 162 tissue culture clones isolated from field populations in Kenya, South Africa, Zimbabwe and Thailand were unique variants of HaSNPV [23,24]. Classification and origin of baculovirus species and strains remain important due to restrictions on import of non-native species and concerns over variation between strains during registration of biopesticides, particularly in Australia [25].

Baculovirus species have been described using restriction endonuclease digestion profile and Sanger sequencing, and more recently by Next Generation Sequencing (NGS) [10,11,16,17,23,26–29]. Previous research has shown that HaSNPV and HzSNPV share sequence similarity of up to 99.9%, but could be distinguished by a small number of nucleotide substitutions and by open reading frame (ORF) insertions and deletions in the published consensus genome [17,30,31]. However, we know little about the strain diversity within these isolates and their taxonomic relationship to the Old and New World wild type strains.

This paper examines the sequence similarity and relationships of two Australian HaSNPV isolates from larvae of unidentified *Helicoverpa* sp. and of nine strains derived by passage in tissue culture and insects. We compare whole genome sequences and sequences of selected hypothetical and functional ORFs to determine patterns of strain selection and evolution [12,17] in comparison to sequences from both Old and New World isolates. Throughout, we use HaSNPV to refer to the *Helicoverpa* SNPV virus species but identify isolates from the insect *H. zea* as HzSNPV to differentiate isolates from that of the host and where sequences use the old nomenclature.

2. Materials and Methods

2.1. Virus Source and Passage

HaSNPV isolate AC53, also known as A44WT [10,16], was obtained from AgBiTech and isolate H25EA1 was selected in vitro by CSIRO from P9/H25WT [15,32–35], and obtained from the University of Queensland [17]. Both were originally isolated from cadavers of an unspecified *Helicoverpa* species in Queensland, Australia in 1973 and 1974, respectively, and passaged once through *H. punctigera* before repeated passage through *H. armigera*. This isolation predates the introduction of New World isolates from *H. zea* and use of commercial biopesticides in Australia [10,16]. Both isolates were passaged once by infection of third instar *H. armigera* larvae using a modified droplet method [36]. Insects were fed a suspension of virus with the addition of 10% blue food dye (Queen Fine Foods©, Brisbane, Queensland, Australia) to visualise ingestion and then maintained in individual cups with fresh modified tobacco hornworm diet at constant $26\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ with 16 h light/8 h dark periods and $70\% \pm 5\%$ humidity until death.

Occlusion bodies were extracted from cadavers by maceration in 0.1% sodium dodecyl sulphate (SDS), filtration through muslin and centrifugation at 500 rpm and $4\text{ }^{\circ}\text{C}$ for 5 min to remove insect debris, followed by centrifugation at 4000 rpm and $4\text{ }^{\circ}\text{C}$ for 20 min in a swing-out rotor (Sorvall Legend RT®, Sorval Heraeus Rotor). The supernatant was discarded and the pellet resuspended in MilliQ water (Merck Millipore, Boston, MA, USA).

2.2. Test for Latent Virus

The possible presence of latent or sub-lethal (covert) HaSNPV infection in the *H. armigera* insects was investigated. A total of 20 instar larvae were collected for examination by PCR [37,38]. A single AC53 infected larvae was used as a positive control. Each larva was homogenized in a 1.5 mL microcentrifuge tube with 1 mL cold buffer (Tris 10 mM, magnesium chloride 1.5 mM, sodium chloride 140 mM and 80 μL of 5% Tergitol) and centrifuged at 3800 rpm (Eppendorf Minispin, Hamburg, Germany) for 10 min. The supernatant was collected and spun at 4000 rpm for a further 10 min to pellet cellular material. The supernatant was then discarded, the pellet resuspended and cells lysed

by addition of 200 μL of 0.05 M sodium chloride, 200 μL of 2 \times Tris/EDTA (TE) buffer and 40 μL of 0.1% SDS, vortexed until clear and incubated at 50 $^{\circ}\text{C}$ for 10 min, then centrifuged at 13,000 rpm for 10 min in order to pellet cellular debris.

Proteins were precipitated by adding 400 μL of supernatant from each sample to 200 μL of ice cold 2.5 M potassium acetate and vortexed thoroughly, then left on ice for 5 min. The samples were centrifuged at 13,000 rpm for 10 min to precipitate the protein. DNA was precipitated from the supernatant by addition of isopropanol and centrifugation at 13,000 rpm for 10 min. The DNA pellet was washed twice with 1 mL of 70% analytical grade ethanol/MilliQ water, centrifuged at 13,000 rpm for 10 min and then air-dried for 10 min. The DNA pellet was resuspended with 60 μL of 1 \times Tris/EDTA. Degenerate rPol and A44-RIX PCR primers and reaction conditions were as previously described [16]. The PCR amplification was carried out using a Mango Taq kit (Meridian Bioscience Inc., Cincinnati, Ohio, USA) and an Eppendorf Pro S thermocycler, and underwent electrophoresis using a 1% *w/v* agarose gel with 0.0001% GelRed (Biotium Inc., Fremont, California, USA) in 1 \times Tris-acetate EDTA buffer for 1 h at 100 volts.

2.3. Strain Isolation by Passage and Selection in Tissue Culture

Strains were isolated from AC53 using a modified tissue culture plaque assay [39–41]. HzAM1 tissue culture cells were obtained from Dr. Steven Reid (University of Queensland, Australia) and cultured in Ex-Cell 420 Serum Free Insect Medium (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% Fetal Bovine Serum (Invitrogen, Thermo-Fisher, Waltham, MA, USA). Second instar *H. armigera* larvae were infected with 1.11×10^5 OB/mL (LC_{90}) of AC53 and hemolymph was harvested at 48 h and 72 h post-infection (pi) by nicking the cuticle of the rear dorsal surface with a scalpel blade. Between 2 and 10 μL was harvested from lots of 10 neonates into glass vials containing 200 μL of Ex-Cell 420 Serum Free Insect Medium (Sigma-Aldrich) with 0.005% phenylthiourea in ethanol (to prevent melanization).

Two methods of tissue culture selection were used (Figure 1). The first method used conventional plaque selection with an agar overlay from infected insect haemolymph. The second used an initial passage of virus from infected haemolymph in tissue culture without agar overlay in order to generate occlusion bodies of strains more adapted to tissue culture. The occlusion bodies generated were used to infect insects from which plaques were selected by the conventional method.

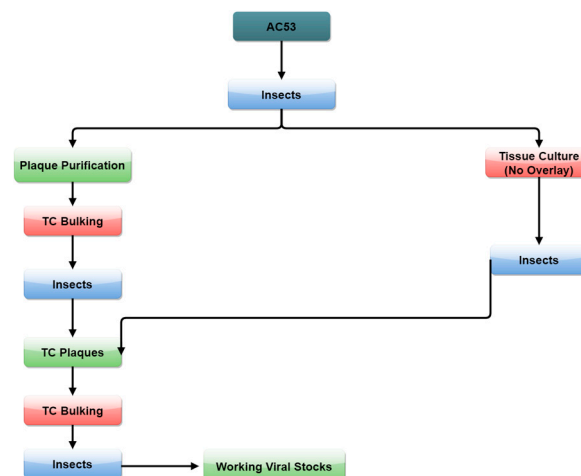


Figure 1. Isolation of *H. armigera* single nucleopolyhedrovirus (HaSNPV) strains in tissue culture (TC).

In Method 1, a total of 100 μL diluted haemolymph was used to infect tissue culture cells at 1.5×10^5 cells per plate (30 mm tissue culture treated petri dishes (Corning Inc., Corning, NY, USA) with an overlay of 1% low temperature gelling agar (SeaPrep LE; Lonza Group, Basel, Switzerland)

in Ex-Cell 420 Serum Free Insect Medium and the addition of 10% Fetal Bovine Serum (Invitrogen) and incubated for 7 days at 28 °C [39,42]. Plaques were visualized by incubation overnight with 25% neutral red in water and picked using a pipette into 200 µL of Ex-Cell 420 Serum Free Insect Medium. Plaque suspensions were then used to infect cells at 2×10^6 cells per plate without overlay and incubated for 7 days as above to produce occlusion bodies. Cells and occlusion bodies were scraped into 2 mL Eppendorf tubes and pelleted at 13,000 rpm for 10 min. Supernatant was poured off and resuspended with 0.05% Tween 80 in MilliQ water and pelleted again. Final pellets were resuspended in 50% glycerol:50% MilliQ and used to infect second instar *H. armigera* larvae as above. The larvae were bled as above, along with a second round of plaque purification and occlusion body production. This was used to infect second instar larvae, from which occlusion bodies were harvested and five strains were produced (“C” strains).

For Method 2, haemolymph, harvested between 48 and 72 h (Table 1) as described above, was first passaged through tissue culture without an agar overlay and occlusion bodies harvested at 7 days. Occlusion bodies were then used to infect second instar larvae that were bled at intervals from 48 to 120 h pi (Table 1). The haemolymph was used for plaque selection and occlusion body production in tissue culture and second instar insects as above (Figure 1), and occlusion bodies were extracted from cadavers to produce four strains (“T” strains).

In the final passage in insects, two strains, T4.1 and T4.2, were selected from two distinct peaks in larval mortality at 168 and 288 h pi in strain T4 to give a total of nine strains (Table 1).

Table 1. HaSNPV-AC53 (AC53) Isolated Strains.

Strain	Time (h) Post-Infection (pi)	First Round Isolation Method
AC53-C1	48 and 48	Agar Overlay
AC53-C5	48 and 48	Agar Overlay
AC53-C6	48 and 48	Agar Overlay
AC53-C9	48 and 48	Agar Overlay
AC53-C3	72 and 72	Agar Overlay
AC53-T2	48 and 48	Tissue Culture—No Overlay
AC53-T4.1 *	72 and 96	Tissue Culture—No Overlay
AC53-T4.2 *	72 and 96	Tissue Culture—No Overlay
AC53-T5	72 and 120	Tissue Culture—No Overlay

Eight isolates were harvested from haemolymph at different times post-infection (passages 1 and 2). Strains C1, C5, C6, C9 and C3 were selected by Method 1, and strains T2, T4 and T5 by Method 2; * Selected from strain T4 at two time points (168 h and 288 h pi) during final passage in neonate larvae.

2.4. DNA Extraction, Next Generation Sequencing Library Preparation, Sequencing and Genome Assembly

DNA was extracted from occlusion bodies using a modification of the method of Doyle et al. [43]. Analytical-grade 0.05 M sodium carbonate was added to the virus pellet to release virions from occlusion bodies. Then, 0.1% SDS in TE buffer was added to disrupt virion membranes. Isolate II Genomic DNA kits (Bioline) were used from Step 4 of the manufacturer’s instructions. DNA concentration was determined, (Qubit assay; Invitrogen) and then diluted to 1 ng/µL in MilliQ water. Library preparation was completed using a Nextera XT kit (Illumina, San Diego, CA, USA) and sequences using 150 base pair (bp) paired-end chemistry on an Illumina NextSeq 500 [11]. The assembly method was completed as previously described [11]. This assembly method was developed into a Bash software pipeline, Invertebrates & Microbiology Group-Assembly Pipeline (IMG-AP).

2.5. Sequence Analysis and Maximum-Likelihood Estimation (MLE)

Twenty-one published full genome sequences of HaSNPV isolates [17,23,24,31,44–48] and three genome sequences from HzSNPV isolates [13,49] were aligned and rooted to the *Autographica californica* MNPV (AcMNPV) [50] using MAFFT (Version 7.222) with the FFT-NS-2 algorithm, default settings [51], and any polymorphism including gaps caused by insertions and deletions were classed as mismatches. Alignments were visualized using Geneious R9.1.5. Maximum-likelihood tree construction was

completed with RAxML (Version 7.2.8) with the GTR GAMMA model, rapid bootstrapping and searching for the best-scoring maximum-likelihood tree with 1000 bootstrap replicates [52]. Tree visualisation and editing was completed using TreeGraph 2.9.2-622 beta [53]. This was repeated for the available baculovirus repeated open reading frame (BRO)-A, BRO-B, ORF42 (ORF43 homolog), ORF61 (HzSNPV ORF62 homolog), ORF78 including ORF78a and ORF78b (HzSNPV ORF79 homolog) *lef-8*, *lef-9* and *polh* sequences on Genbank including *Busseola fusca* SNPV isolate A2-4 [54], *Helicoverpa gelotopoeon* SNPV [55], *Helicoverpa assulta* SNPV [56], *Mamestra configurata* NPV-A (MacoNPV-A) [57,58], *Heliothis virescens* Ascovirus 3e [59] and *Plasmodium falciparum* 3D7. Accession numbers and country of origins of each analysed isolate and ORF are shown in Tables S3–S7.

Comparisons of derived strains ORF and homologous repeat mutations were analysed with MAFFT, with the FFT-NS-2 algorithm, and a local copy of BLAST+ (Version 2.5.0) using the Megablast algorithm, to identify nucleotide mutations and the blast algorithm for amino acid mutations [60–62].

3. Results

3.1. Test for Latent Virus

No latent virus was detected within the colony using both rPol (Figure S1) and A44-RIX (Figure S2) primers. The infected positive control tested positive.

3.2. Strain Isolation

Eight strains were selected, 5 “C” strains from Method 1 and 3 “T” strains from Method 2 (Table 1). Strain T4 was split into two strains, T4.1 and T4.2, from cadavers in two distinct peaks of larval mortality: 40% mortality at 168 h and 60% mortality at 288 h.

3.3. Sequence Analysis

All of the derived strains exhibited differing whole genome sequence lengths of between 130,435 bp and 130,460 bp (Table 2) compared to 130,442 bp for the parent strain AC53 and 130,440 bp for H25EA1. Most of the variation in length was found in non-coding regions, but some variation was found within ORFs (Tables 3 and 4 and Table S1).

Table 2. Sequenced *H. armigera* single nucleopolyhedrovirus (HaSNPV) and *H. zea* single nucleopolyhedrovirus (HzSNPV) nucleotide identity compared to AC53 with sequence identity ranging between 81.692% (L1 strain) and 99.604% (AC53-T5).

Common Name	Genbank Accession	Sequence Length (bp)	Nucleotide Identity to AC53 (%)	Country/Region of Origin
HaSNPV-AC53-C1	KU738896	130,460	99.624	Australia
HaSNPV-AC53-C5	KU738898	130,439	99.600	Australia
HaSNPV-AC53-C6	KU738899	130,435	99.601	Australia
HaSNPV-AC53-T4.1	KU738902	130,440	99.602	Australia
HaSNPV-AC53-T5	KU738904	130,442	99.603	Australia
HaSNPV-AC53-C9	KU738897	130,437	99.599	Australia
HaSNPV-AC53-T2	KU738901	130,440	99.596	Australia
HaSNPV-AC53-T4.2	KU738896	130,443	99.530	Australia
HaSNPV-AC53-C3	KU738897	130,437	99.595	Australia
HaSNPV-H25EA1	KJ922128	130,436	99.423	Australia
HzSNPV-HS18	KJ004000	130,890	99.220	Unknown—Sequenced in Russia
HzSNPV-F16 (Elcar-derived)	AF334030	130,869	99.208	USA—Sequenced in China
HzSNPV-Br/South	KM596835	129,694	98.277	Brazil
HaSNPV-NNg1	AP010907	132,425	96.203	Kenya
HaSNPV-LB1	KJ701029	131,966	96.012	Iberia
HaSNPV-SP1A	KJ701032	132,481	95.961	Iberia
HaSNPV-SP1B	KJ701033	132,265	95.810	Iberia
HaSNPV-LB3	KJ701030	130,949	95.799	Iberia
HaSNPV-LB6	KJ701031	130,992	95.798	Iberia
HaSNPV-C1	AF303045	130,759	95.353	China
HaSNPV-AU	JN584482	130,992	94.860	Australia—Sequenced in China
HaSNPV-G4	AF271059	131,405	94.442	China

AC53-derived, H25EA1, and HzSNPV strains are all within 2% nucleotide identity, whereas the remaining HaSNPV strains are within 5.5% nucleotide identity—excluding the L1 strain, which seems to be an outlier.

Table 3. Amino acid (AA) and nucleotide (N) identity (%) of the regions that are not identical to AC53. The main difference between AC53 and the derived strains occur with both baculovirus repeated open reading frame (BRO)-A and BRO-B, Hr1–Hr5 and HOAR, which is to be expected due to the known hypervariability of the regions. However, open reading frame (ORF) 7 and the hypothetical ORF contains an early stop resulting in a smaller sequence, whereas ORF61 is longer due to the derived strains containing an early stop.

ORF	Protein	AC53-C1		AC53-C3		AC53-C5		AC53-C6		AC53-C9		AC53-T2		AC53-T4.1		AC53-T4.2		AC53-T5		Notes
		AA	N	AA	N	AA	N	AA	N	AA	N	AA	N	AA	N	AA	N	AA	N	
4	HOAR	95.6	95.7	95.8	96.1	96.9	97.1	96.8	96.9	96.9	96.9	96.7	96.9	96.7	96.9	96.9	96.9	96.7	96.9	
5		34.9	97.2	98.3	98.9	34.9	97.2	34.9	97.2	34.9	97.2	34.9	97.2	34.9	97.2	34.9	97.2	34.9	97.2	AC53 and AC53-C3 have identical length
6		99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	
7		94.1	99.3	94.1	99.3	94.1	99.3	94.1	99.3	94.1	99.3	94.1	99.3	94.1	99.3	94.1	99.3	94.1	99.3	AC53 is 85 bp shorter
	Hr1	N.A	99.7	N.A	99.7	N.A	99.7	N.A	99.8	N.A	99.8	N.A	99.7	N.A	99.7	N.A	99.8	N.A	99.7	
	Hr2	N.A	95.5	N.A	95.2	N.A	95.2	N.A	95.2	N.A	95.3	N.A	95.2	N.A	95.2	N.A	95.2	N.A	95.2	
Hypothetical ORF	Hypothetical Protein	70.7	98.0	70.7	98.0	70.7	98.0	70.7	98.0	70.7	98.0	70.7	98.0	70.7	98.0	70.7	98.0	70.7	98.0	AC53 is 24 bp shorter
59	BRO-A	90.9	92.1	90.9	92.1	90.9	92.1	90.9	92.1	90.9	92.1	91.3	92.1	90.9	92.1	90.9	92.1	90.9	92.1	
60	BRO-B	90.4	94.0	90.4	94.0	90.4	94.0	90.4	94.0	90.4	94.0	90.4	94.0	90.4	94.0	90.4	94.0	90.4	94.0	
	Hr3	N.A	99.6	N.A	99.6	N.A	99.6	N.A	99.6	N.A	99.6	N.A	99.69	N.A	99.6	N.A	99.8	N.A	99.6	
61		80.0	86.9	86.8	86.9	80.0	86.9	80.0	86.9	80.0	86.9	80.0	86.9	80.0	86.9	80.0	86.9	80.0	86.9	AC53 is 41 bp longer
68	DNA polymerase	100	100	99.9	99.9	99.9	99.9	100	100	100	100	99.9	99.9	100	100	100	100	100	100	
78a/78b (ORF78 in all other strains)		100	100	100	100	100	100	76.3	99.4	100	100	100	100	100	100	100	100	100	100	Split in two with AC53-C6
	Hr4	N.A	99.5	N.A	99.0	N.A	99.0	N.A	99.0	N.A	99.0	N.A	99.0	N.A	99.0	N.A	99.0	N.A	99.0	
	Hr5	N.A	99.3	N.A	99.1	N.A	99.1	N.A	99.5	N.A	99.2	N.A	99.1	N.A	99.4	N.A	99.1	N.A	99.4	
126	38.7K	100	100	100	100	100	100	100	100	100	100	99.7	99.4	100	100	100	99.8	100	100	
128a/128b (ORF128 in all other strains)		100	100	100	100	100	100	100	100	100	100	100	100	100	100	23.2	87.6	100	100	Split in two with AC53-T4.2
133	PKIP-1	100	100	100	99.8	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
137		97.2	99.3	97.2	99.3	97.2	99.3	97.2	99.3	97.2	99.3	97.2	99.3	97.2	99.3	97.2	99.3	97.2	99.3	
Total regions with sequence polymorphisms		9	14	10	16	10	15	10	15	8	14	11	16	9	14	10	16	9	14	

N.A. = not applicable.

Table 4. Comparison of the nucleotide and amino acid sequence similarity of AC53 derived strains to each other. The greatest diversity was within the five homologous repeat regions, DNA polymerase and HOAR. Only AC53-T2 contained an amino acid difference in BRO-A. ORF5 is shorter in length in all strains except AC53-C3 than in AC53. AC53-C6 and AC53-T4.2 contained unique mutations within ORF78 and ORF128, respectively, due to an inserted stop.

ORF/Region	Nucleotide Similarity and Clusters within Derived Strains	Amino Acid Similarity and Clusters within Derived Strains
HOAR	- AC53-T4.1, AC53-T5 = 100% - AC53-C6, AC53-C9, AC53-T4.2 = 100% - Remaining 4 strains all different at 96.7% to 99.9%	- AC53-T4.1 and AC53-T5 = 100% - AC53-C6, AC53-C9, AC53-T4.2 = 100% - Remaining 4 strains all different at 95.8% to 99.8%
ORF5 *	- AC53-C3 = 96.1% - Remaining strains all identical	- AC53-C3 = 33.3% - Remaining strains all identical
BRO-A	- AC53-T2 = 99.9% - Remaining strains all identical	- AC53-T2 = 99.5% - Remaining strains all identical
DNA-Polymerase	- AC53-T5, AC53-T4.2, AC53-T4.1, AC53-C9, AC53-C6, AC53-C1 = 100% - AC53-C5, AC53-C3 = 100% - AC53-T2 = 99.9%	- AC53-C3, AC53-C5, AC53-T2 = 100% - AC53-T5, AC53-T4.2, AC53-T4.1, AC53-C9, AC53-C6 - AC53-C1 = 100%
ORF78/ORF78a and 78b in AC53-C6	- AC53-C6 = 99.4% - Remaining strains all identical	- AC53-C6 = 77.9% - Remaining strains all identical
38.7K Protein	- AC53-T2 = 99.4% to other 7 strains and 99.6% to AC53-T4.2 - AC53-T4.2 = 99.8% to other 7 strains - Remaining 7 strains all identical	- AC53-T2 = 99.7% - Remaining 8 strains all identical
ORF128/ORF128a and 128b in AC53-T4.2	- AC53-T4.2 = 87.6% - Remaining strains all identical	- AC53-T4.2 = 23.2% - Remaining strains all identical
PKIP-1	- AC53-C3 = 99.8% - Remaining strains all identical	- All strains = 100%
Hr1	- AC53-T4.2, AC53-C9, AC53-C6 = 100% - AC53-T5, AC53-T4.1, AC53-T2, AC53-C5, AC53-C3, AC53-C1 = 100% - 99.9% when both groups compared	- Not Applicable
Hr2	- AC53-C6, AC53-T4.1 = 100% - Remaining strains all identical	- Not Applicable
Hr3	- AC53-T4.2 = 99.8% - Remaining strains all identical	- Not Applicable
Hr4	- AC53-C1 = 99.2% - Remaining strains all identical	- Not Applicable
Hr5	- AC53-T2, AC53-C5, AC53-C3 = 100% - Remaining strains all identical	- Not Applicable

* ORF5 is 87 bp longer in AC53-C3.

Both AC53 and H25EA1, and the derived strains shared sequence similarities with both Old World isolates from *H. armigera* and New World isolates from *H. zea*. Overall sequence similarity between AC53 and other isolates was greatest with *H. zea* isolates (98% to 99%) and 94%–98% with *H. armigera* isolates (Table 2). Sequence similarity to the L1 isolate from India was 81%, which contained a significant rearrangement in the genome and was excluded from whole genome comparisons. AC53 and H25EA1 had 99.4% overall sequence similarity (Table 2). The greatest nucleotide differences within reading frames were in BRO-A (10%) and BRO-B (4%). ORF136 of H25EA1 was 432 bp shorter than in AC53 but had 99% nucleotide sequence similarity to that of AC53.

Comparison of the AC53 genome with the derived strains identified nucleotide base changes in up to 16 different regions, resulting in predicted amino acid changes in 11 ORFs (Table 3, Table S2). Nine ORFs contained predicted amino acid changes in every tissue culture-derived strain (i.e., HOAR, ORF5, ORF6, ORF7, Hypothetical ORF, BRO-A, BRO-B, ORF61, and ORF137). Of these, 6 ORFs (ORF6, ORF7, Hypothetical ORF, BRO-B, ORF61, ORF137) had identical predicted amino acid changes. ORF7 was 85 bp longer in all of the derived strains than in the AC53 parent isolate (Figure S3).

The remaining 3 ORFs (HOAR, ORF5 and BRO-A) had differences between strains as well as from the parent AC53 isolate (Table 4). In the AC53-T2 strain, BRO-A had a single nucleotide change that resulted in a different predicted amino acid sequence from that of the other eight derived strains. Strain AC53-C3 had the full-length ORF5 found in AC53 isolate, but ORF5 was 87 bp shorter in the other eight derived strains resulting in changes in predicted amino acid sequence. HOAR differed from the AC53 parent strain in every derived strain but the strains contained multiple and different polymorphisms resulting in six different genotypes with differences in nucleotide and predicted amino acid sequence (Table 4, Figure S5).

Nucleotide base changes were identified in seven ORFs when comparing the genomes of the derived strains (Table 4), but mutations resulting in amino acid changes were only found in the sequences of six ORFs (HOAR, ORF5, DNA polymerase and BRO-A, ORF61, ORF78 and ORF128). All polymorphisms are listed in Tables 3 and 4 (no polymorphisms were observed in other ORFs). The complete genome sequences of HaSNPV-H25EA1, AC53 and seven derived isolates, contained 139 ORFs and five homologous repeat (hr) regions. Of these, 138 ORFs were shared with the HzSNPV-F16 isolate [13].

Two derived strains (AC53-C6 and AC53-T4.2) contained 140 ORFs as a result of insertions of early stop codons. A single base pair deletion at position 69,728 in ORF78 of AC53-C6 split the ORF into two smaller hypothetical ORFs of 73 bp (ORF78a) and 81 bp (ORF78b). Similarly, in AC53-T4.2, ORF128 splits into two hypothetical ORFs of 267 bp (ORF128a) and 141 bp (ORF128b), and a 26 bp non-coding region resulting from 50 substitutions to five deletions (Figure S6). This split did not occur in AC53-T4.1 or in the AC53 wild type. HaSNPV-H25EA1, AC53 and the nine derived strains all contained an additional hypothetical ORF identified in the antisense direction between ORF54 and *lef-9* (Figure 2). This ORF was 99 bp in the AC53 parent strain and H25EA1, and 120 bp in the derived strains, which contained an additional 21 bp CA-repeat region.

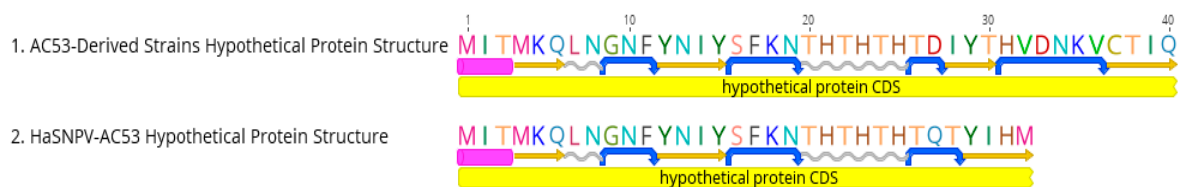


Figure 2. Amino acid sequence and predicted protein structure of hypothetical proteins identified using the EMBOSS garnier [63]. Alpha helix (purple rectangle), beta strands (yellow arrows), coils (gray wavy lines), and the turns (blue curved arrows) are depicted here. AC53-derived strains contain an additional turn and an additional beta strand; CDS, coding DNA sequence.

The Australian isolates contained ORFs similar SNPV isolates from both *H. armigera* and *H. zea*. Both AC53 and H25EA1 and all nine derived strains contained ORF42, which is reported to be found

in some HaSNPV isolates, and as a homolog at ORF43 in some *H. zea* SNPVs [13,48]. Likewise, in both isolates and eight of the derived strains, ORF78 had 99.4% sequence similarity to an ORF79 that is reported to be specific to isolates from *H. zea*, and no similarity to the ORF78 annotated in published sequences of isolates from *H. armigera* [3]. In derived strain AC53-C6, ORF78 is split into ORF78a (73 bp) and ORF78b (81 bp; as above).

There was a pattern of increasing fragmentation of the ORF78 homologs in comparison to published HaSNPV genomes (Figure S7). The full-length *H. zea* ORF 79 homolog in all published *H. zea* isolates, except for HzSNPV-Br/South, is split into two 84 bp and 81 bp hypothetical ORFs (not annotated on Genbank) that are homologs of AC53-C6 ORF78a and ORF78b, respectively. The Iberian (SP1A, SP1B, LB1, LB3 and LB6) and Kenyan (NNg1) HaSNPV isolates contained a similar, unannotated 81 bp homolog of the ORF78b found in AC53-C6 that resulted from either a 16 bp or 14 bp deletion in a 28 bp AT-repeat, as had been previously observed in other strains [3]. The Chinese C1, G4 and AU HaSNPV isolates did not contain a homolog to ORF78b but contained two overlapping, unannotated hypothetical ORFs that are 45 bp and 48 bp in length within the region of the genome analogous to ORF78b of AC53-C6.

Two of the Iberian isolates (LB6 and LB3) also contained an unannotated 73 bp hypothetical ORF78a homolog while the remaining Iberian, NNg1, Chinese C1, G4 and AU isolates contained a much smaller 57 bp hypothetical ORF homolog of ORF78a. The pattern of increased fragmentation of isolates can also be observed in ORF61 (Figure S8). The ORF61 of AC53 and H25EA1 shared 99.4% sequence similarity, with a single substitution in H25EA1 from T to C, at position 52,591. There were again greater similarities with *H. zea* isolates: ORF61 in AC53 had 100% identity to ORF62 reported in *H. zea* isolate 35036 derived from the commercial biopesticide Gemstar, and H25EA1 had 100% sequence similarity to ORF62 in three HzSNPV isolates (isolate F16 from the commercial product Elcar, isolate 35022 from Gemstar, and *H. zea* isolate HS-18 sequenced in Russia), and 100% similarity to ORF66 in the Brazilian HzSNPV isolate Br/South [13,49]. ORF61 of AC53 had 99.4% sequence similarity to ORF62 of the Kenyan isolate NNg1, which has a substitution from T to A at position 52,516 (of AC53). ORF61 of H25EA1 has 98.8% similarity to ORF62 of NNg1, but with a second substitution (T to C) at position 52,591. The ORF61 from AC53 and H25EA1 has 98.9% (two substitutions) and 99.4% (one substitution) sequence similarity to a hypothetical ORF of the samelength (180 bp) in four Iberian *H. armigera* isolates that is not annotated as an ORF in the sequences on Genbank (KJ701030 to KJ701033).

ORF61 in eight of the derived strains has a 27 bp insertion, and one (AC53-C1) has a 20 bp insertion, that truncates ORF61 by 41 bp when compared to AC53. The remaining 72 bp of the ORF have 100% sequence similarity to the same region of ORF61 in AC53. The 72 bp truncated ORF61 also shares 100% sequence similarity to a 72 bp hypothetical ORF in the Iberian HaSNPV isolate LB1, but the truncation was caused by a 1 bp insertion and two substitutions and not the 27 bp insertion. This hypothetical ORF has not been annotated in the LB1 sequence on Genbank (KJ701029). The truncated, 72 bp, hypothetical ORF61 has 98.7% sequence similarity (1 bp substitution) to an unannotated hypothetical ORF in the Chinese *H. armigera* isolate C1, which also contains the 27 bp insertion. Homologs of ORF61 are not found in the Chinese G4 or AU isolates but a 107 bp fragment with 99.1% sequence similarity is found the Hr3 region of those two isolates.

3.4. Maximum Likelihood Estimation

Comparison of the AC53, H25EA1 and derived strain whole genome sequences to all *Helicoverpa* SNPV genomes available on Genbank identified three distinct clusters of isolates based on geographic origin but not on host species (Figure 3). AC53, H25EA1 and the derived strains all clustered with isolates from *H. zea*. This cluster could be further divided (with 77% support) into a cluster containing AC53 and its derived strains, and a second cluster containing H25EA1 and all the HzSNPV isolates. The Old World HaSNPV isolates form two distinct clusters with 100% support: a group of three isolates from China (G4, C1 and the Chinese 'AU' isolate), and a second cluster containing the Iberian and Kenyan isolates.

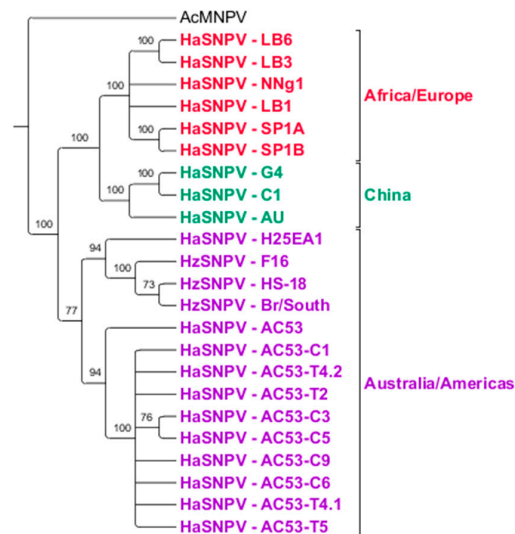


Figure 3. Phylogenetic Relationship of all of the HaSNPV and *H. zea* single nucleopolyhedrovirus (HzSNPV) strains (with bootstrap support as a percentage) and rooted to *Autographica californica* multiple nucleopolyhedrovirus (AcMNPV). Tree has been collapsed based on a minimum of 70% support. Geographically, three distinct groups can be observed; an Australian/American group consisting of AC53, H25EA1, AC53-derived strains and the HzSNPV isolates (purple), a Chinese group containing the HaSNPV C1, G4 and AU strains (green), and a third African/European group containing the Iberian (SP, LB) and Kenyan (NNG1) isolates (red).

BRO-A homologs are only found in AC53, H25EA1, strains derived from AC53, and HzSNPV isolates. Maximum likelihood estimation (MLE) of BRO-A identified four clusters with 100% support: the *H. zea* isolates and H25EA1, isolate AC53, derived strain AC53-T2 and the remaining strains derived from AC53 (Figure 4). BRO-B homologs are found in the majority of other sequenced HaSNPVs except the Chinese AU, Iberian SP1A and Iberian SP1B isolates. Four clusters were identified: *H. zea* isolates, *H. armigera* isolates (including H25EA1), AC53, and the strains derived from AC53, which again clustered separately from all other isolates including AC53.

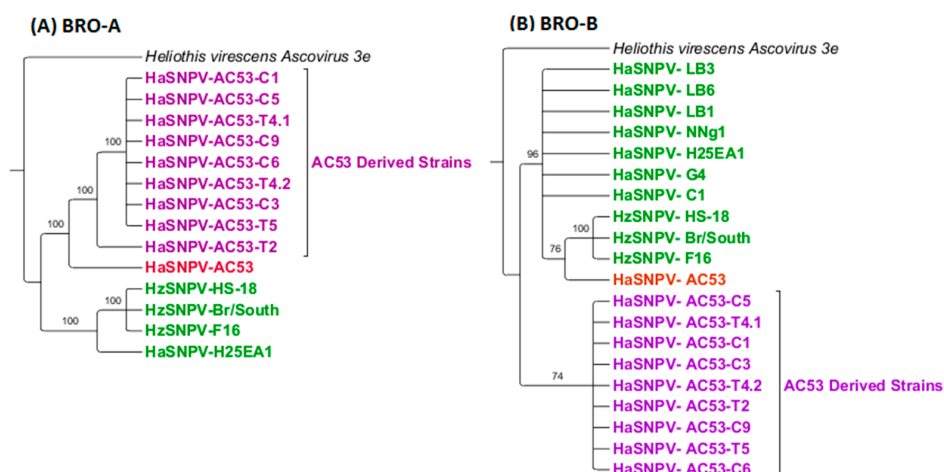


Figure 4. The root for both trees is the *Heliiothis virescens* Ascovirus 3e isolate and have been collapsed based on 70% bootstrap support. (A) baculovirus repeated open reading frame (BRO)-A maximum likelihood estimation (MLE) identified two distinct clusters, AC53 derived and non-AC53 derived; (B) BRO-B also identified two distinct clusters, however the parent AC53 strain was grouped with non-AC53 derived. A third group in BRO-B consisting of HzSNPV isolates could also be identified.

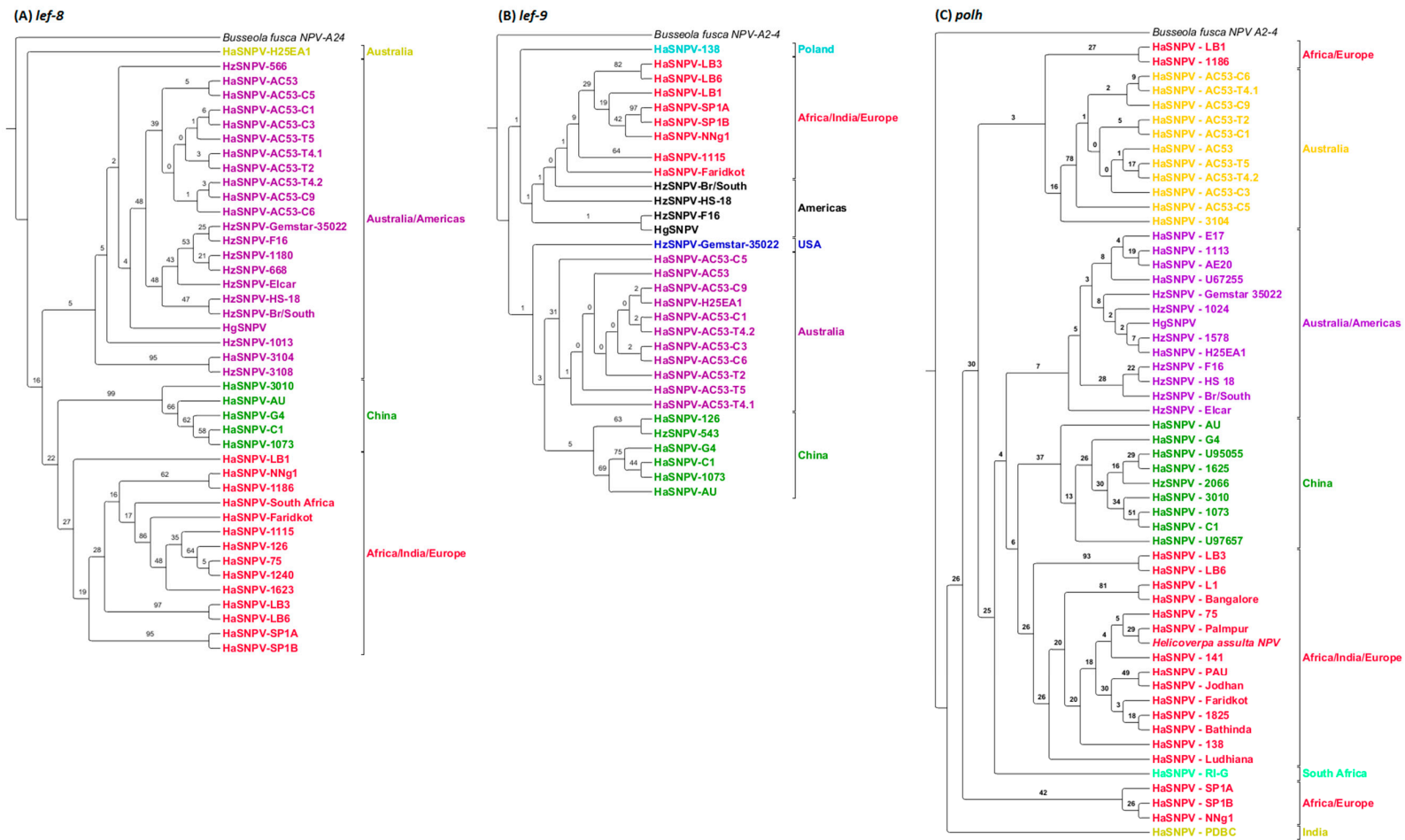


Figure 5. All trees have not been collapsed, due to poor bootstrapping support, and would not be differentiated otherwise, and have all been rooted to *Busseola fusca* NPV A2-4. **(A)** The geographic relationships of *lef-8* have at least three clusters, Africa/India/Europe (red), China (green), Australia/Americas (purple) and the Australian H25EA1 isolate (yellow-green); **(B)** The geographic relationships of *lef-9* have at least four clusters, Africa/India/Europe (red), Poland (aqua), the United States of America (blue) and North and South America (black), Australia (purple) and China (green); **(C)** The geographic relationships of *polh* have at least four clusters, India (yellow-green), Africa/Indian/Europe (red), China (green), South Africa (light green), Australia (yellow) and Australia/Americas (purple).

Analysis of the baculovirus ORFs *lef-8*, *lef-9* and *polh*, commonly used as markers, identified the similar geographic clusters as identified using whole-genome phylogenetic analysis. However, the high sequence similarity in these ORF led to very low levels of support (Figure 5). Isolates from India consistently clustered with the African/European isolates using all three ORFs.

The same three clusters of isolates by geographic origin (Australia/Americas, Europe and Africa, and China) were identified by MLE of *lef-8*, but with very low levels of support. Analysis of *polh* and *lef-9* was less powerful in resolving clusters. North and South American isolates were grouped separately from most of the Australian isolates except H25EA1, but again with very low levels of support. Clusters based on *polh* had particularly low levels of support.

Maximum likelihood analysis of the ORF42 and ORF78 identified the same three geographic clusters: Australia and Americas, Europe/Africa and China (Figure 6). The Indian isolate 'Faridkot' was separate from the clusters using ORF42. Indian isolate L1 was included with the European and African isolates using ORF78. ORF61 did not support resolution based on geographic origin or insect species (Figure 7), although two Indian isolates could be separated with 100% support, and the Iberian isolate LB1 was grouped with the Sudanese (African) isolate.

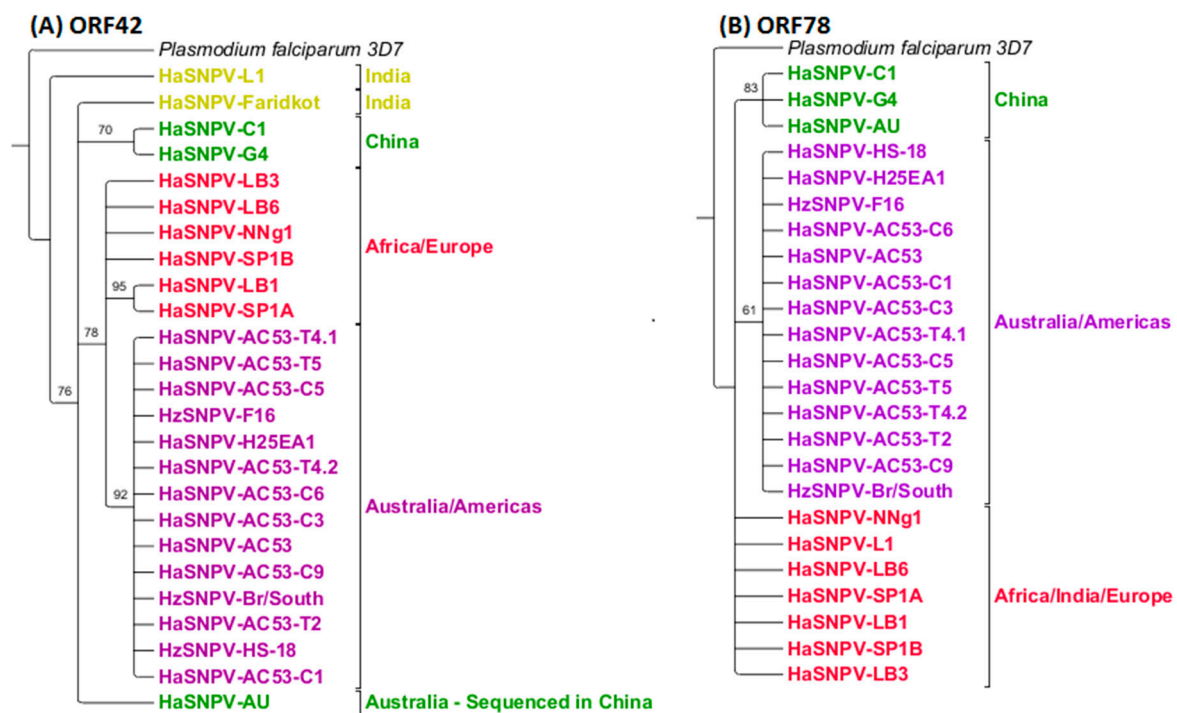


Figure 6. MLEs using (A) open reading frame (ORF) 42 and its homologs; and (B) ORF78 and its homologs identify the same three geographic clusters as the whole-genome sequences. Both trees have been rooted to *Plasmodium falciparum* 3D7 with a similar nucleotide sequence to ORF42 and ORF78, and collapsed to 60% minimum support values. Coloring code is identical to Figure 5.

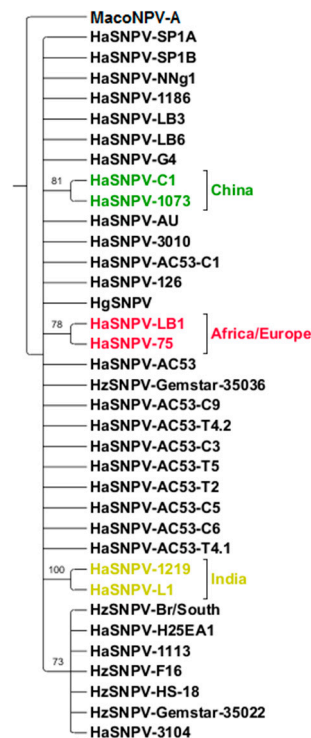


Figure 7. ORF61 (ORF62 homolog), rooted to MacoNPV-A, showing only six strains clustering based on geography.

4. Discussion

The strains derived from AC53 contained a number of polymorphisms but shared high sequence similarity with the parent (over 99.5%) and with each other (over 99.9%; Table S2). These could not be separated into clusters by maximum likelihood analysis.

Previous studies have shown significant deletions to occur in the *ecdysteroid UDP-glucosyltransferase (egt)* gene during passage in tissue culture [64,65]. However, no differences were found within the *egt* gene in the AC53-derived strains, but nine ORFs contained nucleotide sequences that differed from the parent AC53 isolate that resulted in changes in predicted amino acid sequence, of which changes in six ORFs were identical in all nine strains. We speculate that this is a result of selection of strains or of mutations arising as a result of passage of the virus through tissue culture.

The two strains, AC53-T4.1 and AC53T4.2, were isolated from two distinct peaks in time to death during passage in vivo of strain AC53-T4. In AC53-T4.2, ORF128 was split into two hypothetical ORFs of 267 bp (ORF128a) and 141 bp (ORF128b), and a 26 bp non-coding region. This split was not observed in AC53-T4.1 or the AC53 parent isolate. We speculate that changes in ORF128 result in differences in pathogenicity and speed of kill.

Sequencing of a number of ORFs has been used to attempt to differentiate isolates of HaSNPV by geographic or insect of origin [3]. However, maximum likelihood analysis using *polh*, *lef-9* and ORF61/62 homologs did not resolve either insect species or geographic origin of isolation with any significant degree of support, although not all the ORF61/62 isolates described in the published comparisons are available through Genbank [3]. On the other hand, maximum likelihood comparison of whole-genome sequences and sequences of *lef-8*, ORF42 and ORF78 from the Australian isolates and derived strains with sequences of homologs of other isolates available through Genbank identified a consistent pattern of clusters based on geographic origin: Australia and the Americas, Europe, Africa and India, and China. The Chinese 'AU' isolate was produced and sequenced in the same facility as

the other Chinese isolates (G4 and C1) and shares a high degree of sequence similarity with these isolates. We have classified it with these isolates.

There was no support for classification of the viruses based on species of isolation. The Australian isolates and derived strains showed high levels of sequence similarity with isolates from *H. zea*. Several hypothetical ORFs have been used to differentiate between isolates from *H. zea* and *H. armigera*, on the basis of presence or absence, but close inspection of the whole genome sequences on Genbank found that these hypothetical ORFs (ORF42, 62 and 78) were present in all sequences either as whole or truncated hypothetical ORFs that are not annotated [3,13]. A previously unannotated hypothetical ORF located between ORF54 and *lef-9* was identified but not annotated in all published HaSNPV genomes. Clustering of *H. zea* and *H. armigera* isolates based on MLEs from whole-genome and ORF sequences supported the classification of the viruses as a single species [12].

There was a consistent pattern of fragmentation of ORFs 78 and 61 of isolates AC53 and H25EA1 in comparison to homologs in other isolates. All isolates from the Americas had intact homologs with 100% sequence similarity to ORF61 and to ORF78 except for the Brazilian isolate Br/South which had a fragmentation of ORF78 similar to that of the AC53-derived strain AC53-C6.

The fragmentation of these ORFs became more pronounced in Iberian and African isolates. ORF61 was present as a full-length homolog in the African NNg1 and four Iberian strains, but Iberian isolate LB1 was truncated to a 72 bp unannotated hypothetical ORF homolog with 100% sequence similarity to a 72 bp truncated hypothetical ORF found all the strains derived from AC53. All homologs of ORF78 in the Iberian and African isolates were split into two, again similar to that found in AC53-C6. Fragmentation was most strongly found in the Chinese isolates. A truncated homolog of ORF61 with a 27 bp insertion identical to that identified in the AC53 derived strains was identified in isolate C1. In the other Chinese isolates, a 107 bp homolog of ORF61 was found in Hr3. ORF78 was further fragmented into three short, unannotated hypothetical ORFs.

This pattern of fragmentation suggests that strains have been selected following evolutionary bottlenecks. We speculate that the clustering of isolates based on MLE, in combination with the observed pattern of fragmentation suggests an origin of the HaSNPV species in Australia with subsequent movement to the Americas, then to Africa, Europe and India, and more recently to China, a pattern which follows global wind patterns [66–68]. The strains derived from AC53 also contained evidence of selection. Distinct clustering by MLE, using the sequences of ORFs BRO-A and BRO-B, and truncation and fragmentation of ORFs 61 and 78, suggest that plaque selection and passage through tissue culture led to the selection of new dominant genotypes from the AC53 isolate.

In general, the lack of support for clusters identified by MLE, based on ORFs 61, *polh*, *lef-9* and *lef-8* suggest that these regions cannot be used for taxonomic comparison within a virus species. Similarly, ORFs 42 and 78 were able to differentiate HaSNPV isolates based on geographic origin, but only a moderate level of support. Our analysis supports the conclusion that HaSNPV and HzSNPV are variants of the same species, and that host insect cannot be used to predict isolate identity [3,5,13,20], and suggests a possible origin for HaSNPVs isolates in Australia.

Supplementary Materials: The following tables and figures are available online at www.mdpi.com/1999-4915/8/11/280/s1, Figure S1: PCR detection for all NPV using rPol primer set; Figure S2: PCR detection of HaSNPV using the A44-RIX primer set; Figure S3: Nucleotide comparison of ORF7 within AC53 and its derived strains; Figure S4: Nucleotide comparison of ORF5 within AC53 and its derived strains; Figure S5: Nucleotide comparison of the four regions (A, B, C and D) containing mutations within the HOAR nucleotide sequence of AC53 and its derived strains; Figure S6: Nucleotide comparison of ORF128 with AC53 and its derivatives to the AC53-T4; Figure S7: Nucleotide comparison of fragmentation occurring within ORF78/79 with 10 distinct genotypes observed across all HaSNPV and HzSNPV strains; Figure S8: Nucleotide comparison of fragmentation occurring within ORF61/62 with 16 distinct genotypes observed across all HaSNPV and HzSNPV strains; Table S1: Nucleotide and Amino Acid comparison of the AC53 and H25EA1 strains; Table S2: Nucleotide distance matrix of AC53 and its derived strains; Table S3: *Lef-8* analysed strains; Table S4: *Lef-9* analysed strains; Table S5: *polh* analysed strains; Table S6: BRO-A and BRO-B analysed strains; Table S7: ORF42, ORF61 and ORF78 analysed strains.

Acknowledgments: This work was funded in part by the Cotton Research Development Corporation, by funding from the Queensland University of Technology (QUT), Australia; and materials were supplied by AgBiTech. Some of the data reported in this paper was generated in the Central Analytical Research Facility (CARF) operated by

the Institute for Future Environments at QUT. Access to CARF is supported by generous funding from the Science and Engineering Faculty at QUT. We would like to thank staff of the Molecular Genetics Research Facility and the Invertebrate Microbiology Group at QUT for their assistance with sequencing and technical support.

Author Contributions: Christopher Nouné and Caroline Hauxwell contributed equally to this paper.

Conflicts of Interest: The authors declare a conflict of interest. The Cotton Research Development Corporation has funded the work by Christopher Nouné through a post-graduate student scholarship. AgBiTech provided the sample of HaSNPV-AC53 and previously funded consultancy and research work with Caroline Hauxwell, but did not contribute financially to this study.

Script Availability: The IMG-AP is available for download at https://github.com/CNouné/IMG_pipelines.

References

- Rohrmann, G. Introduction to the baculoviruses and their taxonomy. In *Baculovirus Molecular Biology*, 2nd ed.; National Center for Biotechnology Information: Bethesda, MD, USA, 2011.
- Blissard, G.W.; Rohrmann, G.F. Baculovirus diversity and molecular biology. *Annu. Rev. Entomol.* **1990**, *35*, 127–155. [[CrossRef](#)] [[PubMed](#)]
- Rowley, D.L.; Popham, H.J.R.; Harrison, R.L. Genetic variation and virulence of nucleopolyhedroviruses isolated worldwide from the heliothine pests *Helicoverpa armigera*, *Helicoverpa zea*, and *Heliothis virescens*. *J. Invertebr. Pathol.* **2011**, *107*, 112–126. [[CrossRef](#)] [[PubMed](#)]
- Jehle, J.A.; Lange, M.; Wang, H.; Hu, Z.; Wang, Y.; Hauschild, R. Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera. *Virology* **2006**, *346*, 180–193. [[CrossRef](#)] [[PubMed](#)]
- Jehle, J.A.; Blissard, G.W.; Bonning, B.C.; Cory, J.S.; Herniou, E.A.; Rohrmann, G.F.; Theilmann, D.A.; Thiem, S.M.; Vlaskovits, J.M. On the classification and nomenclature of baculoviruses: A proposal for revision. *Arch. Virol.* **2006**, *151*, 1257–1266. [[CrossRef](#)] [[PubMed](#)]
- Herniou, E.A.; Jehle, J.A. Baculovirus phylogeny and evolution. *Curr. Drug Targets* **2007**, *8*, 1043–1050. [[CrossRef](#)] [[PubMed](#)]
- Wardhaugh, K.G.; Room, P.M.; Greenup, L.R. The incidence of *Heliothis armigera* (Hübner) and *H. punctigera* Wallengren (Lepidoptera: Noctuidae) on cotton and other host-plants in the Namoi Valley of New South Wales. *Bull. Entomol. Res.* **1980**, *70*, 113–131. [[CrossRef](#)]
- Daly, J.C.; Gregg, P. Genetic variation in *Heliothis* in Australia: Species identification and gene flow in the two pest species *H. armigera* (Hübner) and *H. punctigera* Wallengren (Lepidoptera: Noctuidae). *Bull. Entomol. Res.* **1985**, *75*, 169–184. [[CrossRef](#)]
- Zhang, G. Commercial viral insecticide *Heliothis armigera* viral insecticide in China. *IPM Pract.* **1989**, *11*, 13.
- Richards, A.R.; Christian, P.D. A rapid bioassay screen for quantifying nucleopolyhedroviruses (Baculoviridae) in the environment. *J. Virol. Methods* **1999**, *82*, 63–75. [[CrossRef](#)]
- Nouné, C.; Hauxwell, C. Complete genome sequences of seven *Helicoverpa armigera* SNPV-AC53-derived strains. *Genome Announc.* **2016**, *4*, e00260-16. [[CrossRef](#)] [[PubMed](#)]
- Harrison, R.L.; Herniou, E.A.; Theilmann, D.A.; Becnel, J.J.; Arif, B.; Jehle, J.A.; Burand, J.P.; Oers, M.V. *Removal of Species Helicoverpa zea Single Nucleopolyhedrovirus from the Genus Alphabaculovirus*; International Committee on Taxonomy of Viruses: Edinburgh, UK, 2013.
- Chen, X.; Zhang, W.-J.; Wong, J.; Chun, G.; Lu, A.; McCutchen, B.; Presnail, J.; Herrmann, R.; Dolan, M.; Tingey, S.; et al. Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedroviruses. *J. Gen. Virol.* **2002**, *83*, 673–684. [[CrossRef](#)] [[PubMed](#)]
- Buerger, P.; Hauxwell, C.; Murray, D. Nucleopolyhedrovirus introduction in Australia. *Virol. Sin.* **2007**, *22*, 173–179. [[CrossRef](#)]
- Nguyen, Q.; Qi, Y.M.; Wu, Y.; Chan, L.C.L.; Nielsen, L.K.; Reid, S. In vitro production of *Helicoverpa baculovirus* biopesticides—Automated selection of insect cell clones for manufacturing and systems biology studies. *J. Virol. Methods* **2011**, *175*, 197–205. [[CrossRef](#)] [[PubMed](#)]
- Christian, P.D.; Gibb, N.; Kasprzak, A.B.; Richards, A. A rapid method for the identification and differentiation of *Helicoverpa* nucleopolyhedroviruses (NPV *Baculoviridae*) isolated from the environment. *J. Virol. Methods* **2001**, *96*, 51–65. [[CrossRef](#)]

17. Noune, C.; Hauxwell, C. Complete genome sequences of *Helicoverpa armigera* single nucleopolyhedrovirus strains AC53 and H25EA1 from Australia. *Genome Announc.* **2015**, *3*, e01083-15. [[CrossRef](#)] [[PubMed](#)]
18. Baillie, V.L.; Bouwer, G. High levels of genetic variation within *Helicoverpa armigera* nucleopolyhedrovirus populations in individual host insects. *Arch. Virol.* **2012**, *157*, 2281–2289. [[CrossRef](#)] [[PubMed](#)]
19. Erlandson, M.A. Genetic variation in field populations of baculoviruses: Mechanisms for generating variation and its potential role in baculovirus epizootiology. *Virol. Sin.* **2009**, *24*, 458–469. [[CrossRef](#)]
20. Gettig, R.R.; McCarthy, W.J. Genotypic variation among wild isolates of *Heliothis* spp. nuclear polyhedrosis viruses from different geographical regions. *Virology* **1982**, *117*, 245–252. [[CrossRef](#)]
21. Crawford, A.M.; Zelazny, B.; Alfiler, A.R. Genotypic variation in geographical isolates of oryctes baculovirus. *J. Gen. Virol.* **1986**, *67*, 949–952. [[CrossRef](#)]
22. Corsaro, B.G.; Fraser, M.J. Characterization of genotypic and phenotypic variation in plaque-purified strains of HzSNPV elkar isolate. *Intervirology* **1987**, *28*, 185–198. [[PubMed](#)]
23. Ogembo, J.G.; Caoili, B.L.; Shikata, M.; Chaeychomsri, S.; Kobayashi, M.; Ikeda, M. Comparative genomic sequence analysis of novel *Helicoverpa armigera* nucleopolyhedrovirus (NPV) isolated from Kenya and three other previously sequenced *Helicoverpa* spp. NPVs. *Virus Genes* **2009**, *39*, 261–272. [[CrossRef](#)] [[PubMed](#)]
24. Ogembo, J.G.; Chaeychomsri, S.; Kamiya, K.; Ishikawa, H.; Katou, Y.; Ikeda, M.; Kobayashi, M. Cloning and comparative characterization of nucleopolyhedroviruses isolated from African bollworm, *Helicoverpa armigera*, (Lepidoptera: Noctuidae) in different geographic regions. *J. Insect Biotechnol. Sericol.* **2007**, *76*, 39–49.
25. Kabaluk, J.T.; Svircev, A.M.; Goettel, M.S.; Woo, S.G. (Eds.) *The Use and Regulation of Microbial Pesticides in Representative Jurisdictions Worldwide*; IOBC Global, 2010; p. 99. Available online: http://www.iobc-global.org/download/Microbial_Regulation_Book_Kabaluk_et_al_2010.pdf (accessed on 17 October 2016).
26. Baillie, V.L.; Bouwer, G. Development of highly sensitive assays for detection of genetic variation in key *Helicoverpa armigera* nucleopolyhedrovirus genes. *J. Virol. Methods* **2011**, *178*, 179–185. [[CrossRef](#)] [[PubMed](#)]
27. Craveiro, S.R.; Inglis, P.W.; Togawa, R.C.; Grynberg, P.; Melo, F.L.; Ribeiro, Z.M.; Ribeiro, B.M.; Bão, S.N.; Castro, M.E. The genome sequence of *Pseudoplusia includens* single nucleopolyhedrovirus and an analysis of p26 gene evolution in the baculoviruses. *BMC Genom.* **2015**, *16*, 127. [[CrossRef](#)] [[PubMed](#)]
28. Chateigner, A.; Bézier, A.; Labrousse, C.; Jiolle, D.; Barbe, V.; Herniou, E.A. Ultra deep sequencing of a baculovirus population reveals widespread genomic variations. *Viruses* **2015**, *7*, 3625–3646. [[CrossRef](#)] [[PubMed](#)]
29. Quail, M.A.; Smith, M.; Coupland, P.; Otto, T.D.; Harris, S.R.; Connor, T.R.; Bertoni, A.; Swerdlow, H.P.; Gu, Y. A tale of three Next Generation Sequencing platforms: Comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genom.* **2012**, *13*, 341. [[CrossRef](#)] [[PubMed](#)]
30. Harrison, R.L. Structural divergence among genomes of closely related baculoviruses and its implications for baculovirus evolution. *J. Invertebr. Pathol.* **2009**, *101*, 181–186. [[CrossRef](#)] [[PubMed](#)]
31. Zhang, C.-X.; Ma, X.-C.; Guo, Z.-J. Comparison of the complete genome sequence between C1 and G4 isolates of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus. *Virology* **2005**, *333*, 190–199. [[CrossRef](#)] [[PubMed](#)]
32. Pedrini, M.R.; Christian, P.; Nielsen, L.K.; Reid, S.; Chan, L.C. Importance of virus—Medium interactions on the biological activity of wild-type *Heliothis* nucleopolyhedroviruses propagated via suspension insect cell cultures. *J. Virol. Methods* **2006**, *136*, 267–272. [[CrossRef](#)] [[PubMed](#)]
33. Nguyen, Q.; Nielsen, L.K.; Reid, S. Genome scale transcriptomics of baculovirus-insect interactions. *Viruses* **2013**, *5*, 2721–2747. [[CrossRef](#)] [[PubMed](#)]
34. Lua, L.H.; Reid, S. Virus morphogenesis of *Helicoverpa armigera* nucleopolyhedrovirus in *Helicoverpa zea* serum-free suspension culture. *J. Gen. Virol.* **2000**, *81*, 2531–2543. [[CrossRef](#)] [[PubMed](#)]
35. Lua, L.H.; Pedrini, M.R.; Reid, S.; Robertson, A.; Tribe, D.E. Phenotypic and genotypic analysis of *Helicoverpa armigera* nucleopolyhedrovirus serially passaged in cell culture. *J. Gen. Virol.* **2002**, *83*, 945–955. [[CrossRef](#)] [[PubMed](#)]
36. Hughes, P.R.; Wood, H.A. A synchronous peroral technique for the bioassay of insect viruses. *J. Invertebr. Pathol.* **1981**, *37*, 154–159. [[CrossRef](#)]
37. Hughes, D.S.; Possee, R.D.; King, L.A. Evidence for the presence of a low-level, persistent baculovirus infection of *Mamestra brassicae* insects. *J. Gen. Virol.* **1997**, *78*, 1801–1805. [[CrossRef](#)] [[PubMed](#)]

38. Hughes, D.S.; Possee, R.D.; King, L.A. Activation and detection of a latent baculovirus resembling *Mamestra brassicae* nuclear polyhedrosis virus in *M. brassicae* insects. *Virology* **1993**, *194*, 608–615. [[CrossRef](#)] [[PubMed](#)]
39. Brown, M.; Faulkner, P. Plaque assay of nuclear polyhedrosis viruses in cell culture. *Appl. Environ. Microbiol.* **1978**, *36*, 31–35. [[PubMed](#)]
40. BDBiosciences. Plaque Assay. Available online: http://www.bdbiosciences.com/br/resources/baculovirus/protocols/plaque_assay.jsp (accessed on 17 September 2012).
41. Hauxwell, I.C. *Evaluation of Potential Baculovirus Insecticides: Studies of the Infection Process and Host Susceptibility*; Imperial College London (University of London): London, UK, 1999.
42. Matsuura, Y.; Possee, R.D.; Overton, H.A.; Bishop, D.H. Baculovirus expression vectors: The requirements for high level expression of proteins, including glycoproteins. *J. Gen. Virol.* **1987**, *68*, 1233–1250. [[CrossRef](#)] [[PubMed](#)]
43. Doyle, C.J.; Hirst, M.L.; Cory, J.S.; Entwistle, P.F. Risk assessment studies: Detailed host range testing of wild-type cabbage moth, *Mamestra brassicae* (Lepidoptera: Noctuidae), nuclear polyhedrosis virus. *Appl. Environ. Microbiol.* **1990**, *56*, 2704–2710. [[PubMed](#)]
44. Zhang, H.; Yang, Q.; Qin, Q.L.; Zhu, W.; Zhang, Z.F.; Li, Y.N.; Zhang, N.; Zhang, J.H. Genomic sequence analysis of *Helicoverpa armigera* nucleopolyhedrovirus isolated from Australia. *Arch. Virol.* **2014**, *159*, 595–601. [[CrossRef](#)] [[PubMed](#)]
45. Arrizubieta, M.; Williams, T.; Caballero, P.; Simón, O. Selection of a nucleopolyhedrovirus isolate from *Helicoverpa armigera* as the basis for a biological insecticide. *Pest Manag. Sci.* **2014**, *70*, 967–976. [[CrossRef](#)] [[PubMed](#)]
46. Arrizubieta, M.; Simón, O.; Williams, T.; Caballero, P. A novel binary mixture of *Helicoverpa armigera* single nucleopolyhedrovirus genotypic variants has improved insecticidal characteristics for control of cotton bollworms. *Appl. Environ. Microbiol.* **2015**, *81*, 3984–3993. [[CrossRef](#)] [[PubMed](#)]
47. Arrizubieta, M.; Simón, O.; Williams, T.; Caballero, P. Genomic sequences of five *Helicoverpa armigera* nucleopolyhedrovirus genotypes from Spain that differ in their insecticidal properties. *Genome Announc.* **2015**, *3*, e00548-15. [[CrossRef](#)] [[PubMed](#)]
48. Chen, X.; IJkel, W.F.; Tarchini, R.; Sun, X.; Sandbrink, H.; Wang, H.; Peters, S.; Zuidema, D.; Lankhorst, R.K.; Vlak, J.M. The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome. *J. Gen. Virol.* **2001**, *82*, 241–257. [[CrossRef](#)] [[PubMed](#)]
49. Ardisson-Araújo, D.M.; Sosa-Gomez, D.R.; Melo, F.L.; Bão, S.N.; Ribeiro, B.M. Characterization of *Helicoverpa zea* single nucleopolyhedrovirus isolated in Brazil during the first old world bollworm (Noctuidae: *Helicoverpa armigera*) nationwide outbreak. *Virus Rev. Res.* **2015**, *20*, 2. [[CrossRef](#)]
50. Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **1994**, *202*, 586–605. [[CrossRef](#)] [[PubMed](#)]
51. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
52. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]
53. Stöver, B.C.; Müller, K.F. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinf.* **2010**, *11*. [[CrossRef](#)] [[PubMed](#)]
54. Lange, M.; Wang, H.; Zhihong, H.; Jehle, J.A. Towards a molecular identification and classification system of lepidopteran-specific baculoviruses. *Virology* **2004**, *325*, 36–47. [[CrossRef](#)] [[PubMed](#)]
55. Ferrelli, M.; Taibo, C.; Fichetti, P.; Sciocco-Cap, A.; Arneodo, J. Characterization of a new *Helicoverpa armigera* nucleopolyhedrovirus variant causing epizootic on a previously unreported host, *Helicoverpa gelatopoeon* (Lepidoptera: Noctuidae). *J. Invertebr. Pathol.* **2015**, *138*, 89–93. [[CrossRef](#)] [[PubMed](#)]
56. Woo, S.-D.; Choi, J.Y.; Je, Y.H.; Jin, B.R. Characterization of the *Helicoverpa assulta* nucleopolyhedrovirus genome and sequence analysis of the polyhedrin gene region. *J. Biosci.* **2006**, *31*, 329–338. [[CrossRef](#)] [[PubMed](#)]
57. Li, Q.; Donly, C.; Li, L.; Willis, L.G.; Theilmann, D.A.; Erlandson, M. Sequence and organization of the *Mamestra configurata* nucleopolyhedrovirus genome. *Virology* **2002**, *294*, 106–121. [[CrossRef](#)] [[PubMed](#)]

58. Li, S.; Erlandson, M.; Moody, D.; Gillott, C. A physical map of the *Mamestra configurata* nucleopolyhedrovirus genome and sequence analysis of the polyhedrin gene. *J. Gen. Virol.* **1997**, *78 Pt 1*, 265–271. [[CrossRef](#)] [[PubMed](#)]
59. Asgari, S.; Davis, J.; Wood, D.; Wilson, P.; McGrath, A. Sequence and organization of the *Heliothis virescens* ascovirus genome. *J. Gen. Virol.* **2007**, *88*, 1120–1132. [[CrossRef](#)] [[PubMed](#)]
60. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
61. Morgulis, A.; Coulouris, G.; Raytselis, Y.; Madden, T.L.; Agarwala, R.; Schaffer, A.A. Database indexing for production MegaBLAST searches. *Bioinformatics* **2008**, *24*, 1757–1764. [[CrossRef](#)] [[PubMed](#)]
62. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinf.* **2009**, *10*, 421. [[CrossRef](#)] [[PubMed](#)]
63. Rice, P.; Longden, I.; Bleasby, A. EMBOSS: The European molecular biology open software suite. *Trends Genet.* **2000**, *16*, 276–277. [[CrossRef](#)]
64. Harrison, R.L. Genomic sequence analysis of the Illinois strain of the *Agrotis ipsilon* multiple nucleopolyhedrovirus. *Virus Genes* **2009**, *38*, 155–170. [[CrossRef](#)] [[PubMed](#)]
65. Simon, O.; Palma, L.; Beperet, I.; Munoz, D.; Lopez-Ferber, M.; Caballero, P.; Williams, T. Sequence comparison between three geographically distinct *Spodoptera frugiperda* multiple nucleopolyhedrovirus isolates: Detecting positively selected genes. *J. Invertebr. Pathol.* **2011**, *107*, 33–42. [[CrossRef](#)] [[PubMed](#)]
66. Parrish, J.T.; Peterson, F. Wind directions predicted from global circulation models and wind directions determined from eolian sandstones of the western United States—A comparison. *Sediment. Geol.* **1988**, *56*, 261–282. [[CrossRef](#)]
67. Trenberth, K.E.; Large, W.G.; Olson, J.G. The mean annual cycle in global ocean wind stress. *J. Phys. Oceanogr.* **1990**, *20*, 1742–1760. [[CrossRef](#)]
68. Parrish, J.T.; Curtis, R.L. Atmospheric circulation, upwelling, and organic-rich rocks in the Mesozoic and Cenozoic eras. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **1982**, *40*, 31–66. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).